S

Scaling Limits of Large Systems of Non-linear Partial Differential Equations

D. Benedetto, M. Pulvirenti
Dipartimento di Matematica, Università di Roma 'La Sapienza', Roma, Italy

Article Outline

Glossary
Definition of the Subject
Introduction
Weak-Coupling Limit for Classical Systems
Weak-Coupling Limit for Quantum Systems
Weak-Coupling Limit in the Bose–Einstein and Fermi–Dirac Statistics
Weak-Coupling Limit for a Single Particle: The Linear Theory
Future Directions
Bibliography

Glossary

Scaling limits A scaling limit denotes a procedure to reduce the degree of complexity of a large particle system. It consists in scaling the space-time variables and, possibly, other quantities (like the interaction potential or the density), in order to obtain a more handable description of the same system. The initial space-time coordinates are called **microscopic**, while the new ones, those well suited for the description of kinetic or hydrodynamical systems, are called **macroscopic**.

Boltzmann equation It is an integro-differential kinetic equation for the one particle distribution function in the classical phase space (see Eq. (64) below). It arises in some physical regimes, namely for a rarefied gas and for a weakly interacting quantum dense gas.

Uehling–Uhlenbeck equation It is a Boltzmann type equation taking into account corrections due to the Bose–Einstein or the Fermi–Dirac statistics (see Eq. (69) below). It holds in the weak-coupling limit.

Fokker–Planck–Landau equation It is a kinetic equation diffusive in velocity (see Eq. (28) below). It arises in the context of a weakly interacting classical dense gas.

Hydrodynamical equations They are evolution equations for macroscopic quantities like density, mean velocity, temperature, and so on.

Low density limit Sometimes called Boltzmann–Grad limit, it is a scaling limit in which the density is vanishing. Applied to classical particle systems it gives the Boltzmann equation.

Weak coupling limit It is a scaling limit in which the density is constant but the interaction vanishes suitably. Applied to a quantum particle systems it gives a Boltzmann equation. Applied to a classical particle systems it gives the Fokker–Planck–Landau equation.

Hydrodynamic limit In this scaling we simply pass from micro to macro variables. We look at the behavior of suitable mean values, which are functions of the space and the time. We expect them to behave in a hydrodynamical way, namely to satisfy a set of hydrodynamical equations.

Wigner transform It is the description of a quantum state as a function in the classical phase space.

Definition of the Subject

Many interesting systems in Physics and Applied Sciences are constituted by a large number of identical components so that they are difficult to analyze from a mathematical point of view. On the other hand, quite often, we are not interested in a detailed description of the system but rather to his global behavior. Therefore it is necessary to look for all procedures leading to simplified models, retaining all the interesting features of the original system, cutting away unnecessary informations. This is exactly the methodology of the Statistical Mechanics and the Kinetic Theory when dealing with large particle systems.

In this contribution we want to approach a problem of this type, namely we try to outline the difficulties in making rigorous the limiting procedure leading the microscopic description of a large particle system (based on the fundamental laws like Newton or Schrödinger equations) to a more handable kinetic or hydrodynamic picture. Such a transition depends on the space-time scale we choose to describe the phenomenon. To fix the ideas we start by considering a system constituted by N identical classical point particles of unitary mass. The Newton equations are:

$$\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} q_i = \sum_{\substack{j=1\dots N: \\ j\neq i}} F(q_i - q_j)\,, \tag{1}$$

where $\{q_1 \dots q_N\}$, $q_i \in \mathbb{R}^3$ are the position of the particles and τ is the time. Here $F = -\nabla\phi$ denotes the interparticle (conservative) force and ϕ the two-body interaction potential.

We now introduce a small parameter $\varepsilon > 0$ expressing the ratio between the macroscopic and the microscopic scales. For instance ε could denote the inverse ratio of a typical distance between two molecules of a gas, measured in microscopic unities, and the same distance measured in meters. It can be as well the ratio between typical macroscopic and microscopic times. We now introduce the new variables

$$x = \varepsilon q \qquad t = \varepsilon\tau\,.$$

Since we are interested in the macroscopic properties of the system, namely in the evolution of macroscopic quantities that are those varying on the (x, t) scale and hence almost constant in the (q, τ) scale, it is natural to rescale also Eq. (1). We write

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} x_i = \frac{1}{\varepsilon} \sum_{\substack{j=1\dots N: \\ j\neq i}} F\left(\frac{x_i - x_j}{\varepsilon}\right). \tag{2}$$

Up to now we did nothing else than an innocent change of variables. However to obtain a non trivial description, N has to diverge when $\varepsilon \to 0$, and we have to specify how. Also additional hypotheses on the strength of the interaction, according to the physical situation at hand, may be necessary. There are various possible scaling limits we shall discuss below.

- *The hydrodynamic limit*
  Here the system is dense, namely $N = O(\varepsilon^{-3})$. For all $x \in \Omega \subset \mathbb{R}^3$, we consider a small box Δ around x. In Δ there are a very large number of particles. We denote the initial density, mean velocity and temperature by $(\rho_0(x), u_0(x), T_0(x))$. At a positive time t the particles starts to interact. For a microscopic time τ the particles in Δ interact practically among themselves only (supposing a short range interaction) and, if an ergodic property would hold, we expect that they reach a thermal equilibrium state. We remind that $t = \varepsilon\tau$ is essentially vanishing, therefore we expect that at macroscopic time $t \approx 0$ the particle are distributed according to the Gibbs measure characterized by the parameters $(\rho_0(x), u_0(x), T_0(x))$, being mass momentum and energy locally conserved. On a larger time scale ($t > 0$) mass momentum and energy are not locally conserved anymore. If a local equilibrium structure is preserved, the parameters of the equilibria $(\rho(x, t), u(x, t), T(x, t))$ will evolve according to five equations which are the exactly the well known Euler equations for compressible gas.
  This complex mechanism is very far to be proven rigorously, even for short times and initial regular data. A heuristic derivation of the Euler equations in terms of particle dynamics was first given by Morrey [35]. We address the reader to [39] and [19] for further references and comments.
- *The low-density (Boltzmann–Grad) limit*
  Now we are dealing with a rarefied gas assuming $N = O(\varepsilon^{-2})$. This is the reason why this scaling is called low-density. We assume also, just for simplicity, that the interaction range of ϕ is one. Consider now a given test particle. If the gas is more or less homogeneous, the number of particles interacting with this particle in a given (macroscopic) time interval is $O(N\varepsilon^2) = O(1)$. Therefore we expect a finite number of collisions per unit time. Moreover each collision is almost instantaneous so that each particle undergoes a jump process in velocity.
  We also assume that initially the particles are independently distributed according to the one-particle distribution on the phase space $f_0(x, v)$ (v is the velocity). Of course the dynamics creates correlations so that, strictly speaking, we expect that the particles, for every positive time t, are not independently distributed anymore. However the probability that a given pair of particles interact is vanishing in the limit $\varepsilon \to 0$. Therefore the statistical independence (called propagation of chaos) is recovered in the same limit. In other words we expect, for the time evolved distribution $f(x, v, t)$, a nonlinear evolution equation, which is the celebrated Boltzmann equation discovered in 1872. This scaling limit, often called the Boltzmann–Grad limit [25], has been proved rigorously by Lanford [31] for a short time

interval (see also [28] for a special case in which the limit holds globally in time). We address the reader to refs. [14,39] and [19] for additional comments and references.

- *The weak-coupling limit*
  Now the gas is dense because we assume $N = O(\varepsilon^{-3})$, however the particles are weakly interacting. This condition is expressed by rescaling the interaction potential by setting

$$\phi \to \sqrt{\varepsilon}\phi \,.$$

  After that the equation of motion (in macroscopic variables) become:

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} x_i = -\frac{1}{\sqrt{\varepsilon}} \sum_{\substack{j=1\ldots N: \\ j\neq i}} \nabla\phi\left(\frac{x_i - x_j}{\varepsilon}\right) . \tag{3}$$

  To have a vague idea of the behavior of a test particle in this limit, we observe first that the velocity change due to a single collision is

$$\Delta v \approx F\Delta t \approx O(\varepsilon/\sqrt{\varepsilon}) = O(\sqrt{\varepsilon}) \,.$$

  Here $\Delta t = O(\varepsilon)$ is the time in which the collision takes place, ϕ is assumed short-range in microscopic variables and then $1/\sqrt{\varepsilon}$ is the order of magnitude of the force F.
  The number of collisions per unit time is $\varepsilon^2\varepsilon^{-3} = \varepsilon^{-1}$, so that

$$\sum |\Delta v|^2 = O(1) \,.$$

  Therefore, provided that the propagation of chaos is ensured, $v = v(t)$ is expected to be not absolutely continuous in the limit and the one-particle distribution function to satisfy a diffusion equation in velocity, the Landau–Fokker–Planck equation.
  The statistical independence at a positive time t is expected because the effect of the interaction between two given particle is going to vanish in the limit $\varepsilon \to 0$. Note that the physical meaning of the propagation of chaos here is quite different from that arising in the contest of the Boltzmann equation. Here two particles can interact but the effect of the collision is small, while in a low-density regime the effect of a collision between two given particles is large but quite unlikely.
  The above scalings make sense as well for quantum systems. What we expect in this case? Notice that the transition from micro to macro variables increase the frequency of the quantum oscillations so that we are always in presence of a semiclassical limit as regards the kinetic energy part of the Hamiltonian. However the potential varies on the same scale of the oscillations. As a consequence we face the following scenario. The hydrodynamic limit of quantum systems yields the Euler equations for the equilibria parameters ρ, u, T. The only difference with the classical case is that the relationship between the pressure and the hydrodynamical fields, is that dictated by the Quantum Statistical Mechanics. Indeed the local equilibrium is achieved at a microscopic level and this is the only quantum effect we expect to survive in the limit.
  The situation is conceptually similar for the low density limit: here we have the classical Boltzmann equation for the one-particle distribution function. The only quantum macroscopic effect is that the cross-section must be computed in terms of the quantum scattering process (see [6]).

Introduction

The hydrodynamical and the low density limits are relatively more popular then the weak-coupling limit, which will be, for that reason, the object of the present contribution. More precisely we analyze large classical and quantum particle systems in the weak-coupling regime. We show how we expect such systems to be well described by the Landau–Fokker–Planck and the Boltzmann equations, for the classical and the quantum case respectively. We also describe the effects of the Fermi–Dirac and Bose–Einstein statistics.

Unfortunately our arguments we present here are largely formal because the rigorous results we know up to now, are few. A rigorous proof of the validity of the weak-coupling limit, or a rigorous derivation of the corresponding kinetic equations, is a challenging and still open problem.

In Sect. "Weak-Coupling Limit for Classical Systems" we introduce and discuss the problem for classical systems, presenting a formal proof of the validity of the Landau equation.

In Sect. "Weak-Coupling Limit for Quantum Systems" we pass to analyze the case of a quantum system, neglecting the correlations due to the statistics. The case of Bosons and Fermions is discussed in Sect. "Weak-Coupling Limit in the Bose–Einstein and Fermi–Dirac Statistics".

Finally, in Sect. "Weak-Coupling Limit for a Single Particle: the Linear Theory" we briefly introduce and discuss the corresponding linear problems, namely the behavior of a single (classical or quantum) particle in a random distribution of obstacles.

Section "Future Directions" is devoted to concluding remarks.

Weak-Coupling Limit for Classical Systems

We consider a classical system of N identical particles of unitary mass. Rescaled positions and velocities are denoted by $x_1 \dots x_N$ and $v_1 \dots v_N$. After having also rescaled the potential in the following way

$$\phi \to \sqrt{\varepsilon}\phi\,,$$

the Newton equations reads as

$$\frac{\mathrm{d}}{\mathrm{d}t}x_i = v_i \quad \frac{\mathrm{d}}{\mathrm{d}t}v_i = -\frac{1}{\sqrt{\varepsilon}}\sum_{\substack{j=1\dots N:\\ j\neq i}} \nabla\phi\left(\frac{x_i - x_j}{\varepsilon}\right)\,. \quad (4)$$

We also assume that $N = O(\varepsilon^{-3})$, namely the density is $O(1)$.

We are interested in the statistical behavior of the system so that we introduce a probability distribution on the phase space of the system, $W^N = W^N(X_N, V_N)$, that is the state at time zero. Here (X_N, V_N) denote the set of positions and velocities:

$$X_N = (x_1, \dots, x_N) \quad V_N = (v_1, \dots, v_N)\,.$$

Then from Eq. (4) we obtain the following Liouville equation

$$(\partial_t + V_N\cdot\nabla_N)W^N(X_N, V_N) = \frac{1}{\sqrt{\varepsilon}}\left(T_N^\varepsilon W^N\right)(X_N, V_N) \quad (5)$$

where $V_N \cdot \nabla_N = \sum_{i=1}^N v_i \cdot \nabla_{x_i}$ and $(\partial_t + V_N \cdot \nabla_N)$ is the usual free stream operator. Also, we have introduced the operator

$$\left(T_N^\varepsilon W^N\right)(X_N, V_N) = \sum_{0<k<\ell\leq N} (T_{k,\ell}^\varepsilon W^N)(X_N, V_N)\,, \quad (6)$$

with

$$T_{k,\ell}^\varepsilon W^N = \nabla\phi\left(\frac{x_k - x_\ell}{\varepsilon}\right)\cdot(\nabla_{v_k} - \nabla_{v_\ell})W^N\,. \quad (7)$$

To investigate the limit $\varepsilon \to 0\,,\ N = \varepsilon^{-3}$, it is convenient to introduce the BBKGY hierarchy (from Bogoliubov, Born, Green, Kirkwood and Yvon, see e. g. [2] and [19]), for the j-particle distributions defined as

$$f_j^N(X_j, V_j) = \int \mathrm{d}x_{j+1}\dots\int \mathrm{d}x_N \int \mathrm{d}v_{j+1}\dots\int \mathrm{d}v_N \; W^N(X_j, x_{j+1}\dots x_N, V_j, v_{j+1}\dots v_N) \quad (8)$$

for $j = 1, \dots, N-1$. Obviously, we set $f_N^N = W^N$.

From now on we shall suppose that, due to the fact that the particles are identical, the function W^N and f_j^N which we have introduced are all symmetric in the exchange of particles.

A partial integration of the Liouville equation (5) and standard manipulations give us the following hierarchy of equations: (for $1 \leq j \leq N$):

$$\left(\partial_t + \sum_{k=1}^j v_k\cdot\nabla_k\right)f_j^N = \frac{1}{\sqrt{\varepsilon}}T_j^\varepsilon f_j^N + \frac{N-j}{\sqrt{\varepsilon}}C_{j+1}^\varepsilon f_{j+1}^N\,. \quad (9)$$

The operator C_{j+1}^ε is defined as:

$$C_{j+1}^\varepsilon = \sum_{k=1}^j C_{k,j+1}^\varepsilon\,, \quad (10)$$

and

$$C_{k,j+1}^\varepsilon f_{j+1}(x_1 \dots x_j, v_1 \dots v_j) = -\int \mathrm{d}x_{j+1}\int \mathrm{d}v_{j+1} F\left(\frac{x_k - x_\ell}{\varepsilon}\right) \cdot\nabla_{v_k} f_{j+1}(x_1, x_2, \dots, x_{j+1}, v_1, \dots, v_{j+1})\,. \quad (11)$$

$C_{k,j+1}^\varepsilon$ describes the "collision" of particle k, belonging to the j-particle subsystem, with a particle outside the subsystem, conventionally denoted by the number $j+1$ (this numbering uses the fact that all particles are identical). The total operator C_{j+1}^ε takes into account all such collisions. The dynamics of the j-particle subsystem is governed by three effects: the free-stream operator, the "recollisions", i. e. the collisions "inside" the subsystem, given by the T term, and the "creations", i. e. the collisions with particles "outside" the subsystem, given by the C term.

We finally fix the initial value $\{f_j^0\}_{j=1}^N$ of the solution $\{f_j^N(t)\}_{j=1}^N$ assuming that $\{f_j^0\}_{j=1}^N$ is factorized, that is, for all $j = 1, \dots N$

$$f_j^0 = f_0^{\otimes j}\,, \quad (12)$$

where f_0 is a given one-particle distribution function. This means that any pair of particles are statistically uncorrelated at time zero. Of course such a statistical independence is destroyed at time $t > 0$ and Eq. (9) shows that the time evolution of f_1^N is determined by the knowledge of f_2^N which turns out to be dependent on f_3^N and so on. However, since the interaction between two given particle is going to vanish in the limit $\varepsilon \to 0$, we can hope that such statistical independence, namely the factorization property (12), is recovered in the same limit.

Therefore we expect that in the limit $\varepsilon \to 0$ the one-particle distribution function f_1^N converges to the solution

of a suitable nonlinear kinetic equation f which we are going to investigate.

If we expand $f_j^N(t)$ as a perturbation of the free flow $S(t)$ defined as

$$(S(t)f_j)(X_j, V_j) = f_j(X_j - V_j t, V_j)\,, \tag{13}$$

we find

$$f_j^N(t) = S(t)f_j^0 + \frac{N-j}{\sqrt{\varepsilon}}\int_0^t S(t-t_1)C_{j+1}^{\varepsilon} f_{j+1}^N(t_1)\mathrm{d}t_1 + \frac{1}{\sqrt{\varepsilon}}\int_0^t S(t-t_1)T_j^{\varepsilon} f_j^N(t_1)\mathrm{d}t_1\,. \tag{14}$$

We now try to keep informations on the limit behavior of $f_j^N(t)$. Assuming for the moment that the time evolved j-particle distributions $f_j^N(t)$ are smooth (in the sense that the derivatives are uniformly bounded in ε), then

$$C_{j+1}^{\varepsilon} f_{j+1}^N(X_j, V_j, t_1) = -\varepsilon^3 \sum_{k=1}^{j}\int \mathrm{d}r \int \mathrm{d}v_{j+1} F(r) \cdot \nabla_{v_k} f_{j+1}(X_j, x_k - \varepsilon r, V_j, v_{j+1}, t_1)\,. \tag{15}$$

Assuming now, quite reasonably, that

$$\int \mathrm{d}r F(r) = 0\,, \tag{16}$$

we find that

$$C_{j+1}^{\varepsilon} f_{j+1}^N(X_j, V_j, t_1) = O(\varepsilon^4)$$

provided that $D_v^2 f_{j+1}^N$ is uniformly bounded. Since

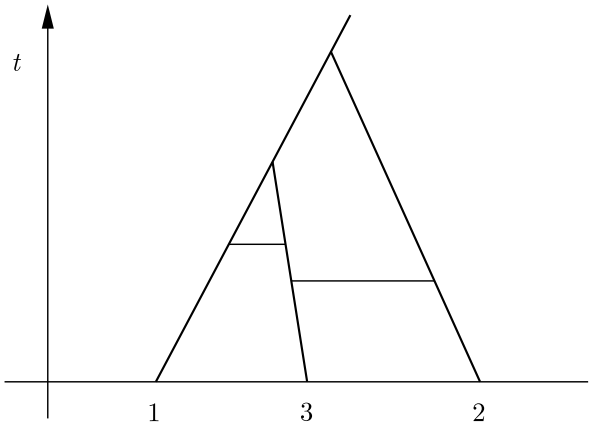
$$\frac{N-j}{\sqrt{\varepsilon}} = O(\varepsilon^{\frac{7}{2}})$$

we see that the second term in the right hand side of (14) does not give any contribution in the limit. Moreover

$$\int_0^t S(t-t_1)T_j^{\varepsilon} f_j^N(t_1)\mathrm{d}t_1 = \sum_{i\neq k}\int_0^t \mathrm{d}t_1 \cdot F\left(\frac{(x_i - x_k) - (v_i - v_k)(t-t_1)}{\varepsilon}\right) g(X_j, V_j, t_1)\,, \tag{17}$$

where g is a smooth function.

Obviously the above time integral is $O(\varepsilon)$ so that also the last term in the right hand side of (14) does not give any contribution in the limit. Then we are facing the alternative: either the limit is trivial or the time evolved distributions are not smooth. However we believe that the limit is not trivial (actually we expect to get a diffusion equation, according to the previous discussion) and a rigorous proof of this fact seems problematic.



Scaling Limits of Large Systems of Non-linear Partial Differential Equations, Figure 1
The collision sequence (1, 2), (1, 3), (1, 3), (2, 3)

The difficulty in obtaining a-priori estimates induce us to exploit the full series expansion of the solution, namely

$$f_1^N(t) = \sum_{n\geq 0}\sum_{G_n} K(G_n)\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \ldots \int_0^{t_{n-1}} \mathrm{d}t_n \cdot \left[S(t-t_1)O_1 S(t_1 - t_2)\ldots O_n S(t_n)\right] f_m^0\,. \tag{18}$$

Here O_j is either an operator C or T expressing a creation of a new particle or a recollision between two particles respectively. G_n is a graph namely a sequence of indices

$$(r_1, l_1), (r_2, l_2), \ldots (r_n, l_n)$$

where (r_j, l_j), $r_j < l_j$ are the pair of indices of the particles involved in the interaction at time t_j. The number of particles created in the process is $m-1$. It is convenient to represent the generic graph in the following way, as in Fig. 1.

The legs of the graph denotes the particles and the nodes the creation of new particles (operators C). Recollisions (operators T) are represented by horizontal link. For instance the graph in figure is $(1,2)$, $(1,3)$, $(1,3)$, $(2,3)$, $m=3$ and the integrand in Eq. (18) is in this case

$$\left[S(t-t_1)C_{1,2}S(t_1-t_2)C_{1,3}S(t_2-t_3)T_{1,3} \cdot S(t_3-t_4)T_{2,3}S(t_4)\right] f_3^0\,. \tag{19}$$

Note that the knowledge of the graph determines completely the sequence of operators in the right hand side of (20).

Finally the factor $K(G_n)$ takes into account the divergences:

$$K(G_n) = O\left(\varepsilon^{-\frac{n}{2}}\varepsilon^{-3(m-1)}\right) . \tag{20}$$

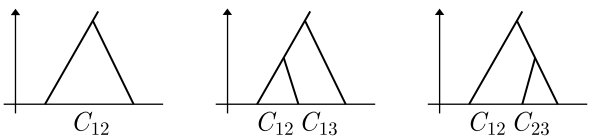
We are not able to analyze the asymptotic behavior of each term of the expansion (18) however we can compute the limit for $\varepsilon \to 0$ of the few terms up to the second order (in time). We have:

$$\begin{aligned} g^N(x_1, v_1, t) &= f^0(x_1 - v_1 t, v_1) + \frac{N-1}{\sqrt{\varepsilon}} \int_0^t S(t-t_1) \\ &\cdot C^\varepsilon_{1,2} S(t_1) f_2^0 \mathrm{d}t_1 + \frac{(N-1)}{\varepsilon}\frac{(N-2)}{\varepsilon} \sum_{j=1,2} \int_0^{t_1} \mathrm{d}t_2 \\ &\cdot S(t-t_1) C^\varepsilon_{1,2} S(t_1-t_2) C^\varepsilon_{j,3} S(t_2) f_3^0 + \frac{N-1}{\varepsilon} \int_0^t \mathrm{d}t_1 \\ &\cdot \int_0^{t_1} \mathrm{d}t_2\, S(t-t_1) C^\varepsilon_{1,2} S(t_1-t_2) T^\varepsilon_{1,2} S(t_2) f_2^0 . \end{aligned} \tag{21}$$

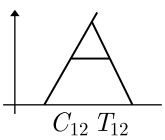
Here the right hand side of (21) defines g^N.

The second and third term in (21) corresponding to the graphs shown in Fig. 2 are indeed vanishing as follows by the use of the previous arguments.

The most interesting term is the last one, shown in Fig. 3.



Scaling Limits of Large Systems of Non-linear Partial Differential Equations, Figure 2
The vanishing terms in the expansion of g^N



Scaling Limits of Large Systems of Non-linear Partial Differential Equations, Figure 3
The first non vanishing terms in the expansion of g^N

To handle this term we denote by $w = v_1 - v_2$ the relative velocity and note that, for a given function u:

$$\begin{aligned} &S(t_1-t_2) T^\varepsilon_{1,2} u(x_1, x_2, v_1, v_2) \\ &= -F\left(\frac{(x_1-x_2) - w(t_1-t_2)}{\varepsilon}\right) \cdot [(\nabla_{v_1} - \nabla_{v_2})u] \\ &\cdot (x_1 - v_1(t_1-t_2), x_2 - v_2(t_1-t_2), v_1, v_2) \\ &= -F\left(\frac{(x_1-x_2) - w(t-t_1)}{\varepsilon}\right) \cdot (\nabla_{v_1} - \nabla_{v_2} + (t_1-t_2) \\ &\cdot (\nabla_{x_1} - \nabla_{x_2})) S(t_1-t_2) u(x_1, x_2, v_1, v_2) . \end{aligned} \tag{22}$$

Therefore the last term in the r.h.s. of (22) is

$$\begin{aligned} &\frac{N-1}{\varepsilon} \int_0^t \mathrm{d}t_1\, S(t-t_1) \int_0^{t_1} \mathrm{d}t_2 \int \mathrm{d}x_2 \int \mathrm{d}v_2 \\ &F\left(\frac{x_1-x_2}{\varepsilon}\right) \cdot \nabla_{v_1} F\left(\frac{(x_1-x_2) - w(t_1-t_2)}{\varepsilon}\right) \\ &\cdot (\nabla_{v_1} - \nabla_{v_2} + (t_1-t_2)(\nabla_{x_1} - \nabla_{x_2})) \\ &\cdot S(t_1) f_2^0(x_1, x_2, v_1, v_2) . \end{aligned} \tag{23}$$

Setting now $r = \frac{x_1-x_2}{\varepsilon}$ and $s = \frac{t_1-t_2}{\varepsilon}$ then

$$\begin{aligned} g^N(x_1, v_1, t) &= (N-1)\varepsilon^3 \int_0^t \mathrm{d}t_1 \int_0^{\frac{t_1}{\varepsilon}} \mathrm{d}s \int \mathrm{d}r \\ &\cdot \int \mathrm{d}v_2\, F(r) \cdot \nabla_{v_1} \\ &F(r - ws) \cdot (\nabla_{v_1} - \nabla_{v_2} + \varepsilon s(\nabla_{x_1} - \nabla_{x_2})) \\ &\cdot S(t_1 - \varepsilon s) f_2^0(x_1, x_2, v_1, v_2) + O(\sqrt{\varepsilon}) . \end{aligned} \tag{24}$$

The formal limit is of (21) is

$$\begin{aligned} g(x_1, v_1, t) = \int_0^t \mathrm{d}t_1 \int \mathrm{d}v_2\, S(t-t_1) \nabla_{v_1} a(v_1 - v_2) \\ \cdot (\nabla_{v_1} - \nabla_{v_2})\, S(t_1) f_2^0 , \end{aligned} \tag{25}$$

where (using $F(r) = -F(-r)$) the matrix a is given by:

$$\begin{aligned} a(w) &= \int \mathrm{d}r \int_0^{+\infty} \mathrm{d}s\, F(r) \otimes F(r - ws) = \frac{1}{2} \int \mathrm{d}r \\ &\cdot \int_{-\infty}^{+\infty} \mathrm{d}s\, F(r) \otimes F(r - ws) \frac{1}{2} \frac{1}{(2\pi)^3} \int_{-\infty}^{+\infty} \mathrm{d}s \\ &\cdot \int \mathrm{d}k\, k \otimes k\, \hat{\phi}(k)^2 \mathrm{e}^{i(w\cdot k)s} = \frac{1}{(8\pi)^2} \int \mathrm{d}k\, k \otimes k\, \hat{\phi}(k)^2 \\ &\cdot \delta(w \cdot k) = \frac{A}{|w|^3}(|w|^2 Id - w \otimes w), \end{aligned} \tag{26}$$

where $\hat{\phi}(k) = \int e^{-il\cdot x}\phi(x)dx$. Here the interaction potential ϕ has been assumed spherically symmetric and:

$$A = \frac{1}{8\pi}\int_0^{+\infty} dr\, r^3\hat{\phi}(r)^2\,. \tag{27}$$

Looking at Eq. (25) we are led to introduce the following nonlinear equation:

$$(\partial_t + v\cdot\nabla_x)f = Q_L(f,f) \tag{28}$$

with the collision operator Q_L given by:

$$Q_L(f,f)(v) = \int dv_1\cdot\nabla_v\Big[a(v-v_1)\,(\nabla_v-\nabla_{v_1})\,f(v)f(v_1)\Big]. \tag{29}$$

Here x plays the role of a parameter and hence his dependence is omitted.

Equation (28) is called the Fokker–Planck–Landau equation (Landau equation in the sequel) and has been introduced by Landau in the study of a dense, weakly interacting gas (see [32]).

From (28) we obtain the following (infinite) hierarchy of equations

$$(\partial_t + V_j\cdot\nabla_{X_j})f_j = C_{j+1}f_{j+1} \tag{30}$$

for the quantities:

$$f_j(t) = f(t)^{\otimes j} \tag{31}$$

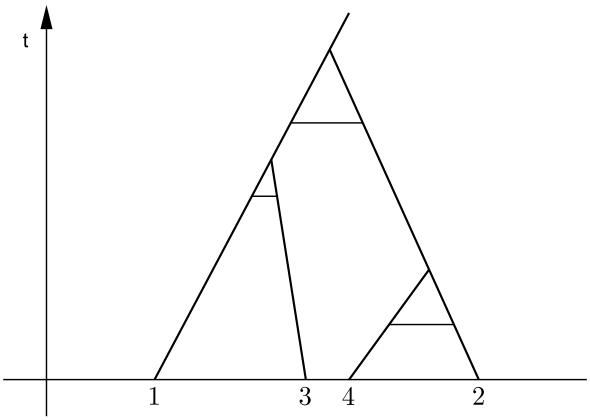
where $f(t)$ solves Eq. (28). Accordingly $C_{j+1} = \sum_k C_{k,j+1}$ where

$$\begin{aligned} &C_{k,j+1}f_{j+1}(x_1\dots x_j, v_1\dots v_j)\\ &= \prod_{r\neq k} f(x_r,v_r)Q_L(f,f)(x_k,v_k)\,. \end{aligned} \tag{32}$$

Therefore f has the following series expansion representation

$$\begin{aligned} f(t) = \sum_{n\geq 0}\int_0^t dt_1\int_0^{t_1} dt_2\,\dots\int_0^{t_{n-1}} dt_n\\ \cdot\Big[S(t-t_1)C_2S(t_1-t_2)C_3\dots C_{n-1}S(t_n)\Big]f^0_{n+1}\,. \end{aligned} \tag{33}$$

The previous calculation shows the formal convergence of g^N to the term with $n = 1$ of the expansion (33), namely we have an agreement between the particle system (18) and the solution to the Landau equation (33) at least up to the first order in time (second order in the potential). Although the above arguments can be made rigorous, under suitable assumption on the initial condition f_0 and the potential ϕ, it seems difficult to show the convergence of the whole series. On the other hand it is clear that the graphs which should contribute in the limit are those formed by a collision-recollision sequence, like that shown in Fig. 4. For those terms it is probably possible to show the convergence. For instance the case in figure has the asymptotics



Scaling Limits of Large Systems of Non-linear Partial Differential Equations, Figure 4
The collision-recollision sequence (1, 2), (1, 2), (1, 3), (1, 3), (2, 4), (2, 4)

$$\begin{aligned} \int_0^t dt_1\int_0^{t_1} dt_2\int_0^{t_2} dt_3\Big[S(t-t_1)C_{1,2}S(t_1-t_2)\\ \cdot C_{1,3}S(t_2-t_3)C_{2,4}S(t_3)\Big]f_4^0\,. \end{aligned} \tag{34}$$

However the proof that all other graphs are vanishing in the limit is not easy. Even more difficult is a uniform control of the series expansion (18), even for short times. As we shall see in the next section, something more can be obtained for quantum systems under the same scaling limit.

Some comments are in order.

In the present section we showed how the Landau equation is expected to be derived in the weak-coupling limit from a particle system. In doing this we essentially followed the monograph [2]. This is not however the usual way in which the Landau equation is introduced in the literature. Indeed it is usually recovered from the Boltzmann equation. when the density increases and the grazing collisions become dominant. In particular, the case of the homogeneous Boltzmann equation has been investigated in [1,24,44] (see also [15] for the Coulomb potential case, and see the general survey [43]). In [1] the authors show that, under suitable assumptions on the cross-section, the diffusion Fokker–Planck–Landau equation (28) can in-

deed be derived. The diffusion operator is the form (29) but with a matrix a given by

$$a(w) = \alpha(|w|)(|w|^2 Id - w \otimes w)\,,$$

with α smooth function. Next in [24] and [44] steps forward were performed to arrive to cover the case $\alpha(|w|) \approx 1/|w|^{\nu}$ for small $|w|$, with $\nu < 1$. We remark that the diffusion coefficient found in this way is different from that derived here in (26).

We conclude by observing that the Landau equation describes a genuine kinetic evolution. Mass, momentum and energy are conserved, while the kinetic entropy is decreasing. The equilibrium is Maxwellian. The difficulties in the validation problem are certainly related to the transition from a reversible (hamiltonian) dynamics to an irreversible one, as for the Boltzmann equation. Here we tried to compare the two series expansions as in the Lanford's validation proof for the Boltzmann equation. This is a sort of Cauchy–Kowalevski brute force argument which works, as well, for negative times. It is clear that, in this way, we cannot go beyond a short time result and even this seems not trivial at all. Other approaches making use of the diffusive nature of the motion of a test particle are, at moment, absent. However the weak-coupling limit of a single particle in a random distribution of scatterers is well understood as we shall discuss later in Sect. "The Weak Coupling Limit for a Single Particle: the Linear Theory".

Weak-Coupling Limit for Quantum Systems

We consider the quantum analog of the system considered in Sect. "Introduction", namely N identical quantum particles with unitary mass in $\mathbb{R}^3$. In the present section the statistical nature of the particles will be ignored.

The interaction between particles is still a two-body potential ϕ so that the total potential energy is taken as

$$U(x_1 \dots x_N) = \sum_{i<j} \phi(x_i - x_j)\,. \tag{35}$$

The associated Schrödinger equation reads

$$i\partial_t \Psi(X_N, t) = -\tfrac{1}{2}\Delta_N \Psi(X_N, t) + U(X_N)\Psi(X_N, t), \tag{36}$$

where $\Delta_N = \sum_{i=1}^N \Delta_i$, Δ_i is the Laplacian with respect to the x_i variables, $X_N = (x_1, \dots, x_N)$ and $\hbar$ is normalized to unity.

As for the classical system considered in Sect." Introduction" we rescale the equation and the potential by

$$x \to \varepsilon x\,, \qquad t \to \varepsilon t\,, \qquad \phi \to \sqrt{\varepsilon}\phi\,. \tag{37}$$

The resulting equation is,

$$i\varepsilon\partial_t \Psi^\varepsilon(X_N, t) = -\frac{\varepsilon^2}{2}\Delta_N \Psi^\varepsilon(X_N, t) + U_\varepsilon(X_N)\,\Psi^\varepsilon(X_N, t), \tag{38}$$

where:

$$U_\varepsilon(x_1 \dots x_N) = \sqrt{\varepsilon} \sum_{i<j} \phi\left(\frac{x_i - x_j}{\varepsilon}\right)\,. \tag{39}$$

We want to analyze the limit $\varepsilon \to 0$ in the above equations, when $N = \varepsilon^{-3}$.

Note that this limit looks, at a first sight, similar to a semiclassical (or high frequency) limit. It is not so: indeed the potential varies on the same scale of the typical oscillations of the wave functions so that the scattering process is a genuine quantum process. Obviously, due to the oscillations, we do not expect that the wave function does converge to something in the limit. The right quantity to look at was introduced by Wigner in 1922 [45] to deal with kinetic problems. It is called the Wigner transform (of Ψ^ε) and is defined as

$$W^N(X_N, V_N) = \left(\frac{1}{2\pi}\right)^{3N} \int \mathrm{d}Y_N\, \mathrm{e}^{iY_N \cdot V_N} \overline{\Psi}^\varepsilon \cdot (X_N + \frac{\varepsilon}{2} Y_N)\, \Psi^\varepsilon(X_N - \frac{\varepsilon}{2} Y_N)\,. \tag{40}$$

The Wigner transform W^N satisfies a transport-like equation, completely equivalent to the Schrödinger equation:

$$(\partial_t + V_N \cdot \nabla_N) W^N(X_N, V_N) = \frac{1}{\sqrt{\varepsilon}} (T_N^\varepsilon W^N)(X_N, V_N). \tag{41}$$

The operator T_N^ε on the right-hand-side of (38) plays the same role of the classical operator denoted with the same symbol in Sect. "Introduction". It is

$$(T_N^\varepsilon W^N)(X_N, V_N) = \sum_{0<k<\ell\le N} (T_{k,\ell}^\varepsilon W^N)(X_N, V_N), \tag{42}$$

where each $T_{k,\ell}^\varepsilon$ describes the interaction of particle k with particle ℓ

$$(T_{k,\ell}^\varepsilon W^N)(X_N, V_N) = \frac{1}{i}\left(\frac{1}{(2\pi)^{3N}} \int \mathrm{d}Y_N \mathrm{d}V'_N \cdot \mathrm{e}^{iY_N \cdot (V_N - V'_N)} \left[\phi\left(\frac{x_k - x_\ell}{\varepsilon} - \frac{y_k - y_\ell}{2}\right)\right) - \phi\left(\frac{x_k - x_\ell}{\varepsilon} + \frac{y_k - y_\ell}{2}\right)\right] W^N(X_N, V'_N)\,. \tag{43}$$

Equivalently, we may write

$$(T^{\varepsilon}_{k,\ell} W^N)(X_N, V_N) = -i \sum_{\sigma=\pm 1} \sigma \int \frac{\mathrm{d}h}{(2\pi)^3} \hat{\phi}(h) \cdot \mathrm{e}^{i\frac{h}{\varepsilon}(x_k - x_\ell)} W^N(x_1, \ldots, x_N; v_1, \ldots, v_k - \sigma\frac{h}{2}, \ldots, v_\ell + \sigma\frac{h}{2}, \ldots, v_N) . \quad (44)$$

Note that $T^{\varepsilon}_{k,\ell}$ is a pseudodifferential operator which formally converge, at fixed ε, for $\hbar \to 0$ (here $\hbar = 1$) to its classical analog given by Eq. (7). Note also that in (44), "collisions" may take place between *distant* particles ($x_k \neq x_\ell$). However, such distant collisions are penalized by the highly oscillatory factor $\exp(ih(x_k - x_\ell)/\varepsilon)$. These oscillations turn out to play a crucial role throughout the analysis, and they explain why collisions tend to happen when $x_k = x_\ell$ in the limit $\varepsilon \to 0$.

The formalism we have introduced is formally similar to the classical case so that we proceed as before by transforming Eq. (38) into a hierarchy of equations. We introduce the partial traces of the Wigner transform W^N, denoted by f^N_j. They are defined through the following formula, valid for $j = 1, \ldots, N-1$:

$$f^N_j(X_j, V_j) = \int \mathrm{d}x_{j+1} \ldots \int \mathrm{d}x_N \int \mathrm{d}v_{j+1} \ldots \int \mathrm{d}v_N \, W^N(X_j, x_{j+1} \ldots x_N, V_j, v_{j+1} \ldots v_N) . \quad (45)$$

Obviously, we set $f^N_N = W^N$. The function f^N_j is the kinetic object that describes the state of the j particles subsystem at time t.

As before, the wave function Ψ, as well W^N and f^N_j, are assumed to be symmetric in the exchange of particle, a property that is preserved in time.

Proceeding then as in the derivation of the BBKGY hierarchy for classical systems, we readily transform Eq. (38) into the following hierarchy:

$$(\partial_t + \sum_{k=1}^{j} v_k \cdot \nabla_k) f^N_j(X_j, V_j) = \frac{1}{\sqrt{\varepsilon}} T^{\varepsilon}_j f^N_j + \frac{N-j}{\sqrt{\varepsilon}} C^{\varepsilon}_{j+1} f^N_{j+1} , \quad (46)$$

where

$$C^{\varepsilon}_{j+1} = \sum_{k=1}^{j} C^{\varepsilon}_{k,j+1} , \quad (47)$$

and $C^{\varepsilon}_{k,j+1}$ is defined by

$$C^{\varepsilon}_{k,j+1} f^N_{j+1}(X_j, V_j) = -i \sum_{\sigma=\pm 1} \sigma \int \frac{\mathrm{d}h}{(2\pi)^3} \int \mathrm{d}x_{j+1} \cdot \int \mathrm{d}v_{j+1} \hat{\phi}(h) \, \mathrm{e}^{i\frac{h}{\varepsilon}(x_k - x_{j+1})} f^N_{j+1}\Big(x_1, x_2, \ldots, x_{j+1}, v_1, \ldots, v_k - \sigma\frac{h}{2}, \ldots, v_{j+1} + \sigma\frac{h}{2}\Big) . \quad (48)$$

As before the initial value $\{f^0_j\}_{j=1}^N$ is assumed completely factorized: for all $j = 1, \ldots, N$, we suppose

$$f^0_j = f_0^{\otimes j} , \quad (49)$$

where f_0 is a one-particle Wigner function, and f^0 is assumed to be a probability distribution.

In the limit $\varepsilon \to 0$, we expect that the j-particle distribution function $f^N_j(t)$, that solves the hierarchy (46) with initial data (49), tends to be factorized for all times: $f^N_j(t) \sim f(t)^{\otimes j}$ (propagation of chaos).

As for the classical case, if f_{j+1} is smooth:

$$C^{\varepsilon}_{k,j+1} f^N_{j+1}(X_j, V_j) = -i\varepsilon^3 \sum_{\sigma=\pm 1} \sigma \int \frac{\mathrm{d}h}{(2\pi)^3} \hat{\phi}(h) \cdot \int \mathrm{d}r \int \mathrm{d}v_{j+1} \, \mathrm{e}^{ih\cdot r} f^N_{j+1}\Big(X_j, x_k - \varepsilon r, \, v_1, \ldots, v_k - \sigma\frac{h}{2}, \ldots, v_{j+1} + \sigma\frac{h}{2}\Big) = O(\varepsilon^4) . \quad (50)$$

Indeed, setting $\varepsilon = 0$ in the integrand, the integration over r produces $\delta(h)$. As a consequence the integrand is independent on σ and the sum vanishes. Therefore the integral is $O(\varepsilon)$. Also

$$\frac{1}{\sqrt{\varepsilon}} \int_0^t \mathrm{d}t_1 \, S(t-t_1) T_{r,k} f^N_j(t_1) = -i \sum_{\sigma=\pm 1} \sigma \int_0^t \mathrm{d}t_1 \frac{\mathrm{d}h}{(2\pi)^3} \hat{\phi}(h) \mathrm{e}^{i\frac{h}{\varepsilon}\cdot(x_r - x_k) - (v_r - v_k)(t-t_1)} \cdot f^N_j(X_j - V_j(t-t_1), V_j, t_1) \quad (51)$$

is weakly vanishing, by a stationary phase argument (see [4]). Therefore, as for the classical case, we analyze

the asymptotics of the collision-recollision term (shown in Fig. 1):

$$-\frac{N-1}{\varepsilon}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}\tau_1 \cdot S(t-t_1) C_{1,2} S(t_1-\tau_1) T_{1,2} S(\tau_1) f_2^0 . \tag{52}$$

Explicitly it looks as:

$$\begin{aligned} &-\frac{N-1}{\varepsilon}\sum_{\sigma,\sigma'=\pm 1}\sigma\sigma'\int_0^t \mathrm{d}t_1\int_0^{t_1}\mathrm{d}\tau_1\int \mathrm{d}x_2\int \mathrm{d}v_2 \\ &\cdot\int\frac{\mathrm{d}h}{(2\pi)^3}\int\frac{\mathrm{d}k}{(2\pi)^3}\hat\phi(h)\,\hat\phi(k)\,\mathrm{e}^{i\frac{h}{\varepsilon}\cdot\left(x_1-x_2-v_1(t-t_1)\right)} \\ &\cdot\mathrm{e}^{i\frac{k}{\varepsilon}\cdot\left(x_1-x_2-v_1(t-t_1)-(v_1-v_2-\sigma h)(t_1-\tau_1)\right)} f_2^0\Big(x_1-v_1 t \\ &\qquad +\sigma\frac{h}{2}t_1+\sigma'\frac{k}{2}\tau_1, x_2-v_2 t_1-\sigma\frac{h}{2}t_1-\sigma'\frac{k}{2}\tau_1; \\ &\qquad v_1-\sigma\frac{h}{2}-\sigma'\frac{k}{2}, v_2+\sigma\frac{h}{2}+\sigma'\frac{k}{2}\Big) . \end{aligned} \tag{53}$$

By the change of variables:

$$t_1-\tau_1=\varepsilon s_1\,,\ (\text{i.e. } \tau_1=t_1-\varepsilon s_1)\,,\ \xi=(h+k)/\varepsilon\,, \tag{54}$$

we have

$$\begin{aligned} (54) &= -(N-1)\,\varepsilon^3\sum_{\sigma,\sigma'=\pm1}\sigma\sigma'\int_0^t \mathrm{d}t_1\int_0^{t_1/\varepsilon}\mathrm{d}s_1 \\ &\cdot\int \mathrm{d}x_2\,\mathrm{d}v_2\,\frac{\mathrm{d}\xi}{(2\pi)^3}\,\frac{\mathrm{d}k}{(2\pi)^3}\,\hat\phi(-k+\varepsilon\xi_1)\,\hat\phi(k) \\ &\cdot\mathrm{e}^{i\xi\cdot\left(x_1-x_2-v_1(t-t_1)\right)}\mathrm{e}^{-is_1 k\cdot(v_1-v_2-\sigma(-k+\varepsilon\xi))} f_{j+1}^0(\dots)\,, \end{aligned} \tag{55}$$

In the limit $\varepsilon\to 0$, the above formula gives the asymptotics

$$\begin{aligned} (52) \underset{\varepsilon\to0}{\sim} &-\sum_{\sigma,\sigma'=\pm1}\sigma\sigma'\int_0^t \mathrm{d}t_1\int \mathrm{d}v_2\,\frac{\mathrm{d}k}{(2\pi)^3}\,|\hat\phi(k)|^2 \\ &\cdot\left(\int_0^{+\infty}\mathrm{e}^{-is_1 k\cdot(v_1-v_2+\sigma k)}\mathrm{d}s_1\right) \\ &\cdot f_2^0\Big(x_1-v_1 t-(\sigma-\sigma')\frac{k}{2}t_1, \\ &x_1-v_1(t-t_1)-v_2 t_1+(\sigma-\sigma')\frac{k}{2}t_1, v_1+(\sigma-\sigma')\frac{k}{2}, \\ &v_2-(\sigma-\sigma')\frac{k}{2}\Big) . \end{aligned} \tag{56}$$

In [4], we completely justify formula (56) and its forthcoming consequences.

Now, we turn to identifying the limiting value obtained in (56). To do so, we observe that symmetry arguments allow us to replace the integral in s by its real part:

$$\mathrm{Re}\int_0^\infty \mathrm{e}^{-is_1 k\cdot(v_1-v_2+\sigma k)}\mathrm{d}s_1 = \pi\delta(k\cdot(v_1-v_2+\sigma k)). \tag{57}$$

Using this we realize that the contribution $\sigma=-\sigma'$ in (56) gives rise to the gain term:

$$\begin{aligned} &\int_0^t \mathrm{d}t_1\int \mathrm{d}v_2\int \mathrm{d}\omega\, B\big(\omega, v_1-v_2\big)\, f_2^0\big(x_1-v_1(t-t_1) \\ &\qquad -v_1' t_1, x_2-v_2(t-t_1)-v_2' t_1, v_1', v_2'\big)\,, \end{aligned} \tag{58}$$

where the integral in ω is on the surface of the unitary sphere in $\mathbb{R}^3$, and

$$B(\omega, v)=\frac{1}{8\pi^2}|\omega\cdot v|\,|\hat\phi(\omega\,(\omega\cdot v))|^2, \tag{59}$$

and the velocities v_1', v_2' are

$$v_1'=v_1-\omega(\omega\cdot(v_2-v_1))\,,\quad v_2'=v_2+\omega(\omega\cdot(v_2-v_1))\,.$$

Similarly, the term $\sigma=\sigma'$ in (5) yields the loss term:

$$\begin{aligned} &\int_0^t \mathrm{d}t_1\int \mathrm{d}v_2\int \mathrm{d}\omega\, B(\omega, v_1-v_2) \\ &\cdot f_2^0(x_1-v_1 t, x_2-v_2(t-t), v_1, v_2)\,. \end{aligned} \tag{60}$$

By the same arguments used in the previous section we can conclude that the full series expansion (20) (of course for the present quantum case) agrees, up to the second order in the potential, with

$$S(t)f_0+\int_0^t \mathrm{d}t_1\, S(t-t_1)Q(S(t_1)f_0, S(t_1)f_0) \tag{61}$$

where

$$\begin{aligned} Q(f,f)=\frac{1}{4\pi^2}\int \mathrm{d}v_1\int \mathrm{d}h\,|\hat\phi(h)|^2\delta((h\cdot(v-v_1+h)) \\ \cdot[f(v+h)f(v_1-h)-f(v)f(v_1)]=\int \mathrm{d}v_1\int \mathrm{d}\omega \\ \cdot B(\omega, v-v_1)[f'f_1'-ff_1]\,, \end{aligned} \tag{62}$$

with

$$\begin{aligned} &f_1=f(v_1)\,,\ f'=f(v')\,,\ f_1'=f(v_1')\,, \\ &v'=v+h=v-\omega(\omega\cdot(v-v_1))\,, \\ &v_1'=v_1-h=v_1+\omega(\omega\cdot(v-v_1))\,. \end{aligned} \tag{63}$$

In other words the kinetic equation which comes out is the Boltzmann equation with cross-section B:

$$(\partial_t + v \cdot \nabla_x) f = Q(f, f)\,, \tag{64}$$

where Q is given by Eq. (62). We note once more that the δ function in Eq. (62) expresses the energy conservation, while the momentum conservation is automatically satisfied. Note also that the cross-section B is the only quantum factor in the purely classical expression (62). It retains the quantum features of the elementary "collisions".

An important comment is in order. Why the kinetic equation for quantum systems is of Boltzmann type in contrast with the classical case where we got a diffusion? The answer is related to the asymptotics of a single scattering (see [36,37] and [8]). For quantum systems we have a finite probability of having any angle scattering, while for a classical particle, we surely have a small deviation from the free motion. Therefore a quantum particle, in this asymptotic regime, is going to perform a jump process (in velocity) rather than a diffusion.

From a mathematical view point we observe that [4] proves more than agreement up to second order. We indeed consider the subseries (of the full series expansion expressing $f_j^N(t)$) formed by *all* the collision–recollision terms (as that shown in Fig. 3). In other words, we consider the subseries of $f_j^N(t)$ given by

$$\sum_{n\geq 1} \sum_{r_1,\ldots,r_n,l_1,\ldots,l_n} \varepsilon^{-4n} \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}\tau_1\, S(t-t_1) C^{\varepsilon}_{r_1,l_1} \cdot S(t_1-\tau_1) T^{\varepsilon}_{r_1,l_1} \int_0^{\tau_{n-1}} \mathrm{d}t_n \int_0^{t_n} \mathrm{d}\tau_n\, S(\tau_{n-1}-t_n) C^{\varepsilon}_{r_n,l_n} \cdot S(t_n-\tau_n) T^{\varepsilon}_{r_n,l_n} S(\tau_n) f^0_{j+n+1}\,. \tag{65}$$

Here the sum runs over all possible choices of the particles number r's and l's, namely we sum over the subset of graph of the form in Fig. 4. We establish in [4] that the subseries (65) is indeed absolutely convergent, for short times, uniformly in ε. Moreover, we prove that it approaches the corresponding complete series expansion obtained by solving iteratively the Boltzmann equation with collision operator given by Eq. (62) extending and making rigorous the above argument.

Under reasonable smoothness hypotheses on the potential and on the initial distribution, assuming in addition that $\hat{\phi}(0) = 0$, we also proved in [7], that all other terms than those considered in the subseries (65) are indeed vanishing in the limit. The condition $\hat{\phi}(0) = 0$ is probably only technical: it takes care automatically of some divergences which are difficult to deal with differently.

Unfortunately this result, even under these severe assumptions on the potential, is not yet conclusive because we are not able to show a uniform bound on the full series. Thus a mathematical justification of the quantum Boltzmann equation is a still an open and difficult problem.

It is not surprising that we know more for the quantum case in comparison with the classical problem introduced in Sect. "Introduction". Now the operators involved are differences while, for the classical case, we have to deal with derivatives. Introducing explicitly the $\hbar$ dependence, we easily realize that the results discussed in the present section are not uniform in $\hbar$.

Weak-Coupling Limit in the Bose–Einstein and Fermi–Dirac Statistics

In this section we approach the same problem as in Sect. "Weak-Coupling Limit for Classical Systems", for Bosons or Fermions, namely for particles obeying the Fermi–Dirac or Bose–Einstein statistics. As we shall see, the kinetic equation we expect to hold in the limit, is the so called Uehling and Uhlembeck equation, which is a Boltzmann type equation, with cubic collision operator. We note that the effects of the quantum correction here, enter also in the structure of the operator.

In this case, the starting point is still the rescaled Schrödinger equation (38), or the equivalent hierarchy (46). The only new point is that we cannot take a totally decorrelated initial datum as in (49). Indeed, the Fermi–Dirac or Bose–Einstein statistics yield correlations even at time zero. In this perspective, the most uncorrelated states one can introduce, and that do not violate the Fermi–Dirac or Bose–Einstein statistics, are the so called quasi-free states. They have, in terms of the Wigner formalism, the following form

$$f_j(x_1, v_1, \ldots, x_j, v_j) = \sum_{\pi \in \mathcal{P}_j} \theta^{s(\pi)}\, f_j^{\pi}(x_1, v_1, \ldots, x_j, v_j)\,, \tag{66}$$

where each f_j^{π} has the value

$$f_j^{\pi}(x_1, v_1, \ldots, x_j, v_j) = \int \mathrm{d}y_1 \ldots \mathrm{d}y_j \int \mathrm{d}w_1 \ldots \mathrm{d}w_j \cdot \mathrm{e}^{i(y_1\cdot v_1 + \cdots + y_j \cdot v_j)} \prod_{k=1}^{j} \mathrm{e}^{-\frac{i}{\varepsilon} w_k \cdot (x_k - x_{\pi(k)})} \cdot \mathrm{e}^{-\frac{i}{2} w_k \cdot (y_k + y_{\pi(k)})} f\left(\frac{x_k + x_{\pi(k)}}{2} + \varepsilon \frac{y_k - y_{\pi(k)}}{4}, w_k\right), \tag{67}$$

and f is a given one-particle Wigner function. Here $\mathcal{P}_j$ denotes the group of all the permutations of j objects and π its generic element.

Note that the uncorrelated case treated in Sect. "Weak-Coupling Limit for Classical Systems", is recovered by the contribution due the permutation $\pi =$ identity.

To verify that (66) and (67) really describe admissible states for the statistics, consider the inverse Wigner transform of (66) and (67). The density matrix obtained in this way does verify the permutation invariance required by the statistics.

Moreover the quasi-free states converge weakly to the completely factorized states as $\varepsilon \to 0$. This is physically obvious because the quantum statistics become irrelevant in the semiclassical limit. However the dynamics take place on the scale ε so that the effects of the statistics are present in the limit. Indeed it is expected that the one-particle distribution function $f_1^N(t)$ converges to the solution of the following cubic Boltzmann equation:

$$(\partial_t + v \cdot \nabla_x) f(x, v, t) = Q_\theta(f)(x, v, t) , \tag{68}$$

$$\begin{aligned} Q_\theta(f)(x, v, t) &= \int \mathrm{d}v_1 \, \mathrm{d}\omega \, B_\theta(\omega, v - v_1) \\ &\cdot \big[f(x, v') f(x, v_1')(1 + 8\pi^3 \theta f(x, v))(1 + 8\pi^3 \theta f(x, v_1)) \\ &- f(x, v) f(x, v_1)(1 + 8\pi^3 \theta f(x, v')(1 + 8\pi^3 \theta f(x, v_1'))\big] , \end{aligned} \tag{69}$$

(for the notations see Eq. (63)). Here $\theta = +1$ or $\theta = -1$, for the Bose–Einstein or the Fermi–Dirac statistics respectively. Finally, B_θ is the symmetrized or antisymmetrized cross-section derived from B (see (59)) in a natural way:

$$\begin{aligned} B_\theta(\omega, v) = \frac{1}{16\pi^2} |\omega \cdot v| \\ \cdot \big[\hat{\phi}(\omega \, (\omega \cdot v)) + \theta \hat{\phi}(v - \omega \, (\omega \cdot v)) \big]^2 . \end{aligned}$$

As we see, the modification of the statistics transforms the quadratic Boltzmann equation of the Maxwell–Boltzmann case, into a cubic one (fourth order terms cancel). Also, the statistics affects the form of the cross-section and B has to be (anti)symmetrized into B_θ.

The collision operator (69) has been introduced by Nordheim in 1928 [38] and by Uehling–Uhlenbeck in 1933 on the basis of purely phenomenological considerations [42].

Plugging in the hierarchy (46) an initial datum satisfying (66), we can follow the same procedure as for the Maxwell–Boltzmann statistics, namely we write the full perturbative series expansion expressing $f_j^N(t)$ in terms of the initial datum and try to analyze its asymptotic behavior.

As we did before, we first restrict our attention to those terms of degree less than two in the potential.

The analysis up to second order is performed in [5]. We actually recover here Eq. (68), (69) with the suitable B_θ. Now the number of terms to control is much larger due to the sum over all permutations that enters the definition (66) of the initial state. Also, the asymptotics is much more delicate. In particular, we stress the fact that the initial datum brings its own highly oscillatory factors in the process, contrary to the Maxwell–Boltzmann case where the initial datum is uniformly smooth, and where the oscillatory factors simply come from the collision operators T and C.

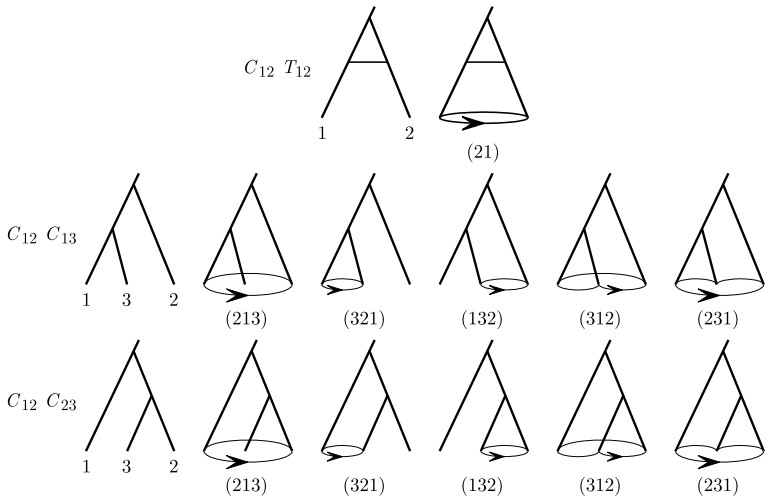
In [5] we consider the graphs of second order in $\hat{\phi}$, show in Fig. 5, which, because of the permutation of initial state, yields various terms: two of them are bilinear in the initial condition f_0 ($C_{12}T_{12}$), and give in the limit the part of Q_θ quadratic in f, twelve are cubic in f_0($C_{12}C_{13}$, $C_{12}C_{23}$). Due to a non-stationary phase argument the two terms with permutation $\pi = id = (123)$ vanish, as that the term with $\pi = (321)$ for $C_{12}C_{13}$, and the term with $\pi = (132)$ for $C_{12}C_{23}$. The two terms with $\pi = (321)$ give rise to truly diverging contributions (negative powers of ε), however their sum is seen to cancel. Last, permutations without fixed point ($\pi = (312)$, (231)) and $\pi = (132)$ for $C_{12}C_{13}$, and $\pi = (321)$ for $C_{12}C_{23}$, give the part of Q_θ cubic in f. This ends up the analysis of terms up to second order in the potential.

The computation is heavy and hence we address the reader to [5] for the details.

We mention that a similar second order analysis, using commutator expansions in the framework of the second quantization formalism, has been performed in [27] (following [26]) in the case of the van Hove limit for lattice systems (that is the same as the weak-coupling limit, yet without rescaling the distances). For more recent formal results in this direction, but in the context of the weak-coupling limit, we also quote [21].

We finally observe that the initial value problem for Eq. (68) is somehow trivial for Fermions. Indeed we have the a priori bounds $f \leq 1/(8\pi)^3$ making everything easy. For Bosons the situation is much more involved even for the spatially homogeneous case. The statistics favor large value of f and it is not clear whether the equation can explain dynamical condensation. See, for the mathematical side, [33,34].

Under the weak coupling limit we derived two very different equations, the Fokker–Planck–Landau equation in the classical case, and a Boltzmann equation with a quantum cross section in the quantum case. For the original large particle system a classical or a quantum description



Scaling Limits of Large Systems of Non-linear Partial Differential Equations, Figure 5
The three graphs of second order with the statistical permutations

depends on the value of the physical scales of the macroscopic variables with respect to $\hbar$. The same happen for the two asymptotic equations. The dependence on $\hbar$ of the collision term (69) is given by

$$Q_{UU}(f)(v) = \frac{1}{16\pi^2\hbar^4}\int \mathrm{d}v_1 \int \mathrm{d}\omega \left[\hat{\phi}\left(\tfrac{\omega\,(\omega\cdot w)}{\hbar}\right) + \theta\hat{\phi}\left(\tfrac{w-\omega\,(\omega\cdot w)}{\hbar}\right)\right]^2 \{(1 \pm (2\pi\hbar)^3 f)(1 \pm (2\pi\hbar)^3 f_1) \cdot f' f'_1 - (1 \pm (2\pi\hbar)^3 f')(1 \pm (2\pi\hbar)^3 f'_1) f f_1\}\,. \tag{70}$$

The quantum-cross section term $\hat{\phi}(\omega\,(\omega\cdot(v-v_1))/\hbar)$ make $\omega\cdot(v-v_1) = O(\hbar)$, so that the collision operator concentrates on grazing collisions. In this sense, the classical limit $\hbar \to 0$ for (70) is a natural grazing collision limit. More formally, the operator (70) tends to the operator Q_L given in Eq. (29) (see [9]); nevertheless no results are available for the limit of the solutions of the equations.

We conclude this section with some consideration on the low-density limit.

In the classical case the low-density limit (or the Boltzmann–Grad limit) yields the usual Boltzmann equation for classical systems and this result has been proved for short times [31]. It is natural to investigate what happens, in the same scaling limit, to a quantum system. Here, due to the fact that the density is vanishing, the particles are too rare to make the statistical correlations effective. As a consequence, we expect that the Bose–Einstein, and Fermi–Dirac statistics, as well fully uncorrelated states, all give rise to the same Boltzmann equation along the low-density limit with the true quantum cross-section, given by the sum of the Born series, under smallness assumption for the potential. The analysis of the partial series of the dominant terms (uniform bounds and convergence as for the weak-coupling limit) has been performed in [6].

Weak-Coupling Limit for a Single Particle: The Linear Theory

Consider the time evolution of a single particle, in the low-density or in the weak-coupling regime, under the action of a random configuration of obstacles $\mathbf{c} = \{c_1 \dots c_N\} \subset \mathbb{R}^{3N}$.

More precisely, after the rescaling, the basic equations are:

$$\dot{x}(t) = v(t)\,, \quad \dot{v}(t) = -\sum_j \nabla\phi_\varepsilon(x(t) - c_j) \tag{71}$$

for a classical particle and, for a quantum particle,

$$i\varepsilon\partial_t\psi = -\frac{\varepsilon^2}{2}\Delta\psi + \sum_j \phi_\varepsilon(x(t) - c_j)\psi\,, \tag{72}$$

where $\phi_\varepsilon = \phi(\frac{x}{\varepsilon})$ and $\phi_\varepsilon = \sqrt{\varepsilon}\phi(\frac{x}{\varepsilon})$ for the low-density and weak-coupling limits respectively. Here ϕ denotes a given smooth potential.

We are interested in the behavior of

$$f_\varepsilon(x, v, t) = \mathbb{E}[f_{\mathbf{c}}(x, v, t)] \tag{73}$$

where $f_{\mathbf{c}}(t)$ is the time evolved classical distribution function or the Wigner transform of ψ according to Eq. (71)

or (72) respectively. Finally $\mathbb{E}$ denotes the expectation with respect to the obstacle distribution, for which, a natural choice could be the Poisson distribution of density μ_ε which is also scaled according to

$$\mu_\varepsilon = \varepsilon^{-2}\,, \quad \mu_\varepsilon = \varepsilon^{-3}$$

for the low-density and weak-coupling respectively.

For the low-density scaling (this is the so called Lorentz model), we obtain, for classical systems, a linear Boltzmann equation (see [3,13,17,23,40]). It is also known that the system does not homogenize to a jump process given by a linear Boltzmann equation in case of a periodic distribution of obstacles [10]. For recent results in this direction see [11].

For the weak-coupling limit of a classical particle we obtain a linear Landau equation as it is shown in [29] and [16].

Let us spend a few words on these results. Each solution $(x(t), y(t)) = (x_{\mathbf{c}}(t), v_{\mathbf{c}}(t))$ of Eq. (71) is a sample of a stochastic process. Looking at the behavior of $v_{\mathbf{c}}(t)$ we suddenly realize that it does not enjoy the Markov property because, once an obstacle produces a change in velocity, we know its existence and this creates time correlations with the future behavior of the particle. However if a recollision with the same obstacle is an unlikely event, we could consider the velocity change due to the collisions, as independent and Poisson distributed (as the obstacles) random variables. In this case it is not difficult to prove the convergence to a diffusion process, solution to the following stochastic differential equation

$$\begin{aligned} \mathrm{d}x &= v\mathrm{d}t\,, \\ \mathrm{d}v &= -2A\tfrac{v}{|v|^3}\mathrm{d}t + \sqrt{\tfrac{2A}{|v|}}\left(I - \tfrac{v\otimes v}{|v|^2}\right)\mathrm{d}w\,, \end{aligned} \tag{74}$$

where $\mathrm{d}w$ is a brownian motion, A is the constant given by Eq. (27), and $\sqrt{\frac{2A}{|v|}}\cdot\left(I - v\otimes v/|v|^2\right)$ is the matrix $\sqrt{2a}$, with a given by Eq. (26). Notice that each collision preserves the energy so that the above process lives on the surface of the sphere $|v| = \text{const}$.

Therefore the main point in the proof of the weak-coupling limit is to show that, in the same limit, the process under consideration is stochastically equivalent to a Markov process with the right properties. Moreover, as a consequence, we have also that $f_\varepsilon(t) \to f(t)$ being f solution to the following linear Landau equation

$$(\partial_t + v\cdot\nabla_x)\, f = \mathrm{div}_v[a(v)\nabla_v f]\,, \tag{75}$$

Alternatively we could proceed differently, in the same (Cauchy–Kovalevskii) spirit of our previous analysis for the nonlinear case. Namely we first represent the time evolved distribution $f_{\mathbf{c}}(t)$ in terms of a series expansion. Then, taking the expectation of each term of the series, we exploit its limit by using the same arguments as in Sect. "Introduction". Finally the convergence of the series (uniformly in ε) should also be proven. This is certainly a possible program which is, however, conceptually and technically weaker than that based on the control of the particle trajectory illustrated above. Indeed the absolute control of the series expansion yielding the solution to Eq. (75) does not use the positivity of the diffusion coefficient so that, at best, we could recover the result for a short time only and in spaces of analytic functions, not at all natural in our context.

This remark applies as well to the nonlinear case for which, however, we do not have any result, even for short times.

As regards the corresponding weak-coupling quantum problem, the easiest case is when ϕ is a Gaussian process. The kinetic equation is still a linear Boltzmann equation. The first result, holding for short times, has been obtained in [41] (see also [30]). More recently this result has been extended to arbitrary times [22]. The technique of [22] can be also applied to deal with a Poisson distribution of obstacles [12] Obviously the cross section appearing in the Boltzmann equation is the one computed in the Born approximation. Finally in [20] the low-density case has been successfully approached. The result is a linear Boltzmann equation with the full cross-section.

Future Directions

The program of obtaining macroscopic equations from the microscopic dynamics is an ambitious and difficult problem. It arised in 1900 with the Hilbert's speech at the congress of mathematicians in Paris, as the sixth problem. After more than hundred years, such a program is far to be achieved.

It is probably true that new ideas and techniques are needed. As we said before in discussing the linear problems, it may be quite possible that classical ideas in the field of partial differential equations are not enough to approach this kind of problems successfully. On the other hand a deeper and more detailed analysis of the particle dynamics is technically difficult and philosophically paradoxical. Indeed from the practical side we introduce a partial differential equation to reduce the difficulties of the study of particle evolution, while its justification may require a deeper control of the underlying microscopic dynamics.

Bibliography

Primary Literature

1. Arseniev AA, Buryak OE (1990) On a connection between the solution of the Boltzmann equation and the solution of the Landau–Fokker–Planck equation (Russian). Mat Sb 181(4):435–446 (translation in Math USSR-Sb 69(2):465–478 1991)
2. Balescu R (1975) Equilibrium and Nonequilibrium Statistical Mechanics. Wiley, New York
3. Boldrighini C, Bunimovich LA, Ya Sinai G (1983) On the Boltzmann equation for nthe Lorentz gas. J Stat Phys 32:477–501
4. Benedetto D, Castella F, Esposito R, Pulvirenti M (2004) Some Considerations on the derivation of the nonlinear Quantum Boltzmann Equation. J Stat Phys 116(114):381–410
5. Benedetto D, Castella F, Esposito R, Pulvirenti M (2005) On The Weak–Coupling Limit for Bosons and Fermions. Math Mod Meth Appl Sci 15(12):1–33
6. Benedetto D, Castella F, Esposito R, Pulvirenti M (2006) Some Considerations on the derivation of the nonlinear Quantum Boltzmann Equation II: the low-density regime. J Stat Phys 124(2–4):951–996
7. Benedetto D, Castella F, Esposito R, Pulvirenti M (2008) From the N-body Schrödinger equation to the quantum Boltzmann equation: a term-by-term convergence result in the weak coupling regime. Commun Math Phys 277(1):1–44
8. Benedetto D, Esposito R, Pulvirenti M (2004) Asymptotic analysis of quantum scattering under mesoscopic scaling. Asymptot Anal 40(2):163–187
9. Benedetto D, Pulvirenti M (2007) The classical limit for the Uehling–Uhlenbeck operator. Bull Inst Math Acad Sinica 2(4):907–920
10. Burgain J, Golse F, Wennberg B (1998) On the distribution of free path lenght for the periodic Lorentz gas. Comm Math Phys 190:491–508
11. Caglioti E, Golse F (2003) On the distribution of free path lengths for the periodic Lorentz gas. III Comm Math Phys 236(2):199–221
12. Chen T (2005) Localization lengths and Boltzmann limit for the Anderson model at small disorders in dimension 3. J Stat Phys 120(1–2):279–337
13. Caglioti E, Pulvirenti M, Ricci V (2000) Derivation of a linear Boltzmann equation for a periodic Lorentz gas. Mark Proc Rel Fields 3:265–285
14. Cercignani C, Illner R, Pulvirenti M (1994) The mathematical theory of dilute gases, Applied Mathematical Sciences, vol 106. Springer, New York
15. Degond P, Lucquin–Desreux B (1992) The Fokker–Planck asymptotics of the Boltzmann collision operator in the Coulomb case. Math Models Methods Appl Sci 2(2):167–182
16. Dürr D, Goldstain S, Lebowitz JL (1987) Asymptotic motion of a classical particle in a random potential in two dimension: Landau model. Comm Math Phys 113:209–230
17. Desvillettes L, Pulvirenti M (1999) The linear Boltzmann eqaution for long-range forces: a derivation for nparticle systems. Math Moduls Methods Appl Sci 9:1123–1145
18. Di Perna RJ, Lions PL (1989) On the Cauchy problem for the Boltzmann equatioin. Ann Math 130:321–366
19. Esposito R, Pulvirenti M (2004) From particles to fluids. In: Friedlander S, Serre D (eds) Handbook of Mathematical Fluid Dynamics, vol 3. Elsevier, North Holland, pp 1–83
20. Eng D, Erdös L (2005) The linear Boltzmann equation as the low-density limit of a random Schrödinger equation. Rev Math Phys 17(6):669–743
21. Erdös L, Salmhofer M, Yau HT (2004) On the quantum Boltzmann equation. J Stat Phys 116:367–380
22. Erdös L, Yau HT (2000) Linear Boltzmann equation as a weak-coupling limit of a random Schrödinger equation. Comm Pure Appl Math 53:667–735
23. Gallavotti G (1972) Rigorous theory of the Boltzmann equation in the Lorentz gas in Meccanica Statistica. reprint Quaderni CNR 50:191–204
24. Goudon T (1997) On Boltzmann equations and Fokker–Planck asymptotics: influence of grazing collisions. J Statist Phys 89(3–4):751–776
25. Grad H (1949) On the kinetic Theory of rarefied gases. Comm Pure Appl Math 2:331–407
26. Hugenholtz MN (1983) Derivation of the Boltzmann equation for a Fermi gas. J Stat Phys 32:231–254
27. Ho NT, Landau LJ (1997) Fermi gas in a lattice in the van Hove limit. J Stat Phys 87:821–845
28. Illner R, Pulvirenti M (1986) Global Validity of the Boltzmann equation for a two-dimensional rare gas in the vacuum. Comm Math Phys 105:189–203 (Erratum and improved result, Comm Math Phys 121:143–146)
29. Kesten H, Papanicolaou G (1981) A limit theorem for stochastic acceleration. Comm Math Phys 78:19–31
30. Landau LJ (1994) Observation of quantum particles on a large space-time scale. J Stat Phys 77:259–309
31. Lanford III O (1975) Time evolution of large classical systems. In: Moser EJ (ed) Lecture Notes in Physics 38. Springer, pp 1–111
32. Lifshitz EM, Pitaevskii LP (1981) Course of theoretical physics "Landau–Lifshit", vol 10. Pergamon Press, Oxford-Elmsford
33. Xuguang LU (2005) The Boltzmann equation for Bose–Einstein particles: velocity concentration and convergence to equilibrium. J Stat Phys 119(5–6):1027–1067
34. Xuguang LU (2004) On isotropic distributional solutions to the Boltzmann equation for Bose–Einstein particles. J Stat Phys 116(5–6):1597–1649
35. CB Jr Morrey (1955) On the derivation of the equations of hydrodynamics from statistical mechanics. Comm Pure Appl Math 8:279–326
36. Nier F (1996) A semi-classical picture of quantum scattering. Ann Sci Ec Norm Sup 29(4):149–183
37. Nier F (1995) Asymptotic analysis of a scaled Wigner equation and quantum scattering. Transp Theory Statist Phys 24(4–5):591–628
38. Nordheim LW (1928) On the Kinetic Method in the New Statistics and Its Application in the Electron Theory of Conductivity. Proc Royal Soc Lond Ser A 119(783):689–698
39. Spohn H (1991) Large scale dynamics of interacting particles, Texts and monographs in physics. Springer
40. Spohn H (1978) The Lorentz flight process converges to a random flight process. Comm Math Phys 60:277–290
41. Spohn H (1977) Derivation of the transport equation for electrons moving through random impurities. J Stat Phys 17:385–412
42. Uehling EA, Uhlembeck GE (1933) Transport Phenomena in Einstein–Bose and Fermi–Dirac Gases. Phys Rev 43:552–561
43. Villani C (2002) A review of mathematical topics in collisional kinetic theory. In: Friedlander S, Serre D (eds) Handbook of

Mathematical Fluid Dynamics, vol 1. Elsevier, North Holland, pp 71–307
44. Villani C (1998) On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations. Arch Rational Mech Anal 143(3):273–307
45. Wigner E (1932) On the quantum correction for the thermodynamical equilibrium. Phys Rev 40:742–759

Books and Reviews

Balescu R (1975) Equilibrium and Nonequilibrium Statistical Mechanics. Wiley, New York
Cercignani C, Boltzmann L (1998) The man who trusted atoms. Oxford University Press, Oxford
Cercignani C, Illner R, Pulvirenti M (1994) The mathematical theory of dilute gases. Applied Mathematical Sciences, vol 106. Springer, New York
Esposito R, Pulvirenti M (2004) From particles to fluids. In: Friedlander S, Serre D (eds) Handbook of Mathematical Fluid Dynamics, vol 3. Elsevier, North Holland, pp 1–83
Spohn H (1991) Large scale dynamics of interacting particles, Texts and monographs in physics. Springer, Heidelberg
Villani C (2002) A review of mathematical topics in collisional kinetic theory. In: Friedlander S, Serre D (eds) Handbook of Mathematical Fluid Dynamics, vol 1. Elsevier, North Holland, pp 71–307

Scaling Properties, Fractals, and the Renormalization Group Approach to Percolation

DIETRICH STAUFFER
Institute for Theoretical Physics, Cologne University, Köln, Germany

Article Outline

Glossary
Definition of the Subject
Introduction
Methods
Quantities and Exponents
Fractal Dimension; Incipient Infinite Cluster
Simple Renormalization Group
Future Directions
Bibliography

Glossary

Cluster Clusters are sets of occupied neighboring sites.

Critical exponent At a critical point or second-order phase transition, many quantities diverge or vanish with a power law of the distance from this critical point; the critical exponent is the exponent for this power law.

Fractals Fractals have a mass varying with some power of their linear dimension. The exponent of this power law is called the fractal dimension and is smaller than the dimension of the space.

Percolation Each site of a large lattice is randomly occupied or empty.

Renormalization A cell of several sites, atoms, or spins is approximated by one single site etc. At the critical point, these supersites behave like the original sites, and the critical point thus is a fixed point of the renormalization.

Definition of the Subject

Percolation theory mostly deals with large lattices where every site is randomly either occupied or empty. In particular it studies the resulting clusters which are sets of neighboring occupied sites.

Introduction

Paul Flory, who later got the Chemistry Nobel prize, published in 1941 the first percolation theory [1], to describe the vulcanization of rubber [2]. Others later applied and generalized it, in particular by dealing with percolation theory on lattices and by studying it with computers. Most of the theory presented here was known around 1980, though in the case of computer simulation with less accuracy than today. But on the questions of universality, of critical spanning probability and of the uniqueness of infinite clusters, the 1990's have shown some of our earlier opinions to be wrong. And even today it is questioned by some that the critical exponents of percolation theory can be applied to real polymer gelation, the application which Flory had in mind two-thirds of a century ago.

On a large lattice we assume that each site independently and randomly is occupied with probability p and empty with probability $1 - p$. Depending on applications, also other words can be used instead of occupied and empty, e. g. Republican and Democrat for the majority party in an electoral district of the USA. A *cluster* is now defined as set of occupied neighboring sites. Percolation theory deals with the number and structure of these clusters, as a function of their size s, i. e. of the number s of occupied sites in the cluster. In particular it asks whether an infinite cluster spans from one side of the lattice to the opposite side. Alternatively, and more naturally if one wants to describe chemical reactions for rubber vulcanization, this site percolation can be replaced by bond percolation, where every site is occupied but the link between neighboring sites is either present with probability p or absent with probability $1 - p$, again independently and ran-

Scaling Properties, Fractals, and the Renormalization Group Approach to Percolation, Table 1
Site and bond percolation thresholds for one dimension, three two-dimensional, four three-dimensional and four hypercubic lattices in higher dimensions [1,3]

p_c	Site	Bond
$d = 1$ chain	1	1
honeycomb	.697043	$1 - 2\sin(\pi/18)$
square	.592746	1/2
triangular	1/2	$2\sin(\pi/18)$
diamond	.4301	.3893
SC	.311608	.248813
BCC	.245691	.180287
FCC	.199236	.120163
$d = 4$ hypercubic	.196885	.160131
$d = 5$ hypercubic	.140797	.118172
$d = 6$ hypercubic	.109018	.094202
$d = 7$ hypercubic	.088951	.078675

domly for each link. A cluster is now a set of neighboring sites connected by links, and the size s of the cluster can be counted as the number of links, or as the number of sites, in that cluster. Because of this ambiguity we discuss here mainly site percolation; bond percolation is similar in the sense that it belongs to the same universality class (same critical exponents). One may also combine both choices and study site-bond percolation where each site is randomly occupied or empty, and where each bond between neighboring occupied sites is randomly present or absent.

Neither temperature nor quantum effects enter this standard percolation model, which is purely geometrical probability theory. However, to understand why percolation works the way it does it is helpful to understand thermal phase transitions like the vapor-liquid critical point; and for magnetic applications it is useful to know that some spins (atomic magnetic moments) have only two states, up or down, according to quantum mechanics. We will explain these physics aspects later.

For small p, most of the occupied sites are isolated $s = 1$, coexisting with only few pairs $s = 2$ and triplets $s = 3$. For large p, most of the occupied sites form one "infinite" cluster spanning the lattice from left to right, with a few small isolated holes in it. Thus there exists one percolation threshold p_c such that for $p < p_c$ we have no spanning cluster and for $p > p_c$ we have (at least) one spanning cluster. Inspite of decades of research in this seemingly simple problem, no exact solution for p_c is proven or guessed for site percolation on the square lattice with nearest-neighbor bonds; only numerically we know it to be about 0.5927462. For site percolation on the triangular lattice or bond percolation on the square lattice, $p_c = 1/2$ exactly. More thresholds are given in Table 1 [1]. They are valid in the limit of $L \to \infty$ for lattices with about L^d sites in d dimensions. For small L instead of a sharp transition at p_c one has a rounded changeover: with a very low probability one chain of L occupied sites at $p = 1/L^{d-1}$ spans from left to right. In one dimension, a small chain can easily be spanned if p is close to one, but for $L \to \infty$ the threshold approaches $p_c = 1$ since at smaller p a hole will appear about every $1/(1 - p)$ sites and prevent any cluster to span.

Methods

This section summarizes some of the methods employed to find percolation properties, first by pencil and paper, and then with the help of computers for which Fortran programs are published e. g. in [4,5]. More details on simulations are reviewed by Ziff in this percolation part of this encyclopedia.

Mean Field Limit

The Bethe lattice or Cayley tree neglects all cyclic links and allows a solution with paper and pencil. We start from one central site, and let z bonds emanate from that. At the end of each bond sits a neighbor. Then from each of these neighbors again z bonds emanate, one back to the central site and $z - 1$ to new sites further outward. They in turn lead again each to $z - 1$ new sites, and so on. None of the newly added sites agrees with one of the already existing sites, and so we can travel along the bonds only outwards or back, but never in a loop. It is quite plausible that an infinite cluster of bond percolation is formed if each site leads to at least one more outward site along an existing bond, that means if $(z - 1)p > 1$. This condition also holds for site percolation. Thus

$$p_c = 1/(z - 1)\,. \tag{1}$$

In this way Flory calculated the threshold and other percolation properties. Today we call this the "mean field" universality class in analogy with thermal phase transitions. The critical exponents, to be discussed below, are integers or simple fractions. To this universality class belong also the Erdös-Rényi random graphs, where we connect in an assembly of N points each pair with a low probability $\propto 1/N$. And the same universality class is reached if we let the dimension d of the hypercubic lattice go to infinity (or at least take it above 6). A disadvantage of the Bethe lattice is its lack of realism: If the length of the bonds is constant, then the exponential increase of the number of sites and bonds with increasing radius leads to an infinite density.

Small Clusters

The probability of a site to be an isolated $s = 1$ cluster on the square lattice is $n_1 = p(1-p)^4$ since the site must be occupied and all its four neighbors be empty. The formula for pairs is $n_2 = 2p^2(1-p)^6$ since the pair can be oriented horizontally or vertically, resulting in the factor 2. Similar, only more difficult, is the evaluation of n_s with a maximum s usually 10 to 20; the general formula is

$$n_s = \sum_t g_{st}\, p^s\, (1-p)^t \qquad (2)$$

where the perimeter t is the number of empty neighbors and g_{st} is the number of configurations (or lattice animals, or polyominoes) of size s and perimeter t. The King's College group in London published these results decades ago. With techniques borrowed from series expansions near thermal critical phenomena, these polynomials allow to estimate not only p_c but also many other quantities (see below) diverging or vanishing near p_c.

Leath Cluster Growth

In the cluster growth method of Leath (1976) one starts with one occupied site in the center of the lattice. Then a cluster is grown by letting each empty neighbor of an already occupied cluster site decide once and for all, whether it is occupied or empty. One needs to keep and to update a perimeter list of undecided neighbors. If that list becomes empty, the cluster growth is finished, and no boundary effects of the lattice influence this cluster. If, on the other hand, the cluster reaches the lattice boundary, one has to stop the simulation and can regard this cluster as spanning (from the center to one of the sides). Repeating many times this growth simulation one can estimate p_c as well as the cluster numbers. More precisely, the cluster statistics obtained in this way is not n_s but $n_s s$ since the original center site belongs with higher probability to a larger than to a smaller cluster.

Hoshen-Kopelman Labelling

To go regularly through a large lattice, which may even be an experimentally observed structure to be analyzed by computer, one could number consecutively each seemingly new cluster, and if no clusters merge later then one has a clear classification: All sites belonging to the first cluster have label 1, all sites of the second cluster have label 2, etc. Unfortunately, this does not work. In the later analysis it may turn out that two clusters which at first seemed separate actually merge and form one cluster:

```
* *     * *        1 1     2 2
  *   * * *          1   3 ? 2
* * * * *          4 ? ? x ?
```

Already in the simple structure shown on the left we have several such label conflicts. The labels to the right come from going through the lattice like a typewriter, from left to right, and after each line to the lower line. When we come to the right neighbor of the 3 we see that 3 is really part of the cluster with label 2. And at the right neighbor of 4 we see that 4 belongs to cluster 1. The stupid method is to go back and to relabel all 3 into 2, and all 4 into 1. If then we come to the site marked with x we see that the whole structure is really one single cluster, and thus all labels 2 have to be relabeled into a 1. This is inefficient for large lattices. Instead, Hoshen and Kopelman (1976) gave each site label $m = 1, 2, 3, \ldots$ another index $n(m)$. This label $n(m)$ of labels equals its argument, $n(m) = m$, if it is still a good "root label", and it equals another number k is the cluster with initial label m later turned out to be part of an earlier cluster k. By iterating the command `m = n(m)` until finally the new m equals $n(m)$ one finds this root label. For the above we make the following assignments and re-assignments to n: $n(1) = 1$, $n(2) = 2$, $n(3) = 3$, $n(3) = 2$, $n(4) = 4$, $n(4) = 1$, $n(2) = 1$. Clusters are now characterized by the same root label for all their labels.

An advantage if this method is that only one line of the square lattice, or one hyperplane of the d-dimensional lattice, needs to be stored at any time, besides the array $n(m)$. And that array can also be reduced in size by regular recycling no longer used labels n, just as beer bottles can be recycled. Lattices with more than 10^{13} sites were simulated, using parallel computers. However, understanding the details of the algorithms and finding errors in them can be very frustrating.

Sometimes one wants to determine the cluster numbers for numerous different p from 0 to 1. Instead of starting a new analysis for each different p one may also fill the lattice with new sites, and make the proper labeling of labels whenever a new site was added [6]. Similarly, one can determine the properties of various lattice sizes L by letting L grow one by one and relabeling the cluster after each growth step [7]. Unfortunately, these two methods came long after most of the percolation properties were already studied quite well by standard Hoshen-Kopelman analysis.

Relation to Ising and Potts Models

The relation between percolation and thermal physics was useful for both sides: Scaling theories for percolation could

follow scaling theories for thermal physics from ten years earlier, and computer simulations for thermal physics could use the Leath and Hoshen-Kopelman algorithms of cluster analysis, leading to the Wolff and Swendsen-Wang methods, respectively, a decade later. A mathematical foundation is given by the Kasteleyn-Fortuin theorem [8] for the partition function Z of the Q-state Potts model at temperature T:

$$Z(Q) = \langle Q^N \rangle \tag{3}$$

where N is the total number $\sum_s n_s$ of clusters for bond percolation at probability $1 = \exp(-2J/k_B T)$, $\langle \dots \rangle$ indicates an average over the configurations at this probability, k_B is Boltzmann's constant and 2J is the energy needed to break a bond between neighboring spins. (Each site i of a Potts lattice carries a variable $S_i = 1, 2, \dots Q$; the energy of a neighbor pair is $-2J$ if the two variables agree, and zero otherwise.) For the special Ising case $Q = 2$ see also this author in this encyclopedia, "Opinion dynamics..." and "Phase transitions...".

Q values of 3 and larger are interesting since for increasing Q a second-order phase transition with a continuous order parameter changes into a first-order phase transition with a jumping order parameter, when T increases. The special case $Q = 2$ is the spin 1/2 Ising model (the model is pronounced EEsing, not EYEsing since Ernst Ising was born in Cologne, Germany, and became US citizen Ernest Ising only after publishing his theory in 1925 and surviving Nazi persecution 1933–1944). The limit $Q \to 0$ corresponds to some tree structures (no cyclic links, as in Flory's percolation theory, [9]). Percolation, on the other hand, corresponds to the limit $Q \to 1$, in the following way: The "free energy" in units of $k_B T$ is in this limit $\ln Z = \ln\langle \exp(N \ln Q)\rangle \simeq \ln\langle \exp[(Q-1)N]\rangle \simeq \ln\langle 1 + (Q-1)N\rangle \simeq (Q-1)N$. Thus for Q near unity this thermal free energy, divided by $Q-1$, is the number of percolation clusters.

In this way thermal physics and percolation are related, and the cluster numbers N correspond to a free energy. In thermal physics, the negative derivative of the free energy with respect a conjugate field gives the order parameter (e.g magnetic field and magnetization), and the field derivative of the order parameter is called the susceptibility. For liquid-gas equilibria, the order parameter is the volume (or the density), the field is the pressure (or chemical potential), and the analog of the susceptibility is the compressibility. This result Eq. (3), not its derivation, we should keep in mind if we now look at the percolation quantities of interest.

Formally we may define for percolation a free energy F as a generating function of a ghost field h:

$$F(h) = \sum_s n_s \exp(-hs) \,. \tag{4}$$

Then its first h-derivative is $-\sum_s n_s s$, and the second one is $\sum_s n_s s^2$, sums which appear below in the percolation probability P_∞ (the order parameter) and the mean cluster size $S = \sum_s n_s s^2 / \sum_s n_s s$ (the susceptibility).

Quantities and Exponents

The basic quantity is n_s, the number (per site) of clusters containing s sites each, and often is an average over several realizations for the same occupation probability p in the same lattice. Several moments

$$M_k = \sum_s n_s s^k \tag{5}$$

are used to define other quantities of interest; in these sums the infinite (spanning) clusters are omitted. The following proportionalities are valid asymptotically in the limit of large lattice size L and for $p \to p_c$:

$$F = M_0 \propto |p - p_c|^{2-\alpha} + \dots \tag{6a}$$

$$P_\infty = p - M_1 \propto (p - p_c)^\beta \tag{6b}$$

$$S = M_2/M_1 \propto |p - p_c|^{-\gamma} \,. \tag{6c}$$

Here F is the analog of the thermal free energy, where the three dots represent analytic background terms whose derivatives are all finite. Since every occupied site must belong either to a finite or to an infinite cluster, $P_\infty = p - \sum_s n_s s$ is the fraction of sites belonging to the infinite cluster and gives the probability that from a randomly selected site we can walk to a lattice boundary along a path of occupied sites. It is thus called the percolation probability but needs to be distinguished from the probability p that a single site is occupied and from the probability R, with $R(p < p_c) = 0$, $R(p > p_c) = 1$, that there is a spanning cluster in the lattice.

The quantity S is usually called the mean cluster size, and we follow this tradition even though it is very bad. There are many ways to define a mean size, and polymer chemists have the much more precise notation of a number average M_1/M_0, a weight average M_2/M_1 and a z average M_3/M_2 for the cluster size (= degree of polymerization). Physicists arbitrarily call the weight-averaged s the mean cluster size S. Numerically, the exponent γ is determined more easily from the "susceptibility" $\chi = M_2 \propto |p - p_c|^{-\gamma}$, since the denominator M_1 in Eq. (6c) approaches very slowly its asymptotic limit of 1.

Scaling Properties, Fractals, and the Renormalization Group Approach to Percolation, Table 2

Critical exponents for percolation clusters. The mean field values are valid for six and more dimensions and also apply to Flory's Bethe approximation and to Erdös-Rényi random graphs. The exponents α, δ, σ, τ can be derived from the scaling laws, Eq. (8)

d	β	γ	ν
2	5/36	43/18	4/3
3	0.41	1.796	0.88
≥ 6	1	1	1/2

The radius of a cluster R_s can be defined as the rms distance r_i, $i = 1, 2, \ldots, s$ of cluster sites from the center of mass r_c of the cluster (radius of gyration):

$$R_s^2 = \langle \sum_i (r_i - r_c)^2 / s \rangle \tag{6d}$$

where the $\langle \ldots \rangle$ average over all cluster configurations at probability p. Then the correlation length ξ is related to the z-average cluster radius through

$$\xi^2 = \sum_s R_s^2 n_s s^2 / \sum_s n_s s^2 \propto |p - p_c|^{-\nu} \tag{6e}$$

with another critical exponent ν.

Finally, right at $p = p_c$, the cluster numbers decay as

$$n_s \propto 1/s^{2+1/\delta} \tag{7}$$

where δ must be positive to allow a finite density $\sum_s n_s s = p$.

These five critical exponents are not independent of each other but are related in d dimensions through the scaling laws:

$$2 - \alpha = \gamma + 2\beta = (\delta + 1)\beta = d\nu \tag{8a}$$

as known from thermal phase transitions; the last equation involving d is not valid in mean field theory (large d) but only for $d \leq 6$. Table 2 gives the numerical estimates of the exponents in three dimensions as well as their mean field values for $d \geq 6$ and their exact two-dimensional results [10,11]. Thus, for six and less dimensions, if you know two exponents you know them all; thus far.

These scaling laws (8a) can be derived by assuming

$$n_s = s^{-\tau} f[(p - p_c)s^{\sigma}] \quad (\tau = 2 + 1/\delta,\ 1/\sigma = \beta\delta) \tag{8b}$$

which was first postulated for the thermal Ising model, and then successfully applied to percolation. Here f is a suitable scaling function, which only in the mean-field limit approaches a Gaussian.

For both thermal critical phenomena and percolation, "universality" asserts that these critical exponents are independent of many details and (for the Potts model) depend only on the dimensionality d and the number Q of possible spin states. Since percolation corresponds to $Q \to 1$ this means that the exponents depend only on d. There are exceptions from this universality for thermal phase transitions, but for random percolation thus far it worked. However, the numerical value of the percolation threshold p_c is not a critical exponent, depends on the lattice structure, and is different for site and bond percolation.

This universality is one of the reasons why the investigation of exponents is important: They allow to classify models and materials. Similarly, in biology we have many birds of different colors, and many types of domestic animals. Biology became a systematic science only when it was found that all mammals share certain properties, which birds no not have. Thus there is the universality class of mammals.

(The proportionality factors in Eq. (6) are not universal, but some of their combinations are; for example, the ratio of the proportionality factors for S above to below p_c is universal. In some sense also the probability $R(p = p_c)$ of a lattice to contain one spanning cluster at the threshold is universal: same for bond and site percolation; however, that probability depends on the boundary conditions and the shape of the sample and thus is far less universal that the mentioned ratio for S.)

Unfortunately, there is another exponent which does not follow from the cluster numbers and radii and for which no scaling law is accepted which relates it to the other exponents above. This refers to the electrical conductivity

$$\Sigma \propto (p - p_c)^{\mu} \tag{9}$$

when each occupied site (or bond) conducts electrical current and each empty site (or deleted bond) is an insulator. The numerical values are 1.30, 2.0 and 3 in two, three and at least six dimensions. If bonds are realized by elastic springs with bending forces, the elastic exponent may be $\mu + 2\nu$ if entropy effects are negligible, or $2 - \alpha$ if entropy effects are dominant. Moreover, μ is less universal: the above lattice values do not hold on a continuum (conducting spheres which may overlap). Similarly, the kinetics of the Ising model determine a critical exponent which differs in different variants of the kinetics and may not be related to the static Ising exponents like β and γ.

Fractal Dimension; Incipient Infinite Cluster

Fractal Dimension D

Typical objects of geometry classes in school are one-dimensional lines, two-dimensional squares or circles, and three-dimensional cubes or spheres. They have a length (radius) L and a mass (volume for unit density) M with $M \propto L^d$ for d dimensions. In reality, mother nature produces much more complex objects, like trees, where the mass varies with a power of the tree height below 3:

$$M \propto L^D \quad (D < d,\ L \to \infty)\,. \tag{10a}$$

D is the fractal dimension, and such objects are called fractals, particularly if they also are self-similar in that a small twig looks like a big branch, etc. Finite-size scaling theory then relates D of the largest (spanning?) cluster at $p = p_c$ to the above percolation exponents through

$$D = d - \beta/\nu = (\gamma+\beta)/\nu = 1/(\sigma\nu) = d/(1+1/\delta) \tag{10b}$$

for $d \le 6$. Thus the critical cluster is about 1.9-dimensional in two and 2.5-dimensional in three dimensions, while in the mean field regime for $d \ge 6$ we have $D = 4$. Why is this so? Any quantity X which is supposed to vary near $p = p_c$ as $|p - p_c|^x$ does so only for infinitely large systems. For a finite lattice size L, the transition is rounded, and neither X nor any of its p-derivatives diverges or becomes exactly zero. In particular, the typical cluster radius or correlation length $\xi \propto |p - p_c|^{-\nu}$ cannot become infinite but becomes of order L. Then the relation $X \propto \xi^{-x/\nu}$ is replaced by

$$X(p = p_c) \propto L^{-x/\nu} \tag{11a}$$

at the threshold, and

$$X(p \simeq p_c) = L^{-x/\nu} g[(p - p_c)L^{1/\nu}] \tag{11b}$$

near the threshold, with a suitable scaling function g. In particular, the fraction P_∞ of sites belonging to the largest cluster at $p = p_c$ vanishes as $L^{-\beta/\nu}$, and the total number M of sites in this cluster as

$$M \propto L^{d-\beta/\nu} \quad \text{or} \quad D = d - \beta/\nu \tag{11c}$$

as asserted in Eq. (10b).

Figure 1 shows the second moment $\chi = M_2 = \sum_s n_s s^2$ in small (curve) and large (+) simple cubic lattices, differing only for $p \simeq p_c$. Figure 2 shows right at $p = p_c$ the variation with lattice size of the number M of sites in the largest cluster and of the second moment M_2 (susceptibility).

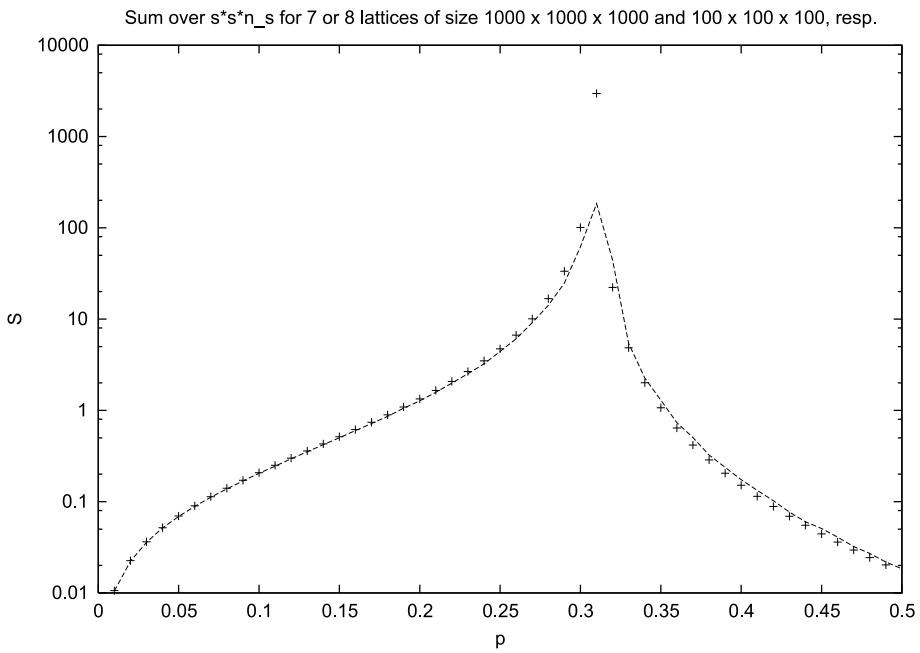
In a finite lattice, the probability $R(p)$ of a spanning cluster to exist goes from nearly zero to nearly unity in a p-interval proportional to $1/L^{1/\nu}$, according to Eq. (11b) with $x = 0$. The derivative dR/dp is the probability that spanning first occurred at probability p. It is plausible that this probability, peaked around p_c, is a Gaussian, i. e. a normal distribution. Unfortunately, the Evil Empire, also known as the Departments of Chemical Engineering, destroyed [12] this beautiful idea: Since for $p \simeq p_c$ and $\xi \sim L$ every part of the lattice is correlated with the rest of the lattice, the central limit theorem does not hold.

(If for $p \ll p_c$ we let the cluster size s go to infinity, which requires a special algorithm, we get into the universality class of lattice animals, Sect. "Small Clusters". Most simply, in the limit $p \to 0$, Eq. (2) simplifies to $n_s/p^s = g_{st}$, that means we look at the distribution of configurations with s sites and perimeter t, where all configurations of a given s are weighted equally, whatever their perimeter t is. An important result for these animals is that in three dimensions their radius R_s varies as $\sqrt{s}$, i,e. their fractal dimension is exactly 2. In two dimensions, only numerical estimates exist with $D \simeq 1.56$. It is highly unusual that a problem has an exact solution in three but not in two dimensions.)

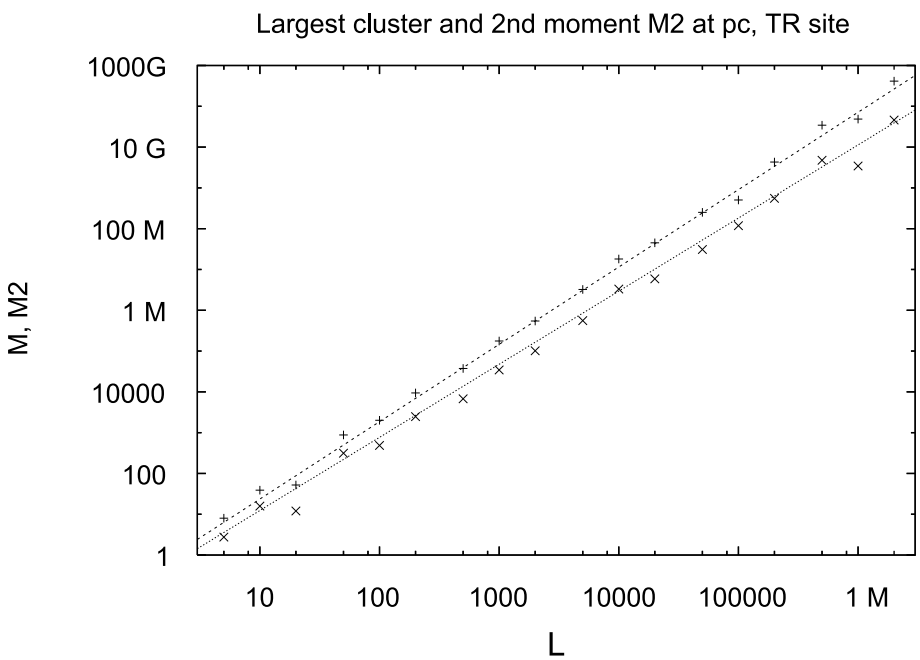
Incipient Infinite Cluster

Right at $p = p_c$ the largest cluster spans the lattice with a pseudo-universal probability $0 < R(p_c) < 1$, and then has a density P_∞ going to zero for L going to infinity. It is also called the incipient infinite cluster IIC. Most of the IIC consists of dangling ends which carry no current if the cluster is interpreted as a random resistor network with conductivity Σ, see Eq. (9) above. The remaining current carrying "backbone" has a fractal dimension 1.643 in two dimensions, 1.7 in three and 2 in at least six dimensions and mostly consists of blobs where current flows along several parallel though connected paths. The few "articulation" sites or bonds, the removal of which cuts the network into two or more parts, are also called "red" since all the current flows through them; they have a fractal dimension of only $1/\nu = 0.75$, 1.14 and 2 in two, three and $\ge$ six dimensions.

How many infinite clusters do we have? The easy answer is: none below, perhaps one at and always one above p_c in an infinite network. Indeed, this is what was claimed mathematically in the 1980's [13]: The number of infinite clusters is zero, one or infinite. Later mathematics excluded the last choice of infinitely many clusters, even though in seven dimensions scaling arguments, confirmed



Scaling Properties, Fractals, and the Renormalization Group Approach to Percolation, Figure 1
"Susceptibility" M_2 in simple-cubic lattice. For the smaller size the maximum is reduced appreciably



Scaling Properties, Fractals, and the Renormalization Group Approach to Percolation, Figure 2
Number M of sites in largest cluster (+) and susceptibility M_2 (x) at $p = p_c = 1/2$ for triangular site percolation. The two straight lines have the exact slopes $D = 91/48$ and $\gamma/\nu = 43/24$ predicted by finite-size scaling. The largest lattice took about 36 h on a workstation with 2 Gb memory. Tiggemann [7] simulated $L = 7 \times 10^6$, 25024, 1305, 225 for $d = 2, 3, 4, 5$ on a large parallel computer

by numerical studies [14], indicated the number of IIC to go to infinity for increasing L in seven dimensions. Only in 1995 and later Aizenman [15] predicted that in all dimensions one may have several spanning clusters at $p = p_c$, in agreement with simulations [16].

Why were the earlier uniqueness theorems irreproducible at p_c and for very elongated rectangles even above p_c [17]? A clear definition of "infinite" is missing in some of the mathematics, although [13] defined a cluster as infinite if its cardinality (= number of sites in it) is infinite for $L \to \infty$ in a hypercubic lattice of L^d sites. Clear definitions of infinity are, of course, needed for reliable proofs [18]. Measure theory as applied in some theorems may be based on some axioms which are not applicable for a fractal IIC. Very simply, imagine each line of an $L \times L$ square lattice to have one randomly selected site occupied and all others empty. The set of occupied site then has cardinality L which is infinite for infinite lattices, but its density becomes zero. Does your measure theory agree with this? More relevant for percolation, even for $p < p_c$ the largest cluster has a size increasing logarithmically with lattice size and thus can be described as infinite, invalidating the percolation threshold as the onset of infinite clusters. Thus infinite might be defined as increasing with a positive power of L, i. e. having a positive fractal dimension. Then we have infinitely many infinite clusters only at $p = p_c$, though in most cases only the largest of them is a spanning cluster. Using "spanning" as a definition of an infinite cluster seems to cause the smallest problems.

Thus one should not regard a question as settled if some mathematical theorem claims to have answered it. The mathematics may not apply to the same problem one is interested in, or (see bootstrap percolation in this encyclopedia) may apply only for unrealistically large lattices. On the other hand, also computer simulations should be relied upon only if confirmed independently. And in the interpretation of simulation results one should be objective and not try to agree with prevailing theories. For example, [14] might already have seen the multiplicity of infinite clusters in five dimensions, not only in seven, had she not followed her obviously incompetent postdoctoral mentor.

(On a more positive side, mathematicians [19] solved biased diffusion on percolating clusters above p_c only a few years after physicists still had controversies about their simulations.)

Simple Renormalization Group

Why are scaling laws and finite-size scaling so simple? Why is universality valid for the exponents? These question arose for thermal critical phenomena as well as for percolation. The main reason is that the correlation length ξ goes to infinity at the critical point. Thus all approximations which restrict the correlations to some finite lengths eventually become wrong, and instead the scaling ideas become correct. They were explained by Ken Wilson through what he called renormalization group, around 1970, and he got the physics Nobel prize for it in 1982. Basically, since correlations extend over long distances, the single atom or lattice point becomes irrelevant and can be averaged over. In politics, we have a similar effect: Many democracies are based on electoral districts, and the candidate winning most votes within this district represents this district in the national parliament. It is the cooperation of many people within the electoral district, not the single vote, which is decisive.

Returning to an $L \times L$ lattice, we can divide it into many blocks of linear dimension b, and treat a block analogously to an electoral district. Thus in an Ising model, if the majority of block spins point upward, the whole block is represented by a superspin pointing up, analogous to the single representative in politics. These block spins then act like the original spins, one can put $b \times b$ superspins into one superblock, and have just one superrepresentative following the majority opinion of the representatives within the superblock. This process can be continued: at each stage $b \times b$ lower representatives are normalized into a single higher representative.

Such a renormalization by majority rule works fine with Ising spins, but percolation deals with connections, not with up and down spins. Thus for percolation a $b \times b$ block is normalized into an occupied supersite if and only if there is a spanning cluster within the block; otherwise the superblock is defined empty. In this way, whole blocks are normalized into single sites via connectedness. And the renormalization is reduced to the standard question which was asked already before Wilson's invention: Does a $b \times b$ lattice have a spanning cluster? The supersite is thus occupied if and only if the block spans, which happens with probability $R_b(p)$. If we call p' the probability of the supersite to be occupied, we thus have

$$p' = R_b(p)\,. \tag{12a}$$

If we are at $p = p_c$, then the renormalization should not change anything drastic since ξ is larger than any b; thus if the renormalization would be exact we would have

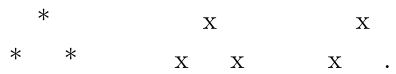
$$p'_c = R_b(p_c)\,. \tag{12b}$$

Practically we determine a fixed point $p = p^*$ such that

$$p^* = R_b(p^*)\,. \tag{12c}$$

and then find p_c as the limit of p^* for $b \to \infty$, which again is similar to what percolation experts did before this renormalization theory.

A particularly simple example is the triangular site percolation problem with $p_c = 1/2$, if we do not divide the lattice into large $b \times b$ blocks but into small triangles of three sites which are nearest neighbors, as shown on the left:



The triangle contains a spanning cluster if either all three sites are occupied (x, central diagram) or two sites are occupied (x) and one site is empty (., right diagram). The first choice appears with probability p^3, the second with probability $p^2(1-p)$. However, this second choice has three possible orientations since each of the three sites can be the single empty site. Thus the total probability of the triangle to have a spanning cluster is

$$p' = p^3 + 3(1-p)p^2 \tag{13a}$$

with three fixed points p^* where $p' = p$:

$$p^* = 0, \quad p^* = 1/2, \quad p^* = 1\,. \tag{13b}$$

The second of these fixed points is the percolation threshold, while the first corresponds to lattice animals (Sect. "Small Clusters" and end of Sect. "Fractal Dimension D") and the third to compact non-fractal clusters. With somewhat more effort one can derive also a good approximation for ν.

This agreement of the fixed point p^* with the true threshold $p_c = 1/2$ is not valid for other lattices or block choices. Nevertheless there was a widespread fixed-point consensus that $R_b(p_c) = p_c$ for sufficiently large b. Regrettably, the Evil Empire [20] again destroyed this beauty and found $R_b(p_c) = 1/2$ for square site percolation where $p_c \simeq 0.593$. In general, $R(p_c)$ is a pseudo-universal quantity depending on boundary conditions and sample shape, while p_c for large samples is independent of these details but is different for site and bond percolation and depends on the size of the neighborhood. Life was much nicer before. Fortunately, if a fixed point is determined by Eq. (12c) and the block size goes to infinity, then the fixed point still approaches p_c.

Future Directions

This review summarized the basic theory, particularly when it was not yet contained in the earlier books [1]. Applications were left to the Sahimi book [1]; even for the very first application [2] there is not yet a complete consensus that the three-dimensional percolation exponents apply to polymer gelation. More recent applications are social percolation [21] for marketing by word-of-mouth, and stock market fluctuations due to herding among traders [22].

Percolation theory, similar to Fortran programming or capitalism, was thought to be finished but seems to be alive and kicking. Nevertheless I think the future is more in its applications.

The manuscript was improved by criticism of A. Aharony.

Bibliography

1. Stauffer D (1979) Phys Rep 54:1; Essam JW (1980) Rep Prog Phys 43:843; Stauffer D, Aharony A (1994) Introduction to Percolation Theory. Taylor and Francis, London (revised second edition); Sahimi M (1994) Applications of Percolation Theory. Taylor and Francis, London; Bunde A, Havlin S (1996) Fractals and Disordered Systems. Springer, Berlin; Grimmett G (1999) Percolation, second edition. Springer, Berlin
2. Flory PJ (1941) J Am Chem Soc 63:3083
3. Grassberger P (2003) Phys Rev E 67:036101
4. Redner S (1982) J Statist Phys 29:309
5. Stauffer D, Jan N (2000) In: Khajehpour MRH, Kolahchi MR, Sahimi M (eds) Annual Reviews of Computational Physics, vol VIII (Zanjan School). World Scientific, Singapore
6. Newman MEJ, Ziff RM (2000) Phys Rev Lett 85:4104
7. Tiggemann D (2006) Int J Mod Phys C 17:1141 and Ph D thesis, Cologne University
8. Kasteleyn PW, Fortuin CM (1969) J Phys Soc Jpn Suppl S 26:11
9. Deng VJ, Garoni TM, Sokal AD (2007) Phys Rev Lett 98:030602
10. Nienhuis B (1982) J Phys A 15:199
11. Smirnov S, Werner W (2001) Math Res Lett 8:729
12. Ziff RM (1994) Phys Rev Lett 72:1942
13. Newman CM, Schulman LS (1981) J Stat Phys 26:613
14. de Arcangelis L (1987) J Phys A 20:3057
15. Aizenman M (1997) Nucl Phys (FS) B 485:551
16. Shchur LN, Rostunov T (2002) JETP Lett 76:475
17. Stauffer D (1999) J Irreproducible Results 44:57
18. Jarai AA (2003) Ann Prob 31:444
19. Berger N, Ganten N, Peres Y (2003) Probab Theory Relat Fields 126:221
20. Ziff RM (1992) Phys Rev Lett 69:2670
21. Weisbuch G, Solomon S (2002) In: Bornholdt S, Schuster HG (eds) Handbook of graphs and networks. Wiley-VCH, Weinheim, p 113
22. Cont R, Bouchaud J-P (2000) Macroeconom Dyn 4:170

Scenario-Driven Planning with System Dynamics

NICHOLAS C. GEORGANTZAS
Fordham University Business Schools, New York, USA

Article Outline

Glossary
Definition of the Subject
Introduction
Environmental Turbulence and Future Uncertainty
SdP with SD: The Modeling Process ≡ Strategic Situation Formulation
Case 1: Cyprus' Environment and Hotel Profitability
Case 2: A Japanese Chemicals Keiretsu (JCK)
Future Directions
Bibliography

Glossary

Mental model how one perceives cause and effect relations in a system, along with its boundary, i. e., exogenous variables, and the time horizon needed to articulate, formulate or frame a decision situation; one's implicit causal map of a system, sometimes linked to the reference performance scenarios it might produce.

Product either a physical good or an intangible service a firm delivers to its clients or customers.

Real option right and obligation to make a business decision, typically a tangible investment. The option to invest, for example, in a firm's store expansion. In contrast to financial 'call' and 'put' options, a *strategic real option* is not tradable. Any time it invests, a firm might be at once acquiring the *strategic real options* of expanding, downsizing or abandoning projects in future. Examples include research and development (abbreviated R&D), merger and acquisition (abbreviated M&A), licensing abroad and media options.

Scenario a postulated sequence or development of events trough time; via Latin *scena 'scene'*, from Greek *σκηνή*, *skēnē* '*tent, stage*'. In contrast to a forecast of what *will* happen in the future, a scenario shows what *might* happen. The term *scenario* must *not* be used loosely to mean situation. *Macro-environmental* as well as *industry-*, *task-* or *transactional-environmental* scenarios are merely inputs to the *strategic objectives* and *real options* a firm must subsequently explore through *strategic scenarios*, *computed* or simulated with an explicit, formal *system dynamics* (abbreviated SD) model of its strategic situation. *Computed strategic scenarios* create the multiple perspectives that strategic thinkers need to defeat the tyranny of dogmatism that often assails firms, governments and other social entities or organizations.

Scenario-driven planning (abbreviated SdP) to attain high performance through strategic flexibility, firms use the SdP *management technology* to create foresight and to anticipate the future with strategic real options, in situations where the business environment accelerates frequently and is highly complex or interdependent, thereby causing uncertainty.

Situation the set of circumstances in which a firm finds itself; its (strategic) state of affairs.

Strategic management process (abbreviated SMP) geared at detecting environmental threats and turning them into opportunities, it *proceeds from* a firm's mission, vision and environmental constraints *to* strategic goals and objectives *to* strategy design or formulation *to* strategy implementation or strategic action *to* evaluation and control *to* learning through feedback (background, Fig. 2).

SMP-1 environmental scanning monitors, evaluates and disseminates knowledge about a firm's internal and external environments to its people. The internal environment contains *s*trengths and *w*eaknesses within the firm; the external shows future *o*pportunities and *t*hreats (abbreviated SWOT).

SMP-2 mission a firm's purpose, *raison d'être* or reason for being.

SMP-3 objectives performance (*P*) goals that SMP often quantifies for some *P* metrics.

SMP-4 policy decision-making guidelines that link strategy design or formulation to action or implementation tactics.

SMP-5 strategy a comprehensive plan that shows how a firm might achieve its mission and objectives. The three strategy levels are: corporate, business and process or functional.

SMP-6 strategy design or formulation the interactive, as opposed to antagonistic, interplay of strategic content and process that creates flexible long-range plans to turn future environmental threats into opportunities; includes internal strengths and weaknesses as well as strategic mission and objectives, and policy guidelines.

SMP-7 strategic action or implementation the process by which strategies and policies are put into action through the development of programs, processes, budgets and procedures.

SMP-8 evaluation and control sub-process that monitors activities and performance, comparing actual results with desired performance.

SMP-9 learning through feedback occurs as knowledge about each SMP element enables improving previous SMP elements (background, Fig. 2).

System an organized group of interrelated components, elements or parts working together for a purpose; parts might be either goal seeking or purposeful.

System dynamics (abbreviated SD) a lucid modeling method born from the need to manage business performance through time. Thanks to Forrester [23], who discovered that all change propagates itself through stock and flow sequences, and user-friendly SD software (*iThink*®, *Vensim*®, etc.), SD models let managers see exactly how and why, like other biological and social organizations, business firms perform the way they do. Unlike other social sciences, SD shows exactly how *feedback loops*, i. e., circular cause and effect chains, each containing at least one time lag or delay, interact within a system to determine its performance through time.

Variable or metric something that changes either though time or among different entities at the same time. An *internal change lever* is a decision or policy variable that a strategy-design modeling, or client, team controls. An *external change trigger* is an environmental or policy variable that a strategy-design modeling team does not control. Both *trigger* and *lever* variables can initiate change and be either endogenous or exogenous to a model of a system.

However certain our expectation, the moment foreseen may be unexpected when it arrives
—T.S. Eliot

Definition of the Subject

Many of us live and work in and about business ecosystems with complex structures and behaviors. Some realize that poor performance often results from our very own past actions or decisions, which come back to haunt us. So business leaders in diverse industries and firms, such as *Airbus, General Motors, Hewlett-Packard, Intel* and *Merck*, use scenario-driven planning (SdP) with system dynamics (SD) to help them identify, design and apply high-leverage, sustainable solutions to dynamically complex strategic-decision situations. One must know, for example, if the effect of an environmental change or strategic action gets magnified through time or is dampened and smoothed away. What may seem insignificant at first might cause major disruption in performance. SdP with SD shows the causal processes behind such dynamics, so firms can respond to mitigate impacts on performance.

Accelerating change and complexity in the global business environment make firms and other social organizations abandon their *inactive, reactive* and *preactive* modes [2]. SdP with SD turns them *proactive*, so they can translate anticipation into action. To properly transform anticipation into action, computed with SD models, 'strategic scenarios' must meet four conditions: *consistency, likelihood, relevance* and *transparency* [37]. Combining SdP with SD for that purpose, with other tools, like actor and stakeholder purposes, morphological methods or probability might help avoid entertainment and explore all possible scenarios. Indeed, SdP with SD

> "does not stand alone… modeling projects are part of a larger effort… modeling works best as a complement to other tools, not as a substitute" (see p. 80 in [75]).

SdP with SD is a systematic approach to a vital top-management job: leading today's firm in the rapidly changing and highly complex global environment. Anticipating a world where product life cycles, technology and the mix of collective- and competitive-strategy patterns change at an unprecedented rate is hard enough. Moving ahead of it might prove larger than the talent and resources now available in leading firms. SdP with SD leads to a decisive integration of strategy design and operations, with the dividing line much lower than at present. As mid-level managers take on more responsibility, senior executives become free to give more time and attention to economic conditions, product innovation and the changes needed to enhance creativity toward strategic flexibility [23].

It is perhaps its capacity to reintegrate strategy content and process that turns SdP with SD into a new paradigm for competitive advantage [42], and simulation modeling in general [28], into a critical fifth tool, in addition to the four tools used in science: observation, logical-mathematical analysis, hypothesis testing and experiment [77]. But full-fledged SD models also allow computing scenarios to assess possible implications of strategic situations. Strategic scenarios are not merely hypothesized plausible futures, but computed by simulating combined changes in strategy and in the business environment [32].

Computed scenarios help managers understand what they do not know, enabling strategy design and implementation through the coalignment of timely tactics to improve long-term performance. Through its judicious use of resources, scenario-driven planning with system dynamics makes the tactics required for implementation clear [27]. And because computed scenarios reveal the required coalignment of tactics through time, SdP with SD helps firms become flexible, dependable and efficient, and save time!

Everyone's mind sees differently, but if there is truth in the adage 'a picture is worth a thousand words', then the complex interrelations that SdP with SD unearth and show must be worth billions. In a world where strategic chitchat dominates, one can only hope that SdP with SD will play

a central role in public and private dialogues about dynamically complex opportunities and threats.

We shape our buildings; thereafter,
our buildings shape us
—Winston Churchill

Introduction

Following on the heels of Ackoff and Emery [3] and Christensen [10], respectively, Gharajedaghi [35] and Raynor [64] show how strategies with the best chances for brilliant success expose firms to debilitating uncertainty. Firms fail as their recipes for success turn bad through time. Gharajedaghi [35] shows, for example, five strategy scenarios that convert success to failure. Each scenario plays a critically different role. Together, however, these scenarios form a dynamically complex system. Through time, as each scenario plays, it enables the context for the next:

1. *Noble ape* or *copycat* strategy imitates and replicates advantage. Also called 'shadow marketing', it lets shadowy copycats instantly *shadow market* product technology, often disruptively.
2. *Patchy* or *sluggish* strategy delays responses to new technology. When this *second* scenario plays, then patching up wastes time, enabling competitors to deliver new technology and to dominate markets. Worse, it causes costs to rise as it drives down product quality.
3. *Satisficing* or *suboptimal* strategy scenarios take many forms. One entails a false assumption: if a policy lever helps produce desired performance, then pulling or pushing on that lever will push performance further.
4. *Gambling* or *changing the game* strategy scenario transforms a strategic situation by playing the game successfully. While dealing with a challenge, firms gradually transform their strategic situation and change the basis for competition, so a whole new game and set of issues emerge. Success marked, for example, the beginning of the *information era*. But competitive advantage has already moved away from having access to information. In our *systems era* [2], creating new knowledge and generating insight is the new game [81].

Lastly, the cumulative effects from these four strategy scenarios trickle down to the:

5. *Archetypal swing* or *paradigm shift* scenario. Both learning and unlearning can cause archetypal swings and paradigm shifts to unfold through time intentionally [76]. These also occur unintentionally when conventional wisdom fails to explain patterns of events that challenge prevailing mental models. The lack of a convincing explanation creates a twilight zone where acceptable ideas are not competent and competent ideas are not acceptable.

Beliefs about the future drive strategies. But the future is unpredictable. Worse, success demands commitments that make it impossible to adapt to a future that turns out surprising. So, strategies with great success potential also bear high failure probabilities. Raynor [64] calls this the *strategy paradox*. Dissolving it requires turning environmental uncertainty into strategic flexibility. To make it so, Raynor urges managers to: anticipate multiple futures with scenarios, formulate optimal strategies for each future, accumulate *strategic real options* [5] and manage the select options portfolio.

SdP with SD helps managers who operate in an uncertain world question their assumptions about how the world works, so they can see it more clearly. To survive, the human mind overestimates small risks and underestimates large risks. Likewise, it is much more sensitive to losses than to gains. So the capability to leverage opportunities and to mitigate risk might have become an economic value driver.

The purpose of computing scenarios is to help managers alter their view of reality, to match it up more closely with reality as is and as it might become. To become a leader, a manager must define reality. The SdP with SD purpose is *not*, however, to paint a more accurate picture of tomorrow, but to improve the quality of decisions about the future. Raynor says that the requisite strategic flexibility, which SdP with SD creates:

> "is not a pastiche of existing approaches. Integrating these tools and grounding them in a validated theory of organizational hierarchy creates something that is quite different from any of these tools on its own, or in mere combination with the others" (see p. 13 in [64]).

Indeed, knowledge of common purposes and the acceptable means of achieving them form and hold together a purposeful hierarchical system. Its members know and share values embedded in their culture, which knits parts into a cohesive whole. And because each part has a lot to say about the whole, consensus is essential to SdP with SD for the co-alignment of diverse interests and purposes.

Ackoff and Emery [3], Gharajedaghi [35] and Nicolis [55] concur that purpose offers the lens one needs to see a firm as a multi-minded social net. A purposeful firm pro-

duces either the same result differently in the same environment or different results in the same or different environments. Choosing among strategic real options is necessary but insufficient for purposefulness. Firms that behave differently but show only one result per environment are goal seeking, not purposeful. Servomechanisms are goal seeking but people are purposeful. As a purposeful system, the firm is part of purposeful sub-systems, such as its *industry value chain* [61] and the society. And firms have purposeful people as members. The result is a dynamically interdependent, i. e., complex, hierarchical purposeful system.

A firm's value chain is, along with its primary and support activities, at once a member of at least one industry value chain and of the society or *macro-environment*. Industry analysis requires looking at value chains independently from the society [61]. But people, the society and firm and industry value chains are so interdependent, so interconnected, that an optimal solution might not exist for any of them independently of the others. SdP with SD helps firms co-align the 'plural rationality' of purposeful stakeholder groups with each other and that of the system as a whole.

Seeing strategic management as a *strategies and tactics net* [27] is in perfect syzygy with the *plural rationality* that SdP with SD accounts for among individuals, groups and organizations. Singer [73,74] contrasts monothematic conventional universes of traditional rationality with the multiverse-directed view of plural rationality. In counterpoint, Morecroft's [52] computed scenarios trace the dysfunctional interactions among sales objectives, overtime and sales force motivation to the intended, i. e., stated, singular rationality that drove action in a large sales organization.

Because their superordinate purpose is neither to compete nor to collaborate, but to develop new wealth-creating capabilities, in unique ways that serve both current and future stakeholder interests, customers and clients included [51], firms can benefit from the multiverse-directed view of strategic management as a net of strategies and tactics. SdP with SD helps firms break free from the tradeoffs tyranny of the mass-production era. Evidently, adherents to tradeoffs-free strategy like *Bell Atlantic*, *Daimler-Benz*, *Hallmark* and *Motorola* "can have it all" [60].

A firm must serve the purposes of its people as well as those of its environment, not as a mindless mechanical system, but as a living, purposeful, *knowledge-bonded* hierarchical system [3,35,55,81]. To clarify, a bike always yields to its rider, for example, regardless of the rider's desire; even if that entails running into a solid brick wall. Ouch**!** But riding a horse is an entirely different story. Horse and rider form a knowledge-bonded system: the horse must know the rider and the rider must know exactly how to lead the horse.

SdP with SD History: Always Back, Always in Style, Always Practical

Herman Kahn introduced scenarios to planning while at RAND Corporation in the 1950s [45]. Scenarios entered military strategy studies conducted for the US government. In the 1960s, Ozbekhan [58] used urban planning scenarios in Paris, France. Organization theorists and even novelists were quick to catch on. The meaning of scenarios became literary. Imaginative improvisation produced flickering apocalyptic predictions of strikingly optimistic and pessimistic futures. Political and marketing experts use scenarios today to jazz up visions of favorable and unfavorable futures.

Wack [78,79] asserts it was Royal Dutch Shell that came up with the idea of scenarios in the early 1970s. Godet [36] points to the French OTAM team as the first to use scenarios in a futures study by DATAR in 1971. Brauers and Weber [8] claim that Battelle's scenarios method [49] was originally a German approach. In connection with planning, however, most authors see scenario methods as typically American.

Indeed, during the 1970s, US researchers Olaf Helmer and Norman Dalkey developed scenario methods at RAND for eliciting and aggregating group judgments via Delphi and cross-impact matrices [4]. They extended cross impact analysis within statistical decision theory [39]. A synthesis of scenario methods began in the 1970s that draws together multiple views, including those of professional planners, analysts and line managers.

Ansoff [6] and other strategy theorists state that the 1970s witnessed the transformation of global markets. Today, changes in the external sociopolitical environment become pivotal in strategy making. Combined with the geographical expansion of markets, they increase the complexity of managerial work. As environmental challenges move progressively faster, they increase the likelihood of strategic surprises. So, strategic thinkers use scenarios to capture the nonlinearity of turbulent environments. Examples are Hax and Majluf [38] and, more clearly so, Porter [61] and Raynor [64]. They consider scenarios instrumental both in defining uncertainty and in anticipating environmental trends.

Huss and Honton [41] see scenarios emerge as a distinct field of study, a hybrid of a few disciplines. They identify multiple scenarios methods that fall into three major categories:

1. Intuitive logics [78,79], now practiced by *SRI International*,
2. Trend-impact analysis, practiced by the *Futures Group* and
3. Cross-impact analysis, practiced by the *Center for Futures Research* using INTERAX (Interactive Cross-Impact Simulation) and by Battelle using BASICS (BAttelle Scenario Inputs to Corporate Strategies).

Similarly, after joining Ozbekhan to advocate reference scenarios, Ackoff [2] distinguishes between:

1. Reference projections as piecemeal extrapolations of past trends and
2. The overall reference scenario that results from putting them together.

Based on Acar's [1] work under Ackoff, Georgantzas and Acar [32] explore these distinctions with a practical managerial technology: *comprehensive situation mapping* (CSM). CSM is simple enough for MBA students to master in their capstone Business Policy course. With the help of *Vensim® PLE* [18], CSM computes scenarios toward achieving a well-structured process of managing ill-structured strategic situations. In their introduction to SD, Georgantzas and Acar (see Chap. 10 in [32]) draw from the banquet talk that Jay Wright Forrester, Germeshausen Professor Emeritus, MIT, gave at the 1989 *International Conference of the System Dynamics Society*, in Germany, at the University of Stuttgart:

After attending the Engineering College, University of Nebraska, which included control dynamics at its core, Forrester went to MIT. There he worked for Gordon S. Brown, a pioneer in feedback control systems. During World War II, Brown and Forrester worked on servomechanisms for the control of radar antennas and gun mounts. This was research toward an extremely practical end, during which Forrester run literally from mathematical theory to the battlefield, aboard the US carrier *Lexington*.

After the war, Forrester worked on an analog aircraft flight simulator that could do little more than solve its own internal idiosyncrasies. So, Forrester invented *random-access magnetic storage* or *core memory*. His invention went into the heart of *Whirlwind*, a digital computer used for experimental development of military combat systems that eventually became the *semi**a**utomatic **g**round **e**nvironment* (SAGE) air defense system for North America.

Alfred P. Sloan, the man who built *General Motors*, founded the *Sloan School of Management* in 1952. Forrester joined the school in 1956. Having spent fifteen years in the science and engineering side of MIT, he took the challenge of exploring what engineering could do for management.

One day, he found himself among students from *General Electric*. Their household appliance plants in Kentucky puzzled them: they would work with three or four shifts for some time and then, a few years later, with half the people laid off. Even if business cycles would explain fluctuating demand, that did not seem to be the entire reason. *GE*'s managers felt something was wrong.

After talking with them about hiring, firing and inventory policies, Forrester did some simulation on a paper pad. He started with columns for inventories, employees and customer orders. Given these metrics and *GE*'s policies, he could tell how many people would be hired or fired a week later. Each decision gave new conditions for employment, inventories and production. It became clear that wholly determined internally, the system had potential for oscillatory dynamics. Even with constant incoming orders, the policies caused employment instability. That longform simulation of *GE*'s inventory and workforce system marked the beginning of system dynamics [23,24,25,26].

SdP with SD Use and Roadmap

Scenarios mostly help forecast alternative futures but, as firms abandon traditional forecasting methods for interactive planning systems, line managers in diverse business areas adopt scenario-driven planning with system dynamics. Realizing that a tradeoffs-free strategy design requires insight about a firm's environment, both business and sociopolitical, to provide intelligence at *all* strategy levels, firms use SdP with SD to design *corporate, business* and *process* or *functional* strategies. SdP with SD is not a panacea and requires discipline, but has been successful in many settings. Its transdisciplinary nature helps multiple applications, namely capital budgeting, career planning, civil litigation [31], competitive analysis, crisis management, decision support systems (DSS), macroeconomic analysis, marketing, portfolio management and product development [65]. SdP with SD is a quest for managers who wish to be leaders, not just conciliators. They recognize that *logical incrementalism*, a piecemeal approach, is inadequate when the environment and their strategy change together.

Top management might see both divisional, i. e., business, and process or functional strategies as ways of implementing corporate strategy. But *active subsidiaries* [43,44] provide both strategic ideas and results to their parent en-

terprise. Drawing too stiff a line between the corporate office and its divisions might be

> "an unhealthy side effect of our collective obsession with generating returns. The frameworks for developing competitive strategy that have emerged over the last thirty years have given us unparalleled insight into how companies can succeed. And competitive strategy remains enormously important, but it should be the preserve of divisional management… corporate strategy should be focused on the management of strategic uncertainty" (see p. 11 in [64]).

Roadmap It is material to disconnect scenarios from unproductive guesswork and to anchor them to sound practices for strategy design. This guided tour through the fascinating but possibly intimidating jungle of scenario definitions shows what the future might hold for SdP with SD. Extensive literature, examples, practical guidelines and two real-life cases show how computed scenarios help manage uncertainty, that necessary disciple of our open market system. Unlike extrapolation techniques, SdP with SD encourages managers to think broadly about the future.

The above sections clarify the required context and provide a glossary. Conceptual confusion leads to language games at best and to operational confusion at worst [15]. SdP with SD helps firms avert both types of confusion. Instead of shifting their focus away from actuality and rationality, managers improve their insight about fundamental assumptions underlying changes in strategy. The mind-set of SdP with SD makes it specific enough to give practical guidance to those managing in the real world, both now and in the future.

The sections below look at *three* SdP with SD facets linked to strategy design and implementation. The *first* facet involves the business environment, the forces behind its texture and future's requisite uncertainty (Sect. "Environmental Turbulence and Future Uncertainty"). The *second* entails unearthing unstated assumptions about changes in the environment and in strategy, and about their potential combined effects on performance. The SdP with SD framework (Sect. "SdP with SD: The Modeling Process ≡ Strategic Situation Formulation") builds on existing scenario methods. Its integrative view delineates processes that enhance institutional learning, bolster productivity and improve performance through strategic flexibility. It shows why interest in computed scenarios is growing.

The *third* facet entails *computing* the combined or mixed effects on performance of changes both in the environment and in strategy. Even in mature economies, no matter how and how frequently said, decision makers often forget how the same action yields different results as the environment changes. The result is often disastrous. Conversely, the tight coupling between computed scenarios and strategic results can create new knowledge. Linking a mixed environmental and decision scenario in a one-to-one correspondence to a strategic result suits the normative inclination of strategic management, placing rationalistic inquiry at par with purely descriptive approaches in strategy research.

The unified treatment of SdP with SD and the strategy-making process grants a practical bonus, accounting for the entry's peculiar nature. It is not only a conceptual or idea contribution, but also an application-oriented entry. Sections "Case 1: Cyprus' Environment and Hotel Profitability" and "Case 2: A Japanese Chemicals Keiretsu (JCK) present two real-life cases of scenario-driven planning with system dynamics. Written with both the concrete and the abstract thinker in mind, the two cases show how firms and organizations build scenarios with a modest investment. SdP with SD provides an effective management technology that serves well those who adopt it. It saves them both time and energy.

Improvements in causal mapping [19,20], and SD modeling and analysis [50,57] contribute to the SdP with SD trend (Sect. "Future Directions"). Behavioral decision theory and cognitive science also help translate the knowledge of managers into SD models. The emphasis remains on small, transparent models of strategic situations and on dialogue between the managers' mental models and the computed scenarios [53].

All prognosticators are bloody fools
—Winston Churchill

Environmental Turbulence and Future Uncertainty

Environmental Turbulence

Abundant frameworks describe the business environment, but the one by Emery and Trist [22], which Duncan [17] abridged, has been guiding many a strategic thinker. It shows four business environments, each more complex and troublesome for the firm than the preceding one (Fig. 1a).

1. *Placid* or *independent-static environment*: infrequent changes are independent and randomly distributed,

i. e., IID. Surprises are rare, but no new major opportunities to exploit either (*cell* 1, Fig. 1a).
2. *Placid-clustered* or *complex-static environment*: patterned changes make forecasting crucial. Comparable to the economist's idea of imperfect competition, this environment lets firms develop distinctive competencies to fit limited opportunities that lead to growth and bureaucracy (*cell* 2, Fig. 1a).
3. *Disturbed-reactive* or *independent-dynamic environment*: firms might influence patterned changes. Comparable to oligopoly in economics, this environment makes changes difficult to predict, so firms increase their operational flexibility through decentralization (*cell* 3, Fig. 1a).
4. *Turbulent field* or *complex-dynamic environment*: most frequent, changes are also complex, i. e., interdependent, originating both from autonomous shifts in the environment and from interdependence among firms and conglomerates. Social values accepted by members guide strategic response (*cell* 4, Fig. 1a).

Ansoff and McDonnell [7] extend the dichotomous environmental uncertainty perceptions by breaking turbulent environments (*cell* 4, Fig. 1a) into *discontinuous* and *surprising*. This is a step in the right direction, but not as helpful as a causal model specific to the system structure of a firm's strategic situation. Assuredly, 2 × 2 typologies help clarify exposition and are most frequent in the organization theory and strategy literatures. The mystical significance of duality affected even Leibniz, who associated one with God and zero with nothingness in the binary system. The generic solutions that dichotomies provide leave out the specifics that decision makers need. No matter what business they are in (Fig. 1b), managers cannot wait until a better theory comes along; they must act now.

It is worth noting that people often confuse the term 'complex' with 'complicated'. Etymology shows that *complicated* uses the Latin ending *-plic*: *to fold*, but *complex* contains the Greek root *πλέξ-* '*plēx-*': *to weave*. A complicated structure is thereby folded, with hidden facets stuffed into a small space (Fig. 1c). But a complex structure has interwoven parts with mutual interdependencies that cause dynamic complexity [46]. Remember: complex is the opposite of independent or untwined (Fig. 1a) and complicated is the opposite of simple (Fig. 1c).

Daft and Weick's [12] vista on firm *intrusiveness* and *environmental equivocality* is pertinent here. They see many events and trends in the environment as being inherently unclear. Managers discuss such events and trends, and form mental models and visions expressed in a fuzzy language and label system [80]. Within an *enactment* process, equivocality relates to managerial assumptions underlying the *analyzability* of the environment. A firm's *intrusiveness* determines how *active* or *passive* the firm is about environmental scanning. In this context, as the global environment gets turbulent, active firms and their subsidiaries construct SdP with SD models and compute scenarios to improve performance.

Managers of active firms combine knowledge acquisition with interpretations about the environment and their strategic situation. They reduce equivocality by assessing alternative futures through computed scenarios. In frequent meetings and debates, some by videoconferencing, managers use the dialectical inquiry process for *s*trategic *a*ssumption *s*urfacing and *t*esting (SAST), a vital strategic loop. Often ignored, the SAST loop gives active firms a strategic compass [47].

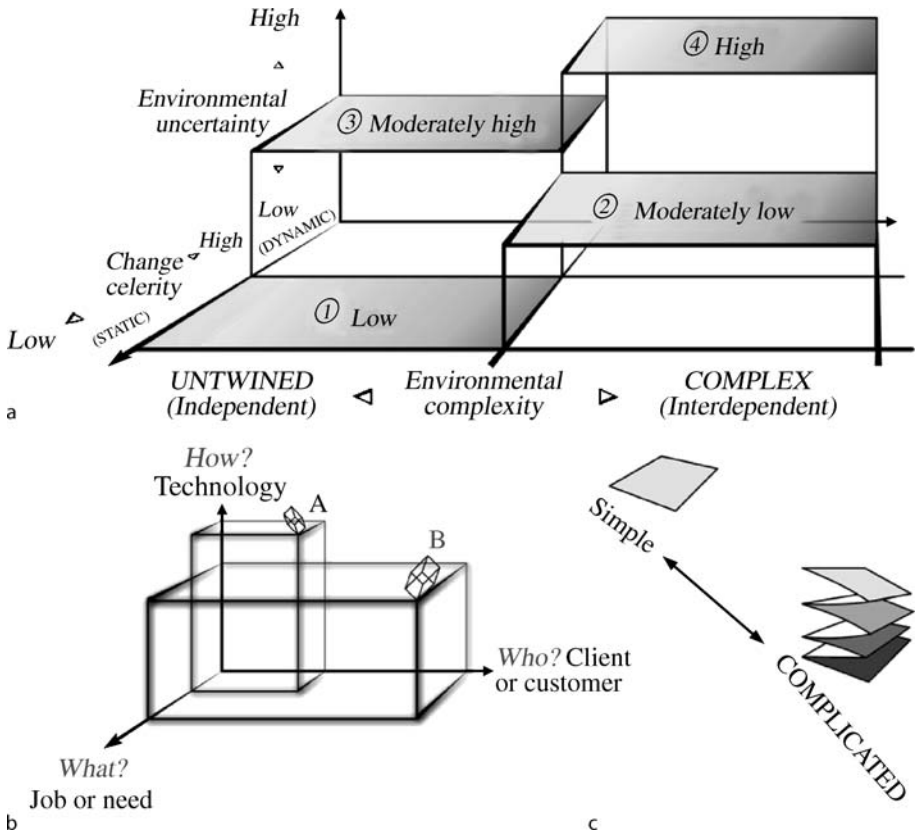
Conversely, passive firms do not actively seek knowledge but reduce equivocality through rules, procedures and regular reports: reams of laser-printed paper with little or no pertinent information. Managers in passive firms use the media to interpret environmental events and trends. They obtain insight from personal contacts with significant others in their environment. Data are personal and informal, obtained as the opportunity arises.

Future Uncertainty

"If we were omnipotent", says Ackoff, then we could get "perfectly accurate forecasts" (see p. 60 in [2]). Thank God the future is unpredictable and we must yet create it. If it were not, then life would have been so boring! Here are some facts about straight forecasting:

1. Forecasts are seldom perfect, in fact, they are always wrong, so a useful forecasting model is one that minimizes error.
2. Forecasts always assume underlying stability in systems.
3. Product family and aggregated forecasts are always more accurate than single product forecasts, so the large numbers law applies.
4. In the short-term, managers can forecast but cannot act because time is too short; in the long term, they can act but cannot forecast.

To offset conundrum #4, SdP with SD juxtaposes the decomposition of performance dynamics into the growth and decline archetypes caused by *balancing* (–) and *reinforcing* (+) recursive causal-link chains or *feedback loops* [33,50]. A thermostat is a typical example of a goal-seeking feedback loop that causes either balancing growth



Scenario-Driven Planning with System Dynamics, Figure 1

a Environmental complexity and change celerity dimensions that cause perceived environmental uncertainty (adapted from [32]). b Scenario-driven planning with system dynamics helps with strategy-design fundamentals, such as, for example, defining a business along the requisite client-job-technology three-dimensional grid. c The simple-complicated dimension must *not* be confused with the environmental complexity dimension (adapted from [46])

or decline. The gap between desired and room temperature causes action, which alters temperature with a time lag or delay. Temperature changes in turn close the gap between desired and room temperature.

Conversely, a typical loop that feeds on itself to cause either exponential growth or decline is that of an arms race. One side increases its arms. The other sides increase theirs. The first side then reacts by increasing its arms, and so on. Price wars between stores, promotional competition, shouting matches, one-upmanship and the wildcard interest rates of the late 1970s are good examples too. Escalation might persist until the system explodes or outside intervention occurs or one side quits, surrenders or goes out of business. In the case of wildcard interest rates, outside intervention by a regulatory agency can bring an end to irrationally escalating rates.

We've never been here before
—Peter Senge

SdP with SD: The Modeling Process ≡ Strategic Situation Formulation

The strategic management process (SMP, Fig. 2) starts with environmental scanning, in order to gauge environmental trends, opportunities and threats. Examples include increasing rivalry among existing competitors and Porter's [62] emphasis on the bargaining power of buyers and suppliers as well as on the threats of new entrants and substitutes. Even if some firms reduce environmental scanning to industry analysis in practice, changes in the environment beyond an industry's boundaries can determine what happens within the industry and its entry, exit and inertia barriers. Internal capability analysis comes next. It examines a firm's past actions and internal policy levers that can both propel and limit future actions. The integrative perspective of the SdP with SD framework on Fig. 2 delineates processes that enhance institutional learning, bolster productivity and improve performance through strategic flexibility.

Strategy design begins by identifying variables pertinent to a firm's strategic situation, along with their interrelated causal links. Changes in these variables can have profound effects on performance. Some of the variables belong to a firm's external environment. Examples are emerging new markets, processes and products, government regulations and international interest and currency rates. Changes either in these or their interrelated causal links determine a firm's performance through time.

It is a manager's job to understand the causal links underlying a strategic situation. SdP with SD helps anticipate the effects of future changes triggered in the external environment. Other variables are within a firm's control. Pulling or pushing on these internal levers also affects performance. To evaluate a change in strategy, one must look at potential results along with changes in the environment, matching resource capabilities, stakeholder purposes, and organizational goals and objectives (Fig. 2).

Most variables interact. Often, the entire set of possible outcomes is obscure, difficult to imagine. But if managers oversimplify, then they end up ignoring the combined effects of chain reactions. Even well-intended rationality often leads to oversimplification, which causes cognitive biases (CBs) that mislead decision makers [21,70,72]. Conversely, computing mixed environmental and decision scenarios that link internal and external metrics can reveal unwarranted simplification.

SdP with SD integrates business intelligence with strategy design, not as a narrow specialty, but as an admission of limitations and environmental complexity. It also uses multiperspective dialectics, crucial for strategic assumption surfacing and testing (SAST). Crucial because the language and labels managers use to coordinate strategic real options are imprecise and fuzzy. Fuzzy language is not only adequate *initially* for managing interdependence-induced uncertainty but required [80]. Decision makers rely on it to overcome psychological barriers and Schwenk's [70] groups of CBs.

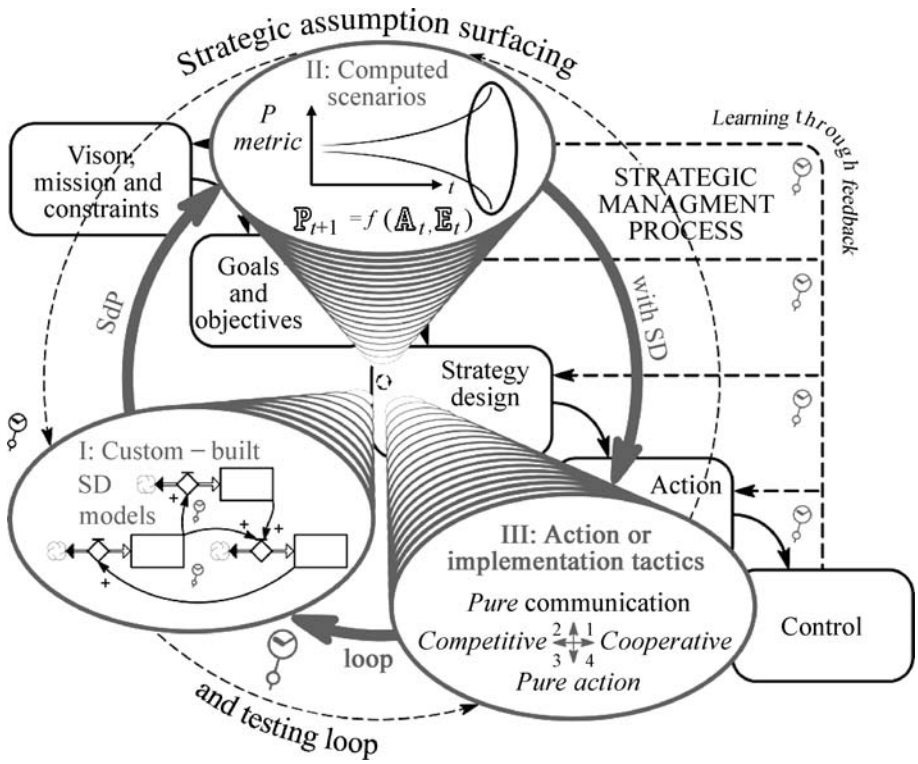
The best-case scenario for a passive firm is to activate modeling on Fig. 2, sometimes unknowingly. When its managers boot up, for example, electronic spreadsheets that contain inside-out causal models, with assumptions hidden deeply within many a formula. At bootup, only the numbers show. So passive-firm managers use electronic spreadsheets to laser-print matrices with comforting numbers. They

> "twiddle a few numbers and diligently sucker themselves into thinking that they're forecasting the future" [69].

And that is only when rapid changes in the environment force them to stop playing *blame the stakeholder*. They stop fighting the last war for a while, artfully name the situation a crisis, roll up their sleeves, and chat about and argue, but quickly agree on some arbitrary interpretation of the situation to generate strategic face-saving options. Miller and Friesen (see pp. 225–227 in [48]) show how for futile firms, rapid environmental changes lead to crisis-oriented decisions. Conversely, successful firms look far into the future as they counter environmental dynamism through strategy design with real options. Together, their options and interpretation of the environment, through the consensus that SdP with SD facilitates, enable a shared logic to emerge: a shared mental model that filters hidden spreadsheet patterns and heroic assumptions clean and clear.

Managers of active firms enter the SdP with SD loop of Fig. 2 both consciously and conscientiously. They activate strategic intelligence via computed scenarios and the SAST loop. Instead of twiddling spreadsheet numbers, *pro*active firm managers twiddle model assumptions. They stake, through SD model diagrams, their intuition about how they perceive the nature and structure of a strategic situation. Computed scenarios quantitatively assess their perceived implications. Having quantified the implications of shared visions and claims about the structure of the strategic situation, managers of active firms are likely to reduce uncertainty and equivocality. Now they can manage strategic interdependence. Because articulated perception is the starting point of all scenarios, computed scenarios give active firms a fair chance at becoming fast strategic learners.

The design of action or implementation tactics requires detailing how, when and where a strategy goes into action. In addition to assuming the form of *pure communication* (III: 1 and 2, Fig. 2) or *pure action* (III: 3 and 4, Fig. 2), in a pragmatic sense, tactics can be either cooperative or competitive and defensive or offensive. Market location tactics, for example, can be either offensive, trying to rob market share from established competitors, or defensive, preventing competitors from stealing one's market share. An offensive tactic takes the form of frontal assault, flanking maneuver, encirclement, bypass attack or guerilla warfare. A defensive tactic might entail raising structural barriers, increasing expected retaliation or lowering the inducement for future attack. Conversely, cooperative tactics try to gain mutual advantage by working with rather than against others. Cooperative tactics take the form of alliances, joint ventures, licensing agreements, mutual service consortia and value-chain partnerships, the co-location of which often creates industrial districts [29].



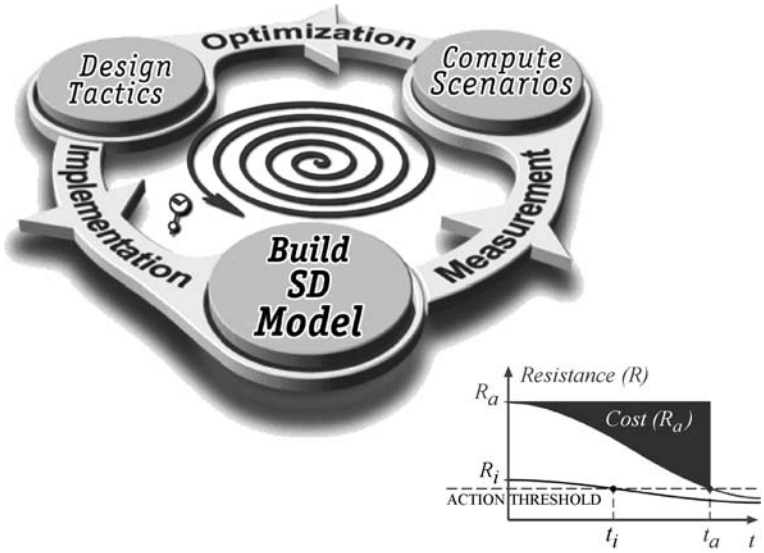
Scenario-Driven Planning with System Dynamics, Figure 2
Cones of resolution show how scenario-driven planning with system dynamics enhances the strategy design component of the strategic management process (SMP; adapted from [32])

The usual copycat strategy retort shows linear thinking at best and clumsy *benchmarking*, also known as shadow marketing, at worst. Its proponents assume performance can improve *incrementally*, with disconnected tactics alone, when strategy design is of primary concern. Piecemeal tactics can undermine strategy, but they are secondary. It might be possible to improve performance through efficient tactics, but is best to design strategies that expel counterproductive tactics. Counterproductive tactics examples are coercive moves that increase rivalry, without a real payoff, either direct or indirect, for the industry incumbent who initiates them. It is atypical of an industry or market leader to initiate such moves.

In strategy, superb action demands superior design. According to the design school, which Ansoff, Channon, McMillan, Porter, Thomas and others lead, logical incrementalism may help implementation, but becomes just another prescription for failure when the environment shifts. Through its judicious use of corporate resources, SdP with SD makes the tactics required for action clear. Also, it reveals their proper coalignment through time, so a firm can build strategic flexibility and save time!

The Modeling Process ≡ Strategic Situation Formulation

SdP with SD (Fig. 2) begins by modeling a business or 'social process' than a business or 'social system'. It is more productive to identify a *social process* first and then seek its causes than to slice a chunk of the real world and ask what dynamics it might generate. Distinguishing between a *social system* and a social process is roughly equivalent to distinguishing between a system's underlying causal structure and its dynamics. Randers (see p. 120 in [63]) defines a social system as a set of cause and effect relations. Its structure is a causal diagram or map of a real-world chunk. A social process is a behavior pattern of events evolving through time. The simulation results of SdP with SD models show such chains of events as they might occur in the real world. An example of a social system (structure) is the set of rules and practices that a firm might enact when dealing with changes in demand, along with the communication channels used for transmitting information and decisions. A corresponding social process (dynamics) might be the stop-and-go pattern of capital in-



Scenario-Driven Planning with System Dynamics, Figure 3

The recursive nature of the modeling process that scenario-driven planning with system dynamics entails a creates a sustainable, ever-expanding vortex of insight and wisdom, needed in strategic real-options valuation, and b saves both time and money as it renders negligible the cost of resistance (*R*) to change

vestment caused by a conservative bias in a firm's culture.

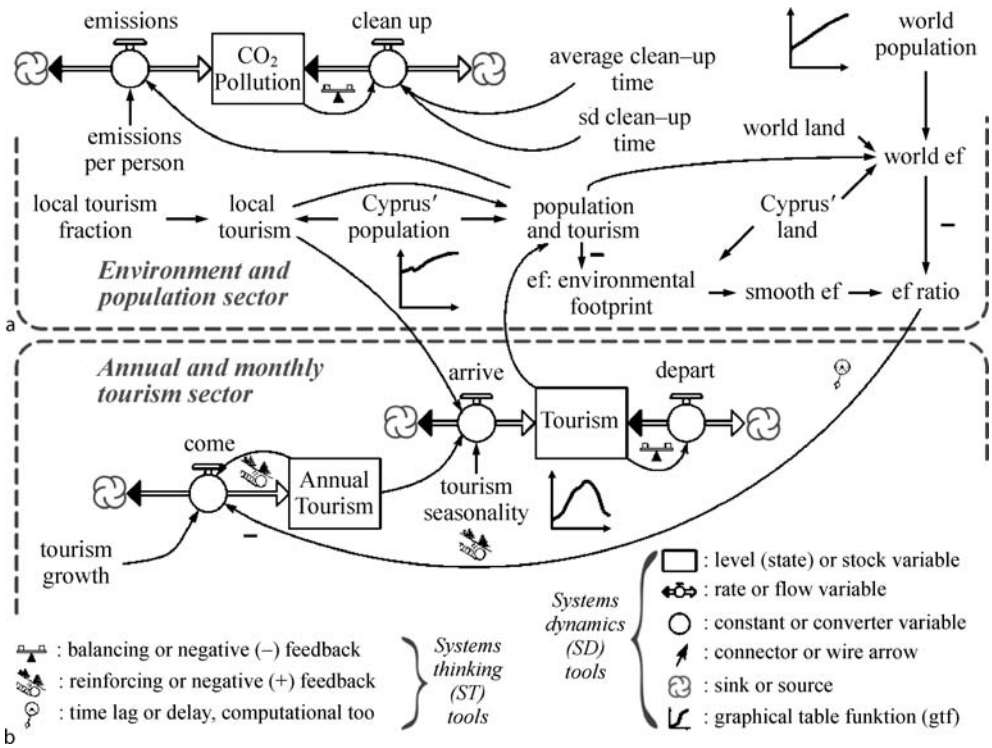
In his model of a new, fast-growing product line, for example, Forrester [24] incorporates such a facet of corporate culture. Causing sales to stagnate, considerable back orders had to accumulate to justify expansion because the firm's president insisted on personally controlling all capital expenditures.

People often jump into describing system structure, perhaps because of its tangible nature as opposed to the elusive character of dynamics or social process fragments. Also, modelers present model structure first and then behavior. Ultimately, the goal in modeling a strategic situation is to link system structure and behavior. Yet, in the early stages of modeling is best to start with system dynamics and then seek underlying causes. Indeed, SD is particularly keen in understanding system performance, "not structure per se" (see p. 331 in [56]), in lieu of SD's core tenet that structure causes performance.

The modeling process itself is recursive in nature. The path from real-world events, trends and negligible externalities to an effective formal model usually resembles an expanding spiral (Fig. 3a). A useful model requires conceptualization; also focusing the modeling effort by establishing both the time horizon and the perspective from which to frame a decision situation. Typically, strategy-design models require a long-term horizon, over which computed scenarios assess the likely effects of changes both in strategy and in the environment.

Computer simulation is what makes SdP with SD models most useful. Qualitative cause and effect diagrams are too vague, tricky to simulate mentally. Produced through knowledge elicitation, their complexity vastly exceeds the human capacity to see their implications. Casting a chosen perspective into a formal SdP with SD model entails postulating a detailed structure; a diagramming description precise enough to propagate images of alternative futures, i. e., computed scenarios, "though not necessarily accurate" (see p. 118 in [63]). But the modeling process must never downplay the managers' mental database and its knowledge content. Useful models always draw on that mental database [24].

Following Morecroft [53], SdP with SD adopters might strive to replace the notion of modeling an objectively singular world *out there*, with the much softer approach of building formal models to improve managers' mental models. The expanding spiral of Fig. 3a shows that the insight required for decisive action increases as the quantity of information decreases, by orders of magnitude. The required quantification of the relations among variables pertinent to a strategic situation changes the character of the information content as one moves from mental to written to numerical data. Perceptibly, a few data remain, but much more pertinent to the nature and structure of the situation. Thanks to computed scenarios, clarity rules in the end. And, if the modeling process stays *i*nteractive (i), as opposed to *a*ntagonistic (a), then clarity means low resistance to change ($R_i < R_a$, Fig. 3b), which helps reach



Scenario-Driven Planning with System Dynamics, Figure 4
Cyprus' a environment and population, and b annual and monthly tourism model sectors (adapted from and extending [30])

a firm's action/implementation threshold quickly ($t_i < t_a$, Fig. 3b). This is how SdP with SD users build strategic flexibility while they save both time and money!

Case 1: Cyprus' Environment and Hotel Profitability

Cyprus' Hotel Association wished to test how Cyprus' year 2010 official tourism strategy might affect tourist arrivals, hotel bed capacity and profitability, and the island's environment [30]. Computed with a system dynamics simulation model, four tourism growth scenarios show what might happen to Cyprus' tourism over the next 40 years, along with its potential effects on the sustainability of Cyprus' environment and hotel profitability. Following is a partial description of the system dynamics model that precedes its dynamics.

Model Description (Case 1)

The SD model highlights member interactions along Cyprus' hotel value chain. The model incorporates a generic value-chain management structure that allows modeling customer-supplier value chains in business as well as in physical, biological and other social systems. Although the structure is generic, its situation specific parameters faithfully reproduce the dynamic behavior patterns seen in Cyprus' hotel value-chain processes, business rules and resources.

Cyprus' Environment, Population and Tourism Model Sectors Within Cyprus' environment and population sector (Fig. 4a), the carbon dioxide (CO_2) pollution stock is the accumulation of Cyprus' anthropogenic emissions less the Mediterranean Sea region's self clean-up rate. The clean-up rate that drains Cyprus' CO_2 pollution depends on the level of anthropogenic pollution itself as well as on the average clean-up time and its standard deviation (sd). Emissions that feed CO_2 pollution depend on Cyprus' population and tourism and on emissions per person [9].

In SD models, rectangles represent stocks, i. e., level or state variables that accumulate through time, e. g., the Tourism stock on Fig. 4b. The double-line, pipe-and-valve-like icons that fill and drain the stocks, often emanating from cloud-like sources and ebbing into cloud-like sinks, represent material flows that cause the stocks to change. The arrive rate of Fig. 4b, for example, shows tourists who flow into the tourism stock per month. Single-line arrows represent information flows, while plain text or circular icons depict auxiliary con-

stant or converter variables, i. e., behavioral relations or decision points that convert information into decisions. Changes in the tourism stock, for example, depend on annual tourism, adjusted by tourism seasonality. Both the diagram on Fig. 4a and Table 1 are reproduced from the actual simulation model, first built on the glass of a computer screen using the diagramming interface of *iThink*® [67], and then specifying simple algebraic equations and parameter values. Built-in functions help quantify policy parameters and variables pertinent to Cyprus' tourism situation.

There is a one-to-one correspondence between the model diagram on Fig. 4a and its equations (Table 1). Like the diagram, the equations are the actual output from *iThink*® too. The equations corresponding to Fig. 8b are archived in [30]. Together, Cyprus' population, local tourism and monthly tourism determine the population and tourism sum (Eq. 1.11, Table 1). According to CYSTAT [11], both *Cyprus' Tourism Organization* and its government attach great importance to local tourism. A study on domestic tourism conducted in 1995 revealed that about 46 percent of Cypriots take long holidays. Of these, 61 percent take long holidays exclusively in Cyprus and eight percent in Cyprus and abroad, while 31 percent chose to travel abroad only. These are precisely the percentages in the model (Eq. 1.10, Table 1).

On Fig. 4a, the world land and population data, minus Cyprus' land, population and tourism co-determine the world EF (environmental footprint, Eq. 1.14, Table 1). Compared to Cyprus' smooth EF, i. e., the smooth ratio of the island's free land divided by its total population and tourism, the world EF gives a dynamic measure of Cyprus' relative attractiveness to the rest of the world. The EF ratio (Eq. 1.6, Table 1), i. e., the ratio of Cyprus' smooth EF (Eq. 1.13, Table 1) divided by the world EF (Eq. 1.14, Table 1), assumes that the higher this ratio is, the more attractive the island is to potential tourists, and vice versa. The EF ratio, which depends on Cyprus' total population and tourism, feeds back to the island's annual tourism via the inflow of foreign visitors who come to visit Cyprus every year (Fig. 4b).

The logistic or Verhulst growth model, after François Verhulst who published it in 1838 [66], helps explain Cyprus' actual annual tourism, a quantity that cannot grow forever (Fig. 4b). Every system that initially grows exponentially eventually approaches the carrying capacity of its environment, whether it is food supply for moose, the number of people susceptible to infection or the potential market for a good or a service. As an 'autopoietic' system approaches its limits to growth, it goes through a non-linear transition from a region where positive feedback dominates to a negative feedback dominated regime. S-shaped growth often results: a smooth transition from exponential growth to equilibrium.

The logistic model conforms to the requirements for S-shaped growth and the ecological idea of carrying capacity. The population it models typically grows in a fixed environment, such as Cyprus' foreign annual tourism has done since 1960 up to 2000. Initially dominated by positive feedback, Cyprus' annual tourism might soon reach the island's carrying capacity, with a nonlinear shift to dominance by negative feedback. While accounting for Cyprus' tourism lost to the summer of 1974 Turkish invasion, officially a very long 'military intervention', further depleting annual tourism is the outflow of Cyprus' visitors (not shown here) who might go as the island's free area reaches its Carrying Capacity, estimated at seventy times the number of Cyprus' visitors in 1960 [30].

Cyprus' Hotel Association listed Cyprus tourism seasonality as one of its major concerns. At the time of this investigation, CYSTAT [11] had compiled monthly tourism data for only 30 months. These were used for computing Cyprus' tourism seasonality (Fig. 4b). Incorporating both the foreign annual tourism and the monthly tourism stocks in the model allows both looking at the big picture of annual tourism growth and assessing the potential long-term effects of tourism seasonality on the sustainability of Cyprus' environment and hotel EBITDA, i. e., *e*arnings *b*efore *i*nterest, *t*axes, *d*epreciation and *a*mortization. The publicly available actual annual tourism data allow testing the model's usefulness, i. e., how faithfully it reproduces actual data between 1960 and 2000 [30].

Cyprus' foreign visitors and local tourists arrive at the island's hotels and resorts according to Cyprus' tourism seasonality, thereby feeding Cyprus' monthly tourism stock. About 11.3 days later, according to CYSTAT's [11] estimated average stay days, both foreign visitors and local tourists depart, thereby depleting the monthly tourism stock. By letting tourism growth = 0 and Cyprus' tourism seasonality continue repeating its established pattern, the model computes a zero-growth or *base-run* scenario. Subsequently, however, tourism growth values other than zero initiate different scenarios.

Cyprus' Tourism Growth Scenarios (Case 1)

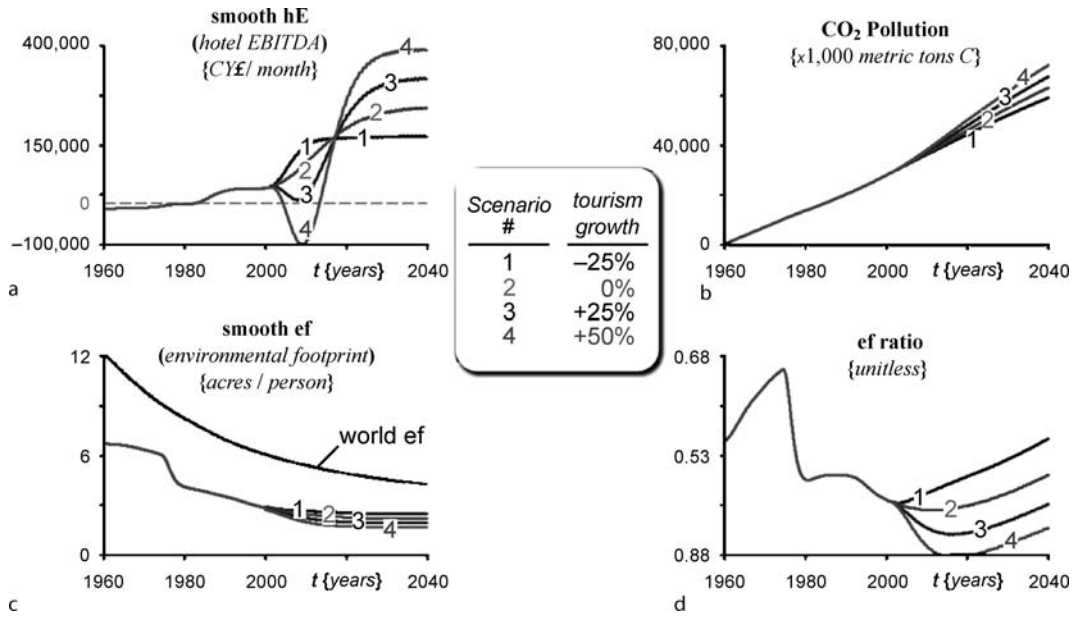
What can Cyprus' hoteliers expect to see in terms of bottom-line dynamics? According to the four tourism-growth scenarios computed on Fig. 5, seasonal variations notwithstanding, the higher Cyprus' tourism growth is, the lower hotel EBITDA (smooth hE) is, in the short term. In the long term, however, higher tourism growth yields higher profitability in constant year 2000 prices.

Scenario-Driven Planning with System Dynamics, Table 1

Cyprus' environment and population (and local tourism) model sector (Fig. 4a) equations, with variable, constant parameter and unit definitions

***Level or state variables* (stocks)**	***Eq.* #**
CO_2Pollution(t) = CO_2Pollution($t - dt$) + (emissions − clean up) ∗ dt	(1.1)
INIT CO_2 Pollution = emissions (Based on 1995 gridded carbon dioxide anthropogenic emission data; unit: 1000 metric ton C per one degree latitude by one degree longitude grid cell)	(1.1.1)
***Rate variables* (flows)**	
Emissions = emissions per person ∗ population and tourism (unit: 1000 metric tons C/month)	(1.2)
Cleanup = max(0, CO_2Pollution/average clean − up time) (unit: 1000 metric tons C/month)	(1.3)
***Auxiliary variables and constants* (converters)**	
Average clean − up time = 1200 (Med Sea region average self clean-up time = 100 years; unit: months)	(1.4)
Cyprus' land = If (time ≤ 168) then (9251 ∗ 247.1052) else ((9251 − 3355) ∗ 247.1052) (Cyprus' free land area; unit: acres; 1 km^2 = 247.1052 acres)	(1.5)
EF ratio = smooth EF/world EF (unit: unitless)	(1.6)
EF: environmental footprint = Cyprus' land/population and tourism (unit: acres/person)	(1.7)
Emissions per person = 1413.4/702000/12 (unit: anthropogenic emissions/person/month)	(1.8)
Local tourism = local tourism fraction ∗ Cyprus' population (unit: persons/month)	(1.9)
Local tourism fraction = 0.46 ∗ (0.61 + 0.08) (Percentages based on a 1995 study on domestic tourism; unit: unitless)	(1.10)
Population and tourism = Cyprus' population + Tourism − local tourism (Subtracts local tourists already included in Cyprus' population; unit: persons)	(1.11)
Sd clean − up time = 240 (clean-up time standard deviation = 20 years; unit: months)	(1.12)
Smooth EF = SMTH3 (EF: environmental footprint, 36) (Third-order exponential smooth of EF)	(1.13)
World EF = (world land − Cyprus' land)/(world population − population and tourism) (unit: acres/person)	(1.14)
World land = 36677577730.80 (unit: acres)	(1.15)
Cyprus' population = GRAPH(time/12) (Divided by 12 since these are annual data; unit: persons) (0.00, 493984), (1.00, 498898), (2.00, 496570), (3.00, 502001), (4.00, 505622), (5.00, 509329), (6.00, 512950), (7.00, 516743), (8.00, 520968), (9.00, 525364), (10.0, 529847), (11.0, 534330), (12.0, 539934), (13.0, 546486), (14.0, 552348), (15.0, 526313), (16.0, 516054), (17.0, 515881), (18.0, 518123), (19.0, 521657), (20.0, 526744), (21.0, 532692), (22.0, 538210), (23.0, 544675), (24.0, 551659), (25.0, 558038), (26.0, 560366), (27.0, 568469), (28.0, 572622), (29.0, 578394), (30.0, 587392), (31.0, 598217), (32.0, 609751), (33.0, 619658), (34.0, 626534), (35.0, 632082), (36.0, 636790), (37.0, 641169), (38.0, 645560), (39.0, 649759), (40.0, 653786), (41.0, 657686), (42.0, 661502), (43.0, 665246), (44.0, 668928), (45.0, 672554), (46.0, 676147), (47.0, 679730), (48.0, 683305), (49.0, 686870), (50.0, 690425), (51.0, 693975), (52.0, 697524), (53.0, 701056), (54.0, 704547), (55.0, 707970), (56.0, 711305), (57.0, 714535), (58.0, 717646), (59.0, 720613), (60.0, 723415), (61.0, 726032), (62.0, 728442), (63.0, 730629), (64.0, 732578), (65.0, 734280), (66.0, 735730), (67.0, 736928), (68.0, 737887), (69.0, 738627), (70.0, 739172), (71.0, 739540), (72.0, 739743), (73.0, 739792), (74.0, 739697), (75.0, 739472), (76.0, 739123), (77.0, 738658), (78.0, 738083), (79.0, 737406), (80.0, 737406)	(1.16)
World population = GRAPH(time/12) (Divided by 12 since these are annual data; unit: persons) (0.00, 3e+09), (1.00, 3.1e+09), (2.00, 3.1e+09), (3.00, 3.2e+09), (4.00, 3.3e+09), (5.00, 3.3e+09), (6.00, 3.4e+09), (7.00, 3.5e+09), (8.00, 3.6e+09), (9.00, 3.6e+09), (10.0, 3.7e+09), (11.0, 3.8e+09), (12.0, 3.9e+09), (13.0, 3.9e+09), (14.0, 4e+09), (15.0, 4.1e+09), (16.0, 4.2e+09), (17.0, 4.2e+09), (18.0, 4.3e+09), (19.0, 4.4e+09), (20.0, 4.5e+09), (21.0, 4.5e+09), (22.0, 4.6e+09), (23.0, 4.7e+09), (24.0, 4.8e+09), (25.0, 4.9e+09), (26.0, 4.9e+09), (27.0, 5e+09), (28.0, 5.1e+09), (29.0, 5.2e+09), (30.0, 5.3e+09), (31.0, 5.4e+09), (32.0, 5.4e+09), (33.0, 5.5e+09), (34.0, 5.6e+09), (35.0, 5.7e+09), (36.0, 5.8e+09), (37.0, 5.8e+09), (38.0, 5.9e+09), (39.0, 6e+09), (40.0, 6.1e+09), (41.0, 6.2e+09), (42.0, 6.2e+09), (43.0, 6.3e+09), (44.0, 6.4e+09), (45.0, 6.5e+09), (46.0, 6.5e+09), (47.0, 6.6e+09), (48.0, 6.7e+09), (49.0, 6.8e+09), (50.0, 6.8e+09), (51.0, 6.9e+09), (52.0, 7e+09), (53.0, 7e+09), (54.0, 7.1e+09), (55.0, 7.2e+09), (56.0, 7.2e+09), (57.0, 7.3e+09), (58.0, 7.4e+09), (59.0, 7.5e+09), (60.0, 7.5e+09), (61.0, 7.6e+09), (62.0, 7.6e+09), (63.0, 7.7e+09), (64.0, 7.8e+09), (65.0, 7.8e+09), (66.0, 7.9e+09), (67.0, 8e+09), (68.0, 8e+09), (69.0, 8.1e+09), (70.0, 8.1e+09), (71.0, 8.2e+09), (72.0, 8.3e+09), (73.0, 8.3e+09), (74.0, 8.4e+09), (75.0, 8.4e+09), (76.0, 8.5e+09), (77.0, 8.5e+09), (78.0, 8.6e+09), (79.0, 8.6e+09), (80.0, 8.7e+09)	(1.17)

High tourism growth implies accommodating overbooked hotel reservations for tourists who actually show up. Free cruises erode Cyprus' hotel EBITDA. The alternative is, however, angry tourists going off in hotel lobbies. Tourists have gotten angry at hotels before, but hotels have made the problem worse in recent years worldwide [16]. They have tightened check-in rules, doubled their renovations and increased the rate of overbooking by about 30 percent. The results can be explosive if one adds the record flight delays that travelers endure. Anyhow, free cruises to nearby Egypt and Israel sound much better than simply training employ-



Scenario-Driven Planning with System Dynamics, Figure 5
Four computed scenarios show how tourism growth might affect Cyprus' hotel EBITDA (smooth hE) and the island's environment, with carbon-dioxide (CO_2) pollution (adapted from and extending [30])

ees to handle unhappy guests that scream in hotel lobbies.

Eventually, as Cyprus' bed capacity increases and thereby catches up with tourism demand, there will be less overbooking and a few free cruises to erode Cyprus' hotel EBITDA. Given enough time for an initial bed capacity disequilibrium adjustment, in the long term, high tourism growth increases both hotel EBITDA (Fig. 5a) and cash [30].

In addition to their profound consequences for its hotel value-chain participants, Cyprus' tourism growth might also determine the fate of the island's environment. Depending on the island's population and emissions per person, high tourism growth implies high anthropogenic emissions feeding Cyprus' CO_2 Pollution. Anthropogenic CO_2 emissions attributed to the upward and downward movements of recurring tourist arrivals create much more stress and strain for the island's natural environment than a consistent stream of tourism with low seasonality would. High tourism growth lowers Cyprus' environmental footprint (EF). The summer 1974 Turkish *military intervention* has had a drastic negative effect on Cyprus' relative attractiveness because it reduced the island's free land by 41 percent.

Although qualitatively similar to the world's average EF after the invasion, Cyprus' environmental footprint is lower than the world's average EF (Fig. 5c), rendering the island's free area relatively less attractive as more foreign tourists visit. Manifested in the EF ratio (Fig. 5d), Cyprus becomes relatively less attractive as more visitors choose to vacation on the island's free area.

Qualitatively, Cyprus' CO_2 pollution scenarios (Fig. 5b) look exactly like the A2 scenario family of harmonized anthropogenic CO_2 emissions, which the *Intergovernmental Panel on Climate Change* (IPCC) computed to access the risks of human-induced climate change [54]. Like in the rest of the world, unless drastic changes in policy or technology alter the emissions per person ratio in the next 40 years, CO_2 pollution is expected to grow proportionally with Cyprus' tourism, degrading the island's environment.

Case 2: A Japanese Chemicals Keiretsu (JCK)

Home of NASA's *Johnson Space Center*, the Clear Lake region in Texas boasts strong high technology, biotechnology and specialty chemicals firms. Among them is JCK, whose recent investment helps the Clear Lake region continue its stalwart role in Houston's regional economic expansion [40].

An active member of a famous Japanese giant conglomerate, JCK's history begun in the late 1800s. Despite its long history, however, it has not been easy for JCK to evade the feedback loop that drives Japanese firms to man-

ufacture outside Japan. Since the 1950s, with Japan still recovering from WWII, the better Japanese companies performed, the better their national currency did. But the better Japan's currency did, the harder it became for its firms to export. The higher the yen, the more expensive and, therefore, less competitive Japan's exports become. This simple loop explains JCK's manufacturing lineage from Japan to USA [34].

But the transition process behind this lineage is not that simple. JCK's use of SdP with SD reveals a lot about its strategy design and implementation tactics. The model below shows a tiny fragment of JCK's gigantic effort to reperceive itself. The firm wants to see its keiretsu transform into an agile, virtual enterprise network (VEN) of active agents that collaborate to achieve its transnational business goals. Although still flying low under the media's collective radar screen, VENs receive increased attention by strategic managers [29].

Sterman (see Chap. 17 and 18 in [75]) presents a generic value-chain management structure that can unearth what VENs are about. By becoming a VEN, JCK is poised to bring the necessary people and production processes together to form *autopoietic*, i. e., self-organizing, customer-centric value chains in the specialty chemicals industry. JCK decided to build its own plant in USA because the net present value (NPV) of the anticipated combined cash flow resulting from a merger with other specialty chemicals manufacturers in USA would have been less then the sum of the NPVs of the projected cash flows of the firms acting independently. Moreover, JCK's own technology transfer cost is so low that the internalization cost associated with a merger would far exceed supplier charges plus market transaction costs. To remain competitive [62], JCK will not integrate the activity but offshoot it as a branch of its VEN-becoming keiretsu. The plant will be fully operational in January 2008. In order to maximize the combined *n*et *p*resent *v*alue of *e*arnings *b*efore *i*nterest, *t*axes, *d*epreciation and *a*mortization, i. e., NPV(EBITDA), of its new USA plant and the existing one in Asia, JCK wishes to improve its USA sales revenue before production starts in USA.

JCK's pre-production marketing tactics aim at building a sales force to increase sales in USA. Until the completion of the new plant (Dec. 2007), JCK will keep importing chemicals from its plant in Asia. Once production starts in USA (Jan. 2008), then the flow of goods from Asia to USA stops, the plant in USA supplies the USA market and the flow of goods from USA to Asia begins.

Strategic scenarios are not new to the chemical industry [82]. SdP with SD helps this specialty chemicals producer integrate its business intelligence efforts with strategy design in anticipation of environmental change. Modeling JCK's strategic situation requires a comprehensive inquiry into the environmental causalities and equivocalities that dictate its actions. Computed strategic and tactical scenarios probe the combined consequences of environmental trends, changes in JCK's own strategy, as well as the moves of its current and future competitors. The section below describes briefly how JCK plans to implement its transnational strategy of balanced marketing and production. This takes the form of a system dynamics simulation model, which precedes the interpretation of its computed scenarios.

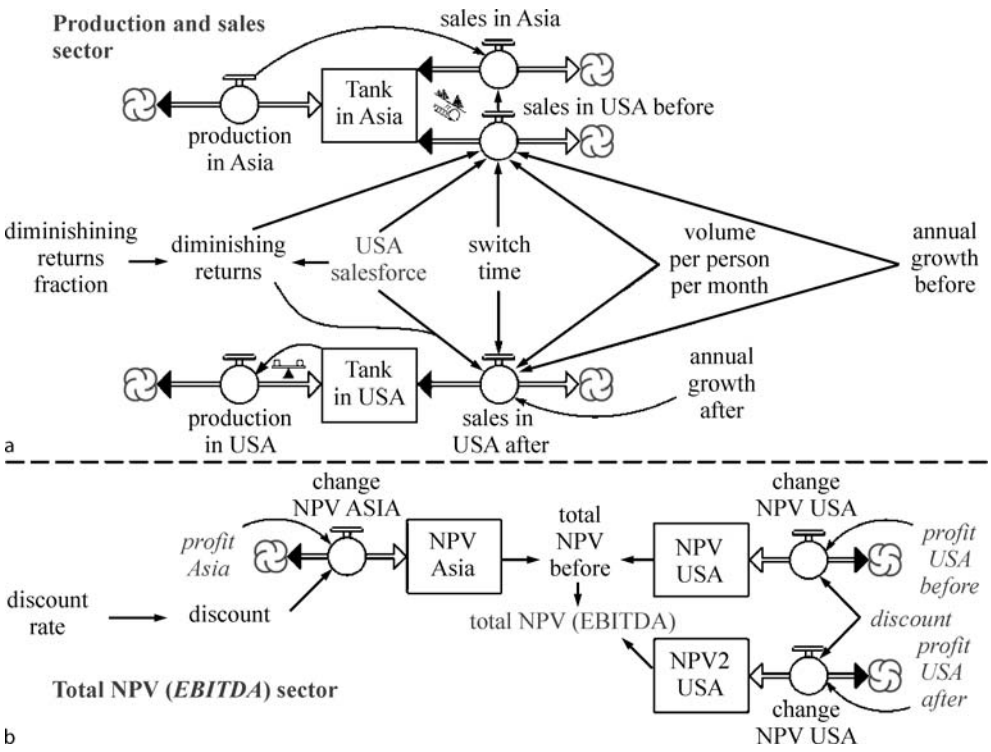
Model Description (Case 2)

The entire model has multiple sectors, four of which compute financial accounting data. Figure 6a shows the production and sales, and Fig. 6b the total NPV(EBITDA) model sectors. The corresponding algebra is in [34]. While JCK is building its USA factory, its factory in Asia makes and sells all specialty chemicals the USA market cannot yet absorb. This is what the *feed-forward* link from the production in Asia flow to the sales in Asia rate shows. The surplus demand JCK faces in Asia for its fine chemicals accounts for this rather unorthodox model structure. The surplus demand in Asia is the model's enabling *safety valve*, i. e., a major strategic assumption that renders tactical implementation feasible.

With the plant in Asia producing at full capacity before the switch, sales in the USA both depletes the tank in Asian stock and reduces sales in Asia. USA sales depend on JCK's USA sales force. But the size of this decision variable is just one determinant of sales in USA.

Sales productivity depends on many parameters, such as the annual growth before the switch rate of specialty chemicals in USA, the average expected volume a salesperson can sell per month as well as on the diminishing returns that sales people experience after the successful calls they initially make on their industrial customers. B2B or business to business, i. e., industrial marketing, can sometimes be as tough as B2C or business to customer, i. e., selling retail.

Time $t = 30$ months corresponds to January 2008, when the switch time converter cuts off the supply of JCK's chemicals from its plant in Asia. Ready by December 2007, the factory in the USA can supply the entire customer base its USA sales force will have been building for 30 months. As production in the USA begins, the sales in the USA before flow stops draining the tank in Asia and sales in Asia resume to match JCK's surplus demand there. Acting both as a production flow and as a continuous-review



Scenario-Driven Planning with System Dynamics, Figure 6
JCK's a production and sales, and b total NPV(EBITDA) model sectors (adapted from [34]; NPV = net present value, and EBITDA = earnings before interest, taxes, depreciation and amortization)

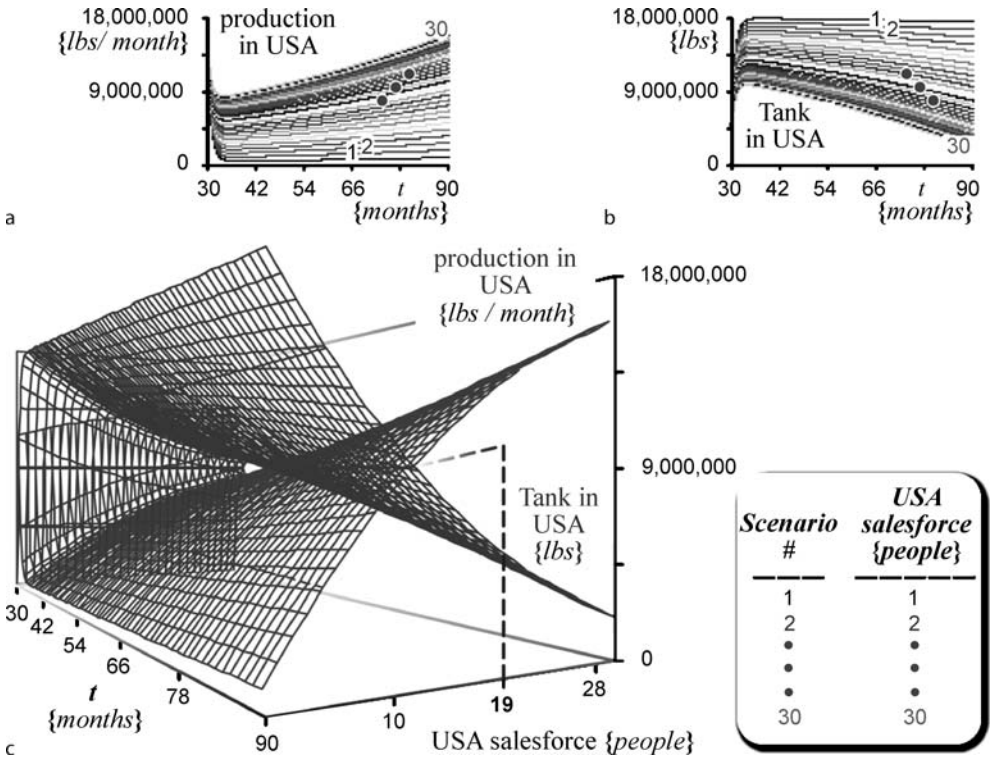
inventory order point, after January 2008, production in USA feeds the tank in USA stock of the rudimentary value-chain management structure on Fig. 6a.

Value chains entail stock and flow structures for the acquisition, storage and conversion of inputs into outputs, and the decision rules that govern the flows. The jet ski value chain includes, for example, hulls and bows that travel down monorail assembly paths. At each stage in the process, a stock of parts buffers production. This includes the inventory of fiberglass laminate between hull and bow acquisition and usage, the inventory of hulls and bows for the jet ski lower and upper structures, and the inventory of jet skis between dealer acquisition and sales. The decision rules governing the flows entail policies for ordering fiberglass laminate from suppliers, scheduling the spraying of preformed molds with layers of fiberglass laminate before assembly, shipping new jet skis to dealers and customer demand.

A typical firm's or VEN's value chain consists of cascades of supply chains, which often extend beyond a firm's boundaries. Effective value chain models must incorporate different agents and firms, including suppliers, the firm, distribution channels and customers. Scenario-driven planning with system dynamics is well suited for value chain modeling and policy design because value chains involve multiple stock and flow chains, with time lags or delays, and the decision rules governing the flows create multiple feedback loops among VEN members or value- and supply-chain partners (see Chap. 17 and 18 in [75]).

Back to JCK, its tank in the USA feeds information about its level back to production in the USA. Acting first as a decision point, production in the USA compares the tank in the USA level to the tank's capacity. If the tank is not full, then production in the USA places an order to itself and, once the USA factory has the requisite capacity, production in the USA refills the tank in the USA, but only until sales in the USA after the switch drains the tank. Then the cycle begins all over again.

Meanwhile, the profit in Asia, profit in the USA before and profit in the USA after sectors [34] perform all the financial accounting necessary to keep track of the transactions that take place in the value chain production and sales sector (Fig. 6a). As each scenario runs, the profit in Asia, the USA before and the USA after sectors feed the corresponding change in net present value (NPV) flows of



Scenario-Driven Planning with System Dynamics, Figure 7
Thirty computed scenarios show JCK's dual, smooth-switch and profitable purpose in production (adapted from [34])

the model's total NPV(EBITDA) sector (Fig. 6b). By adjusting each profit sector's EBITDA according to the discount rate, the change in NPV flows compute the total NPV(EBITDA) both in Asia and in the USA, both before and after JCK's January 2008 supply switch.

JCK's Computed Scenarios (Case 2)

Recall that the SdP with SD modeling-process spiral enabled our modeling team to crystallize JCK's strategic situation into the cyclical pattern that Fig. 3a shows. Although heavily disguised, the JCK measurement data and econometric sales functions let the system dynamics model compute scenarios to answer that razor-sharp optimization question the JCK executives asked:

> What size a USA sales force must we build in order to get a smooth switch in both sales and production in January 2008, and also to maximize the combined NPV(EBITDA) at our two plants in Asia and USA from now through 2012?

Treating the USA sales force policy parameter in the 'Sensi Specs…' menu item of *iThink*® allowed computing a set of 30 strategic scenarios. The 30 scenarios correspond to JCK's hiring from one to 30 sales people, respectively, to sell specialty chemicals to manufacturing firms in the USA, both before and after the January 2008 switch. Figures 7 and 8 show the 30 computed scenarios.

Figure 7c shows the response surfaces the production in USA rate and tank in the USA stock form after January 2008, in response to the 30 computed scenarios. The computed scenario that corresponds to JCK's building a USA sales force of 19 people achieves a smooth balance between sales in Asia and in the USA. Under this scenario, after January 2008, on the line where the two surfaces cross each other, not only the number of pounds of chemicals made and sold in Asia equals the number of pounds of chemicals made and sold in the USA, but as Fig. 7c shows, production in the USA also equals the tank in USA stock. So hiring 19 sales people now meets JCK's smooth switch in sales and production objective. But how?

How does producing and selling in the USA at rates equal to the corresponding rates in Asia constitute a fair response to JCK's smooth switch objective? The JCK executives seemed to accept this at face value. But our team had to clearly explain the dynamics of JCK's rudimen-

tary USA value chain (Fig. 6a), in order to unearth what the USA member of this transnational VEN-becoming keiretsu might be up to.

It looks simple, but the value chain of the production and sales sector on Fig. 6a can show the same amplification symptoms that much more elaborate value chains show when they fall pray to bullwhip effects. Locally rational policies that create smooth and stable adjustment of individual business units can, through their interaction with other functions and firms, cause oscillation and instability. Figure 8a shows the profound consequences of JCK's switch for its value chain in the USA. Because of the sudden switch in January 2008, the computed scenarios cause 30 sudden step changes. Both variables' adjustment rates increase, but the tank in the USA stock's amplification remains almost constant below 50 percent. As customer demand steps up, so do both metrics' new equilibrium points, but in direct proportion to the step increase in customer demand in the USA.

The 30 computed scenarios confirm Sterman's argument that, while the magnitude of amplification depends on stock adjustment times and delivery lags, its existence does not. No matter how drastically customers and firms downstream in a value chain change an orders' magnitude, they cannot affect supply chain amplification. Value chain managers must never blame customers and downstream firms or their forecasts for bullwhip effects. The production in USA amplification is almost double the tank in USA for a small USA sales force, suggesting that JCK's USA factory faces much larger changes in demand than its sales people do. Although temporary, during its disequilibrium adjustment, the tank in the USA consistently overshoots the new equilibrium points that it seeks after the switch (Fig. 7b), an inevitable consequence of stock and flow structure. Customers are innocent, but JCK's value chain structure is not:

First, the tank in the USA stock adjustment process creates significant amplification of the production in the USA rate. Though the tank in the USA relative amplification is 36.18 percent under the USA sales force = 1 scenario, for example, the relative amplification of production in the USA (Fig. 8a) increases by a maximum of more than 90 percent (the peak production in the USA rate, after t = 30 months, divided by the minimum production in the USA rate = 11, 766, 430.01/1, 026, 107.64 = 91.28 percent). The *amplification ratio*, i. e., the ratio of maximum change in output to maximum change in input, therefore is 91.28%/36.18% = 2.52. A one-percent increase in demand for JCK's chemicals causes a 2.52 percent surge in demand at JCK's USA plant. While the amplification ratio magnitude depends on the stock adjustment times and delivery lags, its existence does not (see p. 673 in [75]).

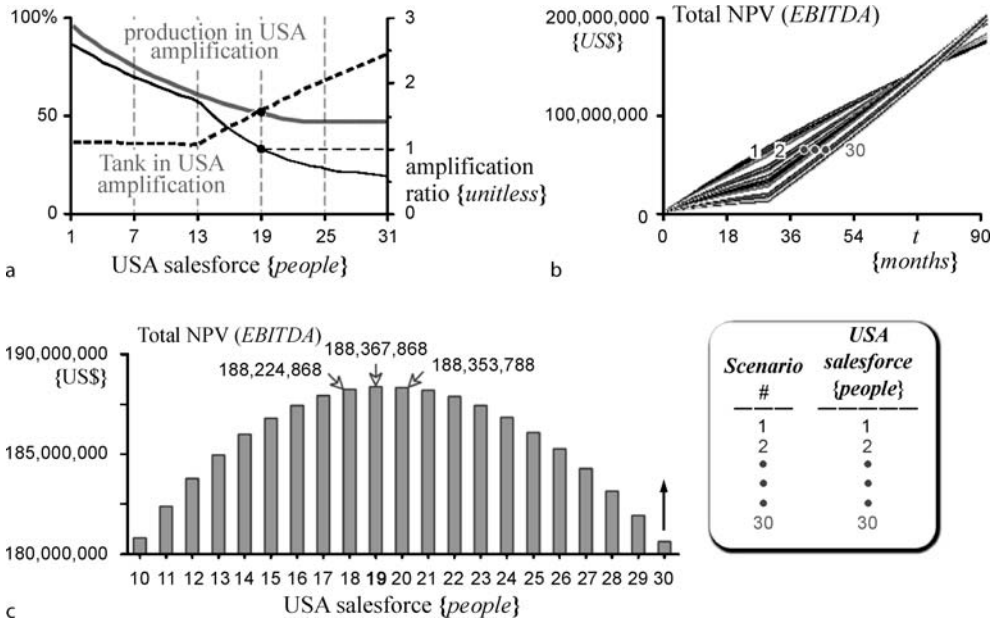
Second, amplification is temporary. In the long run, a one-percent increase in sales in the USA after leads to a one-percent increase in production in the USA. After two-adjustment times, i. e., two months, production in the USA gradually drops (Fig. 7a). During the disequilibrium adjustment, however, production in the USA overshoots its new equilibrium, an inevitable consequence of the stock and flow structure of customer-supplier value chains, no matter how tiny or simple they are. The only way the tank in the USA stock can increase is for its inflow production in the USA rate (order rate) to exceed its outflow rate sales in the USA after (Fig. 6a). Within a VEN's or keiretsu's customer-supplier value chain, supply agents face much larger changes in demand than finished-goods inventory, and the surge in demand is temporary.

The computed scenarios show that as the USA sales force increases, the production in the USA's rate of amplification declines because its new long-term equilibrium point is closer to its initial jump in January 2008. Conversely, as the tank in the USA stock's long-term equilibrium point remains consistently high because of the larger USA sales force, its relative amplification begins to rise. Since the two variables' relative amplification moves in opposite directions, eventually, they meet. What a coincidence! They meet above the USA sales force = 19 people. Now, is this not a much better interpretation of the word 'smooth' in fair response to JCK's smooth-switch performance purpose? The answer to JCK is now pertinent to its balancing its value chain in USA. With a USA sales force = 19, JCK's value chain components show equal relative amplification to sudden changes in demand, attaining nothing less than a magnificent amplification ratio = 1. Now that is smooth!

But what of profitability? JCK's polite executives said: "maximize… combined… NPV". In the time domain (Fig. 8b), total NPV(EBITDA) creates intricate dynamics that obscures the USA sales force effect. But the phase plot on Fig. 8c clearly shows a concave down behavior along the USA sales force: USA sales force = 19 maximizes the two plants' combined total NPV(EBITDA).

Future Directions

The above cases show how scenario-driven planning with system dynamics helps control performance by enabling organizational learning and the management of uncertainty. The strategic intelligence system that SdP with SD provides rests on the idea of a collective inquiry, which translates the environmental 'macrocosm' and a firm's



Scenario-Driven Planning with System Dynamics, Figure 8

Thirty computed scenarios show how hiring a sales force of 19 people in the USA might maximize JCK's NPV(EBITDA), and thereby fulfill its dual, smooth-switch and profitable purpose (adapted from [34])

'microcosm' into a shared causal map with computed scenarios. Informed discussion then takes place. Seeing SdP with SD as an inquiry system might help the outcomes of the situation formulation-solution-implementation sequence, each stage built on successive learning.

Strategic situations are complex and uncertain. Because planning is directed toward the future, predictions of changes in the environment are indispensable components of it. Conventional forecasting by itself provides no cohesive way of understanding the effect of changes that might occur in the future. Conversely, SdP with SD and its computed scenarios provide strategic intelligence and a link from traditional forecasting to modern interactive planning systems. In today's quest for managers who are more leaders than conciliators, the strategists' or executives' interest in scenarios must be welcomed. A clearer delineation of SdP with SD might make it a very rich field of application and research.

The SdP with SD inquiry system on Fig. 2 includes several contributions. *First*, by translating the environmental *macrocosm* and the firm *microcosm* into a common context for conceptualization, the requisites of theory building can be addressed. Planning analysts no longer have to operate piecemeal. A theory and a dominant logic typically emerge from shared perceptions about a firm, its environment and stakeholder purposes through model construction.

Second, the outputs of the strategic management process activities build on each other as successive layers. The SAST loop on Fig. 2 follows the counterclockwise direction of multiperspective dialectics [47]. This process allows adjustment of individual and organizational theories and logic, leading to an evolutionary interpretation of the real system that strategic decisions target.

Third, the inquiry system of Fig. 2 enables flexible support for all phases of strategy design. Problem finding or forming, or situation formulation receives equal attention as problem solving.

SdP with SD helps open up the black box of decision makers' mental models, so they can specify the ideas and rules they apply. That in turn helps enrich their language and label system, organizational capability and knowledge, and strategic decision processing system. Computed scenarios bring about transformation rules not previously thought of as well as new variables and interaction paths.

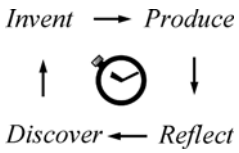
As an entity, each decision maker has a local scope and deals only with specific variables and access paths to other entities. But success factors are not etched in stone. Often, we only observe a representative state of each entity, namely, locally meaningful variables and parts of a scenario. This representativeness changes dynamically in the process of computing scenarios. Beyond the purely technical advantages of computed scenarios, planning becomes interactive, and language and label systems render them-

selves more adequate, effective and precise. Their associated organizational capability develops even more. In addition, the minor and major assumptions in decision makers' mental models surface as computed scenarios specify the conditions under which performance changes.

A line of great immediate concern requires researchers and practitioners alike to explore the modeling process behind SdP with SD. For the sake of realism, to make negotiated perceptions of reality explicit, we need representations where strategic real options and self-interest projections mold the way in which managers incorporate their observations and interpretations into strategy models. This is an unavoidable, most challenging path to tread, if we want to build a dialectical debate into the strategy design process.

Do we really want to? Yes because:

1. The traditional hierarchical organization dogma has been planning, managing and controlling, whereas the new reality of the learning organization incorporates vision, values and mental models. It entails training managers and teams in the IPRD learning cycle conceived by Dewey [14] (cf. Senge and Sterman [71]):



2. In the strategic management process (SMP) evolution, planning is evolving too, from objective-driven to budget-driven to strategy-driven to scenario-driven planning with system dynamics (SdP with SD, see pp. 271–272 in [32]).
3. The inquiry system that mediates the restructuring of organizational *theory in use* [68] determines the quality of organizational learning.

By looking into the dynamics of strategy design and the resulting performance of firms, the SdP with SD framework on Fig. 2 might let managers, planners and business researchers see the tremendous potential of computed strategic scenarios. They might choose to build intelligence systems around SdP with SD to create insight for strategy design. They will be building real knowledge in the process, while developing capability for institutional learning. Both Pascale [59] and de Geus [13] see the capability to speed up institutional learning as a truly sustainable competitive advantage.

Bibliography

Primary Literature

1. Acar W (1983) Toward a theory of problem formulation and the planning of change: Causal mapping and dialectical debate in situation formulation. UMI, Ann Arbor
2. Ackoff RL (1981) Creating the corporate future. Wiley, New York
3. Ackoff RL, Emery FE (1972) On purposeful systems. Aldine-Atherton, Chicago
4. Amara R, Lipinski AJ (1983) Business planning for an uncertain future: Scenarios and strategies. Pergamon, New York
5. Anderson TJ (2000) Real options analysis in strategic decision making: An applied approach in a dual options framework. J Appl Manag Stud 9(2):235–255
6. Ansoff HI (1985) Conceptual underpinnings of systematic strategic management. Eur J Oper Res 19(1):2–19
7. Ansoff HI, McDonnell E (1990) Implanting strategic management, 2nd edn. Prentice-Hall, New York
8. Brauers J, Weber M (1988) A new method of scenario analysis for strategic planning. J Forecast 7(1):31–47
9. Brenkert AL (1998) Carbon dioxide emission estimates from fossil-fuel burning, hydraulic cement production, and gas flaring for 1995 on a one degree grid cell basis. Oak Ridge National Laboratory, Carbon Dioxide Information Analysis Center, Oak Ridge, TN (Database: NDP-058A 2-1998)
10. Christensen CM (1997) The innovator's dilemma: When new technologies cause great firms to fail. Harvard Business School Press, Cambridge
11. CYSTAT (2000) Cyprus key figures: Tourism. The Statistical Service of Cyprus (CYSTAT), Nicosia
12. Daft RL, Weick KE (1984) Toward a model of organizations as interpretation systems. Acad Manag Rev 9:284–295
13. de Geus AP (1992) Modelling to predict or to learn? Eur J Oper Res 59(1):1–5
14. Dewey J (1938) Logic: The theory of inquiry. Holt, Rinehart and Winston, New York
15. Donaldson L (1992) The Weick stuff: Managing beyond games. Organ Sci 3(4):461–466
16. Drucker J, Higgins M (2001) Hotel rage: Losing it in the lobby. Wall Street J (Fri 16 Feb):W1–W7
17. Duncan RB (1972) Characteristics of organizational environments and perceived environmental uncertainty. Adm Sci Q 17:313–327
18. Eberlein RL (2002) Vensim® PLE Software, V 5.2a. Ventana Systems Inc, Harvard
19. Eden C (1994) Cognitive mapping and problem structuring for system dynamics model building. Syst Dyn Rev 10(3):257–276
20. Eden C (2004) Analyzing cognitive maps to help structure issues or problems. Eur J Oper Res 159:673–686
21. Eilon S (1984) The Art of Reckoning: Analysis of performance criteria. Academic Press, London
22. Emery FE, Trist EL (1965) The causal texture of organizational environments. Hum Relat 18:21–32
23. Forrester JW (1958) Industrial dynamics: A major breakthrough for decision makers. Harvard Bus Rev 36(4):37–66
24. Forrester JW (1961) Industrial dynamics. MIT Press, Cambridge
25. Forrester JW (1987) Lessons from system dynamics modeling. Syst Dyn Rev 3(2):136–149

26. Forrester JW (1992) Policies, decisions and information sources for modeling. Eur J Oper Res 59(1):42–63
27. Georgantzas NC (1995) Strategy design tradeoffs-free. Hum Syst Manag 14(2):149–161
28. Georgantzas NC (2001) Simulation modeling. In: Warner M (ed) International encyclopedia of business and management, 2nd edn. Thomson Learning, London, pp 5861–5872
29. Georgantzas NC (2001) Virtual enterprise networks: The fifth element of corporate governance. Hum Syst Manag 20(3):171–188, with a 2003 reprint ICFAI J Corp Gov 2(4):67–91
30. Georgantzas NC (2003) Tourism dynamics: Cyprus' hotel value chain and profitability. Syst Dyn Rev 19(3):175–212
31. Georgantzas NC (2007) Digest® wisdom: Collaborate for win-win human systems. In: Shi Y et al (eds) Advances in multiple criteria decision making and human systems management. IOS Press, Amsterdam, pp 341–371
32. Georgantzas NC, Acar W (1995) Scenario-driven planning: Learning to manage strategic uncertainty. Greenwood-Quorum, Westport
33. Georgantzas NC, Katsamakas E (2007) Disruptive innovation strategy effects on hard-disk maker population: A system dynamics study. Inf Resour Manag J 20(2):90–107
34. Georgantzas NC, Sasai K, Schrömbgens P, Richtenburg K et al (2002) A chemical firm's penetration strategy, balance and profitability. In: Proc of the 20th international system dynamics society conference, 28 Jul–1 Aug, Villa Igiea, Palermo, Italy
35. Gharajedaghi J (1999) Systems thinking: Managing chaos and complexity – A platform for designing business architecture. Butterworth-Heinemann, Boston
36. Godet M (1987) Scenarios and strategic management: Prospective et planification stratégique. Butterworths, London
37. Godet M, Roubelat F (1996) Creating the future: The use and misuse of scenarios. Long Range Plan 29(2):164–171
38. Hax AC, Majluf NS (1996) The strategy concept and process: A pragmatic approach, 2nd edn. Prentice Hall, Upper Saddle River
39. Helmer O (1983) Looking forward. Sage, Beverly Hills
40. Hodgin RF (2001) Clear Lake Area Industry and Projections 2001. Center for Economic Development and Research, University of Houston-Clear Lake, Houston
41. Huss WR, Honton EJ (1987) Scenario planning: What style should you use? Long Range Plan 20(4):21–29
42. Istvan RL (1992) A new productivity paradigm for competitive advantage. Strateg Manag J 13(7):525–537
43. Jarillo JC (1988) On strategic networks. Strateg Manag J 9(1):31–41
44. Jarillo JC, Martínez JI (1990) Different roles for subsidiaries: The case of multinational corporations in Spain. Manag J 11(7):501–512
45. Kahn H, Wiener AJ (1967) The next thirty-three years: A framework for speculation. Daedalus 96(3):705–732
46. Lissack MR, Roos J (1999) The next common sense: The e-manager's guide to mastering complexity. Nicholas Brealey Publishing, London
47. Mason RO, Mitroff II (1981) Challenging strategic planning assumptions. Wiley, New York
48. Miller D, Friesen PH (1983) Strategy making and environment: The third link. Strateg Manag J 4:221–235
49. Millet SM, Randles F (1986) Scenarios for strategic business planning: A case history for aerospace and defence companies. Interfaces. 16(6):64–72
50. Mojtahedzadeh MT, Andersen D, Richardson GP (2004) Using Digest® to implement the pathway participation method for detecting influential system structure. Syst Dyn Rev 20(1):1–20
51. Moore GA (1991) Crossing the chasm. Harper-Collins, New York
52. Morecroft JDW (1985) Rationality in the analysis of behavioral simulation models. Manag Sci 31:900–916
53. Morecroft JDW (1988) System dynamics and microworlds for policymakers. Eur J Oper Res 35:310–320
54. Nakicenovic N, Davidson O, Davis G et al (2000) Summary for policymakers–Emissions scenarios: A special report of working group III of the intergovernmental panel on climate change (IPCC). World Meteorological Organization (WMO) and United Nations Environment Programme (UNEP), New York
55. Nicolis JS (1986) Dynamics of hierarchical systems: An evolutionary approach. Springer, Berlin
56. Oliva R (2004) Model structure analysis through graph theory: Partition heuristics and feedback structure decomposition. Syst Dyn Rev 20(4):313–336
57. Oliva R, Mojtahedzadeh MT (2004) Keep it simple: A dominance assessment of short feedback loops. In: Proc of the 22nd international system dynamics society conference, 25–29 July 2004, Keble College, Oxford University, Oxford UK
58. Ozbekhan H (1977) The future of Paris: A systems study in strategic urban planning. Philos Trans Royal Soc London. 387:523–544
59. Pascale RT (1984) Perspectives on strategy: The real story behind Honda's success. California Manag Rev 26(Spring):47–72
60. Pine BJ-I, Victor B, Boynton AC (1993) Making mass customization work. Harvard Bus Rev 71(5):108–115
61. Porter ME (1985) Competitive advantage: Creating and sustaining superior performance. Free Press, New York
62. Porter ME (1991) Towards a dynamic theory of strategy. Strateg Manag J 12(Winter Special Issue):95–117
63. Randers J (1980) Guidelines for model conceptualization. In: Randers J (ed) Elements of the system dynamics method. MIT Press, Cambridge, pp 117–139
64. Raynor ME (2007) The strategy paradox: Why committing to success leads to failure [And What to Do About It]. Currency-Doubleday, New York
65. Repenning NP (2003) Selling system dynamics to (other) social scientists. Syst Dyn Rev 19(4):303–327
66. Richardson GP (1991) Feedback thought in social science and systems theory. University of Pennsylvania Press, Philadelphia
67. Richmond B et al (2006) iThink® Software V 9.0.2. iSee Systems™, Lebanon
68. Schön D (1983) Organizational learning. In: Morgan G (ed) Beyond method. Sage, London
69. Schrage M (1991) Spreadsheets: Bulking up on data. San Francisco Examiner
70. Schwenk CR (1984) Cognitive simplification processes in strategic decision making. Strateg Manag J 5(2):111–128
71. Senge PM, Sterman JD (1992) Systems thinking and organizational learning: Acting locally and thinking globally in the organization of the future. Eur J Oper Res 59(1):137–150
72. Simon HA (1979) Rational decision making in business organizations. Am Econ Rev 69(4):497–509
73. Singer AE (1992) Strategy as rationality. Hum Syst Manag 11(1):7–21
74. Singer AE (1994) Strategy as moral philosophy. Strateg Manag J 15:191–213

75. Sterman JD (2000) Business dynamics: Systems thinking and modeling for a complex world. Irwin McGraw-Hill, Boston
76. Sterman JD, Wittenberg J (1999) Path dependence, competition and succession in the dynamics of scientific revolution. Organ Sci 10(3, Special issue: Application of complexity theory to organization science, May–Jun):322–341
77. Turner F (1997) Foreword: Chaos and social science. In: Eve RA, Horsfall S, Lee ME (eds) Chaos, complexity and sociology. Sage, Thousand Oaks, pp xi–xxvii
78. Wack P (1985) Scenarios: Uncharted waters ahead. Harvard Bus Rev 63(5):73–89
79. Wack P (1985) Scenarios: Shooting the rapids. Harvard Bus Rev 63(6):139–150
80. Zeleny M (1988) Parallelism, integration, autocoordination and ambiguity in human support systems. In: Gupta MM, Yamakawa T (eds) Fuzzy logic in knowledge-based systems, decision and control. Elsevier Science, North Holland, pp 107–122
81. Zeleny M (2005) Human systems management: Integrating knowledge, management and systems. World Scientific, Hackensack
82. Zentner RD (1987) Scenarios and the chemical industry. Chem Marketing Manag (Spring):21–25

Books and Reviews

Bazerman MH, Watkins MD (2004) Predictable surprises: The disasters you should have seen coming and how to prevent them. Harvard Business School Press, Boston

Bower G, Morrow D (1990) Mental models in narrative comprehension. Science 247(4938):44–48

Mittelstaedt RE (2005) Will your next mistake be fatal? Avoiding the chain of mistakes that can destroy your organization. Wharton School Publishing, Upper Saddle River

Morecroft JDW (2007) Strategic modeling and business dynamics: A feedback systems approach. Wiley, West Sussex

Schnaars SP (1989) Megamistakes: Forecasting and the myth of rapid technological change. The Free Press, New York

Schwartz P (1991) The art of the long view. Doubleday-Currency, New York

Tuchman B (1985) The March of folly: From troy to vietnam. Ballantine Books, New York

Vennix JAM (1996) Group model building: Facilitating team learning using system dynamics. Wiley, Chichester

Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space

Gert Zöller[1], Sebastian Hainzl[2], Yehuda Ben-Zion[3], Matthias Holschneider[1]

[1] Institute of Mathematics and Centre for Dynamics of Complex Systems, University of Potsdam, Potsdam, Germany

[2] GFZ German Research Centre for Geosciences, Potsdam, Germany

[3] Department of Earth Sciences, University of Southern California, Los Angeles, USA

Article Outline

Glossary
Definition of the Subject
Introduction
Modeling Seismicity in Real Fault Regions
Results
Summary and Conclusions
Future Directions
Acknowledgments
Bibliography

Glossary

Bayesian analysis A model estimation technique that accounts for incomplete knowledge. Bayes' theorem is a mathematical formulation of how an a priori estimate of the probability of an event can be updated, if a new information becomes available.

Critical earthquake concept The occurrence of large earthquakes may be described in terms of statistical physics and thermodynamics. In this view, an earthquake can be interpreted as a critical phase transition in a system with many degrees of freedom. The preparatory process is characterized by acceleration of the seismic moment release and growth of the spatial correlation length as in the percolation model. This interpretation of earthquake occurrence is referred to as the critical earthquake process.

Earthquake forecast/prediction The forecast or prediction of an earthquake is a statement about time, hypocenter location, magnitude, and probability of occurrence of an individual future event within reasonable error ranges.

Fault model A fault model calculates the evolution of slip, stress, and related quantities on a fault segment or a fault region. The range of fault models varies from conceptual models of cellular automaton or slider-block type to detailed models for particular faults.

Probability A quantitative measure of the likelihood for an outcome of a random process. In the case of repeating a random experiment a large number of times (e. g. flipping a coin), the probability is the relative frequency of a possible outcome (e. g. head). A different view of probability is used in the → Bayesian analysis.

Seismic hazard The probability that a given magnitude (or peak ground acceleration) is exceeded in a seismic source zone within a pre-defined time interval, e. g. 50 years, is denoted as the seismic hazard.

Self-organized criticality Self-organized criticality (SOC) as introduced by Bak [2] is the ability of a system to organize itself in

the vicinity of a critical point independently of values of physical parameters of the system and initial conditions. Self-organized critical systems are characterized by various power law distributions. Examples include models of sandpiles and forest-fires.

Definition of the Subject

The most fundamental question in earthquake science is whether earthquake prediction is possible. Related issues include the following: Can a prediction of earthquakes solely based on the emergence of seismicity patterns be reliable? In other words, is there a single or several "magic" parameters, which become anomalous prior to a large earthquake? Are pure observational methods without specific physical understanding, like the pattern recognition approach of Keilis–Borok and co-workers [41], sufficient? Taking into account that earthquakes are monitored continuously only for about 100 years and the best available data sets ("earthquake catalogs") cover only a few decades, it seems questionable to forecast earthquakes solely on the basis of observed seismicity patterns. This is because large earthquakes have recurrence periods of decades to centuries; consequently, data sets for most regions include less than ten large events making a reliable statistical testing questionable.

In the studies discussed here, the goal is not to forecast individual earthquakes. Instead, we aim at developing a combined approach based on numerical modeling and data analysis in order to understand seismicity and the emergence of patterns in the occurrence of earthquakes. The discussion and interpretation of seismcity in terms of statistical physics leads to the concept of "critical states", i. e. states in the seismic cycle with an increased probability for abrupt changes involving large earthquakes. A more general goal of this work is to provide perspectives for the understanding of the relevant mechanisms and to give outlines for developments related to time-dependent seismic hazard.

Introduction

Several empirical relationships for the occurrence of seismicity are well-known. The most common one is probably the Gutenberg–Richter law [30] for the relation between frequency and magnitude of earthquakes in a large seismically active region,

$$\log N = a - bM\,, \tag{1}$$

where N is the frequency of earthquakes with magnitude equal to or greater than M; a is a measure of the overall seismicity level in the region and the b value determines the relation between large and small earthquakes. The Gutenberg–Richter law provides an important constraint for the design of physical models and serves as a key ingredient for seismic hazard estimations. Statistical relations for the temporal occurrence of large events are less well known, because the corresponding data records are too short.

Several additional problems exist in the understanding and interpretation of observed seismicity patterns. First, it is important to decide whether an observed pattern has a physical origin or is an artifact, arising for example from inhomogeneous reporting or from man-made seismicity like quarry blasts or explosions [69]. Second, the non-artificial events have to be analyzed with respect to their underlying mechanisms. This leads to an inverse problem with a non-unique solution, which can be illustrated for the most pronounced observed temporal pattern associated with aftershocks. It is empirically known that the earthquake rate $\dot{N}$ after a large event at time t_M follows the Omori–Utsu law [49,67]

$$\dot{N} = \frac{K}{(c + t - t_M)^p}\,, \tag{2}$$

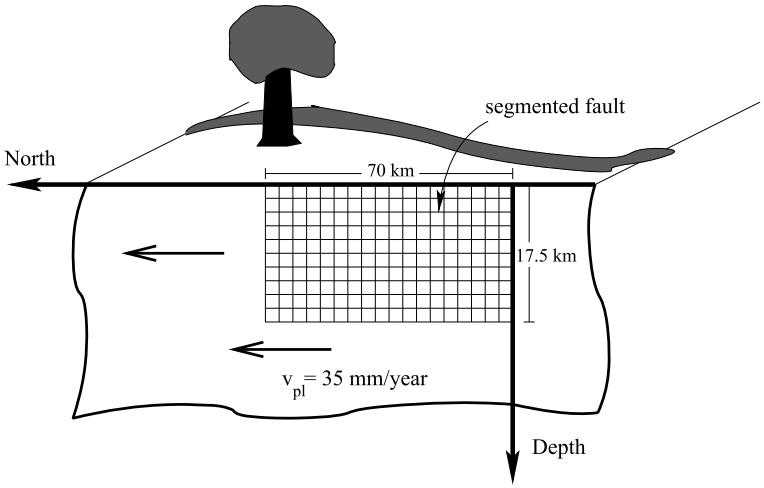
where t is the time, K and c are constants, and the Omori exponent p is close to unity. In particular, aftershocks are an almost universal phenomenon; that is, they are observed nearly after each mainshock. The underlying mechanisms leading to aftershocks are, however, unknown. Various physical models have been designed to explain aftershock occurrence following Eq. (2). These models include viscoelasticity [32], pore fluid flow [46], damage rheology [9,57], and rate-state friction [24]. The question which mechanism or combination of mechanisms is relevant in a given fault zone remains open. Detailed comparisons of observed and modeled seismicity with respect to the aftershock rate, the duration of aftershock sequences, the dependence on the mainshock size, and other features are necessary to address this problem. Additionally, results from lab experiments on rupture dynamics and satellite observations of deformation provide important information for the design of such models.

Apart from aftershock activity, other seismicity patterns are occasionally associated with observations, including foreshocks [39], seismic quiescence [34,72,78], and accelerating moment release [17,38]. These patterns have been documented in several cases before large earthquakes. They occur, however, far less frequently than aftershocks. For example, foreshocks are known to preceed only 20–30% of large earthquakes [71]. Therefore, their predictive power is questionable. Moreover, it is not clear

whether these patterns can be attributed to physical processes or to random fluctuations in the highly sparse and noisy earthquake catalogs. This problem can be addressed by using fault models which simulate long and complete earthquake sequences over thousands of years. If the models capture the main features of the underlying physics, the occurrence of seismicity patterns can be studied with reasonable statistics. The main ingredients of such models are the geometry of a fault region, empirically known constitutive laws, spatial heterogeneities, and stress and displacement functions following dislocation theory [20,47]. In order to allow for detailed studies of the relations between the imposed mechanisms and the observed seismicity functions, it is important that the number of adjustable parameters is limited.

It is emphasized that these models do not aim to reproduce an observed earthquake catalog in detail. Instead, the main goal is to address questions like: Why is the Parkfield segment of the San Andreas fault characterized by relatively regular occurrence of earthquakes with magnitude $M \approx 6$, while on the San Jacinto fault in California the properties of earthquake occurrence are more irregular? Basic models for seismicity are mainly based on one or more solid blocks, which are driven by a plate over a sliding surface. The plate and the blocks are connected with springs. This model can generate stick-slip events considered to represent earthquakes. The slider-block models can produce a wide range of complexity, beginning with a single block model leading to periodic occurrence of events of uniform size, and progressing to an array of connected blocks [18] leading to complex sequences of events with variable size. In order to reduce the computational effort cellular automata are commonly used [42,48]. Mathematically, these models include maps instead of differential equations; physically, this corresponds to instantaneously occurring slip events, neglecting inertia effects. The main ingredients of slider-block and cellular automaton models are (1) external driving (plate motion), and (2) sudden local change of system parameters (stress), when a critical value (material strength) is reached, followed by an avalanche of block slips (stress drop and coseismic stress transfer during an earthquake). While the first process lasts for years to several hundred years, the second occurs on a time scale of a few seconds. The simplest model including these features has been formulated by Reid [52] and is known as *Reid's elastic rebound theory*; in terms of slider-block models, this corresponds to a single block model with constant plate velocity. Accounting for spatial heterogeneity and fault segmentation, many interacting blocks, or fault segments, have to be considered. This leads to a spatiotemporal stress field instead of a single stress value. In general, the material strength will also become space-dependent. Such a model framework can be treated with the methodology of statistical physics similar to the Ising model or percolation models [43]. In this context, large earthquakes are associated with second-order phase transitions [2,59,64]. The view of earthquakes as phase transitions in a system with many degrees of freedom and an underlying critical point, is hereinafter referred to as the "critical point concept". The period before such a phase transition is expected to be characterized by a preparation process including development of power laws and growing spatial correlation length [14]. However, depending on the parameters of a model, different situations are conceivable: the system trajectory can enter the critical state and the critical point frequently ("supercritical") or it may never becomes critical ("subcritical"). A case of special interest is the class of models [2] showing *self-organized criticality* (SOC), which have their origin in a simple cellular automaton model for a sandpile [3]. In this case the system drives itself permanently to the vicinity of the critical point with almost scale-free characteristics. Consequently, each small event can grow into a large earthquake with some probability [28].

Long simulations of earthquake activity can be used to calculate statistical features like the recurrence time distribution of large earthquakes and the frequency-size distribution with high precision. Despite the scaling behavior (Eq. (1)) in the earthquake magnitudes for small and intermediate earthquakes, which is observed for many sets of model parameters, clear deviations become visible for large magnitudes. Such deviations are known from real catalogs, but their statistical significance is not clear in all cases. The model simulations suggest that deviations from scaling for strong earthquakes can be attributed to physical properties. One important property is the spatial disorder of brittle parameters of the fault. The presence of strong heterogeneities suppresses system-wide events with some probability, whereas such events can evolve more easily on smooth faults. The degree of quenched (time-independent) spatial heterogeneity turns out to be a key parameter for statistical and dynamical properties of seismicity [5,12,80]. This includes the temporal regularity of mainshock occurrence, various properties of the stress and displacement fields, and a spontaneous mode-switching between different dynamical regimes without changing parameters. The degree of heterogeneity can act as a tuning parameter that allows for a continuous change of the model dynamics between the end-member cases of supercritical and subcritical behavior. Such a dependence, which is observed also for other parameters, can be visual-



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 1
Sketch of the fault model framework

ized in phase diagrams similar to the phase diagram for the different aggregate states of water [22,79,80]. For increasing complexity of a model, the number of axes of the phase diagram, representing the relevant model parameters, will increase. The above mentioned question of distinguishing different faults like the Parkfield segment and the San Jacinto fault can be rephrased as the problem of assigning the faults to different regions in such a diagram. An important step in this direction is the physical modeling of observed seismicity patterns, including universal patterns like aftershocks (Eq. (2)), common fluctuations like foreshocks and the acceleration of seismic energy release before large earthquakes. The latter phenomenon which sometimes occurs over large regions including more than one fault, can be interpreted in terms of the approach towards a critical point. This view is supported by an observational study of the growth of the spatial correlation length which is a different aspect of the same underlying physics [73,75,76,77].

The establishment of relationships between model parameters and observational features may be used to tune the model towards a specific fault zone, and use the tuned fault model for practical applications of seismic hazard estimations. Toward this end the recurrence time distribution of large earthquakes is needed. Since observational data records are often short and noisy, the use of Bayesian probability theory is helpful for the estimation of uncertain model parameters, and the incorporation of various types of observational data in seismic hazard estimations. The Parkfield segment, as one of the best monitored seismically-active regions, serves as an excellent natural laboratory for such a case study. A discussed example illustrates how partially known parameters like the stress drop and the seismic hazard can be estimated by combining numerical models and observational data [74].

In Section "Modeling Seismicity in Real Fault Regions", the physical fault model used for the discussed studies is described. Results from numerical simulations are presented in Sect. "Results". A summary is given in Sect. "Summary and Conclusions".

Modeling Seismicity in Real Fault Regions

Numerous frameworks have been used to simulate seismicity (see e. g. [6,9,10,18,32,37,48] and references therein). These include slider-block models, cellular automata, "inherently discrete" fault models where the discreteness is an inherent feature of the imposed physics, and continuum models. In this section we illustrate how a fault model (Fig. 1) can be adjusted in order to simulate seismicity of a real fault region, e. g. the Parkfield segment of the San Andreas fault in California.

Fault Geometry and Model Framework

A first constraint for a specific model is to represent the geometry of the fault segment. As shown in Fig. 2, the region of Parkfield is characterized by a distribution of fault segments, which have generally the same orientation. It is therefore reasonable to map these segments in the model on a plane intersecting the surface at a straight line from SE to NW. The dimensions of the fault segment for modeling (Fig. 1) are chosen to be 70 km in length and 17.5 km in depth. As discussed in [10], this geometry corresponds approximately to the San Andreas fault near Parkfield. The

entire fault is an infinite half-plane, but the brittle processes are calculated on the above rectangular section referred to below as the computational grid. The computational grid is discretized to 128×32 cells of uniform size, where stress and slip are calculated. The size of the cells is not related here to observations; rather it depends on the magnitude range under consideration and the computational effort. The failure of a single cell defines the lowest magnitude. A higher resolution of the grid with same overall dimensions increases the magnitude range, because the magnitude is calculated from the slip of all cells during an earthquake. Following Ben–Zion and Rice [10], the material surrounding the fault is assumed to be a homogeneous elastic half space, which is characterized by elastic parameters and a related Green's function:

1. The elastic properties are expressed by the Lamé constants λ and μ, which connect stress and strain in Hooke's law. For many rocks, these constants are almost equal; therefore we use $\lambda = \mu$ with μ being the rigidity. An elastic solid with this property is called a *Poisson solid*. Because the strain is dimensionless, μ has the same dimension as the stress. In the present study, we use $\mu = 30\,\mathrm{GPa}$.
2. The (static) Green's function $G(\vec{y}_1, \vec{y}_2)$ defines the static response of the half space at a position $\vec{y}_1$ to a displacement at $\vec{y}_2$, which may arise from (coseismic) slip or (aseismic) creep motion. Due to the discretization of the fault plane into computational cells, we use the Green's function for static dislocations on rectangular fault patches of width $\mathrm{d}x$ and height $\mathrm{d}z$, which is given in [20] and [47]. The main difference between this Green's function and the nearest-neighbor interaction of most slider-block models and cellular automata is the infinite-range interaction following a decay according to $1/r^3$, where r is the distance between source cell and receiver point.

Interseismic Processes

The motion of the tectonic plates, indicated in Fig. 2, is responsible for the build-up of stress in the fault zone. Geodetic measurements of surface displacements provide estimates of the velocity of the plates. For the San Andreas fault, a value of $v_{\mathrm{pl}} = 35\,\mathrm{mm/year}$ as a long-term average [55] is widely accepted and is adopted for the model. The displacement $\mathrm{d}u$ in the regions surrounding the grid during a time period Δt is simply $\mathrm{d}u = v_{\mathrm{pl}} \cdot \Delta t$. While the average slip rate $\dot{u}$ is independent of the location of a cell, the stress rate $\dot{\tau}$ depends on space. The assumption that the computational grid is embedded in a half-plane which undergoes constant creep, implies that cells at the boundaries of the grid experience higher load than cells in the center of the grid. The Green's function $G(i, j; k, l)$ defines the interaction of points (i, j) and (k, l) in the medium. In particular, the stress response at a position (i, j) to a static change of the displacement field $\mathrm{d}u(k, l)$ is given by

$$\mathrm{d}\tau(i, j) = - \sum_{(k,l)\in \text{ half space}} G(i, j; k, l) \cdot \mathrm{d}u(k, l)\,, \qquad (3)$$

where the minus sign stems from the fact that forward (right-lateral) slip of regions around a locked fault segment is equivalent to back (left-lateral) slip of the locked fault segment. Taking into account that

$$\sum_{(k,l)\in \text{ half space}} G(i, j; k, l) = 0\,, \qquad (4)$$

Eq. (3) can be written as

$$\tau(i, j; t) = - \sum_{(k,l)\in \text{ half space}} G(i, j; k, l) \cdot [u(k, l; t) - v_{\mathrm{pl}} t]\,, \qquad (5)$$

where $u(k, l; t)$ is the total displacement at position (k, l) and time t since the start of the simulation. Because the surrounding regions sustain stable sliding, $u(k, l; t) = v_{\mathrm{pl}} t$ for $(k, l) \notin$ grid, the slip deficit outside the fault region vanishes and it is sufficient to perform the summation on the computational grid:

$$\tau(i, j; t) = \sum_{(k,l)\in \text{ grid}} G(i, j; k, l) \cdot [v_{\mathrm{pl}} t - u(k, l; t)]\,. \qquad (6)$$

Equation (6) can be decomposed to a part for the tectonic loading and a residual part for slip on the computational grid. The tectonic loading follows the formula

$$\tau_{\mathrm{load}}(i, j; t) = \gamma(i, j) \cdot t \qquad (7)$$

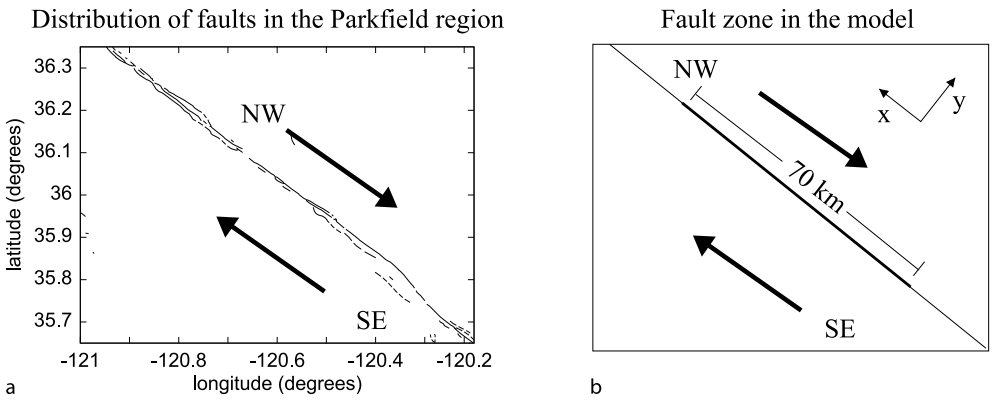
with a space-dependent but time-independent loading rate

$$\gamma(i, j) = v_{\mathrm{pl}} \cdot \sum_{(k,l)\in \text{ grid}} G(i, j; k, l)\,. \qquad (8)$$

The build-up of stress may be reduced by aseismic creep motion, which is implemented by a local constitutive law corresponding to lab-based dislocation creep [5]:

$$\dot{u}_{\mathrm{creep}}(i, j; t) = c(i, j) \cdot \tau^3(i, j; t) \qquad (9)$$

with space dependent but time-independent creep coefficients $c(i, j)$.



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 2
a Distribution of faults in the Parkfield (California) region; b fault region in the model

Friction and Coseismic Stress Transfer; Quasidynamic Approach

It is widely accepted that earthquakes on large faults are due to frictional processes on pre-existing structures. The friction is therefore an important empirical ingredient of a fault model [56]. Numerous laboratory experiments have been carried out to characterize frictional behavior of different materials (see e. g. [19]). An important finding is that the friction coefficient defined as the ratio of shear stress τ_{shear} and compressional normal stress τ_{normal}, $\mu_{\text{f}} = \tau_{\text{shear}}/\tau_{\text{normal}}$ at the initiation of slip is approximately constant for many materials; the value of μ_{f} lies between 0.6 and 0.85. This observation, known as *Byerlee's law*, is related to the Coulomb failure criterion [16] for the Coulomb stress *CS*,

$$CS = \tau_{\text{shear}} - \mu_{\text{f}}\tau_{\text{normal}} \,. \tag{10}$$

The Coulomb stress depends on a plane where shear stress and normal stress are calculated. Neglecting cohesion, the Coulomb criterion for brittle failure is

$$CS \geq \tau_0 \,, \tag{11}$$

which for $CS = 0$ is Byerlee's law.

The North-American plate and the Pacific plate move in opposite directions along the fault plane having strike-slip motion. The absence of normal and thrust faulting reduces the problem to a one-dimensional motion: all parts of the fault move along the fault direction. The stress state of the fault is fully determined by the shear stress τ_{xy} in the coordinates given in Fig. 2b. Slip is initiated if τ_{xy} exceeds $\mu_{\text{f}}\tau_{yy}$. This quantity, which is called the static strength τ_{s} is constant in time if μ_{f} is assumed to be constant. Note that the normal stress on a planar fault in a homogeneous solid does not change [1]. The shear stress τ_{xy} will be denoted simply by τ. In this notation, the failure criterion Eq. (11) reduces to

$$\tau \geq \tau_{\text{s}} \,. \tag{12}$$

When a cell (k, l) fails, the stress drops in this cell to the arrest stress τ_{a}:

$$\tau(k, l) \to \tau_{\text{a}} \,, \tag{13}$$

where the value τ_{a} maybe space-dependent. The stress change produces a corresponding slip

$$\mathrm{d}u(k, l) = \frac{\tau(k, l) - \tau_{\text{a}}}{G(k, l\,;\, k, l)} \tag{14}$$

with the self-stiffness $G(k, l\,;\, k, l)$ of cell (k, l).

The observational effect of dynamic weakening includes also a strength drop from the static strength to a lower dynamic strength:

$$\tau_{\text{s}} \to \tau_{\text{d}} \,. \tag{15}$$

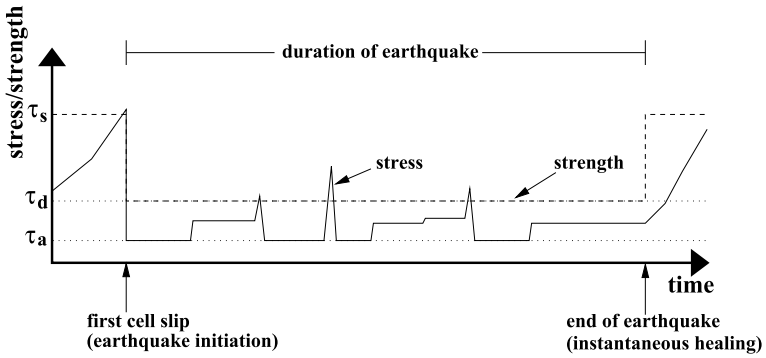
In particular, slipping material becomes weaker during rupture and recovers to the static level at the end of the rupture. This approximation of the strength evolution is known as static-kinetic friction.

The values τ_{s}, τ_{d}, and τ_{a} are connected by the dynamic overshoot coefficient *D*:

$$D = \frac{\tau_{\text{s}} - \tau_{\text{a}}}{\tau_{\text{s}} - \tau_{\text{d}}} \,, \tag{16}$$

or alternatively by the dynamic weakening coefficient ε:

$$\varepsilon = 1 - \frac{\tau_{\text{d}}}{\tau_{\text{s}}} \,. \tag{17}$$



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 3
Pictorial evolution of stress (*solid line*) and strength (*dashed line*) of a hypocenter cell in the quasidynamic approach

Following [10] we use in most simulations $D = 1.25$.

The redistribution of the stress release $\Delta\tau(k,l) = \tau(k,l) - \tau_a$ from cell (k,l) to a point (i,j) at time t is

$$\mathrm{d}\tau(i,j;\,t) = G(i,j;\,k,l)\cdot\delta\left(t - \frac{r(i,j;\,k,l)}{v_s}\right) \cdot \frac{\Delta\tau(k,l)}{G(k,l\,;\,k,l)}\,, \quad (18)$$

where $\delta(x)$ denotes the δ-function, which is 1 for $x = 0$ and 0 else; v_s is the shear-wave velocity, and $r(i,j;\,k,l)$ is the distance between source cell (k,l) and receiver position (i,j). That is, regions far from the slipping cell receive their stress portion later than regions close to the slipping cell. The value of v_s is assumed to be constant. Each "stress transfer event" associated with Eq. (18) gives a transfer of a stress $\mathrm{d}\tau$ from a source cell (k,l) to a receiver cell (i,j) at time t. This time-dependent stress transfer is called the *quasidynamic* approach in contrast to the *quasistatic* approach used in most similar models.

The evolution of stress and strength in a failing cell is shown schematically in Fig. 3. When the slip is initiated, both the stress and the strength drop. Due to coseismic stress transfer during the event, the cell may slip several times before the earthquake is terminated and instantaneous healing takes place in all cells. The piecewise constant failure envelope (dashed line) indicates static-kinetic friction. A model version with gradual healing was employed by [79]. A review of analytical results associated with the basic model in the context of a large universality class is given in ► Jerky Motion in Slowly Driven Magnetic and Earthquake Fault Systems, Physics of.

We note that the Green's function leads to an infinite interaction range. Using open boundary conditions with respect to the computational grid, the stress release from a slipping cell is not conserved on the grid, but on the infinite half plane.

Data

The model produces two types of data, earthquake catalogs and histories of stress and displacement fields. As demonstrated below, all parameters of the model have physical dimensions and can therefore be compared directly with real data, where they are available. This is in contrast to most of the slider-block and cellular automaton models.

Earthquake catalogs include values of the earthquake time, hypocenter, and size. The time of an earthquake is the time of the first slip; the hypocenter is determined by the position of the corresponding cell along strike and depth. The size of an event can be described by different measures: The rupture area A is the total area, which slipped during an earthquake. The potency

$$P = \int u(A)\mathrm{d}A \quad (19)$$

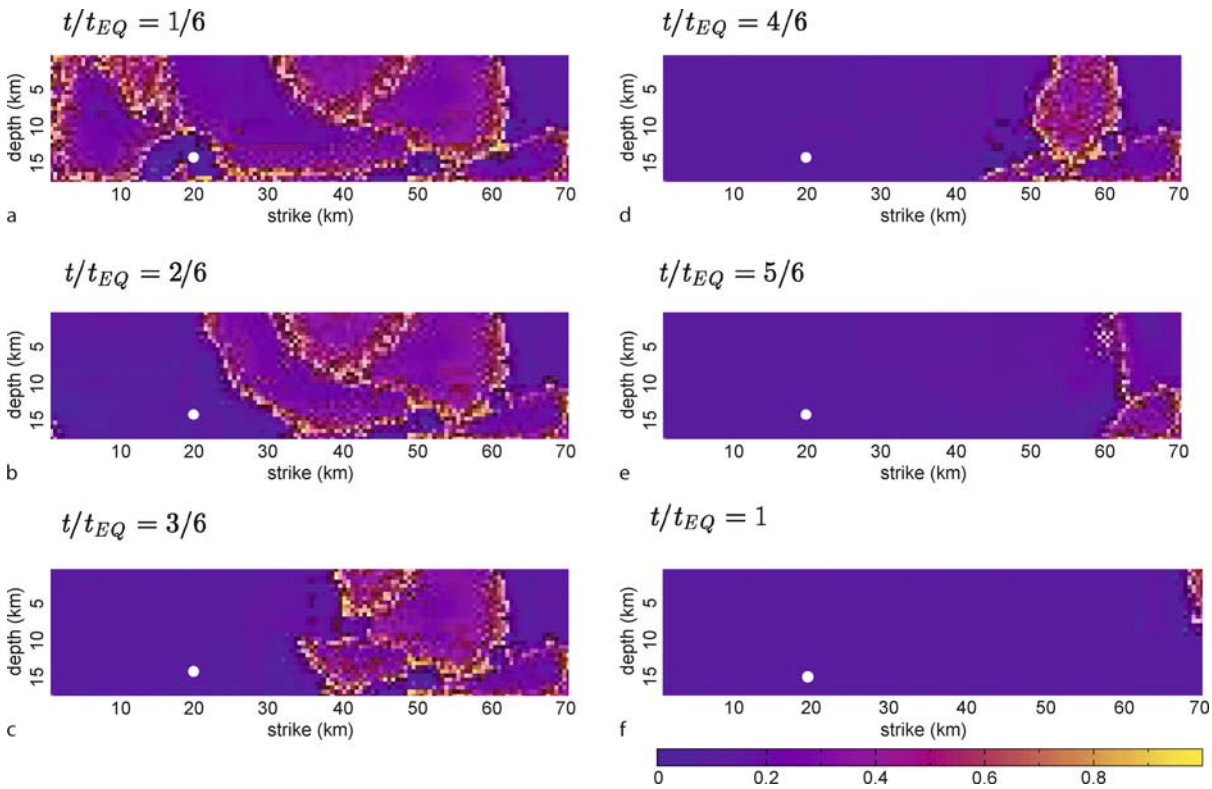
measures [7] the total slip during the event and is related to the seismic moment m_0 by the rigidity: $m_0 = \mu P$. The (moment) magnitude M can be calculated from the potency [10] using

$$M = (2/3)\log_{10}(P) + 3.6\,, \quad (20)$$

where P is given in cm $\cdot$ km^2.

Results

Numerous simulations of the model described in the previous section have been performed. Firstly, simulations have been examined with respect to the spatiotemporal propagation of stress during single earthquakes ("rupture histories"). Then, long deformation histories have been



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 4
Snapshots of rupture evolution for a system-wide event on a smooth fault without creep motion. t denotes the time after the rupture initiation (given in units of the total earthquake duration t_{EQ}). The white circle is the hypocenter of the event. The figure shows the dimensionless stress state $\hat{\tau} = \frac{\tau - \tau_a}{\tau_s - \tau_a}$ of the cells. a $t/t_{EQ} = 1/6$; b $t/t_{EQ} = 2/6$; c $t/t_{EQ} = 3/6$; d $t/t_{EQ} = 4/6$; e $t/t_{EQ} = 5/6$; f $t/t_{EQ} = 1$

simulated in order to search in a large fraction of the parameter space for relationships between input parameters and observed seismicity features. In this section, a selection of key results is presented and discussed in relation to critical states of seismicity.

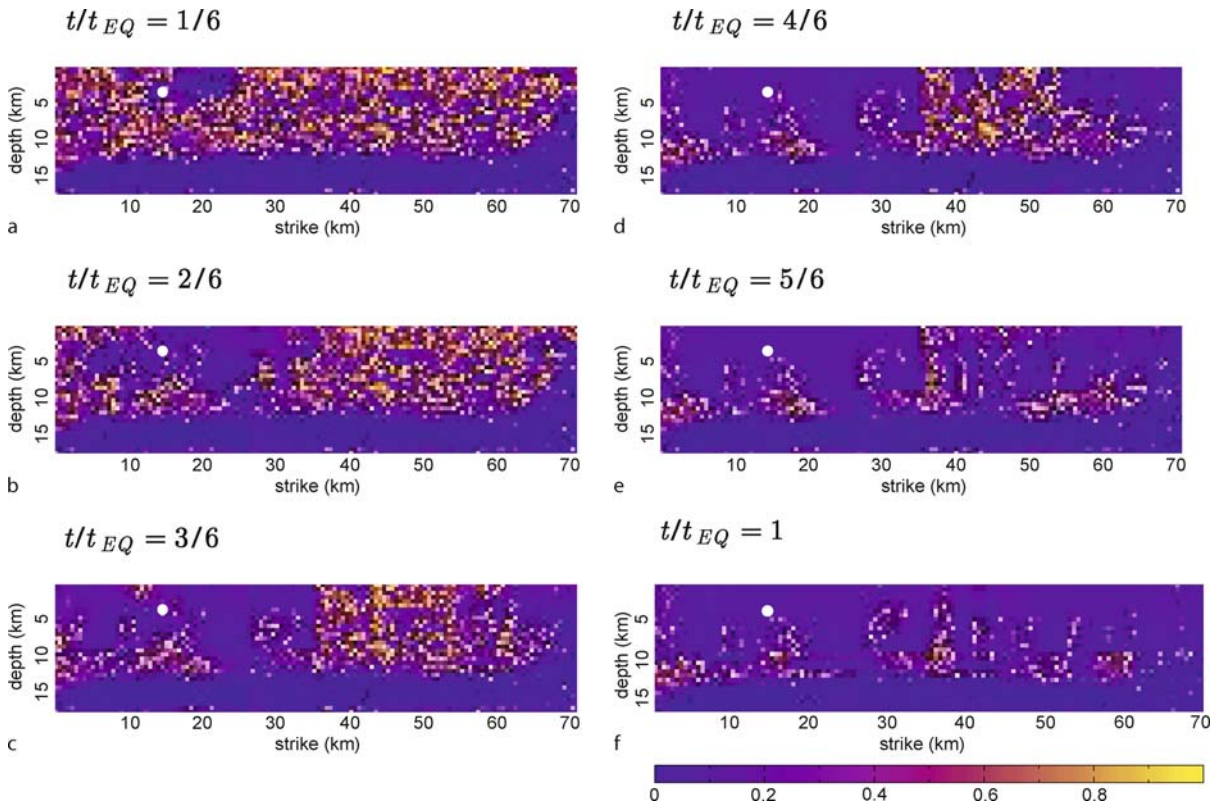
Rupture Histories

We compare qualitatively rupture histories of large earthquakes for three end-member cases in parameter space:

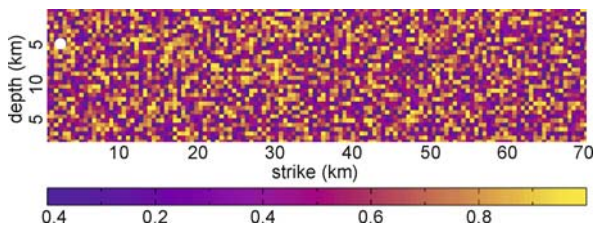
(1) a smooth fault,
(2) a rough fault, and
(3) a fault without dynamic weakening ($\tau_d = \tau_s$ or $D \to \infty$ in Eq. (16)).

Following [5], we vary the degree of quenched spatial disorder for a particular realization by introducing barriers of high stress drop $\tau_s - \tau_a$ in an environment of low stress drop.

The observation that smooth faults show a more regular earthquake occurrence than rough faults, can be explained by the ability of the stress field to synchronize on certain fault patches. On a disordered fault, this type of synchronization is unlikely. Figure 4a shows the stress field (normalized between 0 and 1) immediately before a large earthquake on a smooth fault. The most striking feature is the emergence of clearly defined patches with highly loaded boundaries. During rupture evolution, these patches rupture almost in series until the fault is nearly unloaded (see Fig. 4b–f). A different situation is shown in Fig. 5, corresponding to a rough fault with creep coefficients $c(i, j)$ (Eq. (9)) that increase with depth leading to a brittle-ductile transition zone as in [5] and [80]. Here the stress field in the brittle regime is irregular without obvious pattern formation. Similar behavior is found in a case where dynamic weakening is switched off ($D \to \infty$); in other words, the material heals instantaneously. Figure 6 shows the stress field in this case. In analytical studies, it has been shown that this corresponds exactly to a critical point in a phase diagram spanned by the stress dissipation and dynamic weakening [22,27]. Observational results indicate [68] that irregular slip histories and power



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 5
Same as Fig. 4 for a strongly disordered fault with a brittle-ductile transition at about 15 km depth. a $t/t_{EQ} = 1/6$; b $t/t_{EQ} = 2/6$; c $t/t_{EQ} = 3/6$; d $t/t_{EQ} = 4/6$; e $t/t_{EQ} = 5/6$; f $t/t_{EQ} = 1$



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 6
Same as Fig. 4a for a fault without dynamic weakening ($\tau_d = \tau_s$) corresponding to a dynamic overshoot coefficient $D \to \infty$ (Eq. (16))

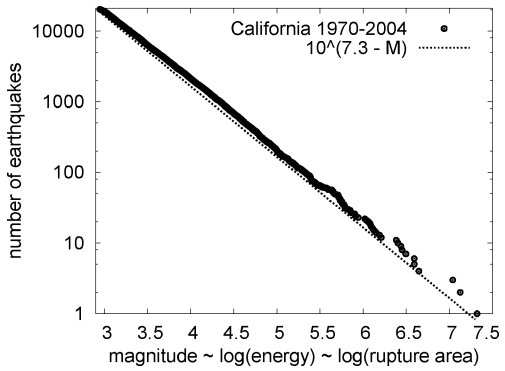
law frequency-size distributions are associated with geometrically disordered fault structures, while characteristic earthquake statistics and overall regular ruptures are found on mature fault with large total displacements.

Although the stress field shows a complex evolution during a simulation, the presence or absence of characteristic length scales indicating the relation to a critical point is easily detected. From an observational point of view, the stress field is not accessible. The evolution of the displacement field may be estimated, e. g. from seismic and geodetic data using slip inversion techniques. Because of the high uncertainties in the calculated slip histories, a quantitative comparison of the simulated data with "natural" data is questionable. However, general features of the quasidynamic ruptures are quite realistic, e. g. the irregular patterns in Fig. 5 resemble the rupture of the Chi–Chi (Taiwan) earthquake on September 21, 1999 ($M_w = 7.6$) [58].

Later we will show that the frequency-size distribution of earthquakes can serve to some extent as a proxy for the degree of disorder of the stress field. Ben-Zion et al. [12] discuss additional seismicity functions that may be used as surrogate variables for the stress.

Frequency–Size Distributions

The frequency-size (FS) distribution is one of the most important characteristics of observed seismicity. For world-



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 7

Frequency-size distribution for California from 1970 to 2004; the dashed line denotes a power law fit to the data

wide seismicity, this distribution is given by the Gutenberg–Richter law (Eq. (1)). Figure 7 shows the FS distribution of California seismicity from 1970 to 2004. Here we use the non-cumulative version of Eq. (1), where N is the number of earthquakes with magnitude between M and $M + \mathrm{d}M$ with a magnitude bin $\mathrm{d}M = 0.1$.

For individual fault zones the FS distribution can deviate from Eq. (1), especially for high magnitudes. Examples are given in Fig. 8, which shows the FS distribution of the Parkfield segment (Fig. 8a) and for the San Jacinto fault (Fig. 8b) in California for a time span of 45 years. The distribution of the Parkfield segment consists of two parts: A scaling regime for $2.2 \leq M \leq 4.5$ and a "bump" for $4.5 < M \leq 6.0$. For the San Jacinto fault, the scaling range is observed for almost all events ($2.2 \leq M \leq 5.0$). The decrease for $M \approx 2$ in both plots is probably due to catalog incompleteness.

The FS distribution as shown in Fig. 8a is called the characteristic earthquake distribution, because of the increased probability for the occurrence of large ("characteristic") events compared to the prediction of the Gutenberg–Richter relation. The latter is an exponential distribution for the earthquake frequency as a function of magnitude, or a power law distribution for the earthquake frequency as a function of potency (Eqs. (19), (20)), moment, energy, or rupture area, over a broad range of magnitudes [66]. The Gutenberg–Richter relation is "scale-free" because a power law distribution indicates the absence of a characteristic scale of the earthquake size [64]. In terms of critical point processes, the absence of a characteristic length scale indicates that the system is close to the critical point. In this state, earthquakes of all magnitudes can occur, or each small rupture can grow into a large one. Therefore, the frequency-size distribution can serve as a proxy for the current state of a system in relation to a critical point.

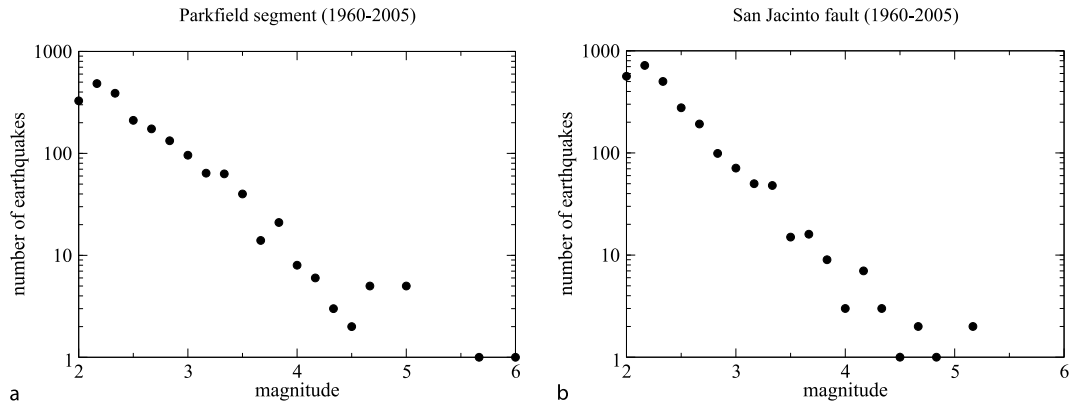
The FS distribution in a model can be tuned by varying the mean stress $\langle\tau\rangle$ on the fault, where $\langle\rangle$ denotes the spatial average. This can be achieved, for instance, by varying brittle properties, e. g. in terms of the dynamic overshoot coefficient D (Eq. (16)), or by introducing dissipation [31,79]. Figure 9 shows FS distributions for two different values of D : $D = 5/4$ (Fig. 9a) and a higher value $D = 5/3$ (Fig. 9b). While Fig. 9a follows a characteristic earthquake behavior similar to the Parkfield case (Fig. 8a), Fig. 9b resembles the shape of the FS distribution of the San Jacinto fault (Fig. 8b).

As an outcome, three cases can be distinguished by means of a critical mean stress τ_{crit}:

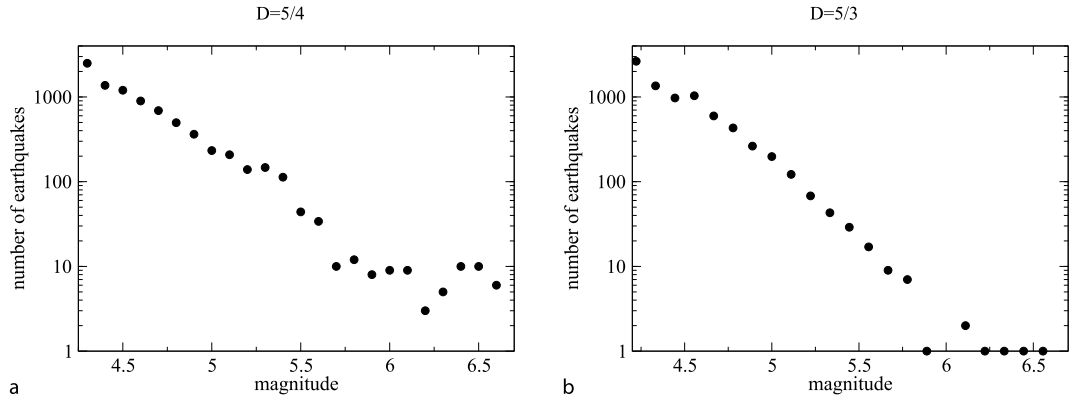
1. subcritical fault ($\langle\tau\rangle < \tau_{\mathrm{crit}}$): the mean stress on the fault is too small to produce large events. The system is always far from the critical point. The FS distribution is a truncated Gutenberg–Richter law.
2. supercritical fault ($\langle\tau\rangle > \tau_{\mathrm{crit}}$): the mean stress is high and produces frequently large events. After a large earthquake (critical point), the stress level is low (system is far from the critical point) and recovers slowly (approaches the critical point). The FS distribution is a characteristic earthquake distribution.
3. critical fault ($\langle\tau\rangle \approx \tau_{\mathrm{crit}}$): the system is always close to the critical point with scale-free characteristics. The FS distribution is a Gutenberg–Richter law with a scaling range over all magnitudes.

If the FS distribution is plotted as a function of the model parameters, the result can be visualized by a phase diagram [22,31,79,80]. An example is provided in Fig. 10, which shows schematically the phase diagram spanned by the degree of quenched spatial disorder and $1 - \varepsilon$ with the dynamic weakening coefficient ε (Eq. (17)). The phase diagram summarizes results from various studies, which demonstrate that the degree of spatial disorder of the stress drop acts as a tuning parameter for the FS distribution [5,36,80].

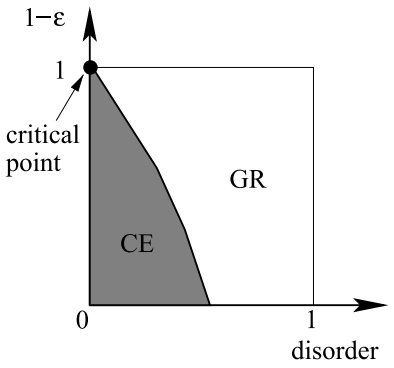
If the model is in the transition regime between Gutenberg–Richter statistics and characteristic earthquake behavior, the ability of the stress field to synchronize on parts of the fault can have additional impact on the dynamics of seismicity: for a model with small cells and high stress fluctuations along the cell boundaries arising from a high degree of spatial disorder, the system can undergo a spontaneous transition from an ordered state and characteristic behavior to a disordered state following Gutenberg–Richter statistics. Due to the high fluctuations in the



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 8
Frequency-size distribution for two faults in California: a the Parkfield segment, and b the San Jacinto fault

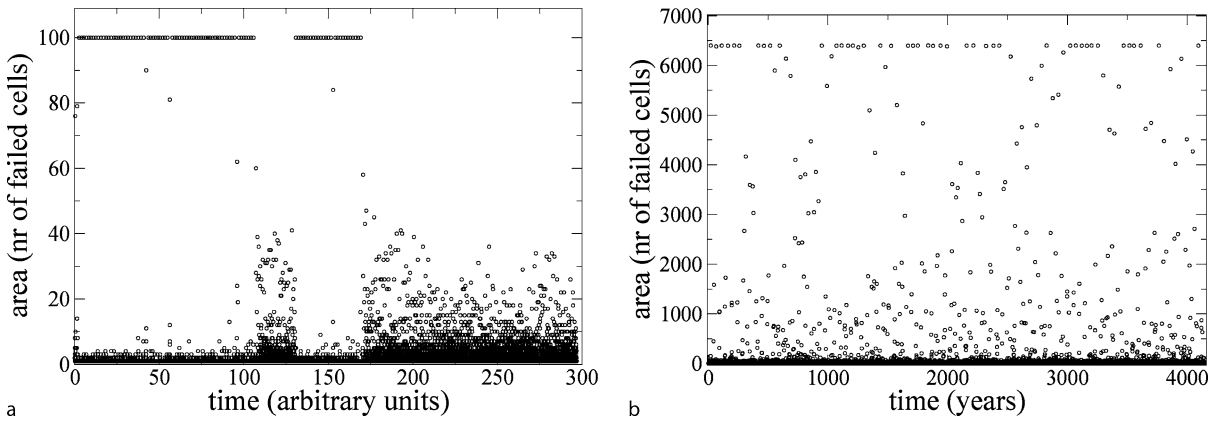


Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 9
Frequency-size distribution for model realizations with different dynamic overshoot coefficients (Eq. (16)): a $D = 5/4$, b $D = 5/3$

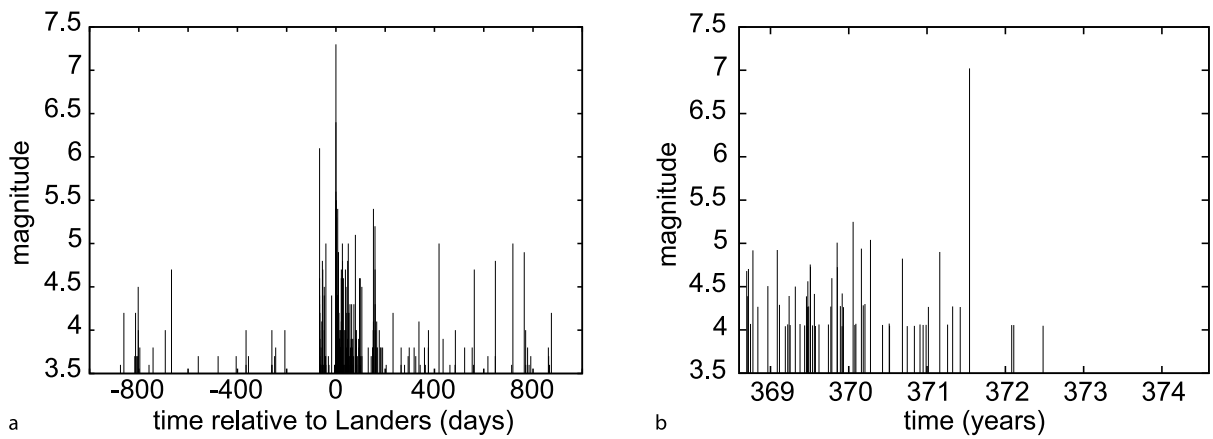


Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 10
Phase diagram for the frequency-size distribution (GR = Gutenberg–Richter distribution, CE=characteristic earthquake distribution) spanned by the degree of quenched spatial disorder and the dynamic weakening represented by ε. The upper left corner corresponds exactly to a critical point [22,27] and results in scale-free characteristics as shown in Fig. 6

stress field, there is some probability that a certain number of cells synchronize by chance, leading to an ordered behavior for some seismic cycles, until the order is destroyed, again resulting from stress fluctuations. This type of mode-switching behavior has been observed earlier in a mean-field model and a damage rheology model [11,22]. Figure 11a gives a corresponding earthquake sequence with spontaneous mode-switching behavior. Figure 11b shows a sequence calculated with a higher grid resolution (128×50 cells). The tendency to mode-switching is less pronounced, but still visible. In [79] it is shown that the emergence of such mode-switching depends both on the spatial range of interaction (given as the decay of the Green's function) and the discretization of the computational grid. In the less realistic model of [22], where the stress redistribution is governed by a constant (space-independent) Green's function, analytical expressions for persistence times have been calculated [27]. In [11] some evidence for mode-switching behavior in long seismic



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 11
Earthquake area (measured as the number of failed cells) as a function of time a in the mean-field model of Dahmen et al. [22] for a fault with 100 cells and b in the elastic model with 128 × 50 cells



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 12
Earthquakes before and after a mainshock: a the *M*7.3 Landers (California) earthquake; b *M*7.3 earthquake in the basic version of the model

records based on paleoseismic and geologic data from the Dead Sea fault and other regions are discussed. However, the relevance of mode-switching to natural seismicity remains unclear due to the general lack of very long data records.

Aftershocks and Foreshocks

The most pronounced temporal pattern in observed seismicity is the emergence of strongly clustered aftershock activity following a large earthquake. Apart from the Omori–Utsu law (Eq. (2)), it is widely accepted that aftershocks are characterized by the following properties:

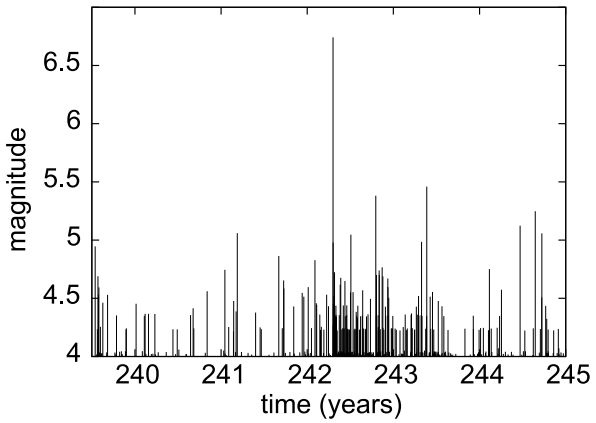
1. The aftershock rate scales with the mainshock size [51].
2. Aftershocks occur predominantly around the edges of the ruptured fault segments [66].
3. Båth's law [4]: The magnitude of the largest aftershock is usually about one unit smaller than the mainshock magnitude.

Deviations from the Omori–Utsu law, especially for rough faults, are discussed in [45]. While aftershocks are observed after almost all large earthquakes, foreshocks occur less frequent [71]. As a consequence, much less is known about the properties of these events. Kagan and Knopoff [40] and Jones and Molnar [39] propose a power law increase of activity according to an "inverse" Omori–Utsu law.

Figure 12a shows an example for the aftershock sequence following the *M*7.3 Landers earthquake in California on June 28, 1992. An earthquake of similar size generated by the model is given in Fig. 12b. The absence of af-

tershocks in the simulation is clearly visible. The reason for the lack of aftershocks is the unloading of the fault resulting from the mainshock: When a large fraction of the fault has ruptured, the stress in this region will be close to the arrest stress after the event. Consequently, the seismicity rate will be almost zero until the stress field has recovered to a moderate level.

Discussions for likely mechanisms of aftershocks are given in [9] and [81]. A common feature is the presence of rapid postseismic stress which generates aftershock activity. In [32], for instance, postseismic stress has been attributed to a viscoelastic relaxation process following the mainshock. In the work discussed here, continuous creep displacement following Eq. (9) is assumed. Additionally, the computational grid is divided by aseismic barriers from the free surface to depth into several seismically active fault segments. As shown in [81], this modification results in a concentration of stress in the aseismic regions during rupture and, subsequently in a release of stress after the event according to the coupled creep process. This stress release triggers aftershock sequences obeying the Omori–Utsu law (Eq. (2)). A typical aftershock sequence after a $M6.8$ event is shown in Fig. 13. In agreement with Båth's law, the strongest aftershock has the magnitude $M = 5.5$ in this sequence. The sequence shows also the effect of secondary aftershocks, namely aftershocks of aftershocks [61]. The stacked earthquake rate as a function of the time after the mainshock is given in Fig. 14. In this case, where the barriers are characterized by creep coefficients higher by a factor of 10^5 than in the other patches, a realistic Omori exponent of $p = 1$ is found.



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 13
Earthquakes before and after a mainshock with $M = 6.8$ in the modified model

Aftershock sequences like in Fig. 13 emerge after all large events in the extended model. In contrast, there is generally no clear foreshock signal visible in single sequences. However, stacking many sequences together unveils a slight increase of the earthquakes rate prior to a mainshock supporting the observation of accelerating foreshock activity. An explanation of these events can be given in the following way: Between two mainshocks the stress field organizes itself towards a critical state, where the next large earthquake can occur. This critical state is characterized by a disordered stress field and the absence of a typical length scale, where earthquakes of all sizes can occur [12]. The mainshock may occur immediately or after some small to moderate events. The latter case can be considered as a single earthquake, which is interrupted in the beginning. This phenomenon of delayed rupture propagation has also provided a successful explanation of foreshocks and aftershocks in a cellular automaton model [33,35].

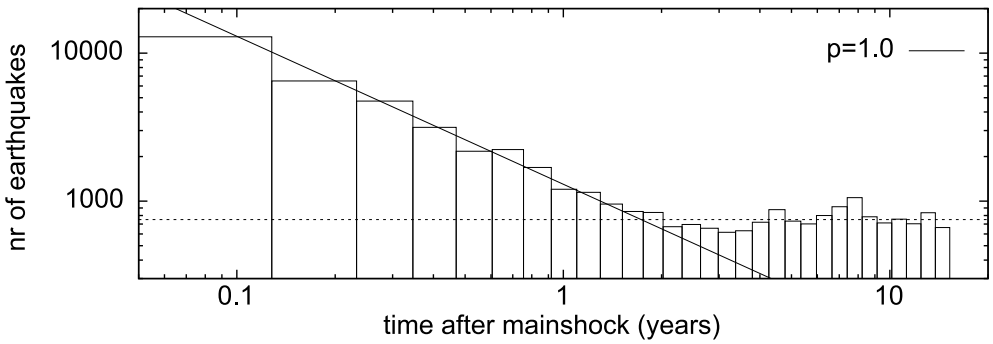
The hypothesis that foreshocks occur in the critical point and belong, in principle, to the mainshock, can be verified by means of the findings from Subsect. "Frequency–Size Distributions". In particular, the frequency-size distribution in the critical point (or close to the critical point) is expected to show scale-free statistics. If an overall smooth fault model following characteristic earthquake statistics is studied over a long time period, the approach of the critical point should be seen in terms of a change of the frequency-size distribution towards Gutenberg–Richter behavior [12]. This change of frequency-size statistics is indeed observed in the model (Fig. 15) and supports the validity of the critical point concept [82].

Accelerating Moment Release

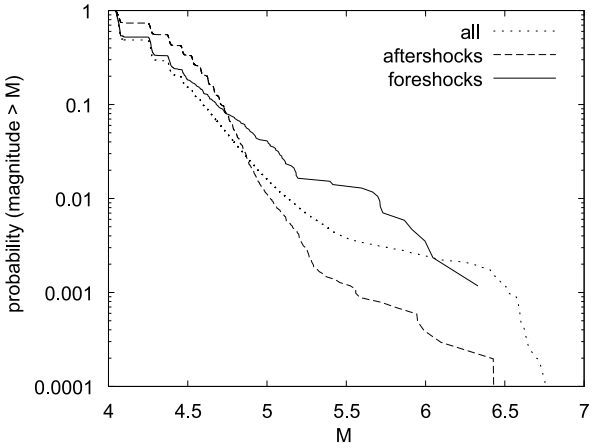
In the previous section, it has been argued that large earthquakes are associated with a critical point and the preparation process is characterized by increasing disorder of the stress field and increasing tendency for scale-free characteristics in the frequency-size distribution. Further support for critical point dynamics has been provided by the observational finding of [17] that the cumulative Benioff strain $\Sigma\Omega(t)$ follows a power law time-to-failure relation prior to the $M7.1$ Loma Prieta earthquake on October 17, 1989:

$$\Sigma\Omega(t) = \sum_{i=1}^{N(t)} \sqrt{E_i} = A - B(t_f - t)^m \tag{21}$$

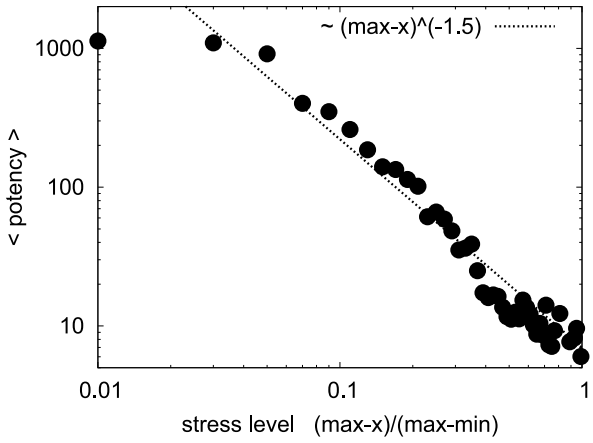
Here, E_i is the energy release of earthquake i, $N(t)$ is the number of earthquakes before time t; t_f is the failure time



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 14
Earthquake rate as a function of time for the model with seismic and aseismic regions. The calculation is based on a simulation with 200,000 earthquakes covering about 5000 years; the earthquake rates are averaged over about 300 mainshocks. A fit of the Omori–Utsu law (Eq. (2)) with $p = 1$ is denoted as a solid line. The dashed line gives the estimated background level of seismicity



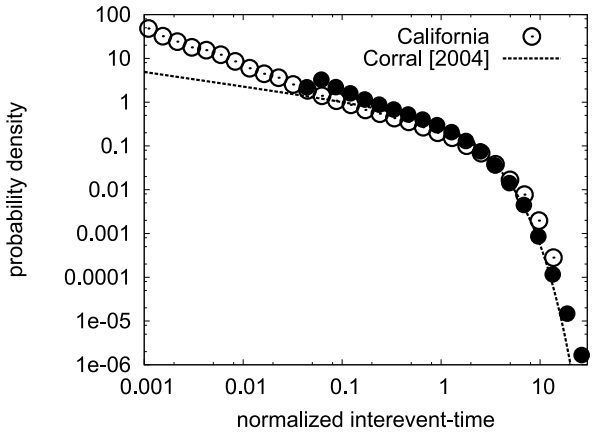
Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 15
Frequency-magnitude distributions of all earthquakes, foreshocks and aftershocks. Foreshocks and aftershocks are defined as earthquakes occurring within one month before and after an earthquake with $M \geq 6$



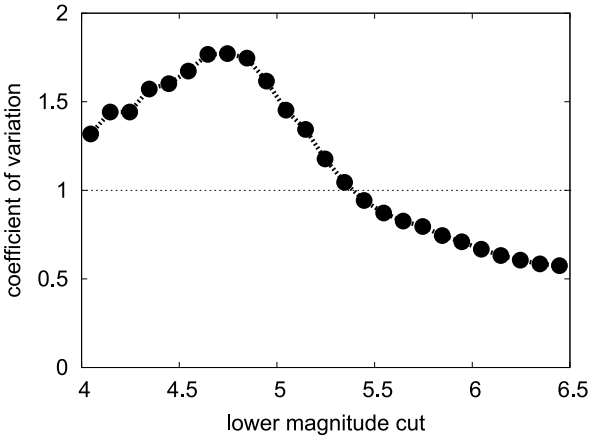
Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 16
Mean potency release (Eq. (19)) as a function of the stress level. The stress level is normalized to the maximum (max) and minimum (min) observed stress

and A, B, and $m > 0$ are constants. Similar studies for numerous seismically active regions followed (see [8,77] and references therein).

The time-to-failure relation Eq. (21) has been explained by [54] and [60] from the viewpoint of renormalization theory and by [8] and [65] in terms of damage rheology. Similar to the findings about foreshocks, the time-to-failure pattern is not universal. Therefore, a stacking procedure is adopted in order to obtain robust results on the validity of Eq. (21) in the model. This is not straightforward, since the interval of accelerating moment release is not known a priori and the duration of a whole seismic cycle, as an upper limit, is not constant. To normalize the time interval for the stacking, the potency release (Eq. (19)) is computed as a function of the (normalized) stress level (Fig. 16). Taking into account that the stress level increases almost linearly during a large fraction of the seismic cycle, the stress level axis in Fig. 16 can effectively be replaced by the time axis leading to a power law dependence of the potency release on time. The best fit is provided with an exponent $s = -1.5$. Transforming the potency release to the cumulative Benioff strain (Eq. (21)), results in an exponent $m = 0.25$ in Eq. (21). This finding is based on a simulation over 5000 years; the exponent is in good agreement with the theoretical work [53], that derives $m = 0.25$ for a spinodal model, and the analytical result of $m = 0.3$ in the damage mechanics model [8]. An observational study of California seismicity finds m between 0.1 and 0.55 [15].



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 17
The normalized interevent time distribution of the model simulations (black dots) compared with the result of [21] and the distribution of earthquakes in California (ANSS catalog of $M \geq 3$ earthquakes occurred between 1970 and 2004 within 29° and 43° latitude and −113° and −123° longitude)



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 18
The temporal earthquake occurrence quantified by the coefficient of variation as a function of the lower magnitude cutoff. Values larger than 1 indicate clustering, whereas lower values point to quasiperiodic behavior

Interevent Times

In recent studies it has been shown that the distribution of interevent times can be described by a universal law. In particular, the distributions from different tectonic environments, different spatial scales (from worldwide to local seismicity) and different magnitude ranges collapse if the time Δt is rescaled with the rate R_{xy} of seismic occurrence in a region denoted by (x, y) [21]. Such rescaling leads to

$$D_{xy}(\Delta t) = R_{xy} \cdot f(R_{xy}\Delta t)\,, \tag{22}$$

where D_{xy} is the probability density for the interevent time Δt, and f can be expressed by a generalized gamma distribution

$$f(\theta) = C\frac{1}{\theta^{\gamma-1}} \exp\left(-\frac{\theta^{\delta}}{B}\right). \tag{23}$$

The parameters C, γ, δ, and B have been determined by a fit to several observational catalogs [21].

In Fig. 17 we compare $D_{xy}(\Delta t)$ from Eq. (22) with two earthquake catalogs: (1) The ANSS catalog of California (catalog ranges are given in the caption), and (2) the model catalog. Due to the universality of Eq. (22) with respect to different spatial scales, the comparison of the model simulating a single fault of 70 km length with a region of hundreds of kilometers including several faults in California does not require coarse graining the ANSS catalog. In the region where the interevent times are calculated, we find a remarkable agreement of the three curves. For small values of Δt, Eq. (22) deviates from the California data; for high values the model has a slightly better correspondence with the observational data than Eq. (22). Thus the results generally support the recent findings of [21].

The degree of temporal clustering of earthquakes can be estimated by the coefficient of variation CV of the interevent time distribution.

$$CV = \sigma/\mu\,, \tag{24}$$

where σ is the standard deviation and μ the mean value of the interevent time distribution. Values of $CV > 1$ denote clustered activity, while $CV < 1$ represents quasiperiodic occurrence of events. The case $CV = 1$ corresponds to a random Poisson process. The studies of [5] and [80] have found that the clustering properties of the large events depend on the degree of quenched spatial disorder of the fault. Figure 18 shows that CV as a function of the lower magnitude cutoff has a characteristic shape. The values of CV are higher than 1 (clustered) for small and intermediate earthquakes ($M \leq 5.4$) and smaller than 1 (quasiperiodic) for larger earthquakes. This corresponds to the case of a low degree of disorder in [80], because the brittle cells which participate in an earthquake have no significant spatial disorder. We note that this behavior resembles the seismicity on the Parkfield segment of the San Andreas fault, which is characterized by a quasiperiodic occurrence of mainshocks. Based on the analysis of 37 earthquake sequences, an estimation of $CV \approx 0.5$ has been found for multiple tectonic environments [26].

A different behavior is observed on the San Jacinto fault in California, where the largest events occur less reg-

ularly and have overall smaller magnitudes. As discussed in [5] and [80], this can be modeled by imposing higher degrees of disorder leading to a broader range of spatial size scales, e. g. by using a higher number of near-vertical barriers. While barriers provide a simple and physically motivated way to tune the degree of disorder, other types of heterogeneities may work as well, as long as they are able to produce strong enough fluctuations of the stress field. As an example, we mention fractal distributions of the stress drop, which can be tuned easily by changing the fractal dimension [63,80].

Recurrence Times of Large Earthquakes

While interevent times are waiting times between successive earthquakes in a given catalog, recurrence times are defined as waiting times between two successive *large* events, typically in the magnitude range $6 \leq M \leq 9$, depending on the region. For example, on the Parkfield segment of the San Andreas fault seven $\sim M6$ earthquakes occurred between 1857 and 2004 with recurrence times $T_1 = 24$ years, $T_2 = 20$ years, $T_3 = 21$ years, $T_4 = 12$ years, $T_5 = 32$ years, and $T_6 = 38$ years.

The distribution of recurrence times of large earthquakes is crucial for the calculation of seismic hazard. Due to a lack of observational data, this distribution is unknown for real fault systems. Commonly used distributions are based on extreme value statistics and on models for catastrophic failure. These include the lognormal distribution [50], the Brownian passage time distribution [44], and the Gumbel distribution [29]. All distributions are characterized by a maximum for a certain recurrence time followed by an asymptotic decay. The Brownian passage time distribution and the lognormal distribution have been used by the Working Group on California Earthquake Probabilities [70], e. g. for calculating earthquake probabilities in the San Francisco Bay area.

Figure 19a shows the probability density function (pdf) of the recurrence times of earthquakes with magnitude $M > 6.2$ in a realization of the numerical model for the Parkfield region [74]. Since we focus on long time-scales, we use here a minimal model without aseismic creep and strong spatial heterogeneities. This model leads to characteristic earthquake statistics and quasiperiodic occurrence of large events, and can therefore serve as a model framework for large earthquakes on the Parkfield segment. However, quantities which are only poorly known from empirical data, e. g. the stress drop, have to be chosen in order to perform a numerical simulation. Starting with an imposed uniform a priori distribution $P(\Delta\tau)$ of stress drops between a lower bound $\Delta\tau_{\min}$ and an upper bound $\Delta\tau_{\max}$, an a posteriori distribution $P(\Delta\tau|T_1, \ldots, T_N)$ can be estimated using observational recurrence times $T_1, \ldots, T_N$ from Parkfield and Bayes' theorem [13],

$$P(\Delta\tau|T_1, \ldots, T_N) = \frac{P(T_1, \ldots, T_N|\Delta\tau)P(\Delta\tau)}{\sum_{s=\Delta\tau_{\min}}^{\Delta\tau_{\max}} P(T_1, \ldots, T_N|s)P(s)}, \tag{25}$$

with the likelihood function

$$P(T_1, \ldots, T_N|\Delta\tau) = \prod_{i=1}^{N} f(T_i|\Delta\tau). \tag{26}$$

The function $f(T_i|\Delta\tau)$ denotes the pdf of recurrence times simulated with a model stress drop $\Delta\tau$. To get an analytic expression of this function, it is fitted by a Gamma distribution $f(t) = \beta^{-1}(\Gamma(\gamma))^{-1}(\frac{t-\mu}{\beta})^{\gamma-1}\exp(-\frac{t-\mu}{\beta})$ with the location parameter μ, the shape parameter $\gamma \equiv 2.0$ and the scale parameter β (with $x \geq \mu\,;\ \gamma, \beta > 0$). For an example see Fig. 19a. In [74] it is shown that the mean value μ_t and the standard deviation σ_t of the fits in this model are related to the average stress drop of a large earthquake $\Delta\tau$ by the simple empirical relations

$$\begin{aligned} \mu_t(\Delta\tau) &= 9.7 \cdot \Delta\tau \\ \sigma_t(\Delta\tau) &= 1.8 \cdot \Delta\tau^2 - 6.8 \cdot \Delta\tau + 11.7 \end{aligned} \tag{27}$$

with μ_t, σ_t in years and $\Delta\tau$ in MPa. Using this approximation in combination with six observational recurrence times from $\sim M6$ earthquakes on the Parkfield segment, we find the a posteriori distribution of stress drops shown in Fig. 19b. The position where this distribution reaches the maximum, $\Delta\tau = (3.04 \pm 0.27)$ MPa, is the most representative value of the stress drop of $\sim M6$ Parkfield events.

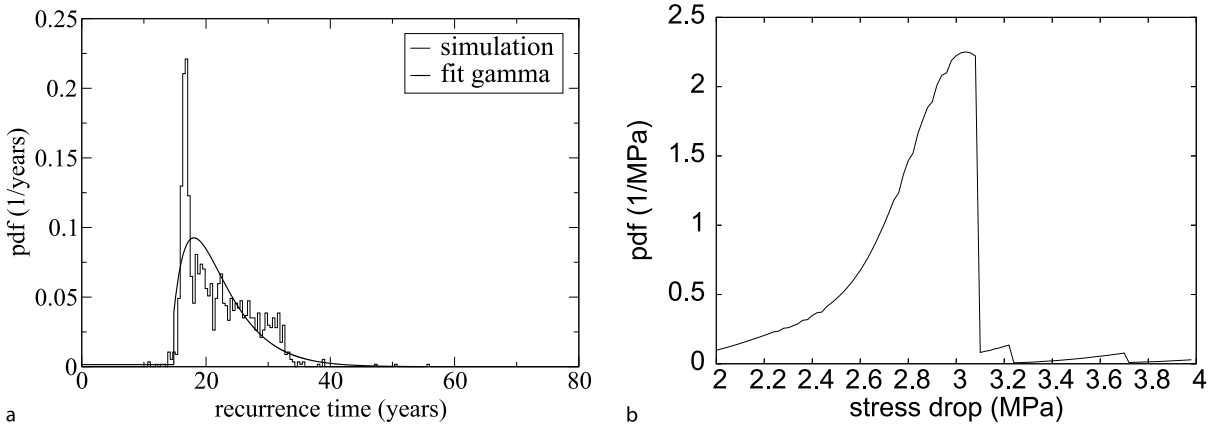
The cumulative probability density function (cdf) of recurrence times based on Eq. (26) and the observational data can now be calculated by

$$C(t) = \int_0^t \int_{\Delta\tau_{\min}}^{\Delta\tau_{\max}} f(t'|\Delta\tau)P(\Delta\tau|T_1, \ldots, T_N)\mathrm{d}\Delta\tau\mathrm{d}t'. \tag{28}$$

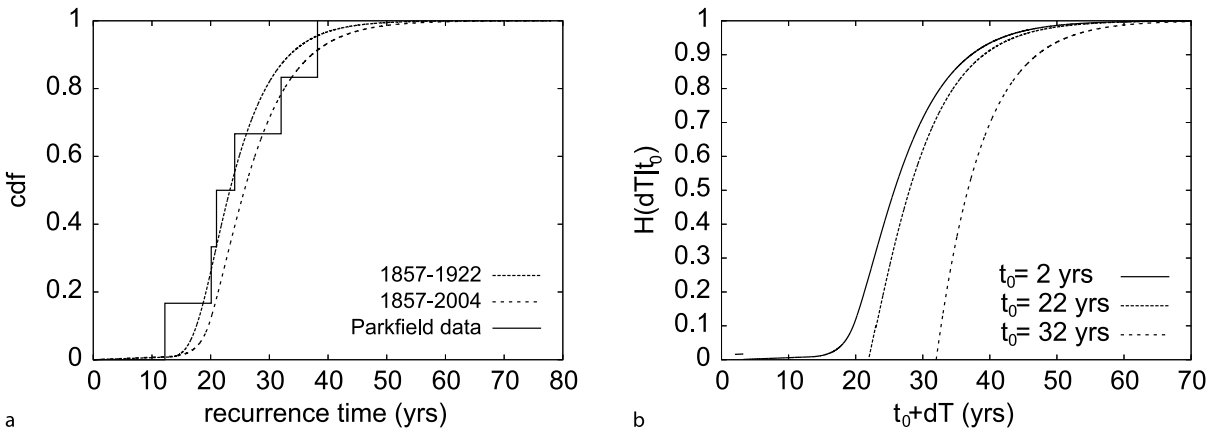
The hazard function

$$H(\Delta t|t_0) = \frac{C(t_0 + \Delta t) - C(t_0)}{1 - C(t_0)} \tag{29}$$

is the conditional probability that the next large earthquake occurs in the interval $[t_0\,;\ t_0 + \Delta t]$ given the time t_0 since the last large event. Results for two choices of observational data (corresponding to two different observa-



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 19
a Approximated probability density function of the recurrence time distribution of large earthquakes ($M > 6.2$) for a simulated earthquake catalog and fit with a truncated Gamma distribution; b A posteriori distribution $P(\Delta\tau|T_1, \ldots, T_N)$ of stress drop $\Delta\tau$ calculated with Bayes' theorem (Eq. (25))



Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space, Figure 20
a Cumulative recurrence time distribution $C(t)$ (Eq. (28)) for (1) the Bayesian approach with three data points: long-dashed line; (2) the Bayesian approach with six data points: short-dashed line. The solid line denotes the cdf of the six Parkfield recurrence times; b Hazard function $H(t_0|\Delta T)$ (Eq. (29)) based on the six observational recurrence times between 1857 and 2004 a as a function of dT for different values of t_0

tional periods) in comparison to the Parkfield cdf are given in Fig. 20a. The hazard function for three fixed values of t_0 and varying Δt is given in Fig 20b.

This approach enables us to calculate the most likely occurrence time of the next (post 2004) Parkfield earthquake by picking the maximum of the (non-cumulative) recurrence time distribution (inner integral in Eq. (28)) after taking all Parkfield earthquakes (1857–2004) into account. Based on the analysis done so far, we may forecast the next $\sim M6$ Parkfield earthquake to occur in May 2027. The error associated with one standard deviation of the pdf is 7.7 years. We note, however, that the probability for the occurrence of a $\sim M6$ earthquake between May 2026 and May 2028 is only about 14%.

Summary and Conclusions

The present review deals with the analysis, the understanding and the interpretation of seismicity patterns with a special focus on the critical point concept for large earthquakes. Both physical modeling and data analysis are discussed. This study aims at practical applications to model data from real fault zones. A point of particular interest is the detection of phenomena prior to large earthquakes and their relevance for a possible prediction of these events. Despite numerous reports on anomalous precursory seismicity changes [62], there is no precursor in sight which obeys a degree of universality that would make it practically useful. It is, therefore, important to study less fre-

quent precursory phenomena by means of long model simulations.

Toward this goal, we discuss a numerical model which is on one hand reasonably physical, and on the other hand simple enough that it allows to obtain some analytical results and perform long simulations. The basic version of the model consists of a segmented two-dimensional strike-slip fault in a three-dimensional elastic half space and is inherently discrete because of the abrupt transition from static to kinetic friction [10]. This paper and ► Jerky Motion in Slowly Driven Magnetic and Earthquake Fault Systems, Physics of summarize a large body of analytical and numerical results associated with the model.

The results of the simulations indicate an overall good agreement of the synthetic seismicity with natural earthquake activity, with respect to frequency-size distributions and various features of earthquake sequences. The degree of spatial heterogeneity on the fault, which is implemented by means of space-dependent rheological properties, has important effects on the resulting catalogs. Smooth faults are associated with the characteristic earthquake statistics, regular occurrence of mainshocks and overall smooth stress fields. On the other hand, rough faults generate scale-free Gutenberg–Richter statistics, irregular mainshock occurrence, and overall rough stress fields. A closer look at the disorder of the stress field shows, however, that even on a smooth fault a gradual roughening takes place when the next large earthquake is approached [12,82]. This is reflected in the frequency-size distribution which evolves towards the Gutenberg–Richter law and other changes of seismicity. The results can be used to establish relations between the proximity of a state on a fault to a critical point, the (unobservable) stress field, and the (observable) seismicity functions. Furthermore, it is demonstrated that the concept of "self-organized criticality" can be folded back to criticality associated with tuning parameters [12,31]. We note that phase diagrams with different dynamic regimes as functions of tuning parameters, in addition to criticality, provide a general and rich description of seismicity. Accelerating seismic release, growing spatial correlation length, changes of frequency-size statistics and evolution of other seismicity parameters may be used to track the approach to criticality [73,75,76,77].

Future Directions

We have demonstrated that numerical fault models are valuable for understanding the underlying mechanisms of observed seismicity patterns, as well as for practical estimates of future seismic hazard. The latter requires model realizations that are tuned to a specific fault zone by assimilating available observational results and their uncertainties. In a case study, the seismic hazard in the Parkfield region has been estimated by combining such a tuned model with few observational data. The use of Bayesian analysis allows us to construct a flexible hazard model for this region which can, in general, incorporate statistical and non-statistical data (e. g. from paleoseismology and geodesy) to improve and update the estimations of the seismic hazard. This approach is particularly promising for less-well monitored regions, and especially for low-seismicity regions like those in central Europe.

Modification of the stress transfer calculations to account for a statistical preference of earthquake propagation direction on a given fault section, e. g. [6,25], can improve the estimates of seismic hazard associated with large faults. It is also possible to extend the discussed framework to other geohazards with even smaller amount of observational data, e. g. the occurrence of landslides. Since the fault model deals with coupled physical processes leading to interacting earthquakes, a challenging future direction will be the design of a more general model for interacting geohazards including earthquakes on different faults as well as landslides triggered by earthquakes, and perhaps tsunamis initiated by (submarine) earthquakes or landslides.

Acknowledgments

We thank Jürgen Kurths, James R. Rice, Frank Scherbaum, Karin Dahmen, Donald L. Turcotte, and many others for useful discussions. GZ and MH acknowledge support from the collaborative research center "Complex Nonlinear Processes" (SFB 555) of the German Research Society (DFG). SH acknowledges support from the DFG-project SCHE280/14 and the EU-project SAFER. YBZ acknowledges support from the National Science Foundation, the United States Geological Survey, and the Southern California Earthquake Center. GZ and YBZ thank the Kavli Institute for Theoretical Physics, UC Santa Barbara, for hospitality during a 2005 program on Friction, Fracture and Earthquake Physics, and partial support based on NSF grant PHY99–0794. We thank James R. Holliday, Vladimir Keilis-Borok and Willie Lee for providing comments on the paper.

Bibliography

1. Aki K, Richards PG (2002) Quantitative seismology. University Science Books, Sansalito
2. Bak P (1996) How Nature Works. The science of self-organised criticality. Springer, New York

3. Bak P, Tang C (1989) Earthquakes as a phenomenon of self-organised criticality. J Geophys Res 94:15635–156637
4. Båth M (1965) Lateral inhomogeneities in the upper mantle. Tectonophysics 2:483–514
5. Ben-Zion Y (1996) Stress, slip, and earthquakes in models of complex single-fault systems incorporating brittle and creep deformations. J Geophys Res 101:5677–5706
6. Ben-Zion Y (2001) Dynamic rupture in recent models of earthquake faults. J Mech Phys Solids 49:2209–2244
7. Ben-Zion Y (2003) Appendix 2, Key Formulas in Earthquake Seismology. In: Lee WHK, Kanamori H, Jennings PC, Kisslinger C (eds) International Handbook of Earthquake and Engineering Seismology, Part B. Academic Press, San Diego, pp 1857–1875
8. Ben-Zion Y, Lyakhovsky V (2002) Accelerated seismic release and related aspects of seismicity patterns on earthquake faults. Pure Appl Geophys 159:2385–2412
9. Ben-Zion Y, Lyakhovsky V (2006) Analysis of aftershocks in a lithospheric model with seismogenic zone governed by damage rheology. J Geophys Int 165:197–210; doi:10.1111/j.1365-246X2006.02878.x
10. Ben-Zion Y, Rice JR (1993) Earthquake failure sequences along a cellular fault zone in a three-dimensional elastic solid containing asperity and nonasperity regions. J Geophys Res 98:14109–14131
11. Ben-Zion Y, Dahmen K, Lyakhovsky V, Ertas D, Agnon A (1999) Self-driven mode switching of earthquake activity on a fault system. Earth Plan Sci Lett 172:11–21
12. Ben-Zion Y, Eneva M, Liu Y (2003) Large Earthquake Cycles and Intermittent Criticality On Heterogeneous Faults Due To Evolving Stress and Seismicity. J Geophys Res 108:2307; doi:10.1029/2002JB002121
13. Bernardo JM, Smith AFM (1994) Bayesian Theory. Wiley, Chichester
14. Binney JJ, Dowrick NJ, Fisher AJ, Newman MEJ (1993) The theory of critical phenomena. Oxford University Press, Oxford
15. Bowman DD, Oullion G, Sammis CG, Sornette A, Sornette D (1998) An observational test of the critical earthquake concept. J Geophys Res 103:24359–24372
16. Brace WF (1960) An extension of the Griffith theory of fracture to rocks. J Geophys Res 65:3477–3480
17. Bufe CG, Varnes DJ (1993) Predicitive modeling of the seismic cycle of the greater San Francisco Bay region. J Geophys Res 98:9871–9883
18. Burridge R, Knopoff L (1967) Model and theoretical seismicity. Bull Seim Soc Am 57:341–371
19. Byerlee JD (1978) Friction of rocks. Pure Appl Geophys 116:615–616
20. Chinnery M (1963) The stress changes that accompany strike-slip faulting. Bull Seim Soc Am 53:921–932
21. Corral Á (2004) Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. Phys Rev Lett 92:108501; doi:10.1103/PhysRevLett.92.108501
22. Dahmen K, Ertas D, Ben-Zion Y (1998) Gutenberg–Richter and characteristic earthquake behavior in simple mean-field models of heterogeneous faults. Phys Rev E 58:1494–1501
23. Daley DJ, Vere-Jones D (1988) An Introduction to the Theory of Point Processes, Springer Series: Probability and its Applications. Springer, Heidelberg
24. Dieterich JH (1994) A constitutive law for earthquake production and its application to earthquake clustering. J Geophys Res 99:2601–2618
25. Dor O, Rockwell TK, Ben-Zion Y (2006) Geologic observations of damage asymmetry in the structure of the San Jacinto, San Andreas and Punchbowl faults in southern California: A possible indicator for preferred rupture propagation direction. Pure Appl Geophys 163:301–349; doi:10.1007/s00024-005-0023-9
26. Ellsworth WL, Matthews MV, Nadeau RM, Nishenko SP, Reasenberg PA, Simpson RW (1999) A physically based earthquake recurrence model for estimation of long-term earthquake probabilities. US Geol Surv Open-File Rept, pp 99–522
27. Fisher DS, Dahmen K, Ramanathan S, Ben-Zion Y (1997) Statistics of earthquakes in simple models of heterogeneous faults. Phys Rev Lett 78:4885–4888
28. Geller RJ, Jackson DD, Kagan YY, Mulargia F (1997) Earthquakes cannot be predicted. Science 275:1616–1617
29. Gumbel EJ (1960) Multivariate Extremal Distributions. Bull Inst Int Stat 37:471–475
30. Gutenberg B, Richter CF (1956) Earthquake magnitude, intensity, energy and acceleration. Bull Seismol Soc Am 46:105–145
31. Hainzl S, Zöller G (2001) The role of disorder and stress concentration in nonconservative fault systems. Phys A 294:67–84
32. Hainzl S, Zöller G, Kurths J (1999) Similar power laws for fore- and aftershock sequences in a spring-block model for earthquakes. J Geophys Res 104:7243–7253
33. Hainzl S, Zöller G, Kurths J (2000) Self-organization of spatio-temporal earthquake clusters. Nonlin Proc Geophys 7:21–29
34. Hainzl S, Zöller G, Kurths J, Zschau J (2000) Seismic quiescence as an indicator for large earthquakes in a system of self-organized criticality. Geophys Res Lett 27:597–600
35. Hainzl S, Zöller G, Scherbaum F (2003) Earthquake clusters resulting from delayed rupture propagation in finit fault segments. J Geophys Res 108:2013; doi:10.1029/2001JB000610
36. Hillers G, Mai PM, Ben-Zion Y, Ampuero JP (2007) Statistical Properties of Seismicity Along Fault Zones at Different Evolutionary Stages. J Geophys Int 169:515–533; doi:10.1111/j.1365-246X2006.03275.x
37. Huang J, Turcotte DL (1990) Are earthquakes an example of deterministic chaos? Geophys Res Lett 17:223–226
38. Jaumé SC, Sykes LR (1999) Evolving towards a critical point: A review of accelerating seismic moment/energy release prior to large and great earthquakes. Pure Appl Geophys 155:279–306
39. Jones LM, Molnar P (1979) Some characteristics of foreshocks and their possible relation to earthquake prediction and premonitory slip on faults. J Geophys Res 84:3596–3608
40. Kagan YY, Knopoff L (1978) Statistical study of the occurrence of shallow earthquakes. J Geophys R Astron Soc 55:67–86
41. Keilis-Borok VI, Soloviev AA (2003) Nonlinear Dynamics of the Lithosphere and Earthquake Prediction, Springer Series in Synergetics. Springer, Heidelberg
42. Lomnitz-Adler J (1999) Automaton models of seismic fracture: constraints imposed by the magnitude-frequency relation. J Geophys Res 98:17745–17756
43. Main IG, O'Brian G, Henderson JR (2000) Statistical physics of earthquakes: Comparison of distribution exponents for source area and potential energy and the dynamic emergence of log-periodic quanta. J Geophys Res 105:6105–6126

44. Matthews MV, Ellsworth WL, Reasenberg PA (2002) A Brownian model for recurrent earthquakes. Bull Seism Soc Am 92:2233–2250
45. Narteau C, Shebalin P, Hainzl S, Zöller G, Holschneider M (2003) Emergence of a band-limited power law in the aftershock decay rate of a slider-block model. Geophys Res Lett 30:1568; doi:10.1029/2003GL017110
46. Nur A, Booker JR (1972) Aftershocks caused by pore fluid flow? Science 175:885–887
47. Okada Y (1992) Internal deformation due to shear and tensile faults in a half space. Bull Seism Soc Am 82:1018–1040
48. Olami Z, Feder HS, Christensen K (1992) Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes. Phys Rev Lett 68:1244–1247
49. Omori F (1894) On the aftershocks of earthquakes. J Coll Sci Imp Univ Tokyo 7:111–200
50. Patel JK, Kapadia CH, Owen DB (1976) Handbook of statistical distributions. Marcel Dekker, New York
51. Reasenberg P (1985) Second-order moment of central California seismicity. J Geophys Res 90:5479–5495
52. Reid HF (1910) The Mechanics of the Earthquake, The California Earthquake of April 18, 1906. Report of the State Investigation Commission, vol 2. Carnegie Institution of Washington, Washington
53. Rundle JB, Klein W, Turcotte DL, Malamud BD (2000) Precursory seismic activation and critical point phenomena. Pure Appl Geophys 157:2165–2182
54. Saleur H, Sammis CG, Sornette D (1996) Discrete scale invariance, complex fractal dimensions, and log-periodic fluctuations in seismicity. J Geophys Res 101:17661–17677
55. Savage JC, Svarc JL, Prescott WH (1999) Geodetic estimates of fault slip rates in the San Francisco Bay area. J Geophys Res 104:4995–5002
56. Scholz CH (1998) Earthquakes and friction laws. Nature 391: 37–42
57. Shcherbakov R, Turcotte DL (2004) A damage mechanics model for aftershocks. Pure Appl Geophys 161:2379; doi:10.1007/s00024-004-2570-x
58. Shin TC, Teng TL (2001) An overview of the 1999, Chichi, Taiwan, earthquake. Bull Seismol Soc Am 91:895–913
59. Sornette D (2004) Self-organization and Disorder: Concepts & Tools, Springer Series in Synergetics. Springer, Heidelberg
60. Sornette D, Sammis CG (1995) Complex critical exponents from renormalization group theory of earthquakes: Implication for earthquake predicitions. J Phys 1(5):607–619
61. Sornette D, Sornette A (1999) Renormalization of earthquake aftershocks. Geophys Res Lett 6:1981–1984
62. Field EH et al. (2007) Special Issue: Regional Earthquake Likelihood Models. Seismol Res Lett 78:1
63. Steacy SJ, McCloskey J, Bean CJ, Ren JW (1996) Heterogeneity in a self-organized critical earthquake model. Geophys Res Lett 23:383–386
64. Turcotte DL (1997) Fractals and chaos in geology and geophysics. Cambridge University Press, New York
65. Turcotte DL, Newman WI, Shcherbakov R (2003) Micro and macroscopic models of rock fracture. J Geophys Int 152:718–728
66. Utsu T (2002) Statistical features of seismicity. In: Int Assoc Seismol & Phys Earth's Interior (ed) International handbook of earthquake and engineering seismology, vol 81A. Academic Press, San Diego, pp 719–732
67. Utsu T, Ogata Y, Matsu'ura RS (1995) The centenary of the Omori formula for a decay law of aftershock activity. J Phys Earth 43:1–33
68. Wesnousky SG (1994) The Gutenberg–Richter or characteristic earthquake distribution, which is it? Bull Seismol Soc Am 90:525–530; 84:1940–1959
69. Wiemer S, Baer M (2000) Mapping and removing quarry blast events from seismic catalogs: Examples from Alaska, the Western United States, and Japan. Bull Seismol Soc Am 90:525–530
70. Working Group on California Earthquake Probabilities (2003) Earthquake probabilities in the San Francisco Bay region. US Geol Survey Open File Report 03–214, US Geological Survey
71. Wyss M (1997) Cannot earthquakes be predicted? Science 278:487
72. Wyss M, Habermann RE (1988) Precursory seismic quiescence. Pure Appl Geophys 126:319–332
73. Zaliapin I, Liu Z, Zöller G, Keilis-Borok V, Turcotte DL (2002) On increase of earthquake correlation length prior to large earthquakes in California. Comp Seismol 33:141–161
74. Zöller G, Ben-Zion Y, Holschneider M, Hainzl S (2007) Estimating recurrence times and seismic hazard of large earthquakes on an individual fault. J Geophys Int 170:1300–1310; doi:10.1111/j.1365-246X200703480.x
75. Zöller G, Hainzl S (2001) Detecting premonitory seismicity patterns based on critical point dynamics. Nat Hazards Earth Syst Sci 1:93–98
76. Zöller G, Hainzl S (2002) A systematic spatiotemporal test of the critical point hypothesis for large earthquakes. Geophys Res Lett 29:1558; doi:10.1029/2002GL014856
77. Zöller G, Hainzl S, Kurths J (2001) Observation of growing correlation length as an indicator for critical point behavior prior to large earthquakes. J Geophys Res 106:2167–2175
78. Zöller G, Hainzl S, Kurths J, Zschau J (2002) A systematic test on precursory seismic quiescence in Armenia. Nat Hazards 26:245–263
79. Zöller G, Holschneider M, Ben-Zion Y (2004) Quasi-static and quasi-dynamic modeling of earthquake failure at intermediate scales. Pure Appl Geophys 161:2103–2118; doi:10.1007/s00024-004-2551-0
80. Zöller G, Holschneider M, Ben-Zion Y (2005) The role of heterogeneities as a tuning parameter of earthquake dynamics. Pure Appl Geophys 162:1027; doi:10.1007/s00024-004-2660-9
81. Zöller G, Hainzl S, Holschneider M, Ben-Zion Y (2005) Aftershocks resulting from creeping sections in a heterogeneous fault. Geophys Res Lett 32:L03308; doi:10.1029/2004GL021871
82. Zöller G, Hainzl S, Ben-Zion Y, Holschneider M (2006) Earthquake activity related to seismic cycles in a model for a heterogeneous strike-slip fault. Tectonophys 423:137–145; doi:10.1016/j.tecto.2006.03.007

Seismicity, Statistical Physics Approaches to

Didier Sornette[1], Maximilian J. Werner[2]

[1] Department of Management, Technology and Economics, ETH Zurich, Switzerland

[2] Swiss Seismological Service, Institute of Geophysics, ETH Zurich, Switzerland

Article Outline

Glossary
Definition of the Subject
Introduction
Concepts and Calculational Tools
Competing Mechanisms and Models
Empirical Studies of Seismicity Inspired by Statistical Physics
Future Directions
Bibliography

Glossary

Chaos Chaos occurs in dynamical systems with two ingredients: (i) nonlinear recurrent re-injection of the dynamics into a finite domain in phase space and (ii) exponential sensitivity of the trajectories in phase space to initial conditions.

Continuous phase transitions If there is a finite discontinuity in the first derivative of the thermodynamic potential, then the phase transition is termed first-order. During such a transition, a system either absorbs or releases a fixed amount of latent heat (e. g. the freezing/melting of water/ice). If the first derivative is continuous but higher derivatives are discontinuous or infinite, then the phase transition is called continuous, of the second kind, or critical. Examples include the critical point of the liquid–gas transition, the Curie point of the ferromagnetic transition, or the superfluid transition [127,235].

Critical exponents Near the critical point, various thermodynamic quantities diverge as power laws with associated critical exponents. In equilibrium systems, there are scaling relations that connect some of the critical exponents of different thermodynamic quantities [32,127,203,216,235].

Critical phenomena Phenomena observed in systems that undergo a continuous phase transition. They are characterized by scale invariance: the statistical properties of a system at one scale are related to those at another scale only through the ratio of the two scales and not through any one of the two scales individually. The scale invariance is a result of fluctuations and correlations at all scales, which prevents the system from being separable in the large scale limit at the critical point [32,203,235].

Declustering In studies of seismicity, declustering traditionally refers to the deterministic identification of fore-, main- and aftershocks in sequences (or clusters) of earthquakes clustered in time and space. Recent, more sophisticated techniques, e. g. stochastic declustering, assign to earthquakes probabilities of being triggered or spontaneous.

Dynamical scaling and exponents Non-equilibrium critical phase transitions are also characterized by scale invariance, scaling functions and critical exponents. Furthermore, some evidence supports the claim that universality classes also exist for non-equilibrium phase transitions (e. g. the directed percolation and the Manna universality class in sandpile models), although a complete classification of classes is lacking and may in fact not exist at all. Much interest has recently focused on directed percolation, which, as the most common universality class of absorbing state phase transitions, is expected to occur in many physical, chemical and biological systems [85,135,203].

Finite size scaling If a thermodynamic or other quantity is investigated at the critical point under a change of the system size, the scaling behavior of the quantity with respect to the system size is known as finite size scaling [32]. The quantity may refer to a thermodynamic quantity such as the free energy or it may refer to an entire probability distribution function. At criticality, the sole length scale in a finite system is the upper cut-off s_c, which diverges in the thermodynamic limit $L \to \infty$. Assuming a lower cut-off $s_0 \ll s_c$, s, a finite size scaling ansatz for the distribution $P(s; s_c)$ of the observable variable s, which depends on the upper cut-off s_c is then given by:

$$P(s; s_c) = a s^{-\tau} G(s/s_c) \quad \text{for} \quad s, s_c \gg s_0 \,, \tag{1}$$

where the parameter a is a non-universal metric factor, τ is a universal (critical) exponent, and G is a universal scaling function that decays sufficiently fast for $s \gg s_c$ [32,36]. Pruessner [163] provides a simple yet instructive and concise introduction to scaling theory and how to find associated exponents. System-specific corrections appear to sub-leading order.

Fractal A deterministic or stochastic mathematical object that is defined by its exact or statistical self-similarity at all scales. Informally, it often refers to a rough or fragmented geometrical shape which can be subdivided into parts which look approximately the same as the original shape. A fractal is too irregular to be described by Euclidean geometry and has a fractal dimension that is larger than its topological dimension but less than the dimension of the space it occupies.

Mean-Field An effective or average interaction field designed to approximately replace the interactions from many bodies by one effective interaction which is constant in time and space, neglecting fluctuations.

Mechanisms for power laws Power laws may be the hallmark of critical phenomena, but there are a host of other mechanisms that can lead to power laws (see Chapter 14 of [203] for a list of power law mechanisms as well as [37,143]). Observations of scale invariant statistics therefore do not necessarily imply SOC, of course. Power laws express the existence of a symmetry (scale invariance) and there are many mechanisms by which a symmetry can be obtained or restored.

Non-equilibrium phase transitions In contrast to systems at equilibrium, non-equilibrium phase transitions involve dynamics, energy input and dissipation. Detailed balance is violated and no known equivalent of the partition function exists, from which all thermodynamic quantities of interest derive in equilibrium. Examples of non-equilibrium phase transitions include absorbing state phase transitions, reaction-diffusion models, and morphological transitions of growing surfaces [85,135].

Phase transitions In (equilibrium) statistical mechanics, a phase transition occurs when there is a singularity in the free energy or one of its derivatives. Examples include the freezing of water, the transition from ferromagnetic to paramagnetic behavior in magnets, and the transition from a normal conductor to a superconductor [127,235].

Renormalization group theory A mathematical theory built on the idea that the critical point can be mapped onto a fixed point of a suitably chosen transformation on the system's Hamiltonian. It provides a foundation for understanding scaling and universality and provides tools for calculating exponents and scaling functions. Renormalization group theory provides the basis for our understanding of critical phenomena [32,216,235]. It has been extended to non-Hamiltonian systems and provides a general framework for constructing theories of the macro-world from the microscopic description.

Self-organized criticality (SOC) Despite two decades of research since its inception by [13] and the ambitious claim by [11] that, as a mechanism for the ubiquitous power laws in Nature, SOC was "How Nature Works", a commonly accepted definition along with necessary and sufficient conditions for SOC is still lacking [93,163,203]. A less rigorous definition may be the following: Self-organized criticality refers to a non-equilibrium, critical and marginally stable steady-state, which is attained spontaneously and without (explicit) tuning of parameters. It is characterized by power law event distributions and fractal geometry (in some cases) and may be expected in slowly driven, interaction-dominated threshold systems [93]. Some authors additionally require that temporal and/or spatial correlations decay algebraically (e. g. [84], but see [163]). Definitions in the literature range from broad (simply the absence of characteristic length scales in non-equilibrium systems) to narrow (the criticality is due to an underlying continuous phase transition with all of its expected properties) (see, e. g., [162] for evidence that precipitation is an instance of the latter definition of SOC in which a non-linear feedback of the order parameter on the control parameter turns a critical phase transition into a self-organized one attracting the dynamics [198]).

Spinodal decomposition In contrast to the slow process of phase separation via nucleation and slow growth of a new phase in a material inside the metasstable region near a first-order phase transition, spinodal decomposition is a non-equilibrium, rapid and critical-like dynamical process of phase separation that occurs quickly and throughout the material. It needs to be induced by rapidly quenching the material to reach a sub-area (sometimes a line) of the unstable region of the phase diagram which is characterized by a negative derivative of the free energy.

Statistical physics is the set of concepts and mathematical techniques allowing one to derive the large-scale laws of a physical system from the specification of the relevant microscopic elements and of their interactions.

Turbulence In fluid mechanics, turbulence refers to a regime in which the dynamics of the flow involves many interacting degrees of freedom, and is very complex with intermittent velocity bursts leading to anomalous scaling laws describing the energy transfer from injection at large scales to dissipation at small scales.

Universality In systems with little or no frozen disorder, equilibrium continuous phase transitions fall into a small set of universality classes that are characterized by the same critical exponents and by certain scaling functions that become identical near the critical point. The class depends only on the dimension of the space and the dimension of the order parameter. For instance, the critical point of the liquid–gas transition falls into the same universality class as the 3D Ising model. Even some phase transitions occurring in high-energy physics are expected to belong to the Ising class. Universality justifies the development and study of extremely simplified models (caricatures) of Nature, since the behavior of the system at the critical point can nevertheless be captured (in some cases exactly). How-

ever, non-universal features remain even at the critical point but are less important, e. g. amplitudes of fluctuations or system-specific corrections to scaling that appear at sub-leading order [32,216,235,239].

Definition of the Subject

A fundamental challenge in many scientific disciplines concerns upscaling, that is, of determining the regularities and laws of evolution at some large scale from those known at a lower scale: biology (from molecules to cells, from cells to organs); neurobiology (from neurons to brain function), psychology (from brain to emotions, from evolution to understanding), ecology (from species to the global web of ecological interactions), condensed matter physics (from atoms and molecules to organized phases such as solid, liquid, gas, and intermediate structures), social sciences (from individual humans to social groups and to society), economics (from producers and consumers to the whole economy), finance (from investors to the global financial markets), Internet (from e-pages to the world wide web 2.0), semantics (from letters and words to sentences and meaning), and so on. Earthquake physics is no exception, with the challenge of understanding the transition from the laboratory scale (or even the microscopic and atomic scale) to the scale of fault networks and large earthquakes.

Statistical physics has had a remarkably successful track record in addressing the upscaling problem in physics. While the macroscopic laws of thermodynamics have been established in the 19th century, their microscopic underpinning were elaborated in the early 20th century by Boltzmann and followers, building the magnificent edifice of statistical physics. Statistical physics can be defined as the set of concepts and mathematical techniques allowing one to derive the large-scale laws of a physical system from the specification of the relevant microscopic elements and of their interactions. Dealing with huge ensembles of elements (atoms, molecules) of the order of the Avogadro number ($\simeq 6 \cdot 10^{23}$), statistical physics uses the mathematical tools of probability theory combined with other relevant fields of physics to calculate the macroscopic properties of large populations.

One of the greatest achievement of statistical physics was the development of the renormalization group analysis, to construct a theory of interacting fields and of critical phase transitions. The renormalization group is a perfect example of how statistical physics addresses the micro-macro upscaling problem. It decomposes a problem of finding the macroscopic behavior of a large number of interacting parts into a succession of simpler problems with a decreasing number of interacting parts, whose effective properties vary with the scale of observation. The renormalization group thus follows the proverb "divide to conquer" by organizing the description of a system scale-by-scale. It is particularly adapted to critical phenomena and to systems close to being scale-invariant. The renormalization group translates into mathematical language the concept that the overall behavior of a system is the aggregation of an ensemble of arbitrarily defined sub-systems, with each sub-system defined by the aggregation of sub-subsystems, and so on [203].

It is important to stress that up to now the term "statistical" has different meanings in statistical physics and in statistical seismology, a field that has developed as a marriage between probability theory, statistics and the part of seismology concerned with empirical patterns of earthquake occurrences [225] (but not with physics). Statistical seismology uses stochastic models of seismicity, which are already effective large-scale representation of the dynamical organization. In contrast, a statistical physics approach to earthquake strives to derive these statistical models or other descriptions from the knowledge of the microscopic laws of friction, damage, rupture, rock-water interactions, mechano-chemistry and so on, at the microscopic scales [200,201]. In other words, what is often missing in statistical seismology is the physics to underpin the stochastic model on physically-based laws, e. g. rate-and-state friction [55].

The previously mentioned successes of statistical physics promote the hope that a similar program can be developed for other fields, including seismology. The successes have been more limited, due to the much more complex interplay between mechanisms, interactions and scales found in these out-of-equilibrium systems. This short essay provides a subjective entry to understand some of the different attempts, underlining the few successes, the problems and open questions. Rather than providing an exhaustive review, we mention what we believe to be important topics and have especially included recent work.

Introduction

Much of the recent interest of the statistical physics community has focused on applying scaling techniques, which are common tools in the study of critical phenomena, to the statistics of inter-event recurrence times or waiting times [14,40,41,42,43,44,46,48,134]. However, the debate over the relevance of critical phenomena to earthquakes stretches back as far as 30 years [7,11,12,47,61,84,93,104,106,109,114,147,151,178,192,193,194,202,203,207,223], ► Jerky Motion in

Slowly Driven Magnetic and Earthquake Fault Systems, Physics of and ► Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space. The current debate on recurrence statistics is thus the latest tack in an evolving string of arguments with a long history. As discussed below, many of the claims made in the recent articles on recurrence statistics have either been challenged, refuted or explained by previously known facts about earthquake statistics [132,133,145,176,177,230]. As will be discussed below, this debate in the literature is important because of the potential consequences for understanding earthquakes, but it needs to be pursued with rigorous scientific arguments accessible to both the seismological and the statistical physics communities.

The debate would almost certainly benefit significantly from testing hypotheses with simulations to establish null hypotheses and benchmarks: seismicity patterns are sufficiently stochastic and earthquake catalogs contain a sufficient amount of observational uncertainties so as to make inference difficult. It is often not straightforward to predict the signal of well-known statistical features such as clustering in new data analysis techniques. Therefore, testing the purported claims by realistic simulations of earthquake catalogs can provide a strong benchmark against which the claims can be evaluated. This view and the corresponding criticism of many studies has been put forward and defended for a long time by Kagan [111].

Such a model-dependent approach may be at odds with the philosophy of a so-called "model-free" analysis, which the community of statistical physicists claim to take in their analysis. For instance, network theory-based approaches, space-time window-based finite size scaling, box-covering methods and other techniques used in the study of critical and fractal phenomena are said to be "model-free" because no assumptions about seismicity are supposedly made at the outset. By using model-free analysis techniques, the often uncertain and sometimes clearly wrong assumptions of flawed models and resulting biased results are meant to be circumvented.

However, as is almost always the case in statistical hypothesis testing, the less assumptions are made about the test, the less powerful the test statistic. More importantly, seismicity is sufficiently stochastic so that well-known features may appear as novel in new analysis methods. Furthermore, to convince the seismological community of new data analysis techniques, the methods need to be tested on established knowledge and show the improvement over traditional methods. These types of initial tests are rarely performed by the statistical physics community.

In the next Sect. "Concepts and Calculational Tools", we present a summary of some of the concepts and calculational tools that have been developed in attempts to apply statistical physics approaches to seismology. Then, Sect. "Competing Mechanisms and Models" summarizes the leading theoretical physical models of the space-time organization of earthquakes. Section "Empirical Studies of Seismicity Inspired by Statistical Physics" presents a general discussion and several examples of the new metrics proposed by statistical physicists, underlining their strengths and weaknesses. Section "Future Directions" briefly outlines expected developments.

Concepts and Calculational Tools

Renormalization, Scaling and the Role of Small Earthquakes in Models of Triggered Seismicity

A common theme in many of the empirical relations in seismology (and in those employed in seismicity models) is the lack of a dominating scale. Many natural phenomena can be approached by the traditional reductionist approach to isolate a process at a particular scale. For example, the waves of an ocean can be described quite accurately by a theory that entirely ignores the fact that the liquid is made out of individual molecules. Indeed, the success of most practical theories in physics depends on isolating a scale [234], although since this recognition, much progress has been made in developing a holistic approach for processes that do not fall into this class. Given current observational evidence, earthquakes seem to belong to the set of processes characterized by a lack of one dominating length scale: fluctuations of many or perhaps a wide continuum of sizes seem to be important and are in no way diminished – even when one is interested solely in large-scale descriptions [206].

The traditional reductionist approach in seismology, which, for instance, attempted to separate large (main) shocks from small (fore- or after-) shocks, is slowly giving way to the holistic approach, in which all earthquakes are created equal and seismicity is characterized by fluctuations of all sizes. This gradual shift is supported, on a conceptual and qualitative level, by the vision of critical phenomena. A particularly strong model of the interactions between earthquakes has emerged in the concept of triggering, which places all earthquakes on the same footing: each earthquake can trigger its own events, which in turn can trigger their own events, and so on, according to the same probability distributions, and the resulting seismicity can be viewed as the superposed cascades of triggered earthquakes that cluster in space and time [77,117,149,150].

From this point of view, it is natural that small earthquakes are important to the overall spatio-temporal patterns of seismicity. Indeed, the scaling of aftershock productivity with mainshock magnitude suggests that small earthquakes are cumulatively as important for the triggered seismicity budget as rarer but larger events [60,76,82]. The importance of small earthquakes has also been documented in, e. g., [73,138,141].

But earthquake catalogs do not contain information (by definition) about the smallest, unobserved events, which we know to exist from acoustic emission experiments and earthquakes recorded in mines. To guarantee a finite seismicity budget, Sornette and Werner argued [209] for the existence of a smallest triggering earthquake, akin to a "ultra-violet cut-off" in quantum field theory, below which earthquakes do not trigger other events. Introducing a formalism which distinguishes between the detection threshold and the smallest triggering earthquake, Sornette and Werner placed constraints on its size by using a simplified version of the popular Epidemic-Type Aftershock Sequence (ETAS) Model [149], a powerful model of triggered seismicity based on empirical statistics, and by using observed aftershock sequences. Sornette and Werner [210] also considered the branching structure of one complete cascade of triggered events, deriving an apparent branching ratio and the apparent number of untriggered events, which are observed when only the structure above the detection threshold is known. As a result of our inability to observe the entire branching structure, inferred clustering parameters are significantly biased and difficult to interpret in geophysical terms. Second, separating triggered from untriggered events, commonly known as declustering, also strongly depends on the threshold, so that it cannot even in theory constitute a physically sound method.

Sornette and Werner [210] also found that a simplified, averaged version of the ETAS model can be renormalized onto itself, with effective clustering parameters, under a change of the threshold. Saichev and Sornette [175] confirmed these results for the stochastic number statistics of the model using a rigorous approach in terms of generating probability functions, but also showed that the temporal statistics could not be renormalized. Furthermore, it can be shown (see Chapter 4 of [229]) that the conditional intensity function of the ETAS model, the mathematical object which uniquely defines the model, cannot be renormalized onto itself under a change of magnitude threshold. It is not a fixed-point of the renormalization process operating via magnitude coarse-graining. The functional form of the model must change under a change in the detection threshold [175]. In other words, if earthquakes occur according to an ETAS model above some cut-off m_0, then earthquakes above m_d cannot be described by the ETAS model in a mathematically exact way. Although in practice, the ETAS model provides an excellent fit.

The issue of how to deal with small earthquakes is thus reminiscent of the decades of efforts that have been invested in physics to deal with the famous ultra-violet cut-off problem, eventually solved by the so-called "renormalization" theory of Feynmann, Schwinger and Tomonaga. In the 1960s and 1970s, this method of renormalization was extended into the "renormalization group" (in fact a semi-group in the strict mathematical sense) for the theory of critical phenomena (see glossary), which we also mention in Sect. "Competing Mechanisms and Models". It is fair to say that there has been limited success in developing a multi-scale description of the physics of earthquakes and, in particular, in addressing the upscaling problem and the impact of the many small earthquakes.

One tantalizing approach, not yet really understood in terms of all its consequences and predictions, is the variant of the ETAS model proposed by Vere-Jones [224], which has the remarkable property of being bi-scale invariant under a transformation involving time and magnitudes. One of the modifications brought in by [224] is to assume that the distribution of the daughter magnitudes is dependent on the mother magnitude m_i through a modification of the Gutenberg–Richter distribution of triggered earthquake magnitudes by a term of the form $\exp(-\delta|m - m_i|)$, where $\delta > 0$ quantifies the distance to the standard Gutenberg–Richter distribution. Remarkably, Saichev and Sornette [174], who studied the Vere-Jones model, found that, due to the superposition of the many magnitude distributions of each earthquake in the cascades of triggered events, the resulting distribution of magnitudes over a stationary catalog is a pure Gutenberg–Richter law. Thus, there might be hidden characteristic scales in the physics of triggering that are not revealed by the standard observable one-point statistical distributions. Simulation and parameter estimation algorithms for the Vere-Jones model are not yet available. If and when these algorithms become available, the study of this bi-scale invariant branching model may be a strong alternative to the ETAS model, as this model is exactly scale invariant with neither ultra-violet nor infra-red cut-offs.

Nevertheless, being empirically based, these stochastic point process models lack a genuine microscopic physical foundation. The underlying physics is not explicitly addressed and only captured effectively by empirical statistics, even at the smallest scales. The physical processes and their renormalization are missing in this approach.

Universality

Universality, as defined in the glossary, justifies the development and study of extremely simplified models (caricatures) of Nature, since the behavior of a studied system at the critical point can nevertheless be captured by toy models (in some cases exactly). For instance, the liquid–gas transition, the ferromagnetic to paramagnetic transition, and the behavior of binary alloys, all apparently different systems, can be described successfully by an extremely simplified picture of Nature (the Ising model) because of universality. The hope that a similar principle holds for earthquakes (and other non-equilbrium systems) underpins many of the models and tools inspired by statistical physics that have been applied to seismicity, to bring about the "much coveted revolution beyond reductionism" [17,67].

Speaking loosely, the appearance of power laws in many toy models is often interpreted as a kind of universal behavior. The strict interpretation of universality classes, however, requires that critical exponents along with scaling functions are identical for different systems. It is interesting to note that slight changes in the sand-pile model already induce new universality classes, so that even within a group of toy models, the promise of universality is not, strictly, fulfilled [13,98,136].

A lively debate in seismology concerns the universality of, on one hand, the frequency-size distribution of earthquake magnitudes (e. g. [47,61,91,108,111,232]), and, on the other hand, the universality of the exponent of the Gutenberg–Richter distribution (e. g. [25,184,212,233]). A spatio-temporally varying critical exponent is not traditionally part of the standard critical phenomena repertoire, although analytical and numerical results based on a simple earthquake model on a fault showed a possible spontaneous switching between Gutenberg–Richter and characteristic earthquake behavior associated with a non-equilibrium phase transition [24,47,61], ► Jerky Motion in Slowly Driven Magnetic and Earthquake Fault Systems, Physics of and ► Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space. Another possible mechanism for the coexistence of and intermittent shifts between different regimes (Gutenberg–Richter scaling and characteristic earthquakes) stems from the competition between several interacting faults which give rise to long intervals of activity in some regions followed by similarly long intervals of quiescence [129,213,214]. Both careful empirical investigations of seismicity parameters and theoretical progress on heterogeneous, spatially extended critical phenomena may help elucidate the controversy.

Intermittent Periodicity and Chaos

As part of the conquest of chaos theory in the 70s and 80s [58,139], its concepts and methods were invariably also applied to seismicity. Huang and Turcotte [87,88] modeled the interaction between two faults by two sliding blocks that are driven by a plate through springs, coupled to one another via another spring and endowed with a velocity-weakening friction law. The dynamical evolution of the blocks showed chaotic behavior and period-doubling (the Feigenbaum route to chaos [58]). Huang and Turcotte [87,88] suggested that the interaction of the Parkfield segment with the southern San Andreas fault may be governed by the kind of chaotic behavior they observed in their model: the lighter block slipped quasi-periodically for several times until both slipped together. Their study explained that apparent quasi-periodicity of earthquakes on fault segments may be a result of chaotic interactions between many fault segments, thereby providing a warning that the extrapolations of quasi-periodic models are not to be trusted (e. g. [15,75] and references therein). However, it is doubtful that models with just a few degrees of freedom can go a long way towards providing deeper physical insights, or predictive tools. One needs to turn to models with a large number of degrees of freedom, for which turbulence appears as the leading paradigm of complexity.

Turbulence

A drastically different approach has been favored by Yan Kagan [104,106,109,114], who described seismicity as the "turbulence of solids" – attesting to the far greater problems in earthquake seismology than the theory of critical phenomena promises to solve. While renormalization group methods and scaling theory have contributed immensely to the study of turbulence [63], the problem of turbulence involves significant additional complications. First, loosely speaking, renormalization group theory helps predict global behavior by coarse-graining over degrees of freedom, which is essentially a bottom–up approach. In turbulence, the enstrophy acts bottom–up, but the energy cascades top down. Secondly, there is a significant spatial and topological aspect to turbulence, for instance involving topological defects, such as filament structures, which are crucial to the dynamical evolution of the system. The existence of the two cascades (top–down and bottom–up) as well as the influence of the dissipation scale all the way within the so-called inertial range makes turbulence the most important problem still unsolved in classical Physics. The importance of addressing the issue of the interplay between the top–down and bottom–up cas-

cades in earthquake toy models has been outlined by [65,66,236,237].

The analogous problem for seismicity lies in the complex fault network, which constrains seismicity through its weak structures but also grows and evolves because of earthquakes redistributing stresses and rupturing fresh surfaces. The statistical description of this tensorial and dynamical problem is only at its beginning [64,99,100,101,102,103,107,116,118,119,142,213]. But it is likely to be a key aspect to the dynamical evolution of faults and seismicity. New physics and approaches are required to tackle the tensorial nature of the stress and strain fields and the complex topological structures of defects, from dislocations to joints and faults, as well as the many different physical processes operating from the atomic scale to the macro-scale [200,201].

Self-Organized Criticality

Self-organized criticality (SOC) refers to the spontaneous organization of a system driven from outside into a globally dynamical statistical stationary state, which is characterized by self-similar distributions of event sizes and sometimes fractal geometrical properties. SOC applies to the class of phenomena occurring in driven out-of-equilibrium systems made of many interactive components, which possess the following fundamental properties: 1) a highly non-linear behavior, 2) a very slow driving rate, 3) a globally stationary regime, characterized by stationary statistical properties, and 4) power-law distributions of event sizes and fractal geometrical properties. The crust obeys these four conditions, as first suggested by [12,193], who proposed to understand the spatio-temporal complexity of earthquakes in this perspective.

The appeal of placing the study of earthquakes in the framework of critical phenomena may be summarized as follows. Power law distributions can be understood as a result of an underlying continuous phase transition into which the crust has organized itself [197,211]. Applying the methods of renormalization group theory may help calculate exponents and scaling functions and rationalize the spatio-temporal organization of seismicity along with its highly correlated structures. For instance, Sornette and Virieux [208] provided a theoretical framework which links the scaling laws exhibited by earthquakes at short times and plate tectonic deformations at large times. Perhaps earthquakes fall into a universality class which can be solved exactly and/or investigated in toy models. Moreover, studying the detailed and highly complicated microphysics involved in earthquakes may not lead to insights about the spatio-temporal organization, because, as a critical phenomenon, the traditional approach of separating length scales to describe systems is inadequate. On the other hand, as mentioned above, there is the possibility of a hierarchy of physical processes and scales which are inter-related [156,157], for which the simplifying approach in terms of critical phenomena is likely to be insufficient.

As another reason for the importance of the topic, interesting consequences for the predictability of earthquakes might be derived, for instance by mapping earthquakes to a genuine critical point (the accelerating moment release hypothesis, e. g. [202,207]) or by mapping earthquakes to SOC (e. g. [70,147]). The latter mapping had led some to argue that earthquakes are inherently unpredictable. In the sandpile paradigm [13], there is little difference between small and large avalanches, and this led similarly to the concept that "large earthquakes are small earthquakes that did not stop," hence their supposed lack of predictability. More than ten years after this contentious proposal, a majority of researchers, including most of the authors of this "impossibility claim," recognize that there is some degree of predictability [83,97]. Actually, the clarifications came from investigators of SOC, who recognized that the long-term organization of sandpiles [51,52] and of toy models of earthquakes and fault networks [142,213,214] is characterized by long-range spatial and temporal correlations. Thus, large events may indeed be preceded by subtle long-range organizational structures, an idea at the basis of the accelerating moment release hypothesis. This idea is also underlying the pattern recognition method introduced by Gelfand et al. [69] and developed extensively by V. Keilis-Borok and his collaborators for earthquake predictions [122]. In addition, Huang et al. [89] showed that avalanche dynamics occurring within hierarchical geometric structures are characterized by significant precursory activity before large events; this provides a clear proof of the possible coexistence between critical-like precursors of large events and a long-term self-organized critical dynamical state.

In summary, self-organized criticality provides a general conceptual framework to articulate the search for a physical understanding of the large-scale and long-time statistical properties of the seismogenic process and of the predictability of earthquakes. Beyond this, it is of little help as many different mechanisms have been documented at the origin of SOC (see, e. g., chapter 15 in [203]). SOC is not a theory, it does not provide any specific calculation tools; it is a concept offering a broad classification of the kinds of dynamics that certain systems, including the Earth crust, seem to spontaneously select.

Competing Mechanisms and Models

It should be noted at this point that the statistical physics approach to earthquake science is not limited to SOC. Over the years, several groups have proposed to apply the concepts and tools of statistical physics to understanding the multiscale dynamics of earthquake and fault systems. Various mechanisms drawn conceptually from statistical mechanics but not necessary even limited to critical (phase transition) phenomena have been proposed and are being pursued. Such approaches include the concept of the critical point earthquake related to accelerated moment release, network theory, percolation and fiber models as models for fracture, and many more, some of which can be found in [84,203,220,221].

In this section, we outline some of the major model classes which underpin distinct views on what are the dominating mechanisms to understand earthquakes and their space-time organization.

Roots of Complexity in Seismicity: Dynamics or Heterogeneity?

The 1990s were characterized by vigorous discussions at the frontier between seismology and statistical physics aimed at understanding the origin of the observed complexity of the spatio-temporal patterns of earthquakes. The debate was centered on the question of whether space-time complexity can occur on a single homogeneous fault, solely as a result of the nonlinear dynamics [23,38,39,128,186,187,188,189], associated with the slip and velocity dependent friction law validated empirically in particular by [53,54,55,56]. Or, is the presence of quenched heterogeneity necessary [21,22,126,168]?

The rediscovery of the multi-slider-block-spring model of [31,33] led to a flurry of investigations by physicists [34,128,170], finding an enticing entry to this difficult field, in the hope of capturing the main empirical statistical properties of seismicity. It is now understood that complexity in the stress field, in co-seismic slips and in sequences of earthquakes can emerge purely from the nonlinear laws. However, heterogeneity is probably the most important factor dominating the multi-scale complex nature of earthquakes and faulting [156,157,181,182]. It is also known to control the appearance of self-organized critical behavior in a class of models relevant to the crust [191,214].

Critical Earthquakes

This section gives a brief history of the "critical earthquake" concept.

We trace the ancestor of the critical earthquake concept to Vere-Jones [223], who used a branching model to illustrate that rupture can proceed through a cascade of damage events. Allègre et al. [7] proposed what is in essence a percolation model of damage/rupture describing the state of the crust before an earthquake. They formulated the model using the language of real-space renormalization group, in order to emphasize the multi-scale nature of the underlying physics, and the incipient rupture as the approach to a critical percolation point. Their approach is actually a reformulation in the language of earthquakes of the real-space renormalization group approach to percolation developed by [165]. Chelidze [35] independently developed similar ideas. In the same spirit, Smalley et al. [192] proposed a renormalization group treatment of a multi-slider-block-spring model. Sornette and Sornette [194] took seriously the concept put forward by [7] and proposed to test it empirically by searching for the predicted critical precursors. Voight [227,228] was probably the first author to introduce the idea of a time-to-failure analysis quantified by a second order nonlinear ordinary differential equation. For certain values of the parameters, the solution of [227,228]'s time-to-failure equation takes the form of a finite time singularity (see [180] for a review and [204] for a mechanism based on the ETAS model). He proposed and did use it later to predict volcanic eruptions. The concept that earthquakes are somehow associated with critical phenomena was also underlying the research efforts of a part of the Russian school [120,222].

The empirical seed for the critical earthquake concept were the repeated observations that large earthquakes are sometimes preceded by an increase in the number of intermediate size events [29,59,92,96,121,123,131,144,164,217]. The relation between these intermediate events and the subsequent main event took a long time to be recognized because the precursory events occur over such a large area. Sykes and Jaumé [217] proposed a specific law $\sim \exp[t/\tau]$ quantifying the acceleration of seismicity prior to large earthquakes. Bufe and Varnes [30] proposed that the finite-time singularity law

$$\epsilon_{\text{Benioff}} \sim 1/(t_c - t)^m \tag{2}$$

is a better empirical model than the exponential law. In (2), $\epsilon_{\text{Benioff}}$ is the cumulative Benioff strain, t_c is critical time of the occurrence of the target earthquake and m is a positive exponent. The fit with this law of the empirical Benioff strain calculated by summing the contribution of earthquakes in a given space-time window is supposed to provide the time t_c of the earthquake and thus constitutes

a prediction. This expression (2) was justified by a mechanical model of material damage. It is important to understand that the law (2) can emerge as a consequence of a variety of mechanisms, as reviewed by [180].

One of these mechanisms has been coined the "critical earthquake" concept, first formulated by Sornette and Sammis [207], who proposed to reinterpret the formula (2) proposed by [30] and previous related works by generalizing them within the statistical physics framework. This concept views a large earthquake as a genuine critical point. Using the insight of critical points in rupture phenomena, Sornette and Sammis [207] proposed to enrich Eq. (2), now interpreted as a kind of diverging susceptibility in the sense of critical phenomena, by considering complex exponents (i. e. log-periodic corrections to scaling). These structures accommodate the possible presence of a hierarchy of characteristic scales, coexisting with power laws expressing the scale invariance associated with a critical phenomenon [199]. This was followed by several extensions [89,94,95,178]. Sornette [202] reviewed the concept of critical "ruptures" and earthquakes with application to prediction. Ike and Sornette [90] presented a simple dynamical mechanism to obtain finite-time singularities (in rupture in particular) decorated by complex exponents (log-periodicity). Bowman et al. [28, 153,154,242,243] proposed empirical tests of the critical earthquake concept. The early tests of [28] have been criticized by [74], while [226] commented on the lack of a formal statistical basis of the accelerating moment release model. This stresses the need for rigorous tests in the spirit of [166,167]. The debate is wide open, especially in view of the recent developments to improve the determination of the relevant spatio-temporal domain that should be used to perform the analyzes [26,27,124,130] (see [62] for a review).

Spinodal Decomposition

Klein, Rundle and their collaborators have suggested a mean-field approach to the multi-slider-block-spring model justified by the long-range nature of the elastic interactions between faults. This has led them to propose that the fluctuations of the strain and stress field associated with earthquakes are technically those occurring close to a spinodal line of an underlying first-order phase transition (see [125,169,172] and references therein). This conceptual view has inspired them to develop the "Pattern Informatics" technique, an empirical seismicity forecasting method based on the idea that changes in the seismicity rate are proxies for changes in the underlying stress [86, 218].

The fluctuations associated with a spinodal line are very similar to those observed in critical phenomena. It is thus very difficult if not impossible in principle to falsify this hypothesis against the critical earthquake hypothesis, since both are expected to present similar if not identical signatures. Perhaps, the appeal of the spinodal decomposition proposal has to be found at the theoretical level, from the fact that first-order phase transitions are more generic and robust than critical phenomena, for systems where heterogeneity and quenched randomness are not too large.

Dynamics, Stress Interaction and Thermal Fluctuation Effects

Fisher et al. [24], Dahmen et al. [47], Ben-Zion et al. [61] and co-workers (see the reviews by ► Jerky Motion in Slowly Driven Magnetic and Earthquake Fault Systems, Physics of and ► Seismicity, Critical States of: From Models to Practical Seismic Hazard Estimates Space) have introduced a mean-field model (resulting from a uniform long-range Green function) of a single fault, whose dynamical organization is controlled by two control parameters, ϵ which measures the dynamic stress weakening and c which is the deviation from stress conservation (due for instance to coupling with ductile layers). The point ($\epsilon = 0$; $c = 0$) is critical in the sense of a phase transition in statistical physics, with its associated scale invariant fluctuations described by power laws. Dynamic stress strengthening ($\epsilon < 0$) leads to truncated Gutenberg–Richter power laws. Dynamic stress weakening ($\epsilon > 0$) is either associated with a truncated Gutenberg–Richter power law for $c > 0$ or with characteristic earthquakes decorating a truncated power law for $c < 0$. The coexistence of a characteristic earthquake regime with a power-law regime is particularly interesting as it suggests that they are not exclusive properties but may characterize the same underlying physics under slightly different conditions. This could provide a step towards explaining the variety of empirical observations in seismology [5,110,185, 232].

Sornette et al. [214] have obtained similar conclusions using a quasi-static model in which faults grow and self-organize into optimal structures by repeated earthquakes. Depending on the value of dynamical stress drop (controlling the coupling strength between elements) relative to the amplitude of the frozen heterogeneity of the stress thresholds controlling the earthquake nucleation on each fault segment, a characteristic earthquake regime with a truncated power law is found for small heterogeneity or large stress drop while the power law SOC regime is recovered by large heterogeneity or small stress drop. The two

approaches of [24,47,61] and ► Jerky Motion in Slowly Driven Magnetic and Earthquake Fault Systems, Physics of on the one hand and of [214] on the other hand can be reconciled conceptually by noting that the dynamic stress weakening of ► Jerky Motion in Slowly Driven Magnetic and Earthquake Fault Systems, Physics of controls the dynamical generated stress heterogeneity while the lack of stress conservation c controls the coupling strength. Fundamentally, the relevant control parameter is the degree of coupling between fault elements seen as threshold oscillators of relaxation versus the variance of the disorder in their spontaneous large earthquake recurrence times. Generically, power law statistics are expected to co-exist with synchronized behavior in a general phase diagram in the heterogeneity-coupling strength plane [152,214].

Let us finally mention a promising but challenging theoretical approach, which has the ambition to bridge the small-scale physics controlled by thermal nucleation of rupture to the large-scale organization of earthquakes and faults [155,205]. Partial success has been obtained with a remarkable prediction on the ("multifractal") dependence of the Omori law exponent of aftershocks on the magnitude of the mainshock, verified by careful empirical analyzes on earthquakes in California, Japan and worldwide [158].

Empirical Studies of Seismicity Inspired by Statistical Physics

> *"False facts are highly injurious to the progress of science, for they often endure long; but false views, if supported by some evidence, do little harm, for everyone takes a salutary pleasure in proving their falseness."*
>
> Charles Darwin, in The Origin of Man, Chap. 6.

Early Successes and Subsequent Challenges

A significant benefit of the statistical physics approach to seismology has been the introduction of novel techniques to analyze the available empirical data sets, with the goal of obtaining new insights into the spatio-temporal organization of seismicity and of revealing novel regularities and laws that may guide the theoretical analysis.

A prominent forerunner is the application of the concept of fractals introduced by Mandelbrot [137] and of the measures of fractal dimensions to describe complex sets of earthquake epicenters, hypocenters and fault patterns. The use of fractals has constituted an epistemologic breakthrough in many fields, and not only in seismology. Indeed, before Mandelbrot, when dealing with most complex systems, one used to say: "this is too complicated for a quantitative analysis" and only qualitative descriptions were offered. After Mandelbrot, one could hear: "this is a fractal, with a fractal dimension equal to xxx!" By providing a new geometrical way of thinking about complex systems associated with novel metrics, Mandelbrot and his fractals have extended considerably the reach of quantitative science to many complex systems in all fields.

However, while there have been some attempts to use fractal dimensions as guidelines to infer the underlying organization processes, as for instance in [195,196], most of the initial reports have lost their early appeal [18,19,183,219] since the complexity of seismicity and faulting is much too great to be captured by scaling laws embodying solely a simple scale invariance symmetry. Among others, multifractal and adapted wavelet tools are needed to quantify this complexity, see for instance [68,146,156,157,159]. It should also be noted that few studies of the fractal dimensions of seismicity address the significant issues of errors, biases and incomplete records in earthquake catalogs – a notable exception being [115].

Since the beginning of the 21st century, a renewal of interest and efforts have burgeoned as groups of statistical physicists, interested in earthquakes as a potential instance of self-organized criticality (SOC), have claimed "novel", "universal" and "robust" scaling laws from their analysis of the spatio-temporal organization of seismicity. The authors purport to have discovered universal and hitherto unknown features of earthquakes that give new insights into the dynamics of earthquakes and add to the evidence that earthquakes are self-organized critical. We now discuss a few of these recent studies to illustrate the existence of potential problems in the "statistical physics" approach. In a nutshell, we show that perhaps most of these "novel scaling laws" can be explained entirely by already known statistical seismicity laws. This claim has been defended by other experts of statistical seismology, the most vocal being perhaps Yan Kagan at UCLA [111].

The flurry of interest from physicists comes from their fascination with the self-similar properties exhibited by seismicity (e. g. the Gutenberg–Richter power law of earthquake seismic moments, the Omori–Utsu law of the decay of aftershock rates after large earthquakes, the fractal and multifractal space-time organization of earthquakes and faults, etc.), together with the development of novel concepts and techniques that may provide new insights. But, and this is our main criticism based on several detailed examples discussed below, many of the new approaches and results do not stand close scrutiny. This failure is rooted in two short-comings: (i) the lack of testing of new methods on synthetic catalogs generated by benchmark mod-

els which are based on well-known statistical laws of seismicity, and (ii) the failure to consider earthquake catalog bias, incompleteness and errors. The latter may cause catalog artifacts to appear as genuine characteristics of earthquakes. Testing the results on a variety of catalogs and considering the influence of various catalog errors can help minimize their influence. The former short-coming often leads to the following scenario: authors fail to realize that a simpler null hypothesis could not be rejected, namely that their "discovery" could actually be explained by just a combination of basic statistical laws known to seismologists for decades.

The well-established laws of statistical seismicity that authors should consider before claiming for novelty include the following:

1. The Gutenberg–Richter law for the distribution of earthquake magnitudes with a b-value close to 1 (corresponding to an exponent $\simeq 2/3$ for the probability density function of seismic moments),
2. The Omori–Utsu law for the decay of the rate of aftershocks following a mainshock,
3. The inverse Omori law for foreshocks,
4. The fact that aftershocks also trigger their own aftershocks and so on, and that aftershocks do not seem to exhibit any distinguishable physical properties,
5. The fact that the distribution of distances between mainshocks and aftershocks has a power law tail,
6. The fertility law (the fact that earthquakes of magnitude M trigger of the order of 10^{aM} aftershocks with $a \sim \leq b \simeq 1$,
7. The fractal distribution of faults which are concentration centers for earthquakes.

This above non-exhaustive list selects "laws" which are arguably non-redundant, in the sense that it is likely not possible to derive one of these laws from the others (a possible exception is the inverse Omori law for foreshock, which can be derived from the direct Omori law for aftershocks in the context of the ETAS model [78,81]). Some experts would argue that we should add to this list other claimed regularities, such as "Båth's law" (see e. g. [190] for a recent discussion emphasizing the importance of this law), that states that the differences in magnitudes between mainshocks and their largest aftershocks are approximately constant, independent of the magnitudes of mainshocks [20]. However, Helmstetter and Sornette [79] and Saichev and Sornette [173] have shown that Bath's law can be accurately recovered in ETAS-type models combining the first, second, fourth, and sixth laws stated above, with the assumption that any earthquake can trigger subsequent earthquakes.

Entropy Method for the Distribution of Time Intervals Between Mainshocks

Mega et al. [140] used the "diffusion entropy" method to argue for a power-law distribution of time intervals between a large earthquake (the mainshock of a seismic sequence or cluster) and the next one. Helmstetter and Sornette [80] showed that all the "new" discoveries reported by [140] (including the supposedly new scaling) can be explained solely by Omori's law for intra-cluster times, without correlation between clusters, thus debunking the claim for novelty.

Scaling of the PDF of Waiting Times

Bak et al. [14] analyzed the scaling of the probability density function of waiting times between successive earthquakes in southern California as a function of "box size" or small regions in which subsequent earthquakes are considered. They found an approximate collapse of the pdfs for different seismic moment thresholds S and box sizes L which suggested the following scaling ansatz for the waiting times T:

$$T^{\alpha} P_{S,L}(T) = f(TS^{-b}L^{d_f}) \,, \tag{3}$$

where $b = 1$ is the Gutenberg–Richter exponent, $d_f \simeq 1.2$ was claimed to be a spatial fractal dimension of seismicity (see [146] and [115] for more in-depth studies), $\alpha = 1$ was identified as the exponent in the Omori law and $f(\cdot)$ is a scaling function which was proposed to be roughly constant up to a constant ("kink") beyond which it quickly decays. The scaling (3) was claimed to be a unified law for earthquakes that revealed a novel feature in the spatio-temporal organization of seismicity in that the Gutenberg–Richter, the Omori law and the spatial distribution of earthquakes were unified into a single picture that made no distinction between fore-, main- and aftershocks. The scaling relations and critical exponents were claimed to be contained in the scaling ansatz. Corral [40,41,42,43] and others broadened the analysis to other regions of the world. Corral [41] proposed a slightly different scaling ansatz for a modified data analysis.

Early criticism came from Lindman et al. [132], who noted that synthetic data generated using a non-homogeneous Poisson process derived from Omori's law was able to reproduce some of the results of [14], indicating a rather trivial origin of the unified scaling law. Molchan [145] showed that, if at least two regions in the data set are independent, then, if a scaling relation were to hold exactly, this scaling function could only be exponential. All other functions could only result in approximate data collapses.

Proponents of the unified scaling law, e. g. [44], argued that indeed all regions were correlated, as expected in systems near a critical point so that the assumption of independence between different regions should not hold. But Molchan [145] also showed that a simple Poisson cluster model (Poissonian mainshocks that trigger Omori-type aftershock sequences) could reproduce the short and long time limits of the observed statistics, indicating that the Omori law, the Gutenberg–Richter relationship and simple clustering were the sole ingredients necessary for the observed short and long time limit, and no spatial correlation was needed.

Saichev and Sornette [176,177] extended Molchan's arguments to show that the approximate data collapse of the waiting times could be explained completely by the Epidemic-Type Aftershock Sequence (ETAS) model of [149]. This provided further evidence that the apparent data collapse was only approximate. Remarkably, the theoretical predictions of the ETAS model seem to fit the observed data better than the phenomenological scaling function proposed by [41] to fit the data. Saichev and Sornette [176,177] thus showed that a benchmark model of seismicity was able to reproduce the apparent unified scaling law and that therefore the distribution of interevent times did not reveal new information beyond what was already known via statistical laws: The combination of the Gutenberg–Richter law, the Omori law, and the concept of clustering suffice to explain the apparent "universal" scaling of the waiting times.

Sornette et al. [215] developed an efficient numerical scheme to solve accurately the set of nonlinear integral equations derived previously in [177] and found a dramatic lack of power for the distribution of inter-event times to distinguish between quite different sets of parameters, casting doubt on the usefulness of this statistics for the specific purpose of identifying the clustering parameter (e. g. [72]).

Scaling of the PDF of Distances Between Subsequent Earthquakes

Davidson and Paczuski [49] claimed evidence contradicting the theory of aftershock zone scaling in favor of scale-free statistics. Aftershock zone scaling refers to the scaling of the mainshock rupture length, along which most aftershocks occur, with the mainshock magnitude [112]. Davidson and Paczuski [49] suggested that the probability density function of spatial distances between successive earthquakes obeys finite size scaling with a novel dynamical scaling exponent, suggesting that the mainshock rupture length scale has no impact on the spatial distribution of aftershocks and that earthquakes are self-organized critical.

Werner and Sonette [230] debunked this claim by showing that (i) the purported power law scaling function is not universal as it breaks down in other regions of the world; (ii) the results obtained by [49] for southern California depend crucially on a single earthquake (the June 28, 1992, M7.3 Landers earthquake): without Landers and its aftershocks, the power law disappears; (iii) a model of clustered seismicity, with aftershock zone scaling explicitly built in, is able reproduce the apparent power law, indicating that an apparent lack of scales in the data does not necessarily contradict aftershock zone scaling and the existence of scales associated with mainshock rupture length scales.

The Network Approach

The recent boom in the statistical mechanics of network analysis has recently extended to applications well beyond physics (for reviews, see [6,16,57,148]). Earthquake seismology is no exception [1,2,3,4,8,9,10,160,161]. The resulting impact has been limited so far for several reasons.

A major concern is the assumption that earthquake catalogs as downloaded from the web are data sets fit for immediate analysis. References [82,113,229] and [231] present modern and complementary assessments of the many issues of incompleteness spoiling even the best catalogs. In particular, we should stress that magnitude determinations are surprisingly inaccurate, leading to large errors in seismic rate estimates [231]. Furthermore, there is no such thing as a complete catalog above a so-called magnitude of completeness, due to the fact that a non-negligible fraction of earthquakes are missed in the aftermath of previous earthquakes [82,113]. One should be concerned that analyses in terms of network metrics could be particularly sensitive to these defects. Nevertheless, Abe and Suzuki [1,2,3,4] applied metrics of network analysis to "raw" catalogs which included events well below the estimated magnitude of completeness. As a result of neglecting to use a (reasonably) homogeneous and trustworthy data set, the results of their analysis may be severely biased, because the reliability of the inferred network structure is probably more sensitive than other metrics to the correct spatio-temporal ordering of the earthquake catalog. No serious study has yet been performed to quantify the usually serious impact of quality issues on the metrics used in network analysis. As a consequence, it is also not clear how to interpret the "success" of [160,161] in reproducing the "features" of Abe and Suzuki's analysis on the synthetic seismicity generated by a spring-block model.

In addition, at best limited attempts have been made to interpret the results of the new network metrics using well-known, established facts in seismology. Many of the claimed novel features are probably very well understood – they are mostly related to scale-invariance and clustering of seismicity, facts documented for decades. The authors should always strive to show that the new metrics that they propose give results that cannot be explained by the standard laws in statistical seismology. Toward this end, there are well-defined benchmark models that incorporate these laws and that can generate synthetic catalogs on which the new metrics can be tested and compared.

A few exceptions are worth mentioning. Motivated by the long-standing and unresolved debate over "aftershock" identification, Baiesi and Paczuski [9,10] and Baiesi [8] provided a new metric for the correlations between earthquakes based on the space-time-magnitude nearest-neighbor distance between earthquakes. The authors compared their results with known statistical laws in seismology and with the predictions of the ETAS model, actually confirming both. While no new law has been unearthed here, such efforts are valuable to validate known laws and continue to test the possible limits. Zaliapin et al. [238] extended their study and investigated the theoretical properties of the metric and its ability to decluster catalogs (i. e., separate mainshocks from aftershocks). They concluded that aftershocks defined from this metric seem to be different from the rest of earthquakes. It will be interesting to see head-to-head comparisons with current state-of-the-art probabilistic declustering techniques that are based on empirical statistical laws and likelihood estimation [105,240,241].

Future Directions

The study of the statistical physics of earthquakes remains wide-open with many significant discoveries to be made. The promise of a holistic approach – one that emphasizes the interactions between earthquakes and faults – is to be able to neglect some of the exceedingly complicated micro-physics when attempting to understand the large scale patterns of seismicity. The marriage between this conceptual approach, based on the successes of statistical physics, and seismology thus remains a highly important domain of research. In particular, statistical seismology needs to evolve into a genuine physically-based statistical physics of earthquakes.

The question of renormalizability of models of earthquake occurrence and the role of small earthquakes in the organization of seismicity is likely to remain an important topic. It connects with the problem of foreshocks and the predictability of large events from small ones and therefore has real and immediate practical applications as well as physical implications.

More detailed and rigorous empirical studies of the frequency-size statistics of earthquake seismic moments and how they relate to seismo-tectonic conditions are needed in order to help settle the controversy over the power-law versus the characteristic event regime, and the role of regime-switching and universality.

Spatially extended, dynamically evolving fault networks and their role in the generation of earthquakes are mostly ignored in the statistical physics approach to seismicity. Akin to the filaments in turbulence, these may provide key insights into the spatio-temporal organization of earthquakes. Novel methods combining information from seismology to faulting will be required (e. g., [71,159,195,196,197]) to build a real understanding of the self-organization of the chicken-and-egg structures that earthquakes-faults constitute. Furthermore, a true physical approach requires understanding the spatio-temporal evolution of stresses, their role in earthquake nucleation via thermally activated processes, in the rupture propagation and in the physics of arrest, both involved in the generation of complex stress fields.

The important debate regarding statistical physics approaches to seismicity would benefit significantly from two points. Firstly, earthquake catalogs contain data uncertainties, biases and subtle incompleteness issues. Investigating their influence on the results of data analyses inspired by statistical physics increases the relevance of the results. Secondly, the authors should make links with the literature on statistical seismology which deals with similar questions. It is their task to show that the new metrics that they propose give results that cannot be explained by the standard laws in statistical seismology. For this, there are well-defined benchmark models that incorporate these laws and that can generate synthetic catalogs on which the new metrics can be tested.

Bibliography

1. Abe S, Suzuki N (2004) Scale-free network of earthquakes. Europhys Lett 65:581–586. doi:10.1209/epl/i2003-10108-1
2. Abe S, Suzuki N (2004) Small-world structure of earthquake network. Physica A: Stat Mech Appl 337:357–362. doi:10.1016/j.physa.2004.01.059
3. Abe S, Suzuki N (2005) Scale-invariant statistics of period in directed earthquake network. Eur Phys J B 44:115–117. doi:10.1140/epjb/e2005-00106-7
4. Abe S, Suzuki N (2006) Complex earthquake networks: Hierarchical organization and assortative mixing. Phys Rev E 74(2):026, 113–+. doi:10.1103/PhysRevE.74.026113

5. Aki K (1995) Earthquake prediction, societal implications. Rev Geophys 33:243–248
6. Albert R, Barabási AL (2002) Statistical mechanics of complex networks, Rev Mod Phys 74(1):47–97. doi:10.1103/RevModPhys.74.47
7. Allègre CJ, Le Mouel JL, Provost A (1982) Scaling rules in rock fracture and possible implications for earthquake prediction. Nature 297:47–49. doi:10.1038/297047a0
8. Baiesi M (2006) Scaling and precursor motifs in earthquake networks. Physica A: Stat Mech Appl 359:775–783. doi:10.1016/j.physa.2005.05.094
9. Baiesi M, Paczuski M (2004) Scale-free networks of earthquakes and aftershocks. Phys Rev E 69(6):066, 106. doi:10.1103/PhysRevE.69.066106
10. Baiesi M, Paczuski M (2005) Complex networks of earthquakes and aftershocks. Nonlin Proc Geophys 12:1–11
11. Bak P (1996) How Nature Works: The Science of Self-Organized Criticality. Springer, New York, p 212
12. Bak P, Tang C (1989) Earthquakes as a self-organized critical phenomena. J Geophys Res 94(B11):15635–15637
13. Bak P, Tang C, Wiesenfeld K (1987) Self-organized criticality: An explanation of the 1/*f* noise. Phys Rev Lett 59(4):381–384. doi:10.1103/PhysRevLett.59.381
14. Bak P, Christensen K, Danon L, Scanlon T (2002) Unified scaling law for earthquakes. Phys Rev Lett 88(17):178,501. doi:10.1103/PhysRevLett.88.178501
15. Bakun, WH, Aagaard B, Dost B, Ellsworth WL, Hardebeck JL, Harris RA, Ji C, Johnston MJS, Langbein J, Lienkaemper JJ, Michael AJ, Murray JR, Nadeau RM, Reasenberg PA, Reichle MS, Roeloffs EA, Shakal A, Simpson RW, Waldhauser F (2005) Implications for prediction and hazard assessment from the 2004 Parkfield earthquake. Nature 437:969–974. doi:10.1038/nature04067
16. Barabási AL, Albert R (1999) Emergence of Scaling in Random Networks. Science 286(5439):509–512. doi:10.1126/science.286.5439.509
17. Barabási AL, Albert R, Jeong H (1999) Mean-field theory fore scale-free random networks. Physica A 272:173–187. doi:10.1016/S0378-4371(99)00291-5
18. Barton CC, La Pointe PR (eds) (1995) Fractals in the Earth Sciences. Plenum Press, New York, London
19. Barton CC, La Pointe PR (eds) (1995) Fractals in petroleum geology and earth processes. Plenum Press, New York, London
20. Båth M (1965) Lateral inhomogeneities in the upper mantle. Tectonophysics 2:483–514
21. Ben-Zion Y, Rice JR (1993) Earthquake failure sequences along a cellular fault zone in a 3-dimensional elastic solid containing asperity and nonasperity regions. J Geophys Res 93:14109–14131
22. Ben-Zion Y, Rice JR (1995) Slip patterns and earthquake populations along different classes of faults in elastic solids. J Geophys Res 100:12959–12983
23. Ben-Zion Y, Rice JR (1997) Dynamic simulations of slip on a smooth fault in an elastic solid. J Geophys Res 102:17771–17784
24. Ben-Zion Y, Dahmen K, Lyakhovsky V, Ertas D, Agnon A (1999) Self-driven mode switching of earthquake activity on a fault system. Earth Planet Sci Lett 172:11–21
25. Bird P, Kagan YY (2004) Plate-tectonic analysis of shallow seismicity: Apparent boundary width, beta, corner magnitude, coupled lithosphere thickness, and coupling in seven tectonic settings. Bull Seismol Soc Am 94(6):2380–2399
26. Bowman DD, King GCP (2001) Stress transfer and seismicity changes before large earthquakes. C Royal Acad Sci Paris, Sci Terre Planetes 333:591–599
27. Bowman DD, King GCP (2001) Accelerating seismicity and stress accumulation before large earthquakes. Geophys Res Lett 28:4039–4042
28. Bowman DD, Oullion G, Sammis CG, Sornette A, Sornette D (1998) An observational test of the critical earthquake concept. J Geophys Res 103:24359–24372
29. Brehm DJ, Braile LW (1998) Intermediate-term earthquake prediction using precursory events in the New Madrid Seismic Zone. Bull Seismol Am Soc 88(2):564–580
30. Bufe CG, Varnes DJ (1993) Predictive modeling of the seismic cycle of the greater San Francisco Bay region. J Geophys Res 98:9871–9883
31. Burridge R, Knopoff L (1964) Body force equivalents for seismic dislocation. Seism Soc Am Bull 54:1875–1888
32. Cardy JL (1996) Scaling and Renormalization in Statistical Physics. Cambridge University Press, Cambridge
33. Carlson JM, Langer JS (1989) Properties of earthquakes generated by fault dynamics. Phys Rev Lett 62:2632–2635
34. Carlson JM, Langer JS, Shaw BE (1994) Dynamics of earthquake faults. Rev Mod Phys 66:657–670
35. Chelidze TL (1982) Percolation and fracture. Phys Earth Planet Interiors 28:93–101
36. Christensen K, Farid N, Pruessner G, Stapleton M (2008) On the finite-size scaling of probability density functions. Eur Phys B 62:331–336
37. Clauset A, Shalizi CR, Newman MEJ (2007) Power-law distributions in empirical data. E-print arXiv:0706.1062
38. Cochard A, Madariaga R (1994) Dynamic faulting under rate-dependent friction. Pure Appl Geophys 142:419–445
39. Cochard A, Madariaga R (1996) Complexity of seismicity due to highly rate-dependent friction. J Geophys Res 101:25321–25336
40. Corral A (2003) Local distributions and rate fluctuations in a unified scaling law for earthquakes. Phys Rev E 68(3):035, 102. doi:10.1103/PhysRevE.68.035102
41. Corral A (2004) Universal local versus unified global scaling laws in the statistics of seismicity. Physica A 340:590–597
42. Corral A (2004) Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. Phys Rev Lett 92:108, 501
43. Corral A (2005) Mixing of rescaled data and bayesian inference for earthquake recurrence times. Nonlin Proc Geophys 12:89–100
44. Corral A (2005) Renormalization-group transformations and correlations of seismicity. Phys Rev Lett 95:028, 501
45. Corral A (2006) Universal earthquake-occurrence jumps, correlations with time, and anomalous diffusion. Phys Rev Lett 97:178, 501
46. Corral A, Christensen K (2006) Comment on "earthquakes descaled: On waiting time distributions and scaling laws". Phys Rev Lett 96:109, 801
47. Dahmen K, Ertaş D, Ben-Zion Y (1998) Gutenberg–Richter and characteristic earthquake behavior in simple mean-field models of heterogeneous faults. Phys Rev E 58:1494–1501. doi:10.1103/PhysRevE.58.1494

48. Davidsen J, Goltz C (2004) Are seismic waiting time distributions universal? Geophys Res Lett 31:L21612. doi:10.1029/2004GL020892
49. Davidsen J, Paczuski M (2005) Analysis of the spatial distribution between successive earthquakes. Phys Rev Lett 94:048, 501. doi:10.1103/PhysRevLett.94.048501
50. Davidsen J, Grassberger P, Paczuski M (2006) Earthquake recurrence as a record breaking process. Geophys Res Lett 33:L11304. doi:10.1029/2006GL026122
51. Dhar D (1990) Self-organized critical state of sandpile automaton models. Phys Rev Lett 64:1613–1616
52. Dhar D (1999) The Abelian sandpile and related models. Physica A 263:4–25
53. Dieterich JH (1987) Nucleation and triggering of earthquake slip; effect of periodic stresses. Tectonophysics 144:127–139
54. Dieterich JH (1992) Earthquake nucleation on faults with rate-dependent and state-dependent strength. Tectonophysics 211:115–134
55. Dieterich J (1994) A constitutive law for rate of earthquake production and its application to earthquake clustering. J Geophys Res 99:2601–2618
56. Dieterich J, Kilgore BD (1994) Direct observation of frictional constacts- New insight for state-dependent properties. Pure Appl Geophys 143:283–302
57. Dorogevtsev SN, Mendes JFF (2003) Evolution of Networks: From Biological Nets to the Internet and WWW. Oxford University Press, New York
58. Eckman JP (1981) Roads to Turbulence in Dissipative Dynamical Systems. Rev Mod Phys 53:643–654
59. Ellsworth WL, Lindh AG, Prescott WH, Herd DJ (1981) The 1906 San Francisco Earthquake and the seismic cycle. Am Geophys Union Maurice Ewing Monogr 4:126–140
60. Felzer KR, Becker TW, Abercrombie RE, Ekstrom G, Rice JR (2002) Triggering of the 1999 Mw 7.1 Hector Mine earthquake by aftershocks of the 1992 Mw 7.3 Landers earthquake. J Geophys Res 107(B09):2190
61. Fisher DS, Dahmen K, Ramanathan S, Ben-Zion Y (1997) Statistics of Earthquakes in Simple Models of Heterogeneous Faults. Phys Rev Lett 78:4885–4888. doi:10.1103/PhysRevLett.78.4885
62. Freund F, Sornette D (2007) Electro-Magnetic Earthquake Bursts and Critical Rupture of Peroxy Bond Networks in Rocks. Tectonophysics 431:33–47
63. Frisch U (1995) Turbulence. The legacy of A.N. Kolmogorov. Cambridge University Press, Cambridge
64. Gabrielov A, Keilis-Borok V, Jackson DD (1996) Geometric Incompatibility in a Fault System. Proc Nat Acad Sci 93:3838–3842
65. Gabrielov A, Keilis-Borok V, Zaliapin I, Newman W (2000) Critical transitions in colliding cascades. Phys Rev E 62:237–249
66. Gabrielov A, Zaliapin I, Newman W, Keilis-Borok V, (2000) Colliding cascades model for earthquake prediction. Geophys J Int 143:427–437
67. Gallagher R, Appenzeller T (1999) Beyond Reductionism. Science 284(5411):79
68. Geilikman MB, Pisarenko VF, Golubeva TV (1990) Multifractal Patterns of Seismicity. Earth Planet Sci Lett 99:127–138
69. Gelfand IM, Guberman SA, Keilis-Borok VI, Knopoff L, Press F, Ranzman EY, Rotwain IM, Sadovsky AM (1976) Pattern recognition applied to earthquake epicenters in California. Phys Earth Planet Interiors 11:227–283
70. Geller RJ, Jackson DD, Kagan YY, Mulargia F (1997) Earthquakes cannot be predicted. Science 275:1616–1617
71. Gorshkov A, Kossobokov V, Soloviev A (2003) Recognition of earthquake-prone areas. In: Keilis-Borok V, Soloviev A (eds) Nonlinear Dynamics of the Lithosphere and Earthquake Prediction. Springer, Heidelberg, pp 239–310 [122]
72. Hainzl S, Scherbaum F, Beauval C (2006) Estimating Background Activity Based on Interevent-Time Distribution. Bull Seismol Soc Am 96(1):313–320. doi:10.1785/0120050053
73. Hanks TC (1992) Small earthquakes, tectonic forces. Science 256:1430–1432
74. Hardebeck JL, Felzer KR, Michael AJ (2008) Improved tests reveal that the accelerating moment release hypothesis is statistically insignificant. J Geophys Res 113:B08310. doi:10.1029/2007JB005410
75. Harris RA, Arrowsmith JR (2006) Introduction to the Special Issue on the 2004 Parkfield Earthquake and the Parkfield Earthquake Prediction Experiment. Bull Seismol Soc Am 96(4B):S1–10. doi:10.1785/0120050831
76. Helmstetter A (2003) Is earthquake triggering driven by small earthquakes? Phys Rev Lett 91(5):058, 501. doi:10.1103/PhysRevLett.91.058501
77. Helmstetter A, Sornette D (2002) Subcritical and supercritical regimes in epidemic models of earthquake aftershocks. J Geophys Res 107(B10):2237. doi:10.1029/2001JB001580
78. Helmstetter A, Sornette D (2003) Foreshocks explained by cascades of triggered seismicity. J Geophys Res (Solid Earth) 108(B10):2457 doi:10.1029/2003JB00240901
79. Helmstetter A, Sornette D (2003) Bath's law Derived from the Gutenberg–Richter law and from Aftershock Properties. Geophys Res Lett 30:2069. doi:10.1029/2003GL018186
80. Helmstetter A, Sornette D (2004) Comment on "Power-Law Time Distribution of Large Earthquakes". Phys Rev Lett 92:129801 (Reply is Phys Rev Lett 92:129802 (2004))
81. Helmstetter A, Sornette D, Grasso J-R (2003) Mainshocks are Aftershocks of Conditional Foreshocks: How do foreshock statistical properties emerge from aftershock laws. J Geophys Res 108(B10):2046. doi:10.1029/2002JB001991
82. Helmstetter A, Kagan YY, Jackson DD (2005) Importance of small earthquakes for stress transfers and earthquake triggering. J Geophys Res 110:B05508. doi:10.1029/2004JB003286
83. Helmstetter A, Kagan Y, Jackson D (2006) Comparison of short-term and long-term earthquake forecast models for Southern California. Bull Seism Soc Am 96:90–106
84. Hergarten S (2002) Self-Organized Criticality in Earth Systems. Springer, Berlin
85. Hinrichsen H (2000) Non-equilibrium critical phenomena and phase transitions into absorbing states. Adv Phys 49:815–958(144)
86. Holliday JR, Rundle JB, Tiampo KF, Klein W, Donnellan A (2006) Systematic procedural and sensitivity analysis of the Pattern Informatics method for forecasting large ($M > 5$) earthquake events in Southern California. Pure Appl Geophys 163(11–12):2433–2454
87. Huang J, Turcotte DL (1990) Evidence for chaotic fault interactions in the seismicity of the San Andreas fault and Nankai trough. Nature 348:234–236
88. Huang J, Turcotte DL (1990) Are earthquakes an example of deterministic chaos? Geophys Rev Lett 17:223–226

89. Huang Y, Saleur H, Sammis CG, Sornette D (1998) Precursors, aftershocks, criticality and self-organized criticality. Europhys Lett 41:43–48
90. Ide K, Sornette D (2002) Oscillatory Finite-Time Singularities in Finance, Population and Rupture. Physica A307(1–2):63–106
91. Jackson DD, Kagan YY (2006) The 2004 Parkfield Earthquake, the 1985 Prediction, and Characteristic Earthquakes: Lessons for the Future. Bull Seismol Soc Am 96(4B):S397–409. doi:10.1785/0120050821
92. Jaumé SC, Sykes LR (1999) Evolving Towards a Critical Point: A Review of Accelerating Seismic Moment/Energy Release Prior to Large and Great Earthquakes. Pure Appl Geophys 155:279–305
93. Jensen HJ (1998) Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological Systems. Cambridge University Press, Cambridge
94. Johansen A, Sornette D, Wakita G, Tsunogai U, Newman WI, Saleur H (1996) Discrete scaling in earthquake precursory phenomena: evidence in the Kobe earthquake, Japan J Phys I France 6:1391–1402
95. Johansen A, Saleur H, Sornette D (2000) New Evidence of Earthquake Precursory Phenomena in the 17 Jan. 1995 Kobe Earthquake, Japan. Eur Phys J B 15:551–555
96. Jones LM (1994) Foreshocks, aftershocks, and earthquake probabilities: accounting for the Landers earthquake. Bull Seismol Soc Am 84:892–899
97. Jordan TH (2006) Earthquake Predictability, Brick by Brick. Seismol Res Lett 77(1):3–6
98. Kadanoff LP, Nagel SR, Wu L, Zhou S-M (1989) Scaling and universality in avalanches. Phys Rev A 39(12):6524–6537. doi:10.1103/PhysRevA.39.6524
99. Kagan YY (1981), Spatial distribution of earthquakes: The three-point moment function. Geophys J R Astron Soc 67:697–717
100. Kagan YY (1981) Spatial distribution of earthquakes: The four-point moment function. Geophys J Roy Astron Soc 67:719–733
101. Kagan YY (1987) Point sources of elastic deformation: Elementary sources, static displacements. Geophys J R Astron Soc 90:1–34
102. Kagan YY (1987) Point sources of elastic deformation: Elementary sources, dynamic displacements. Geophys J R Astron Soc 91:891–912
103. Kagan YY (1988) Multipole expansions of extended sources of elastic deformation. Geophys J R Astron Soc 93:101–114
104. Kagan YY (1989) Earthquakes and fractals. Ann Rev Mater Sci: Fractal Phenom Disordered Syst 19:520–522
105. Kagan YY (1991) Likelihood analysis of earthquake catalogs. Geophys J Int 106:135–148
106. Kagan YY (1992) Seismicity: Turbulence of solids. Nonlinear Sci Today 2:1–13
107. Kagan YY (1992) On the geometry of an earthquake fault system. Phys Earth Planet Interiors 71:15–35
108. Kagan YY (1993) Statistics of characteristic earthquakes. Bull Seismol Soc Am 83(1):7–24
109. Kagan YY (1994) Observational evidence for earthquakes as a nonlinear dynamic process. Physica D 77:160–192
110. Kagan YY (1994) Comment on "The Gutenberg–Richter or char-acteristic earthquake distribution, which is it?" by Wesnousky. Bull Seismol Soc Am 86:274–285
111. Kagan YY (1999) Is earthquake seismology a hard, quantitative science? Pure Appl Geophys 155:33–258
112. Kagan YY (2002) Aftershock Zone Scaling. Bull Seismol Soc Am 92(2):641–655. doi:10.1785/0120010172
113. Kagan YY (2003) Accuracy of modern global earthquake catalogs. Phys Earth Planet Interiors 135:173–209
114. Kagan YY (2006) Why does theoretical physics fail to explain and predict earthquake occurrence? In: Bhattacharyya P, Chakrabarti BK (eds) Modelling Critical and Catastrophic Phenomena in Geoscience: A Statistical Physics Approach. Lecture Notes in Physics, vol 705. Springer, Berlin, pp 303–359
115. Kagan YY (2007) Earthquake spatial distribution: the correlation dimension. Geophys J Int 168:1175–1194. doi:10.1111/j.1365-246X.2006.03251.x
116. Kagan YY, Knopoff L (1980) Spatial distribution of earthquakes: The two-point correlation function. Geophys J R Astron Soc 62:303–320
117. Kagan YY, Knopoff L (1981) Stochastic synthesis of earthquake catalogs. J Geophys Res 86(B4):2853–2862
118. Kagan YY, Knopoff L (1985) The first-order statistical moment of the seismic moment tensor. Geophys J R Astron Soc 81:429–444
119. Kagan YY, Knopoff L (1985) The two-point correlation function of the seismic moment tensor. Geophys J R Astron Soc 83:637–656
120. Keilis-Borok VI (ed) (1990) Intermediate-term earthquake prediction: models, algorithms, worldwide tests. Phys Earth Planet Interiors 61(1–2)
121. Keilis-Borok VI, Malinovskaya LN (1964) One regularity in the occurrence of strong earthquakes. J Geophys Res B 69:3019–3024
122. Keilis-Borok V, Soloviev A (2003) Nonlinear Dynamics of the Lithosphere and Earthquake Prediction. Springer, Heidelberg
123. Keilis-Borok VI, Knopoff L, Rotwain IM, Allen CR (1988) Intermediate-term prediction of occurrence times of strong earthquakes. Nature 335:690–694
124. King GCP, Bowman DD (2003) The evolution of regional seismicity between large earthquakes. J Geophys Res 108(B2):2096. doi:10.1029/2001JB000783
125. Klein W, Rundle JB, Ferguson CD (1997) Scaling and nucleation in models of earthquake faults. Phys Rev Lett 78:3793–3796
126. Knopoff L (1996) The organization of seismicity on fault networks. Proc Nat Acad Sci USA 93:3830–3837
127. Landau LD, Lifshitz EM (1980) Statistical Physics Course on Theoretical Physics, vol 5, 3rd edn. Butterworth-Heinemann, Oxford
128. Langer JS, Carlson JM, Myers CR, Shaw BE (1996) Slip complexity in dynamical models of earthquake faults. Proc Nat Acad Sci USA 93:3825–3829
129. Lee MW, Sornette D, Knopoff L (1999) Persistence and Quiescence of Seismicity on Fault Systems. Phys Rev Lett 83(N20):4219–4222
130. Levin SZ, Sammis CG, Bowman DD (2006) An observational test of the stress accumulation model based on seismicity preceding the 1992 Landers, CA earthquake. Tectonophysics 413:39–52
131. Lindh AG (1990) The seismic cycle pursued. Nature 348:580–581

132. Lindman M, Jonsdottir K, Roberts R, Lund B, Bdvarsson R (2005) Earthquakes descaled: On waiting time distributions and scaling laws. Phys Rev Lett 94:108, 501
133. Lindman M, Jonsdottir K, Roberts R, Lund B, Bdvarsson R (2006) Reply to comment by A. Corral and K. Christensen. Phys Rev Lett 96:109, 802
134. Livina VN, Havlin S, Bunde A (2006) Memory in the occurrence of earthquakes. Phys Rev Lett 95:208, 501
135. Luebeck S (2004) Universal scaling behavior of non-equilbrium phase transitions. Int J Mod Phys B 18:3977
136. Manna S (1991) Critical exponents of the sandpile models in two dimensions. Physica A179(2):249–268
137. Mandelbrot BB (1982) The Fractal Geometry of Nature. W.H. Freeman, San Francisco
138. Marsan D (2005) The role of small earthquakes in redistributing crustal elastic stress. Geophys J Int 163(1):141–151. doi:10.1111/j.1365-246X.2005.02700.x
139. May RM (1976) Simple mathematical models with very complicated dynamics. Nature 261:459–467
140. Mega MS, Allegrini P, Grigolini P, Latora V, Palatella L, Rapisarda A, Vinciguerra S (2003) Power law time distributions of large earthquakes. Phys Rev Lett 90:18850
141. Michael AJ, Jones LM (1998) Seismicity alert probabilities at Parkfield, California, revisited. Bull Seismol Soc Am 88(1):117–130
142. Miltenberger P, Sornette D, Vanneste C (1993) Fault self-organization as optimal random paths selected by critical spatiotemporal dynamics of earthquakes. Phys Rev Lett 71:3604–3607. doi:10.1103/PhysRevLett.71.3604
143. Mitzenmacher M (2004) A Brief History of Generative Models for Power Law and Lognormal Distributions. Internet Math 1(2):226–251
144. Mogi K (1969) Some features of recent seismic activity in and near Japan 2: activity before and after great earthquakes. Bull Eq Res Inst Tokyo Univ 47:395–417
145. Molchan G (2005) Interevent time distribution in seismicity: A theoretical approach. Pure Appl Geophys 162:1135–1150. doi:10.1007/s00024-004-2664-5
146. Molchan G, Kronrod T (2005) On the spatial scaling of seismicity rate. Geophys J Int 162(3):899–909. doi:10.1111/j.1365-246X.2005.02693.x
147. Nature Debates (1999) Nature debates: Is the reliable prediction of individual earthquakes a realistic scientific goal? available from http://www.nature.com/nature/debates/earthquake/equake_frameset.html
148. Newman MEJ (2003) The structure and function of complex networks. SIAM Rev 45(2):167–256. doi:10.1137/S003614450342480
149. Ogata Y (1988) Statistical models for earthquake occurrence and residual analysis for point processes. J Am Stat Assoc 83:9–27
150. Ogata Y (1998) Space-time point-process models for earthquake occurrences. Ann Inst Stat Math 5(2):379–402
151. Olami Z, Feder HJS, Christensen K (1992) Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes. Phys Rev Lett 68(8):1244–1247
152. Osorio I, Frei MG, Sornette D, Milton J, Lai Y-C (2007) Seizures and earthquakes: Universality and scaling of critical far from equilibrium systems. submitted to Phys Rev Lett. http://arxiv.org/abs/0712.3929
153. Ouillon G, Sornette D (2000) The critical earthquake concept applied to mine rockbursts with time-to-failure analysis. Geophys J Int 143:454–468
154. Ouillon G, Sornette D (2004) Search for Direct Stress Correlation Signatures of the Critical Earthquake Model. Geophys J Int 157:1233–1246
155. Ouillon G, Sornette D (2005) Magnitude-Dependent Omori Law: Theory and Empirical Study. J Geophys Res 110:B04306. doi:10.1029/2004JB003311
156. Ouillon G, Sornette D, Castaing C (1995) Organization of joints and faults from 1 cm to 100 km scales revealed by Optimized Anisotropic Wavelet Coefficient Method and Multifractal analysis. Nonlinear Process Geophys 2:158–177
157. Ouillon G, Castaing C, Sornette D (1996) Hierarchical scaling of faulting. J Geophys Res 101(B3):5477–5487
158. Ouillon G, Ribeiro E, Sornette D (2007) Multifractal Omori Law for Earthquake Triggering: New Tests on the California, Japan and Worldwide Catalogs. submitted to Geophys J Int. http://arxiv.org/abs/physics/0609179
159. Ouillon G, Ducorbier C, Sornette D (2008) Automatic reconstruction of fault networks from seismicity catalogs: Three-dimensional optimal anisotropic dynamic clustering. J Geophys Res 113:B01306. doi:10.1029/2007JB005032
160. Peixoto TP, Prado CP (2004) Distribution of epicenters in the Olami–Feder–Christensen model. Phys Rev E 69(2):025101. doi:10.1103/PhysRevE.69.025101
161. Peixoto TP, Prado CPC (2006) Network of epicenters of the Olami–Feder–Christensen model of earthquakes. Phys Rev E 74(1):016, 126 doi:10.1103/PhysRevE.74.016126
162. Peters O, Neelin JD (2006) Critical phenomena in atmospheric precipitation. Nature Phys 2:393–396. doi:10.1038/nphys314
163. Pruessner G (2004) Studies in self-organized criticality, Ph D thesis, Imperial College London, available from http://www.ma.imperial.ac.uk/%7Epruess/publications/thesis_final/
164. Raleigh CB, Sieh K, Sykes LR, Anderson DL (1982) Forecasting Southern California Earthquakes. Science 217:1097–1104
165. Reynolds PJ, Klein W, Stanley HE (1977) Renormalization Group for Site and Bond Percolation. J Phys C 10:L167–L172
166. Rhoades DA, Evison FF (2004) Long-range earthquake forecasting with every earthquake a precursor according to scale. Pure Appl Geophys 161:47–72
167. Rhoades DA, Evison FF (2005) Test of the EEPAS forecasting model on the Japan earthquake catalogue. Pure Appl Geophys 162:1271–1290
168. Rice JR (1993) Spatio-temporal complexity of slip on a fault. J Geophys Res 98:9885–9907
169. Rundle JB, Klein W (1993) Scaling and critical phenomena in a cellular automaton slider block model for earthquakes. J Stat Phys 72:405–412
170. Rundle JB, Klein W (1995) New ideas about the physics of earthquakes. Rev Geophys 33:283–286
171. Rundle PB, Rundle JB, Tiampo KF, Sa Martins JS, McGinnis S, Klein W (2001) Nonlinear network dynamics on earthquake fault systems. Phys Rev Lett 87(14):148, 501. doi:10.1103/PhysRevLett.87.148501
172. Rundle JB, Turcotte DL, Shcherbakov R, Klein W, Sammis C (2003) Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems. Rev Geophys 41(4):1019

173. Saichev A, Sornette D (2005) Distribution of the Largest Aftershocks in Branching Models of Triggered Seismicity: Theory of the Universal Bath's law. Phys Rev E 71:056127
174. Saichev A, Sornette D (2005) Vere-Jones' self-similar branching model. Phys Rev E 72:056, 122
175. Saichev A, Sornette D (2006) Renormalization of branching models of triggered seismicity from total to observable seismicity. Eur Phys J B 51:443–459
176. Saichev A, Sornette D (2006) "Universal" distribution of interearthquake times explained. Phys Rev Lett 97:078, 501
177. Saichev A, Sornette D (2007). Theory of earthquake recurrence times. J Geophys Res 112:B04313. doi:10.1029/2006JB004536
178. Saleur H, Sammis CG, Sornette D (1996) Renormalization group theory of earthquakes. Nonlinear Process Geophys 3:102–109
179. Saleur H, Sammis CG, Sornette D (1996) Discrete scale invariance, complex fractal dimensions and log-periodic corrections in earthquakes. J Geophys Res 101:17661–17677
180. Sammis SG, Sornette D (2002) Positive Feedback, Memory and the Predictability of Earthquakes. Proc Nat Acad Sci USA V99:SUPP1:2501–2508
181. Scholz CH (1991) Earthquakes and faulting: Self-organized critical phenomena with a characteristic dimension. In: Riste T, Sherrington D (eds) Spontaneous Formation of Space Time Structure and Criticality. Kluwer, Norwell, pp 41–56
182. Scholz CH (2002) The Mechanics of Earthquakes and Faulting, 2nd edn, Cambridge University Press, Cambridge
183. Scholz CH, Mandelbrot BB (eds) (1989) Fractals in Geophysics. Birkhäuser, Basel
184. Schorlemmer D, Wiemer S, Wyss M (2005) Variations in earthquake-size distribution across different stress regimes. Nature 437:539–542. doi:10.1038/nature04094
185. Schwartz DP, Coppersmith KJ (1984) Fault behavior and characteristic earthquakes: examples from the Wasatch and San Andreas Fault Zones. J Geophys Res 89:5681–5698
186. Shaw BE (1993) Generalized Omori law for aftershocks and foreshocks from a simple dynamics. Geophys Res Lett 20:907–910
187. Shaw BE (1994) Complexity in a spatially uniform continuum fault model. Geophys Res Lett 21:1983–1986
188. Shaw BE (1995) Frictional weakening and slip complexity in earthquake faults. J Geophys Res 102:18239–18251
189. Shaw BE (1997) Model quakes in the two-dimensional wave equation. J Geophys Res 100:27367–27377
190. Shcherbakov R, Turcotte DL (2004) A modified form of Bath's law. Bull Seismol Soc Am 94(5):1968–1975
191. Shnirman MG, Blanter EM (1998) Self-organized criticality in a mixed hierarchical system. Phys Rev Lett 81:5445–5448
192. Smalley RF Jr, Turcotte DL, Solla SA (1985) A renormalization group approach to the stick-slip behavior of faults. J Geophys Res 90:1894–1900
193. Sornette A, Sornette D (1989) Self-organized criticality and earthquakes. Europhys Lett 9:197–202
194. Sornette A, Sornette D (1999) Earthquake rupture as a critical point: Consequences for telluric precursors. Tectonophysics 179:327–334
195. Sornette A, Davy P, Sornette D (1990) Growth of fractal fault patterns. Phys Rev Lett 65:2266–2269
196. Sornette A, Davy P, Sornette D (1990) Fault growth in brittle-ductile experiments and the mechanics of continental collisions. J Geophys Res 98:12111–12139
197. Sornette D (1991) Self-organized criticality in plate tectonics. In: Proceedings of the NATO ASI. vol 349, "Spontaneous formation of space-time structures and criticality" Geilo, Norway 2–12 April 1991. Riste T, Sherrington D (eds) Kluwer, Dordrecht, Boston, pp 57–106
198. Sornette D (1992) Critical phase transitions made self-organized: a dynamical system feedback mechanism for self-organized criticality. J Phys I France 2:2065–2073. doi:10.1051/jp1:1992267
199. Sornette D (1998) Discrete scale invariance and complex dimensions. Phys Rep 297(5):239–270
200. Sornette D (1999) Earthquakes: from chemical alteration to mechanical rupture. Phys Rep 313(5):238–292
201. Sornette D (2000) Mechanochemistry: an hypothesis for shallow earthquakes. In: Teisseyre R, Majewski E (eds) Earthquake Thermodynamics and Phase Transformations in the Earth's Interior. Int Geophys Series, vol 76. Cambridge University Press, Cambridge, pp 329–366, e-print at http://xxx.lanl.gov/abs/cond-mat/9807400
202. Sornette D (2002) Predictability of catastrophic events: material rupture, earthquakes, turbulence, financial crashes and human birth. Proc Nat Acad Sci USA 99:2522–2529
203. Sornette D (2004) Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and Tools, 2nd edn. Springer, Berlin, p 529
204. Sornette D, Helmstetter A (2002) Occurrence of Finite-Time-Singularity in Epidemic Models of Rupture, Earthquakes and Starquakes. Phys Rev Lett 89(15):158501
205. Sornette D, Ouillon G (2005) Multifractal Scaling of Thermally-Activated Rupture Processes. Phys Rev Lett 94:038501
206. Sornette D, Pisarenko VF (2003) Fractal Plate Tectonics. Geophys Res Lett 30(3):1105. doi:10.1029/2002GL015043
207. Sornette D, Sammis CG (1995) Complex critical exponents from renormalization group theory of earthquakes: Implications for earthquake predictions. J Phys I France 5:607–619
208. Sornette D, Virieux J (1992) A theory linking large time tectonics and short time deformations of the lithosphere. Nature 357:401–403
209. Sornette D, Werner MJ (2005) Constraints on the size of the smallest triggering earthquake from the epidemic-type aftershock sequence model, Båth's law, and observed aftershock sequences. J Geophys Res 110:B08304. doi:10.1029/2004JB003535
210. Sornette D, Werner MJ (2005) Apparent clustering and apparent background earthquakes biased by undetected seismicity. J Geophys Res 110:B09303. doi:10.1029/2005JB003621
211. Sornette D, Davy P, Sornette A (1990) Structuration of the lithosphere in plate tectonics as a self-organized critical phenomenon. J Geophys Res 95:17353–17361
212. Sornette D, Vanneste C, Sornette A (1991) Dispersion of b-values in Gutenberg–Richter law as a consequence of a proposed fractal nature of continental faulting. Geophys Res Lett 18:897–900
213. Sornette D, Miltenberger P, Vanneste C (1994) Statistical physics of fault patterns self-organized by repeated earthquakes. Pure Appl Geophys 142:491–527. doi:10.1007/BF00876052

214. Sornette D, Miltenberger P, Vanneste C (1995) Statistical physics of fault patterns self-organized by repeated earthquakes: synchronization versus self-organized criticality. In: Bouwknegt P, Fendley P, Minahan J, Nemeschansky D, Pilch K, Saleur H, Warner N (eds) Recent Progresses in Statistical Mechanics and Quantum Field Theory. Proceedings of the conference 'Statistical Mechanics and Quantum Field Theory', USC, Los Angeles, May 16–21, 1994. World Scientific, Singapore, pp 313–332
215. Sornette D, Utkin S, Saichev A (2008) Solution of the Nonlinear Theory and Tests of Earthquake Recurrence Times. Phys Rev E 77:066109
216. Stanley HE (1999) Scaling, universality, and renormalization: Three pillars of modern critical phenomena. Rev Mod Phys 71(2):S358–S366. doi:10.1103/RevModPhys.71.S358
217. Sykes LR, Jaumé S (1990) Seismic activity on neighboring faults as a long-term precursor to large earthquakes in the San Francisco Bay Area. Nature 348:595–599
218. Tiampo KF, Rundle JB, Klein W (2006) Stress shadows determined from a phase dynamical measure of historic seismicity. Pure Appl Geophys 163(11–12):2407–2416
219. Turcotte DL (1986) Fractals and fragmentation. J Geophys Res 91:1921–1926
220. Turcotte DL (1997) Fractals and Chaos in Geology and Geophysics, 2nd edn. Cambridge University Press, Cambridge, p 398
221. Turcotte DL, Newman WI, Gabrielov A (2000) A statistical physics approach to earthquakes. In: Rundle JB, Turcotte DL, Klein W (eds) GeoComplexity and the Physics of Earthquake. American Geophysical Union, Washington, pp 83–96
222. Tumarkin AG, Shnirman MG (1992) Computational seismology 25:63–71
223. Vere-Jones D (1977) Statistical theories of crack propagation. Math Geol 9:455–481
224. Vere-Jones D (2005) A class of self-similar random measure. Adv Appl Probab 37(4):908–914
225. Vere-Jones D (2006) The development of statistical seismology: A personal experience. Tectonophysics 413(1–2):5–12
226. Vere-Jones D, Robinson R, Yang W (2001) Remarks on the accelerated moment release model: problems of model formulation, simulation and estimation. Geophys J Int 144:517–531. doi:10.1046/j.1365-246X.2001.01348.x
227. Voight B (1988) A method for prediction of volcanic eruptions. Nature 332:125–130
228. Voight B (1989) A relation to describe rate-dependent material failure. Science 243:200–203
229. Werner MJ (2007) On the fluctuations of seismicity and uncertainties in earthquake catalogs: Implications and methods for hypothesis testing. Ph D thesis, University of California, Los Angeles
230. Werner MJ, Sornette D (2007) Comment on "Analysis of the Spatial Distribution Between Successive Earthquakes" by Davidsen and Paczuski. [Phys Rev Lett 94:048501 (2005)]. Phys Rev Lett 99::179801
231. Werner MJ, Sornette D (2008) Magnitude Uncertainties Impact Seismic Rate Estimates, Forecasts and Predictability Experiments. J Geophys Res 113:B08302. doi:10.1029/2007JB005427
232. Wesnousky SG (1994) The Gutenberg–Richter or characteristic earthquake distribution, which is it? Bull Seismol Soc Am 84(6):1940–1959
233. Wiemer S, Katsumata K (1999) Spatial variability of seismicity parameters in aftershock zones. J Geophys Res 104:13135–13152. doi:10.1029/1999JB900032
234. Wilson K (1979) Problems in physics with many scales of length. Sci Am 241:140–157
235. Yeomans JM (1992) Statistical Mechanics of Phase Transitions. Oxford University Press Inc, New York
236. Zaliapin I, Keilis-Borok V, Ghil M (2003) A Boolean delay equation model of colliding cascades. Part I: Multiple seismic regimes. J Stat Phys 111:815–837
237. Zaliapin I, Keilis-Borok V, Ghil M (2003) A Boolean delay equation model of colliding cascades. Part II: Prediction of critical transitions. J Stat Phys 111:839–861
238. Zaliapin I, Gabrielov A, Keilis-Borok V, Wong H (2008) Clustering analysis of seismicity and aftershock identification. Phys Rev Lett 101:018501. doi:10.1103/PhysRevLett.101.018501
239. Zee A (2003) Quantum Field Theory in a Nutshell. Princeton University Press, Princeton
240. Zhuang J, Ogata Y, Vere-Jones D (2002) Stochastic declustering of space-time earthquake occurrences. J Am Stat Assoc 97:369–380
241. Zhuang J, Ogata Y, Vere-Jones D (2004) Analyzing earthquake clustering features by using stochastic reconstruction. J Geophys Res 109:B05301. doi:10.1029/2003JB002879
242. Zöller G, Hainzl S (2002) A systematic spatiotemporal test of the critical point hypothesis for large earthquakes. Geophys Rev Lett 29:53–1
243. Zöller G, Hainzl S, Kurths J (2001) Observation of growing correlation length as an indicator for critical point behavior prior to large earthquakes. J Geophys Res 106:2167–2176. doi:10.1029/2000JB900379

Seismic Wave Propagation in Media with Complex Geometries, Simulation of

Heiner Igel[1], Martin Käser[1], Marco Stupazzini[2]
[1] Department of Earth and Environmental Sciences, Ludwig-Maximilians-University, Munich, Germany
[2] Department of Structural Engineering, Politecnico di Milano, Milano, Italy

Article Outline

Glossary
Definition of the Subject
Introduction
The Evolution of Numerical Methods and Grids
3D Wave Propagation on Hexahedral Grids: Soil-Structure Interactions
3D Wave Propagation on Tetrahedral Grids: Application to Volcanology
Local Time Stepping: Δt-Adaptation
Discussion and Future Directions
Acknowledgments
Bibliography

Glossary

Numerical methods Processes in nature are often described by partial differential equations. Finding solutions to those equations is at the heart of many studies aiming at the explanation of observed data. Simulations of realistic physical processes requires generally the use of numerical methods – a special branch of applied mathematics – that approximate the partial differential equations and allows solving them on computers. Examples are the finite-difference, finite-element, or finite-volume methods.

Spectral elements The spectral element method is an extension of the finite element method that makes use of specific basis functions describing the solutions inside each element. These basis functions (e. g., Chebyshev or Legendre polynomials) allow the interpolation of functions exactly at certain collocation points. This is often termed spectral accuracy.

Discontinuous Galerkin method The discontinuous Galerkin method is a flavor of the finite-element method that allows discontinuous behavior of the spatial or temporal fields at the element boundaries. The discontinuities – that might be small in the case of continuous physical fields such as seismic waves – then define so-called Riemann problems that can be handled using the concepts from finite-volume techniques. Therefore, the approximate solution is updated via numerical fluxes across the element boundaries.

Parallel algorithms All modern supercomputers make use of parallel architectures. This means that a large number of processors are performing (different) tasks on different data at the same time. Numerical algorithms need to be adapted to these hardware architectures by using specific programming paradigms (e. g., the message passing interface MPI). The computational efficiency of such algorithms strongly depends on the specific parallel nature of the problem to be solved, and the requirement for inter-processor communication.

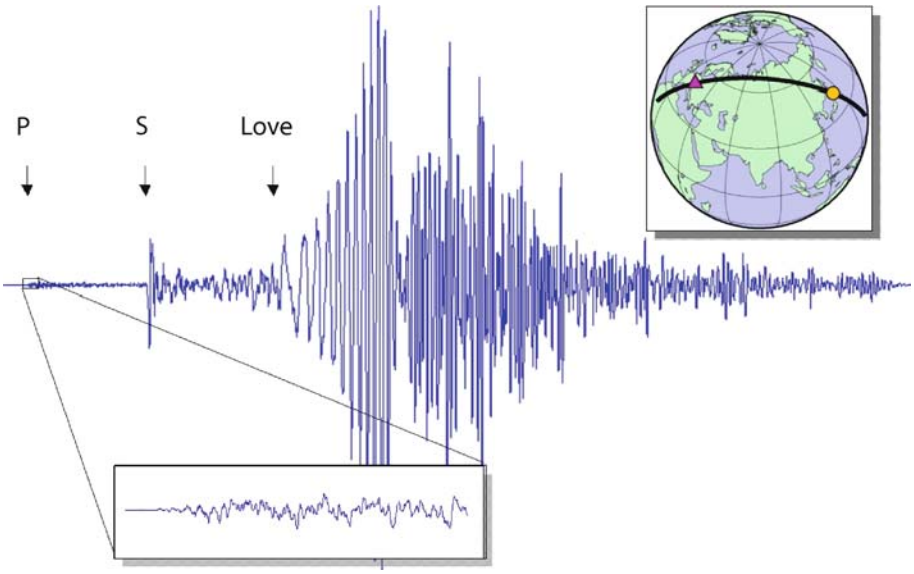
Grid generation Most numerical methods are based on the calculation of the solutions at a large set of points (grids) that are either static or depend on time (adaptive grids). These grids often need to be adapted to the specific geometrical properties of the objects to be modeled (volcano, reservoir, globe). Grids may be designed to follow domain boundaries and internal surfaces. Before specific numerical solvers are employed the grid points are usually connected to form triangles or rectangles in 2D or hexahedra or tetrahedra in 3D.

Definition of the Subject

Seismology is the science that aims at understanding the Earth's interior and its seismic sources from measurements of vibrations of the solid Earth. The resulting images of the physical properties of internal structures and the spatio-temporal behavior of earthquake rupture processes are prerequisites to understanding the dynamic evolution of our planet and the physics of earthquakes. One of the key ingredients to obtain these images is the calculation of synthetic (or theoretical) seismograms for given earthquake sources and internal structures. These synthetic seismograms can then be compared quantitatively with observations and acceptable models be searched for using the theory of inverse problems. The methodologies to calculate synthetic seismograms have evolved dramatically over the past decades in parallel with the evolution of computational resources and the ever increasing volumes of permanent seismic observations in global and regional seismic networks, volcano monitoring networks, and experimental campaigns. Today it is a tremendous challenge to extract an optimal amount of information from seismograms. The imaging process is still primarily carried out using ray theory or extensions thereof not fully taking into account the complex scattering processes that are occurring in nature.

To model seismic observations in their full complexity we need to be able to simulate wave propagation through 3D structures with constitutive relations that account for anisotropic elasticity, attenuation, porous media as well as complex internal interfaces such as layer boundaries or fault systems. This implies that numerical methods have to be employed that solve the underlying partial differential equations on computational grids. The high-frequency oscillatory nature of seismic wave fields makes this an expensive endeavor as far as computational resources are concerned. As seismic waves are propagating hundreds of wavelengths through scattering media, the required accuracy of the numerical approximations has to be of the highest possible order. Despite the fact that the physics of wave propagation is well understood, only recently computational algorithms are becoming available that allow us to accurately simulate wave propagation on many scales such as reservoirs, volcanoes, sedimentary basins, continents, and whole planets.

In addition to the imaging problem for subsurface structure and earthquake sources, the possibilities for 3D wave simulations have opened a new route to forecasting strong ground motions following large earthquakes in seismically active regions. In the absence of any hope to deterministically predict earthquakes, the calculation



Seismic Wave Propagation in Media with Complex Geometries, Simulation of, Figure 1
Transverse velocity seismogram of the M8.3 Tokachi-Oki earthquake near Hokkaido observed at station WET in Germany with a broadband seismometer. The total seismogram length is one hour. Arrival times of body wave phases (P, S) and the onset of transversely polarized surface (Love) waves are indicated

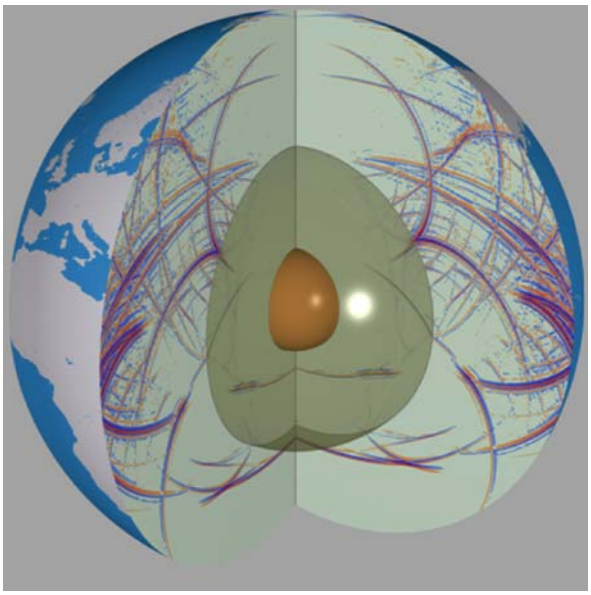
of earthquake scenarios in regions with sufficiently well known crustal structures and fault locations will play an important role in mitigating damage particularly due to potentially amplifying local velocity structures. However, to be able to employ the advanced 3D simulation technology in an efficient way, and to make use of the fast advance of supercomputing infrastructure, a paradigm shift in the concept of wave simulation software is necessary: The Earth science community has to build soft infrastructures that enable massive use of those simulation tools on the available high-performance computing infrastructure.

In this paper we want to present the state of the art of computational wave propagation and point to necessary developments in the coming years, particularly in connection with finding efficient ways to generate computational grids for models with complex topography, faults, and the combined simulation of soil and structures.

Introduction

We first illustrate the evolution of methodologies to calculate and model aspects of seismic observations for the case of global wave propagation. Seismology can look back at almost 50 years of systematic observations of earthquake induced teleseismic ground motions with the standardized global seismic and regional networks. The digital revolution in the past decades has altered the recording culture such that now seismometers are recording ground motions permanently rather than in trigger-mode, observations are becoming available in near-real time, and – because of the required sampling rates – the daily amount of observations automatically sent to the data centers is gigantic. If we take a qualitative look at a seismic observation (Fig. 1) we can illustrate what it takes to model either part or the whole information contained in such physical measurements.

In Fig. 1 a seismogram observed using a broadband seismometer (station WET in Germany) is shown. Globally observed seismograms following large earthquakes contain frequencies up to 1 Hz (P-wave motions) down to periods of around one hour (eigenmodes of the Earth) in which case modeling is carried out in the frequency domain. Seismograms of the kind shown in Fig. 1 contain many types of information. For large earthquakes the first part of the seismogram (inlet) contains valuable information on the spatio-temporal evolution of the earthquake rupture on a finite-size fault. A model of the fault slip history is a prerequisite to model the complete wave form of seismograms as the whole seismogram is affected by it unless severe low-pass filtering is applied. Information on the global seismic velocity structure is contained in the arrival times of numerous body-wave phases (here only P- and S-wave arrivals are indicated) and in the dispersive behavior of the surface waves (here the onset of the low-frequency Love waves is indicated). Further information is contained in the characteristics of the coda to body wave phases in-



Seismic Wave Propagation in Media with Complex Geometries, Simulation of, Figure 2

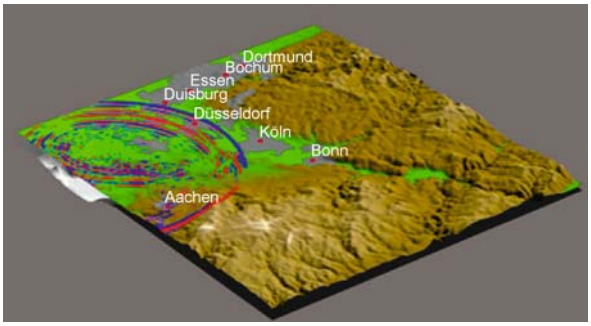
Snapshot of wave propagation inside the Earth approx. 25 minutes after an earthquake occurs at the top part of the model. The radial component of motion is shown (blue and red denote positive and negative velocity, resp.). The simulation was carried out using an axi-symmetric approximation to the wave equation [55,58] and high-order finite-differences. Motion is allowed in the radial and horizontal directions. This corresponds to the P-SV case in 2D cartesian calculations. Therefore the wavefield contains both P- and S-waves and phase conversions

dicative of scattering in various parts of the Earth (see [62] for an account of modern observational seismology).

Adding a temporal and spatial scale to the above qualitative discussion reveals some important insight what it takes to simulate wave propagation on a planetary scale using grid-based numerical methods. Given the maximum frequency of around 1 Hz (P-waves) and 0.2 Hz (S-waves) the minimum wavelength in the Earth is expected to be O(km), requiring O(100 m) type grid spacing at least in the crustal part of the Earth leading to $O(10^{12})$ necessary grid points (or volume elements) for accurate numerical simulations. This would lead to memory requirements O(100 TByte) that are today possible on some of the world's largest supercomputers. The message here is that despite the rapid evolution of computational power, the complete modeling of teleseismic observations using approaches such as spectral elements (e. g., [63,64]) requiring tremendous numbers of calculations to constrain structure and sources will remain a grand challenge for some time to come. However, in many cases it is not necessary or not even desirable to simulate or model the whole seismogram, i. e. the complete observed frequency band. If we lower the cutoff frequency to 0.1 Hz (period 10 s), the required memory drops down to O(100 GByte). Such calculations can be done today on PC-clusters that can be inexpensively assembled and run on an institutional level (e. g., [8]). In addition, it means that the massive use of such forward simulations for imaging purposes and phenomenological investigations of wavefield effects is around the corner. This does not only apply to wave propagation or imaging on a planetary scale but in the same way to problems in volcanology, regional seismology, and exploration geophysics.

An illustration of global wave simulations using the finite difference method (e. g., [14,54,55,58,109,110,114]) is shown in Fig. 2 (more details on the methodologies are given in Sect. "The Evolution of Numerical Methods and Grids"). The snapshot of the radial component of motion at a time when the direct P-wave has almost crossed the Earth reveals the tremendous complexity the wave field exhibits even in the case of a spherically symmetric Earth model (PREM, [37]). The wavefield with a dominant period of ca. 15 seconds also highlights the short wavelengths that need to be propagated over very large distances. This is the special requirement for computational wave propagation that is quite different in other fields of computational Earth Sciences. While the theory of linear elastic wave propagation is well understood and most numerical methods have been applied to it in various forms, the accuracy requirements are so high that – particularly when models with complex geometrical features need to be modeled – there are still open questions as to what works best. One of the main goals of this paper is to highlight the need to focus on the grid generation process for various types of computational grid cells (e. g., rectangular, triangular in 2D, and hexahedral and tetrahedral in 3D) and the interface to appropriate highly accurate solvers for wave propagation problems.

As mentioned above computational modeling of strong ground motions following large earthquakes (see Fig. 3 for an illustration) is expected to play an increasingly important role in producing realistic estimates of shaking hazard. There are several problems that are currently unsolved: (1) to achieve frequencies that are interesting for earthquake engineers in connection with structural damage the near surface velocity structure needs to be known and frequencies beyond 5 Hz need to be calculated. In most cases this structure is not well known (on top of the uncertainties of the lower basin structures) and the required frequencies demand extremely large computational models. (2) In addition to structural uncertain-



Seismic Wave Propagation in Media with Complex Geometries, Simulation of, Figure 3

Snapshot (horizontal component) for a simulation of the M5.9 Roermond earthquake in the Cologne Basin in 1992 [38]. The 3D sedimentary basin (maximum depth 2 km) leads to strong amplification and prolongation of the shaking duration that correlates well with basin depth. Systematic calculations may help mitigating earthquake induced damage

ties, there are strong dependencies on the particular earthquake rupture process that influence the observed ground motions. This suggests that many 3D calculations should be carried out for any characteristic earthquake of interest, to account for such variations (e. g., Wang et al. 2006). (3) The large velocity variations (e. g., 300 m/s up to 8 km/s) require locally varying grid densities which is difficult to achieve with some of the classical numerical methods in use (e. g. finite differences). Some of the potential routes are developed below.

In summary, computational simulation of 3D wave propagation will be more and more a central tool for seismology with application in imaging problems, earthquake rupture problems, questions of shaking hazard, volcano seismology and planetary seismology. In the following we briefly review the history of the application of numerical methods to wave propagation problems and the evolution of computational grids. The increasing complexity of models in terms of geometrical features and range of physical properties imposes the use of novel methodologies that go far beyond the initial approximations based on finite differences.

The Evolution of Numerical Methods and Grids

In this section we give a brief history of the application of numerical methods to the problem of seismic wave propagation. Such a review can not be complete, certainly gives a limited perspective, and only some key references are given. One of the points we would like to highlight is the evolution of the computational grids that are being employed for wave propagation problems and the consequences on the numerical methods of choice now and in the future.

Why do we need numerical approximations to elastic wave propagation problems at all? It is remarkable what we learned about the Earth without them! In the first decades in seismology, modeling of seismic observations was restricted to the calculation of ray-theoretical travel times in spherically symmetric Earth models (e. g., [13,16]). With the advent of computing machines these approaches could be extended to 2D and 3D media leading to ray-theoretical tomography and the images of the Earth's interior that we know today (e. g., [115]). The analytical solution of wave propagation in spherical coordinates naturally leads to spherical harmonics and the possible quasi-analytical solution of wave propagation problems in spherically symmetric media using normal modes. As this methodology leads to complete waveforms the term "waveform inversion" was coined for fitting the waveforms of surface waves by correcting the phase differences for surface waves at particular frequencies (e. g., [118]). This allowed the recovery of seismic velocity models particularly of crust and upper mantle (surface wave tomography). A similar approach in Cartesian layered geometry led to complete solutions of the wave equation in cylindrical coordinates through the summation of Bessel functions, the reflectivity method [46]. This method was later extended to spherical media through the Earth-flattening transformation [85]. Recently, ray-theory was extended allowing the incorporation of finite-frequency effects (e. g., [84]). The impact on the imaging process is still being debated.

Most of these methods are still today extremely valuable in providing first estimates of 2D or 3D effects and are important for the use in standard seismic processing due to their computational efficiency. Nevertheless, with the tremendous improvements of the quality of seismic observations we strive today to extract much more information on Earth's structure and sources from recorded waveforms. As waveforms are in most places strongly affected by 3D structural variations the application of numerical methods that solve "directly" the partial differential equations descriptive of wave propagation becomes mandatory. This necessity was recognized early on and the developments of numerical wave propagation began in the sixties of the 20th century.

Numerical Methods Applied to Wave Propagation Problems

The finite-difference technique was the first numerical method to be intensively applied to the wave propagation

problem (e. g., [1,6,61,77,82,83,88,89,116,117]). The partial differentials in the wave equation are replaced by finite differences leading to an extrapolation scheme in time that can either be implicit or explicit. The analysis of such simple numerical schemes led to concepts that are central to basically all numerical solutions of wave propagation problems. First, the discretization in space and time introduces a scale into the problem with the consequence that the numerical scheme becomes dispersive. This numerical dispersion – for the originally non-dispersive problem of purely elastic wave propagation – has the consequence that for long propagation distances wave pulses are no longer stable but disperse. The consequence is, that in any simulation one has to ascertain that enough grid points per wavelength are employed so that numerical dispersion is reduced sufficiently. Finding numerical schemes that minimize these effects has been at the heart of any new methodology ever since. Second, the so-called CFL criterion [24] that follows from the same theoretical analysis of the numerical scheme basically relates a "grid velocity" – the ratio between the space and time increments dx and dt, respectively – to the largest physical velocity c in the model. In order to have a stable calculation, this ratio has to be smaller than a constant ε that depends on the specific scheme and the space dimension

$$c\frac{\mathrm{d}t}{\mathrm{d}x} \leq \varepsilon\,. \tag{1}$$

This simple relationship has important consequences: When the grid spacing dx must be small, because of model areas with low seismic velocities, then the time step dt has to be made smaller accordingly leading to an overall increase in the number of time steps and thus overall computational requirements. In addition, the early implementations where based on regular rectangular grids, implying that large parts of the model where carrying out unnecessary calculations. As shown below local time-stepping and local accuracy are important ingredients in efficient modern algorithms.

The fairly inaccurate low order spatial finite-difference schemes were later extended to high-order operators [26,48,49,50,51,56,76,103]. Nevertheless, the required number of grid points per wavelength was still large, particularly for long propagation distances. This has led to the introduction of pseudo-spectral schemes, "pseudo" because only the calculations of the derivatives where done in the spectral domain, but the wave equation was still solved in the space-time domain with a time-extrapolation scheme based on finite differences [10,45,47,67]). The advantage of the calculation of derivatives in the spectral domain is at hand: The Fourier theorem tells us that by multiplying the spectrum with ik, i being the imaginary unit and k the wavenumber, we obtain an *exact* derivative (exact to numerical precision) on a regular set of grid points. This sounds attractive. However, there are always two sides to the coin. The calculation requires FFTs to be carried out extensively and the original "local" scheme becomes a "global" scheme. This implies that the derivative at a particular point in the computational grid becomes dependent on any other point in the grid. This turns out to be computationally inefficient, in particular on parallel hardware. In addition, the Fourier approximations imply periodicity which makes the implementation of boundary conditions (like the free surface, or absorbing boundary conditions) difficult.

By replacing the basis functions (Fourier series) in the classical pseudo-spectral method with Chebyshev polynomials that are defined in a limited domain (−1,1) the problem with the implementation of boundary problems found an elegant solution (e. g., [66,107,108]). However, through the irregular spacing of the Chebyshev collocation points (grid densification at the domain boundaries, see section below) new problems arose with the consequence that this approach was not much further pursued except in combination with a multi-domain approach in which the field variables exchange their values at the domain boundaries (e. g., [108]).

So far, the numerical solutions described are all based on the *strong* form of the wave equation. The finite-element method is another main scheme that found immediate applications to wave propagation problems (e. g., [79]). Finite element schemes are based on solving the *weak* form of the wave equation. This implies that the space- and time-dependent fields are replaced by weighted sums of basis (also called trial) functions defined inside elements. The main advantage of finite element schemes is that elements can have arbitrary shape (e. g., triangular, trapezoidal, hexahedral, tetrahedral, etc.). Depending on the polynomial order chosen inside the elements the spatial accuracy can be as desired. The time-extrapolation schemes are usually based on standard finite differences. There are several reasons why finite-element schemes were less widely used in the field of wave propagation. First, in the process a large system matrix needs to be assembled and must be inverted. Matrix inversion in principle requires global communication and is therefore not optimal on parallel hardware. Second, in comparison with the finite-element method, finite- difference schemes are more easily coded and implemented due to their algorithmic simplicity.

A tremendous step forward was the introduction of basis functions inside the elements that have spectral accuracy, e. g., Chebyshev or Legendre polynomials [11,15,39,40,65,86,90,98]. The so-called spectral element scheme became particularly attractive with the discovery that – by using Legendre polynomials – the matrices that required inversion became diagonal [65]. This implies that the scheme does no longer need global communication, it is a local scheme in which extrapolation to the next time step can be naturally parallelized. With the extension of this scheme to spherical grids using the cubed-sphere discretization [63,64] this scheme is today the method of choice on many scales unless highly complex models need to be initiated.

Most numerical schemes for wave propagation problems were based on regular, regular stretched, or hexahedral grids. The numerical solution to unstructured grids had much less attention, despite the fact that highly complex models with large structural heterogeneities seem to be more readily described with unstructured point clouds. Attempts were made to apply finite volume schemes to this problem [31], and other concepts (like natural neighbor coordinates [7] to find numerical operators that are applicable on unstructured grids [72,73,78]). These approaches were unfortunately not accurate enough to be relevant for 3D problems. Recently, a new flavor of numerical method found application to wave propagation on triangular or tetrahedral grids. This combination of a discontinuous Galerkin method with ideas from finite volume schemes [33,70] allows for the first time arbitrary accuracy in space and time on unstructured grids. While the numerical solution on tetrahedral grids remains computationally slower, there is a tremendous advantage in generating computational grids for complex Earth models. Details on this novel scheme are given below.

Before presenting two schemes (spectral elements and the discontinuous Galerkin method) and some applications in more detail we want to review the evolution of grids used in wave propagation problems.

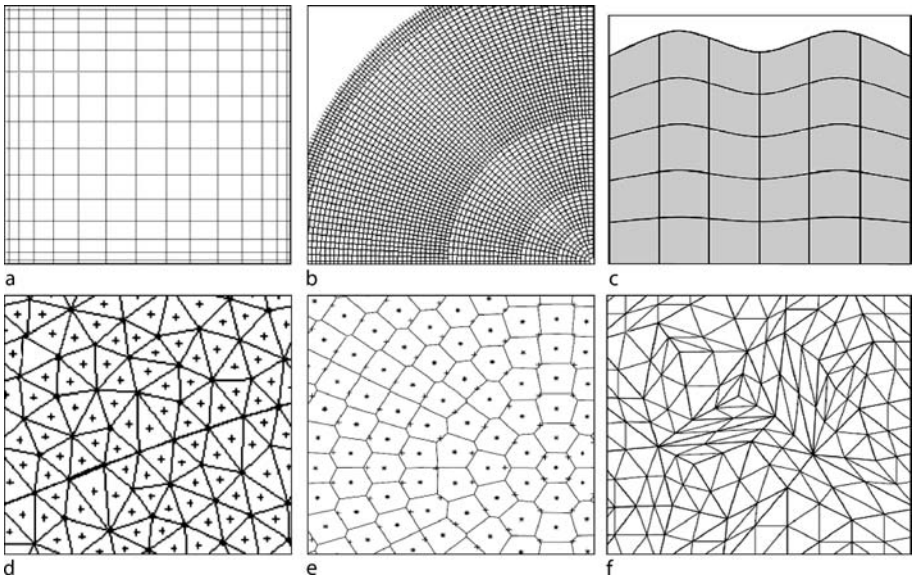
Grids for Wave Propagation Problems

The history of grid types used for problems in computational wave propagation is tightly linked to the evolution of numerical algorithms and available computational resources. The latter in the sense that – as motivated in the introduction – even today realistic simulations of wave propagation are still computationally expensive. This implies that it is not sufficient to apply stable and simple numerical schemes and just use enough grid points per wavelength and/or extremely fine grids for geometrically complex models. Optimal mathematical algorithms that minimize the computational effort are still sought for as the recent developments show that are outlined in the following sections.

In Fig. 4 a number of different computational grids in two space dimensions is illustrated. The simple-most equally-spaced regular finite-difference grid is only of practical use in situations without strong material discontinuities. With the introduction of the pseudospectral method based on Chebyshev polynomials grids as shown in Fig. 4a grids appeared that are denser near the domain boundaries and coarse in the interior. While this enabled a much more efficient implementation of boundary conditions the ratio between the size of the largest to the smallest cell depends on the overall number of grid points per dimension and can be very large. This leads to very small time steps, that can in some way be compensated by grid stretching [9] but overall the problem remains. An elegant way of allowing grids to be of more practical shape is by stretching the grids using analytical functions (Fig. 4c, this basically corresponds to a coordinate transformation, e. g., [50,107]). By doing this either smooth surface topography or smoothly varying internal interfaces can be followed by the grid allowing a more efficient simulation of geometrical features compared to a blocky representation on standard finite difference grids.

The problem of global wave propagation using spherical coordinates (here in the two-dimensional, axi-symmetric approximation) nicely illustrates the necessity to have spatially varying grid density (e. g., [42,43,53,59,89,109]). The grid shown in Fig. 4b demonstrates that in spherical coordinates a regular discretization leads to grid distances that get smaller and smaller towards the center of the Earth. This is in contrast to what is required to efficiently model the Earth's velocity structure: Velocities are small near the surface (requiring high grid density) and increase towards the center of the Earth (requiring low grid density). One way of adjusting is by re-gridding the mesh every now and then, in this case doubling the grid spacing appropriately. This is possible, yet it requires interpolation at the domain boundaries that slightly degrades the accuracy of the scheme.

The problems with grid density and complex surfaces cry for the use of so-called unstructured grids. Let us define an unstructured grid as an initial set of points (a point cloud), each point characterized by its spatial coordinates. We wish to solve our partial differential equations on this point set. It is clear that – with appropriate grid generation software – it is fairly easy to generate such grids that obey exactly any given geometrical constraints be it in connection with surfaces or velocity mod-

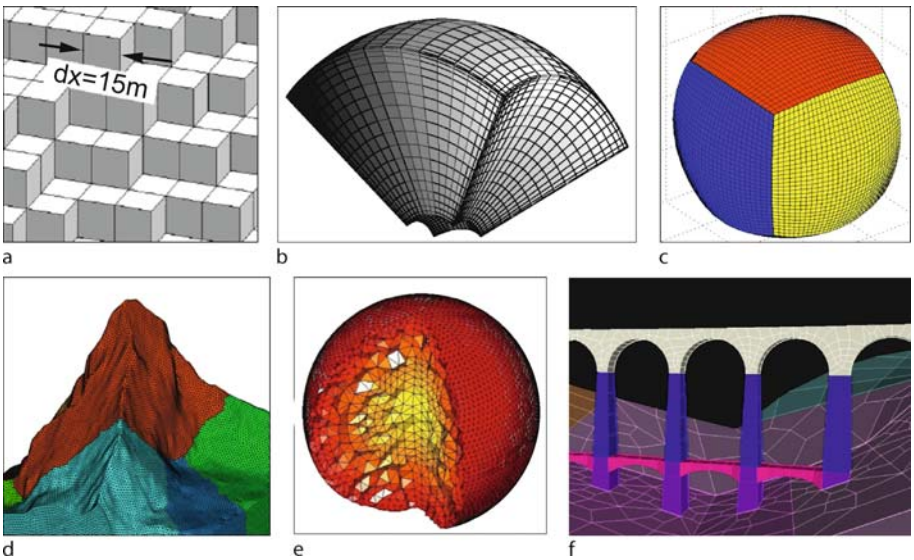


Seismic Wave Propagation in Media with Complex Geometries, Simulation of, Figure 4
Examples of 2D grids used for wave propagation simulations. a Chebyshev grid with grid densification near the domain boundaries. b Multidomain finite-difference grid in regular spherical coordinates. c Stretched regular finite-difference grid that allows following smoothly varying interface or surface boundaries. d Triangular staggered grid following an interface that allows finite-difference type operators. e Unstructured grid with associated Voronoi cells for calculations using the finite-volume method. f Triangular cells for finite-element type calculations. See text for details and references

els (i. e., varying grid density). It is important to note that such point clouds cannot be represented by 2D or 3D matrices as is the case for regular or regular stretched grid types. This has important consequences for the parallelization of numerical schemes. The first step after defining a point set is to use concepts from computational geometry to handle the previously unconnected points. This is done through the idea of Voronoi cells, that uniquely define triangles and their neighbors (Delauney triangulation). In Fig. 4d an example is shown for a triangular grid that follows an internal interface [72]. For finite-difference type operators on triangular grids a grid-staggering makes sense. Therefore, velocities would be defined in the center of triangles and stresses at the triangle vertices. Voronoi cells (Fig. 4e) can be used as volumetric elements for finite volume schemes [31,73]. For finite-element schemes triangular elements (Fig. 4f, e. g., [70]) with appropriate triangular shape functions are quite standard but have not found wide applications in seismology.

If the grid spacing of a regular finite-difference grid scheme in 3D would have to be halved this would result in an overall increase of computation time by a factor of 8 (a factor 2 per space dimension and another factor 2 because of the necessary halving of the time step). This simply means that the accuracy of a specific numerical scheme and the saving in memory or computation time is much more relevant in three dimensions. The evolution of grids in three dimensions is illustrated with examples in Fig. 5. A geometrical feature that needs to be modeled correctly particularly in volcanic environments is the free surface. With standard regular-spaced finite-difference schemes only a stair-step representation of the surface is possible (Fig. 5a, e. g., [87,92]). While the specific numerical implementation is stable and converges to the correct solution a tremendous number of grid points is necessary to achieve high accuracy.

Chebyshev grids and regular grids were applied to the problem of wave propagation in spherical sections (Fig. 5b, e. g., [52,57]). The advantage of solving the problem in spherical coordinates is the natural orthogonal coordinate system that facilitates the implementation of boundary conditions. However, due to the nature of spherical coordinates the physical domain should be close to the equator and geographical models have to be rotated accordingly. A highly successful concept for wave propagation in spherical media was possible through the adoption of the cubed-sphere approach in combination with spectral-elements (Fig. 5c, [63,64]). The cubed-sphere discretization is based on hexahedral grids. Towards the center of the Earth the grid spacing is altered to keep the number of elements per wavelength approximately constant.



Seismic Wave Propagation in Media with Complex Geometries, Simulation of, Figure 5
Examples of 3D grids. a Stair-step representation of a complex free surface with finite-difference cells. b Chebyshev grid in spherical coordinates for a spherical section. c Cubed sphere grid used for spectral-element and multi-domain Chebyshev calculations. d Tetrahedral grid of the Matterhorn. e Tetrahedral grid of the Earth's interior with the grid density tied to the velocity model. f Hexahedral grid of bridge structure and subsurface structure for spectral-element calculations. See text for details and references

Computational grids for wave propagation based on tetrahedra (Fig. 5d,e) are only recently being used for seismic wave propagation in combination with appropriate numerical algorithms such as finite volumes [34] or discontinuous Galerkin (e. g., [70]). The main advantage is that the grid generation process is greatly facilitated when using tetrahedra compared to hexahedra. Generating point clouds that follow internal velocity structures and connecting them to tetrahedra are straight forward and efficient mathematical computations. However, as described in more detail below, tetrahedral grids require more involved computations and are thus less efficient than hexahedral grids. Complex hexahedral grids – even for combined modeling of structure and soil (Fig. 5f) are possible but – at least at present – require a large amount of manual interaction during the grid generation process. It is likely that the combination of both grid types (tetrahedral in complex regions, hexahedral in less complex regions) will play an important role in future developments.

In the following we would like to present two of the most competitive schemes presently under development, (1) the spectral element method and (2) the discontinuous Galerkin approach combined with finite-volume flux schemes. The aim is to particularly illustrate the role of the grid generation process and the pros and cons of the specific methodologies.

3D Wave Propagation on Hexahedral Grids: Soil-Structure Interactions

We briefly present the spectral element method (SEM) based on Lagrange polynomials, focusing only on its main features and on its implementation for the solution of the elasto-dynamic equations. The SEM can be regarded as a generalization of the finite element method (FEM) based on the use of high order piecewise polynomial functions. The crucial aspect of the method is the capability of providing an arbitrary increase in spatial accuracy simply enhancing the algebraic degree of these functions (the spectral degree SD). On practical ground, this operation is completely transparent to the users, who limit themselves to choosing the spectral degree at runtime, leaving to the computational code the task of building up suitable quadrature points for integration and new degrees of freedom. Obviously, the increasing spectral degree implies raising the required computational effort.

On the other hand, one can also play on the grid refinement to improve the accuracy of the numerical solution, thus following the standard finite element approach. Spectral elements are therefore a so-called "$h-p$" method, where "h" refers to the grid size and "p" to the degree of polynomials. Referring to Faccioli et al. [40], Komatitsch and Vilotte [65], Chaljub et al. [15] for further details, we briefly remind in the sequel the key features of the spectral

element method adopted. We start from the wave equation for the displacement $\boldsymbol{u}$:

$$\rho \frac{\partial \boldsymbol{u}^2}{\partial t^2} = \operatorname{div} \sigma_{ij}(\boldsymbol{u}) + f\,, \quad i,j = 1 \ldots \mathrm{d}(\mathrm{d} = 2,3) \quad (2)$$

where t is the time, $\rho = \rho(x)$ the material density, $f = f(x,t)$ a known body force distribution and σ_{ij} the stress tensor. Introducing Hooke's law:

$$\sigma_{ij}(\boldsymbol{u}) = \lambda \operatorname{div} \boldsymbol{u} \delta_{ij} + 2\mu \varepsilon_{ij}(\boldsymbol{u})\,, \quad (3)$$

where

$$\varepsilon_{ij}(\boldsymbol{u}) = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \quad (4)$$

is the strain tensor, λ and μ are the Lamé coefficients, and δ_{ij} is the Kronecker symbol, i. e. $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$, otherwise.

As in the FEM approach, the dynamic equilibrium problem for the medium can be stated in the weak, or variational form, through the principle of virtual work [121] and through a suitable discretization procedure that depends on the numerical approach adopted, can be written as an ordinary differential equations system with respect to time:

$$[M]\,\ddot{\boldsymbol{U}}(t) + [K]\,\boldsymbol{U}(t) = \boldsymbol{F}(t) + \boldsymbol{T}(t) \quad (5)$$

with matrices $[M]$ and $[K]$, respectively the mass and stiffness matrices, and vectors $\boldsymbol{F}$ and $\boldsymbol{T}$ representing the contributions of external forces and traction conditions, respectively. In our SE approach, $\boldsymbol{U}$ denotes the displacement vector at the Gauss–Lobatto–Legendre (GLL) nodes, that correspond to the zeroes of the first derivatives of Legendre polynomial of degree N. The advancement of the numerical solution in time is provided by the explicit 2nd order leap-frog scheme. This scheme is conditionally stable and must satisfy the well known and already mentioned Courant–Friedrichs–Levy (CFL) condition. The key features of the SE discretization are described in the following.

Like in the FEM standard technique, the computational domain may be split into quadrilaterals in 2D or hexahedra in 3D, both the local distribution of grid points within the single element and the global mesh of all the grid points in the domain must be assigned. Many of these grid points are shared amongst several spectral elements. Each spectral element is obtained by a mapping of a master element through a suitable transformation and all computations are performed on the master element. Research is in progress regarding the introduction of triangular spectral elements [80]. The nodes within the element where displacements and spatial derivatives are computed, on which volume integrals are evaluated, are not necessarily equally spaced (similar to the Chebyshev approach in pseudospectral methods mentioned above). The interpolation of the solution within the element is done by Lagrange polynomials of suitable degree. The integration in space is done through Gauss–Lobatto–Legendre quadrature formula.

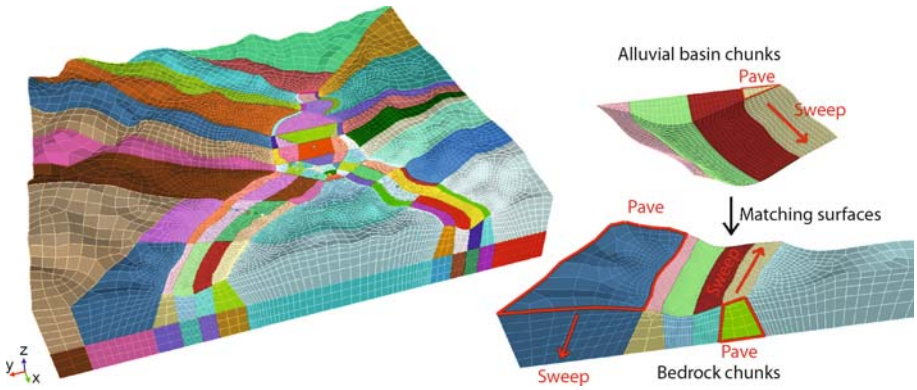
Thanks to this numerical strategy, the exponential accuracy of the method is ensured and the computational effort minimized, since the mass matrix results to be diagonal. The spectral element (SE) approach developed by Faccioli et al. [40] has been recently implemented in the computational code GeoELSE (GeoELasticity by Spectral Elements) [93,102,120] for 2D/3D wave propagation analysis. The most recent version of the code includes: (i) the capability of dealing with fully unstructured computational domains, (ii) the parallel architecture, and (iii) visco-plastic constitutive behavior [30]. The mesh can be created through an external software (e. g., CUBIT [25]) and the mesh partitioning is handled by METIS [81].

Hexahedral Grids

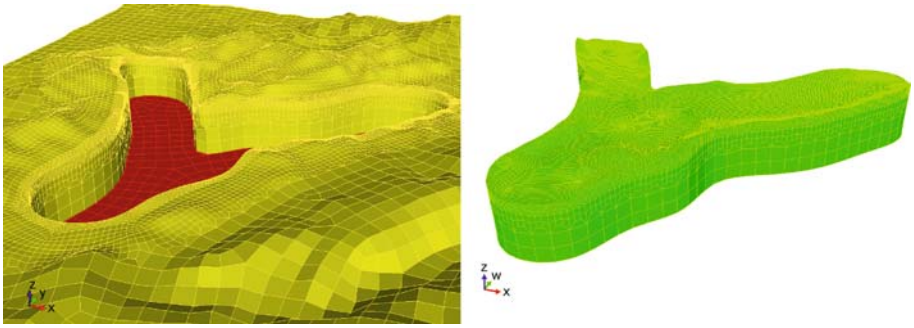
As already mentioned in the SEM here presented the computational domain is decomposed into a family of non overlapping quadrilaterals in 2D or hexahedra in 3D. The grid discretization should be suitable to accurately propagate up to certain frequencies. Obviously, owing to the strong difference of the mechanical properties between soft-soil and stiff-soil (or building construction material) and to the different geometrical details as well, the grid refinement needed in the various parts of the model varies substantially. Therefore, a highly unstructured mesh is needed to minimize the number of elements. While 3D unstructured tetrahedral meshes can be achieved quite easily with commercial or non commercial software, the creation of a 3D non structured hexahedra mesh is still recognized as a challenging problem. In the following paragraph we provide state of the art results concerning the mesh creation.

Grid Generation

Hexahedral grids have more severe restrictions in meshing efficiently. This is basically related to the intrinsic difficulty that arises from the mapping of the computational domain with this particular element. As a consequence automatic procedures have difficulty capturing specific boundaries, create poor quality elements, the assigned size is difficult to be preserved and the generation process is usually much slower compared to the tetrahedral mesh generation algo-



Seismic Wave Propagation in Media with Complex Geometries, Simulation of, Figure 6
3D numerical model used for the simulations of ESG06 "Grenoble Benchmark". "Honoring" technique: The computational domain is subdivided into small chunks and each one is meshed starting from the alluvial basin down to the bedrock. For simplicity only the spectral elements are shown without GLL nodes



Seismic Wave Propagation in Media with Complex Geometries, Simulation of, Figure 7
3D numerical model used for the simulations of ESG06 "Grenoble Benchmark". "Not Honoring" technique: The computational domain is meshed with a coarse mesh and then refined twice approximately in the area where the alluvial basin is located

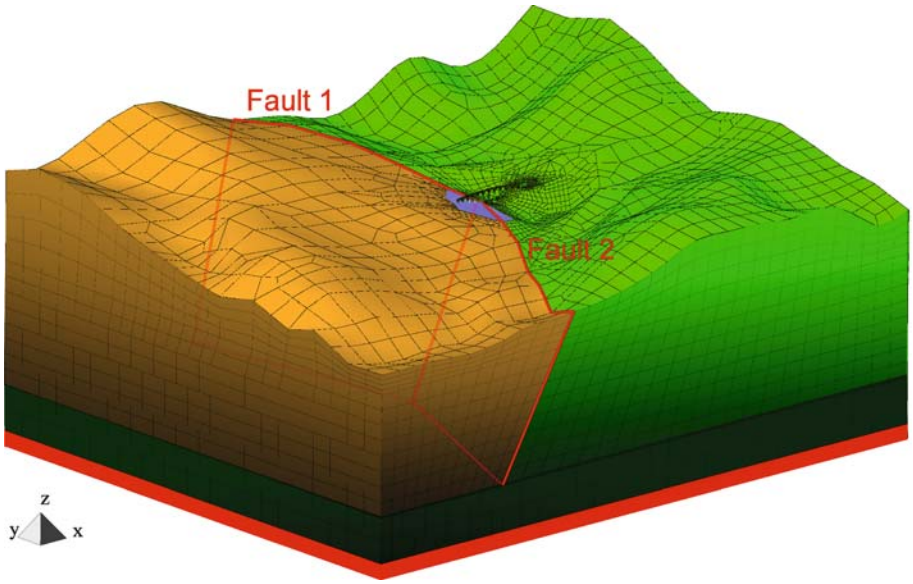
rithms. On the other hand the advantages of hexahedral meshes are usually related to the lower computational cost of the wave propagation solutions with respect to the one based on triangular meshes or hexahedral structured grids (like in the finite difference method).

Nevertheless certain problems can be addressed reasonably well with specific solutions. A quite typical case in earthquake seismology is the study of the alluvial basin response under seismic excitation. In handling this problem, a first strategy is to try to "honor" the interface between the sediment (soft soil) and the bedrock (stiff soil). The two materials are divided by a physical interface and the jump in the mechanical properties is strictly preserved. The major drawback of this approach is that usually it requires strong skills from the user to build-up the mesh and a significant amount of working time (Fig. 6). Given that the "honoring approach" is not always feasible in a reasonable time (or with a reasonable effort) a second strategy is worth to be mentioned: The so called "not honoring" procedure. In this second case the mesh is refined in proximity of the area where the soft deposit are localized but the elements do not respect the interface. On a practical ground the mechanical properties are assigned node by node and the sharp jump is smoothed through the Lagrange interpolation polynomial and substituted with smeared interfaces (Fig. 7). At the present time it is still strongly under debate if it is worth to honor or not the physical interfaces.

Finally, we highlight the fact that meshing software (e. g., CUBIT [25]) is available that seems to be extremely promising and potentially very powerful for the creation of geophysical and seismic engineering unstructured hexahedral meshes. Further very interesting mesh generation procedures based on hexahedral are under investigation [99].

Scale Problem with Structure and Soil

In engineering practice one of the most common approaches to design buildings under seismic load is the im-

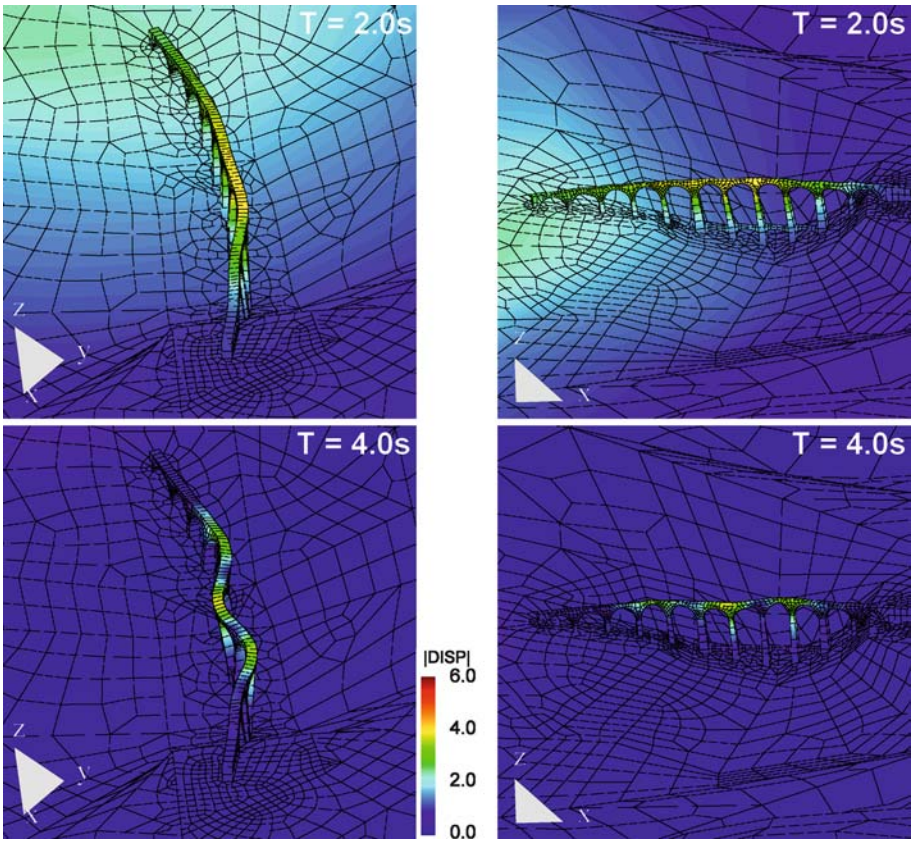


Seismic Wave Propagation in Media with Complex Geometries, Simulation of, Figure 8
3D model of Acquasanta bridge and the surrounding geological configuration. The investigated area is 2 km in length, 1.75 km in width and 0.86 km in depth. The model was designed to propagate waves up to 5 Hz with a SD = 3 (Order 4) and has 38,569 hexahedral elements and 1,075,276 grid points. The contact between calcareous schists (*brown color*) and serpentine rocks (*green color*) is modeled with two sub-vertical faults (*red-line*). *Cyan color* represents the alluvial and weathered deposits

position of an acceleration time history to the structure, basically acting like an external load. An excellent example of this technique can be found in recent publications (e. g., [68,69]) and in the study of the so-called "urban-seismology", recently presented by Fernandez-Ares et al., [44]. In this case the goal is to understand how the presence of an entire city can modify the incident wave-field. Due to the size of the simulation and the number of buildings, the latter are modeled as single degrees of freedom oscillators. The interaction between soil and structure is preserved but the buildings are simplified. For important structure (e. g.: Historical buildings, world heritage buildings, hospitals, schools, theaters, railway and highways) it is worth to provide an ad-hoc analysis capable to take into account the full complexity of the phenomena.

Here we present an example of a fully coupled modeling (Fig. 8): A railway bridge and its geotechnical-topographical surroundings. The Acquasanta bridge on the Genoa-Ovada railway, North Italy, is located in the Genoa district and represents a typical structure the ancestor of which can be traced back to the Roman "Pont du Gard". This structural type did not change significantly along the centuries, thanks to the excellent design achieved no less than 1900 years ago. The Acquasanta bridge structure is remarkable both for the site features and the local geological and geomorphological conditions. The foundations of several of the piers rest on weak rock; moreover, some instability problems have been detected in the past on the valley slope towards Ovada.

Several simulations have been performed with GeoELSE, in order to evaluate the influence of seismic site effects on the dynamic response of the Acquasanta bridge. A fully coupled 3D soil-structure model was designed: The grid is characterized by a "subvertical fault" between calcareous schists and serpentine rocks. This is in accordance with available data, even if further investigations in future should identify more in detail the tectonic structure of the area. The geometry of weathered materials overlaying the calcareous schists on the Ovada side has been assumed according to available information. The dimension of hexahedral elements ranges some tens of centimeters to about 1000 m. With such a model, the problem can be handled in its 3D complexity and we can examine the following aspects that are usually analyzed under restrictive and simplified assumptions: (i) soil-structure interaction, (ii) topographic amplification, (iii) soft soil amplification (caused by the superficial alluvium deposit shown in cyan), (iv) subvertical fault (red line) between the schists, on the Ovada side, and serpentine rock, on the Genoa side. For excitation a shear plane wave (x-direction) was used (Ricker wavelet, fmax = 3 Hz, $t0 = 1.0\,s.$ and amplitude = 1 mm) propagating vertically from the bottom (red elements in Fig. 8).



Seismic Wave Propagation in Media with Complex Geometries, Simulation of, Figure 9
Snapshots of the modulus of the displacement vector and the magnified deformed shape of the bridge (in mm)

In Fig. 9 we present some snapshots of the modulus of the displacement vector and the magnified deformed shape of the bridge. It is worth to note that at $T = 2$ s the motion of the bridge is almost in-plane (direction x), while at $T = 4$ s is clearly evident how the coupling between the in-plane and out-plane (y-direction) motion starts to be important.

The study of the soil-structure interaction problem could be easily enhanced (i) improving the input excitation of the model here presented and (ii) taking into account complex constitutive behavior both from the soil and the structure side. The former is already available in GeoELSE thanks to the recent implementation [41,93] of the domain reduction method (DRM), a methodology that divides the original problem into two simpler ones [4,119], to overcome the problem of multiple physical scales that is created by a seismic source located some kilometers away from the structure with typical element size of the order of meters and located over a relatively small area (less than 1 km^2) on soft deposit. The latter still needs to be improved because of the lack of a complete tool capable to handle in 3D non linear soil behavior, non-linear structural behavior and the presence of the water, that play a crucial role in the failure of buildings. Partial response to this problem can be found in the recent work of Bonilla et al. [5] and in the visco-plastic rheology recently introduced in GeoELSE [30].

3D Wave Propagation on Tetrahedral Grids: Application to Volcanology

As indicated above, the simulation of a complete, highly accurate wave field in realistic media with complex geometry is still a great challenge. Therefore, in the last years a new, highly flexible and powerful simulation method has been developed that combines the Discontinuous Galerkin (DG) Method with a time integration method using Arbitrary high order DERivatives (ADER) of the approximation polynomials. The unique property of this numerical scheme is, that it achieves arbitrarily high approximation order for the solution of the governing seismic

wave equation in space and time on structured and unstructured meshes in two and three space dimensions.

Originally, this new ADER-DG approach [32,35] was introduced for general linear hyperbolic equation systems with constant coefficients or for linear systems with variable coefficients in conservative form. Then, the extension to non-conservative systems with variable coefficients and source terms and its particular application to the simulation of seismic waves on unstructured triangular meshes in two space dimensions was presented [70]. And finally, the further extension of this approach to three-dimensional tetrahedral meshes has been achieved [33]. Furthermore, the accurate treatment of viscoelastic attenuation, anisotropy and poroelasticity has been included to handle more complex rheologies [28,29,71]. The governing system of the three-dimensional seismic wave equations is hereby formulated in velocity-stress and leads to the hyperbolic system of the form

$$\frac{\partial \boldsymbol{Q}_p}{\partial t} + A_{pq}\frac{\partial \boldsymbol{Q}_q}{\partial \xi} + B_{pq}\frac{\partial \boldsymbol{Q}_q}{\partial \eta} + C_{pq}\frac{\partial \boldsymbol{Q}_q}{\partial \zeta} = S_p \,, \quad (6)$$

where the vector $\boldsymbol{Q}$ of unknowns contains the six stress and the three velocity components and S is the source term. The Jacobian matrices A, B and C include the material values as explained in detail in [33,70].

The ADER-DG Method: Basic Concepts

The ADER-DG method is based on the combination of the ADER time integration approach [113], originally developed in the finite volume (FV) framework [96,97,111] and the Discontinuous Galerkin finite element method [18,19,20,21,22,23,91]. As described in detail in [33] in the ADER-DG approach the solution is approximated inside each tetrahedron by a linear combination of space-dependent polynomial basis functions and time-dependent degrees of freedom as expressed through

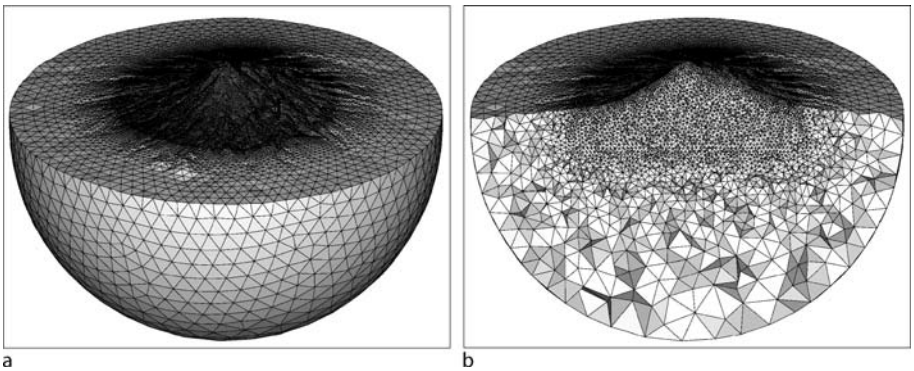
$$(\boldsymbol{Q}_h)_p(\xi,\eta,\zeta,t) = \hat{\boldsymbol{Q}}_{pl}(t)\boldsymbol{\Phi}_l(\xi,\eta,\zeta) \,, \quad (7)$$

where the basis functions Φ_l form an orthogonal basis and are defined on the canonical reference tetrahedron. The unknown solution inside each element is then approximated by a polynomial, whose coefficients – the degrees of freedom $\hat{\boldsymbol{Q}}_{pl}$ – are advanced in time. Hereby, the solution can be discontinuous across the element interfaces, which allows the incorporation of the well-established ideas of numerical flux functions from the finite volume framework [75,112]. To define a suitable flux over the element surfaces, so-called Generalized Riemann Problems (GRP) are solved at the element interfaces. The GRP solution provides simultaneously a numerical flux function as well as a time-integration method. The main idea is a Taylor expansion in time in which all time derivatives are replaced by space derivatives using the so-called Cauchy–Kovalewski procedure which makes recursive use of the governing differential Eq. (6). The numerical solution of Eq. (6) can thus be advanced by one time step without intermediate stages as typical e. g. for classical Runge–Kutta time stepping schemes. Due to the ADER time integration technique the same approximation order in space and time is achieved automatically. Furthermore, the projection of the elements in physical space onto a canonical reference element allows for an efficient implementation, as many computations of three-dimensional integrals can be carried out analytically beforehand. Based on a numerical convergence analysis this new scheme provides arbitrary high order accuracy on unstructured meshes. Moreover, due to the choice of the basis functions in Eq. (7) for the piecewise polynomial approximation [23], the ADER-DG method shows even spectral convergence.

Grid Generation: Unstructured Triangulations and Tetrahedralization

Both tetrahedral and hexahedral elements are effectively used to discretize three-dimensional computational domains and model wave propagation with finite element type methods. Tetrahedra can be the right choice because of the robustness when meshing any general shape. Hexahedra can be the element of choice due to their ability to provide more efficiency and accuracy in the computational process. Furthermore, techniques for automatic mesh generation, gradual mesh refinement and coarsening are generally much more robust for tetrahedral meshes in comparison to hexahedral meshes. Straightforward tetrahedral refinement schemes, based on longest-edge division, as well as the extension to adaptive refinement or coarsening procedures of a refined mesh exist [3,12]. In addition, parallel strategies for refinement and coarsening of tetrahedral meshes have been developed [27].

Less attention has been given to the modification of hexahedral meshes. Methods using iterative octrees have been proposed [74,95], but these methods often result in nonconformal elements that cannot be accommodated by some solvers. Lately also conformal refinement and coarsening strategies for hexahedral meshes have been proposed [2]. Other techniques insert non-hexahedral elements that result in hybrid meshes that need special solvers that can handle different mesh topologies. Commonly, the geometrical problems in geosciences arise through rough surface topography, as shown for the Mer-



Seismic Wave Propagation in Media with Complex Geometries, Simulation of, Figure 10
Tetrahedral mesh for the model of the volcano Merapi. The zone of interest, such as the free surface topography and the volcano's interior are discretized by a fine mesh, whereas the spatial mesh is gradually coarsened towards the model boundaries

api volcano in Fig. 10, and internal material boundaries of complex shape that lead to wedges and overturned or discontinuous surfaces due to folding and faulting. However, once the geometry of the problem is defined by the help of modern computer aided design (CAD) software, the meshing process using tetrahedral elements is automatic and stable. After the mesh generation process, the mesh vertices, the connectivity matrix and particular information about boundary surfaces are typically imported to a solver.

The computational possibilities and algorithmic flexibility of a particular solver using the ADER-DG approach for tetrahedral meshes are presented in the following.

Local Accuracy: *p*-Adaptation

In many large scale applications the computational domain is much larger than the particular zone of interest. Often such an enlarged domain is chosen to avoid effects from the boundaries that can pollute the seismic wave field with possible, spurious reflections. Therefore, a greater number of elements has to be used to discretize the domain describing the entire model. However, in most cases the high order accuracy is only required in a restricted area of the computational domain and it is desirable to choose the accuracy that locally varies in space. This means, that it must be possible to vary the degree p of the approximation polynomials locally from one element to the other [36]. As the ADER-DG method uses a hierarchical order of the basis functions to construct the approximation polynomials, the corresponding polynomial coefficients, i. e. the degrees of freedom, for a lower order polynomial are always a subset of those of a higher-order one. Therefore, the computation of fluxes between elements of different approximation orders can be carried out by using only the necessary part of the flux matrices.

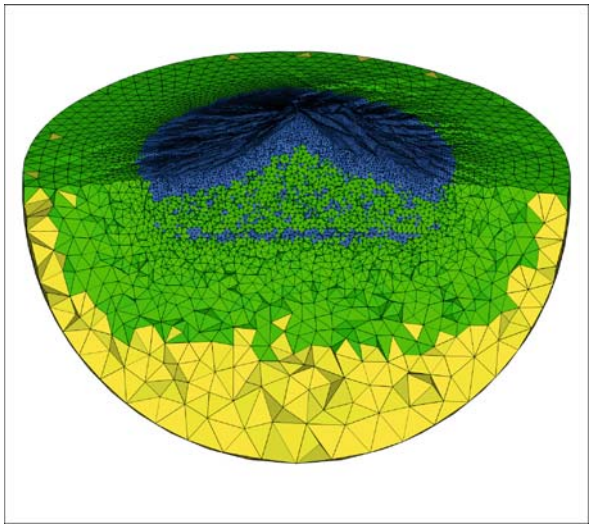
Furthermore, the direct coupling of the time and space accuracy via the ADER approach automatically leads to a local adaptation also in time accuracy, which often is referred to as p_t-adaptivity. In general, the distribution of the degree p might be connected to the mesh size h, i. e. the radius of the inscribed sphere of a tetrahedral element. In particular, the local degree p can be coupled to the mesh size h via the relations

$$p = p_{\min} + (p_{\max} - p_{\min}) \left(\frac{h - h_{\min}}{h_{\max} - h_{\min}} \right)^r , \qquad (8)$$

$$p = p_{\max} - (p_{\max} - p_{\min}) \left(\frac{h - h_{\min}}{h_{\max} - h_{\min}} \right)^r , \qquad (9)$$

where the choice of the power r determines the shape of the p-distribution. Note, that depending on the choice of the first term and the sign the degree p can increase as in Eq. (8) or decrease as in Eq. (9) with increasing h, starting from a minimum degree $p_{\min}$ up to a maximum degree $p_{\max}$. This provides additional flexibility for the distribution of p inside the computational domain. An example of a p-distribution for the volcano Merapi is given in Fig. 11.

Here the idea is to resolve the slowly propagating surface waves with high accuracy, whereas the waves propagating towards the absorbing model boundaries pass through a zone of low spatial resolution. This approach leads to numerical damping due to an amplitude decay that reduces possible boundary reflections. Furthermore, the computational cost is reduced significantly due to the strongly reduced number of total degrees of freedom in the model.



Seismic Wave Propagation in Media with Complex Geometries, Simulation of, Figure 11
The local degree *p* of the approximation polynomial depends on the insphere radius of each tetrahedral element and is given in color code. Close to the surface topography an approximation polynomial of degree $p = 5$ (*blue*) is used, whereas in depth the degree is reduced to $p = 4$ (*green*) and $p = 3$ (*yellow*)

Local Time Stepping: Δt-Adaptation

Geometrically complex computational domains or spatial resolution requirements often lead to meshes with small or even degenerate elements. Therefore, the time step for explicit numerical schemes is restricted by the ratio of the size h of the smallest element and the corresponding maximum wave speed in this element. For global time stepping schemes all elements are updated with this extremely restrictive time step length leading to a large amount of iterations. With the ADER-DG approach, time accurate local time stepping can be used, such that each element is updated by its own, optimal time step [36]. Local time-stepping was used in combination with the finite-difference method [42,106].

An element can be updated to the next time level if its actual time level and its local time step Δt fulfill the following condition with respect to all neighboring tetrahedra n:

$$t + \Delta t \leq \min(t_n + \Delta t_n) \,. \tag{10}$$

Figure 12 is visualizing the evolution of four elements (I, II, III and IV) in time using the suggested local time stepping scheme. A loop cycles over all elements and checks for each element, if condition (10) is fulfilled. At the initial state all elements are at the same time level, however, element II and IV fulfill condition (10) and therefore can be updated. In the next cycle, these elements are already advanced in time (grey shaded) in cycle 1. Now elements I and IV fulfill condition (10) and can be updated to their next local time level in cycle 2. This procedure continues and it is obvious, that the small element IV has to be updated more frequently than the others. A synchronization to a common global time level is only necessary, when data output at a particular time level is required as shown in Fig. 12.

Information exchange between elements across interfaces appears when numerical fluxes are calculated. These fluxes depend on the length of the local time interval over which a flux is integrated and the corresponding element is evolved in time. Therefore, when the update criterion (10) is fulfilled for an element, the flux between the element itself and its neighbor n has to be computed over the local time interval:

$$\tau n = [\max(t, tn) \,, \quad \min(t + \Delta t, tn + \Delta tn)] \,. \tag{11}$$

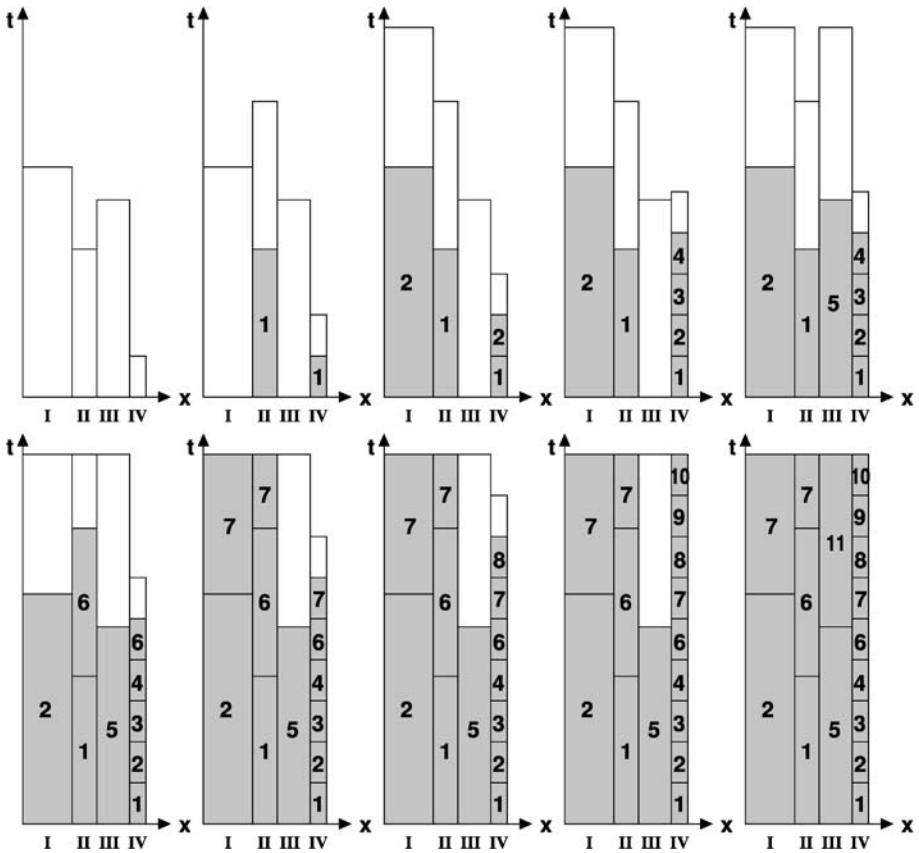
As example, the element III fulfills the update criterion (10) in cycle 5 (see Fig. 12). Therefore, when computing the fluxes only the remaining part of the flux given by the intervals in Eq. (11) has to be calculated. The other flux contribution was already computed by the neighbors II and IV during their previous local updates. These flux contributions have been accumulated and were stored into a memory variable and therefore just have to be added.

Note that e. g. element IV reaches the output time after 10 cycles and 9 local updates, which for a global time stepping scheme would require $9 \times 4 = 36$ updates for the all four elements. With the proposed local time stepping scheme only 16 updates are necessary to reach the same output time with all elements as indicated by the final number of grey shaded space time elements in Fig. 12.

Comparing these numbers leads to a speedup factor of 2.25. For strongly heterogeneous models and local time step lengths this factor can become even more pronounced. However, due to the asynchronous update of elements that might be spatially very close to each other the mesh partitioning for parallel computations becomes an important and difficult issue. Achieving a satisfying load balancing is a non-trivial task and still poses some unresolved problems as explained in the following.

Mesh Partitioning and Load Balancing

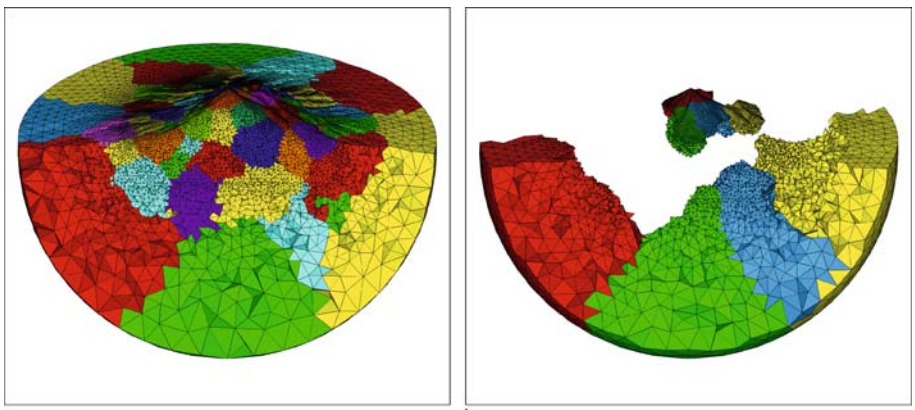
For large scale applications it is essential to design a parallel code that can be run on massively parallel supercomputing facilities. Therefore, the load balancing is an important issue to use the available computational resources efficiently. For global time stepping schemes with-



Seismic Wave Propagation in Media with Complex Geometries, Simulation of, Figure 12
Visualization of the local time stepping scheme. The actual local time level *t* is at the top of the *gray shaded* area with numbers indicating the cycle, in which the update was done. *Dotted lines* indicate the local time step length Δt with which an element is updated

out *p*-adaptation standard mesh partitioning as done e. g. by METIS [60] is sufficient to get satisfying load balancing. The unstructured tetrahedral mesh is partitioned into subdomains that contain an equal or at least very similar number of elements as shown in Fig. 13. Therefore, each processor has to carry out a similar amount of calculations. However, if *p*-adaptation is applied, the partitioning is more sophisticated as one subdomain might have many elements of high order polynomials whereas another might have the same number of elements but with lower order polynomials. Therefore, the parallel efficiency is restricted by the processor with the highest work load. However, this problem can usually be solved by weighted partitioning algorithms, e. g. METIS.

In the case of local time stepping, mesh partitioning is becoming a much more difficult task. One solution is to divide the computational domain into a number of zones, that usually contain a geometrical body or a geological zone that typically is meshed individually with a particular mesh spacing *h* and contains a dominant polynomial order. Then each of these zones is partitioned separately into subdomains of approximately equal numbers of elements. Then each processor receives a subdomain of each zone, which requires a similar amount of computational work as shown in Fig. 13. In particular, the equal distribution of tetraheda with different sizes is essential in combination with the local time stepping technique. Only if each processor receives subdomains with a similar amount of small and large elements, the work load is balanced. The large elements have to be updated less frequently than the smaller elements and therefore are computationally cheaper. Note, that the separately partitioned and afterwards merged zones lead to non-connected subdomains for each processor (see Fig. 13). This increases the number of element surfaces between subdomains of different processors and therefore increases the communication required. However, communication is typically low as the degrees of freedom have to be exchanged only once per



Seismic Wave Propagation in Media with Complex Geometries, Simulation of, Figure 13
Standard partitioning of the computational domain (*left*) and an example of 4 subdomains grouped together for more efficient local time stepping

time step and only for tetrahedra that have an interface at the boundary between subdomains. Therefore, the improvements due to the new load balancing approach are dominant and outweigh the increase in communication.

However, care has to be taken as the distribution of the polynomial degree p or the seismic velocity structure might influence the efficiency of this grouped partitioning technique. A profound and thorough mesh partitioning method is still a pending task as the combination of local time stepping and p-adaptivity requires a new weighting strategy of the computational cost for each tetrahedral element considering also the asynchronous element update. The automatic partitioning of unstructured meshes with such heterogeneous properties together with the constraint of keeping the subdomains as compact as possible to avoid further increase of communication is still subject to future work.

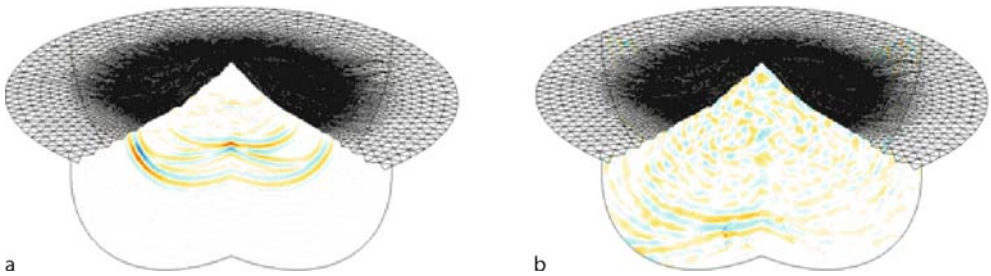
In Fig. 13 an example of a grouped partition of the tetrahedral mesh is shown for 4 processors. Two non-connected subdomains indicated by the same color are assigned to each processor including small – and therefore computationally expensive – tetrahedra that are updated frequently due to their small time step, and much larger elements that typically are cheap due to their large time step. This way, the work load often is balanced sufficiently well over the different processors.

Relevance of High Performance Computing: Application to Merapi Volcano

In recent years the development of the ADER-DG algorithm including the high order numerical approximation in space and time, the mesh generation, mesh adaptation, parameterization, and data visualization created the basis of an efficient and highly accurate seismic simulation tool. Realistic large scale applications and their specific requirements will further guide these developments. On the other hand, the study and incorporation of geophysical processes that govern seismic wave propagation insures, that the simulation technology matches the needs and addresses latest challenges in modern computational seismology. Hereby, the accurate modeling of different source mechanisms as well as the correct treatment of realistic material properties like anelasticity, viscoplasticity, porosity and highly heterogeneous, scattering media will play an important role.

However, only the combination of this state-of-the-art simulation technology with the most powerful supercomputing facilities actually available can provide excellent conditions to achieve scientific progress for realistic, large scale applications. This combination of modern technologies will substantially contribute to resolve current problems, not only in numerical seismology, but will also influence other disciplines. The phenomenon of acoustic, elastic or seismic wave propagation is encountered in many different fields. Beginning with the classical geophysical sciences seismology, oceanography, and volcanology such waves also appear in environmental geophysics, atmospheric physics, fluid dynamics, exploration geophysics, aerospace engineering or even medicine.

With the rapid development of modern computer technology and the development of new highly accurate simulation algorithms computer modeling just started to herald a new era in many applied sciences. The 3D wave propagation simulations in realistic media require a substantial amount of computation time even on large par-



Seismic Wave Propagation in Media with Complex Geometries, Simulation of, Figure 14
Snapshots of the seismic wave field after an explosive event close to the summit of Merapi volcano. The free surface topography introduces strong scattering of the waves making it extremely difficult to invert for the seismic source mechanism or the exact source location

allel computers. Extremely powerful national supercomputers already allow us to run simulations with unrivaled accuracy and resolution. However, using the extremely high accuracy and flexibility of new simulation methods on such massively parallel machines the professional support of experts in supercomputing is absolutely essential. Only professional porting, specific CPU-time and storage optimizations of current software with respect to continuously changing compilers, operating systems, hardware architectures or simply personnel, will ensure the lifetime of new simulation technologies accompanied by ongoing improvements and further developments. Additionally, the expertise and support in the visualization of scientific results using technologies of Virtual Reality for full 3D models not only enhances the value of simulation results but will support data interpretation and awake great interest in the new technology within a wide research community.

As an example, volcano monitoring plays an increasingly important role in hazard estimation in many densely populated areas in the world. Highly accurate computer modeling today is a key issue to understand the processes and driving forces that can lead to dome building, eruptions or pyroclastic flows. However, data of seismic observations at volcanoes are often very difficult to interpret. Inverting for the source mechanism, i. e. seismic moment tensor inversion, or just locating an exact source position is often impossible due to the strongly scattered wave field caused by an extremely heterogeneous material distribution inside the volcano. Furthermore, the rough topography alone can affect the wave field by its strongly scattering properties as shown in Fig. 14.

Therefore, it is fundamental to understand the effects of topography and scattering media and there influence on the seismic wave field. A systematic study of a large number of scenarios computed by highly accurate simulation methods to provide reliable synthetic data sets is necessary to test the capabilities of currently used inversion tools. Slight changes in parameters like the source position, the source mechanism or the elastic and geometric properties of the medium can then reveal the limits of such tools and provide more precise bounds of their applicability in volcano seismology.

Finally, the implementation of the ADER-DG method is still much more expensive than other state-of-the-art implementations of existing methods. However, a fair comparison between accuracy and computational cost is still a pending task. The main reason for the CPU-time difference is the much larger number of tetrahedral elements than hexahedra that have to be used to cover the same volume. Furthermore, due to the choice of the basis functions, the flux computations are expensive, as the matrix-matrix multiplications involved are not sparse.

However, the ADER-DG method is currently implemented on hexahedral meshes to make fair comparisons possible. Preliminary tests show, that the change of mesh topology from tetrahera to hexahedra significantly reduces the computational cost. However, final results are subject to future investigations.

Discussion and Future Directions

As indicated in the introduction and highlighted in the previous sections, computational tools for wave propagation problems are getting increasingly sophisticated to meet the needs of current scientific problems. We are far away from simple finite-difference time schemes that are solving problems on regular grids on serial computers in which case the particular programming approach did not affect dramatically the overall performance. Today, competitive algorithms are results of years of partly highly professional coding. Implementations on high-performance computing hardware requires in-depth knowledge of parallel algorithms, profiling, and many technical aspects of modern computations. To make complex scientific soft-

ware available to other researchers requires implementation and testing on many different (parallel) platforms. This may involve parallelization using different programming paradigms (e. g., the combination of OpenMP and MPI on nodes of shared memory machines), and interoperability on heterogeneous computational GRIDs.

This has dramatic consequences particularly for young researchers in the Earth Sciences who want to use advanced computational tools to model observations. While in the early days a finite-difference type algorithm could be understood, coded, implemented and tested in a few weeks, this is no longer possible. In addition, standard curricula do not offer training in computational methods allowing them to efficiently write and test codes. This suggests that at least for some, well-defined computational problems verified and professionally engineered scientific software solutions should be provided to the community and also professionally extended and maintained in close collaboration with scientists. In seismology we are in a quite fortunate situation. In contrast to many other fields of physical sciences, our constitutive relations (e. g., stress-strain) are fairly well understood, and – as indicated in this paper – numerical solutions for 3D problems and their implementation on parallel hardware are well advanced. Another argument for stable tested "community"-codes for wave propagation is the fact that advancement in many scientific problems (e. g., imaging the Earth's interior, quantifying earthquake-induced shaking hazard) relies on zillions of forward modeling runs with only slight variations of the internal velocity models.

As far as technical developments are concerned, the efficient initialization of complex 3D models on computational grids is still a great challenge. Realistic models may be composed of complex topography, families of overlapping fault surfaces, discontinuous interfaces, and varying rheologies (e. g., elastic, anisotropic, viscoelastic, viscoplastic, porous). This may require the combination of tetrahedral and hexahedral grid in models with strongly varying degree of complexity. Ideally, standards for Earth models (and synthetic data) formats should be established by the communities that allow easy exchange and multiple use of models with different simulation tools (e. g., wave propagation, deformation, earthquake rupture). In addition, the rapid developments towards PetaFlop computing opens new questions about the scalability and efficient parallelization of current and future algorithms.

As the forward problem of wave propagation is at the core of the seismic imaging problem for both source and Earth's structure, in the near future we will see the incorporation of 3D simulation technology into the imaging process. Provided that the background seismic velocity models are fairly well known (e. g., reservoirs, global Earth, sedimentary basins), adjoint methods provide a powerful analytical tool to (1) relate model deficiencies to misfit in observations and (2) quantify the sensitivities to specific aspects of the observations (e. g., [100,104,105]). As the core of the adjoint calculations is the seismic forward problem, the challenge is the actual application to real data and the appropriate parametrizations of the model and the data that optimize the data fitting process.

In summary, while we look back at (and forward to) exciting developments in computational seismology, a paradigm shift in the conception of one of the central tools of seismology – the calculation of 3D synthetic seismograms – is necessary. To extract a maximum amount of information from our high-quality observations scientists should have access to high-quality simulation tools. It is time to accept that "*software is infrastructure*" and provide the means to professionally develop and maintain community codes and model libraries at least for basic Earth science problems and specific focus regions. Developments are one the way along those lines in the SPICE project (Seismic Wave Propagation and Imaging in Complex Media, a European Network [101]), the Southern California Earthquake Center (SCEC [94]) and the CIG Project (Computational infrastructure in geodynamics [17]).

Acknowledgments

We would like to acknowledge partial support towards this research from: The European Human Resources and Mobility Program (SPICE-Network), the German Research Foundation (Emmy Noether-Programme), the Bavarian Government (KOHNWIHR, graduate college THESIS, BaCaTec), and MunichRe. We would also like to thank J. Tromp supporting MS's visit to CalTech. We also thank two anonymous reviewers for constructive comments on the manuscript.

Bibliography

Primary Literature

1. Alterman Z, Karal FC (1968) Propagation of elastic waves in layered media by finite-difference methods. Bull Seism Soc Am 58:367–398
2. Benzley SE, Harris NJ, Scott M, Borden M, Owen SJ (2005) Conformal refinement and coarsening of unstructured hexahedral meshes. J Comput Inf Sci Eng 5:330–337
3. Bey J (1995) Tetrahedral grid refinement. Computing 55:355–378
4. Bielak J, Loukakis K, Hisada Y, Yoshimura C (2003) Domain reduction method for three-dimensional earthquake modeling in localized regions, Part I: Theory. Bull Seism Soc Am 93:817–824

5. Bonilla LF, Archuleta RJ, Lavallée D (2005) Hysteretic and dilatant behavior of cohesionless soils and their effects on nonlinear site response: Field data observations and modelling. Bull Seism Soc Am 95(6):2373–2395
6. Boore D (1972) Finite-difference methods for seismic wave propagation in heterogeneous materials. In: Bolt BA (ed) Methods in Computational Physics, vol 11. Academic Press, New York
7. Braun J, Sambridge MS (1995) A numerical method for solving partial differential equations on highly irregular evolving grids. Nature 376:655–660
8. Bunge HP, Tromp J (2003) Supercomputing moves to universities and makes possible new ways to organize computational research. EOS 84(4):30, 33
9. Carcione JM, Wang J-P (1993) A Chebyshev collocation method for the elastodynamic equation in generalised coordinates. Comp Fluid Dyn 2:269–290
10. Carcione JM, Kosloff D, Kosloff R (1988) Viscoacoustic wave propagation simulation in the earth. Geophysics 53: 769–777
11. Carcione JM, Kosloff D, Behle A, Seriani G (1992) A spectral scheme for wave propagation simulation in 3-D elastic-anisotropic media. Geophysics 57:1593–1607
12. Carey G (1997) Computational grids: Generation, adaptation, and solution strategies. Taylor Francis, New York
13. Cerveny V (2001) Seismic ray theory. Cambridge University Press, Cambridge
14. Chaljub E, Tarantola A (1997) Sensitivity of SS precursors to topography on the upper-mantle 660-km discontinuity. Geophys Res Lett 24(21):2613–2616
15. Chaljub E, Komatitsch D, Vilotte JP, Capdeville Y, Valette B, Festa G (2007) Spectral element analysis in seismology. In: Wu R-S, Maupin V (eds) Advances in wave propagation in heterogeneous media. Advances in Geophysics, vol 48. Elsevier, London, pp 365–419
16. Chapman CH (2004) Fundamentals of seismic wave propagation. Cambridge University Press, Cambridge
17. CIG www.geodynamics.org. Accessed 1 Jul 2008
18. Cockburn B, Shu CW (1989) TVB Runge"Kutta local projection discontinuous Galerkin finite element method for conservation laws II: General framework. Math Comp 52:411–435
19. Cockburn B, Shu CW (1991) The Runge–Kutta local projection P1-Discontinuous Galerkin finite element method for scalar conservation laws. Math Model Numer Anal 25:337–361
20. Cockburn B, Shu CW (1998) The Runge–Kutta discontinuous Galerkin method for conservation laws V: Multidimensional systems. J Comput Phys 141:199–224
21. Cockburn B, Lin SY, Shu CW (1989) TVB Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws III: One dimensional systems. J Comput Phys 84:90–113
22. Cockburn B, Hou S, Shu CW (1990) The Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws IV: The multidimensional case. Math Comp 54:545–581
23. Cockburn B, Karniadakis GE, Shu CW (2000) Discontinuous Galerkin methods, theory, computation and applications. LNCSE, vol 11. Springer, New York
24. Courant R, Friedrichs KO, Lewy H (1928) Über die partiellen Differenzialgleichungen der mathematischen Physik. Mathematische Annalen 100:32–74
25. CUBIT cubit.sandia.gov. Accessed 1 Jul 2008
26. Dablain MA (1986) The application of high-order differencing to the scalar wave equation. Geophysics 51:54–66
27. De Cougny HL, Shephard MS (1999) Parallel refinement and coarsening of tetrahedral meshes. Int J Numer Methods Eng 46:1101–1125
28. de la Puente J, Dumbser M, Käser M, Igel H (2007) Discontinuous Galerkin methods for wave propagation in poroelastic media. to appear in Geophysics
29. de la Puente J, Käser M, Dumbser M, Igel H (2007) An arbitrary high order discontinuous Galerkin method for elastic waves on unstructured meshes IV: Anisotropy. Geophys J Int 169(3):1210–1228
30. di Prisco C, Stupazzini M, Zambelli C (2007) Non-linear SEM numerical analyses of dry dense sand specimens under rapid and dynamic loading. Int J Numer Anal Methods Geomech 31(6):757–788
31. Dormy E, Tarantola A (1995) Numerical simulation of elastic wave propagation using a finite volume method. J Geophys Res 100(B2):2123–2134
32. Dumbser M (2005) Arbitrary high order schemes for the solution of hyperbolic conservation laws in complex domains. Shaker, Aachen
33. Dumbser M, Käser M (2006) An arbitrary high order discontinuous galerkin method for elastic waves on unstructured meshes II: The three-dimensional isotropic case. Geophys J Int 167:319–336
34. Dumbser M, Käser M (2007) Arbitrary high order non-oscillatory finite volume schemes on unstructured meshes for linear hyperbolic systems. J Comput Phys 221:693–723. doi:10.1016/j.jcp.2006.06.043
35. Dumbser M, Munz CD (2005) Arbitrary high order discontinuous Galerkin schemes. In: Cordier S, Goudon T, Gutnic M, Sonnendrucker E (eds) Numerical methods for hyperbolic and kinetic problems. IRMA series in mathematics and theoretical physics. EMS Publishing, Zurich, pp 295–333
36. Dumbser M, Käser M, Toro EF (2007) An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes – V. Local time stepping and p-adaptivity. Geophys J Int 171:695–717
37. Dziewonski AM, Anderson DL (1981) Preliminary reference earth model. Phys Earth Planet Inter 25:297–356
38. Ewald M, Igel H, Hinzen K-G, Scherbaum F (2006) Basin-related effects on ground motion for earthquake scenarios in the lower rhine embayment. Geophys J Int 166:197–212
39. Faccioli E, Maggio F, Quarteroni A, Tagliani A (1996) Spectral-domain decomposition methods for the solution of acoustic and elastic wave equation. Geophysics 61:1160–1174
40. Faccioli E, Maggio F, Paolucci R, Quarteroni A (1997) 2D and 3D elastic wave propagation by a pseudo-spectral domain decomposition method. J Seismol 1:237–251
41. Faccioli E, Vanini M, Paolucci R, Stupazzini M (2005) Comment on "Domain reduction method for three-dimensional earthquake modeling in localized regions, part I: Theory." by Bielak J, Loukakis K, Hisada Y, Yoshimura C, and "Part II: Verification and Applications." by Yoshimura C, Bielak J, Hisada Y, Fernández A. Bull Seism Soc Am 95:763–769
42. Falk J, Tessmer E, Gajewski D (1996) Efficient finite-difference modelling of seismic waves using locally adjustable time steps. Geophys Prosp 46:603–616

43. Falk J, Tessmer E, Gajewski D (1996) Tube wave modelling by the finite differences method with varying grid spacing. Pure Appl Geoph 148:77–93
44. Fernandez A, Bielak J, Prentice C (2006) Urban seismology; City effects on earthquake ground motion and effects of spatial distribution of ground motion on structural response paper presented at 2006 annual meeting. Seism Res Lett 77(2):305
45. Fornberg B (1996) A practical guide to pseudospectral methods. Cambridge University Press, Cambridge
46. Fuchs K, Müller G (1971) Computation of synthetic seismograms with the reflectivity method and comparison with observations. Geophys J Royal Astronom Soc 23(4):417–33
47. Furumura T, Takenaka H (1996) 2.5-D modeling of elastic waves using the pseudospectral method. Geophys J Int 124: 820–832
48. Geller RJ, Takeuchi N (1998) Optimally accurate second-order time-domain finite difference scheme for the elastic equation of motion: One-dimensional case. Geophys J Int 135:48–62
49. Graves RW (1993) Modeling three-dimensional site response effects in the Marina district basin, San Francisco, California. Bull Seism Soc Am 83:1042–1063
50. Hestholm SO, Ruud BO (1998) 3-D finite-difference elastic wave modeling including surface topography. Geophysics 63:613–622
51. Holberg O (1987) Computational aspects of the choice of operator and sampling interval for numerical differentiation in large-scale simulation of wave phenomena. Geophys Prospect 35:629–655
52. Igel H (1999) Wave propagation through 3-D spherical sections using the Chebyshev spectral method. Geop J Int 136:559–567
53. Igel H, Gudmundsson O (1997) Frequency-dependent effects on travel times and waveforms of long-period S and SS waves. Phys Earth Planet Inter 104:229–246
54. Igel H, Weber M (1995) SH-wave propagation in the whole mantle using high-order finite differences. Geophys Res Lett 22(6):731–734
55. Igel H, Weber M (1996) P-SV wave propagation in the Earth's mantle using finite-differences: Application to heterogeneous lowermost mantle structure. Geophys Res Lett 23:415–418
56. Igel H, Mora P, Riollet B (1995) Anisotropic wave propagation through finite-difference grids. Geophysics 60:1203–1216
57. Igel H, Nissen-Meyer T, Jahnke G (2001) Wave propagation in 3-D spherical sections: Effects of subduction zones. Phys Earth Planet Inter 132:219–234
58. Jahnke G, Igel H, Cochard A, Thorne M (2007) Parallel implementation of axisymmetric SH wave propagation in spherical geometry. Geophys J Int (in print)
59. Jastram C, Tessmer E (1994) Elastic modelling on a grid with vertically varying spacing. Geophys Prosp 42:357–370
60. Karypis G, Kumar V (1998) Multilevel k-way Partitioning Scheme for Irregular Graphs. J Parallel Distrib Comput 48(1):96–129
61. Kelly KR, Ward RW, Treitel S, Alford RM (1976) Synthetic seismograms: A finite-difference approach. Geophysics 41: 2–27
62. Kennett BLN (2002) The seismic wavefield, vol I + II. Cambridge University Press, Cambridge
63. Komatitsch D, Tromp J (2002) Spectral-element simulations of global seismic wave propagation, part I: Validation. Geophys J Int 149:390–412
64. Komatitsch D, Tromp J (2002) Spectral-element simulations of global seismic wave propagation, part II: 3-D models, oceans, rotation, and gravity. Geophys J Int 150:303–318
65. Komatitsch D, Vilotte JP (1998) The spectral-element method: An efficient tool to simulate the seismic response of 2D and 3D geological structures. Bull Seism Soc Am 88:368–392
66. Komatitsch D, Coutel F, Mora P (1996) Tensorial formulation of the wave equation for modelling curved interfaces. Geophys J Int 127(1):156–168
67. Kosloff D, Baysal E (1982) Forward modeling by a fourier method. Geophysics 47(10):1402–1412
68. Krishnan S, Ji C, Komatitsch D, Tromp J (2006) Case studies of damage to tall steel moment-frame buildings in Southern California during large San Andreas earthquakes. Bull Seismol Soc Am 96(4A):1523–1537
69. Krishnan S, Ji C, Komatitsch D, Tromp J (2006) Performance of two 18-story steel moment-frame buildings in Southern California during two large simulated San Andreas earthquakes. Earthq Spectra 22(4):1035–106
70. Käser M, Dumbser M (2006) An arbitrary high order discontinuous Galerkin method for elastic waves on unstructured meshes I: The two-dimensional isotropic case with external source terms. Geophys J Int 166:855–877
71. Käser M, Dumbser M, de la Puente J, Igel H (2007) An arbitrary high order discontinuous Galerkin method for elastic waves on unstructured meshes III: Viscoelastic attenuation. Geophys J Int 168(1):224–242
72. Käser M, Igel H (2001) Numerical simulation of 2D wave propagation on unstructured grids using explicit differential operators. Geophys Prospect 49(5):607–619
73. Käser M, Igel H, Sambridge M, Braun J (2001) A comparative study of explicit differential operators on arbitrary grids. J Comput Acoust 9(3):1111–1125
74. Kwak D-Y, Im Y-T (2002) Remeshing for metal forming simulations – part II: Three dimensional hexahedral mesh generation. Int J Numer Methods Eng 53:2501–2528
75. LeVeque RL (2002) Finite volume methods for hyperbolic problems. Cambridge University Press, Cambridge
76. Levander AR (1988) Fourth-order finite-difference P-SV seismograms. Geophysics 53:1425–1436
77. Madariaga R (1976) Dynamics of an expanding circular fault. Bull Seismol Soc Am 66(3):639–66
78. Magnier S-A, Mora P, Tarantola A (1994) Finite differences on minimal grids. Geophysics 59:1435–1443
79. Marfurt KJ (1984) Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations. Geophysics 49:533–549
80. Mercerat ED, Vilotte JP, Sanchez-Sesma FJ (2006) Triangular spectral element simulation of two-dimensional elastic wave propagation using unstructured triangular grids. Geophys J Int 166(2):679–698
81. METIS glaros.dtc.umn.edu/gkhome/views/metis. Accessed 1 Jul 2008
82. Moczo P (1989) Finite-difference techniques for SH-waves in 2-D media using irregular grids – Application to the seismic response problem. Geophys J Int 99:321–329

83. Moczo P, Kristek J, Halada L (2000) 3D 4th-order staggered grid finite-difference schemes: Stability and grid dispersion. Bull Seism Soc Am 90:587–603
84. Montelli R, Nolet G, Dahlen FA, Masters G, Engdahl ER, Hung S (2004) Finite-frequency tomography reveals a variety of plumes in the mantle. Science 303(5656):338–343
85. Müller G (1977) Earth-flattening approximation for body waves derived from geometric ray theory – improvements, corrections and range of applicability. J Geophys 42:429–436
86. Nissen-Meyer T, Fournier A, Dahlen FA (2007) A 2-D spectral-element method for computing spherical-earth seismograms – I. Moment-tensor source. Geophys J Int 168:1067–1092
87. Ohminato T, Chouet BA (1997) A free-surface boundary condition for including 3D topography in the finite-difference method. Bull Seism Soc Am 87:494–515
88. Opršal I, J Zahradník (1999) Elastic finite-difference method for irregular grids. Geophysics 64:240–250
89. Pitarka A (1999) 3D elastic finite-difference modeling of seismic motion using staggered grids with nonuniform spacing. Bull Seism Soc Am 89:54–68
90. Priolo E, Carcione JM, Seriani G (1996) Numerical simulation of interface waves by high-order spectral modeling techniques. J Acoust Soc Am 95:681–693
91. Reed WH, Hill TR (1973) Triangular mesh methods for the neutron transport equation. Technical Report, LA-UR-73-479, Los Alamos Scientific Laboratory
92. Ripperger J, Igel H, Wassermann J (2004) Seismic wave simulation in the presence of real volcano topography. J Volcanol Geotherm Res 128:31–44
93. Scandella L (2007) Numerical evaluation of transient ground strains for the seismic response analyses of underground structures. Ph D Thesis, Milan University of Technology, Milan
94. SCEC www.scec.org. Accessed 1 Jul 2008
95. Schneiders R (2000) Octree-Based Hexahedral Mesh Generation. Int J Comput Geom Appl 10(4):383–398
96. Schwartzkopff T, Munz CD, Toro EF (2002) ADER: A high-order approach for linear hyperbolic systems in 2D. J Sci Comput 17:231–240
97. Schwartzkopff T, Dumbser M, Munz CD (2004) Fast high order ADER schemes for linear hyperbolic equations. J Comput Phys 197:532–539
98. Seriani G, Priolo E, Carcione JM, Padovani E (1992) High-order spectral element method for elastic wave modeling: 62nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1285–1288
99. Shepherd JF (2007) Topologic and geometric constraint-based hexahedral mesh generation. Ph.D. Thesis on Computer Science, School of Computing The Universiy of Utah, Salt Lake City
100. Sieminski A, Liu Q, Trampert J, Tromp J (2007) Finite-frequency sensitivity of surface waves to anisotropy based upon adjoint methods. Geophys J Int 168:1153–1174
101. SPICE www.spice-rtn.org. Accessed 1 Jul 2008
102. Stupazzini M (2004) A spectral element approach for 3D dynamic soil-structure interaction problems. Ph D Thesis, Milan University of Technology, Milan
103. Takeuchi N, Geller RJ (2000) Optimally accurate second order time-domain finite difference scheme for computing synthetic seismograms in 2-D and 3-D media. Phys Earth Planet Int 119:99–131
104. Tape C, Liu Q, Tromp J (2007) Finite-frequency tomography using adjoint methods: Methodology and examples using membrane surface waves. Geophys J Int 168:1105–1129
105. Tarantola A (1986) A strategy for nonlinear elastic inversion of seismic reflection data. Geophysics 51(10):1893–1903
106. Tessmer E (2000) Seismic finite-difference modeling with spatially varying time teps. Geophysics 65:1290–1293
107. Tessmer K, Kosloff D (1996) 3-D elastic modeling with surface topography by a Chebychev spectral method. Geophysics 59:464–473
108. Tessmer E, Kessler D, Kosloff K, Behle A (1996) Multi-domain Chebyshev–Fourier method for the solution of the equations of motion of dynamic elasticity. J Comput Phys 100:355–363
109. Thomas C, Igel H, Weber M, Scherbaum F (2000) Acoustic simulation of P-wave propagation in a heterogeneous spherical earth: Numerical method and application to precursor energy to PKPdf. Geophys J Int 141:307–320
110. Thorne M, Lay T, Garnero E, Jahnke G, Igel H (2007) 3-D seismic imaging of the D″ region beneath the Cocos Plate. Geophys J Int 170:635–648
111. Titarev VA, Toro EF (2002) ADER: Arbitrary high order Godunov approach. J Sci Comput 17:609–618
112. Toro EF (1999) Riemann solvers and numerical methods for fluid dynamics. Springer, Berlin
113. Toro EF, Millington AC, Nejad LA (2001) Towards very high order Godunov schemes, in Godunov methods; Theory and applications. Kluwer/Plenum, Oxford, pp 907–940
114. Toyokuni G, Takenaka H, Wang Y, Kennett BLN (2005) Quasi-spherical approach for seismic wave modeling in a 2-D slice of a global earth model with lateral heterogeneity. Geophys Res Lett 32:L09305
115. Van der Hilst RD (2004) Changing views on Earth's deep mantle. Science 306:817–818
116. Virieux J (1984) SH-wave propagation in heterogeneous media: Velocity-stress inite-difference method. Geophysics 49:1933–1957
117. Virieux J (1986) P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method. Geophysics 51:889–901
118. Woodhouse JH, Dziewonski AM (1984) Mapping the upper mantle: Three dimensional modelling of earth structure by inversion of seismic waveforms. J Geophys Res 89:5953–5986
119. Yoshimura C, Bielak J, Hisada Y, Fernández A (2003) Domain reduction method for three-dimensional earthquake modeling in localized regions, part II: Verification and applications. Bull Seism Soc Am 93:825–841
120. Zambelli C (2006) Experimental and theoretical analysis of the mechanical behaviour of cohesionless soils under cyclic-dynamic loading. Ph D Thesis, Milan University of Technology, Milan
121. Zienckiewicz O, Taylor RL (1989) The finite element method, vol 1. McGraw-Hill, London

Books and Reviews

Carcione JM, Herman GC, ten Kroode APE (2002) Seismic modelling. Geophysics 67:1304–1325

Mozco P, Kristek J, Halada L (2004) The finite-difference method for seismologists: An introduction. Comenius University, Bratislava. Available in pdf format at ftp://ftp.nuquake.eu/pub/Papers

Moczo P, Kristek J, Galis M, Pazak P, Balazovjech M (2007) The finite difference and finite-element modelling of seismic wave propagation and earthquake motion. Acta Physica Slovaca, 57(2)177–406

Wu RS, Maupin V (eds) (2006) Advances in wave propagation in heterogeneous earth. In: Dmowska R (ed) Advances in geophysics, vol 48. Academic/Elsevier, London

Seismic Waves in Heterogeneous Earth, Scattering of

HARUO SATO
Department of Geophysics, Graduate School of Science, Tohoku University, Sendai-shi, Miyagi-ken, Japan

Article Outline

Glossary
Definition of the Subject
Introduction
Radiative Transfer Theory for a Scattering Medium
Wave Envelopes in Random Media and Statistical Characterization
Envelope Broadening of a High-Frequency Seismogram
Spatial Variation of Scattering Characteristics
Temporal Change in the Earth Medium Structure
Future Directions
Acknowledgments
Bibliography

Glossary

Attenuation factor Q^{-1} A measure of attenuation characteristics of a medium caused by intrinsic absorption and scattering loss. The former means the transfer of vibration energy into heat and the latter means the transfer of vibration energy from the direct wave to coda waves caused by scattering due to medium heterogeneity.

Coda waves Wave trains that follow the arrival of the direct S-wave phase are called S-coda waves or simply coda waves. Coda waves are interpreted as a superposition of S waves scattered by distributed heterogeneities. Wave trains between direct P and S wave arrivals are called P-coda waves.

Coda attenuation factor Q_C^{-1} This parameter characterizes the amplitude decay of S coda of a local earthquake with the lapse time increasing based on the S-to-S single scattering. The coda duration shortens for a larger coda attenuation factor.

Envelope broadening The source duration time of a microearthquake is short; however, the apparent duration time of the S-wave seismogram increases with the travel distance increasing because of diffraction and scattering by medium heterogeneities. This phenomenon is called envelope broadening.

Radiative transfer theory A phenomenological theory that describes scattering process of wave energy in a scattering medium on the basis of causality, geometrical spreading and the energy conservation. It neglects the interference of waves but focuses on the intensity only. This theory admits various types of scattering patterns. It is often applied to model the energy propagation of high-frequency seismic-waves in heterogeneous Earth media.

Random media A mathematical model for media whose parameters are described by random functions of space coordinates. The stochastic properties of the ensemble of random media are characterized by their autocorrelation function or the power spectral density function.

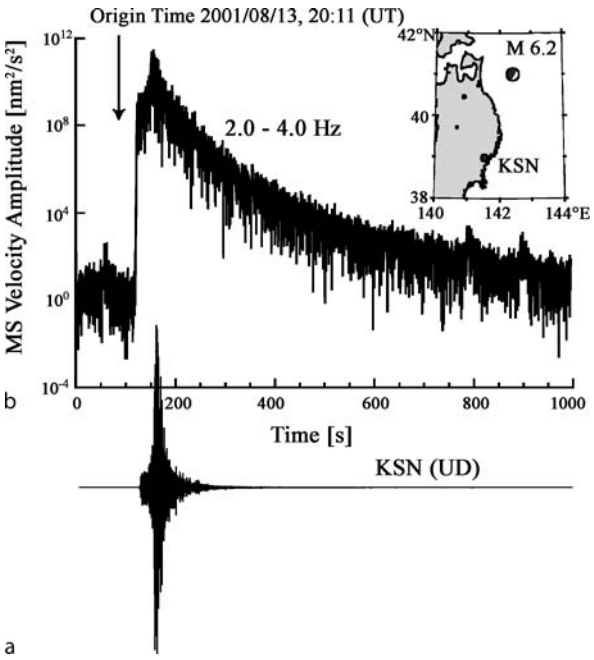
Scattering coefficient g A measure of the scattering power in a unit solid angle at a certain direction by a unit volume of heterogeneous media for the incidence of unit energy flux density. The average of g over the solid angle gives the total scattering coefficient g_0, of which the reciprocal gives the mean free path. This quantity characterizes the coda excitation and the scattering loss in the heterogeneous media.

Definition of the Subject

The structure of the solid Earth was extensively studied by using seismic waves such as travel time analysis based on Snell's law, dispersion analysis of surface waves, and spectral analysis of free oscillation, where the notion of a horizontally stratified structure or a spherical shell structure prevailed among the geophysical community. This means the acceptance of the dominance of gravity in geodynamic process. Velocity tomography revealed that the solid Earth structure is three-dimensionally inhomogeneous with various ranges of scales; however, the resolution of velocity tomography is much coarser than the wavelength of seismic waves. In 1970s, the existence of distributed inhomogeneities having the order of the wavelength of seismic waves was recognized from the observation of coda waves of local earthquakes, which are long-lasting wave trains following the direct S-wave arrival in high-frequency seismograms. Here, we use "high-frequency" for frequency higher than about 1 Hz. The long duration time of coda waves can be interpreted as a direct evidence

of wide-angle scattering caused by distributed small-scale heterogeneities since the source duration time is generally very short. S-wave seismograms of microearthquakes show broadened envelopes with travel distance increasing. This envelope broadening phenomenon is also an evidence of scattering around the forward direction due to random velocity inhomogeneities.

Since then, focusing on the frequency dependence of seismogram envelopes, geophysicists have extensively studied the scattering process of high-frequency seismic waves in relation to the spectral structure of velocity inhomogeneities, where the statistical characterization of the medium inhomogeneity is inevitable. The radiative transfer theory and the stochastic Markov approximation have been developed as mathematical tools for the analyzes of seismogram envelopes. The strength of scattering and/or the spectral structure of random inhomogeneity have been measured in various regions of the solid Earth. The scattering approach is found to be also useful for detecting temporal changes in the crustal medium associated with earthquake occurrences. Thus, scattering of high-frequency seismic waves in the heterogeneous Earth medium is important for understanding the physical structure and the geodynamic process reflecting the evolution of the solid Earth.



Seismic Waves in Heterogeneous Earth, Scattering of, Figure 1
a Seismogram of a local earthquake of M 6.2 in northeastern Honshu, Japan recorded by F-net, NIED. b Bandpass-filtered MS envelope (Courtesy of T. Maeda)

Introduction

Coda Waves

The high-frequency seismogram of a local earthquake has a long tail after the direct S-coda arrival. The tail portion of seismogram is called "S-coda waves" or simply "coda". As an example, Fig. 1a and b show the raw seismogram and the band-pass filtered mean square (MS) trace of an earthquake of magnitude (M) 6.1, respectively. We note that the mean square wave envelope, which is the running mean of the squared trace with characteristic time of a few times the center period, is proportional to the time trace of the wave energy density. The coda wave oscillation lasts more than several hundreds of seconds. The duration of coda waves measured from the P-wave onset until when the coda amplitude decreases to the microseism's level has been used as a quick measure of the earthquake magnitude from a single station observation since the 1960s. Having a motivation to extract the source spectrum of a large earthquake from clipped seismograms, Aki [1] first studied the characteristics of coda waves as scattered waves. Coda envelopes of a local earthquake have a smoothly decaying common curve with lapse time increasing irrespective of epicentral distances and the source radiation pattern. Aki and Chouet [4] interpreted coda waves as single back-scattered S-waves due to heterogeneities randomly distributed in the lithosphere. Their model based on the radar equation for the same location of a source and a receiver can be written as follows.

Point-like isotropic scatterers characterized by total scattering cross-section σ_0 are randomly and homogeneously distributed with number density n in a 3-D medium with background wave velocity V_0. The scattering power per unit volume is characterized by the total scattering coefficient $g_0 = n\sigma_0$, of which the reciprocal gives the mean free path. When the total wave energy W is impulsively radiated from a point source at time $t = 0$, the wave energy density of singly back-scattered waves at the source location is written as

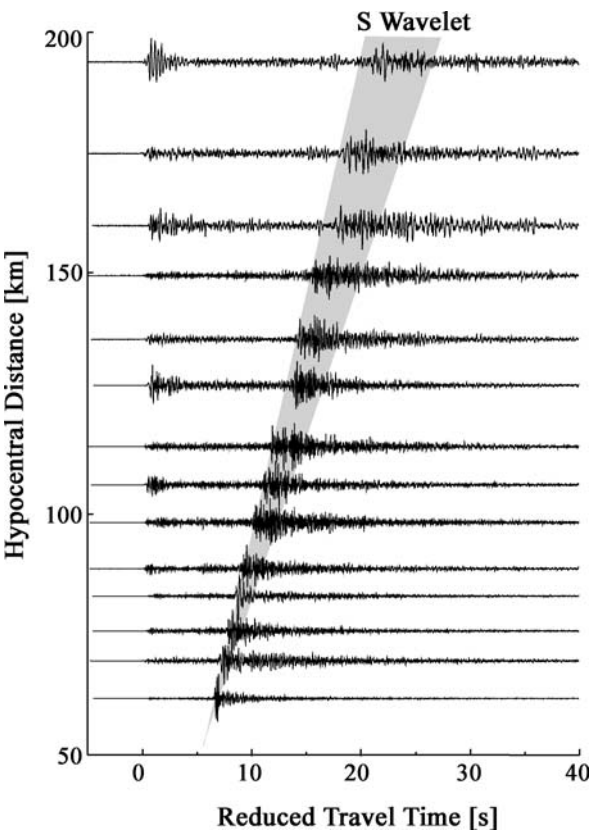
$$E^{SB}(t) \approx \frac{W g_0}{2\pi V_0^2 t^2} e^{-Q_C^{-1} 2\pi f t} \quad (1)$$

since the interference of scattered waves can be neglected because of the random distribution of scatterers. The inverse square of lapse time means geometrical spreading in a 3-D space. Here, an exponential damping term with coda attenuation factor Q_C^{-1} is introduced to represent phenomenological attenuation effect. This simple formula has been widely used for measurements of g_0 and Q_C^{-1} for S-waves in the world since the 1980s. Reported g_0-values

are of the order of 0.01 km^{-1} for 1–30 Hz and Q_C^{-1} values are about 10^{-2} at 1 Hz and decrease to about 10^{-3} at 20 Hz in the lithosphere [74].

Envelope Broadening of S-Seismogram

There is another evidence of scattering due to random inhomogeneity in high-frequency seismograms. Observed S-seismograms of a microearthquake have broadened envelopes around their peaks after the direct arrivals. As shown by an example in Fig. 2, the apparent duration time of the S-seismogram just after the direct S-arrival increases with increasing travel distance. It is more than ten seconds at distances larger than 100 km, where the source duration time is less than one second for an earthquake of M 4.0. Sato [70] called this phenomenon observed in an island arc as "envelope broadening", and Atkinson [6] reported similar phenomenon in a continent. For P-waves of teleseismic events, broadening of the vertical component envelope [35] and the excitation in the transverse component amplitude [52] have been used as a measure of lithospheric heterogeneity. These phenomena can be interpreted by multiple scattering within a narrow angle around the global ray direction due to random velocity inhomogeneities. When the wavelength is much shorter than the characteristic scale of the random velocity inhomogeneity, the scattering process of waves can be represented by successive ray bending processes, where scattering angles are statistically controlled by the spectrum of random velocity inhomogeneity. At a given distance from the source, a small number of rays with large scattering angles arrive long after the direct ray.



Seismic Waves in Heterogeneous Earth, Scattering of, Figure 2
Envelope broadening shown in horizontal component seismograms of a microearthquake with M 4.0 in Japan recorded by Hi-net, NIED, where the abscissa is reduced travel time with move-out velocity 7 km/s (Courtesy of T. Takahashi)

Radiative Transfer Theory for a Scattering Medium

Radiative Transfer Integral Equation for the Isotropic Scattering Process

Disregarding wave interference and focusing on wave power, the radiative transfer theory [8] treats the propagation of wave energy in a scattering medium. Wu [85] first introduced the radiative transfer theory for the stationary state in the synthesis of seismogram envelopes. The non-stationary multiple isotropic scattering process in 1-D was solved by Hemmer [24] and that in 2-D was solved by Shang and Gao [77]. Later, Zeng et al. [92] formulated the time-dependent multiple isotropic scattering process in 3-D as an extension of the single backscattering model [4] as follows.

In a 3-D scattering medium characterized by background velocity V_0 and total scattering coefficient g_0, when the total wave energy W is impulsively radiated isotropically from a source at the origin, the multiple isotropic scattering process is written by the following integral equation for energy density:

$$E(\mathbf{x}, t) = WG_E(\mathbf{x}, t) + g_0 V_0 \cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G_E(\mathbf{x}-\mathbf{x}', t-t')\, E(\mathbf{x}', t')\, d\mathbf{x}' dt' , \tag{2}$$

where the convolution integral in the second term means the propagation of energy from the last scattering point $\mathbf{x}'$ to a receiver at $\mathbf{x}$. The first term is the ballistic term that means the direct propagation of energy from the source with scattering loss, $G_E(\mathbf{x}, t) = \delta(t - r/V_0)\exp(-g_0 V_0 t)/(4\pi V_0 r^2)$, where $r = |\mathbf{x}|$. The solution

is written as [92]

$$E(\mathbf{x}, t) = WG_{\mathrm{E}}(\mathbf{x}, t) + \frac{Wg_0 \mathrm{e}^{-g_0 V_0 t}}{4\pi r^2} \frac{r}{V_0 t} \cdot \ln\left[\frac{V_0 t + r}{V_0 t - r}\right] H\left(t - \frac{r}{V_0}\right) + Wg_0^2 V_0^2 \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathrm{d}\omega \mathrm{d}k \, \mathrm{e}^{-i\omega t - ikr} \cdot \frac{ik}{2\pi r} \frac{\bar{\bar{G}}_{\mathrm{E}}(-k, -i\omega)^3}{1 - g_0 V_0 \bar{\bar{G}}_{\mathrm{E}}(-k, -i\omega)} , \quad (3)$$

where $\bar{\bar{G}}_{\mathrm{E}}(k, s) = (1/kV_0) \tan^{-1} kV_0/(s + g_0 V_0)$ is the Fourier–Laplace transform of G_{E} with respect to coordinate and time, respectively. The second term represents the single scattering process (see Sato [67]), which has a logarithmic divergence at the direct arrival $t = r/V_0$ and decreases according to the inverse square of lapse time at long lapse times as $Wg_0/(2\pi V_0^2 t^2)$. The third term representing multiple scattering converges to a diffusion solution with lapse time increasing as

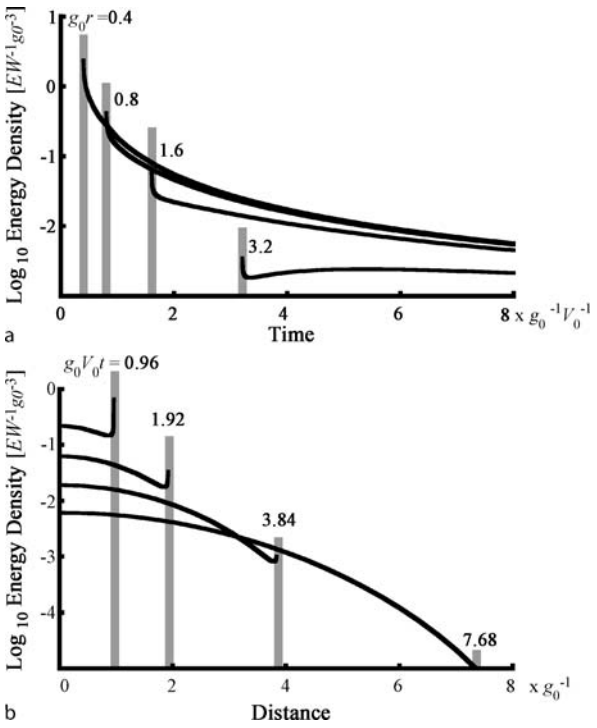
$$E^{\mathrm{Dif.}}(\mathbf{x}, t) = W \left(\frac{3g_0}{4\pi V_0 t}\right)^{3/2} \mathrm{e}^{-\frac{3g_0 r^2}{4V_0 t}} H(t) , \quad (4)$$

where the factor $V_0/(3g_0)$ is the diffusivity. Later, Paaschens [54] proposed an approximation as

$$E^{\mathrm{Paa.}}(\mathbf{x}, t) \approx WG_{\mathrm{E}}(\mathbf{x}, t) + \frac{W\mathrm{e}^{-g_0 V_0 t}}{\left(4\pi V_0 t/(3g_0)\right)^{3/2}} \cdot \left(1 - \frac{r^2}{V_0^2 t^2}\right)^{\frac{1}{8}} M\left(g_0 V_0 t \left(1 - \frac{r^2}{V_0^2 t^2}\right)^{\frac{3}{4}}\right) \cdot H\left(t - \frac{r}{V_0}\right) , \quad (5)$$

where $M(x) \approx \mathrm{e}^x \sqrt{1 + 2.026/x}$. The error of this approximation is of the order of 2% outside the ballistic peak and its tail for $g_0 V_0 t < 6$ and $2 < g_0 r < 4$.

Figure 3 shows spatiotemporal variations in energy density in a scattering medium theoretically predicted by the approximation solution (5) for instantaneous spherical source radiation at the origin. Scattered energy density is shown by a black curve, where the ballistic term is shown by a vertical gray line. At a small distance from the source compared with the mean free path $1/g_0$, the energy density decreases rapidly after the direct arrival as predicted by the single scattering term; however, the decay rate becomes smaller due to multiple scattering according to the power of lapse time $t^{-3/2}$ at long lapse times.



Seismic Waves in Heterogeneous Earth, Scattering of, Figure 3
a Temporal change and b spatial variation of energy density in an isotropic-scattering medium for a point source radiation. Each *black curve* shows the scattering contribution predicted by the Paaschens approximation and each *vertical gray line* shows a ballistic term

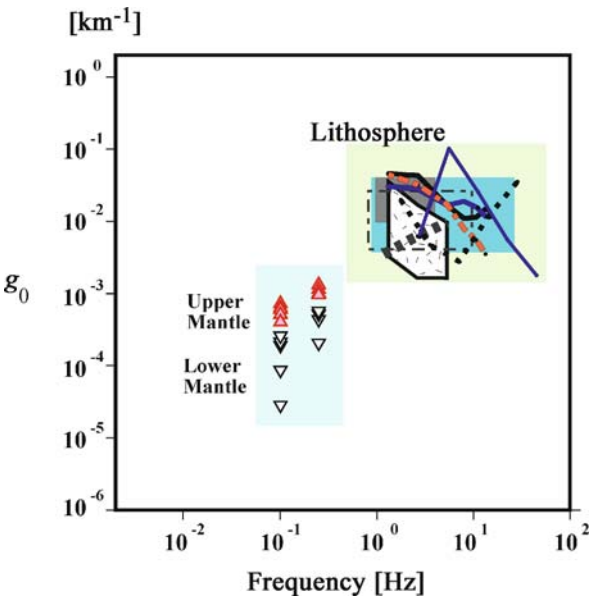
At a long distance, for example at $r = 3.2/g_0$, the energy density has an additional diffusion peak. The spatial distribution of scattered energy density is uniform around the source at a short lapse time compared with the mean free time $1/g_0 V_0$; however, it converges to a Gaussian curve at a long lapse time as theoretically predicted by the diffusion solution (4), for example at $t = 7.48/g_0 V_0$. There is no violation of causality since no signal exists beyond the ballistic peak. The smooth spatial distribution of scattered energy density around the source location gives the physical basis of the coda normalization method for measurements of S-wave attenuation and site amplification factors (e. g. [3,56,89]).

Gusev and Abubakirov [21] and Hoshiba [27] numerically solved the radiative transfer equation for the isotropic scattering process by using the Monte Carlo method. Yoshimoto [88] numerically simulated envelopes in scattering media of which the background velocity decreases with depth. He found the concentration of scattered energy near the surface because of seismic ray bending.

Nonisotropic radiation from a source can be easily introduced in the radiative transfer equation. Sato et al. [75] analytically solved the case of double couple source radiation: the energy density theoretically predicted faithfully reflects the source radiation pattern near the direct arrival; however, the azimuthal dependence diminishes with increasing lapse time. It qualitatively agrees with the observed radiation pattern independence of coda amplitudes at long lapse times. Their solution has been used in the envelope inversion of strong motion records for the spatial distribution of high-frequency wave energy radiation from an earthquake fault (e. g. [47]).

Measurements of Total Scattering Coefficient and Attenuation

For the practical application of the radiative transfer theory to observed seismograms, it is necessary to introduce intrinsic absorption Q_{Int}^{-1} by multiplying an exponential temporal decay factor $\exp[-Q_{\mathrm{Int}}^{-1} 2\pi f t]$ to the resultant energy density. By using the solution (3) of the radiative transfer theory for the isotropic scattering model, total scattering coefficient g_0 and intrinsic absorption factor $Q_{\mathrm{Int.}}^{-1}$ of the S-wave have been measured. Reported g_0 values in the lithosphere are of the order of $10^{-2}\,\mathrm{km}^{-1}$ for frequencies from 1 to 30 Hz as plotted in Fig. 4.



Seismic Waves in Heterogeneous Earth, Scattering of, Figure 4
Total scattering coefficient of S-waves in the Earth. Measurements in the mantle [40] are added to lithospheric inhomogeneity [74]

From the observed lapse time dependence of Q_{C}^{-1}, Gusev [20] quantitatively explained the decrease of g_0 with depth. Lee et al. [40] analyzed coda envelopes of regional earthquakes before and after the ScS arrival around 900s in lapse time from the origin time based on the numerically simulated envelopes for the PREM model, which is characterized by depth-dependent background velocity and total attenuation for S-waves. They reported lower g_0 values in 4s and 10s period bands in the upper and lower mantle compared with those in the lithosphere as illustrated in Fig. 4.

Hoshiba et al. [28] and Fehler et al. [10] developed a method to measure simultaneously g_0 and Q_{Int}^{-1} values for the S-wave from the whole S-envelope analysis based on the synthetic envelope derived from the radiative transfer theory. Their multiple lapse-time window analysis method has been widely used in the world. Estimated scattering loss $g_0 V_0/\omega$ decreases with frequency; however, intrinsic absorption Q_{Int}^{-1} is rather insensitive to frequency. Estimated seismic albedo $B \equiv g_0 V_0 / \left(\omega Q_{\mathrm{Int}}^{-1} + g_0 V_0\right)$, the ratio of scattering loss to the total attenuation, of S-waves in the lithosphere widely distribute from 0.2 to 0.8 for 1–6 Hz, but they are limited between 0.2 and 0.5 for 6–20 Hz.

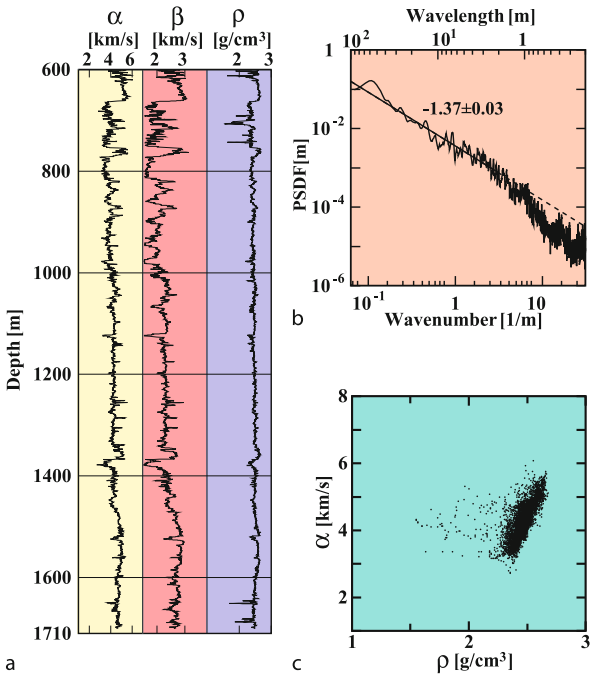
Total scattering cross-section and number density of scatterers appear jointly as the total scattering coefficient in theoretical models; however, Matsumoto [44] proposed a method to separate them from the temporal variation of the semblance coefficient of coda waves recorded by a seismic array. Analyzing data obtained in the aftershock area of the 2000 western Tottori earthquake, Japan, he estimated $n = 0.03\,\mathrm{km}^{-3}$ and $g_0 = 0.001\,\mathrm{km}^{-1}$ at 20 Hz.

It should be noted that lunar seismograms have coda durations exceeding one hour. Applying the diffusion model for the explanation of spindle-like envelopes of lunar-quakes, Dainty and Toksöz [9] estimated g_0 to be as large as $0.05–0.5\,\mathrm{km}^{-1}$ at 0.45 Hz.

Wave Envelopes in Random Media and Statistical Characterization

Statistical Characterization of Random Media

As revealed from tomography analyses (e. g. [93]), velocity structure is three-dimensionally inhomogeneous especially in the lithosphere. Well log data show typical samples of the shallow crust. Figure 5a shows well log data of P- and S-wave velocities and mass density obtained in Kyushu, Japan [80], where wave velocities and mass density are measured in a borehole by using ultrasonic waves and gamma rays, respectively. These well log data clearly show random fluctuation with short wavelengths.



Seismic Waves in Heterogeneous Earth, Scattering of, Figure 5
a Well log data at the YT-2 site, Kyushu, Japan. b Power spectral density function of the fractional fluctuation of P-wave velocity log. c Scattergram of P-wave velocity and mass density. Reproduced from [80]

As a natural consequence, we imagine random inhomogeneities widely distributed in the solid Earth medium.

By using computer power, wave propagation in random media has been numerically studied extensively. Using finite difference simulations of waves in random media, Frankel and Clayton [14] first examined the relation between coda excitation and scattering loss and the spectrum of random inhomogeneity. On the basis of numerical simulations, Frankel and Wennerberg [15] proposed the energy flux model that has a uniform distribution of scattered energy behind the direct waves for the analysis of high-frequency seismogram envelopes. Using a boundary integral method for the simulation of waves in a medium containing many cavities, Yomogida and Benites [87] examined a relation between coda attenuation and the distribution of cavities.

There is an alternative approach to treat statistically randomly inhomogeneous media. The wave-velocity is written as $V(\mathbf{x}) = V_0\{1 + \xi(\mathbf{x})\}$, where V_0 is the average velocity and fractional fluctuation $\xi(\mathbf{x})$ is a homogeneous and isotropic random function of space coordinate $\mathbf{x}$. We imagine an ensemble of random media $\{\xi\}$, which is statistically characterized by the autocorrelation function (ACF) $R(\mathbf{x}) \equiv \langle \xi(\mathbf{x}+\mathbf{x}')\,\xi(\mathbf{x}') \rangle$, where angular brackets mean the ensemble average. The MS fractional fluctuation $\varepsilon^2 \equiv R(0)$ and the correlation distance a are key parameters. The Fourier transform of ACF gives the power spectral density function (PSDF) P. The PSDF of the P-wave velocity fractional fluctuation of well log data shows a power-law characteristic at large wavenumbers as illustrated in Fig. 5b. P-wave velocity and mass density show a good correlation as shown in Fig. 5c. The statistical view is useful for representing geological data, too (e. g. [18,26]).

Scattering Coefficient Based on the Born Approximation

When the medium inhomogeneity is small $|\xi| \ll 1$, scalar wave ϕ is governed by the following wave equation:

$$\left(\Delta - \frac{1}{V_0^2}\partial_t^2\right)\phi + \frac{2}{V_0^2}\xi(\mathbf{x})\partial_t^2\phi = 0\,. \tag{6}$$

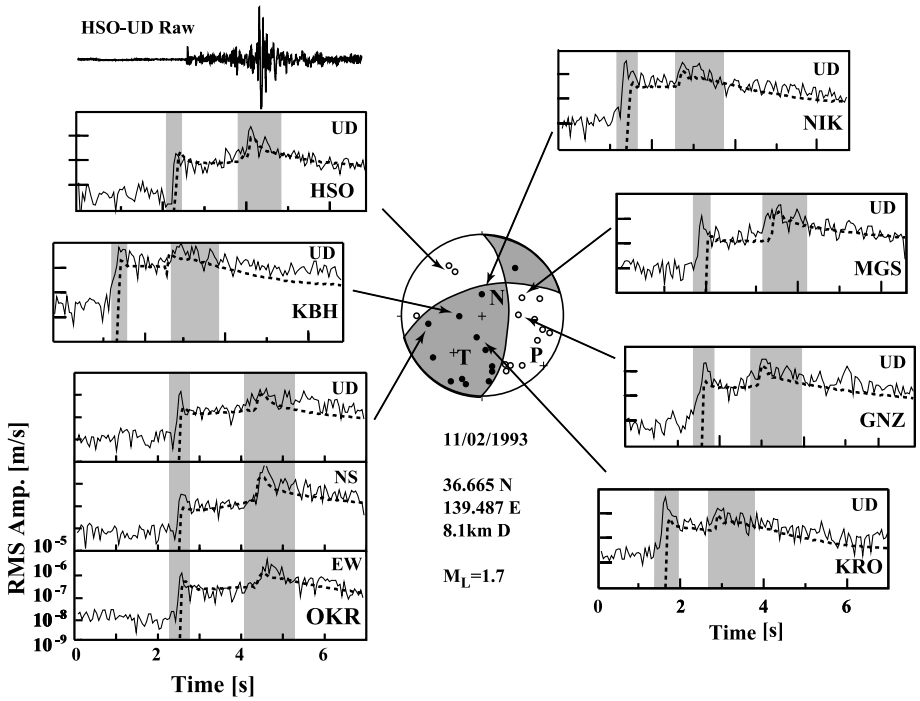
Velocity inhomogeneity is supposed to localize in a volume around the origin, of which the dimension is chosen to be much larger than a. For the incidence of a plane wave of unit amplitude at angular frequency ω as $e^{i(k_0\mathbf{e_z}\mathbf{x}-\omega t)}$, where $\mathbf{e_z}$ is the unit vector to the z direction, we calculate the spherically outgoing scattered waves due to a localized inhomogeneity by using the Born approximation as $\phi^1(\mathbf{x}, t) = -k_0^2 e^{i(k_0 r-\omega t)}\tilde{\xi}(k_0\mathbf{e_r} - k_0\mathbf{e_z})/(2\pi r)$, where the tilde means the Fourier transform with respect to coordinates in 3-D space and $\mathbf{e_r}$ is a radial unit vector (e. g. [74]). According to Aki and Chouet [4], the scattering coefficient defined as the scattering power in a unit solid angle at certain direction by a unit volume of random inhomogeneous media for the incidence of unit energy flux density is statistically written by using its PSDF as

$$g(\psi;\omega) = \frac{k_0^4}{\pi}P(k_0\mathbf{e_r} - k_0\mathbf{e_3}) = \frac{k_0^4}{\pi}P\left(2k_0\sin\frac{\psi}{2}\right), \tag{7}$$

where ψ is the scattering angle measured from the z direction. This functional form means anisotropic scattering depending on frequency. In general, scattering near around the forward direction becomes larger with increasing wavenumber in random media.

Radiative Transfer Theory with Scattering Coefficients Calculated by the Born Approximation

Extending the above formulation to vector wave propagation in random elastic media, we can define scattering coefficients for different scattering modes as PP, PS, SP,



Seismic Waves in Heterogeneous Earth, Scattering of, Figure 6

RMS envelopes of a microearthquake in the shallow crust, Nikko in northern Kanto, Japan. *Fine curves* and *broken curves* are observed and best-fit synthesized envelopes, respectively, where *shades* show time windows used for the estimation of the source radiation energy. The trace on the *top left* shows a raw seismogram. Reproduced from [90]

and SS. For the case of random elastic media characterized by an exponential ACF with $\varepsilon = 10\%$ and $a = 2\,\mathrm{km}$, Sato [68] synthesized three-component seismogram envelopes of a microearthquake of M 3 as a superposition of polarized scattered waves' power at a finite distance from a point shear dislocation source. That is the single scattering approximation of the radiative transfer theory with scattering coefficients calculated by the Born approximation. The SS scattering mode dominates in S coda, and pseudo P and S waves are produced even at a receiver on the null direction of the source radiation. By using a von Kármán-type ACF for describing random elastic media, Sato [71] estimated parameters $\kappa = 0.35, \varepsilon = 8.4\%$ and $a = 2.1\,\mathrm{km}$ from observed frequency dependence of S-wave attenuation and g_0, where the parameter κ controls the role-off of PSDF at large wavenumbers. Extending the above vector wave envelope synthesis to include mode conversions at the free surface, Yoshimoto et al. [90] analyzed three-component seismogram envelopes of microearthquakes in the shallow crust in Nikko, northern Kanto, Japan. A raw seismogram is shown at the top-left of Fig. 6 as an example. Fine curves in Fig. 6 are logarithmic plots of observed root mean square (RMS) envelopes in the 2–16 Hz band. Broken curves are best-fit theoretical envelopes for random elastic media characterized by an exponential ACF with $\varepsilon = 5.7\%$ and $a = 400\,\mathrm{m}$. We find that the fitness is good not only for S coda but also for P coda.

Wave theory in random media predicts that the scattering coefficient has a large lobe in the forward direction in higher frequencies. Gusev and Abubakirov [22] used the Monte Carlo method to simulate envelopes for the multiple nonisotropic scattering process. There have been mathematical developments to derive the radiative transfer equation for multiple nonisotropic scattering from the stochastic averaging of the wave equation in random media (e. g. [13,29,41,64]). Przybilla et al. [58] showed an excellent coincidence of vector envelopes calculated from finite difference simulation in 2-D random elastic media and those synthesized by the radiative transfer theory with scattering amplitudes derived from the Born approximation and the wandering effect of travel time.

Interference of Scattered Waves

The interference effect is neglected in conventional studies of wave scattering in random media; however, it becomes important for a specific case even in random media. When

randomness is strong enough to produce multiple scattering, coda wave intensity at a receiver near a source is enhanced compared to the prediction of conventional radiative transfer theory. Margerin et al. [42] showed that a spot of backscattering enhancement stabilizes in a sphere of radius half a wavelength centered at the source after a transient regime. The enhancement persists in time and should be observable as long as a coda is measurable. From field experiment of seismic waves in a shallow volcanic structure Larose et al. [39] reported the existence of weak localization, where the size of enhancement spot was one wavelength and the estimated mean-free path was 200 m for seismic waves around 20 Hz.

Envelope Broadening of a High-Frequency Seismogram

Markov Approximation for Parabolic Wave Equation

For the study of light propagation through the upper atmosphere and/or acoustic sound propagation through internal waves in oceans, various stochastic methods have been developed in the fields of physics. One of the most attractive methods for explaining the wave envelope around the direct arrival is the Markov approximation for the parabolic wave equation, which is an extension of the phase screen method or the split step Fourier method (e. g. [29,63]). This method is found to be applicable to seismogram envelopes. We imagine an elastic medium composed of a homogeneous half space $z < 0$ and an inhomogeneous half space $z > 0$, where the inhomogeneity is supposed to be small ($\varepsilon^2 \ll 1$) and the randomness is statistically homogeneous and isotropic. When the wavelength is smaller than the correlation distance a, we may neglect conversion scattering between P and S waves. Then, we can describe the principal characteristics of vector wave propagation by using potentials. For the incidence of plane P-wavelet to the z direction from the homogeneous zone, scalar potential is written as a superposition of harmonic waves of angular frequency ω as $\phi = \int_{-\infty}^{\infty} (2\pi i k_0)^{-1} U(\mathbf{x}_\perp, z, \omega)\, \mathrm{e}^{ik_0 z - i\omega t}\, \mathrm{d}\omega$ for $z > 0$, where $\mathbf{x}_\perp = (x, y)$ on the transverse plane. Neglecting the second derivative with respect to z, we have the parabolic-type equation for U as

$$2ik_0 \partial_z U + \left(\partial_x^2 + \partial_y^2\right) U - 2k_0^2 \xi(\mathbf{x})\, U = 0\,. \tag{8}$$

We define the two-frequency mutual coherence function (TFMCF) of field U between two different locations on the transverse plane at a distance z and different angular frequencies at ω' and ω'' as $\Gamma_2(\mathbf{x}_{\perp c}, \mathbf{x}_{\perp d}, z, \omega_c, \omega_d) \equiv \langle U(\mathbf{x}'_\perp, z, \omega') U(\mathbf{x}''_\perp, z, \omega'')^* \rangle$, where ω_c and ω_d are center-of-mass and difference angular frequencies, respectively. In the case of quasi-monochromatic waves $|\omega_d| \ll |\omega_c|$, using causality and neglecting back scattering, we derive the master equation for TFMCF. This derivation is called the Markov approximation. For the ith component, the intensity is defined as the ensemble average of the square of displacement $\langle \partial_i \phi\, \partial_i \phi^* \rangle = 1/(2\pi) \int_{-\infty}^{\infty} \widehat{I}_i^{\mathrm{P}}\, \mathrm{d}\omega_c$. The integrand is the intensity spectral density (ISD) $\widehat{I}_i^{\mathrm{P}}$, which means the time trace of MS amplitude in a band having the central angular frequency ω_c.

Vector Envelopes for a Gaussian ACF

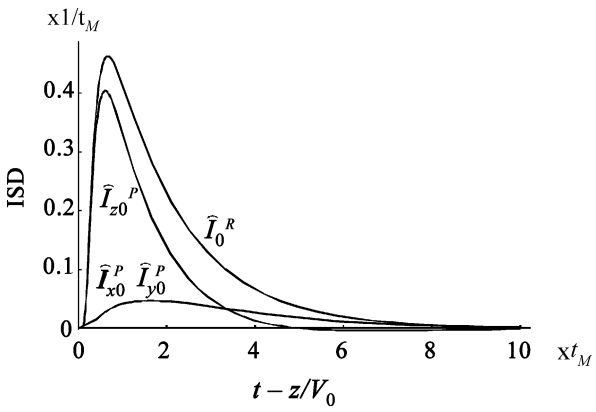
The case of a Gaussian ACF $R(\mathbf{x}) = \mathrm{Exp}\left(-r^2/a^2\right)$ is mathematically tractable. For the initial condition $\widehat{I}_x^{\mathrm{P}} = \widehat{I}_y^{\mathrm{P}} = 0$ and $\widehat{I}_z^{\mathrm{P}} = \delta(t - z/V_0)$ at $z = 0$, ISDs are analytically written as [72]

$$\begin{aligned} \widehat{I}_{x0}^{\mathrm{P}}(z, t, \omega_c) &= \widehat{I}_{y0}^{\mathrm{P}}(z, t, \omega_c) \\ &= 2(V_0/z)(t - z/V_0) \cdot \widehat{I}_0^{\mathrm{R}}(z, t, \omega_c) \\ \widehat{I}_{z0}^{\mathrm{P}}(z, t, \omega_c) &= [1 - 4(V_0/z)(t - z/V_0)] \\ &\quad \cdot \widehat{I}_0^{\mathrm{R}}(z, t, \omega_c)\,, \end{aligned} \tag{9}$$

where subscript "0" means ISD without the wandering effect. The reference ISD is a solution for scalar waves for the initial condition $\Gamma_2(\mathbf{x}_\perp, z = 0) = 1$ [82]:

$$\widehat{I}_0^{\mathrm{R}}(z, t, \omega_c) = \frac{1}{t_M} \frac{\pi}{8} \vartheta_1' \left(0, \mathrm{e}^{-\frac{\pi^2}{4} \frac{(t - z/V_0)}{t_M}}\right) \cdot H\left(t - \frac{z}{V_0}\right)\,, \tag{10}$$

where $t_M = \sqrt{\pi} \varepsilon^2 z^2/(2V_0 a)$ is the characteristic time and function ϑ_1' is the derivative of the elliptic theta function of the first kind. Function $\widehat{I}_0^{\mathrm{R}}$ shows a broadened envelope having a delayed peak and a smoothly decaying tail as illustrated by a chained curve in Fig. 7, where solid and broken curves show three-component ISDs $\widehat{I}_{0z}^{\mathrm{P}}$ and $\widehat{I}_{x0}^{\mathrm{P}} (= \widehat{I}_{y0}^{\mathrm{P}})$, respectively, for $V_0 t_M/z = 0.05$ as an example. All the three component envelopes have broadened traces; however, the peak height of the transverse component is smaller than that of the longitudinal component and the peak delay of the transverse component is larger than that of the longitudinal component. When $\varepsilon^2 z/a \ll 1$, the peak height of $\widehat{I}_{z0}^{\mathrm{P}}$ approximately decays according to the square of travel distance and the peak ratio of the trans-



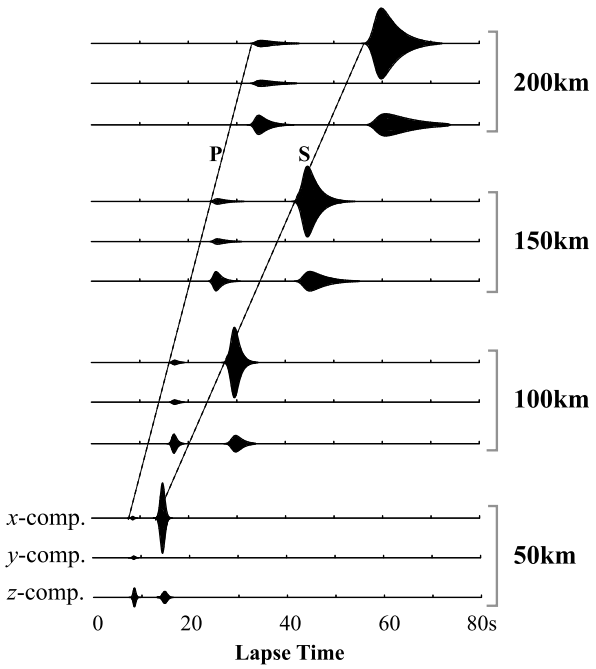
Seismic Waves in Heterogeneous Earth, Scattering of, Figure 7
Chained curve shows the reference ISD without the wandering effect in 3-D random elastic media characterized by a Gaussian ACF for the incidence of a plane P-wavelet. *Solid* and *broken curves* show three-component ISDs without the wandering effect for $V_0 t_M/z = 0.05$. Reproduced from [72]

verse component to longitudinal component is proportional to $\varepsilon^2 z/a$. ISDs $\widehat{I}_x^{\mathrm{P}}$, $\widehat{I}_y^{\mathrm{P}}$, and $\widehat{I}_z^{\mathrm{P}}$ can be calculated by using the convolution of (9) with the travel-time wandering effect $\exp\left[-(V_0 t - z_0)^2/2\sqrt{\pi}\varepsilon^2 az\right] V_0/\sqrt{2\pi\sqrt{\pi}\varepsilon^2 az}$ in time domain. For 2-D cases, the validity of the Markov approximation was numerically confirmed by a comparison with the finite difference simulations [11,36].

The above synthesis can be simply extended to S wave envelopes. Extension from plane wave incidence to impulsive radiation from a point source is also possible [73]. Figure 8 shows simulated three-component RMS envelopes along the z axis for a point source radiation, where random media are characterized by average P and S wave velocities, 6 km/s and 3.46 km/s, respectively, and a Gaussian ACF with $\varepsilon = 5\%$ and $a = 5$ km. We assume that P-wavelet radiation is isotropic and S-wavelet radiation is axially symmetric around the y axis with polarization to the x axis, where the ratio of S to P-wave source energy is chosen to be 23.3. Envelope broadening is common to both P and S waves in the syntheses. Excitation of the transverse component for P-waves and that of the radial component for S-waves are prominent. The appearance of scattered S-waves having long duration in synthesized envelopes at large travel distances qualitatively well explains observed characteristics shown in Fig. 2.

Randomness in the Lithosphere

Applying the envelope broadening model to S-wave seismograms recorded in Kanto, Japan, Sato [70] estimated the ratio $\varepsilon^2/a \approx 10^{-3}\,\mathrm{km}^{-1}$ with the assumption of a Gaussian ACF and S-wave attenuation $Q^{-1} = 0.014 f^{-1}$.



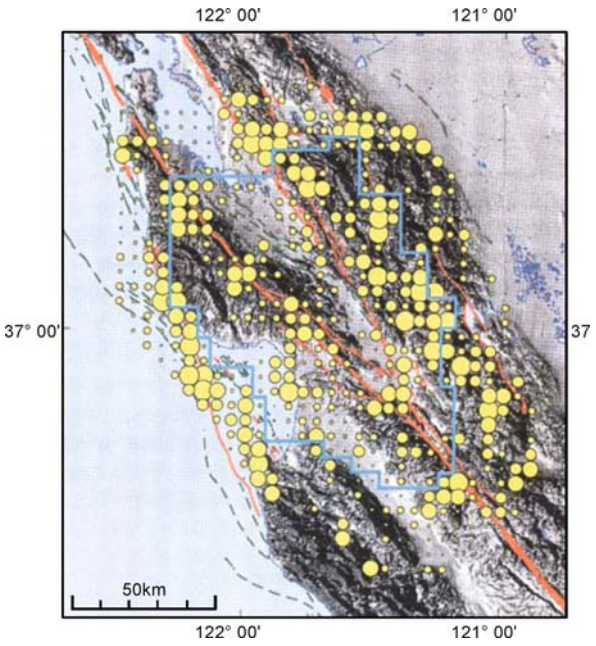
Seismic Waves in Heterogeneous Earth, Scattering of, Figure 8
Synthesized vector envelopes in random media characterized by a Gaussian ACF for radiation of P wavelet and S wavelet with a polarization to the *x*-axis from a point source. Reproduced from [73]

Saito et al. [66] studied the case of a von Kármán-type random media having a power-law spectrum at large wavenumbers, which are more appropriate for the real Earth inhomogeneity. The resultant envelope shows frequency dependence, which is controlled by the roll-off of the PSDF. Analyzing the hypocentral-distance dependence and frequency dependence of S-wave seismogram envelopes in northern Honshu, Japan for 2–32 Hz, Saito et al. [66] estimated parameters of von Kármán-type ACF as $\kappa = 0.6$ and $\varepsilon^{2.2}/a \approx 10^{-3.6}\,\mathrm{km}^{-1}$ with $Q^{-1} = 0.009 f^{-1}$. It means the PSDF decreases as wavenumber to the power of −4.2. Petukhin and Gusev [55] averaged S-wave seismogram envelopes of small earthquakes recorded in Kamchatka and compared the shapes with those numerically calculated for various types of random media. They concluded that random media whose PSDF decreases as the wavenumber to the power of −3.5 to −4 are appropriate.

Spatial Variation of Scattering Characteristics

Scattering Coefficient and Active Faults

Precisely examining coda envelopes of local earthquakes against lapse time measured from the origin time, we find



Seismic Waves in Heterogeneous Earth, Scattering of, Figure 9
Distribution of relative scattering coefficient at a depth of 0–5 km in central California revealed from the coda envelope inversion. *Circles* with larger diameter indicate stronger scattering and *solid lines* represent active faults. Reproduced with permission from [49]

temporal fluctuations around the smoothly decaying master curve. We may interpret that swellings and dips around the master curve are caused by stronger and weaker scatterers, respectively, distributed in the subsurface. By using a single isotropic scattering model, Nishigami [48] proposed an inversion scheme from coda envelopes of local earthquakes recorded at multiple stations for estimating the spatial variation of the scattering coefficient. Applying this inversion scheme to coda records obtained in central California, Nishigami [49] mapped the distribution of relative scattering coefficient in the shallow crust as in Fig. 9. A good correlation is found between sub-parallel active faults and relatively stronger scattering zones marked by larger circles, where some large circles are caused by topographic roughness. He also suggested that segment boundaries of the San Andreas Fault are characterized by relatively stronger scattering.

Stacking forward scattered energy in the coda of teleseismic P waves observed by a local seismographic network, Revenaugh [59] proposed a Kirchhoff coda migration method, which puts a focus on small-angle scattering from the forward direction. He made a map of *P*-wave scatterers in the upper mantle beneath southern California. Between depths of 50 km and 200 km, the southern flank of the slab subducting beneath the Transverse Ranges was marked by strong scattering. Using the same method, Revenaugh [60,61] estimated geographic variation of the statistical significance of scattering potential in the upper crust in California, where the scattering potential is a measure of the likelihood that scattering strength locally exceeds the regional mean. In the region surrounding the 1992 Landers earthquake of M 7.3, he found a noticeable tendency for aftershocks to cluster in regions of strong scattering potential.

There were more precise mappings of scattering coefficient. Slant-stacking records of 12 explosions in the Awaji island, Japan registered by a dense seismic array for a 6–10 Hz band, Matsumoto et al. [45] mapped the spatial distribution of PP single scatterers. The resultant distribution of scatterers shows higher strengths beneath the initiation point of the mainshock rupture and in the southwestern part of the fault plane of the 1995 Kobe earthquake (M 7.2). Analyzing precisely aftershock records of the 2000 western Tottori earthquake (M 7.3), Japan registered by a dense seismic network, Asano and Hasegawa [5] found strong scattering along and around the fault zone of 20 km in length.

Coda Attenuation and Deformation Zone

Regional variation of coda attenuation Q_c^{-1} has been measured from the decay gradient of coda amplitude envelopes of small earthquakes in various areas in the world. Jin and Aki [31] made a map of Q_c^{-1} at 1 Hz in China. They reported that Q_c^{-1} is as large as 0.01 in Tibet at the active continental collision. They found that large historical earthquakes took place in large Q_c^{-1} regions. Jin and Aki [33] made precise analysis of Q_c^{-1} for 1–32 Hz in Japan. They found significant spatial variation up to a factor of 3 for the lower frequency bands, as well as its strong frequency dependence. They found conspicuous large Q_c^{-1} for 1–4 Hz in a narrow belt from Niigata towards the south-west to the Biwa lake along the Japan Sea coast, which coincides with a high-deformation rate zone revealed from the GPS observation. For frequency bands 4–16 Hz (2–4 Hz in Kyushu), large Q_c^{-1} areas agree with volcanic and geothermal areas.

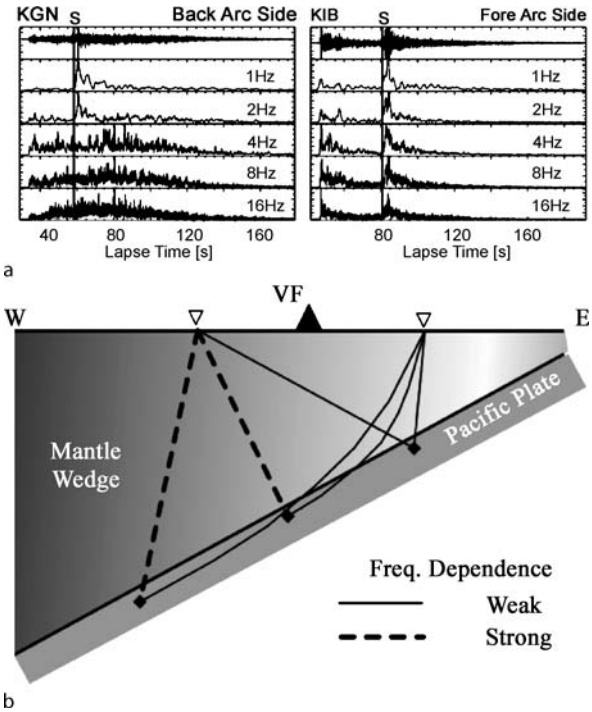
Attenuation and Volcanoes

Yoshimoto et al. [91] studied the spatial variation of MS amplitude of *S*-coda at a fixed lapse time across the volcanic front in northeastern Honshu, Japan: *S*-coda energy is uniformly distributed in the fore-arc side, whereas an exponential decrease with horizontal offset to the west from the volcanic front was found in the back-arc side. The

decay rate increases with increasing frequency. They interpreted this variation by a diffusion–absorption model, where the intrinsic absorption factor of S-wave $Q_{\mathrm{Int}}^{-1} = 0.02$ at a frequency of 10 Hz beneath the back-arc side, which is about twice as large as those reported for the fore-arc side.

Scattering and Volcanoes

Medium heterogeneity is strong beneath volcanoes. Applying the diffusion model to seismogram envelopes beneath Merapi volcanoes, Friedrich and Wegler [16] estimated the total scattering coefficient as large as $5\,\mathrm{km}^{-1}$ as shown in Fig. 4. Nishimura et al. [50] applied an envelope inversion method based on the isotropic scattering model for PP and PS scattering to artificial explosion records obtained in Jemez volcanic field, New Mexico. They found that the mid-crust under most of the region is fairly transparent but that the lower crust is heterogeneous. The strongest scattering occurs at shallow depths beneath the center of the caldera, where the medium is highly heterogeneous.



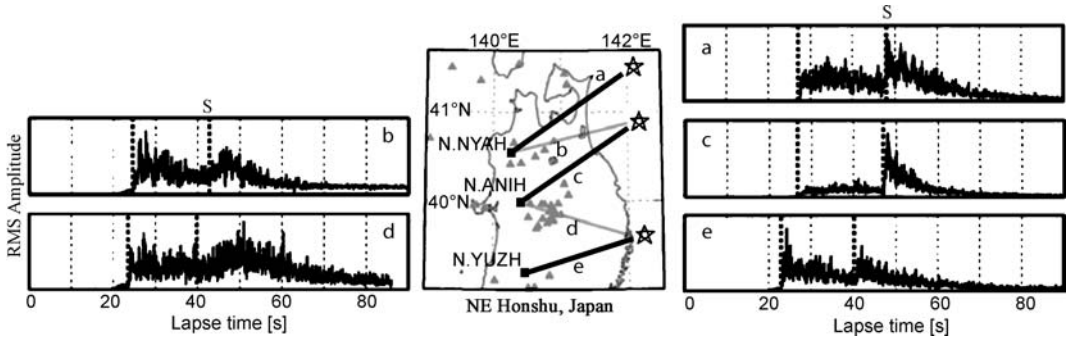
Seismic Waves in Heterogeneous Earth, Scattering of, Figure 10
a Seismogram envelopes observed in the back-arc side and fore-arc side in Kanto-Tokai, Japan and b a schematic illustration of seismic rays. Reproduced from [53]

Obara and Sato [53] analyzed S-wave envelopes of microearthquakes in Kanto-Tokai, Japan, where the Pacific plate is subducting from east to west, to examine regional differences in their envelope broadening. As shown by examples in Fig. 10a, envelope broadening is typically stronger for higher frequencies in records at stations on the back-arc side of the volcanic front but weaker and frequency independent in records on the fore-arc side. These regional differences in the envelope broadening mean that PSDF of velocity inhomogeneity is rich in short-wavelength components in the mantle wedge on the back-arc side and poor on the fore-arc side as schematically illustrated in Fig. 10b. Takahashi et al. [83] precisely examined how the peak delay from the S-wave onset depends on the ray path in northern Japan. They found that peak delays observed in the back-arc side of the volcanic front are larger for rays which propagate beneath Quaternary volcanoes (see Fig. 11b and d); however, peak delays for rays which propagate between Quaternary volcanoes are as short as those observed in the fore-arc side (see Fig. 11a, c, and e). Large peak delay suggests strong scattering due to medium inhomogeneity. That is, the structure beneath Quaternary volcanoes is not only characterized by low velocity and large intrinsic absorption revealed from tomography studies but also by strong inhomogeneity.

Nonisotropic Random Medium Oceanic Slab

If random media are statistically nonisotropic, scattering contribution depends on the propagation direction. Saito [65] studied the envelope broadening in nonisotropic random media based on the Markov approximation. His simulations show that the envelope of scalar wavelet propagating in parallel to the longer correlation direction has longer duration compared to that with the shorter correlation direction. The effective envelope broadening in the elongated direction shows the wave trap phenomenon of nonisotropic random media.

An intensity anomaly is observed on the eastern seaboard of northern Japan for deep earthquakes. The waveform records in the region of high intensity show a low-frequency ($f < 0.25$ Hz) onset for both P and S waves, followed by large-amplitude high-frequency ($f > 2$ Hz) later arrivals with long coda. A simple subduction zone model comprising a high-velocity plate with low attenuation cannot explain quantitatively these observed facts. Furumura and Kennett [17] proposed a scattering slab model that the nonisotropic random structure in the Pacific plate works as a wave-guide for high-frequency seismic waves. Their preferred random medium is characterized by a von Kármán-type ACF with elongated cor-



Seismic Waves in Heterogeneous Earth, Scattering of, Figure 11
Path dependence of RMS envelopes (16–32 Hz) in northeastern Honshu, Japan, where *stars* and *triangles* indicate earthquake epicenters and Quaternary volcanoes, respectively. Ray paths b and d travel beneath Quaternary volcanoes (*triangles*), and a, c, and e travel between Quaternary volcanoes. Reproduced from [83]

relation distance of about 10 km parallel to the plate margin and much shorter correlation length of about 0.5 km in thickness and ε of about 2%. They clearly demonstrated the scattering waveguide effects and frequency selectivity for seismic waves traveling through the Pacific plate by using 3-D numerical simulations.

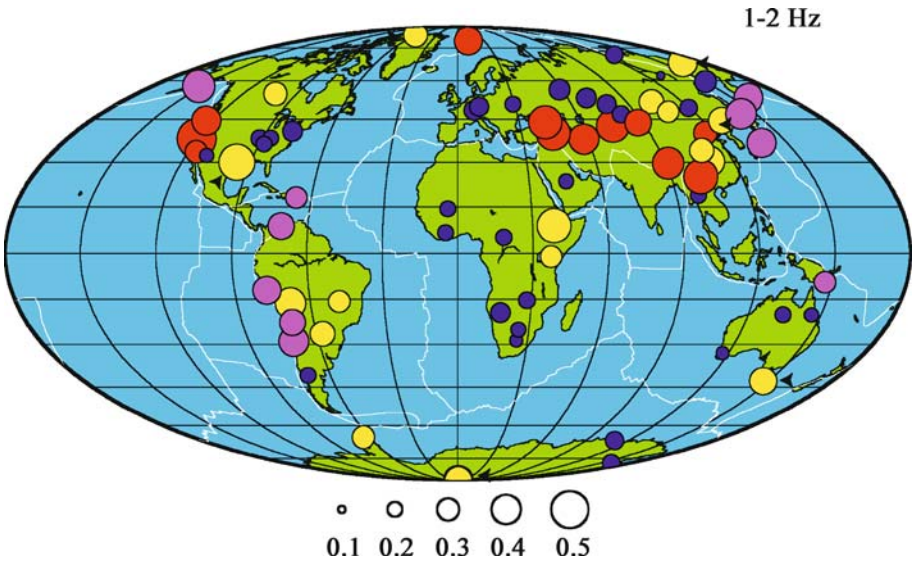
Lateral Variation of Lithospheric Heterogeneity

Korn [34] developed the energy flux model [15] appropriate for the wave front of teleseismic P and P-coda waves propagating through a scattering layer. During the propagation the primary wave loses energy due to scattering and intrinsic absorption, then the scattered energy appears as coda energy behind the wave front. From the analysis of vertical component trace envelopes observed in the world, Korn [35] found strong scattering at island arcs and smaller scattering on stable continental areas like Australia. Nishimura et al. [52] analyzed the energy partition of teleseismic P and P-coda into the transverse component to evaluate the lithospheric heterogeneity in the western Pacific region. They showed the presence of strong heterogeneity in and around the tectonically active regions. Kubanza et al. [37] systematically characterized the medium heterogeneity of the lithosphere by analyzing the partition of P-wave energy into the transverse component for 0.5–4 Hz. They found significant regional differences as shown in Fig. 12. The energy partition to the transverse component is small at stations on stable continents while the partition is large at stations in tectonically active regions such as island arcs or collision zones.

Random Inhomogeneity in the Lithosphere and Mantle

Records of earthquakes registered by arrays of seismographs are useful for the statistical measurement of the Earth inhomogeneity. Aki [2] first analyzed teleseismic P-waves centered on about 0.6 Hz registered by a seismic array in Montana for the quantification of the lithospheric inhomogeneity. Measuring transverse correlation functions of teleseismic P-waves arriving from near vertical incidence, he found a positive correlation between log-amplitude and phase fluctuations as theoretically predicted for a Gaussian ACF. From plots of the ratio of RMS log-amplitude to RMS phase fluctuations against the correlation between log-amplitude and phase fluctuations, he estimated the thickness of the inhomogeneous lithosphere to be 60 km, $a = 10$ km, and $\varepsilon = 4\%$. Flatté and Wu [12] measured the transverse correlation of log-amplitude and phase fluctuations of teleseismic P-wave beams with 2 Hz center frequency recorded at NORSAR. They also introduced the new concept of angular correlation functions, which are based on measurements of two rays with different incident angles. They proposed a model for lithospheric and asthenospheric inhomogeneities that consists of two overlapping layers: the upper layer extending from the surface to about the 200 km depth has a white PSDF, however, the lower layer extending from 15 to 250 km has poor amplitude in the short wavelength spectrum. There are more small-scale inhomogeneities near the surface compared with the deeper portions.

The radiative transfer theory with scattering coefficients calculated from the Born approximation was also used for the study of mantle inhomogeneity. Analyzing stacked P and P coda envelopes of teleseismic (>10°) events at 1 Hz, Shearer and Earle [79] concluded that most scattering occurs in the lithosphere and upper mantle, but that some lower mantle scattering is likely required. They estimated ε to be 3–4% and $a = 4$ km in the upper mantle and 0.5% and 8 km in the lower mantle. Analyzing envelopes of precursors to PKP, Margerin and Nolet [43] found that inhomogeneity can not be restricted to the D″ layer and a small inhomogeneity spread over the whole



Seismic Waves in Heterogeneous Earth, Scattering of, Figure 12
Plots of the square root of the relative partition of wave energy into the transverse component (1–2 Hz) by a diameter of the circle revealed from the teleseismic P-waves. Reproduced from [37]

lower mantle. They proposed a von Kármán ACF of $\kappa = 0$ for random media, which has a power law PSDF rich in short wavelengths compared with an exponential ACF. They mentioned that ε of 0.1–0.2% in the whole lower mantle is enough to explain the observation even though correlation distance is irresolvable because of the limited range of observations.

Imaging of Subsurface Heterogeneity

There have been developments in deterministic imaging of medium inhomogeneity from the analysis of array records of P-coda waves. When the structures of interest are characterized by laterally variable stratification, the receiver-function technique [38] is useful since it is based on the deconvolution of the horizontal component trace in the radial direction by the vertical component trace for measuring the Ps conversion depth. On the other hand, scattering from localized volume inhomogeneity is most readily treated by using the Born approximation. Analyzing array records of teleseismic P coda in central Oregon by using the Born approximation for both forward scattered waves and backscattered free-surface reflected waves, Rondenay et al. [62] successfully imaged the precise structure of the Cascadia subduction zone, which is consistent with the consequences of prograde metamorphic reactions occurring within the oceanic crust. Analyzing array records of P coda waves of regional earthquakes at Izu-Oshima volcano, which erupted in 1986, Mikada et al. [46] deterministically imaged PP and PS scatterers on the basis of diffraction tomography. They interpreted a cloud of scatterers centered at about 10 km depth beneath the volcano crater as a primary magma reservoir and smaller and shallower patches of high scattering strength with sub-magma reservoirs.

Temporal Change in the Earth Medium Structure

There were reports on the temporal changes in the Earth medium structure revealed from the analyses of scattered waves. One is coda amplitude envelope analysis, which gives information about the change in intrinsic absorption and scattering strength of the crustal heterogeneity. Another is coda phase interferometry, which offers information about the change in background velocity.

Change in Coda Characteristics

Monitoring coda envelopes of local earthquakes, Gusev and Lemzikov [23] reported temporal change in Q_c^{-1} before and after the 1971 Ust-Kamchatsk earthquake (M 7.8), and Jin and Aki [30] reported temporal change in Q_c^{-1} associated with the 1976 Tangshan earthquake (M 7.8) in China. Their observation attracted the interest of geophysicists to the temporal variation of coda characteristics because of a potential for monitoring the stress accumulation process preceding an earthquake occurrence. Analyzing high-frequency seismograms recorded at River-

side, California for 55 years, Jin and Aki [32] found a temporal variability in Q_c^{-1} at about 1.6 Hz having a positive correlation with the seismic b-value calculated for M > 3 earthquakes within a 180 km radius. The seismic b-value is a measure of the ratio between the numbers of small to large earthquakes; smaller b-values mean that there are relatively fewer small earthquakes compared to the number of larger ones. They interpreted these changes in Q_c^{-1} and the b-value by creep fractures in the ductile part of the lithosphere. Hiramatsu [25] precisely examined the temporal variation in Q_c^{-1} and b-value for 10 years before and after the 1995 Hyogo-ken Nanbu earthquake (M 7.2) in Japan. At frequencies between 1.5 and 4.0 Hz the temporal variation in Q_c^{-1} increased after the mainshock occurrence, where the variation in b-value was opposite.

Sato [69] analyzed the relation between coda duration time and earthquake magnitude of small earthquakes before and after an M 6.8 earthquake in central Japan. He found that coda durations were anomalously longer than usual for 16 months before the earthquake occurrence. From 24-year observation of coda at 0.5 Hz in Kamchatka Gusev [19] reported prominent anomalies in coda level residual from the mean coda excitation level at 100s lapse time associated with two M 8 earthquakes and a volcanic eruption.

Sawazaki et al. [76] measured the temporal variation of the spectral ratio of coda waves registered on the ground surface to that at the bottom of a borehole of 100 m depth in Japan, which experienced strong ground motion of several hundred gals. They reported a sudden drop of the site amplification factor caused by earthquake strong motion and gradual recovery for a few years approaching to the original ratio. They suggested crack formation and ground water movement for explaining the site factor weakening observed.

Coda Interferometry

Pairs of earthquakes with almost identical focal mechanisms are called earthquake doublets. The cross-correlation function of earthquake doublet records allow us to detect differences in the background velocity of the Earth medium that took place in between the pair of earthquakes. Applying the phase spectral analysis to coda wave records of earthquake doublets before and after the 1979 Coyote earthquake of M 5.9 in California, Poupinet et al. [57] found that the coda wave arrivals for some stations are progressively delayed for the second earthquake in the doublet. They interpreted systematic variation along the coda as a decrease of background S-wave velocity by 0.2% in an oblong region 5–10 km in radius at the south end of the aftershock zone. Applying the phase spectral analysis to records of repeated artificial explosions, Nishimura et al. [51] found that the average seismic velocity of the crust in the frequency range of 3–6 Hz decreased by about 1% around the focal region of an M 6.1 earthquake at Iwate volcano in northeastern Honshu, Japan in 1998. They interpreted this velocity drop by the dilatation caused by the M 6.1 earthquake with stress sensitivity of the velocity change $(\delta V/V)/\delta\sigma$ of the order of 0.1 MPa^{-1}. From the set of successive artificial explosion experiments, they observed gradual recovery of the seismic velocity towards its original value over the next four years. While interferometry detected a change in velocity of the order of 1%, it was unidentifiable from travel time analysis of first arrivals. Using coda waves is superior to direct waves since coda waves volumetrically sample the Earth medium.

Snieder et al. [81] demonstrated detection of the nonlinear dependence of the seismic velocity in granite on temperature and the associated acoustic emissions from the interference measurement of coda waves in rock samples as a laboratory experiment. They named this method "coda interferometry" and proposed to use it for detecting the presence of temporal changes in the medium, or in diagnostic mode. There is an idea to retrieve the Green function from the stacked cross-correlation function (CCF) of multiple scattered waves or microseisms at a pair of stations on condition that the propagation directions of those waves are randomly isotropic. Stacking CCFs of coda waves at several pairs of stations for regional earthquakes in Mexico, Campillo and Paul [7] estimated the surface wave velocity between each station pair from the peak delay. The idea was extended for monitoring the temporal change in the crustal structure. Wegler and Sens-Schönfelder [84] computed the ACF of microseisms recorded at a site in the vicinity of the source region of the 2004 Mid-Niigata earthquake (M 6.6) in Japan for three months. They detected a sudden decrease of relative seismic velocity in the crust of 0.6% at the occurrence of the earthquake from the temporal variation of stacked ACFs.

Future Directions

In addition to classic parameterization as a layered structure with sharp edges and smooth velocity perturbation, we introduced new approaches using scattered waves that reflect solid Earth heterogeneity. For high-frequency seismograms of earthquakes, envelope characteristics such as the excitation level and the decay gradient of coda envelopes and the envelope broadening of the direct wavelet

are useful for the study of small-scale inhomogeneities. The lithospheric inhomogeneity is phenomenologically well characterized by the scattering coefficient and coda attenuation factor as a function of frequency. Furthermore, the power spectral density function of random velocity inhomogeneity is estimated from the frequency dependence of high-frequency seismogram envelopes of local earthquakes or the array analysis of teleseismic waves. The radiative transfer theory with scattering coefficients calculated from the Born approximation and the Markov approximation for the parabolic wave equation are useful mathematical tools for the analyses.

Scattering characteristics are found to vary spatially reflecting seismotectonic settings. It will be necessary for us to make a classification of seismogram-envelope patterns in various regions in the world under different tectonic conditions. It is interesting to model how such a variation of medium inhomogeneity was created through the geodynamic process. Compared to the lithospheric inhomogeneity, there were insufficient numbers of studies on the mantle inhomogeneity. It will be necessary to map the distribution of inhomogeneities deep in the mantle, which is useful for the study of the evolution of the planet Earth.

For mathematical simplicity; however, most approaches assume homogeneity and isotropy of randomness and a constant background velocity, which are somewhat different from reality. It will be necessary to mathematically develop the envelope synthesis in inhomogeneous media that are a superposition of small-scale random inhomogeneities and a gradually varying background velocity. As revealed from the ray path dependence of S-wave envelope broadening, randomness varies from place to place. It is also necessary to develop the envelope synthesis for random media having spatially varying statistical parameters. In addition, it is important to examine how conversion scattering between P and S waves contributes to form spindle-like envelopes in highly scattering media as shown in high-frequency seismograms observed in volcanoes and on the Moon.

For further understanding, there are monographs that treat the discussed subjects as follows: Sato and Fehler [74] review seismological observation and mathematical models; Shapiro and Hubral [78] put special focus on wave propagation through stratified random media; Goff and Holliger [18] summarize the crustal heterogeneity; Wu and Maupin [86] compile recent developments in mathematical modeling of wave propagation in inhomogeneous media; Chandrasekhar [8] is a classic text for radiative transfer theory; Ishimaru [29] and Rytov et al. [63] offer advanced mathematical tools for the study of wave propagation in random media.

Acknowledgments

The author thanks Michael Korn, Heiner Igel, and an anonymous reviewer. Their comments were helpful for improving the readability. The author is grateful to NIED, Japan for providing digital seismic data.

Bibliography

1. Aki K (1969) Analysis of seismic coda of local earthquakes as scattered waves. J Geophys Res 74:615–631
2. Aki K (1973) Scattering of P waves under the Montana LASA. J Geophys Res 78:1334–1346
3. Aki K (1980) Attenuation of shear-waves in the lithosphere for frequencies from 0.05 to 25 Hz. Phys Earth Planet Inter 21:50–60
4. Aki K, Chouet B (1975) Origin of coda waves: Source, attenuation and scattering effects. J Geophys Res 80:3322–3342
5. Asano Y, Hasegawa A (2004) Imaging the fault zones of the 2000 western Tottori earthquake by a new inversion method to estimate three-dimensional distribution of the scattering coefficient. J Geophys Res 109:B06306. doi:10.1029/2003JB002761
6. Atkinson GM (1993) Notes on ground motion parameters for Eastern North America: Duration and H/V ratio. Bull Seismol Soc Am 83:587–596
7. Campillo M, Paul A (2003) Long-Range Correlations in the Diffuse Seismic Coda. Science 299:547–549. doi:10.1126/science.1078551
8. Chandrasekhar S (1960) Radiative Transfer. Dover, New York
9. Dainty AM, Toksöz MN (1981) Seismic codas on the earth and the moon: A comparison. Phys Earth Planet Inter 26:250–260
10. Fehler M, Hoshiba M, Sato H, Obara K (1992) Separation of scattering and intrinsic attenuation for the Kanto-Tokai region, Japan, using measurements of S-wave energy versus hypocentral distance. Geophys J Int 108:787–800
11. Fehler M, Sato H, Huang LJ (2000) Envelope broadening of outgoing waves in 2-D random media: A comparison between the Markov approximation and numerical simulations. Bull Seismol Soc Amer 90:914–928
12. Flatté SM, Wu RS (1988) Small-scale structure in the lithosphere and asthenosphere deduced from arrival time and amplitude fluctuations at NORSAR. J Geophys Res 93:6601–6614
13. Foldy LL (1945) The multiple scattering of waves- I General theory of isotropic scattering by randomly distributed scatterers. Phys Rev 67:107–119
14. Frankel A, Clayton RW (1986) Finite difference simulations of seismic scattering: Implications for the propagation of short-period seismic waves in the crust and models of crustal heterogeneity. J Geophys Res 91:6465–6489
15. Frankel A, Wennerberg L (1987) Energy-flux model of seismic coda: Separation of scattering and intrinsic attenuation. Bull Seismol Soc Am 77:1223–1251
16. Friedrich C, Wegler U (2005) Localization of seismic coda at Merapi volcano (Indonesia). Geophys Res Lett 32:L14312. doi:10.1029/2005GL023111
17. Furumura T, Kennett BLN (2005) Subduction zone guided waves and the heterogeneity structure of the subducted plate: intensity anomalies in northern Japan. J Geophys Res 110:B10302. doi:10.1029/2004JB003486

18. Goff JA, Holliger K (2002) Heterogeneity in the Crust and Upper Mantle – Nature, Scaling and Seismic Properties. Kluwer Academic/Plenum Publishers, Dorderecht, pp 1–358
19. Gusev AA (1995) Baylike and continuous variations of the relative level of the late coda during 24 years of observation on Kamchatka. J Geophys Res 100:20311–20319
20. Gusev AA (1995) Vertical profile of turbidity and coda Q. Geophys J Int 123:665–672
21. Gusev AA, Abubakirov IR (1987) Monte-Carlo simulation of record envelope of a near earthquake. Phys Earth Planet Inter 49:30–36
22. Gusev AA, Abubakirov IR (1996) Simulated envelopes of non-isotropically scattered body waves as compared to observed ones: Another manifestation of fractal heterogeneity. Geophys J Int 127:49–60
23. Gusev AA, Lemzikov VK (1985) Properties of scattered elastic waves in the lithosphere of Kamchatka: Parameters and temporal variations. Tectonophysics 112:137–153
24. Hemmer PC (1961) On a generalization of Smoluchowski's diffusion equation. Physica A 27:79–82
25. Hiramatsu Y, Hayashi N, Furumoto M (2000) Temporal changes in coda *Q*21 and *b* value due to the static stress change associated with the 1995 Hyogo-ken Nanbu earthquake. J Geophys Res 105:6141–6151
26. Holliger K, Levander A (1992) A stochastic view of lower crustal fabric based on evidence from the Ivrea zone. Geophys Res Lett 19:1153–1156
27. Hoshiba M (1991) Simulation of multiple-scattered coda wave excitation based on the energy conservation law. Phys Earth Planet Inter 67:123–136
28. Hoshiba M, Sato H, Fehler M (1991) Numerical basis of the separation of scattering and intrinsic absorption from full seismogram envelope – A Monte-Carlo simulation of multiple isotropic scattering. Pa Meteorol Geophys, Meteorol Res Inst 42:65–91
29. Ishimaru A (1978) Wave Propagation and Scattering in Random Media, vol 1 and 2. Academic, San Diego
30. Jin A, Aki K (1986) Temporal change in coda *Q* before the Tangshan earthquake of 1976 and the Haicheng earthquake of 1975. J Geophys Res 91:665–673
31. Jin A, Aki K (1988) Spatial and temporal correlation between coda *Q* and seismicity in China. Bull Seismol Soc Am 78:741–769
32. Jin A, Aki K (1989) Spatial and temporal correlation between coda Q^{-1} and seismicity and its physical mechanism. J Geophys Res 94:14041–14059
33. Jin A, Aki K (2005) High-resolution maps of Coda Q in Japan and their interpretation by the brittle-ductile interaction hypothesis. Earth Planets Space 57:403–409
34. Korn M (1990) A modified energy flux model for lithospheric scattering of teleseismic body waves. Geophys J Int 102:165–175
35. Korn M (1993) Determination of site-dependent scattering Q from P-wave coda analysis with an energy-flux model. Geophys J Int 113:54–72
36. Korn M, Sato H (2005) Synthesis of plane vector-wave envelopes in 2-D random elastic media based on the Markov approximation and comparison with finite difference simulations. Geophys J Int 161:839–848
37. Kubanza M, Nishimura T, Sato H (2006) Spatial variation of lithospheric heterogeneity on the globe as revealed from transverse amplitudes of short-period teleseismic P-waves. Earth Planets Space 58:45–e48
38. Langston CA (1979) Structure under Mount Rainer, Washington, inferred from teleseismic body waves. J Geophys Res 84:4749–4762
39. Larose E, Margerin L, van Tiggelen BA, Campillo M (2004) Weak Localization of Seismic Waves. Phys Rev Lett 93:048501-4. doi:10.1103/PhysRevLett.93.048501
40. Lee WS, Sato H, Lee KW (2003) Estimation of S-wave scattering coefficient in the mantle from envelope characteristics before and after the ScS arrival. Geophys Res Lett 30:2248. doi:10.1029/2003GL018413
41. Margerin L (2005) Introduction to radiative transfer of seismic waves. In: Levander A, Nolet G (eds) Seismic Earth: Array Analysis of Broad-band Seismograms, Geophysical Monograph Series, vol 157, chap 14. AGU, Washington, pp 229–252
42. Margerin L, Campillo M, van Tiggelen BA (2001) Coherent backscattering of acoustic waves in the near field. Geophys J Int 145:593–603
43. Margerin L, Nolet G (2003) Multiple scattering of high-frequency seismic waves in the deep Earth: PKP precursor analysis and inversion for mantle granularity. J Geophys Res 108, B11:2514. doi:10.1029/2003JB002455
44. Matsumoto S (2005) Scatterer density estimation in the crust by seismic array processing. Geophys J Int 163:622–628
45. Matsumoto S, Obara K, Hasegawa A(1998) Imaging P-wave scatterer distribution in the focal area of the 1995 M7.2 Hyogo-ken Nanbu (Kobe) Earthquake. Geophys Res Lett 25:1439–1442
46. Mikada H,Watanabe H, Sakashita S (1997) Evidence for subsurface magma bodies beneath Izu-Oshima volcano inferred from a seismic scattering analysis and possible interpretation of the magma plumbing system of the 1986 eruptive activity. Phys Earth Planet Inter 104:257–269
47. Nakahara H, Nishimura T, Sato H, Ohtake M (1998) Seismogram envelope inversion for the spatial distribution of high-frequency energy radiation from the earthquake fault: Application to the 1994 far east off Sanriku earthquake, Japan. J Geophys Res 103:855–867
48. Nishigami K (1991) A new inversion method of coda waveforms to determine spatial distribution of coda scatterers in the crust and uppermost mantle. Geophys Res Lett 18:2225–2228
49. Nishigami K (2000) Deep crustal heterogeneity along and around the San Andreas fault system in central California and its relation to the segmentation. J Geophys Res 105:7983–7998
50. Nishimura T, Fehler M, Baldridge WS, Roberts P, Steck L (1997) Heterogeneous structure around the Jemez Volcanic Field, New Mexico, USA, as inferred from the envelope inversion of active-experiment seismic data. Geophys J Int 131:667–681
51. Nishimura T, Tanaka S, Yamawaki T, Yamamoto H, Sano T, Sato M, Nakahara H, Uchida N, Hori S, Sato H (2005) Temporal changes in seismic velocity of the crust around Iwate volcano, Japan, as inferred from analyses of repeated active seismic experiment data from 1998 to 2003. Earth Planets Space 57:491–505
52. Nishimura T, Yoshimoto K, Ohtaki T, Kanjo K, Purwana I (2002) Spatial distribution of lateral heterogeneity in the upper mantle around the western Pacific region as inferred from analysis of transverse components of teleseismic P-coda. Geophys Res Lett 29:2089–2137. doi:10.1029/2002GL015606

53. Obara K, Sato H (1995) Regional differences of random inhomogeneities around the volcanic front in the Kanto-Tokai area, Japan, revealed form the broadening of S wave seismogram envelopes. J Geophys Res 100:2103–2121
54. Paaschens JCJ (1997) Solution of the time-dependent Boltzmann equation. Phys Rev E 56:1135–1141
55. Petukhin AG, Gusev AA (2003) The Duration-distance Relationship and Average Envelope Shapes of Small Kamchatka Earthquakes. Pure Appl Geophys 160:1717–1743
56. Phillips WS, Aki K (1986) Amplification of coda waves from local earthquakes in Central California. Bull Seismol Soc 76:627–648
57. Poupinet G, Ellsworth WL, Frechet J (1984) Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras fault, California. J Geophys Res 89:5719–5731
58. Przybilla J, Korn M, Wegler U (2006) Radiative transfer of elastic waves versus finite difference simulations in two-dimensional random media. J Geophys Res 111:B04305. doi:10.1029/2005JB003952
59. Revenaugh J (1995) A scattered-wave image of subduction beneath the Transverse Ranges. Science 268:1888–1892
60. Revenaugh J (1995) Relationship of the 1992 Landers, California, earthquake sequence to seismic scattering. Science 270:1344–1347
61. Revenaugh J (1999) Geologic Applications of Seismic Scattering. Annu Rev Earth Planet Sci 27:55–73
62. Rondenay S, Bostock MG, Shragge J (2001) Multiparameter two-dimensional inversion of scattered teleseismic body waves 3. Application to the Cascadia 1993 data set. J Geophys Res 106:30795–30807
63. Rytov SM, Kravtsov YA, Tatarskii VI (1987) Principles of Statistical Radio Physics, vol 4, Wave Propagation Through Random Media. Springer, Berlin
64. Ryzhik LV, Papanicolaou GC, Keller JB (1996) Transport equations for elastic and other waves in random media. Wave Motion 24:327–370
65. Saito T (2006) Synthesis of scalar-wave envelopes in two-dimensional weakly anisotropic random media by using the Markov approximation. Geophys J Int 165:501–515. doi:10.1111/j.1365-246X2006.02896.x
66. Saito T, Sato H, Ohtake M (2002) Envelope broadening of spherically outgoing waves in three-dimensional random media having power-law spectra. J Geophys Res 107:2089. doi:10.1029/2001JB000264
67. Sato H (1977) Single isotropic scattering model including wave conversions: Simple theoretical model of the short period body wave propagation. J Phys Earth 25:163–176
68. Sato H (1984) Attenuation and envelope formation of three-component seismograms of small local earthquakes in randomly inhomogeneous lithosphere. J Geophys Res 89:1221–1241
69. Sato H (1987) A precursor-like change in coda excitation before the western Nagano earthquake (Ms = 6.8) of 1984 in central Japan. J Geophys Res 92:1356–1360
70. Sato H (1989) Broadening of seismogram envelopes in the randomly inhomogeneous lithosphere based on the parabolic approximation: Southeastern Honshu, Japan. J Geophys Res 94:17735–17747
71. Sato H (1990) Unified approach to amplitude attenuation and coda excitation in the randomly inhomogeneous lithosphere. Pure Appl Geophys 132:93–121
72. Sato H (2006) Synthesis of vector wave envelopes in three-dimensional random elastic media characterized by a Gaussian autocorrelation function based on the Markov approximation: Plane wave case. J Geophys Res 111:B06306. doi:10.1029/2005JB004036
73. Sato H (2007) Synthesis of vector-wave envelopes in 3-D random elastic media characterized by a Gaussian autocorrelation function based on the Markov approximation: Spherical wave case. J Geophys Res Solid Earth 112:B01301. doi:10.1029/2006JB004437
74. Sato H, Fehler M (1998) Seismic Wave Propagation and Scattering in the Heterogeneous Earth. AIP Press/Springer, New York
75. Sato H, Nakahara H, Ohtake M (1997) Synthesis of scattered energy density for non-spherical radiation from a point shear dislocation source based on the radiative transfer theory. Phys Earth Planet Inter 104:1–281
76. Sawazaki K, Sato H, Nakahara H, Nishimura T (2006) Temporal Change in Site Response Caused by Earthquake Strong Motion as Revealed from Coda Spectral Ratio Measurement. Geophys Res Lett 33:L21303. doi:10.1029/2006GL027938
77. Shang T, Gao L (1988) Transportation theory of multiple scattering and its application to seismic coda waves of impulsive source. Sci Sin 31B:1503–1514
78. Shapiro SA, Hubral P (1999) Elastic Waves in Random Media – Fundamentals of Seismic Stratigraphic Filtering. Springer, Berlin
79. Shearer PM, Earle PS (2004) The global short-period wavefield modeled with a Monte Carlo seismic phonon method. Geophys J Int 158:1103–1117
80. Shiomi K, Sato H, Ohtake M (1997) Broad-band power-law spectra of well-log data in Japan. Geophys J Int 130:57–64
81. Snieder R, Gret A, Douma A, Scales J (2002) Coda wave interferometry for estimating nonlinear behavior in seismic elocity. Science 295:2253–2255
82. Sreenivasiah I, Ishimaru A, Hong ST (1976) Two-frequency mutual coherence function and pulse propagation in a random medium: An analytic solution to the plane wave case. Radio Sci 11:775–778
83. Takahashi T, Sato H, Nishimura T, Obara K (2006) Strong inhomogeneity beneath Quaternary volcanoes revealed from the peak delay analysis of S-wave seismograms of microearthquakes in northeastern, Japan. Geophys J Int 168:90–99. doi:10.1111/j.1365-246X2006.03197.x
84. Wegler U, Sens-Schönfelder C (2007) Fault zone monitoring with passive image interferometry. Geophys J Int 168:1029-1033. doi:10.1111/j.1365-246X2006.03284.x
85. Wu RS (1985) Multiple scattering and energy transfer of seismic waves – separation of scattering effect from intrinsic attenuation – I Theoretical modeling. Geophys J R Astron Soc 82:57–80
86. Wu RS, Maupin V (eds) (2007) Advances in Wave Propagation in Heterogeneous Earth. In: Dmowska R (ed) Advanced in Geophysics, vol 48. Academic Press, San Diego, pp 561–596
87. Yomogida K, Benites R (1995) Relation between direct wave *Q* and coda *Q*: A numerical approach. Geophys J Int 123:471–483
88. Yoshimoto K (2000) Monte-Carlo simulation of seismogram envelope in scattering media. J Geophys Res 105:6153–6161
89. Yoshimoto K, Sato H, Ohtake M (1993) Frequency-dependent attenuation of P and S waves in the Kanto area, Japan, based on the coda-normalization method. Geophys J Int 114:165–174

90. Yoshimoto K, Sato H, Ohtake M (1997) Short-wavelength crustal inhomogeneities in the Nikko area, central Japan, revealed from the three-component seismogram envelope analysis. Phys Earth Planet Inter 104:63–73
91. Yoshimoto K, Wegler U, Korn M (2006) A volcanic front as a boundary of seismic attenuation structures in northeastern Honshu, Japan. Bull Seismol Soc Am 96:637–646
92. Zeng Y, Su F, Aki K (1991) Scattering wave energy propagation in a random isotropic scattering medium I Theory. J Geophys Res 96:607–619
93. Zhao D, Hasegawa A, Horiuchi S (1992) Tomographic imaging of P and S wave velocity structure beneath Northeastern Japan. J Geophys Res 97:19909–19928

Self-assembled Materials

Aatto Laaksonen[1], Lennart Bergström[2]
[1] Division of Physical Chemistry, Dept. of Physical, Inorganic and Structural Chemistry, Arrhenius Laboratory, Stockholm University, Stockholm, Sweden
[2] Materials Chemistry Research Group, Dept. of Physical, Inorganic and Structural Chemistry, Arrhenius Laboratory, Stockholm University, Stockholm, Sweden

Article Outline

Glossary
Definition of the Subject
Introduction
First-Principles Calculations and Simulations
Atomistic Molecular Dynamic Simulations
Monte Carlo Simulations
Kinetic Monte Carlo Simulations
Phase-Field Modeling
Random Sequential Adsorption
Cellular Automata
Multi-Scale Modeling
Future Directions
Acknowledgments
Bibliography

Glossary

AI Ab initio
AM1 Austin model 1
AO Atomic orbital
B3LYP Becke's three parameter hybrid LYP functional
BKL Bortz–Kalos–Lebowitz
BO Born–Oppenheimer
CA Cellular automata
CE Cluster expansion
CC Coupled cluster
CI Configuration interaction
CP Car-Parrinello
DCA Dynamic cellular automata
DFT Density functional theory
DLVO Derjaguin–Landau–Verwey–Overbeek
GGA Generalized gradient approximation
HF 1) Hartree–Fock 2) Harris–Foulkes functional
kMC kinetic Monte Carlo
KS Kohn–Sham
LbL Layer-by-layer
LCAO Linear Combination of Atomic Orbitals
LDA Local density approximation
LYP Lee–Young–Parr
MC Monte Carlo
MD Molecular Dynamics
MNDO Modified neglect of differential overlap
MO Molecular Orbital
MP Møller–Plesset
NDDO Neglect of diatomic differential overlap
NDO Neglect of differential overlap
NPT Isobaric ensemble: constant N, P and T
NVE Microcanonical ensemble: constant N, V and E
NVT Canonical ensemble: constant N, V and T
PBC Periodic boundary conditions
PF Phase-field
PM3 Parametric model 3
PP Pseudo-potential
PW Perdew–Wang
QCA Quantum-dot cellular automata
RSA Random sequential adsorption
SA Self-assembly
SAM Self-assembled monolayers
SCF Self-consistent field
STM Scanning tunneling microscopy
TB Tight-binding
XC Exchange correlation

Definition of the Subject

The concept of self-assembly is today highly popular and frequently used to describe a wide range of phenomena. It is also a concept that has a possibility to change the way we produce various types of materials [1]. Self-assembly can broadly speaking be defined as a process with the following features [2], where: (i) it involves pre-existing components, i. e. the components are not formed by the reaction itself. (ii) the process should be reversible to some extent. (iii) it can be controlled by design.

The supramolecular chemistry approach, pioneered by Jean-Marie Lehn [3], where molecular recognition is used to assemble supramolecular materials is clearly an impor-

tant starting point for much of this work. The pre-existing components are usually molecules, amphiphilic species, or oligomeric or polymeric species. However, self-assembly can also be observed on larger length scales involving e. g. nanoparticles (superlattices) and larger colloidal particles (colloidal crystals). The final self-assembled material should display certain features that can be related to the building blocks in addition to the novel features that are created by the self-assembled material itself. Reversibility is a key feature since self-assembly processes commonly scan a rich energy landscape with many metastable states. This is commonly regulated by the balance between the attractive and repulsive interactions between the self-assembling components, involving e. g. van der Waals, electrostatic, hydrophobic, hydrogen bonding and entropic forces (molecules and mesoscale objects), and magnetic, capillary, and gravity forces (meso- or macroscale objects).

The degree of versatility, simplicity and flexibility of the various self-assembly methods and the ability to introduce specific functions with a high degree of spatial accuracy are important features for self-assembled materials to move beyond academic beauty and be interesting for applications. The preparation of self-assembled coatings can be traced back to the seminal work by Irving Langmuir and Katherine Blodgett prior to the Second World War. The development of a theory for adsorption on surfaces by Langmuir together with the design of a technique that both could measure the pressure-area isotherms of monolayers of amphiphilic molecules assembled at the air-liquid interface, and also be used to transfer these monolayers onto different substrates was a very important discovery that has had a profound effect on future work on self-assembled layers. The observation by Nuzzo and coworkers [4] that alkanethiolates can self-assemble on gold surfaces and result in well-ordered films sparked a large research interest in these so called self-assembled monolayers (SAMs). The relatively non-local nature of the thiol bond to the gold surface allows the much weaker, attractive interaction *between* the amphiphilic chains in the monolayer to control the film structure. The SAMs are a beautiful example of self-assembly as a crucial process in the formation of a dense monolayer.

Self-organization of nanoparticles into two- and three-dimensional superlattices has attracted much interest since the early work on iron oxide "super crystals" [5]. Indeed, understanding and optimizing the structures, at all length scales, of nanocrystal superlattices is an important step towards controlled design of novel nanostructured materials and devices. This work did not really gain momentum until methods to achieve shape and size control with a high fidelity was established in the end of 1980s by e. g. Brus, Steigerwald et al. [6] and Moerup, Thölen and Koch [5] mentioned above. The ability to assemble different nanocrystals with size-tunable optical, electronic and magnetic properties into well-defined structures opens up the possibility to study and develop new materials with tailored couplings between the constituent units [7,8]. Layer-by-layer (LbL) self-assembly of charged polymers has evolved as probably the most versatile technique to create multifunctional coatings on a wide range of substrates [9]. These are but a limited set of examples of self-assembly processes to produce novel materials that has been demonstrated in the last decades [10,11,12,13,14,15,16,17,18,19,20]. While the examples of self-assembly are plentiful, the attempts to model the self-assembly processes and systems is still in its infancy. This review attempts to be a guide and introduction to the field of computer modeling of self-assembly.

Introduction

Computer modelling of structural properties and dynamical processes in matter is based on physical and mathematical models. Ideally, we would prefer to use first-principles methods of quantum mechanics with no other input than atomic nuclei and the electrons spontaneously forming atoms, molecules and more complicated systems through local intermolecular interactions. Unfortunately this is possible only for systems of very modest sizes using rather drastic approximations and simplifications. We can always blame the computers for not been powerful enough but the truth is that the "first-principles" models would grow too complicated and out of hands very rapidly. We therefore use different physical models to describe matter at different length and time scales together with an increasing amount of empirical data as input to the models.

Atomistic computer simulations such as the classical Molecular Dynamics (MD) are currently run for molecular systems consisting of the order of 10^5 atoms, corresponding typically to a system size (length scale) of 5–10 nm. These simulations are typically extended to 10–100 ns (time scale). This may be enough for simulations of isotropic liquids of simple molecules, whereas for many complex systems, the requirements for the length and time scales are much more extensive. To extend the simulations to cover longer scales requires either simplifications of used models or choosing more suitable modelling technique [21]. In the first strategy the resolution of the model is reduced by removing certain degrees of freedom of the studied system which are not important or can be considered as fluctuations while keeping those degrees of freedom which are operational on the longer scales. This is

Self-assembled Materials, Table 1
Time and length scales covered by modelling methods reviewed in this article

Scale	Typical length	Typical time
sub-atomic	< 0.1 nm	∼ 1 fs
atomic	0.1–1 nm	1 fs–1 ns
meso	1 nm–1 μm	1 ns–1 ms
macro	1 μm–	1 ms–

generally called coarse-graining of the model leading to a much fewer interaction sites thereby speeding up the calculations. An alternative approach is to use a hybrid by keeping an accurate description for a selected area of importance and the environment is merely made to a fluctuating medium with more or less specific interactions. These techniques may include combining quantum and classical mechanics [22] or classical mechanics and hydrodynamic/continuum models.

In Table 1 there is a somewhat artificial division of matter on scales from sub-atomic (with nuclei and electrons) to macroscopic. There are modelling and simulation methods designed separately to these scales. Going from one scale to the next in the table requires either coarse-graining or fine-graining of the used model by building a bridge between two techniques. This means taking as much vital information as possible to the next level. Coarse-graining is obviously easier than fine-graining as the latter requires re-activating non-existing degrees of freedom from an under-defined system. Multiscale modelling is currently a hot topic and in a rapid development with the purpose to bridge several scales in a more systematic way.

The aim of this chapter is to describe a box of tools containing some selected modelling methods to study self-assembly. The tools should together cover the scales in the table in Table 1. We will give a short description of each of them together with some examples of applications. At the end of the chapter we show examples of how to combine several techniques to extend length and time scales in modelling. We will work from the "bottom-up" by starting with first-principles calculations on electrons and nuclei. Classical Molecular Dynamics (MD) and Monte Carlo (MC) are mainly used in atomistic simulations with empirically parameterized potential functions. Meso-scale simulations of aggregates and soft particles use coarse-grained models. We will discuss methods equally applicable on both short and long scales such as the kinetic Monte Carlo (kMC) and Phase Field (PF) modelling. All these techniques so far are examples of off-lattice modelling methods. We then discuss two lattice methods with some further development towards more off-lattice character, namely Random Sequential Adsorption (RSA) and Cellular Automata (CA). Finally we will discuss multiscale modeling before the paper ends with a discussion on future directions.

First-Principles Calculations and Simulations

Down in the bottom of the world of atoms and molecules, quantum mechanics should be used to treat the moving nuclei and electrons building up the matter. The electronic structure in atoms and molecules and molecules gives the first level of properties to all matter around us. Theoretical calculations utilize the so-called Born–Oppenheimer (BO) approximations which keeps the much heavier nuclei in fixed positions while the electrons adapt themselves pair wise in binding and non-binding molecular orbitals (MO), constructed as linear combinations of atomic orbitals (LCAO) using suitable basis sets (normally Gaussian functions or plane waves). Computational quantum chemistry is largely based on solving the Schrödinger equations interactively within the self-consistence field (SCF) theory until suitable convergence criteria are reached. The previously commonly used Hartree–Fock (HF) method was made computationally convenient in matrix form by Roothaan [23] in early 50s and allowed accurate calculations for larger and larger molecules at the same time as computers were made more and more powerful [24]. As there is no exact solution to many-body problems ($N > 2$) all modelling methods use approximations to reduce the problem to independent particle or pair problem [25]. Within the HF approximation the full N-electron problem is reduced to an independent N single-electrons moving in a mean-field of the other electrons. By increasing the size of the basis set a so called HF limit can be reached where the electron-electron repulsions are treated uncorrelated. As the electron correlation effects are very important for many molecular properties a number of post-HF schemes have been developed, for example Møller–Plesset (MPn ($n = 2, 4, \ldots$)), configuration interaction (CI) or coupled cluster (CC) [25].

The most affordable quantum chemistry methods are the semi-empirical MO methods [26,27,28]. Among the most popular are those based on the neglect of diatomic differential overlap (NDDO) approximation, such as Austin Model 1 (AM1) [29], Parametric Model 3 (PM3) [30,31] and modified NDO (MNDO) [32]. All these schemes start from HF-Roothaan equations [25] introducing the frozen-core approximation and using only the valence AOs. Some electron-electron integrals are simply neglected, some calculated approximately and some are replaced by empirical parameters. These methods are very

fast compared to *ab initio* methods but give often unreliable results; the errors tend to be unsystematic and therefore both relative energies and conformational energies become less reliable when calculated using semi-empirical schemes. Also hydrogen bonds are not accurately described. These shortcomings are found to be largely due to an unbalanced treatment of core-electron and electron-electron interactions. Orthogonalization corrections have been worked out to improve the NDDO schemes [33,34] but have given only a slight improvement. This unbalance is corrected in the recently published Extended NDDO scheme [35] where the old NDDO is expressed as zeroth order term. However, it still remains to be parameterized and packaged to a new semi-empirical scheme, which is a major undertaking.

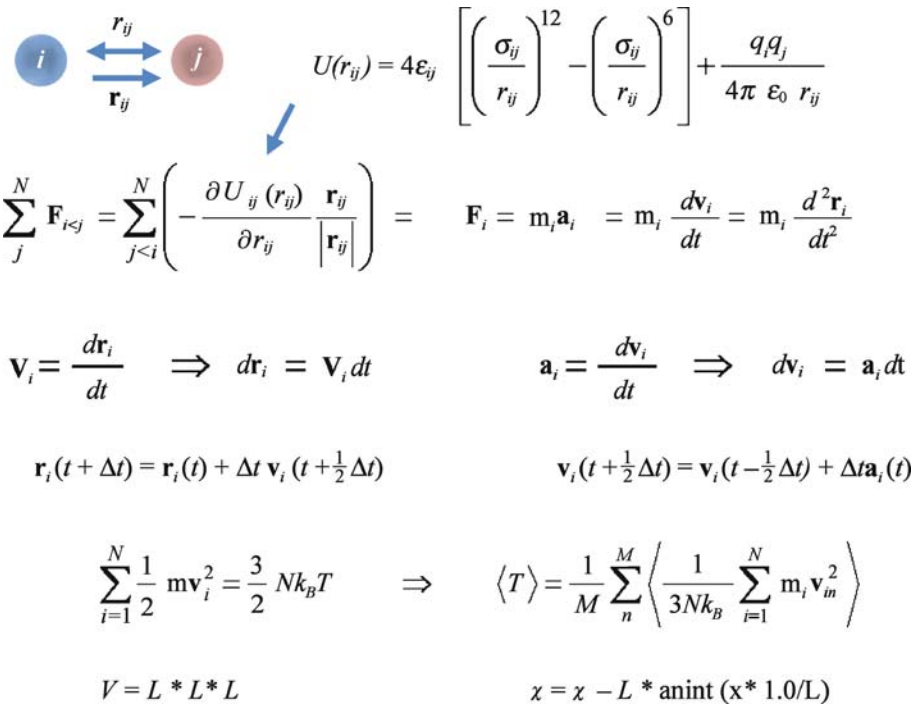
In both materials science and chemistry, the density functional theory (DFT) has recently gained the position as the most popular method [25]. It is basically as fast as the HF method but can incorporate a large part of the electron correlation effects. In DFT the energy is given as a functional of the electron density $\rho(\mathbf{r})$. The contributions to total energy come from the kinetic energy of the electrons, nuclear-electron attraction, electron-electron Coulombic repulsion and finally from combined exchange and correlation energy. The most difficult problem in DFT is to derive the exchange-correlation (XC) term. In materials science for solids and metals the so called local density approximation (LDA) is used as it is suitable in cases of slowly fluctuating electron densities [25]. LDA is not a good choice for chemical problems and so called gradient corrected methods like the generalized gradient approximation (GGA) are used where the XC term is not just functional of the electron density but also derivatives of the density with respect of the coordinates (x, y, z) [25]. The most popular XC functionals are currently those by Lee, Yang and Parr (LYP) [36] and by Perdew and Wang (PW91) [37]. In chemical applications the hybrid methods where the HF exchange energy is mixed the corresponding DFT term have gained popularity. Becke's three parameter hybrid [38] with the LYP functional, the so called B3LYP, is frequently used. Current DFT methods are based on formulation proposed by Kohn and Sham (KS) [39,40] by constructing the total electron density from a KS orbital. This gives conceptually similar equations as in HF method. Car and Parrinello (CP) presented a DFT based Molecular Dynamics method using plane-wave basis functions [41] which is the most popular *ab initio* simulation tool both in chemistry and materials science.

Recently a robust approximate DFT method is introduced where the electron-electron interactions in the total energy density functional in DFT are expanded with respect to a reference density and keeping the first order correction and neglecting the higher order corrections. This is called Harris–Foulkes functional [42,43] and is stable with respect to the used reference density as a small change in the reference density introduces an error only to second order term. Computational schemes based on Harris–Foulkes functionals are normally called *ab initio* tight-binding density functional theory (AB-TBDFT) methods, and were first introduced by Sankey et al. [44]. A highly efficient AB-TBDFT method is recently presented by Tu and Laaksonen [45,46] applicable on simulations of large systems. Various methods to incorporate quantum mechanics into molecular dynamics simulations are reviewed in [22].

First principles electronic structure calculations and simulations have a firm position in materials science. First-principles MD simulations to materials properties are reviewed in [47]. In the following a few examples are chosen from literature. Materials simulations using VASP gives a quantum perspective to materials science according to [48] using pseudo potentials (PP) and a plane wave basis set based on a finite-temperature local-density approximation. Density functional theory meeting statistical physics [49]: from the atomistic to the mesoscopic properties of alloys by combining DFT calculations with the so-called Cluster Expansion (CE) methods and Monte-Carlo (MC) simulations is reported. Determination of solid-state nanostructure from *ab initio* structure calculations of nano-structured materials using diffraction data in combination with distance geometry methods is explained in [50]. Tools, results and perspectives of quantum software interfaced with crystal structure databases are given [51]. Density-functional theory electronic structure calculations of static and elastic properties and *ab initio* molecular dynamics simulations for poly-atomic systems made possible in package reported in [52] and first-principles computation of material properties using the ABINIT software package in [53]. A primer to efficient tight-binding molecular dynamics is given by Colombo [54]. CONQUEST code for large-scale *ab initio* calculations in materials science is presented in [55]. Prediction of materials properties by *ab initio* computer simulations is reported [56].

Atomistic Molecular Dynamic Simulations

Molecular Dynamics (MD) [57,58,59,60,61,62,63,64,65] is, in its simplest formulation, numerical integrations of Newton's second law (acceleration ($\mathbf{a}(t)$) of a particle is equal to the force ($\mathbf{F}(t)$) acting on a particle divided by the mass ($\mathbf{m}$) of the particle) applied on a collection of inter-



Self-assembled Materials, Figure 1
Molecular Dynamics simulations in a nutshell

acting particles. Integrating the acceleration of each particle gives the velocity ($\mathbf{v}(t)$) and a subsequent integration of the velocity gives the actual position ($\mathbf{r}(t)$) of the particle. Thereby the trajectory of the particle can be determined. The forces momentarily acting on the particles (masses) are calculated as negative spatial derivatives of the potential functions $U(\mathbf{r})$ (force fields) describing the attractive and repulsive interactions between the particles (including the internal degrees of freedom): $\mathbf{F} = -\nabla U(\mathbf{r})$. Normally only the pairwise intermolecular interactions are included as an approximation in classical MD simulations of liquids and solutions. The number of particles (N) in current simulations is from thousands to hundreds of thousands and is a compromise set by the available computing power. Particles are inserted into a cell (box) with a volume V. The shape of the cell is preferably chosen to have translational symmetry so that replicas of the cell make a pseudo-infinite periodic system where particles are synchronously able to leave and enter the cells according to periodic boundary conditions (PBC). This keeps the particle density constant. So called minimum image convention connected to the PBC is applied to avoid artificial condensation effects (for more details see Table 1 and the text in connection to the figure). In Newtonian dynamics the energy (E) is expected to be conserved at equilibrium conditions and the simulation corresponds to a microcanonical (NVE) ensemble. By using suitable Hamiltonians containing thermostats canonical (NVT) ensembles can be constructed and with corresponding barostats isobaric (NPT) ensembles. Several other ensembles can be constructed for MD simulations as well.

Historically, Alder and Wainwright introduced the method of Molecular Dynamics 50 years ago by simulating a system of hard spheres [66,67] observing a phase transition, while Rahman [68] carried out the first simulation for a liquid system using argon atoms in 1964 giving fairly accurate diffusion coefficient. Verlet was able to obtain thermodynamic state functions noble gases [69]. A very important pioneering work for molecular systems was that of Rahman and Stillinger for liquid water in 1974 obtaining radial distribution functions in close agreement with experiments [70], which also started a real boom in computer simulations. MD simulation is a powerful method to study structural, thermodynamic and dynamical properties in gases, liquids and solids. Using MD the motion of molecules can be followed to gain a better understanding of chemical reactions, fluid flow, phase transformations, droplet formation, and other physical phenomena. MD is a statistical mechanical method (both equilibrium and non-equilibrium).

Figure 1 gives all the information needed to write a simple computer program to run MD simulations of N

charged particles in a cubic volume V. *First row*: Pairwise interaction potential consisting of Lennard–Jones (van der Waals) and Coulombic terms. Particle sizes are determined by the σ parameter and strength of the van der Waals attractions is given by ε. Particle charges are given by q in the Coulombic term and pairwise distances are given by r_{ij} (scalar) and the $\mathbf{r}_{ij}$ (vector). *Second row*: Forces acting on each particle are obtained as the negative spatial derivative of the potential function given in first row and are equal to particles mass times its acceleration (Newton's II law). *Third row*: By integrating the fluctuating time-dependent acceleration gives particle's velocity and a subsequent integration of the velocity gives the momentary position of the particle. *Fourth row*: Example of a numerical finite difference algorithm called the "Verlet leap-frog" to carry out the integration. Δt is the time step and is normally of order of a few femto seconds depending of the particle masses and the density of the system. Observe that the velocities in the Verlet leap-frog scheme are obtained as half time steps to increase the accuracy and stability of the algorithm. *Fifth row*: Example of how compute the average temperature from simulation data using the kinetic energy and the equipartition principle (each translational degree of freedom contributes 1/2 kT to the kinetic energy). M is the total number of time steps in the simulation. Other average quantities, for example the total energy (potential+kinetic energy) is calculated similarly. *Sixth row*: For a cubic simulation cell of a volume V (containing the N particles) the side length is L. Periodic boundary conditions (PBC) are easily applied on the positions (to keep the particles inside the cell) by the simple formula where the anint is a function which truncates the floating point number keeping only the integer digit. Example here is given for the x-components and should be done same way for y- and z-components. PBCs are applied on the distances as well according to the minimum distance convention to avoid artificial condensation effects during the simulations. Large scale simulations are carried out using parallel computers with a large number of processors. This requires computer software adapted to parallel architectures [71,72]. A large number of simulation packages are available from the authors developing them. Most of them are publicly available while some (most often those who are made particularly user-friendly by graphical interfaces) require a license to be purchased.

MD is an example of N-particle simulation techniques, a wide-spread method found virtually in all areas dealing with interacting dynamical particles from solvated electrons to protein-lipid interactions and from orientations of liquid crystals to formation of galaxies [58,59,60,61,62,63,64,65]. It is also the main simulation method in materials science from metals to zeolites and from polymers to colloids.

Self-assembled thiol monolayers on Au(111) made hydrophilic with polar end groups are studied with STM and MD simulations [73] to examine the role of the formation of hydrogen bonds between the molecules in the layer and with polar co-adsorbates (water and solvent). MD have been performed [74] to investigate the two-dimensional structure of organosilane self-assembled monolayers (SAMs). Unlike alkanethiol SAMs, the arrangement of molecules in organosilane SAMs is not crystalline. Simulations performed for structures with different bonding networks in the polysiloxane layer shows that the ratio of hydrogen bonds has a profound effect on conformations and strain energies. Water density profile close to spherical and planar hydrophobic objects is studied [75] using MD simulations. A substantial increase of the depletion layer thickness was found with the temperature. High electrostatic surface potential presumably plays an important role in the presence of charged solutes possibly promoting adsorption into the interfacial layer. MD investigations [76] of the morphology and structure of monolayer and multilayer of chiral molecule *N*-stearoyl-*L*-glutamic acid (C-18-L-Glu) self-assembled on a mica surface show that hydrogen-bonding effects are the major driving forces in the layer formation proposing a multilayered model for the self-assembling. Static and dynamic properties of 2:1 layered silicates ion exchanged with alkyl-ammonium surfactants are studied using MD simulations [77] providing the structure and dynamics of the intercalated surfactant in agreement with experiments. Structure and properties of self-assembled monolayers (SAMs) of a bi-stable[2]rotaxane on Au(111) surfaces as a function of surface coverage are reported based on atomistic molecular dynamics (MD) studies with a force field optimized from DFT calculations with several experiments validating the predictions [78]. To develop strategies to self-assembly of nanostructured materials MD simulations for telechelic molecules composed of two polyhedral oligomeric silsesquioxane cages connected by one hydrocarbon backbone dissolved in liquid normal hexane were carried out [79]. MD simulations of nanoparticle self-assembly at a liquid-liquid interface are carried out showing in situ formation of clusters and migration of both single particles and clusters from the water phase to the trichloroethylene phase [80].

Among similar modelling works related to self-assembly could be mentioned recent advances of classical density functional theory with emphasis on applications to quantitative modeling of the phase and interfacial behavior of condensed fluids and soft materials, including

colloids, polymer solutions, nanocomposites, liquid crystals, and biological systems are summarized [81]. Molecular mechanical methods are used to calculate strain distribution in self-assembled interfacial misfit dislocation arrays in highly mismatched III–V semiconductor materials in good agreement with experimental data using cross section high-resolution transmission electron micrograph images and also with other theoretical values [82]. Molecular self-assembly at equilibrium is studied using a family of simple directional and short-range van der Waals potentials giving rise to the self-assembly of linear polymeric, random surface, tubular, and hollow icosahedral structures. Dipolar potential with its continuous rotational symmetry about the dipolar axis contributes to chain formation, while higher multipoles lead to the self-assembly of open sheet, nanotube, and hollow icosahedral geometries [83]. To further understand the mechanism behind Layer-by Layer assembly processes requires even theoretical work, including thermodynamics calculations and molecular dynamics simulation as many kinds of physicochemical molecular interactions, including hydrogen bonding, charge transfer interactions, and stereocomplex formation are involved [84].

Monte Carlo Simulations

With modern computers the Monte Carlo technique has evolved from a classical numerical algorithm to solve multi-dimensional definite integrals based on statistical sampling to a very powerful simulation technique to study a large variety of complex systems with interacting particles [59]. The name "Monte Carlo", given by Stanislaw Ulam [85], refers to the capitol of Monaco at Mediterranean, famous for its casinos and roulette tables. The corresponding "roulette" in the MC method is the random number generator supplying a sequence of numbers normally between 0 and 1 from a uniform random distribution (rand[0, 1]). MC methods can be used, for example, to compute equilibrium properties of classical many-particle systems. MC was in fact among the first applications run in the first computers in early 50s. The first liquid simulation using MC was carried out by Metropolis and coworkers [86]. As molecular computer simulation methods MC and MD are very similar until the particles start to move. The simulated systems in both MC and MD are built in the same way by inserting the N particles in a simulation cell of volume V, applying the periodic boundary conditions and choosing a force field to describe the interactions between the particles and the cut-off distances for short and long-range interactions. Ewald summation can be employed if found necessary. By fixing the temperature T we can perform the MC simulation in a canonical (NVT) ensemble which is also the simplest choice in molecular MC simulations. Recall that the NVE (microcanonical) ensemble is the natural choice in the case of Newtonian MD simulations. While the MD simulation method is completely deterministic in the sense that the particles always follow exactly same trajectories if same input is given, the Monte Carlo (MC) method is stochastic and the particles are moved as a random walk. How the particles are moved depends on the studied systems. In the case of small and rigid molecules (as a liquid) the molecules (center-of-mass) is translated (displaced) in x, y and z direction of a small distance one molecule at a time rotating it slightly around its principal axes. For larger and flexible molecules the individual atoms are slightly moved and in addition different types of specific movements are performed mimicking somehow the real motion of the molecule. For example MC simulations of polymers are undergoing residual "flips", "rotations", "crankshaft", "slithering snake" moves and "end-bridging" plus many more. Probably the simplest MC technique to simulate interacting particles at a desired temperature is Metropolis Monte Carlo [86].

The mother of all Monte Carlo methods is the Markovian master equation:

$$\frac{\mathrm{d}P(\vec{r}_i,t)}{\mathrm{d}t} = \sum_{\vec{r}_j} W(\vec{r}_j,\vec{r}_i)P(\vec{r}_j,t) - \sum_{\vec{r}_j} W(\vec{r}_i,\vec{r}_j)P(\vec{r}_i,t) \quad (1)$$

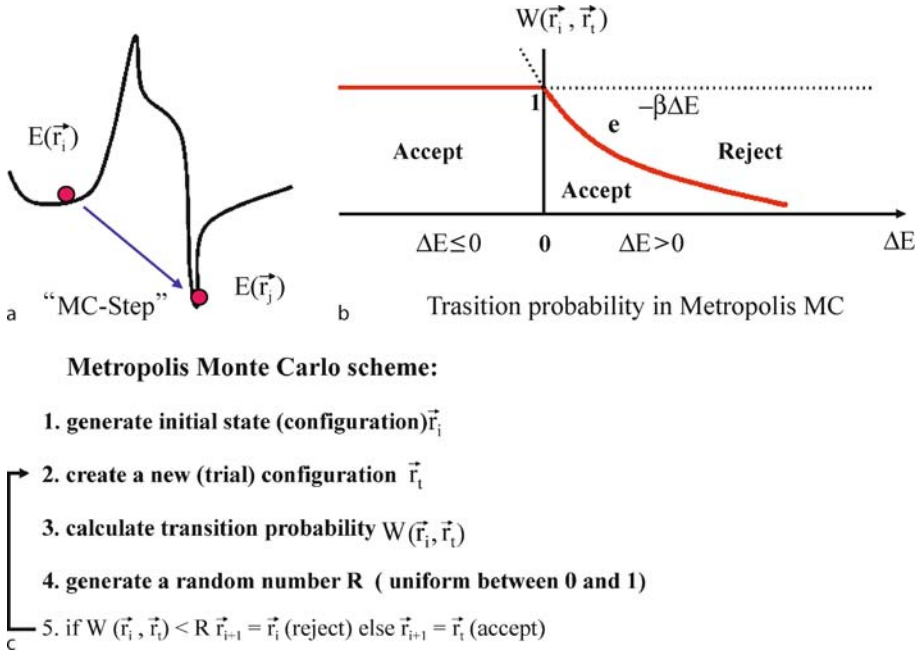
where $P(\vec{r},t)$ is the probability that the system is in state $\vec{r}$ at time t (time-dependent distribution of states/configurations) and $W(\vec{r}_i,\vec{r}_j)$ is the transition probability/rate per unit time for the system to undergo a transition from i-state to j-state. All MC techniques can be seen as methods for solving the master equation. At steady state (stationary or equilibrium):

$$\frac{\mathrm{d}P(\vec{r}_i,t)}{\mathrm{d}t} = 0; \quad t\to\infty \wedge P(\vec{r}_i,t)\to P(\vec{r}_i) \Rightarrow \sum_{\vec{r}_j} W(\vec{r}_i,\vec{r}_j)P(\vec{r}_i) = \sum_{\vec{r}_j} W(\vec{r}_j,\vec{r}_i)P(\vec{r}_j)\,. \quad (2)$$

The detailed balance (microscopic reversibility) is valid. The canonical contribution is:

$$\sum_{\vec{r}_j} W(\vec{r}_i,\vec{r}_j)\mathrm{e}^{-\frac{E(\vec{r}_i)}{k_\mathrm{B}T}} = \sum_{\vec{r}_j} W(\vec{r}_j,\vec{r}_i)\mathrm{e}^{-\frac{E(\vec{r}_j)}{k_\mathrm{B}T}} \quad (3)$$

$$W(\vec{r}_i,\vec{r}_t) = \min\left(1,\mathrm{e}^{-\frac{E(\vec{r}_t)-E(\vec{r}_i)}{k_\mathrm{B}T}}\right)\,. \quad (4)$$



Self-assembled Materials, Figure 2

a Schematic illustration of a Monte Carlo move. b Transition probability *W* in Metropolis Monte Carlo. c Computational scheme for Metropolis Monte Carlo

However, it is not possible to calculate W from this equation. MC methods therefore use different functional forms for W. Transition probabilities in thermodynamical MC (like Metropolis) do not need to have *any* relationship to the dynamics (energy barriers) of the system. To assign real time to MC steps (Fig. 2a) is not possible. Conventional MC methods are only applied to sample systems in (or close to) thermal equilibrium. Functional form of transition probability in Metropolis MC:

In standard Metropolis MC the looping should not be done until accepted. All moves with probability W larger than 1 are accepted (the energy goes down) while if W is less or equal to 1 the move is accepted if the random number is less than the Boltzmann factor in Eq. (4) (see also Figs. 2). A rule of thumb is to choose the size of the displacements in such a way that the acceptance/rejection ration is close to 1/2 which might be considered as the upper limit while the lower limit is roughly 1/3 [87]. Several particles should not be moved simultaneously in Metropolis MC. There is no direct correspondence between a Monte Carlo step and real time but it is possible to use experimental information of for example diffusion to calibrate the MC simulations to obtain a rough estimate of real time.

Some examples of MC works in connection to self-assembly are for example the MC (both multi-scale NVT and grand canonical) simulations carried out to investigate nano patterns an self-assembly of surfactants inside SWCNT systems [88] and Monte Carlo simulations of gold nanocrystals and (111) slabs covered with alkyl thiols are carried out with and without explicit solvent (*n*-hexane) at $T = 300$ K. Inclusion of explicit solvent is found important in [89]. Self-organization and chain-forming of large ensembles of nanoparticles is studied by combining DLVO theory and MC simulations [90]. The growth of linear agglomerates is kinetically controlled by a high activation barrier from all of the directions except one at end of the chain. Adsorption of different model amphiphiles in apolar and polar solvents is investigated using MC simulations with a coarse grained model. As coating agents the surfactants with a single hydrocarbon tail or two branches are found to protect better the particle surfaces than amphiphiles with three or more branches [91]. Lattice Monte Carlo simulation is used to investigate the equilibrium between free surfactant molecules in aqueous solution and those adsorbed layers on structured solid. The solid surfaces are composed of hydrophilic and hydrophobic surface regions [92]. Lattice MC simulation of self-assembled ordered hybrid materials is reported and a comparison of structural characterization of the different phases using aggregate size distribution, density profiles, and radial distribution functions [192]. MC simulations are used to study self-assembly of symmetric diblock copolymers

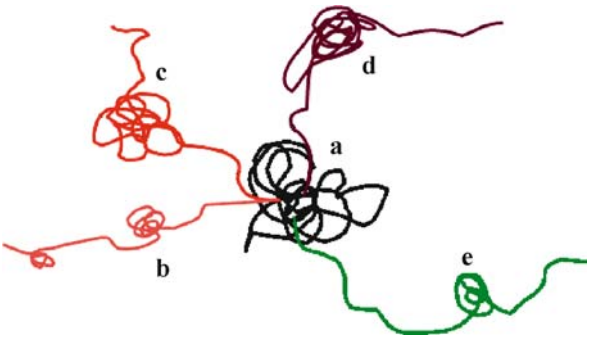
in confined state resulting novel self-assembled strip, circle, core-multishell, and multi-barrel layer structures [94]. Self-assembly and adsorption on hydrophobic surfaces are studied by MC methods in [95,96,97,98].

Kinetic Monte Carlo Simulations

Ideally we would like to have a general computer simulation method operational on atomistic level of matter which we can used for both short and long length and time scales. We wish to use it to simulate general dynamical processes with many different types of transitions between various states the system populates during its evolution towards a more permanent equilibrium state. The method of molecular dynamics (MD) is often the first choice. Using MD we model our system by applying realistic physical conditions. We also have correct real time in our simulation procedure. However, the time step Δt in atomistic MD simulation to numerically solve the equations of motion has to be chosen according to the frequency ($\omega_{\max}$) of the fastest dynamical event in the molecular system this being in practice $\Delta t \ll (\omega_{\max})^{-1}$. This makes atomistic MD simulations very inefficient if the events of interest occur beyond nanosecond time scales. Unfortunately a vast amount of motional modes and dynamical events take place beyond the capability of MD simulations even with the fastest computers around.

We can use Monte Carlo (MC), such as the common Metropolis method discussed above. We can build the molecular system similarly as in MD and impose the desired conditions onto it. We start simulations and pick up the particles randomly and one at a time moving them in small displacements based on the Boltzmann probability. Again, we realize that as the size of the system increases together with the number of motional degrees of freedom our simulation starts to take too long time. Besides, as mentioned in the previous chapter, we do not even know how to add the time into a Metropolis simulation as different events have different characteristic times for their dynamical behavior. MC as we perform it is normally only for sampling ensembles.

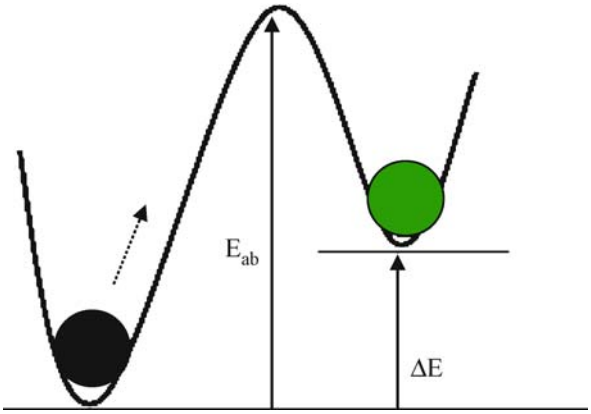
We now consider using a method called the kinetic Monte Carlo (kMC). Using kMC dynamical processes can be simulated with real time incorporated and all possible time scales can be covered easily. This sounds simply too good to be true so what is the "catch" here? Before going into details of kMC method (in fact there are many variants of it) we try to trace its origin by going back almost four decades ago and to computer simulations of vapor deposition on two-dimensional lattices [99] and crystal growth with surface diffusion [100]. These two works



Self-assembled Materials, Figure 3
Accessible stationary states (b–d) from state a

carry much of the basic underlying idea behind several kMC schemes developed later. The paper by Bortz, Kalos and Lebowitz (BKL) presents elegantly the theory and algorithms for a method finding immediately a strong position among physicists [101]. The BKL scheme is the solid framework in many kMC programs used today. In almost parallel and from purely chemical point of view Gillespie presented his stochastic simulation method for coupled chemical reactions [102] with the starting point being that on atomistic level the time evolution of chemical systems such as chemical reactions can be seen as both discrete and stochastic rather than continuous and deterministic as the text book kinetic equations are presented for us. Although the BKL method and the Gillespie method appear to have been developed in separate communities they both lead to same family of kMC methods [103]. The name kinetic Monte Carlo appears first as late as in 1992 [104].

Kinetic Monte Carlo is also based on solving the Markovian master Eq. (1) discussed in the previous Chap-



Self-assembled Materials, Figure 4
Transition from state a to state b

ter. In kMC we assume that the system is in a (stationary) state and that a number of other states are accessible from the current state (Fig. 4). A new state is randomly chosen among the neighboring states and a transition is made to this new state. The real time spent in the state before transition is added to the time development and can be seen as the time step in kMC. The time spent in a state (Fig. 4) is related to the transition rate:

$$r_{ab} = \upsilon e^{-\frac{E_{ab}}{k_B T}} \tag{5}$$

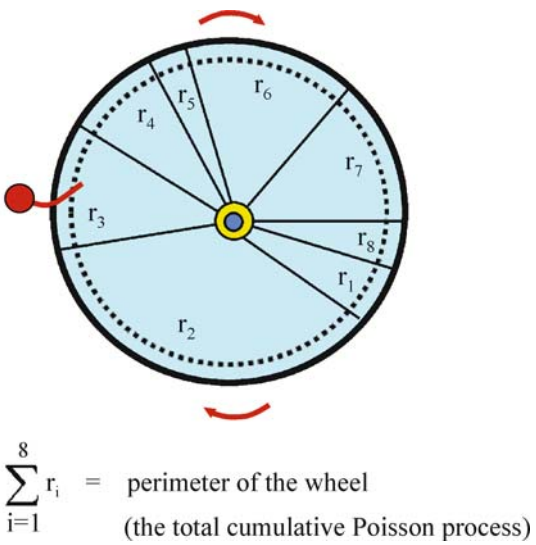
– where υ is the "attempt frequency". The time step is then calculated as:

$$\Delta t = -\frac{\ln(\text{rand}[0,1])}{\sum_{i=b,c,d,\ldots} r_{ai}} \tag{6}$$

kMC requires that the simulated events are Poisson processes. There should also be a dynamic hierarchy among the processes. The real time increments (time steps) should be able to be calculated based on Eqs. (5)–(6). In practice the detailed balance (Eq. (2)) is not strictly required as the systems can be far from equilibrium. If the transition probabilities are: (i) independent on previous history, (ii) the same at all times and (iii) a uniform function of time, the the transitions are Poisson processes. A kMC transition should not have any effect on any other transition. Also the time when the next transition will occur should be independent from the occasion of the previous transition. To illustrate kMC we may consider a "wheel of fortune" (Fig. 5) where the sectors correspond to individual Poisson processes (events) and are not necessarily evenly divided. The perimeter of the wheel is the cumulative process (also a Poisson process). The corresponding dynamical hierarchy in Fig. 5 is as:

$$[0 \leqslant r_5/r_{\max} \leqslant r_8/r_{\max} \leqslant r_1/r_{\max} \leqslant \cdots \leqslant r_{\max}/r_{\max} \leqslant 1]\,.$$

The whole spectrum of rates exists and the fastest rate (r_2 in Fig. 5) gives the highest probability. Note, the Metropolis MC does not fullfill dynamic hierarchy as all $\Delta E \leqslant 0$ transitions are given probability 1. After a kMC time step we are in a new state and we have a new "wheel" like the one in Fig. 5 but possibly with a different number of sectors corresponding to the new states with a finite probability to make transitions to. A new ensemble of processes is obtained after each transition. This also means that the time step in kMC changes continuously from one step to another. There are no restrictions how short or long the time step is: it can be for example a few femtoseconds or several hours. It simply follows the evolution of the system during the simulation. Time steps can be shorter in the beginning of simulation due to (initially) fast moving objects. Also, slow moving objects may break to smaller ones moving faster. Kinetic MC sounds like the optimal simulation method which can cover all possible time and length scales. The "catch" here is that the *user* has to supply both the states and the transition rates to the new states after each time steps. The kMC is not a self-going scheme in the same way as the other simulation methods are. This of course limits its use currently. Transitions rates can be obtained for example from quantum chemical calculations, MD simulations, transition state theory and experiments. As all the information of the system and rates is available the method has many attractive features: No thermodynamical equilibration is needed and no simulation of "dead dynamics" like in MD (event/step is always guaranteed). kMC is suitable for "driven systems" (irradiation, high pressure, etc.). kMC has been used in studies surface diffusion, crystal growth, molecular beam epitaxy growth, defect mobility, vacancy diffusion, reactions of surfaces, diffusion in zeolites, clustering dynamics, nucleic acid hybridization, protein folding, etc. The number of new application areas is growing rapidly. kMC can be easily incorporated in multiscale simulations schemes.



Self-assembled Materials, Figure 5
The Poisson "wheel of fortune" to illustrate the kMC

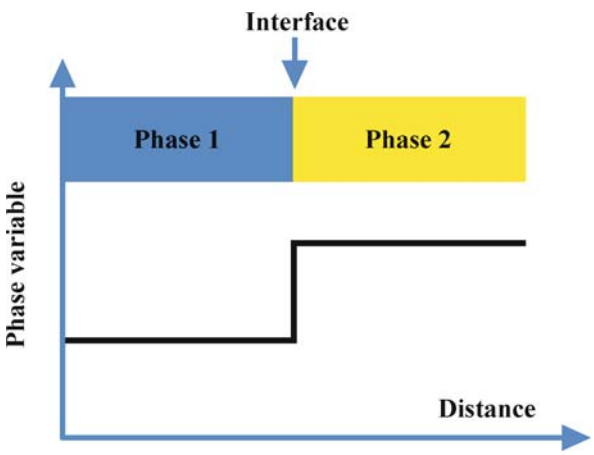
Kinetic MC is used to determine the critical layer thickness for misfit dislocations using deposition flux, temperature and a pairwise interaction potential between the particles as parameters [105]. Kinetic MC simulations are used to study the deposition rate dependence of nanopattern formation on periodically strained surfaces. The optimum nanopattern quality depends on surface strain field, temperature and deposition rate and the amount of material deposited [106]. Self-assembly of the

elementary building blocks of nanophase materials are studied using both a kinetic mean-field model and a mesoscopic approach in which self-assembly is viewed as an encounter-controlled process on a discrete lattice [107]. A 3D kMC model with a Green's function-based long-range strain energy contribution and with an up-down ratio for atoms to jump out of the plane of the surface is developed to simulate the growth of self-assembled quantum dot islands [108].

Phase-Field Modeling

Classical (macroscopic) theories & models assume sharp interfaces (infinitely thin with a specific surface tension) between phases and thermodynamical variables such as T, P, concentration, etc. are used independently in each phase obeying the phenomenological rules (schematically shown in Fig. 6). Phase diagrams are common tools in chemistry text books and traditionally the compositional and structural evolution of the phase regions have been treated mathematically using distinct interfaces with suitable boundary conditions at the interfaces. The idea of a *sharp* interface is old going back to early 19th century and Young, Laplace and Gauss. Gibbs later presented the interface as a dividing surface so that the properties were distinctly different on both sides of the interface (Fig. 7). Subtraction of the upper curve (real system) from the idealized system below gives the corresponding excess quantity (here concentration) and sharp mathematical surface line. The Gibbs dividing surface has zero volume but non-zero excess quantities.

Also the idea of a *diffuse* continuous interface (Fig. 8) goes back to 19th century when it was thought that gas and liquid states were simply distant points of the same condition of matter and it was possible to move from one phase to another along a continuous path. This picture was established both experimentally by Andrews and theoretically by Thompson and van der Waals [109]. Compared to singularly sharp interface model there are many benefits with using a diffuse inteface with a finite interface volume. All field variables and corresponding equations can be defined simultaneously for the whole system. Diffuse interface allows modelling of various types of physical variables, systems and effects in a flexible and efficient way. Diffuse inteface exists on atomic scale as the surface layers are rough on atomic scale rather than mono-molecular. The density profile at the interface becomes larger than the atomic length scale due to the roughness.



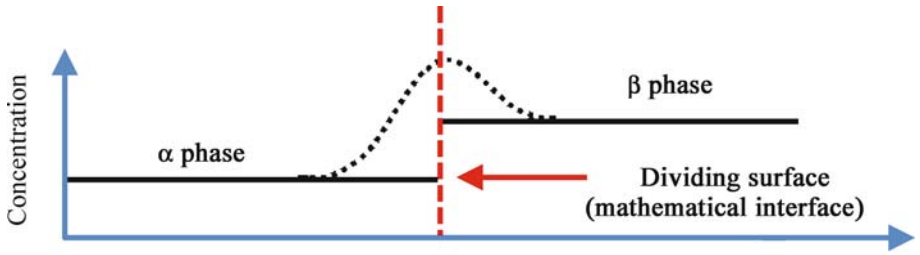
Self-assembled Materials, Figure 6
Sharp limit between two phases

Phase field modelling gathers all categories of diffuse interface models to describe a wide variety of materials phenomena (for excellent reviews, see [110,111,112,113,114,115,116,117]). It can describe easily a large number of phenomena from nucleation to solute trapping as different physical effects can be easily included [115]. It is based on two assumptions: (i) There exists a continuous phase field variable $\phi(x, t)$ which characterizes the phases and interfaces of the system at any point point in space and time. In bulk phases it is given a constant value. (ii) The total (Helmholtz) free energy, $F[\phi, \dots]$ can be given as a functional of the phase field variable ϕ together with any other thermodynamical variable, for example temperature, concentration etc. Even other variables can be introduced depending on the studied system. A general form of F is given as [113]:

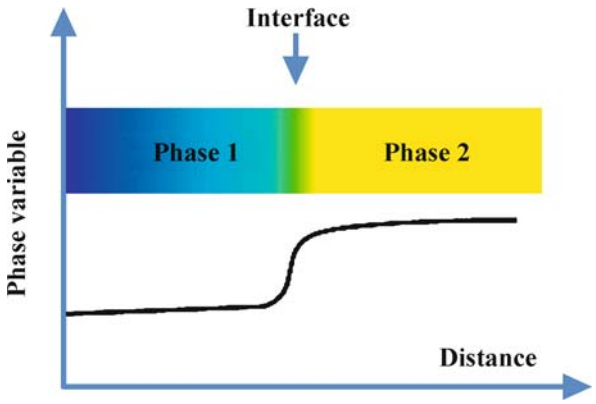
$$F = \int \Big[f(c_1, c_2, \dots, c_n, \eta_1, \eta_2, \dots, \eta_p) + \sum_{i=1}^{n} \alpha_i (\nabla c_i)^2 + \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{p} \beta_{ij} \nabla_i \eta_k \nabla_j \eta_k \Big] \mathrm{d}^3 r + \iint G(r - r') \mathrm{d}^3 r \mathrm{d}^3 r' \tag{7}$$

c denotes *conserved* variables (satisfying local conservation conditions)
η denotes *non-conserved* variables.
f is local free-energy density,
α and β are coefficients for energy gradients.

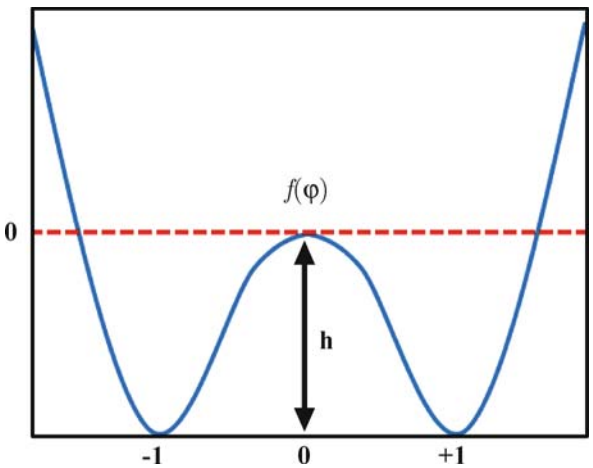
The first integral in Eq. (7) contains the local short-range interactions to the free energy, while the second integral represents the non-local long-range interactions, such as, elastic, electrostatic, polar, etc. Various phase-field models differ in their ways to treat all these contributions



Self-assembled Materials, Figure 7
Gibbs dividing surface



Self-assembled Materials, Figure 8
Diffuse continuous interface between two phases



Self-assembled Materials, Figure 9
Simple double-well free energy density function

to the total free energy functional. The local free energy f is the most important term in the phase field models:

$$f = f(c_1, c_2, \ldots, c_n, \eta_1, \eta_2, \ldots, \eta_p) \, .$$

A large variety of these functions have been proposed. In many phase-field models (like for solidification) simple double-well potential functions are used for local free density like the one given in Eq. (8) and in Fig. 9:

$$f(\varphi) = 4h\left(-\tfrac{1}{2}\varphi^2 + \tfrac{1}{4}\varphi^4\right) \, . \tag{8}$$

The variable ϕ in Eq. (8) is normally called order parameter. It varies smoothly from one phase to another, distributing interfacial forces and other interactions over the interface region [8]. We can often use it in a simple double-well potential for a case where the values $\phi = -1$ and $+1$ represent liquid and solid state, respectively, h is the barrier height between the two states. We may also have the phase field $\phi = -1$ and $+1$ representing two different heterogenous mixtures during an isostructural (spinodal) decomposition while 0 is a single homogeneous phase and h is the driving force for the transformation. Other types of similar simple local free energy density functions can also be found in literature.

Some other forms of local free energy profiles are for example the "double-obstacle" potential:

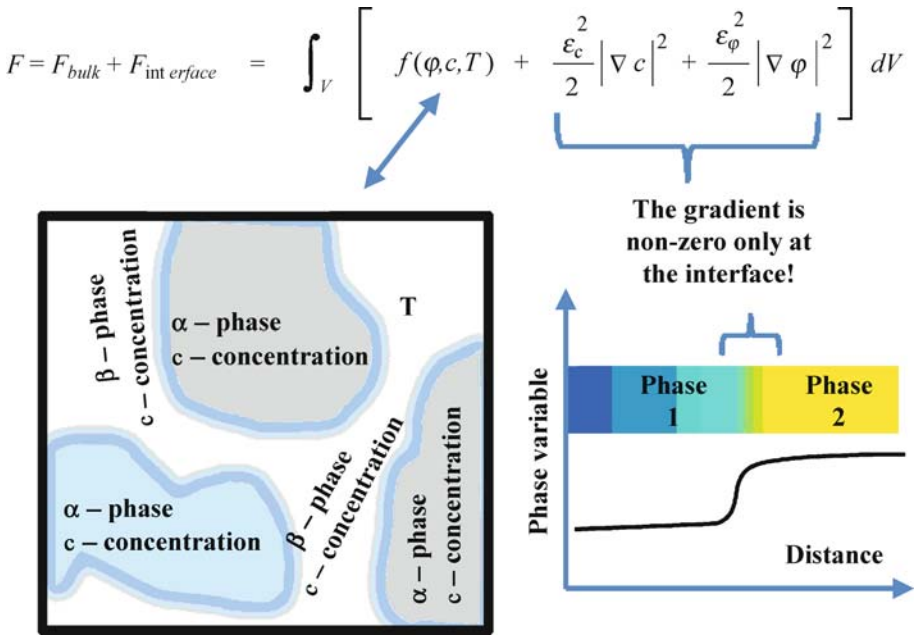
$$f(\varphi) = h(1-\varphi^2) + I(\varphi)$$
$$I(\varphi) = \begin{cases} \infty & \text{if } |\varphi| > 1 \\ 0 & \text{if } |\varphi| \leqslant 1 \, . \end{cases} \tag{9}$$

Or a "crystalline" potential having infinite number of minima:

$$f(\varphi) = h \sin(\pi\varphi) \, . \tag{10}$$

Often additional phase field parameters like temperature are needed, T = temperature, T_m melting point and α is a positive constant:

$$f(\varphi, T) = 4h\left(-\tfrac{1}{2}\varphi^2 + \tfrac{1}{4}\varphi^4\right) + \frac{15\alpha}{8}\left(\varphi - \tfrac{2}{3}\varphi^2 + \tfrac{1}{5}\varphi^4\right)(T - T_\mathrm{m}) \, . \tag{11}$$



Self-assembled Materials, Figure 10
Two main contributions to free energy: bulk and interface

Often coupling of phase field variables are needed – like in the following function to describe the spatial distribution for grain growth (for grains with different orientations):

$$f(\varphi_1,\varphi_2,\ldots) = 4h\left(-\tfrac{1}{2}\sum_i \varphi_i^2 + \tfrac{1}{4}\sum_i \varphi_i^4\right) + \alpha \sum_i \sum_{j>i} \varphi_i^2 \varphi_j^2 \,. \quad (12)$$

When $\alpha > 2h$ (in Eq. (12)), infinite number of of minima are located at $(1,0,\ldots)$, $(0,1,\ldots)$, $(-1,0,\ldots)$, etc., representing possible orientations of grains in a polycrystal. Functions in Eqs. (9)–(12) are discussed in [113] where several other types local free density functions are found. In general there are a large number of free energy functions like those above, all designed for specific applications and studies of phenomena. For many studies like solid-state transformations the energy functions are made of well-defined physical order parameters and use required symmetry operations of the studied high temperature phases. In general, the interfacial energies should be seen as anisotropic due to the crystalline feature of solid phase [113]. Multiple order parameter models have been proposed using both thermodynamical and geometrical formulations [114].

There are two main contributions to the total free energy; bulk and interface as illustrated in the example in Fig. 10.

Consider the isothermal process from the example in Fig. 10 with a decreasing total free energy F. At equilibrium, if the gradient coefficients are constants, the variational derivatives of F satisfy following equations [114]:

$$\frac{\delta F}{\delta \varphi} = \frac{\delta f}{\delta \varphi} - \varepsilon_\varphi^2 \nabla^2 \varphi = 0 \quad (13)$$

$$\frac{\delta F}{\delta c} = \frac{\delta f}{\delta c} - \varepsilon_c^2 \nabla^2 c = \text{constant} \,. \quad (14)$$

The Eq. (13) is for non-conservative variables while Eq. (14) is for conservative variables. Equation (14) is constant as the amount of solute is constant in the volume (concentration is one of the conserved quantities).

Time evolution of the phase field is assumed to be proportional to variation of the total free energy functional with respect to the order parameter ϕ [110]: $\partial\varphi/\partial t \propto L(\delta F/\delta\varphi)$. Where L = partial differential operator ($L(0) = 0$). The phase field equations are given:

Cahn–Hillard equation [118,119]:

$$\frac{\partial c}{\partial t} = \nabla M \nabla \frac{\delta F}{\delta c} = \nabla M \nabla \left(\frac{\delta f}{\delta c} - \varepsilon_c^2 \nabla^2 c \right) \quad (15)$$

– for conserved variables such as: concentration, molar fraction, mass, etc.

Allen–Cahn equation [120,121]:

$$\frac{\partial \eta}{\partial t} = -L \frac{\delta F}{\delta \eta} = -L \left(\frac{\delta f}{\delta \eta} - \varepsilon_\eta^2 \nabla^2 \eta \right) \quad (16)$$

– for non-conserved variables such as: order parameters, phase fields, etc. This equation is also known as the time-dependent Ginzburg–Landau equation.

The evaluation of the phase field variables can be obtained by solving the kinetic Eqs. (15) and (16) above. Finite difference methods with explicit time-stepping are normally used by constructing a uniform grid. Discussion of numerical methods and stability of the computational schemes are given in [122].

The areas of applications of phase field simulations are increasing rapidly. Solidification and solid-state transformations are examples of the use phase field modelling as well as coarsening, grain growth and dislocation dynamics. A list of several applications can be found in Table 1 of [113].

A further development of phase field theory, called phase-field crystal (PFC) method, was present by Elder and Grant [123] for elastic and plastic deformations, free surfaces and multiple crystal orientations.

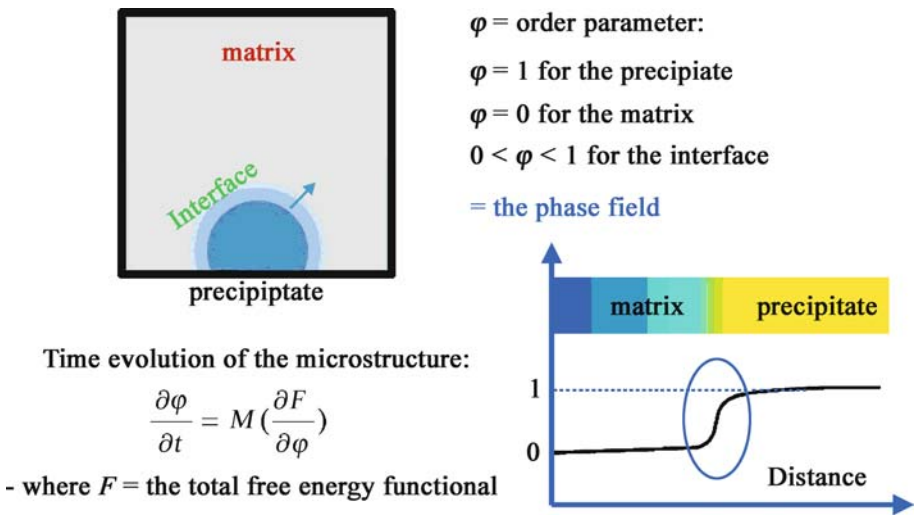
Continuous phase field model by Lu and Suo have been found applicable on self-assembly and monolayer growth on surfaces describing all required components: the phase separation, phase coarsening and phase refining processes [112]. Lu and Salac [124] have developed a phase field model to simulate pattern formation due to electric dipole interactions among adsorbate molecules revealing a unique self-assembly behavior. Wang [125] has developed a diffuse interface field approach (DIFA) which is capable to explicitly describe short- and long-range interactions and treat arbitrarily shaped particle packing processes and other related applications.

Molecular dynamics simulations can be combined with phase field modelling. Two basic approaches have been used where MD simulations carried out at the interface region are connected the phase field simulations using finite element methods [126,127] or where (moving region) MD simulations are carried out at the interface region to supply parameters, such as the strain response, kinetic coefficient (velocity vs. under-cooling behavior) and diffusivity, to the phase field model [128]. As the microstructure in many phase-field models is assumed to be coherently uniform, which is, in many cases incorrect, atomistic MD simulations may be of help in supplying anisotropic information about structure and dynamics of interface regions. A very interesting theoretical model for binary crystal nucleation was developed by Gránásy and coworkers [129,130,131,132,133,134,135,136,137,138,139], and incorporated into simulations and applied on Lennard–Jones and ice-water systems [129,130,131], polymeric dendrites [132,134,136], hydrate formation by Gránásy, Kvamme and coworkers [133,135,137,138] and heterogeneous crystal nucleation [139]. Davidchack and Laird [140,141] have used MD simulations with hard spheres to study interfacial planes between crystal and fluid-like domains.

Plapp and Karma have proposed a robust hybrid multi-scale scheme applicable in one, two and three dimensions by combining diffusion Monte Carlo (similar to the one used to solve Schrödinger equation) and phase-field models to perform simulations of dendrictic solidification. The method is applicable for other applications than solidification [142]. General critical discussion of microstructure modelling is given in [143] highlighting phase field and cellular automata methods. More examples of multi-scale methods will be given below in a dedicated chapter.

Random Sequential Adsorption

Random sequential adsorption (RSA) is a family of models for a variety of phenomena in chemistry, physics, biology, ecology, sociology and in many other fields as it is a general and highly interesting mathematical and physical packing problem [144]. It is applicable on all kinds of moving particles from atoms to motor vehicles arriving randomly as a Poisson process selecting a free space based on a distribution of sites or slots. Particles become bound if they find a vacant place (with no overlap with other particles already occupying the actual space), or they leave the system in case of no luck (possibly to try again). The process continues for a certain period of time until the distribution of sites becomes saturated (complete packing) following the random packing scheme, or if no more particles are arriving although there may still be sites available. For excellent reviews see [145,146,147,148,149]. Colloids and Surfaces A, vol. 165 is a special issue containing several articles of RSA. The one-dimensional RSA model is known as the "car parking" problem as it can be seen as an illustrative and familiar analog. Imagine a long street with apartment houses where people and families live. During the day (office hours) there is plenty of space to park but in the evening the parking spots become quickly occupied (adsorption). In the next morning cars one after another are driven away (desorption). The solution of this 1D RSA problem would give us an answer of how many drivers on average (driving similar cars) can park at random positions along a street of a certain length? This can be solved analytically if the cars are equally long and the parking slots are discrete. For continuous space along the street the exact solution was given by a Hungarian mathematician Alfréd Rényi [150]. Higher dimensional problems are complicated and approximate solutions are given



Self-assembled Materials, Figure 11
A phase field model for a growth of a precipitate

based on series expansion, or as a sequential Markovian process solved using Monte Carlo simulations although the detailed balance does not apply for irreversible processes.

The two dimensional RSA is an important and reliable model for molecules adsorbing on a surface reversible or binding irreversibly (with covalent bonds) to specific sites on the surface. Molecules can typically be monomers, dimers (occupying two neighboring sites which are vacant). Polymers can be adsorbed if they find successive empty slots on the surface. Various rules can be imposed on gas molecules approaching already occupied site. The interactions can be both specific and non-specific depending on the studied systems. Orchestration of van der Waals, electrostatic, hydrophilic/phobic interactions as well as hydrogen bonds covalent and ionic bonds leads to adsorption of molecules onto surfaces (consisting of atoms and molecules). Collective sequential adsorption (CSA) is a further generalization of RSA taking into account the local environment around the adsorption site [146]. Even large biomolecular systems (e. g. proteins, DNA, colloids, cells) can be coarse-grained and treated with RSA methods. RSA have been used even for translocation of biopolymers through pores [151] and nucleosomes on a stretched single strand DNA [152]. Generalized RSA methods can used to simulate multilayer buildup in a self-assembled layer-by-layer (LbL) process [153]. Good agreement was obtained between the simulations and experiment for colloids and poly-electrolytes suggesting that the method could be extended to polymers and proteins. Kinetics, jamming limit and structural phase behavior of polydisperse tethered nanoparticles are studied by Gray et al. [154]. Dynamics of self-assembled polyelectrolyte multilayers on glass substrates have been studied by Breit and coworkers [155]. The kinetics could be quantitatively understood using a RSA model for the buildup of a film consisting of polyelectrolyte disks with polydisperse sizes. RSA types of methods in modelling of self-assembled monolayers of charged colloidal particles are reviewed in [156]. Erban and Chapman have presented a diffusion driven random sequential adsorption simulation model operating in real physical time containing features for both reaction kinetics between the molecules and the surface and geometrical constraints of the molecular surface. It is illustrated by an assembly of reactive polymers on a virus surface [157]. Pre-patterning is used as a tool in RSA-based lattice Monte Carlo simulations to improve the self-assembly in nano and micro-scale structure engineering by Cadilhe et al. [158]. RSA model is applied for gelatin self-assembly in binary mixture of water and ethanol [159]. Short-time and long-time kinetics in colloidal adsorption to a monolayer are studied by combining RSA model and Brownian dynamics by Gray and Bonnegaze [160] in the case of self-assembly to meso-structures.

Cellular Automata

The Cellular Automaton (CA) concept was created by John von Neuman and Stanislaw Ulam more than a half century ago to originally solve problems in evolutionary biology [161]. CAs are discrete dynamical systems

where space and time variables and all properties have finite discrete values. CAs start from simple identical (or nearly identical) individual systems interacting locally but evolve to complex systems following simple rules in a synchronous manner. CAs are similar to Petri nets [162] in providing a less mathematical way to differential equations and calculus. Other similar automata are the Turing systems, presented by Alan Turing in 1952 [163] showing how a simple mathematical model can describe spontaneously spreading of reacting chemical specie giving stationary concentration patterns. Turing patterns can also be found as stripes on zebras and spots on cheetahs and elsewhere in nature. Maybe the most widely known application of CAs is the tic-tac-toe like "Game of Life" by John Conway from 1970, a very simple model simulating birth and death of cells in a square lattice interacting according to a set of Boolean conditions. It was easily adapted from paper sheets to computer screen. Dynamic Cellular Automata (DCA) by Wishart and coworkers [164] is a simulation approach allowing a Brownian type of stochastic motion of individual molecules. DCA is more flexible as it allows several cells/molecules to move in a single step. Lattice gas automata are another common use of CA which together with the Boltzmann equation and kinetic theory has led to development of Lattice Boltzmann simulations [165], a powerful method to do fluid dynamics and to solve the Navier–Stokes equations in a from-bottom-up way. The obvious drawback in the current CA methods is that the objects do not have realistic features.

The quantum-dot cellular automata (QCA) approach is proposed as an alternative to molecular electronics. Theoretical behavior of QCA (arrays of Coulomb- coupled quantum-dot cells) is studied in [166] and the state of each QCA cell is determined by its interaction with neighboring cells through the Coulomb interaction and demonstrated experimentally [167] and the field was reviewed [168,169,170]. A single-molecule implementation of a QCA cell is presented combined with *ab initio* electronic structure calculations [171] using simple prototype molecular systems with a molecule in which charge is localized on specific sites and can tunnel between those sites while the role of the dots is played by redox sites, with tunneling paths provided by bridging ligands. Possibilities to coarse-grain cellular automata, emergence, and the predictability of complex systems are studied by Israeli and Goldenfeld [172]. As molecular self-assembly is driven by local, short-range forces and therefore the dynamics is solely based on local interactions and as atomistic simulations of self-assembly become quickly complicated, a cellular automata based simulation, in which data structures, representing different molecular entities such as water and hydrophilic and hydrophobic monomers, share locally propagated force information on a hexagonal, two-dimensional lattice is proposed, with the purpose of this level of description to gain insight about entropy-driven processes in molecular many-particle systems [173].

Nilsson and Rasmussen present a lattice gas technique for simulating molecular self-assembly of amphiphilic polymers in aqueous environments, where water molecules, hydrocarbons tail-groups and amphiphilic head-groups are explicitly represented on a three dimensional discrete lattice [174]. A Cellular Automata Model of Diffusion in Aqueous Systems show that lipophilic solutes diffuse faster than do polar solutes (in agreement with experiments) [175]. Cellular automaton and MC simulations for the case of CO oxidation on a catalytic surface with simultaneous adsorption, reaction, and diffusion, including the Eley–Rideal step in the reaction mechanism is presented [176]. As properties of diffusing species in microporous materials are strongly influenced by the confining framework, providing the energy landscape for the transport process. Because of the simple topology and the cellular nature of the cages of a zeolite Demontis et al. [177] suggest that it is appropriate to apply to the study of the problem of diffusion in tight confinement a time-space discrete model such as a lattice-gas cellular automaton.

Multi-Scale Modeling

Molecules are moving in a world where "nanometer" could be used as a standard length unit and "femtosecond" is a convenient time unit. For us humans "meter" and "second" are the familiar units to which we all can relate our daily activities. In measuring the time from the big bang and distances between galaxies, both meter and second turn out immediately not to be very convenient to use. Obviously phenomena occur at different length and time scales. Connecting time and length scales has rapidly become a very vital area in from-bottom-up strategies. It will certainly still take quite some time for multi-scale modelling of matter to mature. Below we give examples of both concurrent and hierarchical approaches to multiscale modelling, the two main approaches.

"Concurrent" multiscale modelling can be accomplished by making geometric network of grid points describing the system at a detailed level successively coarser while losing more and more details. For example a crystal could be described with atoms or molecules in crystal sites connected together in 3D net. By applying a scheme with successive coarsening of connecting grid points approaching a continuum description can be coupled to the atomistic scheme by finite element methods [178]. A spe-

cific hand-shake region is created between the both scales. This is used to model cracks and fractures both at atomistic and macroscopic scale.

"Hierarchical" multiscale modelling can be performed by successively reducing the degrees of freedom which are not needed in the next less accurate description of the studied system. This makes the modelling faster and allows longer length and time scales to be covered. A powerful method called the Inverse Monte Carlo (IMC) was presented by us [179] where the "Inverse Problem" was solved completely to obtain interaction potentials (force fields) from radial distribution functions (RDF). The IMC constructs effective potentials consistent with the original RDFs. It can be used for example by starting from *ab initio* MD simulations and construct atomistic interaction potentials to be used in classical all-atom simulations. In all-atom simulations IMC can construct coarse-grained potentials for example by grouping together functional groups in large molecular systems and eliminate the solvent by incorporating it into the coarse-grained potentials. These coarse-grained potentials reproduce the RDFs in the all-atom simulations carried out with explicit solvent. This scheme can be used even beyond this kind of meso-scale description. There are other related methods presented [180,181].

Innovative computational schemes for multiscale modelling are reported continuously based on varying strategies. Examples below are given of studies (i) to couple first-principles simulations and other quantum schemes (ii) to use results from atomistic simulations to carry out subsequent coarse-graining (iii) to connect atomistic schemes with continuum models (iv) to apply kinetic Monte Carlo and other kinetic schemes subsequent to particle simulations, as well as (v) other combinations of various methods discussed earlier in this article are listed below, including a beautiful text-book example (vi) of multi-scale modelling of human hair.

(i) Transparent interface is created between classical molecular dynamics and first-principles molecular dynamics combining first-principles Born–Oppenheimer local spin density molecular dynamics (BO-LSD-MD) with classical molecular dynamics in [182]. Effective quantum mechanical classical mechanical (QM/CM) partitioning method for multiscale modeling is proposed in [183]. Quantum mechanics at the core of multi-scale simulations with neglect of diatomic differential overlap theory (used in common semi-empirical schemes such as AM1, PM3 and MNDO) is considered in [184]. Potential parameterization based on extended model systems where the force data is calculated from QM methods on a limited range of applications is shown to be essential for a consistent and ultimately, predictive embedding approach to concurrent multi-scale materials simulation in [185].

(ii) Automatic coarse-graining of polymers by deriving effective potentials for multi-atom units or super-atoms from atomistic simulations is reported in [186]. Hierarchical multi-scale modelling of plasticity of submicron thin metal films allowing modelling of thicker films a discrete dislocation model of "diffusional creep" is presented in [187]. Materials modeling platform bridging the molecular characteristics of polymers with macroscopic properties is given in [188]. Linkage between atomistic and meso-scale coarse-grained simulation based on model where the force acting on the CG particles is divided into the mean force (calculated from constrained MD simulations), friction force and random force is discussed in [189,190]. Combined atomistic and meso-scale simulation of grain growth in nanocrystalline thin films with input materials parameters obtained by MD simulation in [191]. Fuzzy clustering approach to hierarchical MD simulation of multiscale materials phenomena by combining a hierarchy of sub-dynamics: (i) rigid-body cluster dynamics for global conformational changes; (ii) implicit integration of Newton's equations for the coalescence of the clusters; and (iii) normal-mode analysis of fast atomic oscillations is used to facilitate the seamless integration of the multiple levels of abstraction in [192]. A multi-scale simulation of tungsten film delamination from silicon substrate where MD simulations of a single crystal *W* block under tension are used to calibrate a new decohesion model to investigate the effect of specimen size and loading rate on the material properties in [193].

(iii) Coupling of atomistic and continuum models in computational materials science using finite element methods is studied in [194] and finite-element method multi-scale atomistic-continuum modelling of crack propagation in a two-dimensional macroscopic plate by coupling the crack dynamics at the macro scales and nano-scales via an intermediate meso-scale continuum is carried out with molecular dynamics simulation driving the crack tip forward in [195].

(iv) Molecular simulations in zeolitic process design using configurational-biased Monte Carlo simulations combined with transition rate theories kinetic Monte Carlo and molecular dynamics simulations [196].

Multi-scale modeling of hydrogen isotope diffusion in graphite is carried out where molecular dynamics calculations resolve the microscopic length scale and deliver reliable input data for kinetic Monte Carlo calculations (jump frequencies, migration energies, jump step-sizes) from meso-scale up to the macroscopic system length [197]. An MD study of the carbon-catalyst interaction energy for multi-scale modelling of single wall carbon nano-tube growth is presented [198]. A multi-scale approach to radiation-induced segregation at various grain boundaries based on a new rate equation model and MD simulations is reported in [199]. Multi-Scale Computational Framework by integrating a Computational Fluid Dynamics software, a Kinetic Monte Carlo solver and an MD simulator for the self-assembly of atoms into molecular structures is presented in [200]. A novel polycrystalline thin film growth simulator with an atomic level one-dimensional kinetic lattice Monte Carlo model and a real time feature scale two-dimensional facet nucleation and growth model is demonstrated in [201]. A multi-scale atomistic study of the interstitials agglomeration in crystals using a hierarchy of atomistic approaches: the tight-binding molecular dynamics, molecular dynamics based on environment dependent inter-atomic potentials and a lattice kinetic Monte Carlo (LKMC) I reported in [202]. Computer simulations, spanning across different time and length scales, are used to study thin film growth morphology in organic self-assembled monolayers using thiophenes on gold. *Ab initio* calculations created a catalog of the energetics in vacuum and interactions in three orthogonal orientations to a Au (111) surface in [203]. This information was supplied as the input for kMC simulations of dimer and trimer representations of small organic molecules to describe both sub-monolayer and multilayer growth. Finally, MD studies were used to understand the packing structures of stable polymorphs of thiophene SAMs.

(v) Studies of behavior and structure of charged surfaces on different length scales is carried out using a different computational schemes yielding different levels of description of charged surfaces. (a) Coarse grained MC simulations of idealized surfaces incorporate large-length-scale fluctuation and correlation effects in the counter-ion cloud at a charged surface. (b) Brownian dynamics simulations of more realistic and structured surfaces give modified counter-ion distributions, also allowing estimations of mobility of counter-ions at charged surfaces. (c) All-atomistic MD simulations reproduce water structuring effects at surfaces such as hydration and hydrophobic de-wetting. (d) *Ab initio* calculations finally give the effective interactions between oppositely charged groups in vacuum and in solution in [204]. Multiscale simulations using generalized interpolation material point method and Structured Adaptive Mesh Refinement Application Infrastructure parallel processing as a multiple length scale tool from nanometer to millimeter is presented in [205]. To simulate the agglomeration carbonaceous nanoparticle assembly was studied using a multiscale coarse-graining by starting with an atomistic ensemble of 10 000 nanoparticles (or effectively 2 million total carbon atoms). The coarse-graining was accomplished applying a force-matching procedure. The results show rich and varied clustering behaviors for different particle morphologies [206]. A multiscale coarse-graining to derive coarse-grained models is applied to C60 and to carbonaceous nano-particles produced in combustion environments. The coarse-graining of the inter-particle force field is accomplished applying a force-matching procedure from all-atom MD simulations reproduce accurately the structural properties of the nano-particle systems [207].

(vi) A very impressive multiscale modelling approach to the mechanics of human hair fibres is given by Akkermans in [208] providing a very concrete example of multiscale modelling work-flows at their best where mesoscale models are constructed from atomistic simulations and where meso-scale simulation methods are used as input to finite-element calculations.

There are several excellent reviews discussing a great variety of invented multi-scale modelling schemes in materials science listed below and well worth of consulting. Simulation methods including broad areas of quantum mechanics, molecular dynamics and multiple-scale approaches, based on coupling the atomistic and continuum models are discussed in [209] and continuum/quasi-continuum approaches, the kinetic Monte Carlo technique and accelerated molecular dynamics simulation are gathered in [210]. Recent advances in bridging scale between quantum mechanical and continuum coupling are briefly described in [211] and multiscale hybrid simulation methods for material systems based on tight-binding DFT, MD and continuum models for deformation and diffusion on surfaces are reviewed in [212]. A seamless coupling of quantum to statistical to continuum mechanics involving models for unifying finite elements, molecular dynamics and semi-empirical tight-binding and coarse-grained Molec-

ular Dynamics as an effective model in solid state is explained in [213]. Current trends from atomistic simulation towards multiscale modelling of materials based on quantum mechanical, especially density functional theory for electronic properties linked to atomic/molecular dynamics and kinetic Monte Carlo simulations where coarse graining leads to lattice gas and cellular automata, and also to continuum equations solved by finite-element and finite-difference techniques are discussed in [214]. Vvedensky presents a comprehensive outlook and an excellent review of current multiscale modelling methods for nanostructures across a vast range of length and time scales clearly stating that a complete understanding of the behavior of materials thereby requires theoretical and computational tools that span from the atomic-scale detail of first-principles methods to the more coarse-grained description provided by continuum equations with the ultimate aim to systematically couple the scales from the atomistic to the continuum level in [215].

Multi-scale mechanics of nano-composites including an interface to better understand the phenomenological changes across multiple length and time scales are reviewed in [216] and current multi-scale modeling and simulation of nano-structured materials in [217]. Modelling the nano-scale phenomena in condensed matter physics via computer-based numerical simulations focusing on the adhesive and indentation properties of the solid surfaces in nano-contacts, the nucleation and growth of nano-phase metallic and semi-conducting atomic and molecular films on supporting substrates, and the nano- and multi-scale crack propagation properties of metallic lattices are discussed in [218]. Heterogeneous multiscale methods, applicable on complex fluids, micro-fluidics, solids, interface problems, stochastic problems, and statistically self-similar problems are reviewed in [219] and the emerging role of multiscale modeling in nano- and micro-mechanics of materials in [220].

Hierarchical modeling of amorphous polymers where a broad spectra of length and time scales governing the behavior of these materials is based on the use of connectivity-altering Monte Carlo algorithms for rapid equilibration of atomistic models of long-chain polymer systems, calculation of their conformational, packing and volumetric properties, and assessment of their entanglement structure, and self-consistent field calculations of morphology development in complex systems containing block copolymers, coupled with rubber elasticity theory for the prediction of the stress-strain behavior of these systems are discussed in [221]. Linking various length scales via materials informatics is reviewed in [222]. Recent progress in the simulations of liquid crystals across a range of length and time scales is reviewed in [223] with three material properties of liquid crystals (the archetypal self-assembled materials in Nature) are discussed in detail: elastic constants, rotational viscosity and helical twisting powers. Computer simulations of surfactant solutions are reviewed by Shelley & Shelley discussing the importance of connecting the inherently disparate length scales treating the systems using a single coherent multiscale simulation [224]. In an excellent review dealing with self-assembly from nano-scale to micro-scale colloids, issues like the effect of the shape and composition on assembly and the role of these factors in the self-organization of particles into ordered assemblies, scale dependent effects on assembly related to interparticle forces, vitrification and gelation and also building block design rules from computer simulations using anisotropic interactions are discussed in [225]. Karakasisis and Charitidis give an excellent review of methods in multiscale computational materials science how to connect electronic structure calculations with classical atomistic simulation using molecular dynamics or Monte Carlo methods at the nano/micro scale and further with kinetic Monte Carlo for larger system/time scales and finite elements for very large scales, presenting both hierarchical and hybrid strategies in [226]. A multiscale modelling scheme is reviewed, starting the initial nucleation processes using MD simulation and inputing the data into a cellular automaton (CA)-based model of the microstructure formation at the micro-scale, as well as to the macroscopic heat flow equation in [227]. Thermodynamics of self-assembly of surfactants in solution through simulations is now being expanded to include phenomena in the fluid dynamic regime a smooth link from MD to mesoscopic and macroscopic length and time scales. Recent trends in this area along with new results based on classical approaches are reviewed in [228].

Future Directions

Self-assembly is a process of pivotal importance that should be very attractive to study using modelling and simulations. The intermolecular interactions that drives the process can be directly introduced into common forms of the interaction potentials and force-fields. Many self-assembly processes can indeed also be rationalized and understood both molecularly and macroscopically by means of thermodynamics.

Self-assembly is typically a relatively slow process that proceeds through several metastable states. The extended time-scale and structure evolution that usually extends of several length scales requires a combination of approaches. The rapid expansion of multi scale modeling approaches

show much promise but it is still a long way to go before the fine details of self assembly processes with durations of hours up to weeks can be captured accurately. In principle the kinetic Monte Carlo method could cover arbitrary long time scales but it requires detailed knowledge about the system, its states and barriers between them.

Another great challenge in modern computational molecular science is the transition from a descriptive to a predictive level of molecular modeling. Despite several examples of successful predictions of molecular properties *in silico*, current molecular modeling remains largely on a descriptive stage: simulations are used to describe already known experimental phenomena, in order to give a better understanding, for testing the theories describing experiment, or for correct interpretation of the experimental data. Ability of simulations to predict, to discover something new is still rather limited. While some properties are predicted, others are not, successful investigations are published while failures are not, etc., so it is still difficult to rely on simulation results and experimental verification of simulation is therefore always needed. One of the reasons to a limited predictability of molecular models is that the currently used force fields were developed originally at the early stages of molecular modeling in the 80s with rather limited computing resources at hands and later only refined in a few occasions. In the case of meso-scale modeling on a coarse-grained level, even a functional form of the interaction potential is not known and therefore many *ad hoc* models are used. In order to substantially improve reliability and predictability of molecular simulations, a new generation of force fields would be necessary, which have to be derived in a more consecutive manner with more emphasis on the fundamental *ab-initio* approaches. To replace experiment with a computer experiment surely lies far in the future. However, if modelling can point the right direction to an experimentalist, much guesswork, effort and resources can be saved.

Acknowledgments

Both authors are connected to the Berzelii centre EXSELENT on porous materials at Stockholm University (http://www.exselent.su.se/).

Bibliography

1. Ozin GA, Arsenault AC (2005) Nanochemistry, A Chemical Approach to Nanomaterials. RSC Publishing, London
2. Whitesides GM, Grzybowski B (2002) Science 295:2418
3. Lehn J-M (1988) Angew Chem Int Ed 27:89
4. Nuzzo RG, Fusco FA, Allara DL (1987) J Am Chem Soc 109:2358
5. Bentzon MD, van Wonterghem J, Moerup S, Tholen A, Koch CJW (1989) Phil Mag B 60:169
6. Steigerwald ML, Alivisatos AP, Gibson JM, Harris TD, Kortan R, Muller AJ, Thayer AM, Duncan TM, Douglass DC, Brus LE (1988) J Am Chem Soc 110:3046
7. Redl FX, Cho K-S, Murray CB, O'Brien S (2003) Nature 423:968
8. Shevchenko EV, Talapin DV, Kotov NA, O'Brien S, Murray CB (2006) Nature 439:55
9. Decher G (1997) Science 277:1232
10. Whitesides GM, Mathias JP, Seto CT (1991) Science 254:1312
11. Fendler JH (1996) Chem Matter 8:1616
12. Lawrence DS, Jiang T, Levett M (1996) Chem Rev 95:2229
13. Zhang S (2003) Nature Biotech 21:1171
14. Tirrell M (2005) AIChE J 51:2386
15. Zhao X, Zhang S (2006) Chem Soc Rev 35:1105
16. Dubois LH, Nuzzo RG (1992) Ann Rev Phys Chem 43:437
17. Ulman A (1996) Chem Rev 96:1533
18. Laibinis PE, Whitesides GM, Allara DL, Tao YT, Parikh AN, Ruzzo RG (1991) J Am Chem Soc 113:7152
19. Tanev PT, Pinnavaia TJ (1995) Science 267:865
20. Zeng FW, Zimmerman SC (1997) Chem Rev 97:1681
21. Lyubartsev AP, Laaksonen A (2003) In: Karttunen M, Vattulainen I, Lukkarinen A (eds) Novel methods in soft matter simulations. Lecture Notes in Physics, vol 640. Springer, Berlin
22. Laaksonen A, Tu Y (1999) In: Balbuena PB, Seminario JM (eds) Molecular Dynamics: from classical to quantum methods. Elsevier, Amsterdam
23. Roothaan CCJ (1951) Rev Mod Phys 23:69
24. Hehre WJ, Radom L, Schleyer PVR, Pople JA (1986) Ab initio molecular orbital theory. Wiley, New York
25. Leach A (2001) Molecular Modelling: Principles and Applications, 2nd edn. Pearson Education, New Jersey
26. Pople JA, Santry DP, Segal GA (1965) J Chem Phys 43:S129
27. Pople JA, Segal GA (1965) J Chem Phys 43:S136
28. Pople JA, Beveridge DL, Dobosh PA (1967) J Chem Phys 47:2026
29. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) J Am Chem Soc 107:3902
30. Stewart JJP (1989) J Comput Chem 10:209
31. Stewart JJP (1989) J Comput Chem 10:221
32. Dewar MJS, Thiel W (1977) J Am Chem Soc 99:4899
33. Kolb M, Thiel W (1993) J Comput Chem 14:775
34. Weber W, Thiel W (2000) Theor Chem Acc 103:495
35. Tu Y, Jacobsson SP, Laaksonen A (2003) Mol Phys 101:3009
36. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785
37. Burke K, Perdew JP, Wang Y (1998) In: Dobson JF, Vignale G, Das MP (eds) Electronic Density Functional Theory: Recent Progress and New Directions. Plenum Press, New York
38. Becke AD (1988) Phys Rev A 38:3098
39. Hohenberg P, Kohn W (1964) Phys Rev B 136:864
40. Kohn W, Sham LJ (1965) Phys Rev A 140:1133
41. Car RR, Parrinello M (1985) Phys Rev Lett 55:2471
42. Foulkes WMC, Haydock R (1989) Phys Rev B 39:12520
43. Harris J (1985) Phys Rev B 31:1770
44. Sankey OF, Niklewski DJ (1989) Phys Rev B 40:3979
45. Tu Y, Laaksonen A (2005) In: Chipot C, Elber R, Laaksonen A, Leimkuhler B, Mark A, Schlick T, Schuette C, Skeel R (eds) Advances in Algorithms for Macromolecular Simulation. Lecture Notes in Computational Science & Engineering, vol 49. Springer, Heidelberg
46. Tu Y, Jacobsson SP, Laaksonen A (2006) Phys Rev B 74:205104

47. Madden PA, Heaton R, Aguado A, Jahn S (2006) J Mol Struct – Theochem 771:9
48. Hafner J (2007) Comp Phys Commun 177:6
49. Muller S (2006) Surf Interf Anal 38:1158
50. Juhas P, Cherba DM, Duxbury PM, Punch WF, Billinge SJL (2006) Nature 440:655
51. Le Page Y, Rodgers JR (2005) Acta Appl Cryst 38:697
52. Bockstedte M, Kley A, Neugebauer J, Scheffler M (1997) Comp Phys Commun 107:187
53. Gonze X, Beuken JM, Caracas R, Detraux F, Fuchs M, Rignanese GM, Sindic L, Verstraete M, Zerah G, Jollet F, Torrent M, Roy A, Mikami M, Ghosez P, Paty JY, Allan DC (2002) Comp Mater Sci 25:478
54. Colombo L (2005) Riv Nuovo Cim 28:1
55. Bowler DR, Choudhury R, Gillan MJ, Miyazaki T (2006) Phys Stat Solidi B 243:989
56. Kawazoe Y (2003) Bull Mater Sci 26:13
57. Allen MP, Tildesley D (1987) Computer Simulation of Liquids. Clarendon, Oxford
58. Rapaport DC (2004) The Art of Molecular Dynamics Simulation, 2nd edn. Cambridge University Press, Cambridge
59. Frenkel D, Smit B (2002) Understanding Molecular Simulation: from Algorithms to Applications, 2nd edn. Computational Science Series, vol 1. Academic Press, London
60. Schlick T (2002) Molecular Modeling: An Interdisciplinary Guide. Springer, New York
61. Hinchliffe A (2003) Molecular Modelling for Beginners. Wiley, New York
62. Field MF (2007) A Practical Introduction to the Simulation of Molecular Systems. Cambridge University Press, Cambridge
63. Seminario JM, Balbuena P (eds) (1999) Molecular Dynamics (Theoretical and Computational Chemistry). Elsevier Science, Amsterdam
64. Haile JM (1992) Molecular Dynamics Simulation: Elementary Methods. Wiley, New York
65. Leimkuhler B, Reich S (2004) Simulating Hamiltonian Dynamics. Cambridge University Press, Cambridge
66. Alder BJ, Wainwright TE (1957) J Chem Phys 27:1208
67. Alder BJ, Wainwright TE (1959) J Chem Phys 31:459
68. Rahman A (1964) Phys Rev 136:405
69. Verlet L (1967) Phys Rev 159:98
70. Stillinger FH, Rahman A (1974) J Chem Phys 60:1545
71. Hedman F, Laaksonen A (1999) In: Balbuena PB, Seminario JM (eds) Molecular Dynamics: from classical to quantum methods. Theoretical Chemistry Series, vol 7. Elsevier Science BV, Amsterdam
72. Hedmann F (2006) http://www.diva-portal.org/diva/getDocument?urn_nbn_se_su_diva-1008-4__fulltext.pdf
73. Sprik M, Demarchee E, Michel B, Rothlisberger U, Klein ML, Wolf H, Ringsdorf H (1994) Langmuir 10:4116
74. Yamamoto H, Watanabe T, Nishiyama K, Tatsumura K, Ohdomari I (2006) J Phys IV 132:189
75. Mamatkulov SI, Khabibullaev PK, Netz TR (2004) Langmuir 20:4756
76. Yu KQ, Li ZS, Sun JZ (2002) Langmuir 18:1419
77. Hackett E, Manias E, Giannelis EP (1998) J Chem Phys 108:7410
78. Jang SS, Jang YH, Kim Y-H, Goddard III WA, Flood AH, Laursen BW, Tseng H-R, Stoddart JF, Jeppesen JO, Choi JW, Steuerman DW, Delonno E, Heath JR (2005) J Am Chem Soc 127:1563
79. Striolo A, NcCabe C, Cummings PT (2006) J Chem Phys 125:104904
80. Luo M, Mazyar OA, Zhu Q, Vaughn MW, Hase WL, Dai LL (2006) Langmuir 22:6385
81. Wu J (2006) AIChE J 52:1169
82. Jallipalli A, Balakrishnan G, Huang SH, Khoshakhlagh A, Dawson LR, Huffaker DL (2007) J Cryst Growth 303:449
83. van Workum K, Douglas JF (2006) Phys Rev E 73:031502
84. Ariga K, Hill JP, Ji Q (2007) Phys Chem Chem Phys 9:2319
85. Cooper N (1987) Stanislaw Ulam 1909–1984. Los Alamos Sci, vol 15. Special Issue. See: http://www.fas.org/sgp/othergov/doe/lanl/pubs/number15.htm
86. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AN, Teller E (1953) J Chem Phys 21:1087
87. Jayaram B (1996) J Math Chem 20:395
88. Ni R, Cao D, Wang WC (2007) J Phys Chem C 111:11802
89. Pool R, Schapotschnikow P, Vlugt TJH (2007) J Phys Chem 111:10201
90. Sinyagin AY, Belov A, Tang Z, Kotov NA (2006) J Phys Chem B 110:7500
91. Reimer U, Zehl T, Wahab M, Schiller P, Moegel H-J (2006) Coll Surf A Phys Chem Eng Asp 290:25
92. Reimer U, Wahab M, Schiller P, Moegel HJ (2006) Langmuir 21:1640
93. Patti A, Mackie AD, Siperstein FR (2007) Langmuir 23:6771
94. He XH, Song M, Liang HJ, Pan CY (2001) J Chem Phys 114:10510
95. Wijmans C, Linse P (1995) Langmuir 11:3748
96. Wijmans C, Linse P (1996) J Phys Chem 100:12583
97. Wijmans C, Linse P (1997) J Chem Phys 106:328
98. Linse P (2000) In: Alexandridis P, Lindman B (eds) Amphiphilic Block Copolymers: Self-Assembly and Application. Elsevier Science, Amsterdam
99. Abraham FF, White GM (1970) J Appl Phys 41:1841
100. Gilmer GH, Bennema P (1972) J Appl Phys 43:1347
101. Bortz AB, Kalos MH, Lebowitz JL (1975) J Comput Phys 17:10
102. Gillespie DT (1977) J Phys Chem 81:2340
103. Chatterjee A, Vlachos DG (2007) J Computer-Aided Matter Des 14:253
104. Metiu H, Lu Y-T, Zhang Z (1992) Science 255:1088
105. Much F, Ahr M, Biehl M, Kinzel W (2001) Europhys Lett 56:791
106. Larsson MI (2004) Surf Sci 551:69
107. Kozak JJ, Nicolis C, Nicolis G (2007) J Chem Phys 126:154701
108. Zhu R, Pan E, Chung PW (2007) Phys Rev B 75:205339
109. JS Rowlinson Notes (2003) Rec R Soc Lond 57:143
110. Wheeler AA, Boettinger WJ, McFadden GB (1992) Phys Rev A 45:7424
111. Anderson DM, McFadden GB, Wheeler AA (1998) Ann Rev Fluid Mech 30:139
112. Lu W, Suo Z (2001) J Mech Phys Solids 49:1937
113. Chen LQ (2002) Annu, Rev Mater Res 32:113
114. Boettinger WJ, Warren JA, Beckermann C, Karma A (2002) Ann Rev Mater Res 32:163
115. Thornton K, Ågren J, Voorhees PW (2003) Acta Mater 51:5675
116. Beckermann C, Diepers H-J, Steinbach I, Karma A, Tong X (1999) J Comput Phys 154:468
117. Boettinger WJ, Coriell SR, Greer AL, Karma A, Kurz W, Rappaz M, Trivedi R (2000) Acta Mater 48:43
118. Cahn JW, Hillard JE (1957) J Chem Phys 28:258
119. Cahn JW, Hillard JE (1959) J Chem Phys 31:688
120. Cahn JW (1961) Acta Metall 9:795

121. Allen SM, Cahn JW (1979) Acta Metall Mater 27:1085
122. Chen LQ, Shen J (1998) Comput Phys Commun 108:147
123. Elder KR, Grant M (2004) Phys Rev E 70:051605
124. Lu W, Salac D (2005) Phys Rev Lett 94:146103
125. Wang YU (2007) Acta Mater 55:3835
126. Inoue T (2004) J Phys IV 120:3
127. Inoue T (2004) Mater Sci Res Intl 10:1
128. Hoyt JJ, Sadigh B, Asta M, Foiles SM (1999) Acta Mater 47:3181
129. Gránásy L, Börsönyi T, Pusztai T (2002) Phys Rev Lett 88: 206105
130. Gránásy L, Börsönyi T, Pusztai T (2002) J Cryst Growth 237–239:1813
131. Gránásy L, Pusztai T, James PF (2002) J Chem Phys 117:6157
132. Gránásy L, Pusztai T, Warren JA, Douglas JF, Börsönyi T, Ferrero V (2003) Nat Mater 2:92
133. Kvamme B, Graue A, Aspenes E, Kuznetsova T, Gránásy L, Tóth G, Pusztai T, Tegze G (2004) Phys Chem Chem Phys 6:2327
134. Gránásy L, Pusztai T, Börsönyi T, Warren JA, Douglas JF (2004) Nat Mater 3:645
135. Svandal A, Kvamme B, Gránásy L, Pusztai T (2005) J Phase Equil Diff 26:534
136. Pusztai T, Bortel G, Gránásy L (2005) Mater Sci Eng A 413–414:412
137. Svandal A, Kvamme B, Gránásy L, Pusztai T, Buanes T, Hove J (2006) J Cryst Growth 486:1813
138. Tegze G, Pusztai T, Tóth G, Gránásy L, Svandal A, Buanes T, Kuznetsova T, Kvamme B (2006) J Chem Phys 124:234710
139. Gránásy L, Pusztai T, Saylor D, Warren JA (2007) Phys Rev Lett 98:035703
140. Davidchack RL, Laird BB (1998) J Chem Phys 108:9452
141. Davidchack RL, Laird BB (2000) Phys Rev Lett 85:4751
142. Plapp M, Karma A (2000) J Comput Phys 165:592
143. Raabe D (2002) Adv Mater Progr Rept 14:639
144. Penrose MD (2001) Commun Math Phys 218:153
145. Bartelt MC, Privman V (1991) Intern J Mod Phys B5:2883
146. Evans JW (1993) Rev Modern Phys 65:1281
147. Privman V (2000) Colloids Surf A 165:231
148. Talbot J, Tarjus G, van Tassel PR, Viot P (2000) Colloids Surf A 165:287
149. Senger B, Voegel J-C, Schaaf P (2000) Colloids Surf A 165:255
150. Rényi A (1963) Trans Math Stat Prob 4:205
151. D'Orsogna MR, Chou T, Antal T (2007) J Phys A Math Theor 40:5575
152. Ranjith P, Marko JF (2006) Phys Rev E 74:041602
153. Adamczyk Z, Weroński P, Barbasz J, Kolańsiska M (2007) Appl Surf Sci 253:5776
154. Gray JJ, Klein DH, Korgel BA, Bonnecaze RT (2001) Langmuir 17:2317
155. Breit M, Gao M, von Plessen G, Lemmer U, Feldmann J, Cundiff ST (2002) J Chem Phys 117:3956
156. Yuan Y, Oberholzer MR, Lenhoff AM (2000) Colloids Surf A165:125
157. Erban R, Chapman SJ (2007) Phys Rev E 75:041116
158. Cadilhe A, Aráujo N, Privman V (2007) J Phys Condens Matter 19:065124
159. Bohidar HB, Mohanty B (2004) Phys Rev E 69:021902
160. Gray JJ, Bonnecaze RT (2001) J Chem Phys 114:1366
161. Wolfram S (2002) A new kind of science. Wolfram Media, Champaign
162. Petri CA (1962) Kommunikation mit Automaten. Ph D Thesis, University of Bonn
163. Turing AM (1952) Phil Trans R Soc Lond B237:37
164. Wishart DS, Yang R, Arndt D, Tang P, Cruz J (2005) In Silico Biol 5:139
165. Succi S (2001) The Lattice Boltzmann Equation for Fluid Dynamics and Beyond. Oxford University Press, Oxford
166. Tougaw PD, Lenta CS (1996) J Appl Phys 80:4722
167. Amlani I, Orlov AO, Snider GL, Lent CS, Bernstein GH (1998) Appl Phys Lett 72:2179
168. Orlov AO, Amlani I, Bernstein GH, Lent CS, Snider GL (1997) Science 277:928
169. Snider GL, Orlov AO, Amlani I, Zuo X, Bernstein GH, Lent CS, Merz JL, Porod W (1999) J Appl Phys 85:4283
170. Cole T, Lusth JC (2001) Progr Quant Electr 25:165
171. Lent CS, Isaksen B, Lieberman M (2003) J Am Chem Soc 125:1056
172. Israeli N, Goldenfeld N (2006) Phys Rev E 73:026203
173. Mayer B, Köhler G, Rasmussen S (1997) Phys Rev E 55:4489
174. Nilsson M, Rasmussen S (2003) Discret Math Theor Comput Sci AB(DMCS):31
175. Kier LB, Cheng C-K, Testa B, Carrupt P-A (1997) J Pharmac Sci 86:774
176. Sreekumar P, Jayaraman VK, Kulkarni BD (1998) Ind Eng Chem Res 37:2188
177. Demontis P, Pazzona FG, Suffritti GB (2007) J Chem Phys 126:194709
178. Gumbsch P (1995) J Mater Res 10:2897
179. Lyubartsev AP, Laaksonen A (1995) Phys Rev E 52:3730
180. Muller-Plathe F (2002) Chem Phys Chem 3:754
181. Izvekov S, Parrinello M, Burnham CJ, Voth GA (2004) J Chem Phys 120:10896
182. Du MH, Cheng HP (2003) Int J Quant Chem 93:1
183. Mallik A, Taylor DCE, Runge K, Dufty JW (2004) Int J Quant Chem 100:1019
184. Bartlett RJ, McClellan J, Greer JC, Monaghan S (2006) J Comp Aided Mater Des 13:89
185. Zhu WM, Runge K, Trickey SB (2006) J Comp Aided Mater Des 13:75
186. Faller R (2004) Polymer 45:3869
187. Buehler MJ, Hartmaier A, Gao H (2004) Model Sim Mater Sci 12:S391
188. Doi M (2001) J Comp Appl Math 149:13
189. Kinjo T, Hyodo S (2007) Mol Simul 33:417
190. Hyodo S (2007) Mol Sim 33:279
191. Haslam AJ, Moldovan D, Phillpot SR, Wolf D, Gleiter H (2002) Comp Mater Sci 23:15
192. Nakano A (1997) Comp Phys Commun 105:139
193. Shen LM, Chen Z (2005) Int J Solids Struct 42:5036
194. Curtin WA, Miller RE (2003) Model Simul Mater Sci Eng 11:R33
195. Rafii-Tabar H, Hua L, Cross M (1998) J Phys Cond Matter 10:2375
196. Smit B, Krishna R (2003) Chem Eng Sci 58:557
197. Warrier M, Schneider R, Salonen E, Nordlund K (2004) Contr Plasma Phys 44:307
198. Shibuta Y, Elliott JA (2006) Chem Phys Lett 427:365
199. Sakaguchi N, Watanabe S, Takahashi H, Faulkner RG (2004) J Nucl Mater 329:1166
200. Vasenkov AV, Fedoseyev AI, Kolobov VI, Choi HS, Hong KH, Kim K, Kim J, Lee HS, Shin JK (2006) J Comp Theor Nanosci 3:453

201. Zhang J, Adams JB (2004) Comp Mater Sci 31:317
202. La Magna A, Coffa S, Libertino S, Brambilla L, Alippi P, Colombo L (2001) Nucl Instr Meth Phys Res B 178:154
203. Haran M, Goose JE, Clote NP, Clancy P (2007) Langmuir 23:4897
204. Netz RR (2004) J Phys Cond Mat 16:S2353
205. Ma J, Lu H, Wang B, Roy S, Hornung R, Wissink A, Komanduri R (2005) Comp Model Eng Sci 8:135
206. Izvekov S, Violi A (2006) J Chem Theory Comput 2:504
207. Izvekov S, Violi A, Voth GA (2005) J Phys Chem B Lett 109:17019
208. Akkermans RLC, Warren PB (2004) Phil Trans Royal Soc A 362:1783
209. Liu WK, Karpov EG, Zhang S, Park HS (2004) Comp Meth Appl Mech Eng 193:1529
210. Starrost F, Carter EA (2002) Surf Sci 500:323
211. Lium WK, Park HS, Qian D, Karpov EG, Kadowaki H, Wagner GJ (2006) Comp Meth Appl Mech Eng 195:1407
212. Csanyi G, Albaret T, Moras G, Payne MC, De Vita A (2005) J Phys Cond Matter 17:R691
213. Rudd E, Broughton JQ (2000) Physica, Status Solidi B – Basic Res 217:251
214. Nieminen RM (2002) J Phys Cond Matter 14:2859
215. Vvedensky DD (2004) J Phys Condens Matter 16:R1537
216. Buryachenko VA, Roy A, Lafdi K, Anderson KL, Chellapilla S (2005) Compos Sci Tech 65:2435
217. Gates TS, Odegard GM, Frankland SJV, Glancy TC (2005) Compos Sci Tech 65:2416
218. Rafii-Tabar H (2000) Phys Rept Rev Sect Phys Lett 325:240
219. Weinan E, Engquist B, Li X, Ren WQ, vanden-Eijnden E (2007) Commun Comp Phys 2:367
220. Ghoniem NM, Cho K (2002) Comp Model Eng Sci 3:147
221. Theodorou N (2005) Comp Phys Commun 169:82
222. Liu ZK, Chen LQ, Rajan K (2006) JOM 58:42
223. Wilson M (2005) Intl Rev Phys Chem 24:421
224. Shelley JC, Shelley MY (2000) Curr Opin Coll Interf Sci 5:101
225. Glotzer SC, Solomon MJ, Kotov NA (2004) AIChE J 50:2978
226. Karakasidis TE, Charitidis CA (2007) Mater Sci Eng C 27:1082
227. Rafii-Tabar H, Chirazi A (2002) Phys Rep 365:145
228. Rajagopalan R (2001) Curr Opin Coll Interf Sci 6:357

Self-Organization and the City

Juval Portugali
ESLab (Environmental Simulation Lab), Department of Geography and the Human Environment, Tel Aviv University, Tel Aviv, Israel

Article Outline

Glossary
Definition of the Subject
Introduction
Complexity Theories of Cities – an Overview
Self-Organization and the City
Future Directions
Bibliography

Glossary

Self organization A property of open and complex systems that achieve their order spontaneously, that is, by means of "self-organization".

City A form of settlement that first emerged in the Near East (the core being Mesopotamia) some 5500 years ago. Since its first appearance in Mesopotamia it has diffused in space and time. With colonialism the Western–European form of city has diffused to the entire world, suppressing on the way other forms of cities (e. g. in South America, East Asia, etc.). In the last few decades the city has become the most dominant form of settlement: for the first time in human history more than half of world population lives in cities.

Urbanism The term refers to the totality of life in cities: the interrelations between the social structure, culture, economy, politics, architecture, physical morphology, … associated with life in cities. The first appearance of cities is thus termed *the urban revolution*. Most students of urbanism would agree that 21st century society is undergoing a major urban transformation; some describe it as a new urban revolution.

Planning Planning is, on the one hand, a basic cognitive capability of humans while on the other, a profession and research domain termed interchangeably *city planning*, *urban and regional planning* or *environmental planning*. City planners work and act in the context of planning law and administration the aim of which is to regulate and control life in cities.

SIRN (synergetic inter-representation network) An approach to cognitive mapping and urban dynamics suggesting that cities emerge, maintain their order and change again as a consequence of an on-going interaction between cognitive maps that are constructed in the mind/brain of humans as internal representations and the city as a collective external representation. This on-going interaction gives rise to a network some of whose elements are in the mind/brain, while others in the world.

Information compression, inflation and adaptation A view suggesting that Shannon's notion of information is a property of closed systems and that in complex, self-organizing systems one has to take into consideration the role of *semantic information*. Due to semantic information, the process of self-organization often entails *information compression*; in some cases it entails *information inflation*. The suggestion is that information compression and inflation are two facets of the process of *information adaptation*.

Definition of the Subject

Self-organization is a central property of open and complex systems. While the concept had already appeared in the 1940s, its modern use was pioneered in the 1960s by people such as Haken [50,51] with his theory of *synergetics*, Prigogine with his notion of *dissipative structures* [87,109,110] and others (see review in Chap. 3.1 [95]). Such systems are typically in "a far from equilibrium condition" and exhibit phenomena of chaos, fractal structure and the like. For a long time, the term "self-organization" was used also as an umbrella name for these theories; nowadays it is common to refer to these theories as *complexity theories*.

The notions of self-organization and complexity originated in the sciences, specifically in physics, as a property of natural systems. However, as we shall see below, from the start they were associated with the city – at the beginning the city was used as a metaphor to convey the notion of "self-organization" [87] and at a later stage it was studied as a genuine self-organization system in its own sake [5].

Most theories and methodologies of complexity developed in the last three decades have been applied also to the study of cities with the result that we now have a rich body of research on *fractal cities* [20], *self-organization and the city* [95], *cities and complexity* [16], cellular automata and agent base urban simulation models [22], studies on cities from the perspective of Bak's self-organized criticality [20], studies on cities as networks [16] and much more. This growing body while enriching our understanding of cities and providing sophisticated tools to city planning, also exposes problems that will become the challenges for the next generation of studies on complexity, self-organization and the city.

Introduction

The title "Self-Organization and the City" enfolds *two notions* – 'Self-organization' and 'City', *a thesis* suggesting that cities are complex self-organizing systems, and an *inconsistency* – the view of cities as complex systems that achieve their order spontaneously contradicts the traditional view of cities as symbols of organized order and planning. From the title thus follow four questions: what is self-organization? What is a city? In what sense are cities self-organizing systems? What is the meaning of planning in a self-organizing system? In an *Encyclopedia of Complexity and System Sciences* such as this, there is no need to introduce the term self-organization beyond what has been said about it above; we are thus left with three introductory tasks, namely, to clarify the notion "City", elaborate the thesis that cities are self-organizing systems, and, solve the contradiction between self-organization and planning in the realm of cities.

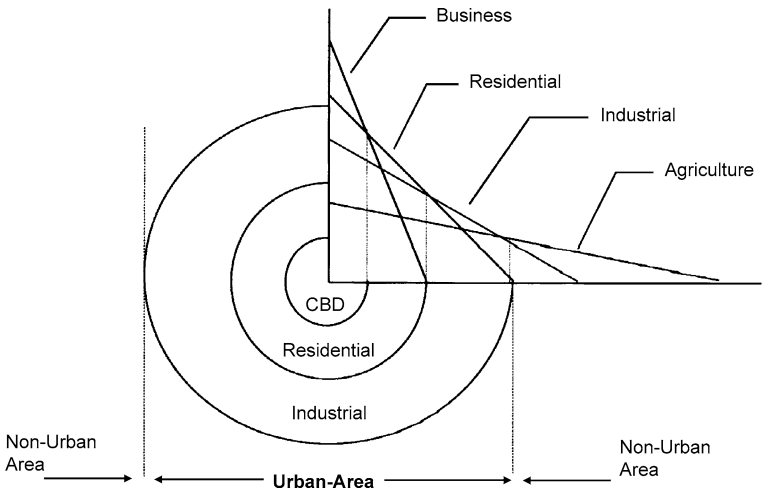
What is a City?

There are two ways to answer this question: first, by looking at explicit attempts to define a city. Second, by exposing the way the different theories of cities explicitly describe, or implicitly perceive, a city.

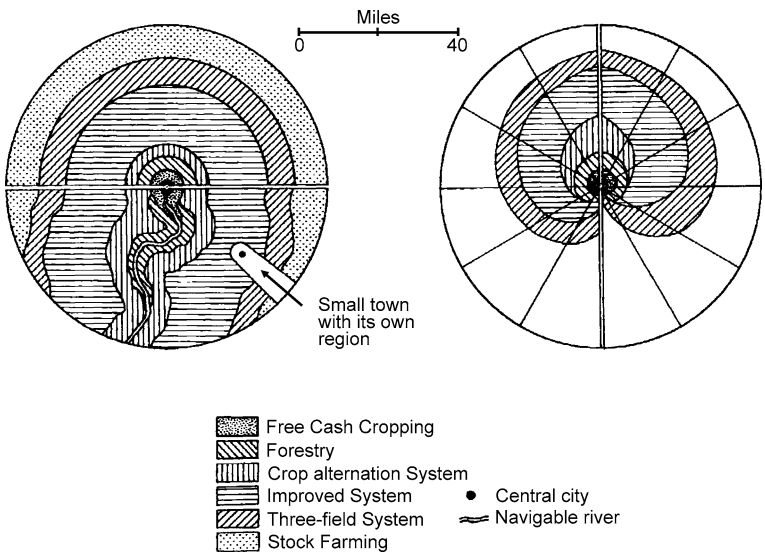
Explicit Attempts to Define a City The history of the many attempts to define "a city" is rather confusing: Whenever a definition was proposed, it was always possible to falsify it (in Popper's [90] sense) by putting forward cities that do not comply with the definition. The main reason for the failure to define cities is that the various attempts to do so were always made with reference to what in cognitive science [115] is called *classical categories*. That is, groups composed of entities sharing some necessary and sufficient conditions that define them as a category and distinguish them from other categories. Students of urbanism have implicitly treated cities as classical categories, and yet, cities are not classical categories – they form a category due to what Wittgenstein [136] has termed *family resemblance*. As a consequence, attempts to define them in terms of a classical category ended up with a failure [95].

A family resemblance category becomes a category not when its elements share some common denominators, but when they form a *network* of partial links and similarities. Further research and experiments have found that many family resemblance categories have a *core-periphery* structure, in the sense that some instances of the category are more *prototypical* of the category than others and they thus form its center while the rest of the instances form the category's periphery [74,78,115].

The city is a good example of a family resemblance category with a core periphery structure. On the one hand, there are no common elements between the "first" cities of some 5,500 years ago and the cities of today except for the name. On the other hand, the first cities had space-time links and similarities with subsequent cities, which in turn had common elements with subsequent cities, and so on until the global cities of today. The result of this process is that cities form a huge space-time family resemblance network extending in time and space from the ancient cities of 5,500 years ago to the cities of today. In this network, one can identify space-time moments during which certain cities became more characteristic or proto-



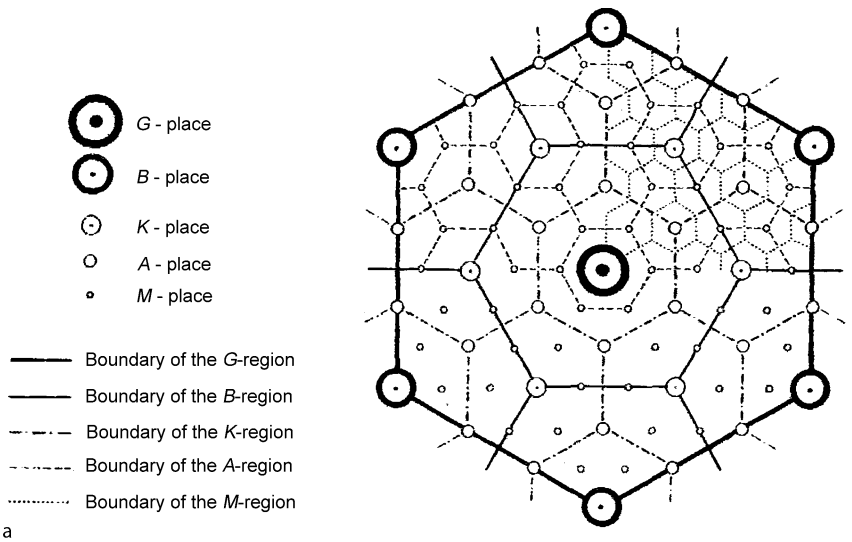
Self-Organization and the City, Figure 1
Thünen's type land-use system as transformed into an urban land-use system by location theory. Businesses are prepared to pay high rent at the center of the city, but are reluctant to "live" far from it. Their spatial demand curves (or rbc – rent bid curves) are thus the highest and steepest. Industrialists, in this exposition, are exactly the opposite and residents are in between: they cannot afford to pay the high prices at the center, but are prepared to live far from it, and so on. Each land use thus occupies a ring were it can pay (bid for) the highest rent. Note that the principle of marginal utility which is implicit in Thünen's landscape, here appears explicitly as the central economic principle



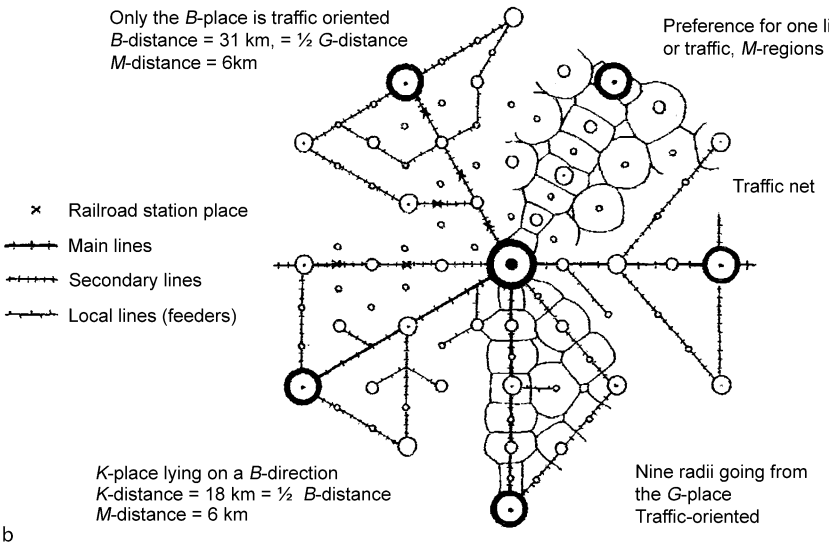
Self-Organization and the City, Figure 2
The diagrams of Thünen's *Isolated State*: *Left, upper part of the diagram* "This shows the Isolated State in the shape it must take from the assumptions made in Section One…". *Left, lower part* "Here we see the Isolated State crossed by a navigable river. Here the ring of crop alternation become very much larger, stretching along the river … The effect of constructing a highway is similar, … " (Par. 385). *Right* "The diagram illustrates the effect of the Town grain price on the extension of the cultivated plain" (Par. 386). Source: [129]

typical of the category than others. Such cities have temporarily captured the center of the category city, pushing to the periphery the rest of the instances, only to be replaced in subsequent space-time moments by other prototypical cities, other centers and other peripheries. How does this huge network evolve in time and space? The answer is: "by means of self-organization" [95].

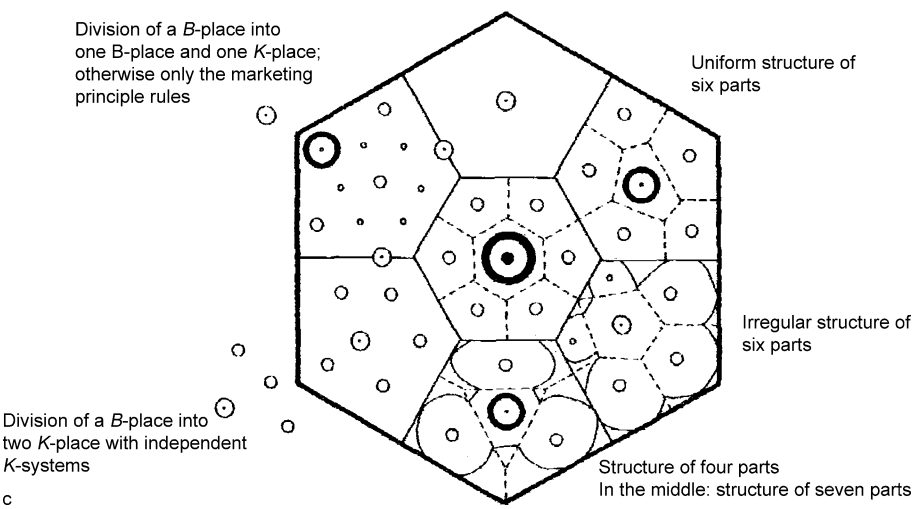
Images of Cities This section discusses images of cities that are implicit or explicit in several of the urban theories.


G - place
B - place
K - place
A - place
M - place
Boundary of the G-region
Boundary of the B-region
Boundary of the K-region
Boundary of the A-region
Boundary of the M-region

a


Only the B-place is traffic oriented
B-distance = 31 km, = ½ G-distance
M-distance = 6km
Preference for one line
or traffic, M-regions
Traffic net
Railroad station place
Main lines
Secondary lines
Local lines (feeders)
K-place lying on a B-direction
K-distance = 18 km = ½ B-distance
M-distance = 6 km
Nine radii going from
the G-place
Traffic-oriented

b


Division of a B-place into
one B-place and one K-place;
otherwise only the marketing
principle rules
Uniform structure of
six parts
Irregular structure of
six parts
Division of a B-place into
two K-place with independent
K-systems
Structure of four parts
In the middle: structure of seven parts

c

◄ **Self-Organization and the City, Figure 3**
Christaller's systems of central places according to the three locational principles. a The marketing regions in a system of central places. b A system of central places developed according to the traffic principle. c A system of central places developed according to the separation principle. (Source: Figs. 2, 4, 6 in [36])

Only theories that facilitate subsequent discussion will be discussed.

The Economic City The city is portrayed as a center-periphery structure with several rings that come into being by an economic competition between land uses that differ in their spatial demand function (Fig. 1). The land uses with the highest and steepest *spatial demand curves* or *rent bid curves* [8,69] capture the central ring – the most accessible area of the city – thus forming the CBD (central business district); the rest of the land uses occupy the peripheral rings. The origin of this city image is von Thünen's [129] *Isolated State* (Fig. 2): Originally formulated as a theory of agricultural land uses it became the founding theory of all location theories including urban land use theories.

The City as a Central Place The city is perceived as a central place that mediates between the city's complementary region and other cities that form a hierarchical network of central places (Fig. 3). The origin of this view is Christaller's [36] *central place theory* that perceived the city as a central place for tertiary activities (a market place, transportation node and administrative center). Losch's [81] central place theory was more ambitious and complicated and portrayed the city as a central place for all production, consumption, transportation and political activities (Fig. 4).

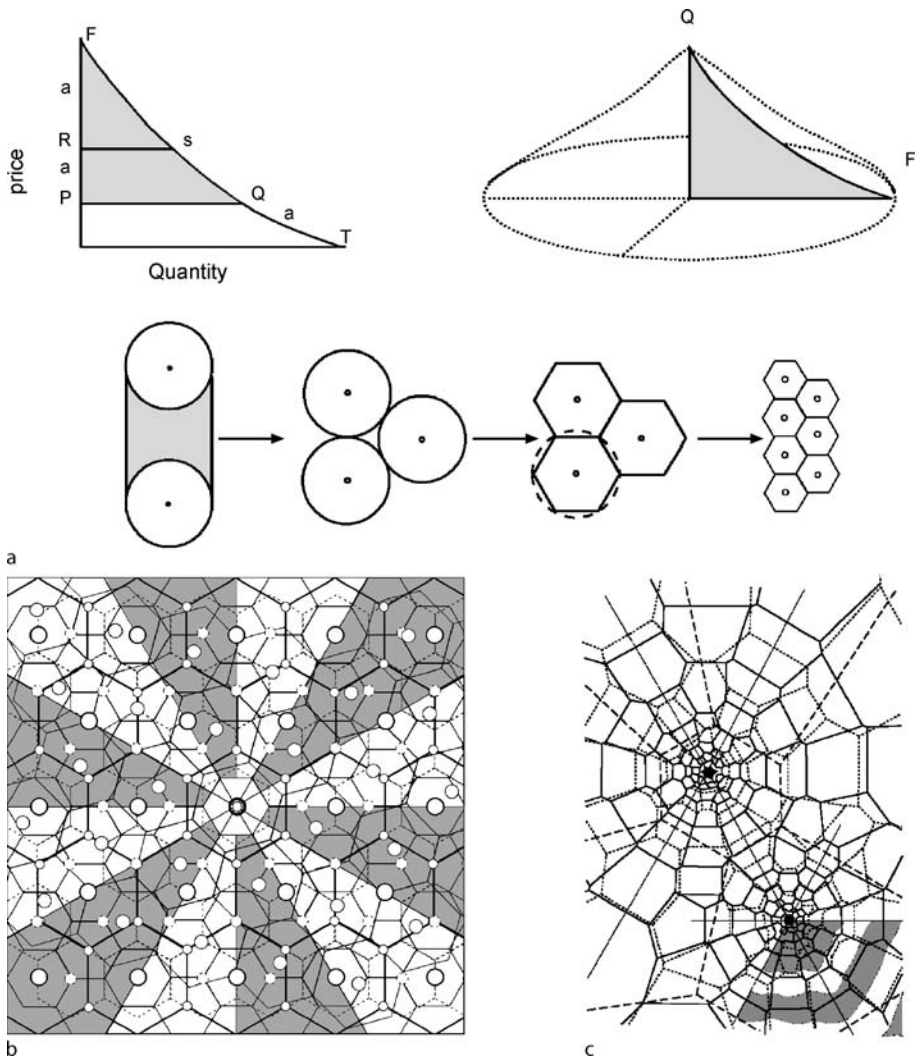
The City as a Node in a System of Cities The city here "looses" its autonomy in the sense that it is perceived as a node in a system of cities – no attention is paid to the city's role or function; the focus of interest is on the system as a whole. This view is due to Auerbach [10] who already at the turn of the 20th century showed that the rank-size distribution of cities obeys the power law. In a famous work from 1949 Zipf showed that this rank-size distribution typifies not only cities (Fig. 5), but a whole range of phenomena [138]. Zipf's work provided a source of inspiration to a long list of subsequent studies on systems of cities [111].

The Ecological City As in the economic city, here too, the city is portrayed as a center-periphery ring structure. However, here the city's structure emerges out of a competition between cultural and socio-economic groups in a way similar to competition between species in natural ecology. The various studies in this domain are termed *urban ecology*; their source of inspiration is the Chicago school of social ecology. Several urban landscapes were suggested the most dominant one is Burgess' [30,31,32] ring structure (Fig. 6).

The City as a Representation of Society The city is here perceived as a spatial representation of society as a whole. This image of the city emerged in the early 1970s as a consequence of a paradigm shift the study of cities underwent – from liberal social and economic theories to more radical ones with Marxism being the most dominant view. Two Marxist interpretations of the city can give the flavor of this approach. The first is Castells' [33] view according to which the city is a spatial representation of the structure of society as perceived by structuralist-Marxist theory (Fig. 7). The second is Harvey's view according to which the city's landscape emerges, as a logical consequence, out of internal contradictions inherent in the capitalist *mode of production* that according to this view, dominates world society of the 20th and 21st centuries. Namely, between forces of spatial agglomeration and processes of spatial dispersion. As illustrated in Fig. 8, this tension can be resolved only by *the urbanization of capital* [47,63].

The City as a Socio-cultural Force The city is here perceived as a force that is shaping the life of the people living in it. In urban societies it implies that the city is in fact shaping society. This view is due to the study of Wirth's [135] *Urbanism as a way of life* and also of Park's study *The City* [88]. In 1970 Lefebvre has published a monograph *La Révolution Urbaine* suggesting from a Marxist point of view that society is reaching a stage of being completely urban so that urbanism is replacing industrialism as the major force of society [79].

The Postmodern City The city of 21st century is described as the postmodern city [95]: Untamed, shrew, capricious, ever-changing; actually it is not a city but a text written by millions of unknown writers, unaware that they are writers, read by millions of readers, each reading his or her own personal and subjective story in this ever-changing chaotic text, thus changing and recreating and further complicating it. Today's urbanism is a big theater at the center of whose stage we see a kaleidoscope of shapes, forms, high-tech science-fiction structures, cultures and sub-cultures, Italians, Chinese, Japanese, Jews, Indians, Gays, Lesbians. Yapese; nothing is stable, nothing



Self-Organization and the City, Figure 4

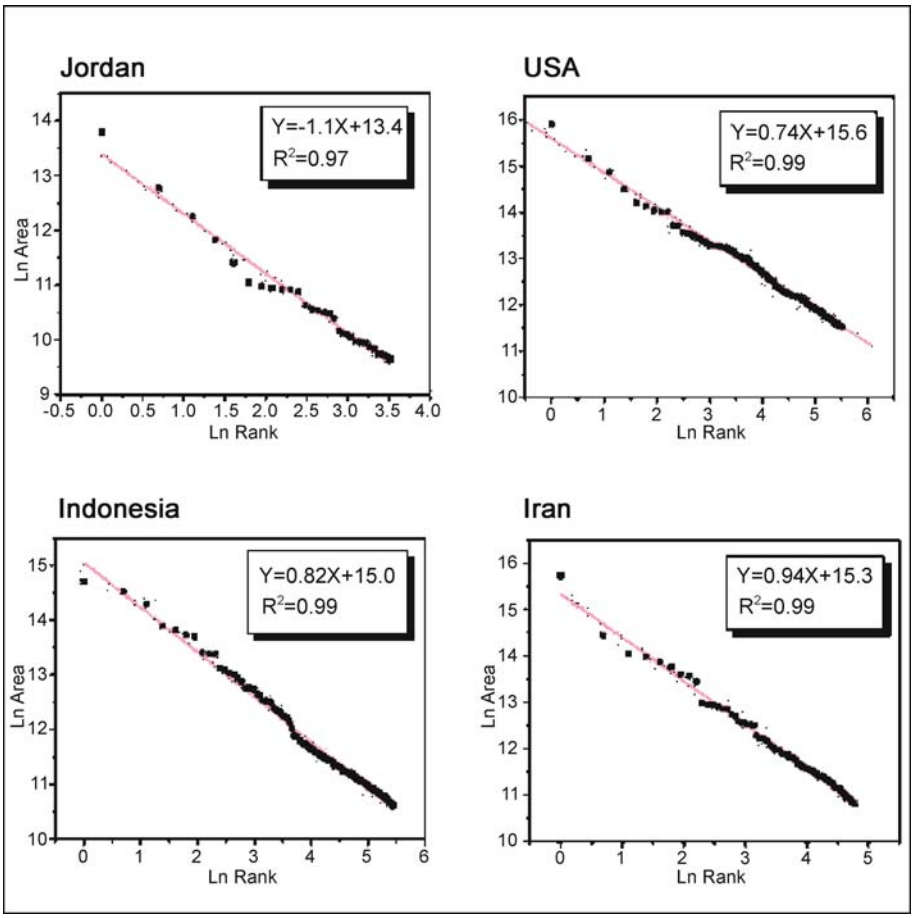
a The derivation of Lösch's system of central places. *Top* The derivation of a *spatial demand cone* with its market area (*right*) out of an "ordinary" demand curve (*left*). *Bottom* Development of market areas from the large circle to the final small hexagon. Source: Figs. 20–23 in [81]. b Lösch's derived system of central places with their market areas, divided into "city-poor", "city-rich" sectors. Source: Fig. 28 in [81]. c A Lösch system of central places modified by Isard [69] so as to be consistent with the resulting population distribution

is true nor matter for more than a second, not the Marxist urban categories, nor any other grand theory or truth; all must go, must move, clear the way to the new next whatever it is.

The Self-organizing City Strangely enough, an image of the city similar to the postmodern one is emerging out of complexity studies of cities. This is a seemingly similarity, however; a closer look reveals, first, that theories of complexity made a direct and explicit link to the views of cities as central places, to the studies on systems of cities and to the ecological views on the city. Second, that the notion of self-organizing city has several important resemblances with modern social theory oriented urban studies that perceive the city as the representation of society.

In What Sense Are Cities Self-organized Systems?

Self-organization is a property of systems that are open and complex. No one plan such systems, no one fully controls them and yet they have order, rules, organization and all these emerge spontaneously by means of self orga-



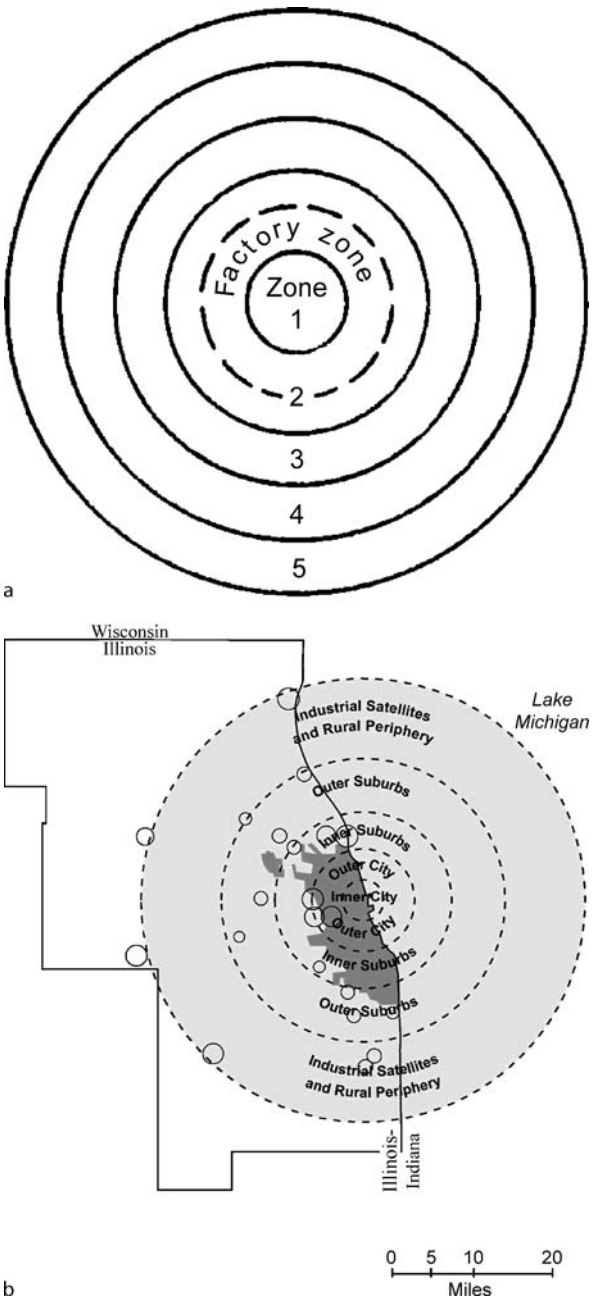
Self-Organization and the City, Figure 5
Rank size distribution of cities with more than 100,000 inhabitants in four different countries. Source: Fig. 2.7 in [27]

nization. A nice example of self-organization is provided by human languages (Chinese, Hebrew, English . . .). Each such language is an open system; each is a complex system; each is a system that emerged out of synergetic interaction between a huge number of people (the "parts" of such systems); no one has ever fully controlled languages; no one has fully planned a language; and yet each of the human languages has order, rules, organization and all these emerged spontaneously by means of self organization.

Similarly to human languages, each city is an artifact; each is an open system; each is a complex system; each city is a system that emerged out of interaction between a huge number of people; no one fully controls it; and yet it has order, rules, organization and all these spontaneously by means of self organization.

But cities are not languages. For one thing, their products are stand-alone objects such as buildings, roads, bridges, etc. that exist and survive independently of their producers. The products of languages are humans' voices and gestures that have no existence independent of their producers. Cities, in this respect, are akin to writing and texts – the external, stand-alone, representations of languages. The appearance of cities, some 5,500 years ago, hand in hand with writing, is, to my mind, not accidental.

A second difference concerns planning: There are no language planners and the attempt to "plan" the international language of Esperanto ended in failure. But there are many city planners – much more than appreciated in conventional planning theory. This is so because planning is a basic human property with the implication that each agent operating in the city (person, family, company) is a planner on a certain level. In certain cases, because of the nonlinearities that typify the complexity of cities, the planned action of a single individual might influence the city more than that of the official planners and their plans. Urban dynamics can thus be seen as an on-going interaction between planners and their plans when none of them



Self-Organization and the City, Figure 6
a Burgess concentric zone model. (1) Central Business District, (2) Zone in transition, (3) Zone of working men's homes, (4) Residential zone, (5) commuters' zone. b Regional differentiation of Chicago as commonly presented in several geographical textbooks. Source: [95]

can fully determine the final form and structure of the city. They are all *participants* in a big city-planning game (see Part III in [95]).

The Inconsistency Between Self-Organization and Planning

All theories of cities were associated with city planning. The basic idea is that the city is an artifact and as such a product of humans' intentions and needs. Planning is needed in order to implement human needs and intentions in a rational way. The notion of complexity suggests that the city is a product of self-organization; if this is so, who needs city planning? The answer to this inconsistency has already been given above: the plans produced by city planners, like those produced by "ordinary" urban agents, are *participants* in a big city-planning game.

Complexity Theories of Cities – an Overview

The discussion in this section proceeds under the titles of eight "cities" that are related to general theories or specific methodologies. It starts with 'dissipative cities' to indicate that Prigogine's was the first complexity theory applied to the study of cities

Dissipative Cities

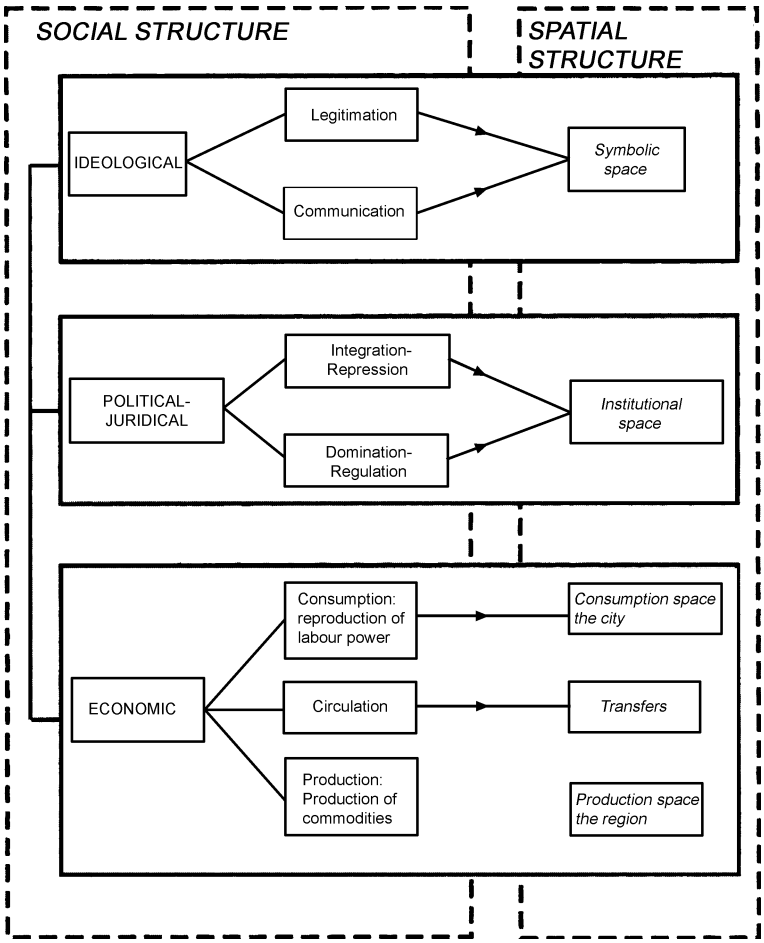
In their introduction to *Self-Organization in Nonequilibrium Systems* Nicolis and Prigogine [87] use the example of a city as a metaphor to convey to their fellow physicists what they mean by "self-organization."

> "An appropriate illustration would be a town that can only survive as long as it is a center for inflow of food, fuel … and sends out products and wastes."

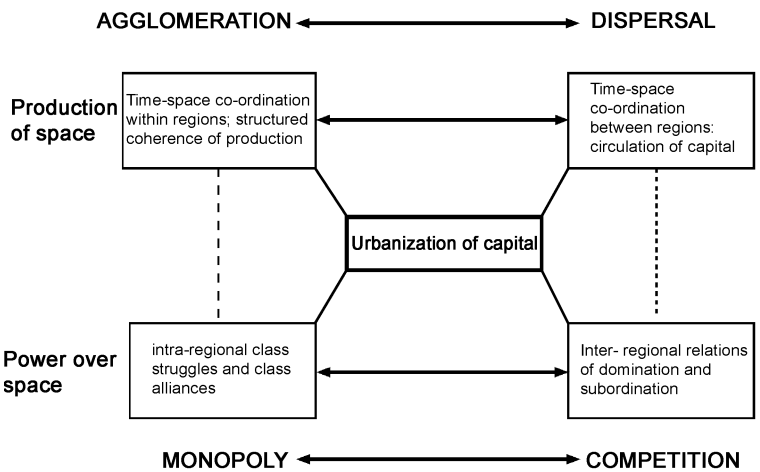
Peter Allen [5] – Prigogine's student – showed that towns and cities are not just metaphors, but genuine self-organizing systems. He did so by reformulating *central place theory* (above Sect. "Images of Cities") in terms of Prigogine's theory. (Note the resemblance between the hexagonal landscapes of *central place theory* and the hexagonal Bénard cells – one of the canonical experiments of the paradigm of self-organization).

Allen and co-workers' have developed a sequence of several models which elaborated their theoretical treatment of hierarchical landscapes of central places, first with respect to systems of cities in a given region and later at the intra-urban scale in connection with a single city. At a later stage they have also applied their models to real case studies of Brussels and the Belgian provinces [121], see also [109].

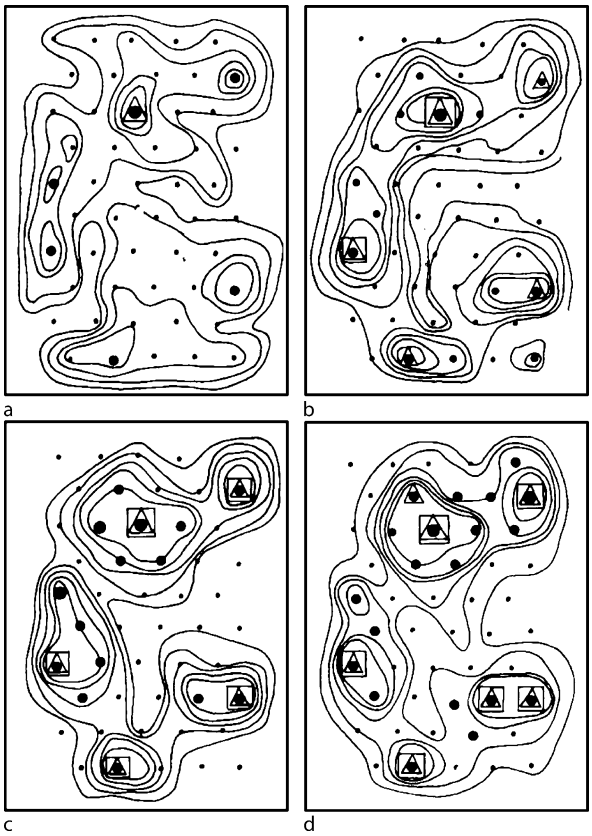
A typical model of Allan's starts with an infrastructure of localities in a region, each with its residents and jobs. The actors are individuals who migrate in order to get employment, and employers who offer or take away



Self-Organization and the City, Figure 7
The Marxist city as a spatial representation of social structure [33]. Source: [47]



Self-Organization and the City, Figure 8
The Marxist city as an outcome of basic tensions in the landscape of capitalism [63]. Source: [47]



Self-Organization and the City, Figure 9
Allen and Sanglier's simulated evolution of a dissipative system of cities. a at time (*t*) $t = 4$; b at $t = 12$; c at $t = 20$; d at $t = 34$

jobs depending on the market's situation. The migration/interaction between localities and the introduction and extraction of economic activities (i. e. employment opportunities), create for each locality a kind of local "carrying capacity" and for the system as a whole nonlinearities and feedback loops which link population growth and manufacturing activities. An example for a simulated scenario produced by the model is Fig. 9. It starts (Fig. 9a) with a hypothetical region characterized by a rectangular lattice of homogeneous localities. Then, the mere play of chance factors, such as the place and time where different enterprises and migrations start, produce symmetry breakings which entail an uneven distribution of population, employment and so on (Figs. 9b–d). The result is an evolutionary process by which new urban centers emerge, grow, and form the whole of the regional system of central places; as the system evolves, some old localities grow, others decline or even disappear, thus constructing the specific history of this region.

Allen and co-workers' approach exposes the similarity and difference between the "old" static approaches of Christaller and Lösch, and the new treatment by means of self-organization. In both, economic activities and interactions give rise to cities as central places. However, while in the old formulations the landscape reflects an equilibrium state which is the optimized sum of the properties of the various economic forces, the new landscape reflects a far-from-equilibrium situation in which the spatial hierarchical order among the central places is obtained, maintained and then transformed, by means of interplay between interaction, fluctuations and dissipation.

Synergetic Cities

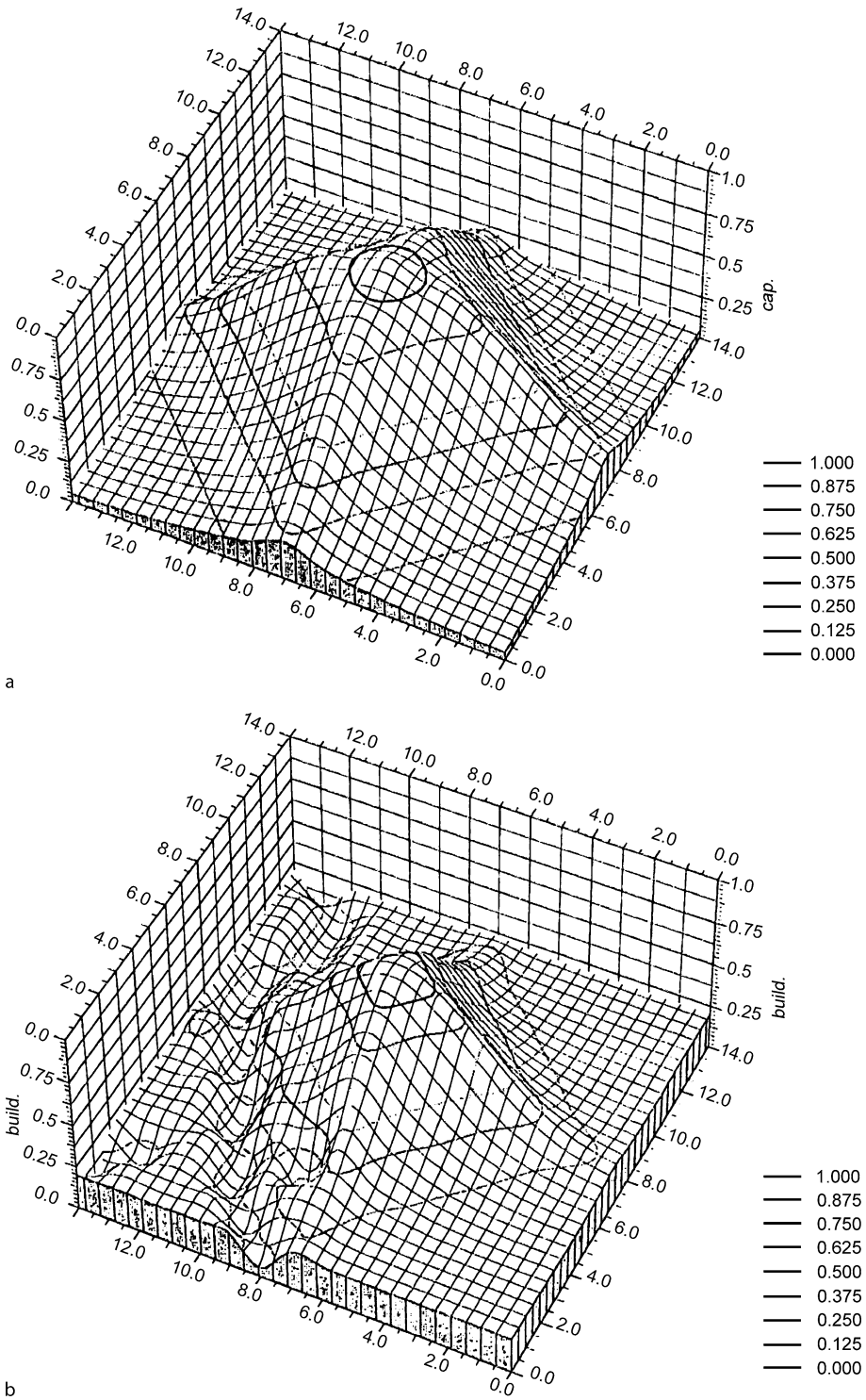
Two main approaches of synergetics have been applied to the study of cities. The first is the master-equation approach that is characteristic of Weidlich and co-workers studies in sociology, economics and urban dynamics [131,132,133,134]. For many years this was the main synergetic approach to cities, and most applications thus far have been within this conceptual frame [48,112,119]. The second, the pattern recognition approach, typifies the synergetics' treatment of pattern formation, cognition, pattern recognition and brain activities, as developed in the last three decades by Haken and co-workers [56]. Since the 1990s this approach has been applied to the study of cities as self-organizing systems [58,95].

Slow Cities and Fast Regions One way to look at Haken's synergetics and its slaving principle is in terms of interplay between slow and fast processes:

> If in a system of nonlinear equations of motion for many variables these variables can be separated into slow ones and fast ones, a few of the slow variables … are predestined to become "order parameters" dominating the dynamics of the whole system on the macro-scale [134].

This perspective stands at the basis of Weidlich's and co-workers studies on sociodynamics and cities [131,132,133,134]. According to this perspective, fast and slow processes are easily identifiable in processes of settlement and urbanism. The fast ones typify the local microlevel of building sites, streets, subways, etc., whereas the slow processes typify the macrolevel of whole regions which are often described as systems of cities. The relations between the slow and the fast processes are described by the slaving principle: on the one hand, the regional system

> serves as the environment and the boundary condition under which each local urban microstructure evolves. On the other hand, the … regional macrostructure is … the global resultant of many local structures [134].



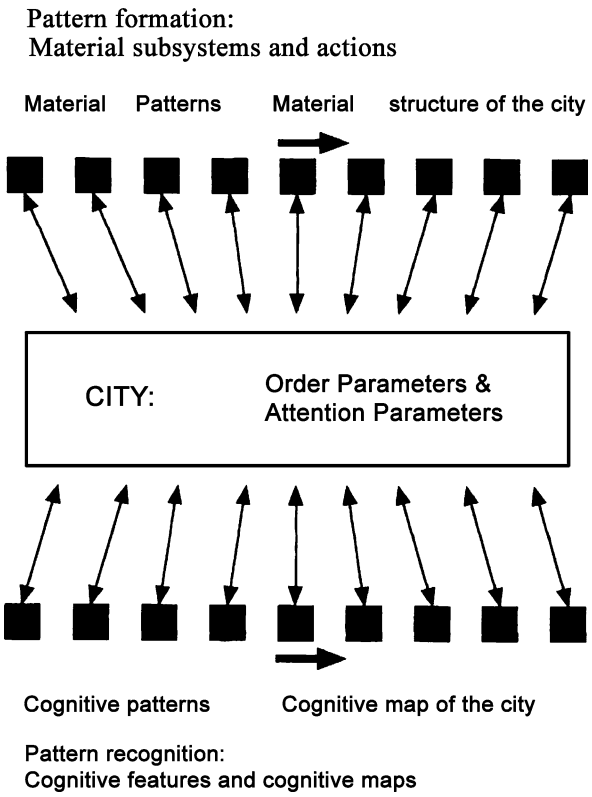
Self-Organization and the City, Figure 10
Building and development under population pressure [134]. *Top*: on a uniform urban plain. *Bottom:* on an urban plain with disturbances

This circular causality between the local and the global, allows one to study global regional systems by assuming that local processes adapt to the slow regional ones, and to study local urban processes by treating the regional context as given, and of course to study the complex interplay between the local and the global. In all three cases Weidlich has recently prescribed a four stages approach: stage 1 concerns the configuration space of the variables; stage 2, measures the utility of each configuration; stage 3, defines transition rates between configurations which are in fact utility differences; stage 4 derives stochastic or quasi-deterministic evolution equations for the system under consideration. The central evolution equation is the master equation which defines the probability that the configuration under examination is realized at a certain time.

The above theoretical procedure has been used to study the role of population pressure in "fast and slow processes in the evolution of urban and regional settlement structures". Figure 10 brings some results from these studies, in which the city capacity for building and development is related to population pressure. Figure 10a shows the evolving city capacity when the urban plain is uniform, and Fig. 10b, when it is disturbed in one of its sites.

Pattern Formation and Pattern Recognition in the City

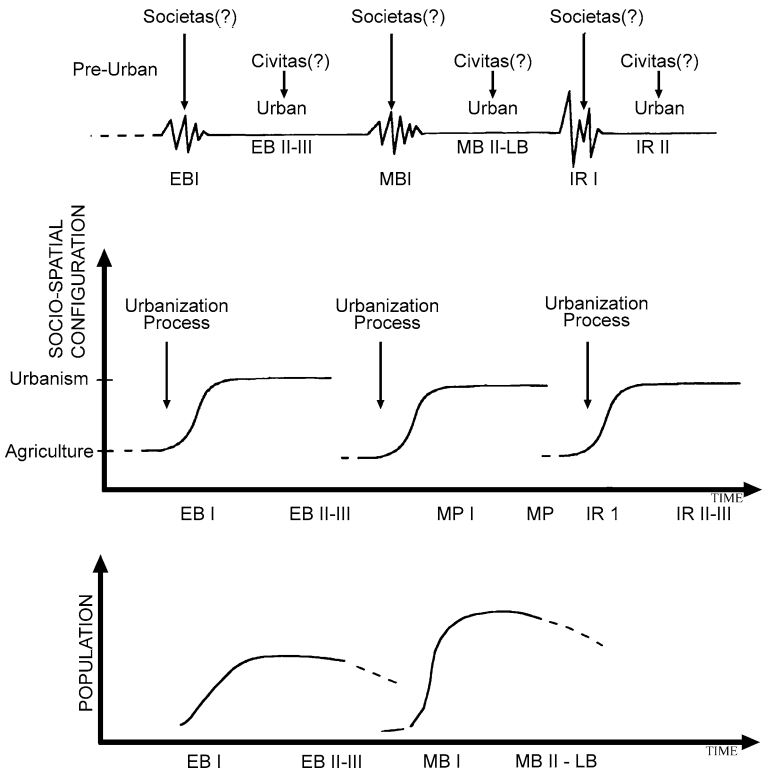
The paradigm of pattern recognition was derived by an analogy to the material process of pattern formation [54]. Haken and Portugali suggested that the synergetic pattern recognition paradigm is specifically attractive for the study of cities [58]. The latter can be perceived as self-organizing systems which are both physical and cognitive: individuals' cognitive maps determine their location and actions in the city, and thus the physical structure of the city and the latter simultaneously affects individuals' cognitive maps of the city. In their preliminary mathematical model Haken and Portugali construct the city as a hilly landscape which is evolving, changing and moving as a consequence of the movement and actions of individuals (firms etc.). The latter give rise to the order parameters which compete and enslave the individual parts of the system and thus determine the structure of the city. The new feature of this exposition is that the order parameters enslave and thus determine two patterns (Fig. 11): one is the material pattern of the city, and the other is the cognitive pattern of the city – its cognitive map(s).



Self-Organization and the City, Figure 11
The city as self-organizing systems which is at the same time both physical and cognitive. Its emerging order- and attention-parameters enslave the city's cognitive and material patterns. Source: [95]

Chaotic Cities

Self-organization is often regarded as a theory about *Order out of Chaos* [110] and yet, with a few exceptions [39,95] chaotic behavior is rarely studied in cities. Most complexity studies of cities perceive cities as ordered structures which the theory of complexity explains just how their order state was created. According to Batty (see p. 29 in [16]) this is "because the required growth rates [for chaotic behavior to appear in cities] are far too large". My view is that this is due, firstly, to the tendency of most students of complexity to focus on the short-term dynamics of Western cities from which perspective cities are indeed structurally stable. However, when the focus of interest turns to the long-term rural urban migration process in a country such as China, or to the archaeological record of the rise and fall of urban cultures [95,105], chaos suddenly appears. Looking at this *longue durée* [26] of cities, their evolution exhibits a very distinct and routinized path: a long period of "steady state", followed by a short period of strong fluctuations or chaos, from which the system re-emerges to a new level of steady state and structural stability, and so on (Fig. 12). As can be seen, the urban system moves from one structurally stable state to another, via bifurca-



Self-Organization and the City, Figure 12

The evolution of the settlement system in Palestine from the Early Bronze Age period to the Iron Age. Source: [100]. *Top*: A description in terms of chaos and order. *Middle*: A description of the process as a rhythm between agriculture and urbanism, interrupted by global collapses of the urban system. *Bottom*: Calculated population changes in the Early Bronze and Middle Bronze periods. Source: [93,95]

tions, when every evolutionary move is a transition from a microscopic chaotic state to an ordered, macro, steady state.

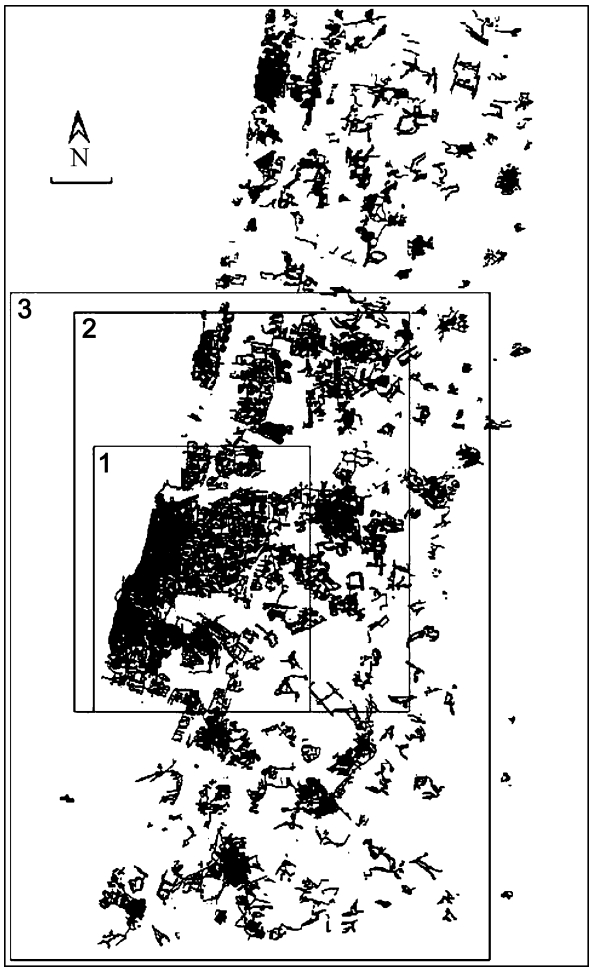
Secondly, this is due to the fact that phenomena of chaos and their role in cities during their short-term structurally stable periods have not as yet been fully studied. In a preliminary attempt to do so it has been found that often, when the city as a whole evolves stably, a few local unstable chaotic areas are found captive within the otherwise stable city. This phenomenon has been termed *the captivity principle* with the suggestion that it might play a supplementary role to Haken's slaving principle (see Chap. 5.8 in [95], [61]), namely, that these local islands of instability are needed in order to maintain the overall global stability of the city. Figure 15 below provides a hypothetical example simulated by means of cellular automata.

The play between chaos and order might show up also in the daily routines of cities. The movement of cars on the roads, or of pedestrians on pavements, are characterized by shifts between instable and stable motions and might thus be candidates for this kind of interpretation.

Fractal Cities

Mandelbrot's theory of fractals is based on the notions of *self similarity* and the *fractal dimension*, and, on the idea that a rather simple iterative process might produce highly complex geometrical shapes. Using these principles, several scholars have demonstrated, first, that the complex geometries of urban form, growth and evolution, on intra-urban and inter-urban regional scales, can be generated by means of a simple iterative process with a few and simple rules. Second, that many urban structures are self-similar and have fractal structure. The most comprehensive study in this domain is Batty and Longley's [19] *Fractal Cities*, to which one can add studies on urban structure, on the fractal structure of transportation networks [23], on the question "when and where is a city fractal?" [24] and more (For updated survey of studies see [16]). Figure 13 illustrates the evolving fractal structure of the Tel-Aviv metropolitan area from 1935 onwards.

Another important insight implied by fractal cities studies is that a city, or a system of cities, in a steady state

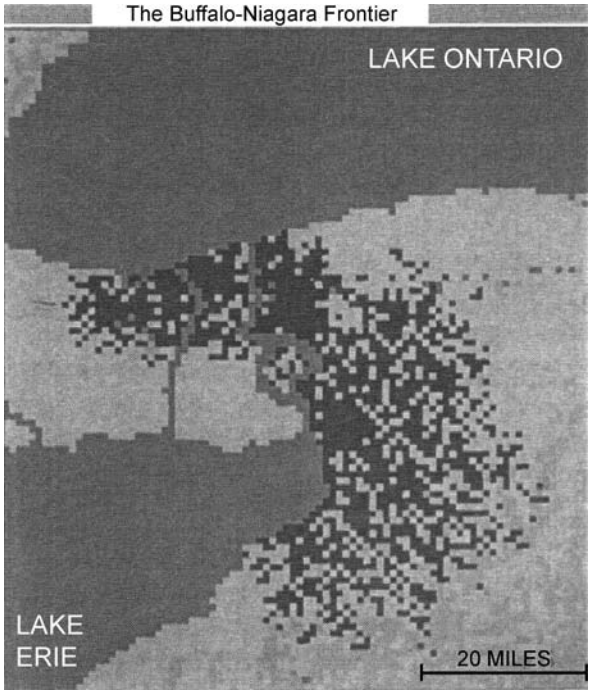


Self-Organization and the City, Figure 13
The evolving fractal structure of the Tel-Aviv metropolitan area from 1935 onwards: The central parts 1 and 2 were fractal during the entire period, while their fractal dimension increased with time. The entire metropolitan area became fractal only after 1985. In 1991 the fractal dimension of the Tel Aviv metropolitan area was found to be 1.667 with error of 0.037. Source: [23]

does not mean equilibrium and stability, as is the case of Christaller's and Lösch's central place theories, for example, but rather a rich and complex evolution and change according to a given ordering principle.

Cellular Automata Cities

The attraction of CA (cellular automata) models to the study of cities is almost self-evident. Real cities are built of discrete spatial units such as houses, lots, city-blocks and the like. CA models are also built of discrete spatial units – the cells. In real cities the properties of local spatial units (e. g. land value) are determined, to a large extent, in relation to their immediate neighbors; so are the

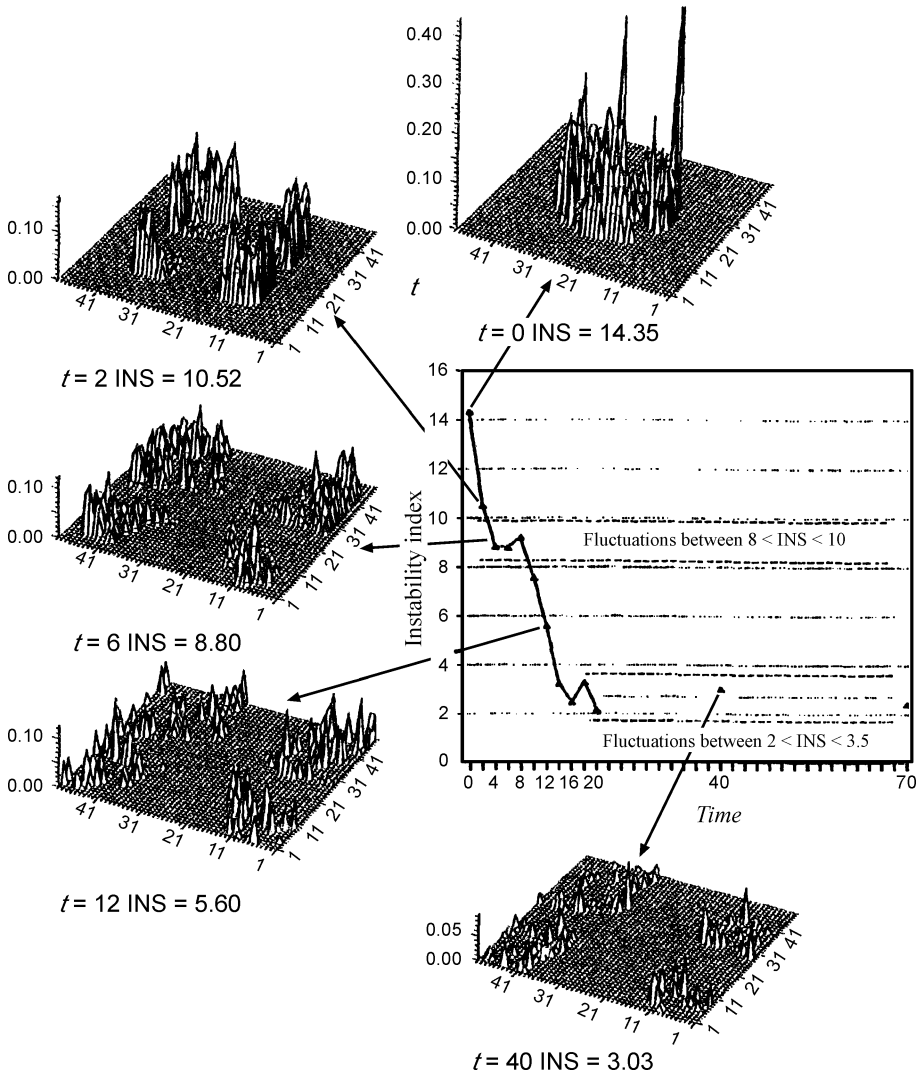


Self-Organization and the City, Figure 14
Cellular automata simulation of the Buffalo–Niagra frontier. Source: [17]

properties of the cells in CA models. These resemblances, together with the mathematical simplicity of CA models, make them natural tools to simulate urban processes. In the last few years CA urban simulation models are among the most popular approach to simulate the dynamic of cities [16,22,95].

One can divide the various models of CA cities into implicit and explicit self-organized CA cities. The first group refers to studies the aim of which is to use the simulation capabilities of the CA city in order to best-fit a certain simulated pattern to an existing one [16]. Figure 14 is an example. The general motivation here is to explain an existing or historical pattern, or alternatively to predict a future one for the purpose of planning. The fact that the model has properties of self-organization, just adds more realism and sophistication to the simulation.

The second group concerns explicit self-organized CA cities. Here the central motivation is to use the model as means to investigate the self-organization properties inherent in cities and urbanism. For example, how micro decisions and behavior of individuals and firms, taken at the local scale, are related to the global behavior and structure of the city. Such models are essentially heuristic and they regard the simulated CA city as essentially a learning device (Fig. 15).



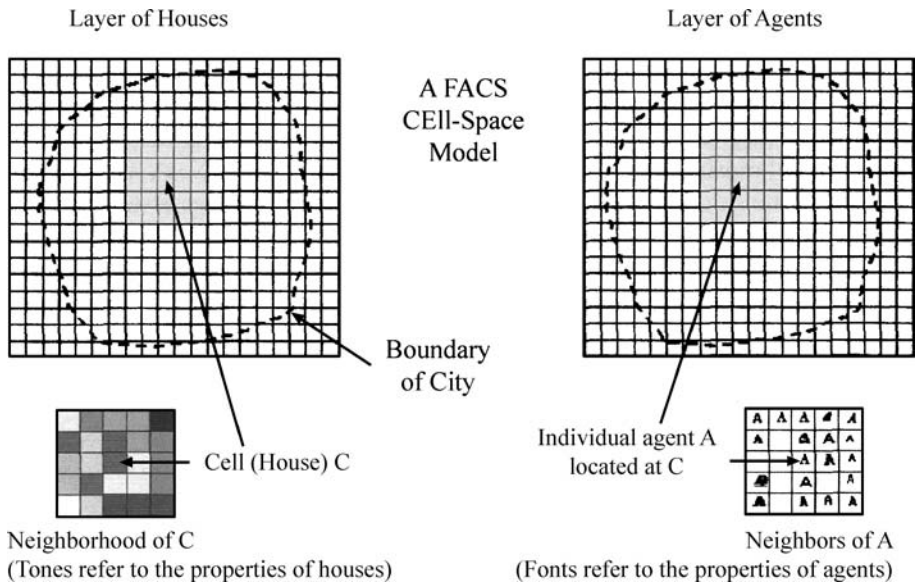
Self-Organization and the City, Figure 15
Time evolution of consecutive stages of SIS (stability-instability surface) in the development of a city with 33% neutral Greens when the rest of the Greens and all the Blues are segregatives. Source: [95]

Because of their iterative structure, CA models can be used as convenient tools to generate fractal structures [16,17] and the insight they add to our understanding of cities is similar: an iterative process guided by a few simple rules, can generate complex structures such as cities [16].

AB and FACS Cities

CA is an efficient tool to model the relations between infrastructure objects of the city. Unlike infrastructure object, urban agents have aims and plans, can learn and move in the city, and see and know beyond their nearest neighbors. *Agent base* (AB) models that are built to imitate such cognitive entities were applied to the city dynamic too. An important source of inspiration here was Schelling's model that demonstrated how local and simple behavior of urban agents can give rise to complex residential segregation in cities – even when their tendency for segregation is minimal [122]. Subsequent agent base studies have supported Schelling's finding [16] and added that a small minority of agents with a tendency for segregation is sufficient to turn the whole city into a segregative structure [95].

Free agents on a cellular space (FACS) models combine CA and AB models [95]. A typical such model is built as a superposition of a CA layer simulating the relationship



Self-Organization and the City, Figure 16
A typical FACS model is constructed of two-layers: a population layer of human agents describing the migratory and interaction activities of individuals (*right*), superimposed on a CA infrastructure describing the urban landscape (*left*). Source: [95]

between the city's infrastructure objects (buildings, roads …) and AB layer that simulates the activities of the urban agents (Fig. 16). At each model iteration new agent(s) come to the city with a certain intention in mind – say to find a house to live in. The agent then examines the available empty cells/buildings, ranks them according to its set of preferences and picks the best one. Once the agent located itself in a certain cell, the CA dynamics starts: The properties of each cell are determined by reference to the properties of its neighbors; and if the cell is occupied by a certain agent, by some mix between the properties of the agent and its neighbors. Figure 17 presents typical results.

Sandpile Cities

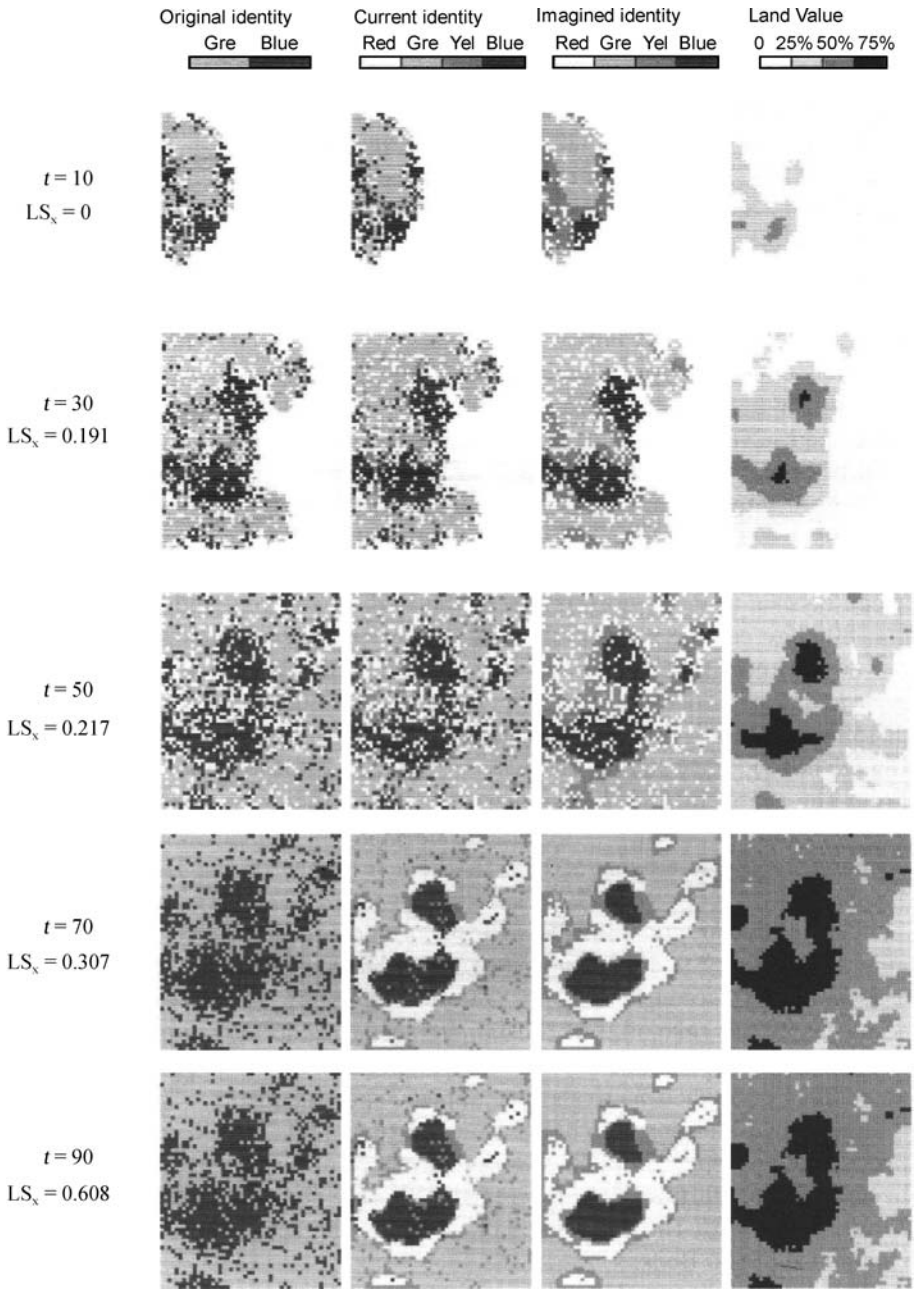
The sandpile, the canonical example of *self-organized criticality* [11,12,16,20], has two incongruous features: the system is unstable in many of its local locations; nevertheless its global state is absolutely robust: The local configurations of the sand change all the time because of the avalanches, while the statistical properties, such as the size distribution of the avalanches, remain essentially the same. Similarly to the sandpiles, cities appear volatile and fast moving at their local scales while at their global scale they appear absolutely robust [16,20]. For example, the size-distribution of many cities and systems of cities remains essentially the same under circumstances such as ongoing population growth (Above Sect. "Image of Cities").

Compared to the "grand" synergetic and dissipative cities, the sandpile city is a kind of a zooming-in to the internal dynamics of self-organized cities in their steady state periods – when they are controlled by what in synergetics is called order parameters. Sandpile cities show how complex and rich is the internal dynamics of a city in steady state (Fig. 18).

Small World Cities

The notion *network* is implicit in all theories of complexity. Recently, Watts and Strogatz [130] showed that complex networks have 'small world' characteristics [84] and Barabasi and Alberst [14] demonstrated that complex networks are scale free thus following the *power law* that according to Barabasi [13] is a mark of self organization.

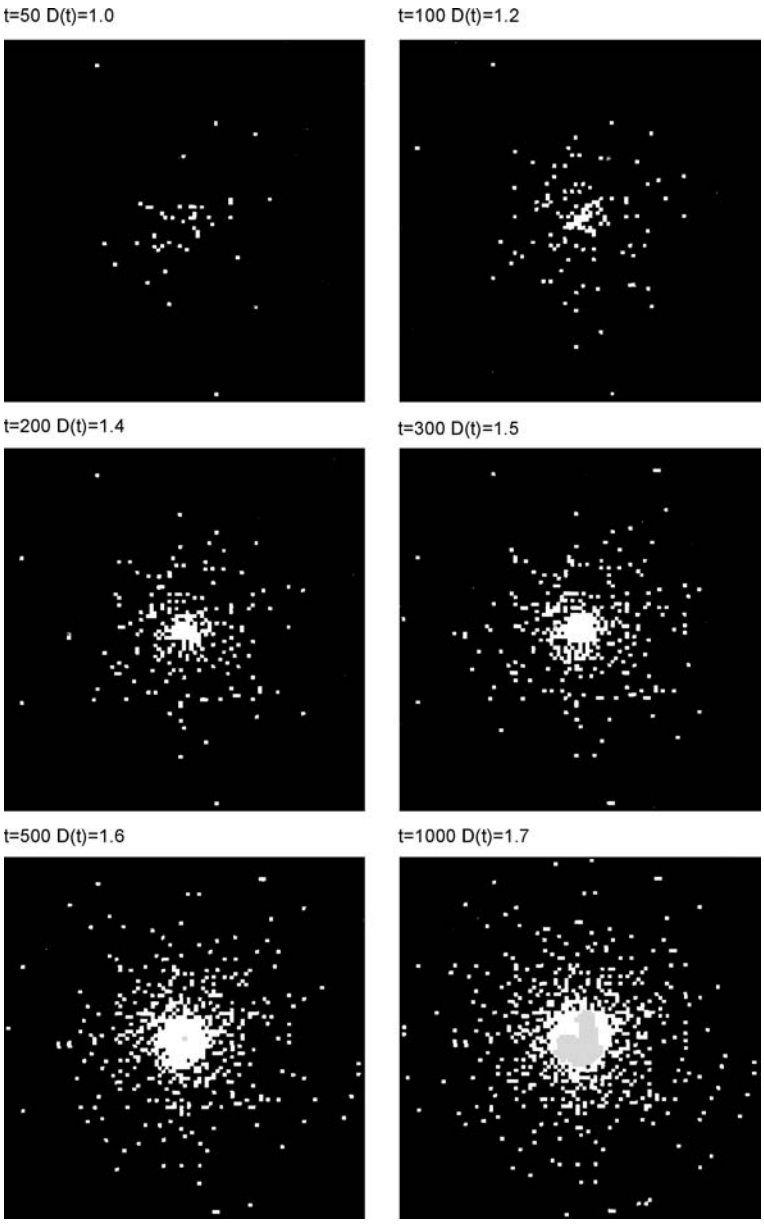
The link to cities as complex systems was just natural: The view of systems of cities as networks characterized by the power law was indicated above. Single cities too were described as networks. Thus Alexander's classic "a city is not a tree" [1] demonstrated that cities are typified not by a simple *tree network*, but by a complex *semi-lattice network* (Fig. 19). Alexander's view was recently reformulated in terms of the new science of networks [117,118]. Another example is Hillier's *space syntax* that analyzes the morphology of urban spaces in terms of networks [65,66]. Space syntax exposes the way society



Self-Organization and the City, Figure 17
Several snapshots from an evolving FACS city. Source: [95]

determines the urban morphology and the way the letter feeds back and re-shapes society. The link between space syntax and network analysis has already produced several useful results [38,41,67,89,114].

In the domain of transportation one can mention studies that characterize roads' traffic dynamics in terms scale-free networks (Fig. 20) [68,72], the same was found for the transit system in Beijing [137], pedestrian movement [73] and for the canal networks of Venice [25]. Andersson et al. [9] showed that the market dynamics generates land values that can be represented as a growing scale-free network. Finally, Batty [16] has suggested viewing cities and their dynamics from the combined perspectives of networks, fractals, self-organized criticality and AB.



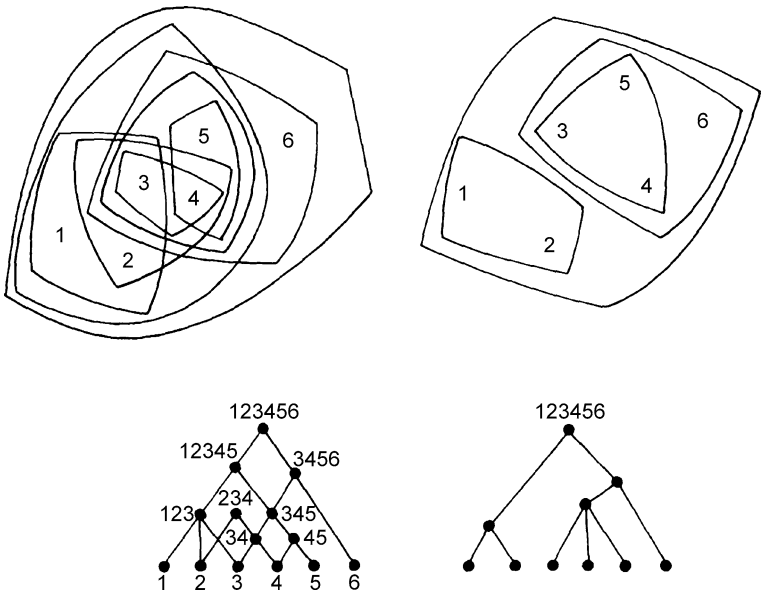
Self-Organization and the City, Figure 18
Self-organized criticality: "Simulation of a hypothetical urban growth pattern in its critical level". Source: [16]

Self-Organization and the City

Self-Organization and the City is an ongoing project that explores the city as a complex system from two interrelated perspectives: Haken's (1983) synergetic theory of complex systems, in particular from the perspective of the pattern recognition paradigm [50,51,55] and IRN – inter-representation nets [94]. The link between the two is termed SIRN – synergetic inter-representation nets.

SIRN – Synergetic Inter-representation Networks

IRN commences with a distinction between cognitively *simple tasks* that can be performed by working memory (e. g. $2 \times 3 = 6$) and *complicated tasks* (e. g. $257 \times 389 = 99\,973$) that are the result of the "magic number seven" that constraints our ability to process information in working memory [85]. One way to overcome this limitation is by means of IRN: We first externalize the task



Self-Organization and the City, Figure 19
A tree network (*right*) versus a semi-latice network (*left*). Source: [1]

(write it down on a paper); then we solve part of it internally ($9 \times 7 = 63$); externalize it again and so on in a sequence until the task is completed.

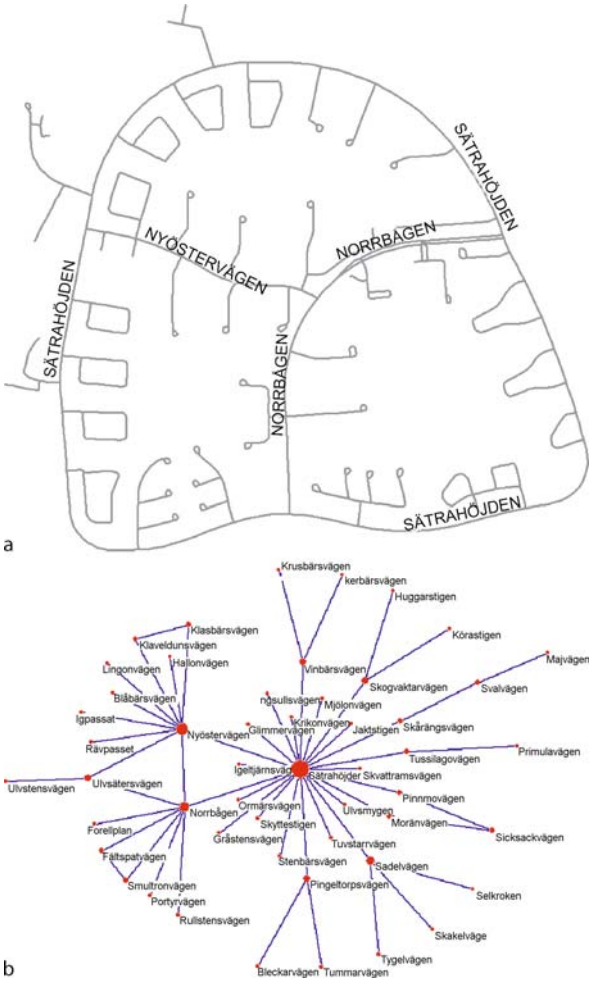
Complex tasks refer to creative cognitive tasks, when a person writes, paints, designs etc. Such a task often starts with a vague idea in mind that the person then externalizes by writing it down or painting … Here too the process proceeds by interplay between internal and external representations, but with one important addition – it involves emerging properties. It is here were synergetics gets in and the process becomes SIRN. More specifically, the process might start with a preliminary internal idea (or external cue that entails internal idea), that the person then externalizes and so on. After a few internal-external iterations an order parameter (in the sense of synergetics) emerges and enslaves subsequent iterations.

The development of the notion of SIRN was inspired by Bartlett's serial reproduction scenarios in his study *Remembering* [15]. A typical such scenario starts when a test person is shown a text or a figure and is asked to reproduce it out of memory (Fig. 21). The result is offered to a second person that is asked to do the same and so on. As shown by Bartlett, at the beginning the reproductions change from person to person, however, at certain stage they stabilize and become a *scheme*. Stadler and co-workers [126] demonstrated that the scenarios proceed as synergetic self-organized process. The focus of interest in the above studies was on the way schemata are created. Haken and Portugali have used the Bartlett scenarios as illustration of the play between internal and external representations [59,94].

It is important to emphasis, first, that external representations are media that enable communication between persons and the emergence of collective SIRNs – e. g. a brain storm. Second, that internal and external representations are *generative* – once produced, they generate new ideas and properties not seen before in previous representations.

The Basic SIRN Model Haken and Portugali [59] have cast the SIRN process into the formalism of synergetics. They started with Haken's [54] 'synergetic computer' (Fig. 22, *Top*), composed as it is of an input layer with model neurons representing the initially given input activity; a middle layer representing the order parameters, and an output layer with neurons representing the final activity of each neuron. The first step is to look at this network from the side, as indicated by the arrow. The result is shown in Fig. 22, *bottom, left*. Adding to the latter external inputs and outputs, we arrive at our basic SIRN model (Fig. 22, *bottom, right*) that has two kinds of inputs, internal and external and two kinds of outputs, again internal and external. The middle node symbolizes the order parameters that emerge out of the interaction between internal and external representations.

The basic SIRN model can be seen as symbolizing a self-organizing active agent that is subject to two flows of information: internal and external (Fig. 23). The first



Self-Organization and the City, Figure 20

A small street network (a) and its connectivity graph (b). Every node in b is labeled by the corresponding street name, and the size of nodes shows the degree of connectivity of individual streets. Source: [73]

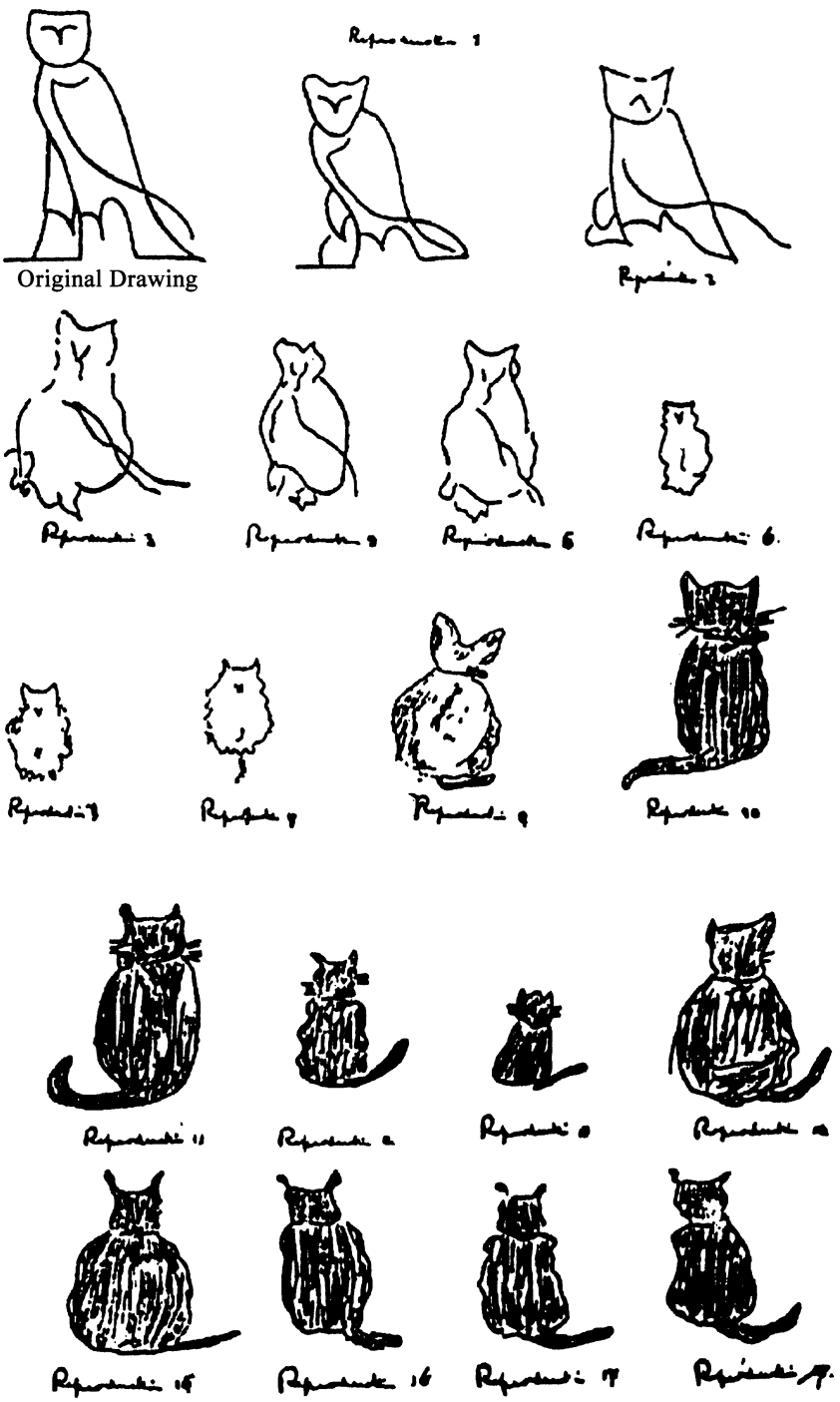
is coming from the mind/brain, in the form of ideas, fantasies, dreams, thoughts, and the like, while the second from the 'world' – via the senses, the agent's body and/or artifacts. The interaction between these two flows gives rise to an order parameter that governs the agent's action and behavior, as well as the feedback information flow to the agent's mind. 'Action or behavior' may refer to a single individual executing exploratory behavior, reproducing texts or drawing, as well as to several individuals collectively reproducing a large-scale artifact such as a city. In an analogous fashion, the 'feedback information flow' refers to the formation of internal representations, such as images or learned patterns. The order parameters are determined by a competition in line with the synergetics' pattern recognition paradigm noted above. Note that all the above steps (and below), can and have been, performed by a computer so that the approach is entirely operational.

In order to apply the basic SIRN model to specific case studies, Haken and Portugali [59] reformulated it in terms of three prototype sub-models: the *intrapersonal*, the *interpersonal collective*, and the *interpersonal with a common reservoir* sub-models (Fig. 24). The first refers to a solitary agent, the second to a sequential dynamics of several agents whereas the third to a simultaneous interaction. The third sub-model is, in fact, a theory of urban dynamic. The intrapersonal is typical to the way of an artist, for instance, develops her/his work (Fig. 25), whereas the interpersonal to the Bartlett scenario that provided a source of inspiration to IRN (above, Fig. 21).

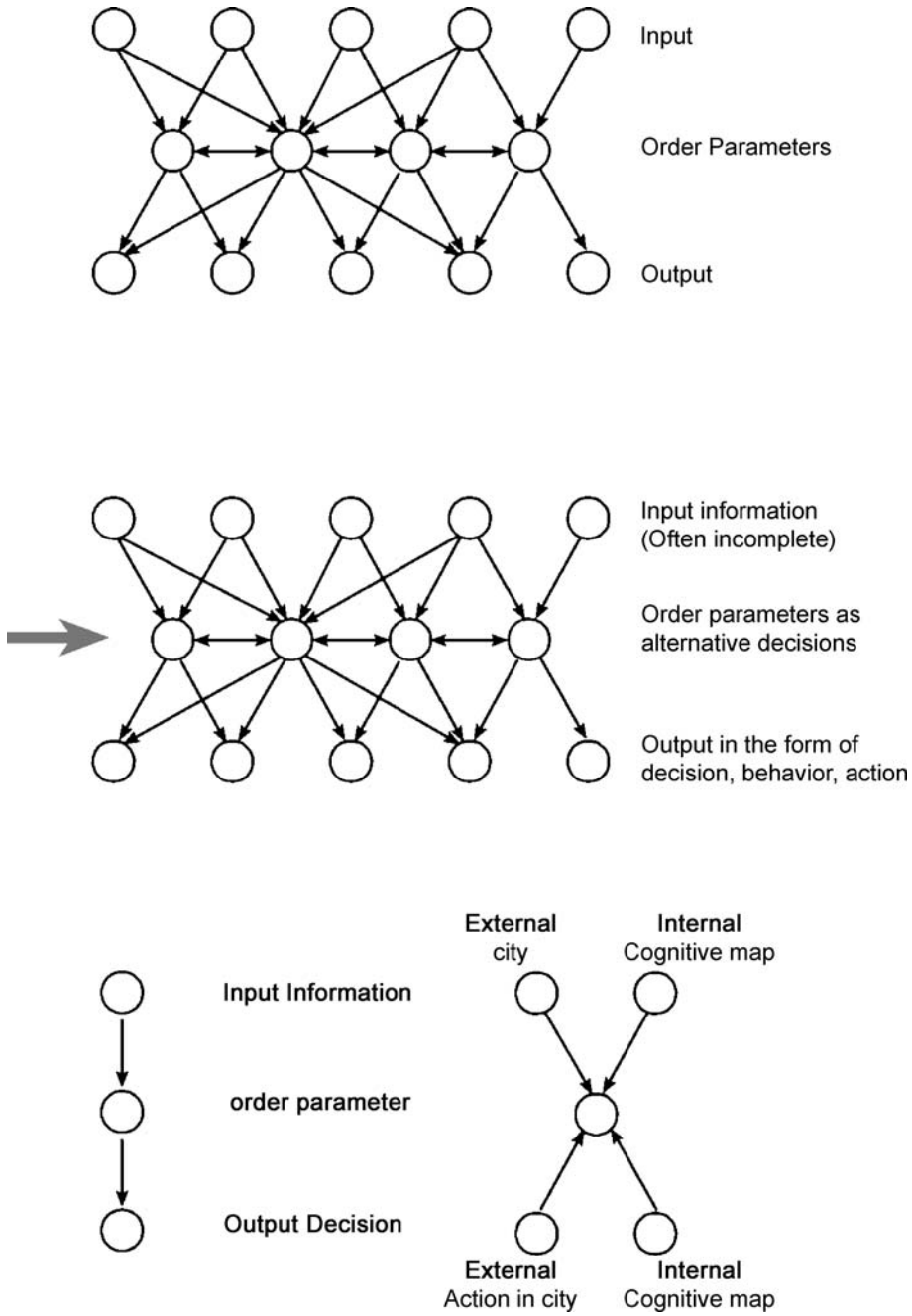
In the first two sub-models the process depends fully on the biological memories of individuals. In the third sub-model the process depends partly on biological memories, as before, but partly also on externalized non-biological memory that we term a *common reservoir*. This common reservoir of external, artificial and non-biological memory, might take the form of texts, Internet, buildings or whole cities.

Figures 24 (*bottom*) and 26 illustrate graphically this public-collective SIRN sub-model. Each individual agent is subject to internal input constructed by the mind/brain, and external input which is the legible information coming from the common reservoir, that is, the city. The interaction between these two forms of input gives rise to a competition between alternative decision rules that ends up when one or a few decision rules "wins". The winning rule(s) is/are the order parameter(s) that enslave(s) the system. The emerging order parameter governs an external output, which in the case of a city is the agent's behavior and action in the city, and an internal output, which is an information feedback loop back to the agent's mind/brain.

Both the previous sub-model and the present one involve a two-scale self-organization process: an individual-local scale referring to each individual agent as a self-organizing system, and a collective-global scale, referring to the whole city as a self-organizing system. The individual agents by their action and behavior determine the city, which by means of its emerging order parameter(s) enslaves the minds of the individual agents. In the language of synergetics this process is termed *circular causality*. In terms of social theory it is close to notions of socio-spatial *reproduction* and *structuration*. Recent applications show that the common reservoir might be a non-biological externalized memory such as a city[95,96,97,98], a planning



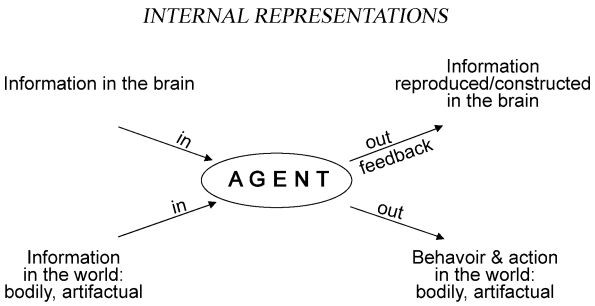
Self-Organization and the City, Figure 21
A Bartlett's scenario of serial reproduction: an Egyptian 'Mulak' (owl) transformed into a cat (see pp. 180–181 of [15])



Self-Organization and the City, Figure 22
The derivation of the SIRN model. See text

textual report or an urban planning policy emerging out of a discourse among the members of a planning team [105]. Note that as in the previous model, here too, due to circular causality, as the process evolves the subjective cognitive maps of the individual agents are becoming more similar to each other and an inter-subjective, collective cognitive map emerges. Both private–subjective cognitive maps and public-collective ones are thus *constructions*.

The City Game A simple and effective way to illustrate the SIRN view on the dynamics of cities is by means of a set of experiments termed *city games* [94]. A city game



Self-Organization and the City, Figure 23
The SIRN model symbolizes a self-organizing agent that is subject to two forms of information: internal and external, and is actively constructing two forms of information, again internal and external. Graphically, Fig. 23 corresponds to Fig. 22 (*bottom, right*) turned 180° on its NW-SE axis

can be described as a group dynamics that involves some 40 to 70 participants. Their aim is to build a city on a floor, representing the site for a new city. Each player is given a 1:100 mock-up of a building and in his/her turn is asked to locate it in the virtual city on the floor, in what s/he considers as the best location for that building. In a typical game (Fig. 27a and Fig. 27b), the players observe the city as it develops, and in the process also learn the spontaneously emerging order on the ground. It is typical in such games that, after a few initial iterations, an observable urban order emerges. The participants internalize this emerging order and tend to locate their buildings in line with it. As can be seen, the main features of such a game are the main ingredients of SIRN, namely, a sequential interplay between internal and external representations, the emergence of a collective complex city as an artifact and a typical synergetic process of self-organization. Needless to say that the city game is not a 1:1 description of reality, but an illustration of the dynamics of cities as *dual* self-organizing systems.

Cognition and the City

SIRN is at once a theory of cognition, cognitive mapping and urban dynamics. This emphasis on cognition is a direct consequence of complexity theory; a major achievement of complexity theory was to show how local behavior and interaction between urban agents give rise to the global structure of the city. The agent is thus the main and most important actor. Given this, one would assume that practitioners of complexity theory and urban simulation models will have an elaborated theory of agents' perception, behavior, decision making and action; especially so, in light of the fact that a whole body of research on agents' behavior was readily available. I'm referring to studies on spatial cognition, spatial behavior and cognitive mapping that were developed on the interface between cognitive science and urban studies [45,77,92,98,99,103]. And yet, with few exceptions such as SIRN, this body of theoretical and empirical studies is largely overlooked by students of complexity theory of cities. Researchers in this field tend to follow economists by assuming that individuals behave in space as simple "economic persons". The result is that the rather *simple* behavior of agents in the models contradicts the *complex* behavior revealed by studies on cognitive mapping and spatial behavior.

In *The Sciences of the Artificial* Simon [124] suggested that the observed complex behavior of human agents, guided as it is by aims, plans, intentions, needs, policies and so on, misleads us as it is only an external appearance of innately simple behaving entities: Similarly to simple animals, we humans as

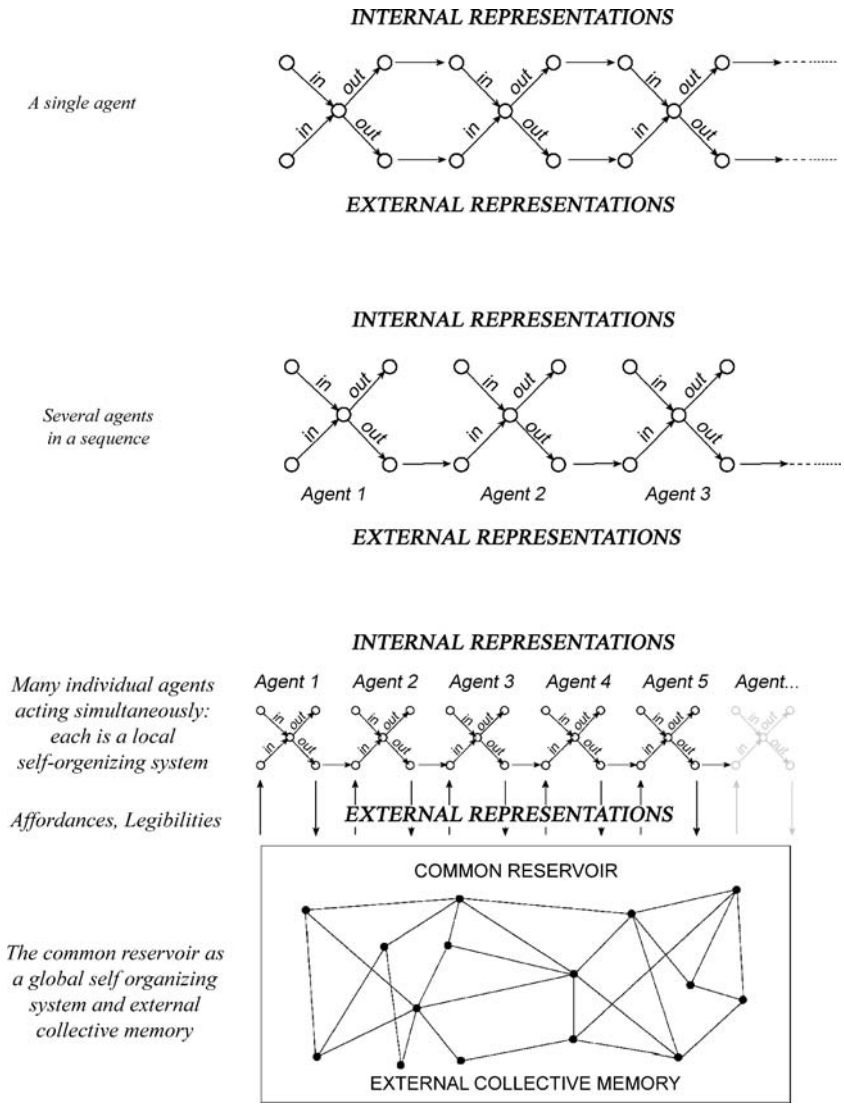
> "behaving systems, are quite simple. The apparent complexity of our behavior over time is largely a reflection of the complexity of the environment in which we find ourselves" (ibid. 53).

Most AB/CA urban simulation models are built in line with Simon's logic. They typically start with local interactions between agents having a few simple aim(s) "in mind". This interaction gives rise to an urban system, which from iteration to iteration becomes increasingly complex. Complexity is thus understood as a property of the whole global system, but not of its individual parts.

The efficiency of the *simple cause→complex effect* model is apparent. But there is a catch here: Several empirical studies, of animals' and humans' exploratory behavior, for example, falsify Simon's view [96,97,98]. Furthermore, the property of the city as a *dual* self-organizing system implies that the initial conditions of such complex systems are relatively large numbers of interacting parts, each of which is itself a complex system exhibiting complex behavior. Can there then be a science of cities that is not based on Simon's model? The answer is yes! To see how we shall look at the relations between self organization and information.

Information Compression, Inflation, Adaptation

Complexity is a property of systems that exchange matter and *information* with their environment and that their huge number of parts forms networks characterized by complex feedback and feedforward loops that allow intensive flow of *information* inside the system.



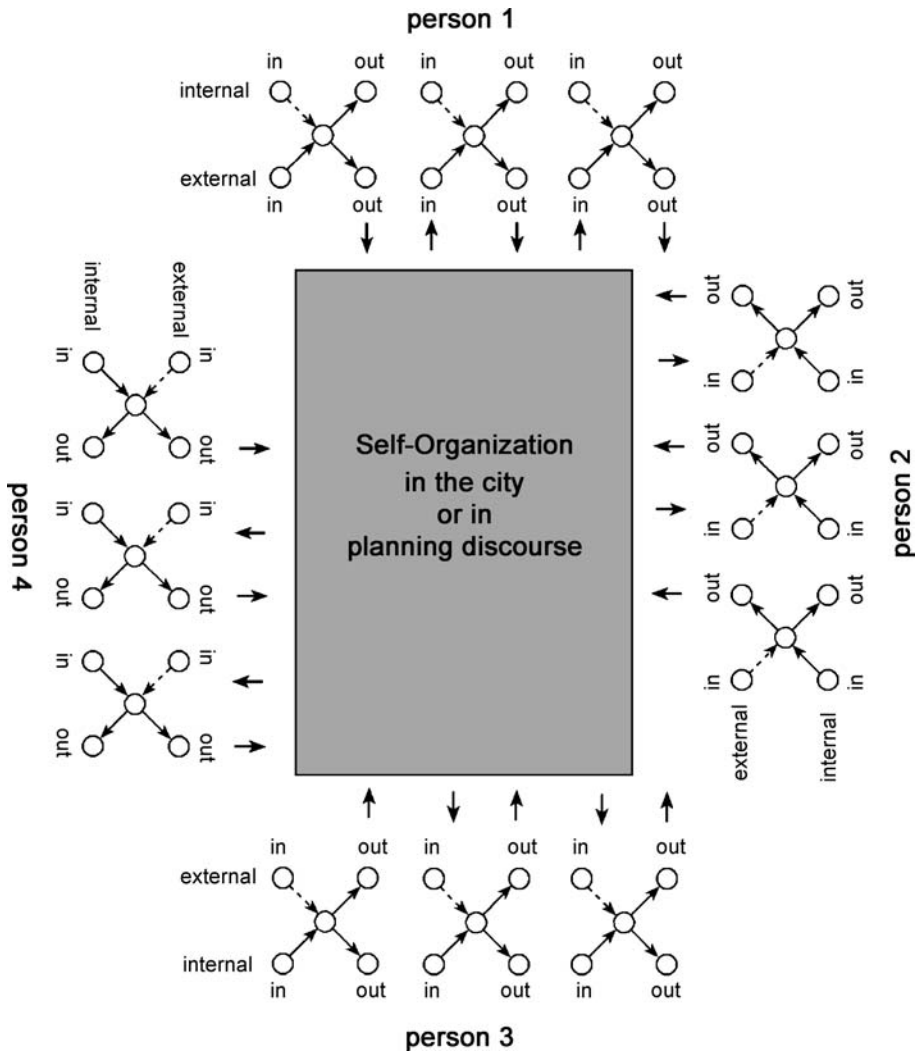
Self-Organization and the City, Figure 24

Top: **The intrapersonal SIRN submodel of a single person.** ***Middle:*** **The interpersonal submodel: serial reproduction of several persons.** ***Bottom:*** **The interpersonal with a common reservoir submodel. Note that in the intrapersonal submodel information is transmitted via external and internal outputs, in the interpersonal via external output only (action and behavior), while in the third submodel information and interaction between the agents are mediated by the common reservoir (e. g. a text, a city, Internet, etc.). Source: [96]**

The notion *Information* is associated with Shannon's theory of information [123] that has played a seminal role in the development of system thinking. In Shannon's theory the notion of information is a pure quantity (usually measured by *bits*) devoid of any meaning. Such a concept of information makes sense only in closed systems where the number of possible states the system can take is finite and a-priori known; hence the link between information and the notion of Entropy, which is a property of closed systems. For example, the information conveyed by throwing a die is 2.5 bits, that is, the logarithm to the base of 2 of the six possible states the system can take. But the complex systems we are dealing with are by definition *open*. So what is the meaning of information in complex systems?

In *Information and Self-Organization* Haken [52] suggested that complex systems 'self-organize', that is, 'interpret', the information that comes from the environment. In other words, the meaning assigned to the message depends on the receiver (the receiving system) and not just on the message itself as in Shannon's theory. Haken (see

Self-Organization and the City, Figure 25
'The Kiss' by Brancusi: from a figurative kiss in 1907, to the geometrical 'Gate of the Kiss' in 1937: An intrapersonal SIRN process in sculpturing

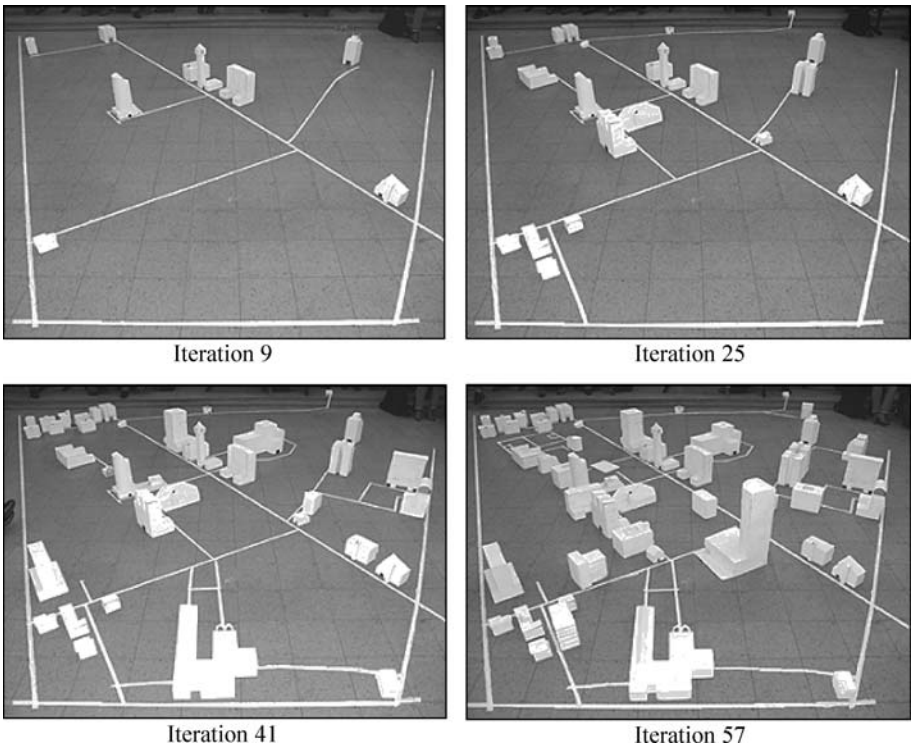


Self-Organization and the City, Figure 26

Another conceptualization of the public-collective SIRN sub-model. Each individual agent is subject to internal input – a cognitive map constructed by the mind/brain and external input –the legible information coming from the common reservoir, that is, the city of a planning team. The interaction between these two forms of input gives rise to a competition between alternative decision rules that ends up when one or a few decision rules "wins". The winning rule(s) is/are the order parameter(s) that enslave(s) the system. The emerging order parameter governs an external output, which in the case of a city is the agent's behavior and action in the city, and an internal output, which is an information feedback loop back to the agent's mind/brain

p. 15 in [53]) has consequently suggested two forms of information: *semantic information* which is information *with* meaning, versus *Shannonian information* which is "information with meaning exorcised". Haken further emphasizes that the process of self-organization implies "an enormous compression of information" (see pp. 25, 35, 151 in [52]).

Haken and Portugali [60] have studied information in the context of the city. They show that different elements of the city transmit different quantities of Shannonian information that can be practically measured by means of information *bits*, for example (Fig. 28). They further show that cognitive processes such as pattern recognition and categorization entail an enormous information compression thus affecting the quantity of the Shannonian information conveyed by the city (ibid) and that information compression is just one facet of the process – the other facet is *information inflation* [62]: In certain urban situations categorization might entail information compression while in others information inflation (Fig. 29). Infor-



Self-Organization and the City, Figure 27a
Four snapshots from a typical city game (at iterations 1, 15, 35, 57)

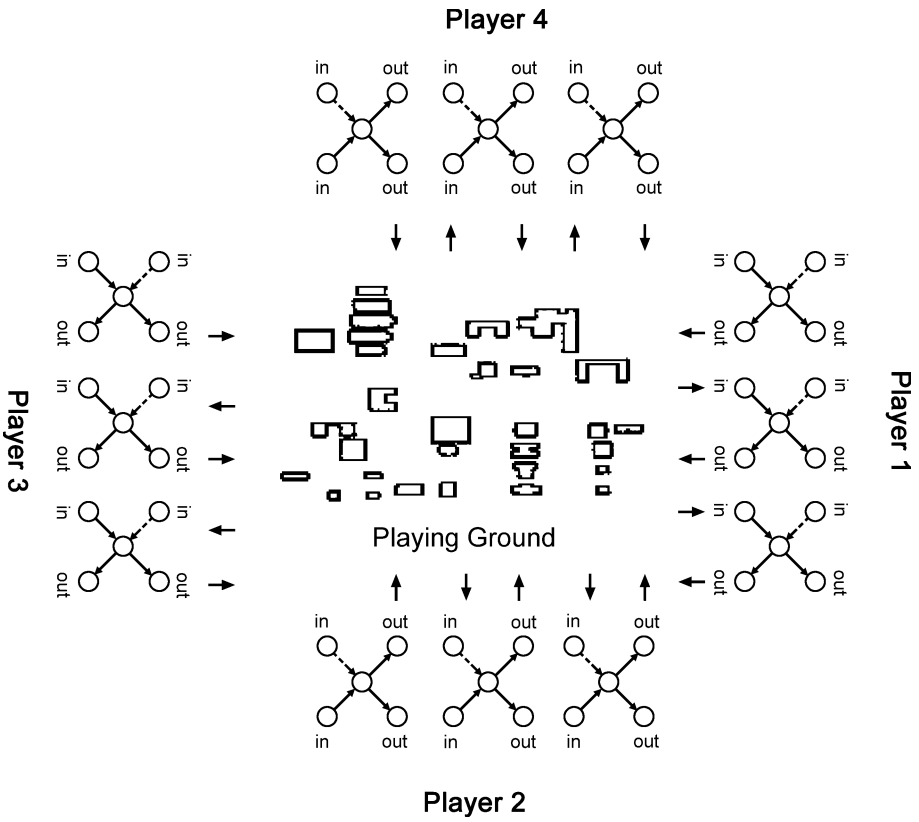
mation inflation and compression are thus two aspects of the process of *information adaptation* by which individuals and collectivities shape the city in a way that is adapted to the inhabitants' cognitive capabilities.

The notion information adaptation has far reaching implications to the above discussion about the scientific method and the science of cities: Self-organization as information compression implies a *complex* → *simple* model and thus an alternative Simon's *simple* → *complex* model. The process of information inflation, on the other hand, is in line with Simons' model. These two models are thus two facets of a single process of information adaptation that in some cases requires inflations while in others compression. Complexity theory shows that whatever are the opening conditions (complex or simple) a scientific approach is possible.

CogCity A central property of complex systems is the process of *circular causality* that typifies also the dynamics of cities: Thus in the SIRN model the interaction between the local/micro urban agents gives rise to the global structure of the city, which then feeds back and prescribes the behavior, interaction and action of the agents, and so on. Guided by Simon's *simple* → *complex* model, standard urban simulation models have become excellent tools to simulate the first part of this loop – the way local interactions give rise to a global structure – but they fail to describe the second, feedback part of the loop. CogCity (cognitive city) is a model that attempts to simulate the dynamic of cities as a process of circular causality [98].

CogCity is essentially a FACS model (above Sect. "AB and FACS cities") with several additions that make it an explicit SIRN, cognitive, urban simulation model. It differs from standard AB/CA urban simulation models in that the latter are essentially bottom-up in their structure (Fig. 30, *left*). CogCity, per contra, is characterized by an on-going interaction between top-down and bottom-up. Figure. 30, *right* describes a typical scenario: It starts top-down when an agent arrives to the city with a global cognitive map in mind; compares it to the global structure of the city and selects a local sub-area. Now starts the bottom-up process: the agent selects the empty cells in that local area; evaluates the appropriateness of the cells and their nearest neighbors in light of its needs and then takes a decision and action. In parallel, the properties of every cell are determined according to its relations to its nearest neighbors and so on.

In a regular AB/CA simulation the process ends here: the global outcome is recorded and mapped as the output



Self-Organization and the City, Figure 27b
A conceptualization of the city game

of this specific iteration and the model is ready for a new iteration. In a SIRN-CogCity model the process continues and feeds back to the global structure of the city that allows the top-down process in the next iteration: Firstly, the state of the various central places is determined. Secondly, peripheries are determined around central places. Thirdly, areas are defined or redefined. Fourthly, subareas are redefined. Fifthly, given areas and subareas, the global state of the city as a whole and its rank-size structure, is defined. The latter changes redefine the local membership state of each cell in the various infrastructure objects and become the externally represented input for a new agent in the next iteration, and so on in circular causality (Figs. 31, 32).

Planning

The link between self-organization and cities is contradictory. Firstly, since cities were always regarded as symbols of planned action – walls, roads, castles, fortresses, indicated a central authority that is capable of planning. Secondly, since planning as means to achieve a controlled order diametrically opposes self-organization as the spontaneous emergence of order. This is a seemingly contradiction, however, since cities are *dual* self-organizing systems with the implication that every urban agent is a planner at a certain scale. This view is supported by psychology and cognitive science that consider planning as one of the basic cognitive capabilities of humans [86].

Cognitive planning, that is, the ability to think, decide and act ahead, must be based on information about the future which by definition is partial and insufficient – a situation that according to Haken typifies also the process of pattern recognition as conceptualize by synergetics. Based on this analogy Haken [57] described decision situation in the context of planning as in Fig. 33.

This decision situation raises the question of 'How do people complement the unknown data?' According to Haken and Portugali, as in pattern recognition tasks here too, the unknown data is being supplied by means of associative memory [57,95], conceptual cognitive maps [98,99] and decision heuristics [95,127,128]. Table 1 specified several decision heuristics and their interpretation in the context of cities.

Self-Organization and the City, Table 1
Seven heuristics and their interpretation in the context of cities. For sources see [95]

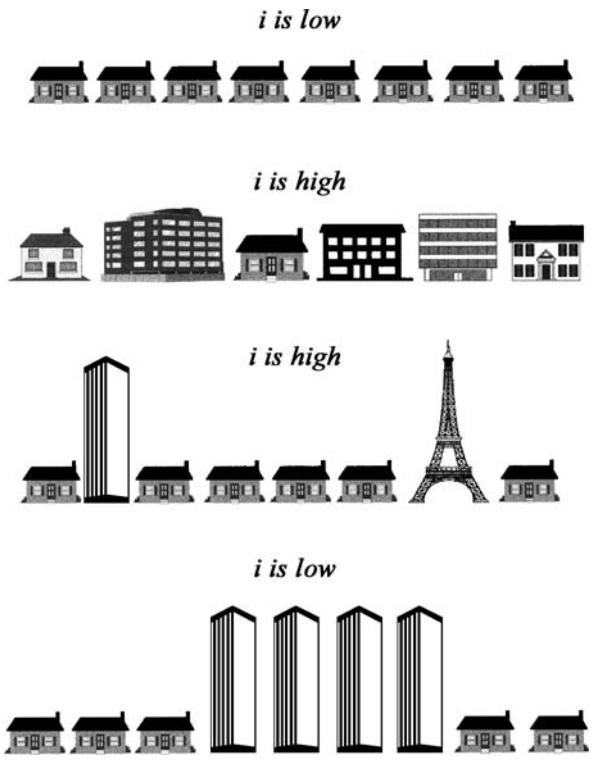
Heuristic	Description	Urban interpretation
Similarity	The similarity of two items is expressed as a function of their common and distinctive features.	The recognition of fundamental rules of urban composition, such as a grid layout, built form, or urban fabric, is performed through similarity.
Representativeness	The probability that an object or event belongs to a particular class is judged by the degree to which the description is representative of a stereotype.	Urban categories are identified on the basis of architectural stereotypes, such as church, skyscraper, boulevard, tower, park, arch, etc.
Availability	The probability of an event, or frequency of a class, is assessed by the ease with which instances or occurrences can be brought to mind, or recalled.	Availability would make universal symbols (such as Macdonald's signs, Stop signs, etc.) more easily identified and recalled.
Decision frame	The frame that a decision maker formulates of the problem (gain versus loss, etc.) is influenced by norms, habits, and personal characteristics of the decision maker.	Urban frames for decisions (congested versus free, public versus privet, etc.) depend on the cultural code of each agent (e. g. a tourist, a taxi driver, a policeman, etc.).
Anchoring	The tendency of people to make estimates by starting from an initial base value that is adjusted to yield the final answer.	City's internal representations can contain certain categories such as style, urban violence, town size, etc., which can be 'fired on' early in the process of cognition; once switched on, it stays on and is only eventually reprocessed.
Synergetic I: Collective effects	When facing complex decision situations people tend to rely on what other people are doing or saying.	Drivers, pedestrians, intra- and inter-urban immigrants, tend to 'follow the stream'. That is: to take decisions in line with what others are doing.
Synergetics II: Attention parameters effect	When facing complex decision situations people often employ several heuristics in a sequence. First, the attention parameter calls into use a heuristic. Then, when exhausted, another attention parameter heuristic emerges and so on.	Intra- and inter-urban immigrants, for example, often start with a given location decision heuristic (say synergetic I); if it doesn't work, they switch to an alternative heuristic and so on.

Haken and Portugali [95] have further suggested that complex processes of decision making in the context of city planning evolve according to their SIRN model. This suggestion was further elaborated by Portugali and Alfasi [105] who demonstrated empirically how this SIRN process practically takes place in the reality of *planning discourse* as it evolved among members of a planning team engaged in formulating urban policies concerning the development of the city of Beer Sheva, Israel (Fig. 34).

As a basic cognitive capability planning is intimately associated with the fact that humans are social creatures – people tend to plan together (e. g. families, friends, firms etc). Some planning decisions are thus made solitarily while others collectively. Planning is also a profession that is closely linked to the central authorities of society (municipal, regional, national governments etc.). We thus have three forms of planning – *solitary*, *collective* and *professional*.

The notion that cities are complex self-organizing systems thus implies a novel view on planning the essence of which is, first, that all three forms of planning (solitary, collective and professional) participate in the dynamic of cities. Second, that due to non-linearities that typify cities as complex systems, the act of a single solitary planner might affect the evolution of a city more than the planning act of a professional planning team. (For an example see [101]). Does that mean that due to self-organization there is no need for city planning? Not at all! – It means that we have to adopt a new perception of *plans as participants* in the overall urban dynamics. It also means that we have to adopt a new perception of urban dynamics as a complex interaction between many plans at different scales, or more specifically, between solitary, collective and professional, planning agents, each with its specific plan.

Can there be an administrative planning process that is built in line with the above? The answer suggested by Portugali and Alfasi is positive [105]: In a sequence of studies they have portrayed the principles of such self-planned city [95] and the way it can be applied to the reality of city planning law and structure of Israel [2]. Similarly to current planning systems it is a 3-layers system: the legislative, the judiciary and the executive. It differs in the following: First, its planning law refers to the qualitative relations between the various city objects (Fig. 35) and not to land use plans that assume to determine top-down the urban landscape. Second, it suggests a novel planning judicature



Self-Organization and the City, Figure 28
Different configurations and categorizations of buildings convey different quantities of Shannonian information. When all building are similar, information (*i*) is low. When they are different, *i is high* but difficult to memorize. When landmarks are added *i is high*, provided that they are located apart from each other; otherwise, *i is low*. Source: [60]

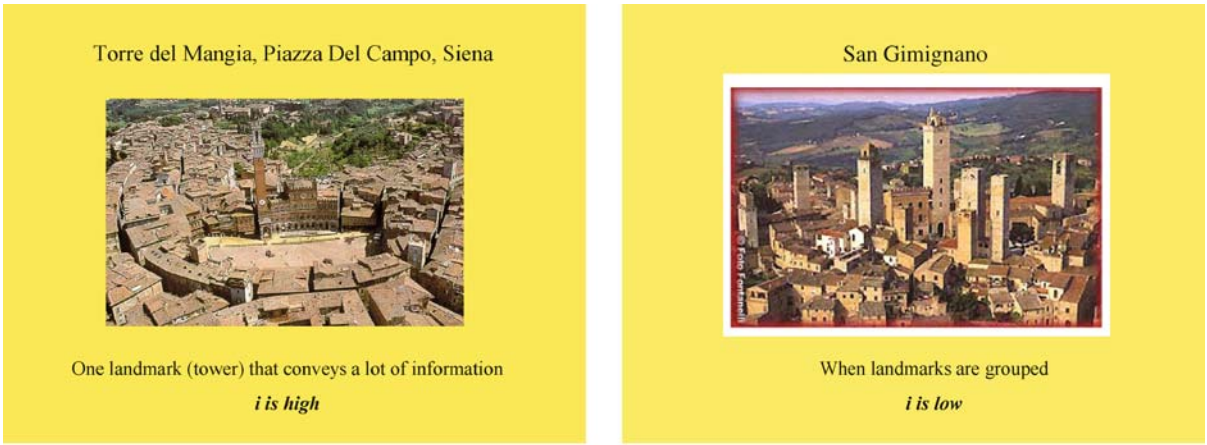
composed of spatially distributed "planning courts" conducted by professionals who have specialized in both law and planning. Their aim is to evaluate, accept or reject the plans proposed by all planning agents – solitary, collective or professional. Third, it suggests a separation of authorities that doesn't exist today in standard planning administrations. Fourth, it suggests a process of hermeneutic planning that enables phase transition and adaptation to new situations. Figure 36 describes the structure and operation of this self-organized planning system.

Prediction, Planning, Self-Organization and Cities
The first principle of the above planning system is that its planning laws refer to qualitative relations between the various city objects and not to land use plans. The reason is that land use plans are commonly based on predictions. This is problematic since prediction in the context of complex systems such as cities is associated with four fundamental properties. First, the nonlinearities that typify cities imply that one cannot establish predictive cause-effect relationships between some of the variables. Second, many of the triggers for change in complex systems have the nature of unpredictable mutations [4]; not because of lack of data, but because of their very nature. Third, unlike closed systems, in complex systems, the observer, with his/her predictions, is part of the system – a point made by Jantsch [71] more than two decades ago and largely ignored since then. In such a situation, predictions are essentially feed-forward loops, affecting the system and its future evolution with implications that include self-fulfilling and self-falsifying or self-defeating predictions [101].

From the above follows a dilemma: complex systems are in essence unpredictable and yet, the current practice of planning as well as planning administration and law are based on the ability to predict. In a recent paper (ibid) it was shown that this situation leads to planning paradoxes that are the result of phenomena of self-fulfilling and self-falsifying predictions. It was further shown that these phenomena are the result of the *feed-forward* and *feedback loops* that are typical of complex systems in general and of cities and regions in particular. The existence of such loops is one of the properties that make systems complex. Such loops are responsible to the situation by which a prediction or a plan, once produced, becomes a *participant* in the system's dynamic.

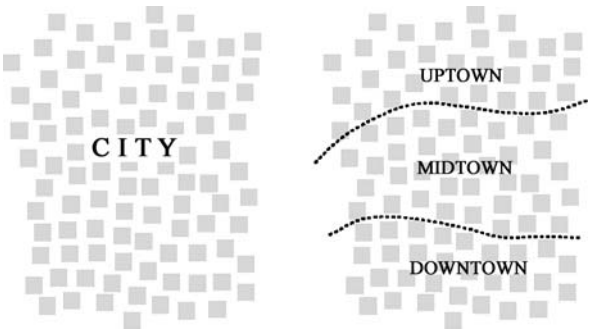
Another way to look at this issue is from the point of view of the distinction between Shannonian and semantic information (above, Sect "Information Compression, Inflation, Adaptation"): Predictions and plans are essentially kinds of information transmission. One can thus speak of Shannonian prediction and semantic prediction. In the first, the outcome of the prediction is independent of the receiver(s) while in the second it depends on the meaning attached to it by a receiver or receivers. A weather forecast is a good example for both: it has no effect on the climatic system, but it might affect the urban system – following the prediction people might behave in different ways that might entail phenomena of self-falsifying and self-fulfilling predictions as described above.

Planning theory has not as yet internalized the implications of complexity theory to city planning. For example, in the planning and decision support systems (PSS, DSS) that are currently discussed and built by proponents of the complexity paradigm, urban simulation models are assumed to function as sophisticated prediction devises [28,29,43]. The result is a discrepancy that to my mind characterizes the domain of urban and regional planning: On the one hand, planning theory, as well as the structure of planning law, practice and administration, are all based on the (usually implicit) assumption that



a b

Self-Organization and the City, Figure 28a, b
a An example of a good landmark. b An example of not very effective landmark



Self-Organization and the City, Figure 29
If all buildings in the city are different from each other, categorization entails information compression; if they are similar – information inflation

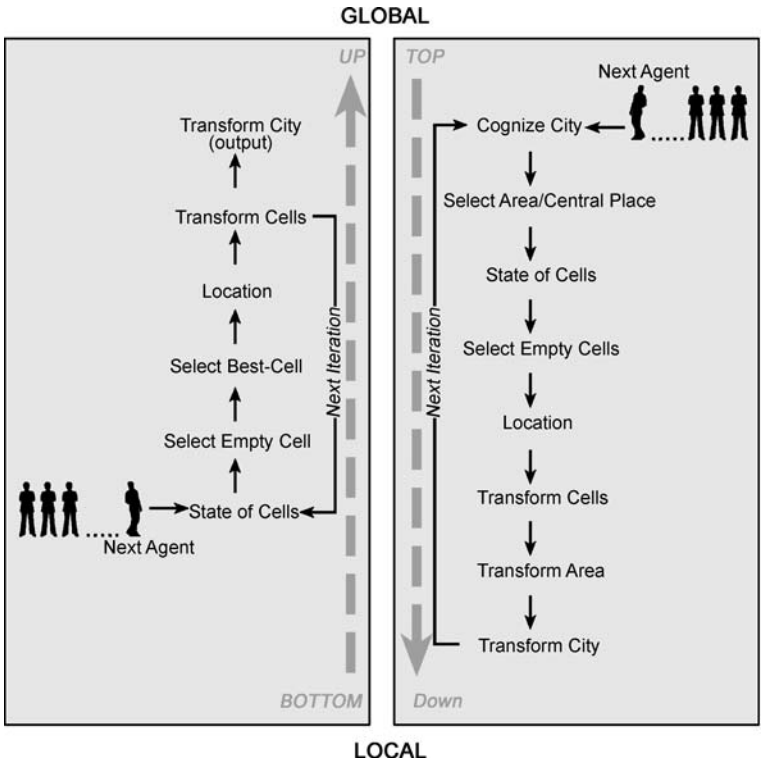
cities are essentially predictable entities; that given sufficient data, information and models, their future behavior is in essence predictable. On the other hand, current urban theory suggests that cities are complex, self-organizing and non-linear systems and that as a consequence their future behavior is in essence not predictable; even if sufficient information and data is collected and available [95].

Urbanism

In Sect. "Explicit Attempts to Define a City" we've defined the category "city" in terms of a family resemblance according to which a settlement becomes a "city" not by having some necessary and sufficient properties, but by being a member in a network of cities that has center, periphery etc. This view is in line with cognitive science's approach to concepts and categories. In the latter it is common also to distinguish between *basic level* categories and *super-ordinate* categories [116]. A 'chair', for instance, is a basic level category whereas 'furniture' a super-ordinate. The suggestion here is that a *city* is a basic level category whereas *urbanism* is a super-ordinate one, referring to the totality of cities ranging from their physical structure, architecture, economics, politics, social and cultural composition, and so on. The term 'urban revolution' (coined by Childe, [35]) thus implies a major transformation in society with the basic level category 'city' at its center.

Most complexity studies of cities have traditionally focused on specific aspects of cities – land-use, morphology, transportation, social segregation etc. – but not on the totality of city life which is what urbanism is all about. Why? Because they evolved mainly out of regional sciences' attempt to develop a scientific approach to cities and the consequent tendency to choose research issues that can be analyzed by reference to 'real world' data. The study of urbanism was thus left to the "soft" social theory approaches to cities[33,34,63,64].

This is rather unfortunate because 21st century world society is undergoing a major transformation with urbanism at its center: Massive rural-urban migration and demographic processes entailed a situation by which cities such as Mexico City, Bombay (Mumbai) and Sao Paolo grew from 8.8 million, 6.2 million and 8.3 million respectively in 1970 to over 20 million, over 16 million and again over 18 million today; for the first time in human history, the number of people living in cities is crossing 50% of the world's population and the process is still on. In the last



Self-Organization and the City, Figure 30
A cognitive (*right*) vs non-cognitive (*left*) AB/CA urban simulation model. Source: [98]

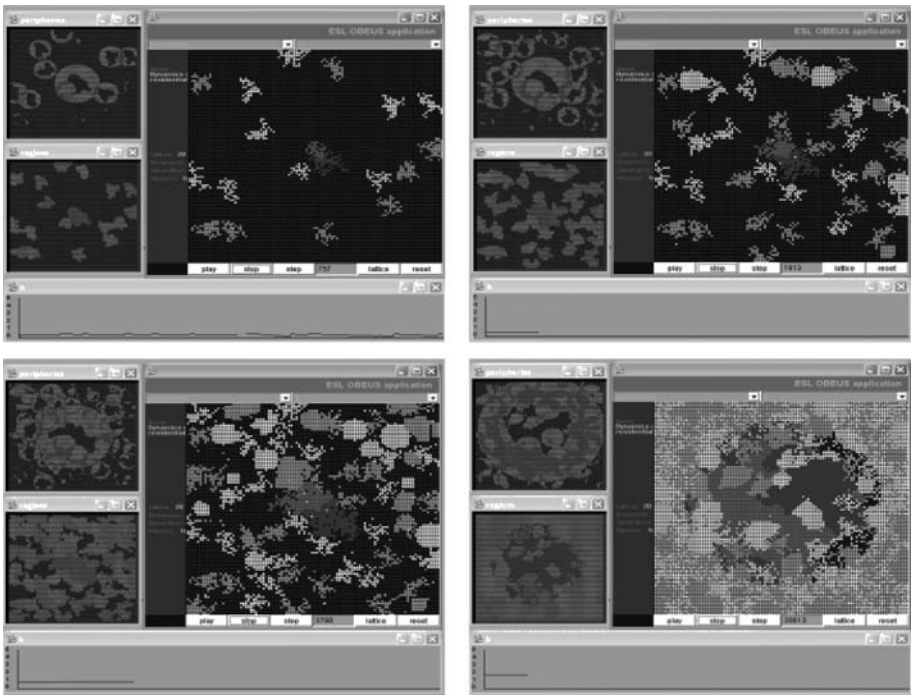
few decades we've witnessed the emergence of *world cities*, or *global cities*, that form the centers for the globalization process.

These quantitative processes are associated with several qualitative processes: a process of privatization that leads to the decline of the welfare nation-state; the emergence of a civil society that takes over many of the past duties and functions of the nationalist welfare state; the crucial problems of many (post)modern counties are no longer classical national problems (e. g. national self-determination, national boundaries etc.), but rather the problems of cities. The events of September 11 and the ensuing wars in Afganistan and the Middle East are tragic indications to the *urbanization of war*. Finally, the process of globalization is making some world cities more dominant than the states within which they exist thus repressing the nation states.

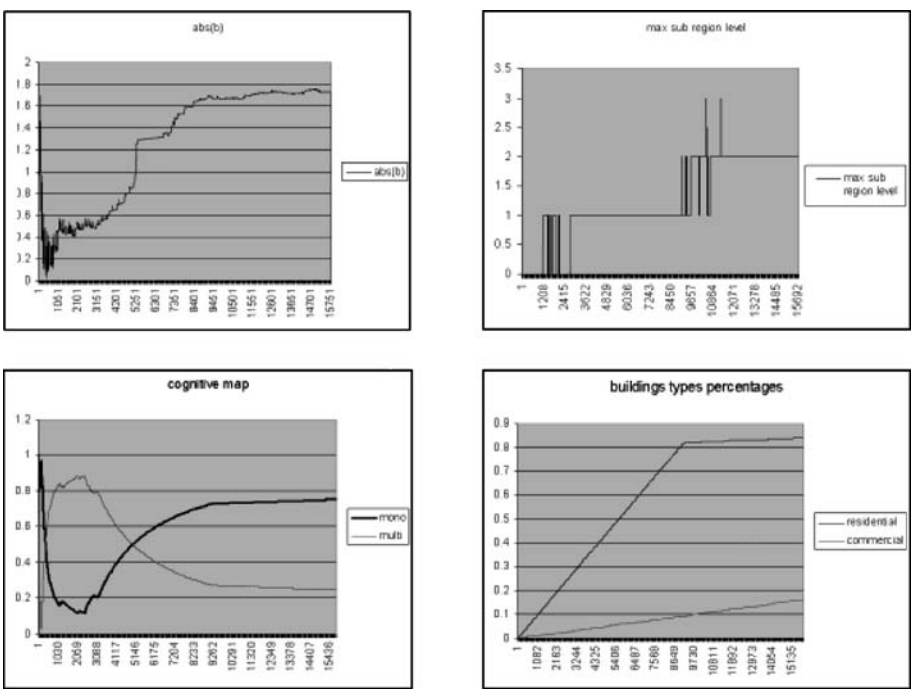
All the above indicates the more fundamental change: According to Lefebvre [79], its essence is that urbanism is replacing industrialization as the dominant force in society. My view is that the essence of this change is that urbanism is challenging nationalism as the order parameter of modern society [101].

Complexity and Urbanism Complexity studies of cities, with their focus on the short-term dynamics of cities and of national systems of cities are indeed highly advanced in terms of mathematical formalism and data analysis but rather anachronistic in terms of the issues studied; as such they have so far said very little on the dramatic urban phenomena of 21st century. Can they say more about the issue of urbanism? The answer is yes! And for several reasons: To my mind the "deeper messages" of complexity theories is that they have discovered properties in matter hitherto assigned to life, art and society [91]. It is not surprising therefore that complexity theories, particularly synergetics, bear many similarities to social theory and philosophy and, as a consequence, several of the notions that originated in the study of complex systems can be related to similar notions that originated in the domain of social theory [100]:

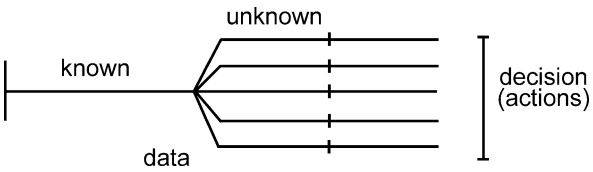
- Both are essentially systemic and even holistic.
- Both tend to conceptualize 'development' and 'evolution' in terms of abrupt changes rather then a smooth progression. In social theory the common terms for an abrupt change is (social/political/cultural) 'revolution',



Self-Organization and the City, Figure 31
Preliminary results from an evolving scenario simulated by CogCity: The central screen in each of the four snapshots shows the evolving spatial distribution of various kinds of agents, the top left screen the evolution of centers and sub-centers, while the bottom left, the evolving cognitive maps of agents. Source: [98]



Self-Organization and the City, Figure 32
Preliminary results from an evolving scenario simulated by CogCity – the graphs. Source: [98]



Self-Organization and the City, Figure 33
Decisions in the city are characterized by insufficient data. In such a reality the known data may be complemented in a variety of ways. Each of theses ways might entail a different decision and action. Source: [57,95]

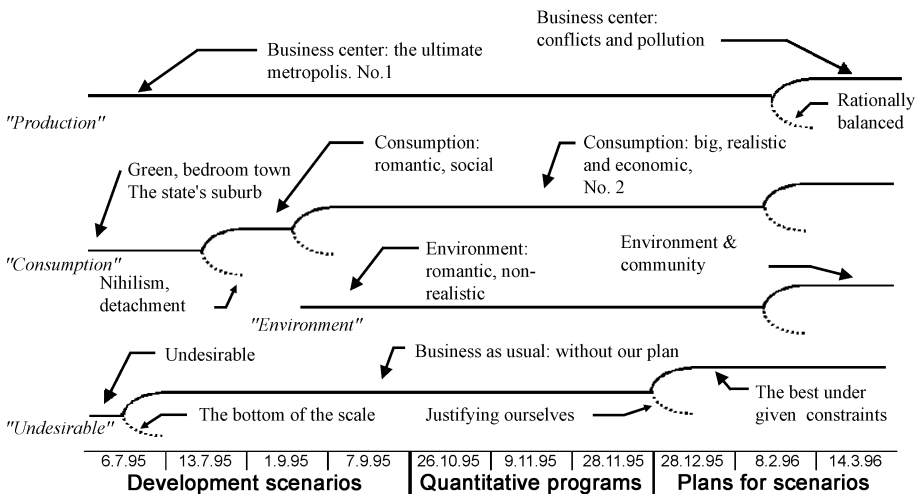
while in the language of complexity 'bifurcations' and 'phase transitions' (that reminds one of Gould's and Eldredge's, [46] *punctuated equilibrium*).

- Synergetics' notion of 'order parameter' is similar to social theory's notion of 'mode of production'.
- Synergetics' notions of 'enslavement' and 'circular causality' are close to social theory's notions of 'social reproduction' and 'socio-spatial reproduction' [44,80].
- Complexity's view of systems in 'a far from equilibrium condition' comes close to postmodernism's recent emphasis on viewing reality as ever changing and transforming; hence the general popularity of notions such as 'chaos' and 'butterfly effect'.

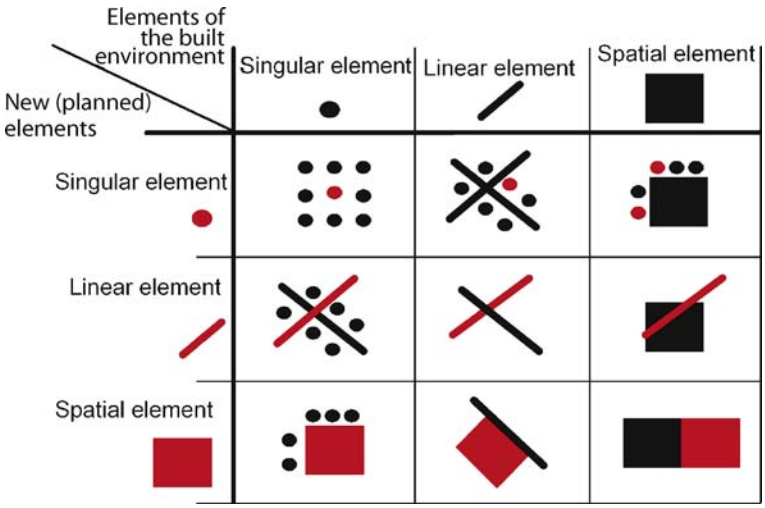
Several writers have already responded to these similarities from the perspective of the sciences, philosophy, media/cultural critics and modern and postmodern social theory [37,75,76,100,113]. A preliminary attempt has also been made to employ synergetics as a complexity theory of urbanism [95,101]. That is, to interpret the current changes in cities and urbanism in terms of synergetics along the following scenario: the combined force of rapid population growth, urban expansion and technological change throughout the 20th century acted as a control parameter. Toward the end of the 20th century and at the beginning of the 21st, we are witnessing a bifurcation and phase transition followed by a competition between the newly emerging *urban order parameter* and the old nationalist one. My personal view is that what we see emerging today out of this competition is not the replacement of nationalism as an order parameter by urbanism as an order parameter, but the *urbanization of nationalism*.

Future Directions

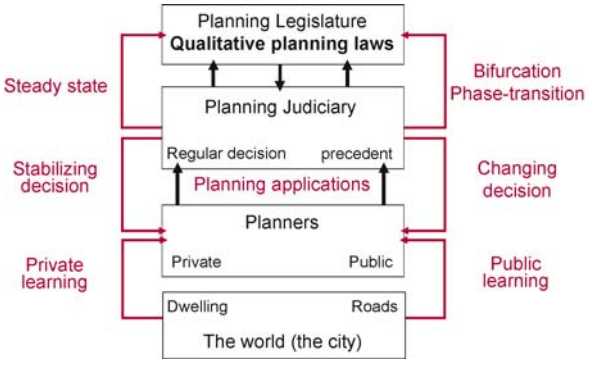
Looking in retrospect at more than two and a half decades of complexity theory studies of cities, one can now appreciate some of its major achievements: First, the link between cities and complexity theory gave urban studies a strong theoretical basis it never had before. The fact that complexity theory was applied to a large number of domains gave urban studies a wide context and many sources of inspiration. The fact that complexity theory comes with a rich and strong mathematical formalism gave urban studies a sound methodological background. The attempt



Self-Organization and the City, Figure 34
Bifurcation diagram of the planning discourse. Each alternative is represented with a continuous line along the time axis (the X-axis). A horizontal line represents an order state during which the alternative scenario maintains a certain image and possesses certain attributes. Bifurcation points indicate a shift from one order state to another. The broken lines represent optional order states that were not actualized. Source: [105]



Self-Organization and the City, Figure 35
Proposition for a new planning structure for a self-planned city: The interrelations between urban elements provide the basis for planning law. A singular urban element might be a building, a linear one might be a road, while an example for a spatial urban element is a park



Self-Organization and the City, Figure 36
Proposition for a new planning structure for a self-planned city: The system is built of private planners (the inhabitants of the city) and professional planners. Each of them might submit a plan to the planning judiciary. In the latter the "planning judge" takes decisions according to the planning law as determined by the planning legislature. Unlike the current structure, there is a clear separation of authorities

to transform the study of cities into a science of cities is today closer than ever.

Complexity theory has given us a new insight to our understanding of the dynamics of cities. According to Batty [16] the most important contribution is that complexity studies of cities have verified the intuitive views of Jean Jacobs [70] and Alexander [1], namely, that the complex entity "city", with its variety of different land-uses, socio-spatially and culturally segregated communities, transportation networks and all the rest, is an outcome of "bottom-up" processes: The local interaction between agents at local scale, conducted by very few and simple rules gives rise the complexity we term 'city'.

At the same time, however, it must be admitted that the potential contribution of complexity theories to urbanism, planning and urban design has yet to be realized. The complexity approach has indeed given the bottom-up views on the nature of cities a strong mathematical formalism that can be quantified by real data. But this focus on the local, the bottom-up and the quantifiable was not without price: Cities and urbanism of the 21st century are in the midst of a dramatic transformation, new forms of cities are emerging – world cities, global cities, megacities, and yet the vast majority of complexity studies still focus on the old traditional quantitative urban questions, leaving the qualitative grand urban issues to the "non-scientific" social theory oriented urban studies. Can complexity theories of cities contribute? The answer suggested above is yes! SIRN is one approach in this direction and the field is ripe for others.

The same applies to planning. In the literature on planning theory it is common to make a distinction between *planning theory* versus *theory in planning* [40]. That is, between a theories about how to plan and theories about urban and regional dynamics that planners can use during the planning process. Examining complexity theories of cities from this perspective we see that they are very innovative with respect to theory in planning, but very conservative when it comes to theory of planning: The vast

majority of studies simply ignore the implications of complexity to urban, regional and environmental planning.

Why do we need a complexity theory of planning? The answer is twofold: First, standard planning theory was developed in the 1950s and 1960s hand in hand with what we consider today as anachronistic urban theory: Both are based on the (usually implicit) assumption that cities are in essence simple, mechanistic systems that given sufficient data and advanced technologies, their future behavior is predictable and hence controllable. As we've seen above, complexity theories tell us a different story: Cities are complex self-organized systems that are in essence unpredictable and controllable; even if sufficient data and the most advanced technologies are at hand. From here follow a whole set of new and interesting questions: what is the role of planning in a complex system? Are all parts and components of the system unpredictable? In the above we've suggested some preliminary answers – but new ones must still come.

Finally it is important to mention the issue of extreme events in cities. The rapid processes of urbanization cities underwent in the last few decades and the 'urbanization of war' made cities rather vulnerable areas in cases of extreme events. The question of how cities and their inhabitants behave and respond to extreme events, is a pressing social issue that already started to capture the attention of students of complexity theory of cities and urbanism.

Bibliography

1. Alexander C (1965) A city is not a tree. Architectural Forum April-May, 58–62
2. Alfasi N, Portugali J (2007) Planning rules for a self-planned city. Planning Theory 6(2):164–182
3. Allen PM, Strathern M (2004) Complexity: The integrated framework for integrated models of urban and regional systems. A talk delivered at a conference on The Dynamics of Complex Urban Systems, November 4–6, 2004, Monte Verita
4. Allen PM (1997) Cities and Regions As Self-Organizing Systems: Model of Complexity. Routledge, London
5. Allen PA (1981) The evolutionary paradigm of dissipative structures. In: Jantsch E (ed) The Evolutionary Vision. Westview Press, Boulder, pp 25–71
6. Allen P, Sanglier M (1981) Urban evolution, self-organization and decision making. Environ Planning A 13:169–183
7. Allen P, Sanglier M, Engelen G, Boon F (1985) Towards a new synthesis in the modeling of evolving complex systems. Environ Planning B Planning Des 12:65–84
8. Alonso W (1965) Location and Land Use. Harvard University Press, Cambridge, MA
9. Andersson C, Hellervik A, Lindgren K, Hagson A, Tornberg J (2003) Urban economy as a scale-free network. Phys Rev E 68(3):036124
10. Auerbach F (1913) Das Gesetz der Bevolkerungskoncentration. Petermanns Geogr Mitt 59:74–76
11. Bak P, Chen K, Creutz M (1989) Self-organized criticality in the game of life. Nature 342:780–782
12. Bak P, Chen K (1991) Self-organized criticality. Sci Am 28: 26–33
13. Barabási A-L (2002) Linked: The new science of networks. Perseus, Cambridge
14. Barabási A-L, Réka A (1999) Emergence of scaling in random networks. Science 286:509–512
15. Bartlett FC (1932/1961) Remembering: A Study in Experimental and Social Psychology. Cambridge University Press, Cambridge
16. Batty M (2005) Cities and Complexity: Understanding Cities with Cellular Automata, Agent Based Models and Fractals. MIT Press, Cambridge, Mass
17. Batty M (1997) Cellular automata and urban form: a primer. J Am Planning Assoc 63(2):266–274
18. Batty M (1996) Urban Evolution on the Desktop: Simulations Using Extended Cellular Automata. Unpublished paper, Centre for Academic Spatial Analysis, UCL
19. Batty M, Longley P (1994) Fractal Cities. Academic Press, London
20. Batty M, Xie Y (1999) Self-organized criticality and urban development. In: Portugali J (ed) Population, Environment and Society on the Verge of the 21st Century, pp 109–124
21. Batty M, Xie Y (1994) From cells to cities. Environ Planning B: Planning Des 21:531–548
22. Benenson I, Torrens PM (2004) Geosimulation: Automata based modeling for urban phenomena. Wiley, London
23. Benguigui L (1995) A fractal analysis of the public transportation system of Paris. Environ Planning A 27:1147–1161
24. Benguigui L, Czamanski D, Marinov M, Portugali J (2000) When and where is a city fractal. Environ Planning B 27(4): 507–519
25. Blanchard P, Volchenkov D (2007) Scale-free Segregation in Transport Networks. arXiv:0710.1592
26. Braudel F (1993) A History of Civilizations, translated by Richard Mayne. Penguin Books, New York
27. Blumenfeld-Lieberthal E (2005) Dynamics of Urban Morphology in the Tel-Aviv Metropolitan Area. Ph.D thesis, Technion, Haifa
28. Brail RK (2006) Planning support systems evolving: When the rubber hits the road. In: Portugali J (ed) Complex Artificial Environments, Complexity series. Springer, Heidelberg, pp 307–317
29. Brail RK, Klosterman RE (eds) (2001) Planning Support Systems. ESRI Press, New York
30. Burgess B (1925) The growth of the city: An introduction to a research project. In: Park R, Burgess EW, McKenzie RD (eds) The City. Chicago University Press, Chicago, pp. 47–62
31. Burgess EW (1926/1968) The Urban Community. Greenwood Press, New York
32. Burgess EW (1927) The determination of gradients in the growth of the city. American Sociological Society Publications 21:178–184
33. Castells M (1977) The Urban Question. MIT Press, Cambridge, MA
34. Castells M (1996) The Rise of the Network Society. Blackwell Publishers, Malden, MA
35. Childe VG (1950) The urban revolution. Town Planning Rev 21:3–17

36. Christaller W (1933/1966) Central Places in Southern Germany. Prentice Hall, Englewood Cliffs, NJ
37. Cilliers P (1998) Complexity and Postmodernism. Routledge, London
38. Dalton N, Peponis J, Conroy-Dalton R (2003) To tame a TIGER one has to know its nature: extending weighted angular integration analysis to the description of GIS road-centerline data for large scale urban analysis. In: Hanson J (ed) Proceedings of the 4th International Space Syntax Symposium. University College London, London, pp 65.1–65.10
39. Dendrinos DS, Sonis M (1990) Chaos and Socio–Spatial Dynamics. Springer, New York, Berlin
40. Faludi A (1973) A Reader in Planning Theory. Pergamon Press, Oxford
41. Figueiredo L, Amorim L (2005) Continuity lines in the axial system. In: Van Nes A (ed) 5th International Space Syntax Symposium. TU Delft, Faculty of Architecture, Section of Urban Renewal and Management, Delft, pp 161–174
42. Garling T, Golledge RG (eds) (1993) Behavior and Environment. North-Holland, Amsterdam, London, New York, Tokyo
43. Geertman S, Stillwell J (eds) (2003) Planning Support Systems in Practice. Springer, Heidelberg
44. Giddens A (1984) The Constitution of Society: Outline of the Theory of Structuration. University of California Press, Berkeley
45. Golledge RG, Timmermans H (eds) (1988) Behavioral Modeling in Geography and Planning. Croom Helm, London, New York, Sydney
46. Gould SJ (1980) The Panda's Thumb. Norton, New York
47. Gregory D (1994) Geographical Imaginations. Blackwell, Cambridge, MA
48. Haag G, Muntz M, Pumain D, Saint-Julien T, Sanders L (1992) Interurban migration and the dynamics of a system of cities. Environment and Planning A 24:181–198
49. Haken H (1979) Pattern formation and pattern recognition – an attempt at a synthesis. In: Haken H (ed) Pattern Formation by Dynamical Systems and Pattern Recognition. Springer, Berlin, pp 2–13
50. Haken H (1983) Synergetics, an Introduction, 3rd edn. Springer, Berlin
51. Haken H (1987) Advanced Synergetics: An Introduction (2nd print). Springer, Berlin, New York
52. Haken H (1988/2000) Information and Self-Organization: A Macroscopic Approach to Complex Systems. Springer, Berlin
53. Haken H (1990) Synergetics of Cognition. Springer, Berlin
54. Haken H (1991/2004) Synergetic Computers and Cognition. Springer, Berlin
55. Haken H (1993) Synergetics as a strategy to cope with complex systems. In: Haken H, Mikhailov A (eds) Interdisciplinary Approaches to Non-Linear Complex Systems. Springer, Berlin
56. Haken H (1996) Principles of Brain Functioning: A Synergetic Approach to Brain Activity, Behavior and Cognition. Springer, Berlin
57. Haken H (1998) Decision making and optimization in regional planning. In: Beckmann MJ, Johansson B, Snickars F, Thord R (eds) Knowledge and Networks in a Dynamic Economy. Springer, Berlin
58. Haken H, Portugali J (1995) A synergetic approach to the self-organization of cities. Environ Planning B: Planning Des 22:35–46
59. Haken H, Portugali J (1996) Synergetics, inter-representation networks and cognitive maps. In: Portugali J (ed) The Construction of Cognitive Maps. Kluwer, Dordrecht, pp 45–67
60. Haken H, Portugali J (2003) The face of the city is its information. Environ Psychol 23:385–408
61. Haken H, Portugali J (in preparation) Synergetics and captivity in a city
62. Haken H, Portugali J (in preparation) Information adaptation
63. Harvey D (1985) The Urbanization of Capital. Basil Blackwell, Oxford
64. Harvey D (1996) Justice, Nature and Geography of differences. Basil Blackwell, Oxford
65. Hillier B (1999) Space is the Machine: A Configurational Theory of Architecture. Cambridge University Press, Cambridge
66. Hillier B, Hanson J (1984) The social logic of space. Cambridge University Press, Cambridge
67. Hillier B, Iida S (2005) Network effects and psychological effects: a theory of urban movement. In: Cohn A, Mark D (eds) Spatial Information Theory. Lecture Notes in Computer Science, vol 3603. Springer, Berlin, pp 473–490
68. Hu MB, Wang WX, Jiang R, Wu QS, Wang BH, Wu YH, (2006) Urban Traffic Dynamics: A Scale-Free Network Perspective, http://arxiv.org/abs/physics/0606086v1
69. Isard W (1956) Location and Space Economy. Published jointly by the Technology Press of Massachusetts Institute of Technology and Willey, New York
70. Jacobs J (1961) The Death and Life of Great American Cities. Penguin Books, England
71. Jantsch E (ed) (1981) The Evolutionary Vision. Westview Press, Boulder, CO
72. Jiang B (2007) A topological pattern of urban street networks: Universality and peculiarity. Physica A 384:647–655
73. Jian B (2006) Small world modeling for complex geographic environments. In: Portugali J (ed) Complex Artificial Environments: simulation, cognition and VR in the study and planning of cities. Springer, Berlin
74. Johnson M (1987) The body in the mind: The bodily basis of meaning, imagination, and reason. The University of Chicago Press, Chicago
75. Johnson (2001) media/cultural critics – postmodern complexity
76. Kellert SH (1993) In the Wake of Chaos. University of Chicago Press, Chicago
77. Kitchin R, Freundschuh S (2000) Cognitive Mapping: Past, Present and Future. Outledge, London
78. Lakoff G (1987) Women, Fire and Dangerous Things: What Categories Reveal About the Mind. The University of Chicago Press, Chicago, London
79. Lefebvre H (1970) La Révolution Urbaine. Gallimard, Paris
80. Lefebvre H (1974) The Production of Space. English translation, 1995. Blackwell, Oxford
81. Lösch A (1954) The Economics of Location. Yale University Press, New Haven
82. Lynch K (1960) The Image of the City. MIT Press, Cambridge, MA
83. Mandelbrot BB (1983) The Fractal Geometry of Nature. Freeman, San Francisco
84. Milgram S (1967) The Small World Problem. Psychol Today May:60
85. Miller GA (1956) The magic number seven, plus or minus two:

Some limits on our capacity for processing information. Psychol Rev 63(2):81–97
86. Morris R, Ward G (eds) (2005) The cognitive psychology of planning. Psychology Press, Hove, East Sussex, New York
87. Nicolis G, Prigogine I (1977) Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order Through Fluctuations. Wiley-Interscience, New York
88. Park RE (1925/1967) The City. Chicago University Press, Chicago
89. Porta S, Crucitti P, Latora V (2005) The Network Analysis of Urban Streets: A Primal Approach, http://arxiv.org/abs/physics/0506009
90. Popper KR (1959) The Logic of Scientific Discovery. Hutchinson, London
91. Portugali J (1985) Parallel currents in the natural and social Sciences. In: Portugali J (ed) Links Between Natural and Social Sciences. A special theme issue. Geoforum 16(2), pp 227–238
92. Portugali J (1990) Preliminary notes on social synergetics, cognitive maps and environmental recognition. In: Haken H, Stadler M (eds) Synergetics of Cognition. Springer, Berlin, pp 379–392
93. Portugali J (1994) Theoretical speculations on the transition from nomadism to monarchy. In: Finkelstein I, Na'aman N (eds) From Nomadism to Monarchy. Yad Izhak Ben-zvi, Jerusalem, pp 203–217
94. Portugali J (1996) Inter-representation networks and cognitive maps. In: Portugali J (ed) The Construction of Cognitive Maps. Kluwer, Dordrecht, Boston, London, pp 11–43
95. Portugali J (2000) Self-Organization and the City. Springer, Heidelberg
96. Portugali J (2002) The Seven Basic Propositions of SIRN (Synergetic Inter-Representation Networks). Nonlinear Phenom Complex Syst 5(4):428–444
97. Portugali J (2003) SIRN (Synergetic Inter-Representation Networks), Artifacts and Snow's two cultures. In: Tschacher W, Dauwalder J-P (eds) Dynamical System Approaches to Embodied Cognition. World Scientific, Singapore, pp 277–294
98. Portugali J (2004) Toward a cognitive approach to urban dynamics. Environ Planning B: Planning Des 31:589–613
99. Portugali J (2005) Cognitive Maps Are over 60. In: Cohn AG, Mark DM (eds) COSIT 2005. LNCS, vol 3693. Springer, Berlin, pp 251–264
100. Portugali J (2006) Complexity theory as a link between space and place. Environ Planning A 38:647–664
101. Portugali J (2006) The scope of complex artificial environments. In: Portugali J (ed) Complex Artificial Environments: Simulation, cognition and VR in the study and planning of cities. Springer, Heidelberg
102. Portugali J (ed) (1992) Geography, Environment and Cognition. A special theme issue. Geoforum 23(2)
103. Portugali J (ed) (1996) The Construction of Cognitive Maps. Kluwer, Dordrecht
104. Portugali J (ed) (2006) Complex Artificial Environments: Simulation, cognition and VR in the study and planning of cities. Springer, Heidelberg
105. Portugali J, Alfasi N (2008) An approach to planning discourse analysis. Urban Studies, Forthcoming
106. Portugali J, Benenson I, Omer I (1994) Socio–spatial residential dynamics: stability and instability within a self-organizing city. Geograph Anal 26(4):321–340
107. Portugali J, Benenson I, Omer I (1997) Spatial cognitive dissonance and sociospatial emergence in a self-organizing city. Environ Planning B: Planning Des 27:263–285
108. Portugali J, Haken H (1992) Synergetics and cognitive maps. In: Portugali J (ed) Geography, Environment and Cognition. A special issue. Geoforum 23(2), pp 111–130
109. Prigogine I (1980) From Being to Becoming. Freeman & Co, San Francisco, CA
110. Prigogine I, Stengers I (1984) Order Out of Chaos. Bantam, New York
111. Pumain D (ed) (2005) Hierarchy in Natural and Social Sciences, Methodos Series (Hardcover). Springer, Heidelberg
112. Pumain D, Saint-Julien T, Sanders L (1987) Application of a dynamic urban model. Geograph Anal 19(2):152–166
113. Rasch W, Wolfe C (2000) Observing Complexity. University of Mennesota Press, Minneapolis
114. Ratti C (2004) Space syntax: some inconsistencies. Environ Planning B Planning Des 31:487–499
115. Rosch E (1999) Reclaiming concepts. J Consciousness Stud 6(11–12):61–77
116. Rosch E, Mervis C, Gray W, Johnson D, Boyes-Braem P (1976) Basic objects in natural categories. Cogn Psychol 8:382–439
117. Salingaros N (2006) A Theory of Architecture. Umbau-Verlag, Solingen, Germany
118. Salingaros N (2005) Principles of Urban Structure. Techne Press, Amsterdam, Holland
119. Sanders L (1992) Systems de Villes et Synergetique. Anthropos, Paris
120. Sanders L, Pumain D, Mathian H, Guerin-Pace F, Bura S (1997) SIMPOP: a multiagent system for the study of urbanism. Environ Planning B Planning Des 24:287–306
121. Sanglier M, Allen P (1989) Evolutionary models of urban systems: an application to the Belgian provinces. Environ Planning A 21:477–498
122. Schelling T (1974) On the ecology of micro-motives. In: Marris R (ed) The Corporate Society. Macmillan, London
123. Shannon CE, Weaver W (1959/1963) The Mathematical Theory of Communication. University of Illinois Perss, Illinois
124. Simon H (1979) The Science of the Artificial. MIT Press, Cambridge, MA
125. Snow CP (1964) The Two Cultures and a Second Look. Cambridge University Press, Cambridge
126. Stadler M, Kruse P (1990) The self organization perspective in cognition research: historical remarks and new experimental approaches. In: Haken H, Stadler M (eds) Synergetics of Cognition. Springer, Berlin, pp 32–52
127. Tversky A, Kahneman D (1974) Judgement under uncertainty: Heuristics and biases. Science 185:1124–1131
128. Tversky A, Kahneman D (1981) The framing of decision and psychology of choice. Science 211:4538
129. von Thünen JH (1826/1966) von Thünen's Isolated State. An English translation (Hall P (ed)) Pergamon, Oxford
130. Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. Nature 393:440–442
131. Weidlich W, Haag G (1983) Concepts and Models of a Quantitative Sociology. Springer, Berlin, New York
132. Weidlich W (1987) Synergetics and social science. In: Graham R, Wunderlin A (eds) Lasers and Synergetics. Springer, Berlin, pp 238–256
133. Weidlich W (1994) Synergetic modeling concepts for sociodynamics with application to collective political opinion formation. J Math Sociol 1(4):267–291

134. Weidlich W (1999) From fast to slow processes in the evolution of urban and regional settlement structures: the role of population pressure. In: Portugali J (ed) Population, Environment and Society on the Verge of the 21st Century. A special theme issue, Discrete Dynamics in Nature and Society, pp 137–147
135. Wirth L (1938) Urbanism as a Way of Life. Am J Sociol 64: 1–24
136. Wittgenstein L (1953) Philosophical Investigations. Blackwell, Oxford. Translated by G. E. M. Anscombe
137. Wu J, Gao Z, Sun H, Huang H (2004) Urban Transit System as a Scale-Free Network. Mod Phys Lett B 18(19–20):1043–1049
138. Zipf GK (1949) Human Behaviour and the Principle of Least-Effort. Addison-Wesley, Cambridge

Self-Organization in Clinical Psychology

Günter Schiepek[1], Volker Perlitz[2]
[1] Institute of Synergetics and Psychotherapy Research, Paracelsus Medical University, Salzburg, Austria
[2] Klinik für Psychosomatik, Universitätsklinikum der RWTH Aachen, Aachen, Germany

Article Outline

Definition of the Subject
Introduction
Dynamic Diseases
Self-Organized Synchronization Patterns in Peripheral Physiological Systems
Nonlinear Dynamics in the Communication of Patient and Therapist
Self-Organization in Human Change Processes
The Concept of Self-Organization Promotes New Information Technologies in Clinical Psychology – The Synergetic Navigation System
The Self-Organizing Brain
Future Directions
Bibliography

Definition of the Subject

Clinical psychology is a sub-discipline of psychology engaged in the description, classification, explanation, and treatment of mental disorders. The primary focus is on psychological methods, models, and topics such as behavior, cognition, emotion, and social interaction with substantial overlap with related areas in psychiatry, psychosomatics, or behavioral medicine. Yet, the main stream in clinical psychology views the etiology of mental disorders, their time courses and susceptibility to psychological treatment still through the magnification glass of linear input-output philosophy of human functions. Owing to this paradigm, linear combinations of variables (as inner conflicts, irrational cognitions, or stressors) trigger the development of psychiatric diseases or disorders in genetically predisposed individuals. Therefore, linear multivariate regression models are assumed be able to predict the probability of falling ill or suffering from disorders. As an important field in clinical psychology, psychotherapy research defines randomized controlled trials as the golden standard of outcome research. Here, patients randomly assigned to different treatment modalities are being compared with respect to the outcome of different tests. In this regard, the input (treatment) is thought to determine the outcome (treatment effects).

Contrary, or rather, supplementing this line of research is the scientific paradigm of self-organization, i. e. the functioning of complex nonlinear systems with circular causality at its center. Gestalt psychology, traditionally concerned with patterns ("Gestalts") in perception, human behavior and interaction (e. g., those prevalent in group dynamics, Lewin [25]) focuses on such self-organization processes. Gestalt psychologists like Wolfgang Köhler (e. g. [28]), Wolfgang Metzger, Max Wertheimer, Kurt Lewin and others can be seen as direct predecessors of modern complexity researchers in psychology [67]. Another root of this development is Jean Piaget's equilibration theory of action-cognition patterns (schemata) describing assimilation-accommodation-cycles of these schemata [49]. During these processes, input from the inner and outer environment assumes the role of disturbing stimulation of individual system dynamics. A third important line of thinking in circular causality comes from anthropological medicine. The "Gestaltkreis" integrates feedback loops between sensorial and actional systems on the one side, and individual and environmental systems on the other side (ecosystemic approach) [84].

Introduction

During the past decades in clinical psychology, it was particularly the transdisciplinary approach of *synergetics* [16] which inspired a specific nonlinear and complexity research on cognition [17,79]), social interaction [41,77], etiology and dynamics of mental diseases (e. g., [57,78]), and psychotherapy (for an overview see Haken and Schiepek [21]). Synergetics describes, measures, and explains the autonomous processes of pattern formation and pattern transitions in complex nonlinear systems. Founded on Haken's fundamental discovery that these processes do not depend of the matter of the sys-

tems they occur in, synergetics became one of the most important inspirations to many scientific fields and topics. Especially, Haken early transferred synergetics to brain research (e. g., [3]), since the brain is an outstanding example of a complex, self-organizing system. Today, it is widely accepted that the brain and a serial computer not only differ profoundly, but there is almost nothing they share: No wonder in light of the more than 10^{11} nonlinear interconnected neurons forming a dynamic mega-network of neural networks with essential features like arrays of emerging and submerging synchronizations, its flexibility and ever changing pattern formation, working at the edge of chaos, or realizing combined (activating and inhibiting) feedback mechanisms following the principles set forth by synergetics which describes the laws of self-organizing systems [18,19].

Taking a closer look at most of the phenomena clinical psychology is concerned with it becomes obvious that they appear to be of dynamic nature. Human development processes, human change and learning processes, the dynamics and prognosis of mental disorders, problems manifesting in social systems like couples, families, teams, or the question of how psychotherapy works: Self-organization is a ubiquitous entity.

Dynamic Diseases

Mental disorders are characterized by specific dynamic patterns, mirroring "endogenous" and common features of a disorder (like the repetitive phases of unipolar major depression or the bipolar phases of bipolar disorders, oscillating between mania and depression), as well as the effects of an individual life-style including individual coping and treatment efforts. Mental disorders can be conceived as highly structured and coherent states which enslave and thus impair the individual's mental and social functioning. Following the "enslaving principle", emerging order parameters reduce the degrees of freedom in the behavior of the single parts of a system. There is phenomenal evidence that this is the case in many mental disorders. Obsessive-compulsive disorder patients coerced to repeat unwished thoughts or rituals are just but one most impressive example. On the brain level, such pathological states correspond with abnormal synchronization in specific neural networks impairing brain functions. In obsessive-compulsive disorders, cortico-striato-thalamo-cortical feedback-loops are thought to be at the center of the dysfunctional network [55,59], while abnormal synchronization in highly similar neural populations is the source of Parkinsonian resting tremor [34,43].

At times, transitions between different pathological states or between states of health and disease are linear and balanced, at other times they are discontinuous and abrupt, such as in nonlinear phase transitions accompanied by critical fluctuations described by synergetics to occur in physical systems. Such transitions have been reported for unipolar or bipolar cyclic depression (e. g., [1]), and also for schizophrenia [68]. The usefulness of the concept of attractors in psychopathology is best reflected by the final common pathway of different disorders with similar phenomenology and syndromal patterns. Different initial conditions and different qualities and degrees of stressors and vulnerability factors may result in similar pathological end-states on the one hand. But on the other hand, small fluctuations within the intrapsychic or environmental conditions or small differences of some boundary or threshold conditions may result quite different disorders or may decide between health and disease (for a dynamical simulation of major depression see Schaub and Schiepek [56]). The encouraging message synergetics delivers is that while the structure of a generic system may stay unchanged, small changes in control parameters, threshold conditions, and internal or external fluctuations are able to trigger dramatic changes in the behavior of the system. As a consequence, therapy exerting changes of these parameters is thus able to trigger return of the system to a healthy state.

For illustrative purposes we present results of a computer simulation of different chronic courses of schizophrenia [57]. A qualitative network model of five macroscopic variables was transformed into a set of nonlinear difference equations, with each equation describing and determining the change rate of each variable from t to $t + 1$. The empirical references of the simulation model were empirical studies of the chronic course of schizophrenic patterns, mostly mixed psychotic episodes, healthy functioning, and chronic states. For example, Ciompi and Müller [12] report on eight different patterns in the long-term evolution of schizophrenia, most of them reproduced by our model. These patterns result from various combinations of slow vs. acute onset, acute episodes vs. progressive deterioration, and remission vs. chronic end-state.

Variables taken into account were chosen from reviews in psychiatric and psychological schizophrenia research (e. g., [6,7,11]). Selected were (1) degree of cognitive disorders, (2) emotional and interpersonal stress, (3) withdrawal and social isolation, (4) degree of expressed (negative) emotions in the social environment of the patient, and (5) positive symptoms like delusions and hallucinations. The parameters mediating the nonlinear in-

terplay of these macroscopic variables or order parameters were (a) diffuseness of affective-cognitive schemata as a central long-term vulnerability of mental functioning, (b) dopamine and serotonin metabolism, (c) social deficits and lack of competencies, (d) genetic risk for schizophrenia, and (e) some parameters mediating mixed feedback processes, especially the negative feedback responsible for antipsychotic damping effects of the pathology. Results of the simulation runs are indicated in Fig. 1. The simulation reproduces most precisely (a) episodic patterns with prodromal symptoms and acute onset, (b) acute onset, but continuous evolution with chronic end-state, and (c) slow and smooth onset with chronic long-term course (see also [76]).

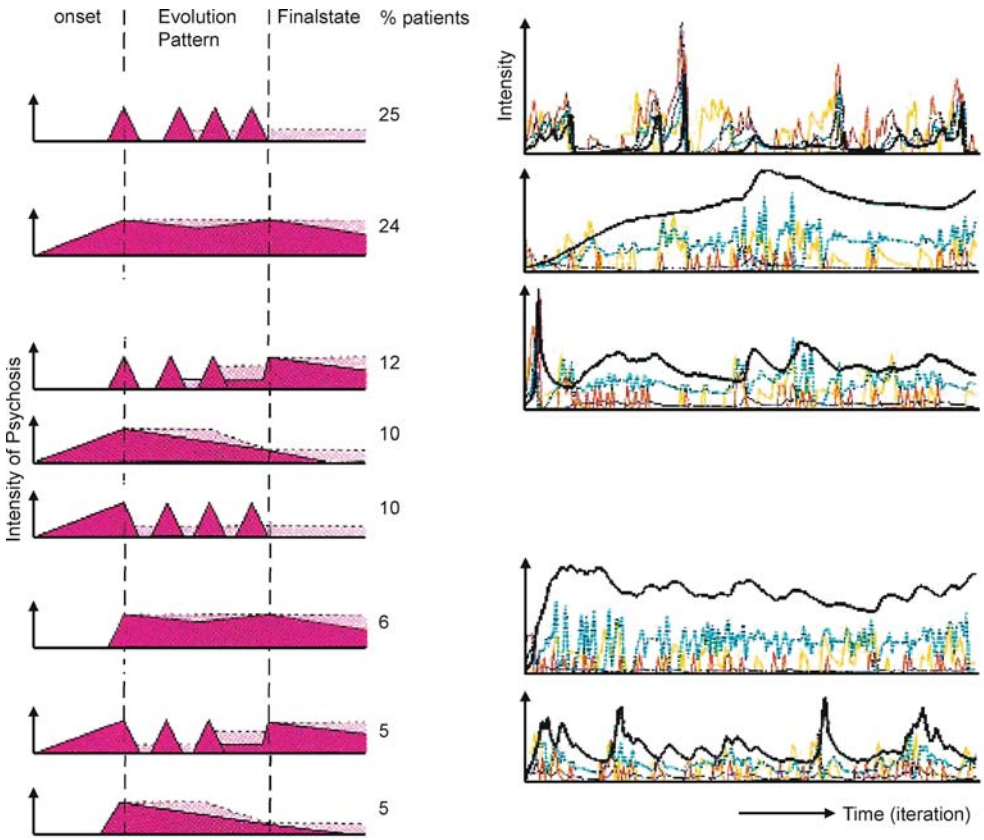
Other simulation models are effective on the microscopic level of neural networks. Kruse et al. [27] introduced a model focusing on the coupling dynamics between neurons processing brain correlates of social experiences. If unable to learn from cues delivered by the relevant environment, this system will fail to establish adaptive and coherent structures. When inducing fluctuations which promote re-learning and self-healing processes, the neural network causes incoherent and chaotic behavior. Most current models of schizophrenia take into account the neural circuits of relevant brain regions (cortical areas, basal ganglia like striatum and pallidum, thalamic areas, brain stem centers) and particularly the equilibria between different neurotransmitters and neuromodulators (e. g. [9]). The complicated local balances and their (non-) equilibria states are in the focus of the strongly evolving field of computational or systems neuroscience [15,44].

Not only central neuroscience has benefitted from concepts introduced by synergetics, however, but also physiology studying the effects of the autonomic nervous system (ANS) activities on peripheral systems, such as the cardiovascular, the respiratory, or the microcirculation system. Such activities are most prominent as the ANS engages in the mediation of emotions.

Self-Organized Synchronization Patterns in Peripheral Physiological Systems

For decades, the study of the ANS involvement in emotional arousal and its impact on the cardiovascular system has attracted clinical and scientific attention in psychology and psychophysiology (e. g. [65]). Ever since the seminal findings of W. B. Cannon [8] and H. Selye [63] on the general adaptation syndrome, colloquially condensed as stress, the clinical relevance of emotional responses became ultimately clear. This obvious clinical relevance is thwarted by the fact that direct observations of ANS activity in humans are restricted not only because of ethical constraints but also because of the fear to provoke what they strive to detect. Therefore, the study of the ANS in humans had to rely for a long time preferably on indirect measures, such as the power spectral density (PSD), a linear computation method. Based on the fast Fourier transform (FFT) which extracts periodic components in the frequency domain, the PSD was favored by many researchers due to its computational ease to analyze frequencies inherent in the two branches of the ANS, the parasympathetic (PNS) and the sympathetic nervous system (SNS). This allowed to divide the effects of the PNS and SNS activity on variations of the heart rate, the so-called heart rate variability (HRV), into three major variance components, the very low frequency (VLF) band below 0.04 Hz, the low frequency (LF) band between 0.04–0.15 Hz, and the high frequency (HF) band between 0.15–0.45 Hz. While the origin of both VLF and the HF bands is not debated, controversy reigns whether the origin of the LF band is attributable to SNS activity or whether it represents a mixture of SNS and PNS activity. Subsequently, some authors propose calculating the LF to HF ratio assumed to reflect the sympathovagal balance (for an overview see [73]). There has been growing discontent and criticism as to the validity of such drawer style classifications based on the consideration that the PSD, or FFT resp., as a linear routine is only able to detect linear properties, that are to some extent included in most physiological signals [23,86]. That, however, should restrict and limit its use since an increasing body of scientific evidence is demonstrating the obvious: In times of adaptation and rapid changes – a hallmark of life and its living systems – healthy ANS activity exhibits nonlinear dynamics necessary to mediate responses appropriate to those change processes. This is particularly the case for emoting as one of the most volatile change patterns.

However, this is not only true for discrete emotion transitions but also for a process crucial for the maintenance of health, namely psychophysical relaxation. Contrary to the rigid scheme depicted above, Perlitz and coworkers have introduced a relaxation model which takes into account adaptive, self-organizing characteristics of the central and peripheral subsystems involved in the psychophysical relaxation process. They scrutinized the physiological conditions and interactions observed with the emergence of a frequency at ca. 0.15 Hz, which in terms of the classical scheme is attributed to the transition between parasympathetic and sympathetic nervous activity. This frequency prevailed at different amplitudes in HRV, blood pressure and respiration, but foremost in the microcirculation of the forehead skin. Using several nonlinear methods, such as wavelet time frequency distribu-



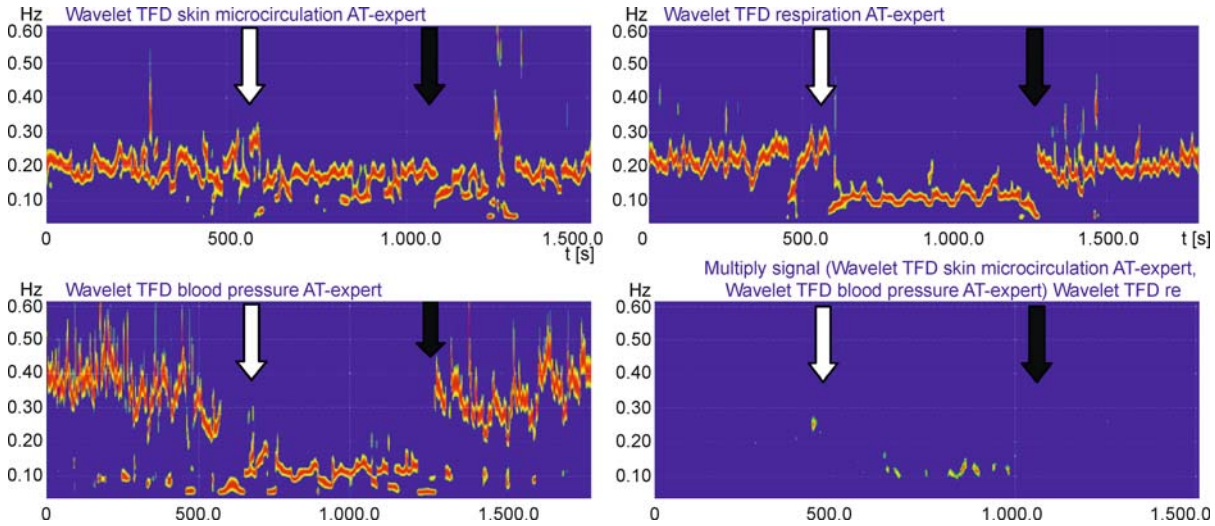
Self-Organization in Clinical Psychology, Figure 1
Different patterns of the long-term evolution of schizophrenia (empirical data from a study by Ciompi and Müller [12]) (*left*) were reproduced by simulations based on a set of five coupled nonlinear difference equations with different parameter values [57] (*right*)

tions (TFD) or post-event-scan (PES) analysis, this 0.15 Hz frequency band (range 0.12–0.18 Hz) emerged or erupted with amplified oscillations and periods of 6–7 s in all time series of subsystems under study. The emergence clearly depended on psychomotor drive reduction which can be either reduced by taking naive relaxation maneuvers (such as closing the eyes), or be enhanced using auto-suggestive means, such as autogenic training. Their zest to elaborate the origin of this frequency was supported by invasive observations with anesthetized dogs made by Lambertz and colleagues who had presented their findings earlier. They found a rhythm at a similar frequency which originated in reticular brainstem neurons of freely breathing dogs when administering narcotics to reduce drive. Followed by the emergence in those unspecific reticular neurons, this frequency also emerged in arterial blood pressure, HRV, and respiration [32,47,48]. This reticular rhythm, termed retR, was unaffected by changes in the frequency of respiration or arterial blood pressure which could both be presumed to exert distinct influences owing to linear models. Rather, in these experiments respiration and HRV were entrained to the 0.15 Hz band at 1:1, 2:1 and 1:2 integer number ratios which are, according to Bethe [5], an outflow of central-peripheral order–to–order transitions. With regard to parallels in frequency and dynamics observed in man and dog, Perlitz and coworkers suggested that also in humans the 0.15 Hz band most likely originates from reticular neurons of the lower brainstem network [46,47,48].

In summary, the findings presented in Fig. 2 underpin the theory of synergetics, since there is reason to regard the ca. 0.15 Hz frequency as an order parameter and the level of mental drive as control parameter. The ca. 0.15 Hz frequency is a prominent example of biological pattern formation lacking external or macroscopic control.

Nonlinear Dynamics in the Communication of Patient and Therapist

As mentioned above, psychotherapy is usually conceptualized as the application of psychological treatments to pa-



Self-Organization in Clinical Psychology, Figure 2

Wavelet time frequency distributions (TFD) of peripheral noninvasively obtained recordings of a female expert in autogenic training (AT, 56 yrs., healthy, non-smoker, 15 yrs practice AT). *Top left*: TFD of glabella skin microcirculation photoplethysmography; *top right*: TFD of chest respiration related movements; *bottom left*: TFD of peripheral systemic arterial blood pressure; *bottom right*: So-called "joined TFD" of PPG-, respiration- and blood pressure-TFD, a novel method by Besting and colleagues (2005) (multiplying TFDs yielding only frequencies prominent at identical times and identical frequencies, www.Simplana.de) used to compute the intersection of TFD time series. *White arrows* mark the start of AT, *black arrows* mark the end of AT. In the TFD of PPG, the main frequency is at ca. 0.21 Hz prior to the onset of AT and is clearly stabilized at ca. 0.18 Hz with the start of AT, with signs of dissociation when terminating AT. The TFD of respiration supplies ample evidence of an order–order transition triggered by the practice of AT: The main frequency plummets from ca. 0.25 to 0.15 and 0.07 Hz to be maintained at ca. 0.12 Hz. With termination of AT, the main frequency skips back to frequencies shown beforehand. The TFD of systemic arterial blood pressure exhibits an intersection of approx. 90% during the AT section, but also few minor intersections before and after AT (data not shown); the joined TFD intersection shows merely few frequency "spots" at ca. 0.12 Hz during AT

tients in order to change their problem states and diseases. However, as different research programs revealed during the last decades, psychological change processes show all important features of nonlinear systems – like deterministic chaos, nonstationary phase transitions, and nonlinear coupling between patient and therapist. Physiological synchronization appears to be realized at an interpersonal level (between therapist and patient) as well as between different phenomenological levels of the interpersonal system (speech qualities and psycho-physiological variables). In a study of Villmann et al. [83] heart rate, respiratory frequency, muscular tension, and skin conductivity were measured from both, therapist and patient, during 37 therapy sessions. Speech production was analyzed by the Mergenthaler model focusing on emotional feeling and cognitive referential activity/abstraction [38]. Physiological data were analyzed by an artificial neural network approach (growing self-organizing map), which uses a kernel smoothing for improved data density estimation. It was possible to generate an entropy model of psycho-physiological variability detecting emotionally instable phases during the therapy process. The entropy reflecting psycho-physiological and emotional variability was related to the dramatic value of speech analysis according to the cycle model of Mergenthaler.

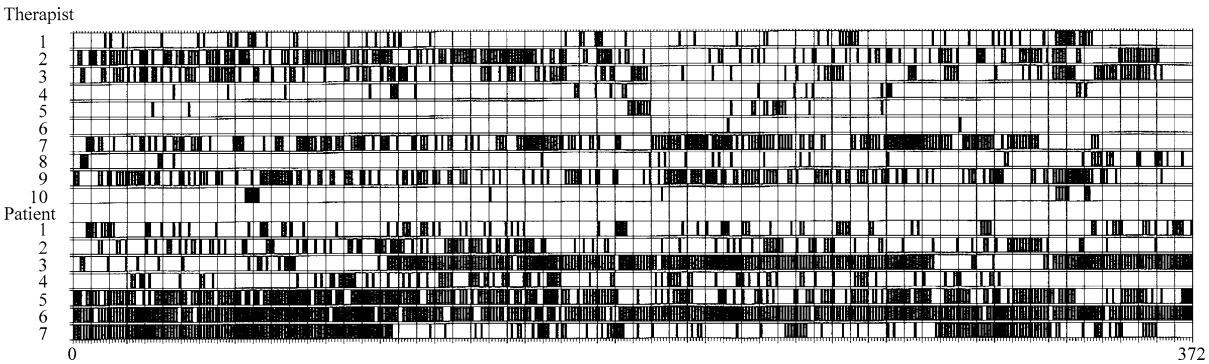
Empirical evidence exists also for synchronized chaoto-chaotic phase transitions in the brains of therapist and patient during a therapeutic interview, measured by local largest Lyapunov exponents in the EEGs of both interaction partners [52].

Taking into account the importance of the therapeutic relationship for the treatment outcome the attention of a study realized by Schiepek and co-workers focused on the interactional process between therapist and patient [26,58]. The authors used the method of sequential plan analysis, which is a development of the hierarchical plan analysis proposed by Grawe and Caspar (e. g. [10]). Plans in this sense are verbally or non-verbally communicated intentions of self-presentation in a social situation. Patient's and therapist's interactional behavior was analyzed on the basis of video recordings. Two complete therapies (13 and 9 therapy sessions, resp.) were encoded with a sampling rate of 10 s (Fig. 3). The construction of an inclusive hierarchical plan-analysis leads to an

Self-Organization in Clinical Psychology, Table 1

Second-order plans and categories of self-presentations as identified by the hierarchical plan analysis of a complete 13-session psychotherapy. Encoding of therapist and patient. Plans and categories are used as ideographic observation categories for the sequential plan analysis

	Second-order plans	Categories of self-presentation
Therapist	1 show competence 2 encourage a trusting relationship 3 show understanding 4 motivate her	I encourage trust/create a secure atmosphere
	5 encourage her to reflect on her patterns of thinking 6 confront her with her avoidance and problem behavior	II confrontation/exposing to insecurity
	7 activate her 8 show her that she is responsible	III encourage self-responsibility of the patient
	9 guide her focus of attention 10 give her structure	IV activate structuring work
Patient	1 demonstrate strength and competence 2 make it clear that things are or have been difficult 3 be a good patient/create a good relationship to the therapist	I search for sympathy/appreciation/good relationship
	4 show that your suffering is strongly influenced by external causes 5 ask for help from the therapist	II externalization/demonstration of helplessness
	6 show interest and willingness in solving your problems 7 protect yourself from threatening changes	III problem-oriented work (self-relatedness vs. avoidance)



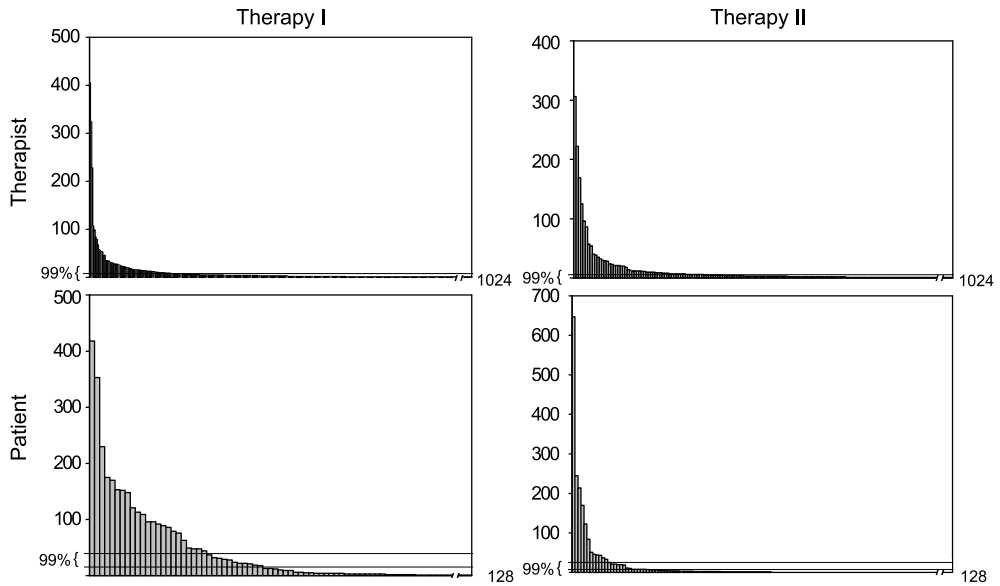
Self-Organization in Clinical Psychology, Figure 3

Nominal sequences of interactional plans of the therapist (*top*) and the patient (*bottom*) during a psychotherapy session. The sampling rate is 10 s. Different plans can be realized simultaneously. The pattern looks like a music score with the plans representing the different instruments of an orchestra. A sonification of the score of plans coded from a 13-session psychotherapy is recorded on a DVD added to the textbook of Haken and Schiepek [21]

ideographic categorical system for the observation of the client–therapist interaction (Table 1).

The first hints of order in the dynamics came from the distribution of simultaneous configurations (on-off-patterns) of plans in the scores. This distribution follows a power law ($1/f^a$) demonstrating a distinct structure order within the data (Fig. 4). Following Bak et al. [2], power law-distributions as demonstrated in Fig. 4 emerge from self-organized criticality within dynamic systems.

Further data analysis was based on the time series of the highest-level categories, the so-called categories of self-presentation (see Table 1). Since in the hierarchical system of the plan analysis the operators at the lowest observation level were quantified by intensity ratings, the plans and the self-presentation categories at the top level integrating the lower level categories were also quantified. The time series were analyzed by methods which are sensitive to the nonlinearity as well as the nonstationarity of the time series [21,26,58,70]. Nonlinearity was proofed by surrogate data tests [51] using random surrogates and FFT-based phase-randomized surrogates [69]. Whereas fractal dimensionalities of the empirical time series (based on



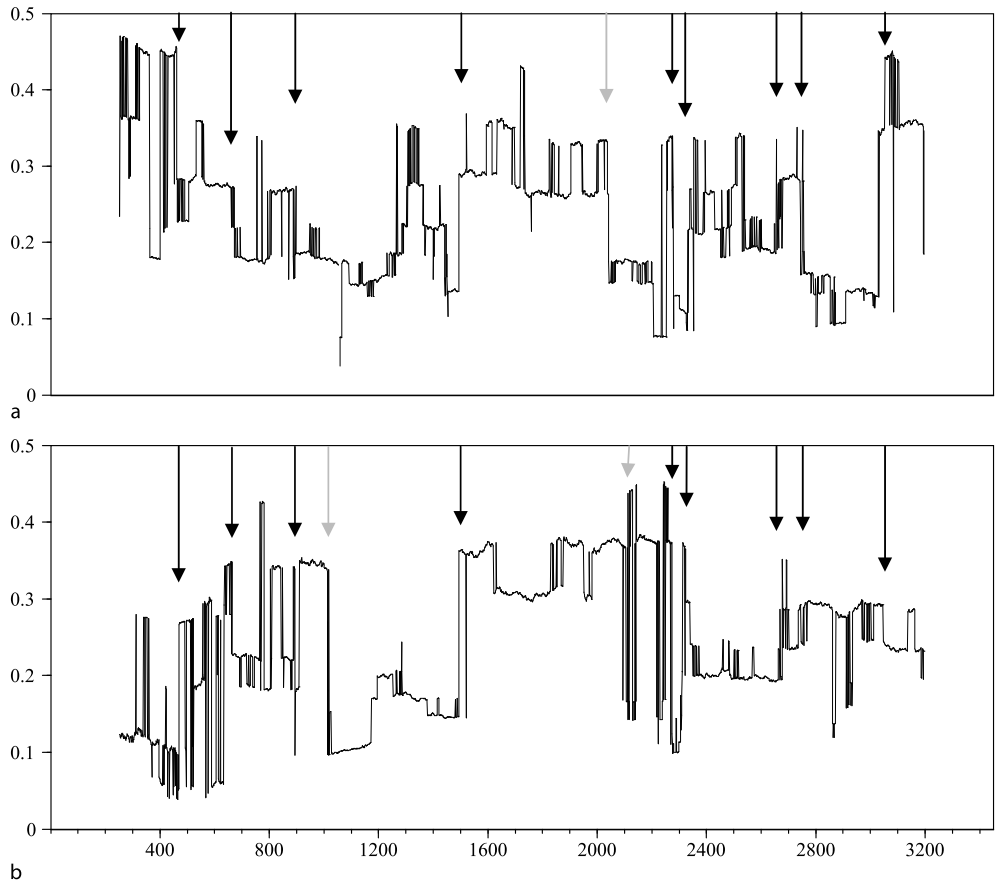
Self-Organization in Clinical Psychology, Figure 4

Empirical frequencies of constellations of interactional plans realized by therapists (10 plans) and patients (7 plans) within two psychotherapies (therapy I: 13 sessions, therapy II: 9 sessions). *X*-axis: Number of all possible configurations of plans (therapist: $2^{10} = 1024$, patient: $2^7 = 128$) ordered by the frequency of their realization. *Y*-axis: Frequencies of plan configurations. The distributions follow a power-law ($1/f^a$) distribution

the correlation dimension D2 as well as mean Pointwise D2 [66]) saturated at finite values (convergence to a fractal dimensionality of about 6), random and FFT-surrogates did not. The methods of PD2 [66] and of the local largest Lyapunov exponents (algorithm from Rosenstein et al. [53]) were used to identify phase-transition like discontinuities. Following the evolution of PD2 dimensionalities, both therapies realized nonstationarities, and both therapies showed periods of strongly synchronized (with correlations from 0.80 to 1.00) and anti-synchronized PD2-processes (with correlations from −0.80 to −1.00) between patient and therapist. Quite similar and even more pronounced dynamical jumps were to be seen in the development of the local largest Lyapunov exponents (Fig. 6), representing changes in the chaoticity of a time signal [26]. An important part of the discontinuities of the LLLE were exactly synchronized between patient and therapist. Obviously both persons create a dynamic self-organizing communication system, which allows for the individual change processes of the patient.

These results get support from nonlinear coupling measures between the time series of the interaction partners. Pointwise transinformation as well as pointwise coupling conditional divergence [33,80] were applied to the data, and both indicate changing and time-dependent coupling strengths between the time series of the interaction partners. There is no priority to the therapist's influence on the patient, which contradicts the classical idea that input from the therapist should determine the client's output. The other way round is also true and both constitute the circular causality of psychotherapeutic self-organization.

In other studies, sequential plan analysis was applied to the microdynamics of group interaction [21]. In a group of five persons a creativity and problem solving task was to be solved within 2,5 h (creation of ideas, rules, and physical handicraft realization of a prototype board game from different materials). Similar to the psychotherapy study the sampling rate was 10 s. The superordinate plans which could be identified for all five persons were (1) spontaneity and emotional engagement vs shyness, restricted behavior, and orientation to social norms, (2) engagement in the group interaction and in positive social climate, (3) task orientation. Length of time series was about 810 coding points (= intervals). D2 as well as mean PD2 estimates saturated at a fractal dimensionality of about 5 for all categories. The embedding of the time series was realized by two ways: (1) The phase space was constituted by the three dimensions of superordinated plans with five trajectories representing the five group members, or (2) the phase space was constituted by the five persons with three trajectories representing the time course of the three plans



Self-Organization in Clinical Psychology, Figure 5

Synchronized jumps in the dynamics of local largest Lyapunov exponents (*black arrows*). *Grey arrows* indicate not clearly synchronized changes. a Therapist, b Patient

Self-Organization in Clinical Psychology, Table 2

Factors (principal component analysis) of the Therapy Process Questionnaire (TPQ). Factor analysis was based on 94 therapy processes (mean stay = 66 days, daily ratings). Seven first-order factors (*right*) are related to three second-order factors (*left*). Numbers behind the first-order factors indicate factor loadings on second-order factors (for details see [21])

I(2) Change involvement	I Therapeutic progress/confidence in treatment effects/self-efficacy (.571) VI Intensity of therapeutic work/motivation to change (.596) V Opening of perspectives/personal innovations (.649)
II(2) Relationship/Social climate	III Quality of the therapeutic relationship/openness/confidence in the therapist (.705) II Ward atmosphere, social relationship to other inpatients (.692)
III(2) Emotionality	IV Dysphoric emotions/self-relatedness (.732) VII Impairment by symptoms and problems

(additional embedding dimensions result from time delay coordinates). In both cases PD2 results show an evolving pattern of quasi-attractors with changing complexity, and LLLEs (algorithm from Rosenstein et al. [53]) portray chaoto-chaotic phase transitions with clear-cut and interpersonally synchronized jumps – similar to the dyadic interaction of the psychotherapy study.

Self-Organization in Human Change Processes

A quite different approach to human change processes focuses on inpatient treatments at a hospital of psychosomatics. In a study by Schiepek and coworkers (results in [21]) 94 change processes were investigated, realized by 91 inpatients with different diagnoses (depression, anxiety

disorders, posttraumatic stress disorders, eating disorders, somatoform disorders, and others). The time series data was produced by patients' self-ratings which were completed once a day in the evening. For this purpose a 53-item rating sheet was developed (*Therapy Process Questionnaire [TPQ]*, [21]) whose factor analysis resulted in a solution of seven factors defining the subscales of the questionnaire (Table 2). The ratings combined seven-step Likert scales and visual analogue scales especially for ratings of emotions. TPQ measurements reflect important aspects of the patient's experience of progress and goal attainment, emotional involvement, self-efficacy, therapeutic relationship, social relations with other inpatients, and the ward atmosphere.

The inclusive outcome criterion integrated the following measures: Inventory of Interpersonal Problems (IIP), Gießener Beschwerdebogen (GBB), Hospital Anxiety and Depression Scale (HADS), Questionnaire for Social Support (F-SOZU), a life-quality questionnaire (Münchener Lebensqualitäts-Dimensionenliste), a self-efficacy questionnaire (Fragebogen zur Generalisierten Kompetenzerwartung), the Sense-of-Coherence Questionnaire, and an interview-based assessment of personal resources. Additionally, therapists and patients scored the overall treatment effectiveness and treatment quality.

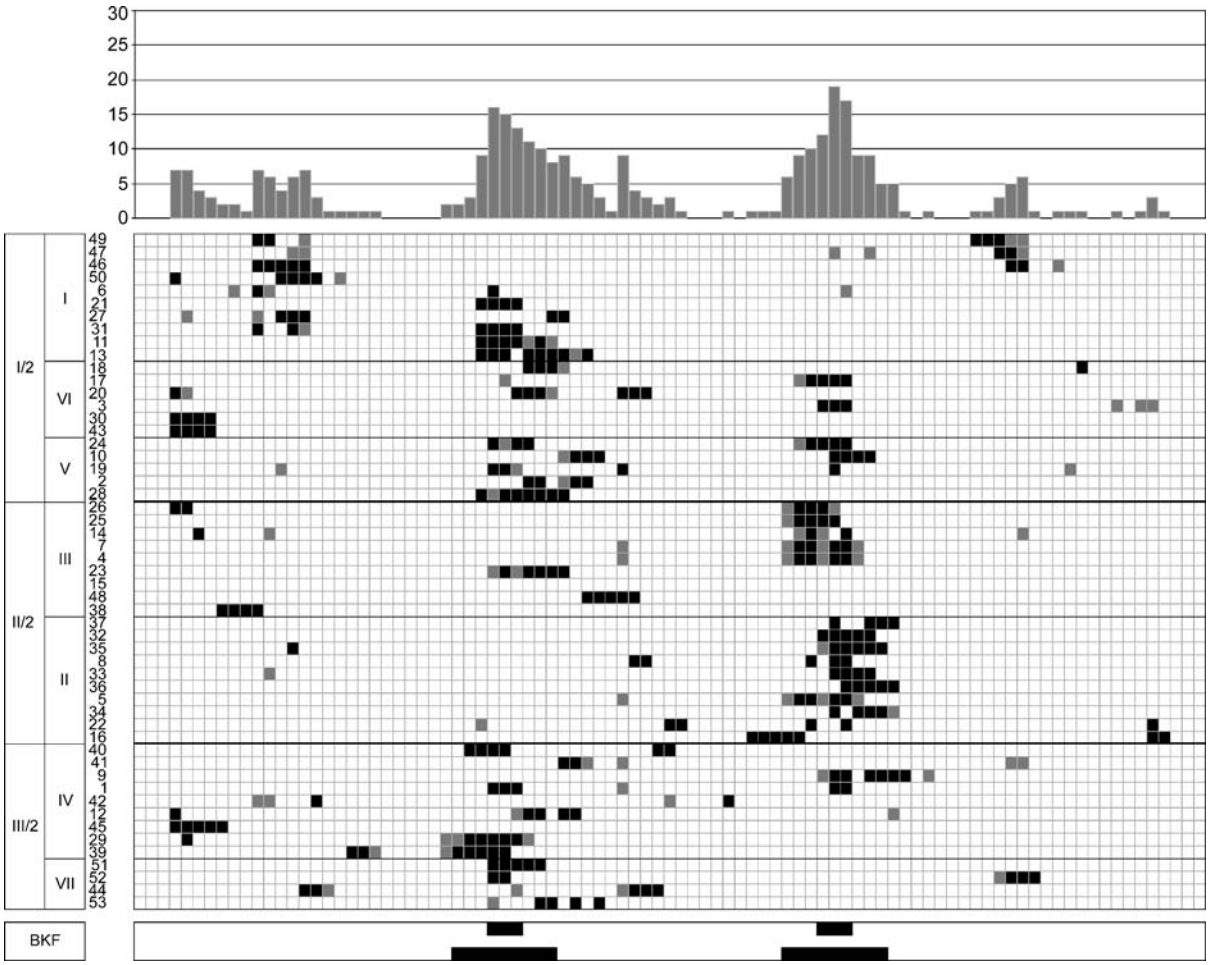
Results confirmed synergetic conceptualizations of how psychotherapy works and corroborated hypotheses drawn from this model. Here therapy is supposed to provide support for the patient's own self-organization processes, which should be characterized by cascades of order-to-order transitions accompanied by critical instabilities of the process. Pathological and restrictive order should be transformed into more flexible and adaptive patterns of behavior, and the synchronization of the different aspects of the patient's experience should undergo some transformations. Exactly this could be observed.

Significant correlations exist between the local maxima of critical fluctuations and the outcome of psychotherapy. The local maxima were defined by the difference between the mean dynamic complexity of the whole psychotherapy process and the maximum of the complexity which was observed during the process. Correlations were −0.455 (second-order factor I: "Change involvement" of the TPQ, $p = 0.002$), −0.431 (second-order factor 2: "Relationship/social climate", $p = 0.003$), and −0.572 (second-order factor 3: "Emotionality", $p = 0.000$) (compare Table 2). Negative correlations result from the fact that increased local maxima of dynamic complexity correspond to reduced problems, symptoms, and impairment.

The *dynamic complexity* combines a fluctuation index with a distribution index. The fluctuation index measures the frequency and amplitude of the change rates of a time series between the reversals of the development within a scanning window gliding over the whole time series. For analysis purpose a window width of seven measurement points (= days) was introduced. The distribution index measures the scattering of realized values within a given scanning window. The more scores are restricted to only narrow intervals of the available scale range, the smaller the distribution index becomes. The score of this index increases as the interval filled by the realized values grows. The algorithm solves the problem of value distribution independently of the scale resolution, the width of the scanning window, and of any combination of these parameters.

In order to answer the question if the observed intensities of dynamic complexity reach critical values, intra-item calibration procedures were used in order to define adequate thresholds fitting to the actual dynamics. The time series of dynamic complexity were standardized by z-transformations, providing significance thresholds of 5% or 1%. Applying this threshold method to all items of the TPQ reduces the quantitative complexity signals of each time series to a three-step signal (not significant, complexity exceeds a 5% threshold, complexity exceeds a 1% threshold). A synopsis of these qualitative signals referring to all items of the TPQ gives an impression of the localization of critical fluctuations during the whole process. Dynamic complexities seem to be synchronized over many items and factors of the TPQ, resulting in the structure of columns of grey (<5%) or black (<1%) dots. In a large part of the investigated therapies such column-like structures could be identified. In an item-by-time synopsis they indicate phases of intensified as well as synchronized fluctuations and entropies of quite different aspects of the process. Consequently, these item-by-time synopses are called *complexity resonance diagrams* (Fig. 6).

In order to confirm the structures found within the complexity-resonance-diagrams, surrogate tests were realized based on random as well as on FFT-based surrogates of the time series. The empirical patterns are impressively different from the surrogate-based patterns (all realized comparisons with $p = 0.000$). Further support for phase-transition like phenomena in the change processes came from recurrence plots representing similarities and dissimilarities of dynamic segments of a whole time series [13,80,85]. This method is based on the embedding of time series into a phase space constructed by time-delay coordinates, a method which is also crucial in the algorithms for the estimation of dimensional complexity or chaoticity (e. g., Kolmogorov–Sinai-Entropy, Lyapunov Exponents). Neighbors in the time-delay phase space rep-

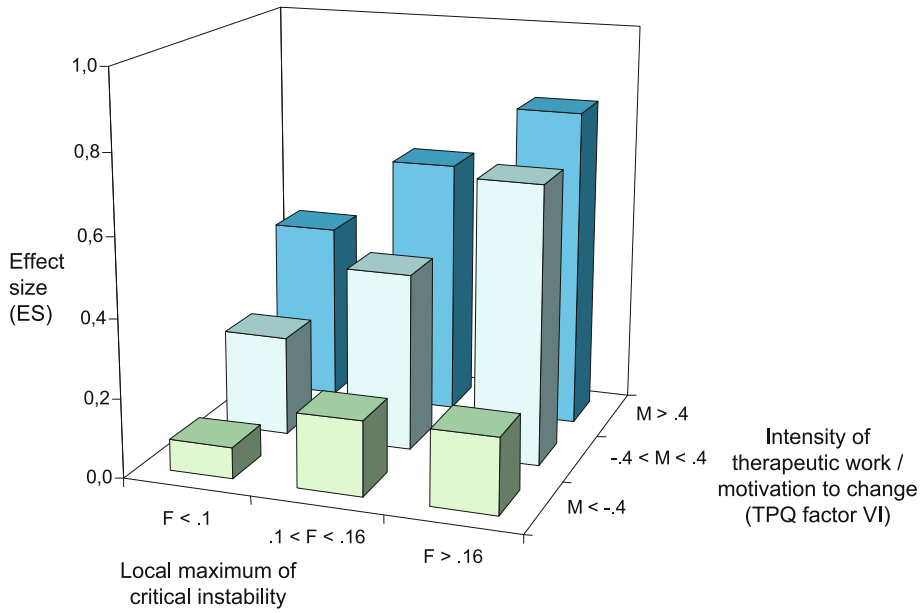


Self-Organization in Clinical Psychology, Figure 6

Complexity resonance diagram of a psychotherapy process. Such diagrams portray the threshold exceeding dynamic complexities of a process encoded by the 53 items of the Therapy Process Questionnaire (TPQ). *Gray dots*: 5% threshold of significance; *black dots*: 1% threshold of significance. *X*-axis: Days of hospital stay, *Y*-axis: Items of the TPQ arranged by the order of the factors as reported in Table 2. Window width for the calculation of dynamic complexities is 7. Column-like structures indicate phases of critical instabilities during the process

resent similar dynamic segments and are plotted by a dot in the recurrence plot. Dissimilarities are represented by empty columns in the recurrence plots, which in many cases exactly correspond to the columns of dots in the complexity-resonance diagrams. The overall correlation is −0.45, if small shifts (lags of + or −3 measurement points at maximum) will be allowed. This means that periods of critical instability correspond to transient dynamics outside of the quasi-attractors established by the self-organizing system under consideration. These different ways to identify critical phase transitions are further validated by the time frequency distribution (TFD) of the time series. The TFD method uses wavelet spectra in order to scan the evolution of the frequency distributions within a signal [33,80]. It is a dynamic counterpart to the static fast Fourier transformation and allows for the identification of pronounced frequency amplitudes or changes in the frequency distributions. In the data set of the referred study these often appear exactly during the phase transitions which can be identified by other methods (see the synoptical representations of different time series analysis methods on the DVD in the textbook of Haken and Schiepek [21]).

An overall result of the study is shown in Fig. 7. It portrays the evidence that in order to bring forth change processes within self-organizing systems at least two conditions should be realized. The first condition: The degree of the control parameter energizing the system and push-



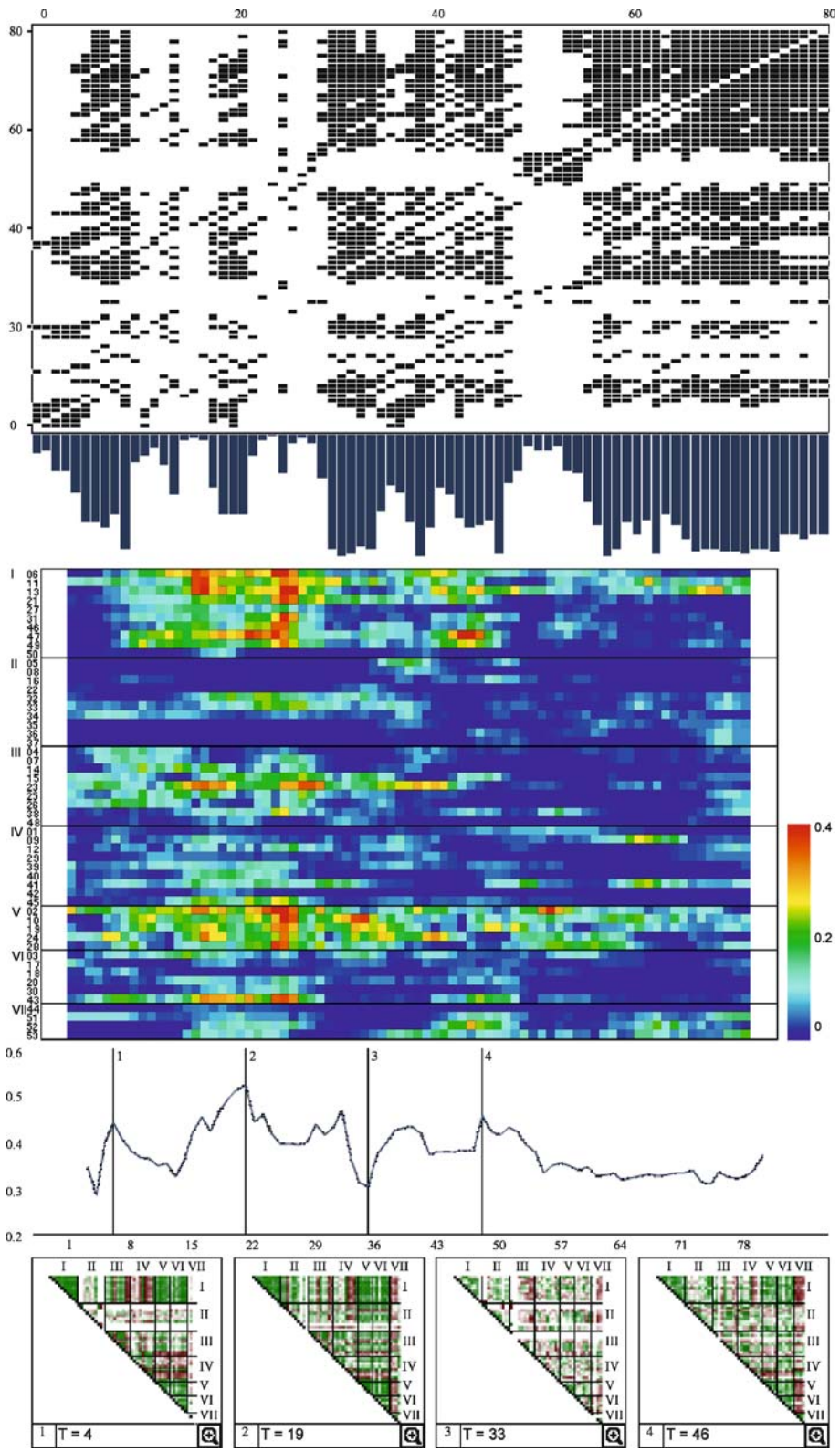
Self-Organization in Clinical Psychology, Figure 7

The effect size (ES) (mean ES of all outcome measures introduced in the study, see text) of inpatient psychotherapy is produced by an interaction between the local maximum of critical fluctuations and the intensity of the control parameter realized during the change process. The local maxima of fluctuations were defined by the difference between the mean dynamic complexity of the whole therapy process and the maximum of the complexity observed during the process. The diagram is based on the mean of the local maxima of all items. The control parameter was defined by the overall mean of the TPQ factor VI: Intensity of therapeutic work/motivation to change

ing it away from its actual equilibrium state should exceed a certain intensity level. With respect to psychotherapies this control parameter could be the patient's motivation to change including his engagement into the therapeutic work. Second condition: The degree of instability the system attains during its change process. This instability during emerging symmetries and symmetry breaking transitions is given by the local maximum of dynamic complexity during the hospital stay. The interaction of both conditions results in treatment effectiveness. A third important condition is not represented in Fig. 7: It is the experienced stability of the outer environment (context at the ward or therapeutic bond) or of the inner environment (as self-esteem, self-confidence, or activated resources). This context of stability is a prerequisite for a system to undergo critical instabilities.

The Concept of Self-Organization Promotes New Information Technologies in Clinical Psychology – The Synergetic Navigation System

Since self-organization and nonlinear dynamics seem to be ubiquitous in human change processes, it should be helpful to go beyond the diagnostics of steady states to an assessment of dynamics. Practitioners should get information on the therapy and its features *during* the ongoing process in order to use this information for an adequate placement of interventions and a control of the dynamics. "Controlling" self-organization processes in psychotherapies means the generation and co-creation (together with the patient) of adequate boundary conditions, the decision to do or to retain certain interventions, and to support the dynamics which the system is creating by itself. The patient takes an active and cooperative role in this understanding of data-based and co-creative change processes. Another important motivation for the development of real-time assessment comes from the evidence that most of the empirically identified specific and non-specific factors driving therapeutic change processes are connected with specific persons (the concrete therapist who meets a concrete patient in a concrete setting) and evolve by its nonlinear interactions in specific systems. These factors are (i) personal features of the patient like his motivation to change, his premorbid adaptation and degree of social functioning, personality integration, ego-strength, or comorbidities, (ii) personal and professional features of the therapist like his own personality integration, social and professional competencies, allegiance to his approach of


0
20
40
60
80
80
60
40
30
0
I
II
III
IV
V
VI
VII
0.4
0.2
0
0.6
0.5
0.4
0.3
0.2
1
2
3
4
1
8
15
22
29
36
43
50
57
64
71
78
I II III IV V VI VII
1 T = 4
2 T = 19
3 T = 33
4 T = 46

◄ **Self-Organization in Clinical Psychology, Figure 8**
Synopsis of a psychotherapy process as monitored by the Synergetic Navigation System. The time course of the inpatient treatment of a patient with eating disorders portrays a clear cut phase-transition associated with critical instabilities. *Top*: Recurrence plot of the item "Today I was successful to do steps towards my personal goals". *Dots* represent recurrent segments of the time series, empty spaces represent transitions. *Middle*: Complexity resonance diagram of all items of the TPQ. Different from Fig. 6, the intensities of the dynamic complexity of each item is transformed into colors. Items are arranged by the order of the first- and second-order factors of TPQ. *Bottom*: Mean of all inter-item correlations irrespective of the sign (absolute values). This is a measure of the overall synchronisation of the patient's experiences as represented by the items of the TPQ. The correlation structure is shown at four measurement points (days) of the psychotherapy process ($t = 4$, $t = 19$, $t = 33$, $t = 46$). Intensity of green represents positive correlations, intensities of red represents negative correlations

doing therapy, stress-resistance, and so on, and (iii) factors of the professional and social context (see the so-called generic model of psychotherapy [29,42]). In consequence, *evidence-based treatments* should be based on the evidence of concrete data mirroring the ongoing change process and on the professional decisions reflecting this insight.

Real-time monitoring actually uses internet-based presentations (including PDA or cell phone technology) of outcome and process questionnaires. Data are sent to a server, where they are stored and analyzed. Professionals and patients can inspect the results whenever they want. Experiences with real-time feedback to therapists (based on an outcome questionnaire the patient fills out during the therapy sessions in an ambulatory or outpatient context) are encouraging. Lambert and co-workers (e. g., [31]) were able to identify processes on the way of getting difficult or unsuccessful ("not on track" therapies, compared to more promising "on track" therapies), and helped therapists to correct these not-on-track dynamics by specific interventions. By this, threatening drop-outs could be avoided, bad results could be corrected, and on-track processes could be optimized and even shortened.

More sophisticated than the distinction between "on-track" and "not-on-track" courses is the feedback on self-organization features realized by a system based on synergetics [21]. The *Synergetic Navigation System* uses the therapy process questionnaire for daily ratings and applies methods from nonlinear time series analysis in order to identify important qualities of the change process. This are:

- Stability or instability of the dynamics as represented by the subscales (factors) of the TPQ (see Table 2), which is measured by the dynamic complexity
- Recurrence plots indicating transitions or repeating patterns
- Intensity of synchronization and time-dependent synchronization patterns between the items and the factors of the TPQ (realized by the cross-correlations of all items of the TPQ, calculated within a running window).

Figure 8 shows a synopsis of these analysis methods applied to a specific change process. Preceding the inspection of all analysis results the raw data series of the items and the time courses of the factors (z-transformed values) are available. Additionally patients can write an electronic diary after filling out the questionnaire. The diary entries can be presented within a gliding tip-tool running over the time series. By this, corresponding qualitative and quantitative information completes the picture.

The Self-Organizing Brain

The human brain is one of the most outstanding examples of a complex nonlinear system producing self-organized patterns of functioning. Since function corresponds to structure and vice versa, structural changes (changes of intersynaptic coupling strengths and network configurations, (re-)wiring patterns following the synchronized co-activity of neurons) can be explained by functional self-organization of neural populations. Perception, action and transition of action patterns, decision making, and cognitive, behavioral, as well as emotional learning are psychological functions following the principles of self-organization [21]. At a neural level they correspond to and are based on nonlinear brain dynamics. The emergence of order parameters and the occurrence of phase transitions can be described and measured on a psychological as well as on a neural level.

One of the phenomena modeled by synergetics is Gestalt perception – the construction of percepts and the switching of ambiguous visual patterns (e. g., Necker cube or stroboscopic alternative motion). These processes of Gestalt perception constitute the link between Gestalt psychology and actual mathematical modeling in synergetics [17]. The binding of different perceptual features or components to coherent structures or "qualia" seems to be due to synchronization processes of extended brain regions and converging integrative areas [64]. Pattern perception corresponds to pattern formation – as H. Haken puts it into pointed words. Tallon-Baudry et al. [71,72] measured enhanced gamma-band activity (30–50 Hz) in the EEG of the primary and secondary visual cortex while subjects identified a triangle within the offered stimulus material. This could be a fingerprint of correspond-

ing neural synchronization processes. This activity occurred when subjects saw a real object (triangle) as well as a figural illusion of the object (Kanizsa triangle), but not if geometrical components could not be composed to a true Gestalt. The research group of Basar–Eroglu and Stadler [4] measured significant gamma-band activity in EEG during states of perceptual switching triggered by stroboscopic alternative motions. In summary: Perception of multistability is one of the multifold cognitive processes giving rise to 40 Hz enhancement in the cortex, and coherent oscillations reflect an important mechanism of feature linking in the visual cortex which corresponds to the emergence of a neural order parameter. Changing order parameter dynamics during different cognitive activities was shown by Schupp et al. [62]. Mental imagery of an object could be differentiated from its concrete perception. The dimensional complexity of prefrontal EEG was increased during sensory imagery compared to the real perception of the same object (compare [36]).

The well-known movement coordination paradigm modeled by Haken et al. [22] was used to demonstrate neural correlates of instability and symmetry breaking processes in the motor brain. The order parameter in this finger movement experiment is the relative phase of the index fingers of both hands. Metronome-pacing – with movement frequency as the control parameter – triggers the system from parallel (out-of-phase) to mirror (in-phase) movement. Meyer-Lindenberg et al. [39] showed that the emergence of patterns in open, nonequilibrium systems like the brain is governed by their stability in response to small disturbances. Transitions could be elicited by interference at the neural level. Functional neuroimaging (PET) identified premotor (PMA) and supplementary motor (SMA) cortices as having neural activity linked to the degree of behavioral instability, induced by increasing frequency of the finger movement. These regions then were transiently disturbed with graded transcranial magnetic stimulation (TMS), which caused sustained and macroscopic behavioral transitions from the less stable out-of-phase to the stable in-phase movement, whereas the stable pattern could not be affected. Moreover, the strength of the disturbance needed (a measure of neural stability) was linked to the degree of the control parameter (movement frequency) and thereby to the behavioral stability of the system.

Synergetic research in clinical psychology is now reaching the brain level. The aim of an actual fMRI-study [60] is the investigation of phase transitions of brain activity and related subjective experiences of patients during their psychotherapy process. Repeated fMRI scans are related to the degree of stability or instability of the ongoing dynamics (measured by the dynamic complexity of daily TPQ-ratings) as well as to the therapy outcome. Real-time monitoring by the *Synergetic Navigation System* allows for the identification of stable or unstable periods and by this for a decision on the appropriate moments of fMRI acquisitions. Three or four scans are realized during each of the psychotherapy processes of 15 patients. The study includes only patients with obsessive-compulsive disorder (OCD) of the washing/contamination fear subtype (DSM IV: 300.3), without any medication or comorbid psychiatric or somatic diagnoses. Patients are matched to healthy controls. (This research is a multi-center study of the Ludwig-Maximilians-University Munich, Institute of Psychology (Prof. Dr. Günter Schiepek, head of the project), and Clinic of Psychiatry (PD Dr. Oliver Pogarell, Dipl. Psych. Susanne Karch, Dr. Christoph Mulert), Hospital of Psychosomatic Medicine Windach/Ammersee and Day Treatment Centre Munich/Westend (Dr. Igor Tominschek, cand. Psych. Stephan Heinzel, Prof. Dr. Michael Zaudig), University Hospital Vienna/Astria, Clinic of Psychiatry (Prof. Dr. Martin Aigner, Prof. Dr. Gerhard Lenz, cand. med. Markus Dold, Dr. Annemarie Unger), MR Centre of Excellence, Medical University Vienna/Austria (Prof. Dr. Ewald Moser, Dr. Christian Windischberger).

OCD seems to be an appropriate model system for synergetic studies in clinical psychology, since the pathological order parameter is phenomenologically quite evident, the disease has an obvious and quite stable time course, and therapeutic phase transitions – if they do occur at all – are easy to be observed. OCD-specific functional neuroanatomy is partially known: Friedlander and Desrocher [14] report on an executive dysfunction model corresponding to the cortico-striato-thalamo-cortical feedback-loops involved in perseverations and compulsions, and on a modulatory control model involved in the pathological mechanisms of anxiety and distress provoking obsessions.

The visual stimulation paradigm of the study uses symptom provoking, disgust provoking, and neutral pictures. The disgust and the neutral pictures are taken from the International Affective Picture System, whereas the OCD-related pictures are photographed in the home setting of the patients, showing specific and individualized symptom provoking stimuli.

For illustrative purposes we report on the results of a single case. It is a female patient, whose fMRI scans were taken three times during the 59 days of their hospital stay at days 9, 30, and 57. The healthy control was also scanned three times at identical time intervals as the patient. The second acquisition was done after an intensive period of critical instability of the TPQ-based time se-

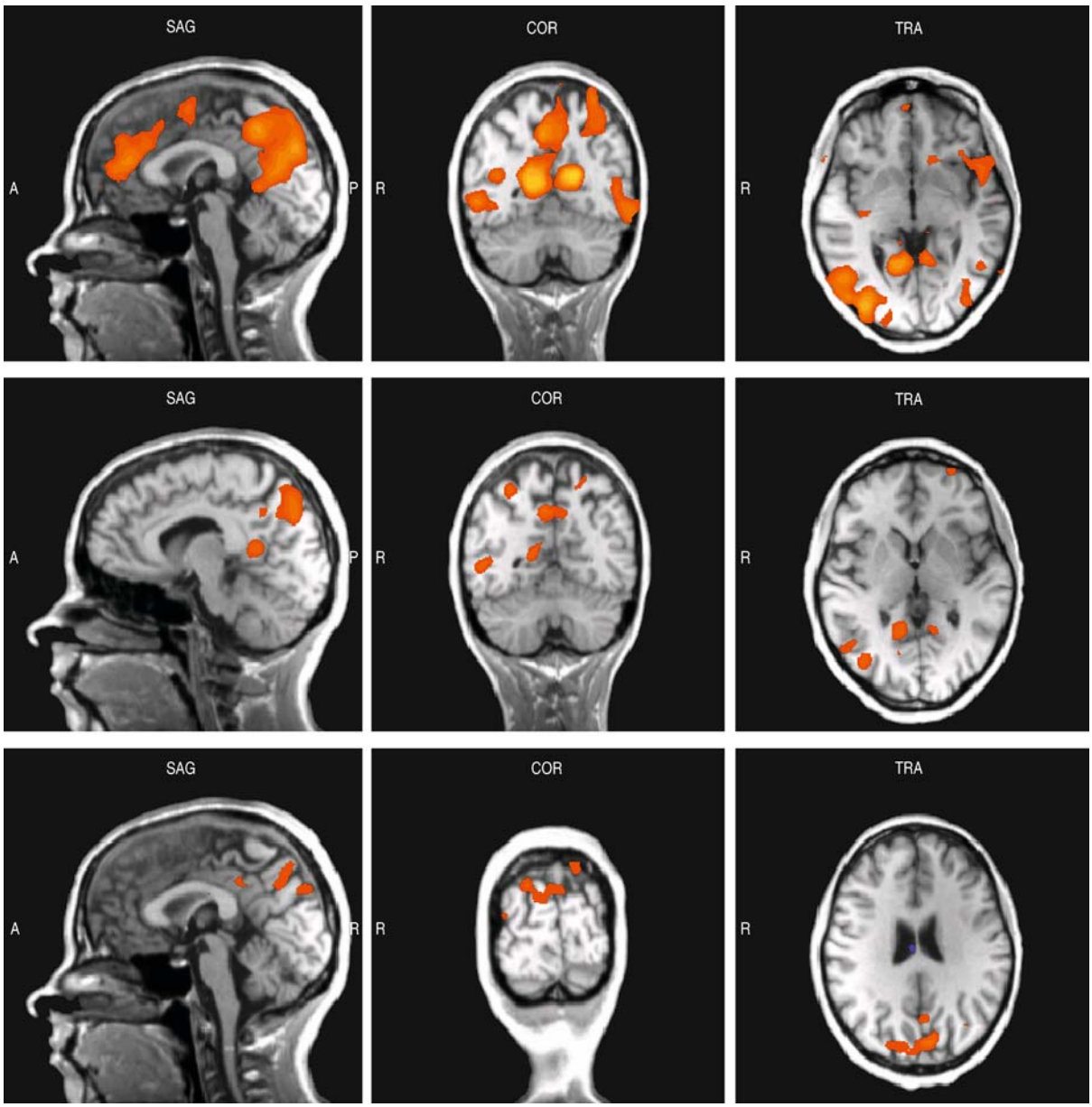
ries, but just before the flooding was started. (Flooding or response prevention is an essential therapy technique in the treatment of OCD, where patients are confronted with symptom provoking stimuli but abstain from performing compulsive rituals.) The instability of the patient's process was the precursor of an important personal decision to divorce from her husband. (It should be noted that the development of her OCD symptoms was in the context of a long-lasting marital conflict.) This decision was the essential phase transition of the therapy.

Indeed, the most pronounced changes in brain activity occurred from the first to the second fMRI scan, whereas BOLD response differences from the second to the third session were only slight. They perhaps represent the neural correlates of an important personal phase transition related to the resolution of a severe personal conflict. Because these changes occurred before the flooding procedure was started, this can be seen as indicator of an early rapid response in the therapy [29]. Additionally, marked alternations in brain activity were to be observed before or during symptom reduction took place (measured by the Y-BOCS), not afterwards.

Alternations in brain activity involved widespread areas, e. g. the medial frontal brain regions including anterior cingulate cortex, superior and middle frontal gyrus, inferior frontal and precentral gyrus, superior temporal gyrus, superior parietal lobe, cuneus, thalamus and caudate nucleus in both hemispheres, as well as the right fusiform gyrus (see Fig. 9 for a OCD to disgust contrast). Thalamic and basal ganglia activation is part of the dorsolateral-caudate-striatum-thalamus circuitry of OCD. Especially the caudate nucleus takes a role within the executive dysfunction model of compulsions, and its activity has been found to be reduced after treatment [40].

The function of the anterior cingulate cortex is interesting with regard to synergetics. The cingulate cortex comprises various functions like somatosensoric integration, mediation of affective and cognitive processes, control of attention, and processing of painful stimuli. Additionally, it plays an important role as conflict monitoring system: It is sensitive to ambiguous or conflicting information [81,82], is involved in decision making [24,54], and its activation is predictive to treatment outcome in depression (e. g. [37]). This is true especially for the dorsal (cognitive) structures of the ACC. It could be an indicator of symmetry states of brain functioning, which is characterized by two or more dynamic patterns or attractors in competition. In the present case, the ACC activation at the beginning of the therapy could be either part of the pathology or could be indicative for the critical instability of the cognitive-affective system of the patient, preparing her important decision. The second fMRI measure was conducted during a local minimum of critical fluctuations. If the impressive change in cingulate activation could be attributed to a changed critical symmetry state of the neural self-organization before vs. after the phase-transition or to changes in symptom severity cannot be decided within a single case study, but seems to be an interesting question to further research. Perhaps the fact that during the second fMRI measure the Y-BOCS score was nearly on the same level as during the first measure – only 14% reduction, compared to 50% reduction in dynamic complexity – could be a first argument in favor of the instability hypothesis.

The paradigm of self-organization is a very promising approach to clinical as well as other fields of psychology. Its interdisciplinary is due to the fact that the laws and principles of self-organization are true for neural, mental, and behavioral processes (and the corresponding data qualities). Interdisciplinary cooperation is underpinned by the unifying terminology as well as by the unifying formalism and modeling tools of synergetics. This opens new perspectives for basic and applied research, but also for the treatment of mental disorders. New developments in the real-time monitoring of human change processes based on synergetics and nonlinear science have been mentioned. Another field of encouraging developments is deep brain stimulation (DBS), which apply to neurological diseases as Parkinsonian or essential tremor, but also to psychiatric disorders as OCD or mayor depression [75]. The difference between new technologies (applying the mathematical instruments and concepts of synergetics as well as methods from stochastic phase resetting) and classical electrical deep brain stimulation is that normal DBS at high frequencies has a blocking effect on the stimulated target and mimics the effect of tissue lesioning. New technologies are demand-controlled, working with low stimulation frequencies, and avoid the suppression of neurons' firing. Its effect is a desynchronization of pathologically synchronized populations of neurons, using multi-site coordinated reset stimulation [74] or nonlinear delayed feedback stimulation [50]. Both methods counteract abnormal interactions and detune the macroscopic frequency of the collective oscillators – that is the abnormally established order parameters of neural synchronization. Thereby they restore the natural frequencies of the individual oscillatory units. Neurons get in the range of physiological functioning and can engage in changing and varying synchronization patterns. If altered synchronization patterns also change the coupling strength connecting synapses, a rewiring of neural nets could be reached. Changed function triggers the emergence of healthy at-



Self-Organization in Clinical Psychology, Figure 9
Brain activation patterns of a patient with OCD during psychotherapy. BOLD signals from a 1.5 Tesla fMRI scanner. *Top*: First scan (9th day of hospital stay; $x = 0$, $y = -55$, $z = -2$; p(uncor) <0.001). *Middle*: Second scan (30th day of hospital stay; $x = 8$, $y = -54$, $z = 5$; p(uncor) < 0.001). *Bottom*: The third scan (57th day of hospital stay; $x = 0$, $y = -85$, $z = 26$; p(uncor) < 0.001). Activations during the presentation of OCD-related pictures compared to activations during the presentation of neutral pictures (OCD > disgust)

tractors and by this changes the structure of neural networks. Perhaps in the future technologies of DBS or even non-invasive brain stimulation could be combined with psychotherapy and psychological navigation instruments developed to optimize self-organizing change processes.

Future Directions

The future developments of self-organization and complexity research in clinical psychology and psychotherapy will be interconnected to its acceptance in practice and

training. Perhaps this sounds paradoxically, since in most other scientific fields the future depends on the investigations to basic research and to new technologies. Of course this holds also for synergetics and its applications to clinical psychology. However, it should be noted that complexity research and nonlinear dynamics are done since more than two decades in European academic psychology with poor impact to mainstream science. So, the future will depend on a greater number of new arriving and highly qualified students in this topic who do not avoid the touch with mathematics. Self-organization and complexity research including its mathematical backgrounds should become part of the training curricula in psychology and psychotherapy. Since the Synergetic Navigation System waits for its broad application in clinical and psychotherapeutic practice, a new decade of practice-based research can be started. But these developments depend on its acceptance by practitioners because of the competencies required for the widespread use of sophisticated methods. This integration of science with practice will open huge sources and new dimensions of data gathering on dynamic systems. An important database for outcome and time series data (including biomarkers) of human change processes is actually prepared.

Another stream of development is concerning the integration of psychological and biological/physiological data. Since human self-organization takes place on synchronized mental, social, and biological system levels, all of them should be taken into account in further research. One research paradigm was suggested in this chapter: The investigation of individual and social processes by the Synergetic Navigation System, and in parallel repeated brain scans using fMRI technology or other methods to get insight into brain dynamics (EEG, gene expression markers [25], immune or endocrine markers [61], or others). Two final remarks: First, future developments of synergetic-based minimal invasive DBS could be combined with psychotherapy and psychological interventions – as pharmacological and psychological treatments are combined nowadays. Second, the nonlinear networks underlying psychological as well as neural self-organization will not be understood without applying appropriate mathematical tools, giving raise to a new systemic psychology and neuroscience.

Bibliography

1. an der Heiden U (1992) Chaos in health and disease – phenomenology and theory. In: Tschacher W, Schiepek G, Brunner EJ (eds) Self-organization and clinical psychology. Springer, Berlin, pp 55–87
2. Bak P, Chen K, Creutz M (1989) Self-organized criticality and the 'Game of Life'. Nature 342:780–782
3. Basar E, Flohr H, Haken H, Mandell AJ (1983) Synergetics of the brain. Springer Series in Synergetics, vol 23. Springer, Berlin
4. Basar-Eroglu C, Strüber D, Kruse P, Basar E, Stadler M (1996) Frontal gamma-band enhancement during multistable visual perception. Int J Psychophysiol 24:113–125
5. Bethe A (1940) Die biologischen Rhythmusvorgänge als selbständige und erzwungene Kippvorgänge betrachtet. Pflügers Arch 244:1–42
6. Böker W, Brenner HD (1996) Stand systemischer Modellvorstellungen zur Schizophrenie und Implikationen für die Therapieforschung. In: Böker W, Brenner HD, Genner RM (eds) Integrative Therapie der Schizophrenie. Huber, Bern, pp 17–32
7. Böker W, Brenner HD (eds) (1989) Schizophrenie als systemische Störung. Huber, Bern
8. Cannon WB (1915) Bodily changes in pain, hunger, fear, and rage: An account of recent researches into the function of emotional excitement. Appleton, New York
9. Carlsson A (2006) The neurochemical circuitry of schizophrenia. Pharmacopsychiatry 39(Suppl 1):S10–S14
10. Caspar F (1996) Beziehungen und Probleme verstehen. Eine Einführung in die psychotherapeutische Plananalyse. Huber, Bern
11. Ciompi L (1989) Zur Dynamik komplexer biologisch-psychosozialer Systeme: Vier fundamentale Mediatoren in der Langzeitentwicklung der Schizophrenie. In: Böker W, Brenner HD (eds) Schizophrenie als systemische Störung. Huber, Bern, pp 27–38
12. Ciompi L, Müller C (1976) Lebensweg und Alter der Schizophrenen. Springer, Berlin
13. Eckmann JP, Oliffson Kamphorst S, Ruelle D (1987) Recurrence plots of dynamical systems. Europhys Lett 4:973–977
14. Friedlander L, Desrocher M (2006) Neuroimaging studies of obsessive-compulsive disorder in adults and children. Clin Psychol Rev 26:32–49
15. Friston KJ, Harrison L, Penny WD (2003) Dynamic causal modelling. NeuroImage 19:1273–1302
16. Haken H (1990) Synergetics – an introduction. Nonequilibrium phase transitions in physics, chemistry, and biology. Springer, Berlin (first edition 1977)
17. Haken H (1990) Synergetics as a tool for the conceptualization and mathematization of cognition and behavior – How far can we go? In: Haken H, Stadler M (eds) Synergetics of cognition. Springer, Berlin, pp 2–31
18. Haken H (1996) Principles of brain functioning. A synergetic approach to brain activity, behavior, and cognition. Springer, Berlin
19. Haken H (2002) Brain dynamics. Springer, Berlin
20. Haken H (2004) Synergetics. Introduction and advanced topics. Springer, Berlin
21. Haken H, Schiepek G (2006) Synergetik in der Psychologie. Selbstorganisation verstehen und gestalten. Hogrefe, Göttingen
22. Haken H, Kelso JAS, Bunz H (1985) A theoretical model of phase transition in human hand movements. Biol Cybern 51:347–356
23. Kettunen J, Keltikangas-Järvinen L (2001) Intraindividual analysis of instantaneous heart rate variability. Psychophysiology 38:659–668

24. King-Casas B, Tomlin D, Anen C, Camerer CF, Quartz SR, Montague PR (2005) Getting to know you: Reputation and trust in a two-person economic exchange. Science 308:78–83
25. Koch JM, Kell S, Hinze-Selch D, Aldenhoff JB (2002) Changes in CREB-phosphorylation during recovery from major depression. J Psychiat Res 36:369–375
26. Kowalik ZJ, Schiepek G, Kumpf K, Roberts LE, Elbert T (1997) Psychotherapy as a chaotic process II: The application of nonlinear analysis methods on quasi time series of the client-therapist-interaction: A nonstationary approach. Psychother Res 7:197–218
27. Kruse P, Carmesin HO, Stadler M (1997) Schizophrenie als Korrespondenzproblem plastischer neuronaler Netze. In: Schiepek G, Tschacher W (eds) Selbstorganisation in Psychologie und Psychiatrie. Vieweg, Braunschweig, pp 171–190
28. Köhler W (1947) Gestalt psychology. Liveright, New York
29. Lambert MJ, Ogles BM (2004) The efficacy and effectiveness of psychotherapy. In: Lambert MJ (ed) Bergin and Garfield's handbook of psychotherapy and behavior change. Wiley, New York, pp 139–193
30. Lambert MJ, Whipple JL, Smart DW, Vermeersch DA, Nielsen SL, Hawkins EJ (2001) The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? Psychother Res 11:49–68
31. Lambert MJ, Whipple JL, Vermeersch DA, Smart DW, Hawkins EJ, Nielsen SL, Goates M (2002) Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. Clin Psychol Psychother 9:91–103
32. Lambertz M, Vandenhouten R, Grebe R, Langhorst P (2000) Phase transitions in the common brainstem and related systems investigated by nonstationary time series analysis. J Autonom Nerv Syst 78:141–157
33. Lambertz M, Vandenhouten R, Langhorst P (2003) Transiente Kopplungen von Hirnstammneuronen mit Atmung, Herzkreislaufsystem und EEG: Ihre Bedeutung für Ordnungsübergänge in der Psychotherapie. In: Schiepek G (ed) Neurobiologie der Psychotherapie. Schattauer, Stuttgart, pp 302–324
34. Lenz F, Kwan H, Martin R, Tasker R, Dostrovsky J, Lenz Y (1994) Single unit analysis of the human ventral thalamic nuclear group. Tremor-related activity in functionally identified cells. Brain 117:531–543
35. Lewin K (1951) Field theory in social psychology. Harper, New York
36. Lutzenberger W, Elbert T, Birbaumer N, Ray WJ, Schupp H (1992) The scalp distribution of the fractal dimension of the EEG and its variation with mental tasks. Brain Topogr 5:27–34
37. Mayberg HS, Brannan SK, Mahurin RK, Jerabek PA, Brickman JS, Tekell JL, Silva JA, McGinnis S, Glass TG, Martin CC, Fox PT (1997) Cingulate function in depression: A potential predictor of treatment response. Neuroreport 8:1057–1061
38. Mergenthaler E (1998) Cycles of emotion-abstraction patterns: A way of practice oriented process research? Br Psychol Soc – Psychother Sect Newsl 24:16–29
39. Meyer-Lindenberg A, Ziemann U, Hajak G, Cohen L, Faith Berman K (2002) Transitions between dynamical states of differing stability in the human brain. PNAS USA 99:10948–10953
40. Nakao T, Nakagawa A, Yoshiura T, Nakatani E, Nabeyama M, Yoshizato C, Kudoh A, Tada K, Yoshioka K, Kawamoto M, Togao O, Kanba S (2005) Brain activation of patients with obsessive-compulsive disorder during neuropsychological and symptom provocation tasks before and after symptom improvement: A functional magnetic resonance imaging study. Biol Psychiatry 57:901–910
41. Nowak A, Vallacher RR (1998) Dynamical social psychology. Guilford, New York
42. Orlinsky DE, Ronnestad MH, Willutzki U (2004) Fifty years of psychotherapy process-outcome research: Continuity and change. In: Lambert MJ (ed) Bergin and Garfield's handbook of psychotherapy and behavior change. Wiley, New York, pp 307–389
43. Pare D, Curro'Dossi R, Steriade M (1990) Neuronal basis of the parkinsonian resting tremor: A hypothesis and its implications for treatment. Neurosci 35:217–226
44. Penny WD, Stephan KE, Mechelli A, Friston KJ (2004) Modelling functional integration: A comparison of structural equation and dynamic causal models. NeuroImage 23(Suppl 1):264–274
45. Perlitz V, Cotuk B, Haberstock S, Kahn N, Grebe R, Petzold ER, Schmid-Schönbein H (2003) Differentiation of cutaneous haemo- and neurodynamics using multiscale Time-Frequency-Distribution portrays. In: Blazek V, Schultz-Ehrenburg U (eds) Proceedings of the 10th international symposium CNVD 2001 at Aachen. Germany
46. Perlitz V, Cotuk B, Lambertz M, Grebe R, Schiepek G, Petzold ER, Schmid-Schönbein H, Flatten G (2004) Coordination dynamics of circulatory and respiratory rhythms during psychomotor relaxation. Autonom Neurosci 115(1–2):82–93
47. Perlitz V, Cotuk B, Schiepek G, Sen A, Haberstock S, Schmid-Schönbein H, Petzold ER, Flatten G (2004) Synergetik der hypnoiden Relaxation. Psychother Psych Med 54:250–258
48. Perlitz V, Lambertz M, Cotuk B, Grebe R, Vandenhouten R, Flatten G, Petzold ER, Schmid-Schönbein H, Langhorst P (2004) Cardiovascular rhythms in the 0.15 Hz band: Common origin of identical phenomena in man and canine in the reticular formation of the brain stem? Pflügers Arch – Europ J Physiol 448(6):579–592
49. Piaget J (1976) Die Äquilibration der kognitiven Strukturen. Klett-Cotta, Stuttgart
50. Popovych OV, Hauptmann C, Tass PA (2006) Control of neural synchrony by nonlinear delayed feedback. Biol Cybern 95:69–85
51. Rapp PE, Albano ME, Zimmerman ID et al (1994) Phase-randomized surrogates can produce spurious identifications of non-random structure. Phys Lett A 192:27–33
52. Rockstroh B, Watzl H, Kowalik ZJ, Cohen R, Sterr A, Müller M, Elbert T (1997) Dynamical aspects of the EEG in different psychopathological states in an interview situation. A pilot study. Schizophr Res 28:77–85
53. Rosenstein MT, Collins JJ, de Luca CJ (1993) A practical method for calculating Largest Lyapunov Exponents from small data sets. Physica D 65:117–134
54. Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD (2003) The neural basis of economic decision-making in the Ultimatum Game. Science 300:1755–1758
55. Saxena S, Rauch SL (2000) Functional neuroimaging and the neuroanatomy of obsessive-compulsive disorder. Psychiatr Clin North Am 23:563–586
56. Schaub H, Schiepek G (1992) Simulation of psychological processes: Basic issues and an illustration within the etiology of a depressive disorder. In: Tschacher W, Schiepek G, Brunner EJ (eds) Self-organization and clinical psychology. Springer, Berlin, pp 121–149

57. Schiepek G, Schoppek W, Tretter F (1992) Synergetics in psychiatry: Simulation of evolutionary patterns of schizophrenia on the basis of nonlinear difference equations. In: Tschacher W, Schiepek G, Brunner EJ (eds) Self-organization and clinical psychology. Springer, Berlin, pp 163–194
58. Schiepek G, Kowalik ZJ, Schütz A, Köhler M, Richter K, Strunk G, Mühlnickel W, Elbert T (1997) Psychotherapy as a chaotic process I. Coding the client-therapist-interaction by means of sequential plan analysis and the search for chaos: A stationary approach. Psychother Res 7:173–194
59. Schiepek G, Tominschek I, Karch S, Mulert C, Pogarell O (2007) Neurobiologische Korrelate der Zwangsstörungen – Aktuelle Befunde zur funktionellen Bildgebung. Psychother Psych Med 57:379–394
60. Schiepek G, Tominschek I, Karch S, Lutz J, Mulert C, Born C, Pogarell O (2008) A controlled single case study with repeated fMRI measures during the treatment of a patient with obsessive-compulsive disorder: Testing the nonlinear dynamics approach to psychotherapy. World J Biol Psychiatry. doi:10.1080/15622970802311829
61. Schubert C, Schiepek G (2003) Psychoneuroimmunologie und Psychotherapie: Psychosozial induzierte Veränderungen der dynamischen Komplexität von Immunprozessen. In: Schiepek G (ed) Neurobiologie der Psychotherapie. Schattauer, Stuttgart, pp 485–508
62. Schupp HAT, Lutzenberger W, Birbaumer N, Miltner W, Braun C (1994) Neurophysiological differences between perception and imagery. Cog Brain Res 2:77–86
63. Selye HA (1936) Syndrome produced diverse nocuous agents. Nature 138:32
64. Singer W, Gray CM (1995) Visual feature integration and the temporal correlation hypothesis. Ann Rev Neurosci 18:555–586
65. Sinha R, Lovallo WR, Parsons OA (1992) Cardiovascular differentiation of emotions. Psychosom Med 54:422–435
66. Skinner JE, Molnar M, Tomberg C (1994) The point correlation dimension: Performance with nonstationary surrogate data and noise. Int Physiol Behav Sci 29:217–234
67. Stadler M, Kruse P (1990) The self-organization perspective in cognition research. Historical remarks and new experimental approaches. In: Haken H, Stadler M (eds) Synergetics of cognition. Springer, Berlin, pp 32–52
68. Strauss JS (1989) Intermediäre Prozesse in der Schizophrenie: Zu einer neuen dynamisch orientierten Psychiatrie. In: Böker W, Brenner HD (eds) Schizophrenie als systemische Störung. Huber, Bern, pp 39–50
69. Strunk G (2004) Organisierte Komplexität. Mikroprozess-Analysen der Interaktionsdynamik zweier Psychotherapien mit den Methoden der nichtlinearen Zeitreihenanalyse. Dissertation, Universität Bamberg
70. Strunk G, Schiepek G (2006) Systemische Psychologie. Eine Einführung in die komplexen Grundlagen menschlichen Verhaltens. Spektrum Akademischer Verlag, Heidelberg
71. Tallon-Baudry C, Bertrand O (1999) Oscillatory gamma activity in humans and its role in object representation. Trends Cog Sci 3:151–162
72. Tallon-Baudry C, Bertrand O, Wienbruch C, Ross B, Pantev C (1997) Combined EEG and MEG recordings of visual 40 Hz resonses to illusory triangles in human. Neuroreport 8:1103–1107
73. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996) Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Eur Heart J 17:354–381
74. Tass P, Hauptmann C (2007) Therapeutic modulation of synaptic connectivity with desynchronizing brain stimulation. Int J Psychophysiol 64:53–61
75. Tass PA, Klosterkötter J, Schneider F, Lenartz D, Kouluosakis A, Sturm V (2003) Obsessive-compulsive disorder: Development of demand-controlled deep brain stimulation with methods of stochastic phase resetting. Neuropsychopharmacology 28:S27–S34
76. Tretter F, Scherer J (2006) Schizophrenia, neurobiology, and the methodology of systemic modelling. Pharmacopsychiatry 39(Suppl 1):S26–S35
77. Tschacher W (1997) Prozessgestalten. Hogrefe, Göttingen
78. Tschacher W, Kupper Z (2002) Time series models of symptoms in schizophrenia. Psychiatry Res 113:127–137
79. Tschacher W, Dauwalder JP (1999) Situated cognition, ecological perception, and synergetics: A novel perspective for cognitive psychology? In: Tschacher W, Dauwalder JP (eds) Dynamics, synergerics, autonomous agents. World Scientific, Singapore, pp 83–104
80. Vandenhouten R (1998) Analyse instationärer Zeitreihen komplexer Systeme und Anwendungen in der Physiologie. Shaker, Aachen
81. van Veen V, Carter CC (2002) The anterior cingulate as a conflict monitor: fMRI and ERP studies. Physiol Behav 77:477–482
82. van Veen V, Carter CC (2002) The timing of action-monitoring processes in the anterior cingulate cortex. J Cog Neurosci 14:593–602
83. Villmann T, Liebers C, Bergmann B, Gumz A, Geyer M (2008) Investigation of psycho-physiological interactions between patient and therapist during a psychodynamic therapy and their relation to speech in terms of entropy analysis using a neural network approach. New Ideas in Psychology 26:309–325
84. von Uexküll T, Wesiack W (1996) Wissenschaftstheorie: Ein biopsycho-soziales Modell. In: Adler RH, Herrmann JM, Köhle K, Schonecke OW, von Uexküll T, Wesiack W (eds) Thure von Uexküll. Psychosomatische Medizin. Urban Schwarzenberg, München, pp 13–52
85. Webber CL, Zbilut JP (1994) Dynamical assessment of physiological systems and states using recurrence plot strategies. J Appl Physiol 76:965–973
86. Yuru Z, Jan KM, Ju KH, Chon KH (2006) Quantifying cardiac sympathetic and parasympathetic nervous activities using principal dynamic modes analysis of heart rate variability. Am J Physiol Heart Circ Physiol 291:H1475–H1483

Self-Organization in Magnetohydrodynamic Turbulence

Pierluigi Veltri, Vincenzo Carbone, Fabio Lepreti, Giusy Nigro
Dipartimento di Fisica, Università della Calabria, Arcavacata di Rende, Italy

Article Outline

Glossary
Definition of the Subject
Introduction
MHD Equations and Quadratic Invariants
Self-Organization at Large Scales
Self-Organization at Small Scales and Intermittent Structures in MHD Turbulence
Future Directions
Bibliography

Glossary

Alfvén speed The Alfvén speed c_A is the propagation speed of Alfvén waves and is given by $c_A = B_0/(4\pi\rho)^{1/2}$, where $\mathbf{B}_0$ is mean magnetic field and ρ the plasma mass density. Alfvèn waves are transverse, incompressible magnetohydrodynamic waves that propagate along $\mathbf{B}_0$ and originate from the tension of magnetic field lines.

Elsässer variables Elsässer variables $\mathbf{z}^\sigma$ are defined by $\mathbf{z}^\sigma = \mathbf{v} + \sigma\mathbf{B}/(4\pi\rho)^{1/2}$, with $\sigma = \pm 1$, $\mathbf{v}$ the velocity field, $\mathbf{B}$ the magnetic field and ρ the plasma mass density. The equations of incompressible MHD are often written in terms of these variables in order to describe the propagation of Alfvén waves and the non-linear couplings occurring in MHD turbulence.

Magnetohydrodynamics Magnetohydrodynamics (abbreviated, MHD) represents a one-fluid mathematical model which describes plasma dynamics at low frequencies: The main dynamical variables are the velocity of the fluid and the magnetic field. The vector equations for these variables are the fluid momentum conservation and the induction Maxwell equation.

MHD turbulence MHD turbulence is that turbulence which develops inside plasmas at macroscopic level, when viscosity and resistivity are low. Apart from velocity fluctuations which are also present in ordinary fluids, it is characterized also by the presence of magnetic field fluctuations.

Reverse field pinch Reverse field pinch (abbreviated, RFP) are plasma fusion toroidal devices whose conception is based on the idea that non-linear interactions in plasmas spontaneously give rise to magnetic structures where the Laplace force is zero (force free structures).

Shell models Shell models of turbulence are dynamical systems consisting of a set of ordinary differential equations representing a simplified version of the Navier–Stokes or MHD equations in the wavevector space. These models provide the possibility to investigate turbulence at very high Reynolds number regimes at the cost of neglecting information about spatial structures.

Solar corona The solar corona is the region extending from the solar surface up to one million of kilometers in the space, which can be visible to the naked eye during the eclipses. It is constituted mainly by a hydrogen plasma (proton and electrons) at a temperature of about two million degrees. The corona is highly structured by the magnetic field generated at the sun surface.

Solar wind The solar wind is a stream of plasma mainly composed of protons and electrons (hydrogen plasma), which flows out of the sun, due to the fact that plasma pressure associated to the very high coronal temperature overcomes the sun gravity. The flow velocity ranges from 250 km/s in the equatorial plane to about 900 km/s in the polar regions. Solar wind represents an extremely efficient plasma laboratory where the turbulence associated with the supersonic flow can be studied using space experiment data.

Definition of the Subject

Plasma dynamics at low frequency, i. e. at frequencies lower then the ion cyclotron frequency, can be described using a one-fluid model usually called magnetohydrodynamics (MHD), where the main dynamical variables are represented by fluid velocity and magnetic field, which evolve non-linearly being coupled to each other. This description applies both to laboratory plasma devices (tokamaks, reverse field pinch etc.), devoted to realize controlled nuclear fusion, and to space and astrophysical plasmas (Solar corona, Solar wind). Very often, when viscosity and resistivity are sufficiently small, the plasma behavior is characterized by the presence of a developed turbulence. In the last 20 years huge progress in understanding the properties of such turbulence has been realized, both by the use of high resolution computer simulations and by analysis of space and laboratory data. One of the most fascinating results of these studies concerns the evidence of self organization processes which have been shown to be very effective inside this kind of turbulence. Other aspects of turbulence and dynamical complexity in space plasmas are considered in the review by Chang ▶ Space Plasmas, Dynamical Complexity in in this Encyclopedia.

Introduction

Dynamical systems, whose time evolution is described by non-linear equations, can often give rise to chaotic behavior, i. e. to a behavior characterized by a strong dependence on initial conditions, which, in all practical cases,

does not allow to predict the final fate of such systems, or at least limits considerably the possibility to make predictions. Nevertheless, inside these systems, sometimes new complex, coherent structures, without any need of external intervention, can develop. This phenomenon is called self-organization.

Self-organization requires decreasing the entropy inside the coherent organized structures and exporting such entropy to the surroundings. For this reason the coherent structures can be called, following Prigogine, *dissipative structures*, in that their organization is always associated with an increased efficiency of dissipative processes around them.

MHD turbulence, as well as ordinary fluid turbulence, is characterized by a phenomenology, first described some centuries ago around the year 1500 by Leonardo da Vinci (Piumati 1894, fo. 74,v), where energy injected at large scales (injection range) is transferred by non-linear interactions (inertial range) towards smaller and smaller scales where it is finally dissipated (dissipative range). The physical phenomenon by which coherent organized structures are produced through an increased dissipation inside MHD turbulence has been evidenced both at large scale and at small scales. In the latter case it is usually called *intermittency*. We will discuss separately the two cases.

The occurrence of self organization at large scale seems to be directly related to the existence in MHD equations of some quadratic invariants (energy, magnetic helicity, cross helicity), which are conserved inside inviscid flows and are dissipated at very different rates in dissipative flows. Actually the existence of some long-living non-trivial states, towards which dissipative flows are attracted, has been evidenced for a long time. Moreover it has been suggested that these states could be derived from a variational principle, i. e. by minimizing an energy integral subject to some constraints on the other quadratic invariants. Woltjer [58] first explored the astrophysical implications of this possibility. The conjecture by Taylor [51] that an MHD system relaxes towards a state where the energy is minimum, subject to the constraint that magnetic helicity is conserved, allowed one to obtain a quite particular solution, usually called *Taylor's vortex*. This solution was extremely useful to explain the large scale behavior of reverse field pinch (RFP) devices [51]. The same conjecture has also been used to estimate the energy release in coronal structures [25,42] in connection to the problem of coronal heating.

The discovery in the solar wind of quite peculiar *Alfvénic fluctuations*, characterized by a high degree of correlation between velocity and magnetic field [3], which was interpreted by Dobrowolny et al. [20] as the result of a self organization of MHD turbulence, has suggested the formulation [10,38,53] of a different minimum principle: Alfvénic solutions can be obtained by minimizing the energy subject to the condition that cross-helicity is conserved.

In any case the energy minimization is dynamically realized through a *selective decay process*, i. e. a process during which some energy is dissipated into heat and the appearance of the coherent large scale structures is associated with redistribution of entropy to the surroundings.

The self organization processes discussed above, and related to the spontaneous creation of large scale coherent structures during the MHD turbulence dynamics, have their counterpart also at small scales. Actually, the nonlinear energy cascade process in fully developed MHD turbulence is characterized by the formation at the smallest dissipative scale of coherent structures. These structures are strictly related to the intermittency phenomenon; that is, to the breakdown of global self-similarity in the turbulent cascade, which represents a typical signature of nonlinear interactions.

In the following sections we first present the MHD equations (Sect. "MHD Equations and Quadratic Invariants") properties and discuss the characteristics of the self organized structures at large scales (Sect. "Self-Organization at Large Scales"). Then in Sect. "Self-Organization at Small Scales and Intermittent Structures in MHD Turbulence" we show that at small scales coherent dissipative structures are spontaneously formed in a physical process usually called *intermittency*. Finally in Sect. "Future Directions" we discuss the possible future developments and perspectives of such kind of studies.

MHD Equations and Quadratic Invariants

The equations describing the time evolution of an incompressible magnetofluid can be written

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} = -\frac{1}{\rho}\nabla p + (\nabla\times\mathbf{b})\times\mathbf{b} + \nu\nabla^2\mathbf{v}\,, \quad (1)$$

$$\frac{\partial \mathbf{b}}{\partial t} = \nabla\times(\mathbf{v}\times\mathbf{b}) + \mu\nabla^2\mathbf{b} \quad (2)$$

where $\mathbf{v}$ represents the fluid velocity and p the fluid pressure, $\mathbf{b} = \mathbf{B}/\sqrt{4\pi\rho}$ ($\mathbf{B}$ being the magnetic field and ρ the mass density), ν is the kinematic viscosity and $\mu = (c^2\eta)/(4\pi)$ the magnetic diffusivity (η being the fluid resistivity). The velocity and magnetic field must also satisfy the incompressibility conditions

$$\nabla\cdot\mathbf{v} = 0 \qquad \nabla\cdot\mathbf{b} = 0\,. \quad (3)$$

Quadratic Invariants

It can be easily shown that during their ideal time evolution, i. e. in absence of viscosity and resistivity, these equations conserve the following quadratic invariants: Total energy (kinetic + magnetic) per mass unit

$$E = \frac{1}{2}\int (\mathbf{v}^2 + \mathbf{b}^2)\mathrm{d}V \tag{4}$$

and cross helicity

$$H_c = \int \mathbf{v}\cdot\mathbf{b}\mathrm{d}V\,. \tag{5}$$

Moreover Woltjer [58] showed that in the ideal 3D MHD, also magnetic helicity

$$H_m = \int \mathbf{a}\cdot\mathbf{b}\mathrm{d}V \tag{6}$$

(**a** being the magnetic vector potential defined through $\mathbf{b} = \nabla\times\mathbf{a}$) remains invariant during the evolution of any closed magnetic flux system. In 2D MHD on the contrary, the third conserved quantity is the square of the magnetic field vector potential [24]

$$A = \int \mathbf{a}^2\mathrm{d}V \tag{7}$$

In Eqs. (4), (5), (6) and (7) the integrals are extended to the whole volume of the magnetofluid.

Elsässer's Variables

In the following we will find it useful to rewrite these equations in terms of the Elsässer [21] variables $\mathbf{z}^\sigma$ defined by $\mathbf{z}^\sigma = \mathbf{v} + \sigma\mathbf{b}$ with $\sigma = \pm 1$. Using such variables the equations governing incompressible MHD are recast to the more compact form

$$\frac{\partial \mathbf{z}^\sigma}{\partial t} + (\mathbf{z}^{-\sigma}\cdot\nabla)\mathbf{z}^\sigma = -\frac{1}{\rho}\nabla\left(p + \frac{B^2}{8\pi}\right) + \nu^+\nabla^2\mathbf{z}^\sigma + \nu^-\nabla^2\mathbf{z}^{-\sigma} \tag{8}$$

with $\nu^\sigma = \nu + \sigma\mu$. Obviously also the fields $\mathbf{z}^\sigma$ must satisfy the incompressibility conditions

$$\nabla\cdot\mathbf{z}^\sigma = 0\,. \tag{9}$$

In terms of these variables, the conservation of total energy per mass unit and cross helicity can be written as the conservation of the two pseudo-energies

$$E^\sigma = \frac{1}{2}\int \mathbf{z}^{\sigma^2}\mathrm{d}V\,. \tag{10}$$

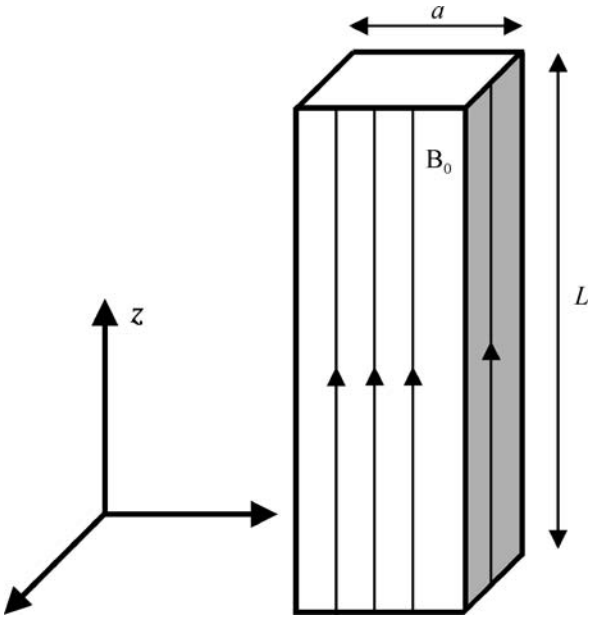
MHD equations display and infinite number of ideal non-quadratic invariants, but quadratic invariants (4), (5) and (6) (or in 2D (7)) play a key role. Actually, let us introduce the Fourier transform of the field $\mathbf{z}^\sigma$ in a cubic periodic box of size a

$$\mathbf{z}^\sigma(\mathbf{r}, t) = \sum_{\mathbf{k}} \mathbf{z}^\sigma(\mathbf{k}, t)\exp(i\mathbf{k}\cdot\mathbf{r})$$

where $\mathbf{k} = 2\pi\mathbf{m}/a$, and $\mathbf{m}$ is vector of integer numbers. In terms of these Fourier transforms (taking into account the incompressibility condition (9) and neglecting dissipative terms), MHD Equations (8) can be written

$$\frac{\partial z_i^\sigma(\mathbf{k}, t)}{\partial t} = \sum_{l,s=1}^{3}\sum_{\mathbf{p},\mathbf{q}} M_{i,l,s}(\mathbf{k},\mathbf{p},\mathbf{q}) z_l^\sigma(\mathbf{p}, t)\, z_s^{-\sigma}(\mathbf{q}, t)\delta_{\mathbf{k},\mathbf{p}+\mathbf{q}} \tag{11}$$

with $M_{i,l,s}(\mathbf{k},\mathbf{p},\mathbf{q}) = -i\left(\delta_{il} - \frac{k_i k_l}{k^2}\right)k_s$ and where $\delta_{\mathbf{k},\mathbf{p}+\mathbf{q}}$ is the usual Kronecker symbol. The peculiar form of these equations shows that non-linear terms are characterized by triad interactions among wave vectors **k**, **p** and **q** such that $\mathbf{k} = \mathbf{p} + \mathbf{q}$. Quadratic invariants are also called *rugged* invariants, because they are conserved in any single interaction among a triad of wave vector and for this reason



Self-Organization in Magnetohydrodynamic Turbulence, Figure 1
Sketch of the reduced MHD configuration

they survive also when a Galerkin approximation is performed on MHD equations. In other words, even if the r. h. s. of (11) contains an infinite number of non-linear interactions, $E(t)$, $H_c(t)$ and $H_m(t)$ remain invariant also by retaining any finite number of interacting triads in Equation (11).

Reduced Magnetohydrodynamics

In some fusion plasma devices, or in the solar corona, the plasma β parameter (that is the ratio between the kinetic and magnetic pressure) is low $\beta \simeq 10^{-2}$, and MHD equations can be simplified into the so-called Reduced MHD [47,61]. This approximation is valid, for example, for a plasma column with a low aspect ratio $a/L \ll 1$ (a and L being respectively the width and the height of the plasma column), with a strong magnetic field $\mathbf{B}_0$ along the z-direction (Fig. 1).

In this case the plasma dynamical variables reduce to the velocity field $\mathbf{v}_\perp$ and the magnetic field $\mathbf{b}_\perp$ perpendicular to $\mathbf{B}_0$, so that starting from MHD Equs. (8), the following set of RMHD equations [61] is obtained:

$$\frac{\partial \mathbf{z}^\sigma}{\partial t} - \sigma c_A \frac{\partial \mathbf{z}^\sigma}{\partial z} + (\mathbf{z}^{-\sigma} \cdot \nabla_\perp)\mathbf{z}^\sigma = -\frac{1}{\rho}\nabla(p + \frac{B^2}{8\pi}) + \nu^+ \nabla_\perp^2 \mathbf{z}^\sigma + \nu^- \nabla_\perp^2 \mathbf{z}^{-\sigma} \quad (12)$$

$$\nabla_\perp \cdot \mathbf{z}^\sigma = 0 \quad (13)$$

where $\mathbf{z}^\sigma = \mathbf{v}_\perp + \sigma \mathbf{b}_\perp$ (with $\sigma = \pm 1$), $\nabla_\perp$ is the spatial gradient perpendicular to $\mathbf{B}_0$, $c_A = B_0/(4\pi\rho)^{1/2}$ is the Alfvén velocity associated with B_0.

Equation (12) shows that, in this approximation, nonlinear interactions are retained only in the directions perpendicular to $\mathbf{B}_0$, while only propagation at the Alfvén speed takes place parallel to $\mathbf{B}_0$.

Let us now Fourier transform the fields with respect to the perpendicular variable $\mathbf{r}_\perp = (x, y)$, in a 2D periodic box of size a

$$\mathbf{v}_\perp(\mathbf{r}, z, t) = \sum_{\mathbf{k}} v(\mathbf{k}, z, t)\mathbf{e}(\mathbf{k}) \exp(i\mathbf{k} \cdot \mathbf{r}) , \quad (14)$$

$$\mathbf{b}_\perp(\mathbf{r}, z, t) = \sum_{\mathbf{k}} b(\mathbf{k}, z, t)\mathbf{e}(\mathbf{k}) \exp(i\mathbf{k} \cdot \mathbf{r}) , \quad (15)$$

where $\mathbf{k} = 2\pi\,\mathbf{m}/a$, ($\mathbf{m}$ is a couple of integers) and $\mathbf{e}(\mathbf{k})$ is a unit vector perpendicular to $\mathbf{k}$. After some algebra, it can be shown that in this approximation, for each value of z, the MHD equations reduce to [49]

$$\begin{aligned} \frac{\partial v(\mathbf{k})}{\partial t} &= c_A \frac{\partial b(\mathbf{k})}{\partial z} + \sum_{\mathbf{pq}} c(k, p, q)(p^2 - q^2) \\ &\quad \times \left[v(\mathbf{p})v(\mathbf{q}) - b(\mathbf{p})b(\mathbf{q})\right] \delta_{\mathbf{k},\mathbf{p}+\mathbf{q}} \\ \frac{\partial b(\mathbf{k})}{\partial t} &= c_A \frac{\partial v(\mathbf{k})}{\partial z} + \sum_{\mathbf{pq}} c(k, p, q)k^2 \\ &\quad \times \left[b(\mathbf{p})v(\mathbf{q}) - v(\mathbf{p})b(\mathbf{q})\right] \delta_{\mathbf{k},\mathbf{p}+\mathbf{q}} \end{aligned} \quad (16)$$

(for simplicity we omit the time and z dependence of the Fourier amplitudes) where

$$c(k, p, q) = \frac{p_x q_y - p_y q_x}{2kpq}$$

means that the sum is extended over all wave vectors $\mathbf{p}$ and $\mathbf{q}$ which satisfy the triad-interaction relation $\mathbf{k} = \mathbf{p} + \mathbf{q}$. Let us note that, if we chose a box of size $a = 2\pi$, each wave vector $\mathbf{k}$ turns out to be represented by a couple of integers.

Equations (16) have quadratic invariants namely the total energy

$$E(t) = \sum_{\mathbf{k}} [|v(\mathbf{k}, t)|^2 + |b(\mathbf{k}, t)|^2]$$

and the cross-helicity

$$H_c(t) = \sum_{\mathbf{k}} \mathrm{Re}[v(\mathbf{k}, t)b^*(\mathbf{k}, t)] .$$

When the background magnetic field is set to zero $B_0 = 0$, also the mean-square of the vector potential

$$A(t) = \sum_{\mathbf{k}} |b(\mathbf{k}, t)|^2/k^2$$

is conserved. On the contrary, when $B_0 \neq 0$, the last quantity is almost constant, with relative fluctuations of about $\Delta A/A \simeq 10^{-2}$ [49].

This means that 2D MHD has a much wider field of physical applications then those situations where no explicit dependence on the z components exists. Actually 2D MHD represents a good approximation for the low-beta and low aspect ratio plasmas as for example some laboratory devices (tokamaks, RFP, etc.) or some magnetic structures (loops) in the solar corona. In the next section we will take advantage of this result.

Self-Organization at Large Scales

Relaxation Processes in MHD

In MHD, the quadratic invariants (4, 5, 6) play a crucial role in the problem of predicting the final fate of solutions starting from quite general initial conditions. When

studying the ideal relaxation, the quadratic invariants remain constant during time evolution, and the problem of predicting the final state reduces to calculating the ensemble–averaged equilibrium spectra of these invariants. Using a standard statistical approach it has been shown by Frisch et al. [23] that these spectra are determined by the ensemble-averaged initial values of invariants. Dissipative relaxation processes in MHD are much more complicated, since the values of the ideal invariants change during the evolution. It seems however that some long-living nontrivial states toward which dissipative flows are attracted exist, and that these states could be derived by minimizing an energy integral subject to some constraints.

Dissipative relaxations have been investigated in connection with measurements in laboratory plasmas, and observations in the solar wind turbulence. Taylor [51], inspired by laboratory plasma experiments, conjectured that an MHD system relaxes towards a state where the energy tends to a minimum, subject to the constraint that magnetic helicity is conserved. The relaxation can then be seen as a *selective decay* between the two invariants. From a mathematical point of view, the solution of the problem can be obtained through a variational principle

$$\delta\left\{\int(\mathbf{v}^2+\mathbf{b}^2)\mathrm{d}V-\lambda\int\mathbf{a}\cdot\mathbf{b}\mathrm{d}V\right\}=0\,,$$

where λ is a Lagrange multiplier. By imposing the conservation of magnetic helicity, we get the so-called force-free solution, characterized by $\mathbf{v}=0$ and

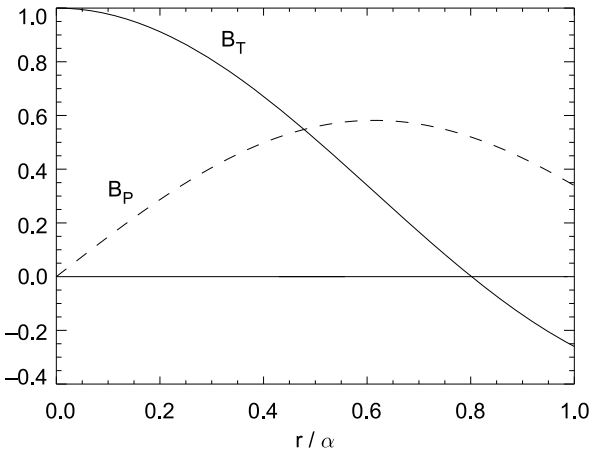
$$\nabla\times\mathbf{b}=\alpha\mathbf{b} \tag{17}$$

(α being a constant), i. e. a solution where the velocity field and the Laplace force are both null. This means that the kinetic energy decays to zero and the magnetic energy occupies the largest scale. Taylor's hypothesis allows one to circumvent in an elegant way the complex plasma relaxation dynamics (involving turbulence and reconnection).

When looking at (17) in cylindrical coordinates, the following solution can be obtained

$$B_T=B_0J_0(\alpha r)\quad B_P=B_0J_1(\alpha r) \tag{18}$$

where B_T and B_P are respectively the axial and poloidal components of the magnetic field (B_0 is a constant and J_0 and J_1 are Bessel functions). This solution, which can display a change in the sign of the magnetic field axial component (Fig. 2), was able to explain the puzzling phenomenon of field reversal spontaneously observed in some laboratory experiments, where no current necessary to produce such reversal was present in the initial state. This theory has also been the basis for the development



Self-Organization in Magnetohydrodynamic Turbulence, Figure 2

Axial (*solid line*) and poloidal (*dashed line*) components of magnetic field in a force free cylindrical plasma column as function of the radial coordinate normalized to α^{-1} ($B_0=1$)

of RFP devices (Fig. 3), toroidal devices which at a first order approximation can be considered periodic cylinders (Fig. 4). Applications to spheromak and other plasma fusion devices [4,52,59] have been considered more recently.

Belcher and Davis [3] analyzing solar wind data discovered velocity and magnetic field fluctuations were extremely well-correlated (Fig. 5). Dobrowolny et al. [20] suggested that this correlation could be the result of a self organization of MHD turbulence and set up a phenomenological explanation for this self organization, based on the idea that between the two modes of propagation of *Alfvénic* fluctuations in solar wind only that initially dominating survives to the non-linear cascade process towards dissipative scales.

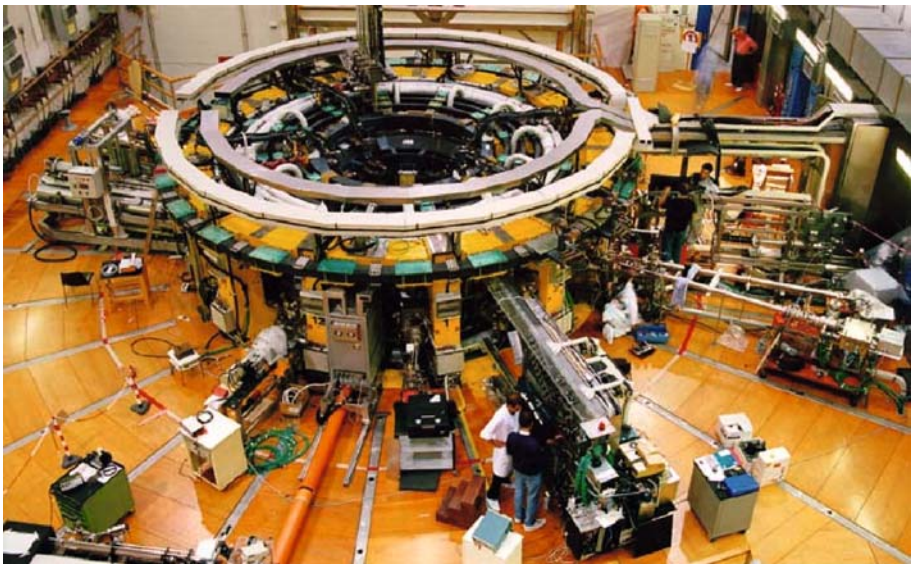
This kind of self organization can also be obtained through a variational principle. Let us impose that energy assumes a minimum value, constrained to the conservation of cross-helicity, we have then to write

$$\delta\left\{\int(\mathbf{v}^2+\mathbf{b}^2)\mathrm{d}V-\lambda\int\mathbf{v}\cdot\mathbf{b}\mathrm{d}V\right\}=0\,,$$

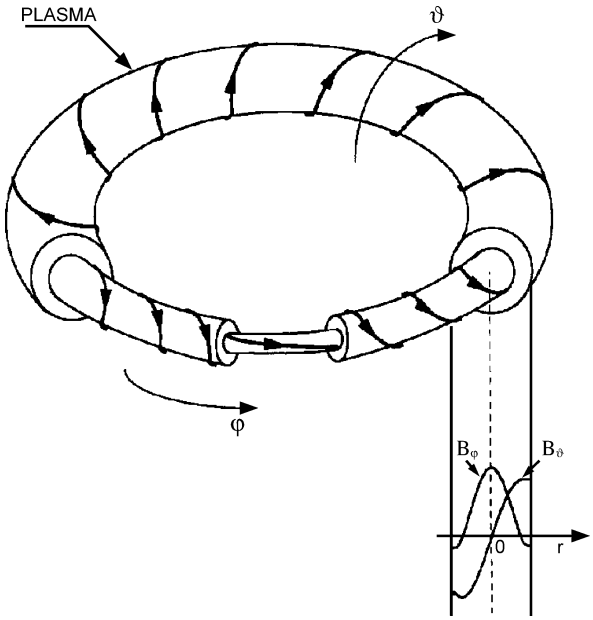
once again λ is a Lagrange multiplier. By imposing the conservation of cross helicity, we finally get the two solutions

$$\mathbf{v}=\sigma\mathbf{b} \tag{19}$$

with $\sigma=\pm1$. These solutions, which can also be written $\mathbf{z}^\sigma=0$ for either $\sigma=1$ or $\sigma=-1$ in terms of the Elsässer [21] variables, correspond to annihilating nonlinear interactions in (1, 8). This phenomenon has been



Self-Organization in Magnetohydrodynamic Turbulence, Figure 3
The RFX: A reverse field pinch plasma device in Padua (Italy) (Consorzio RFX is a research organization promoted by CNR, ENEA, Università di Padova, Acciaierie Venete S.p.A. and INFN, within the framework of the Euratom – ENEA Association)



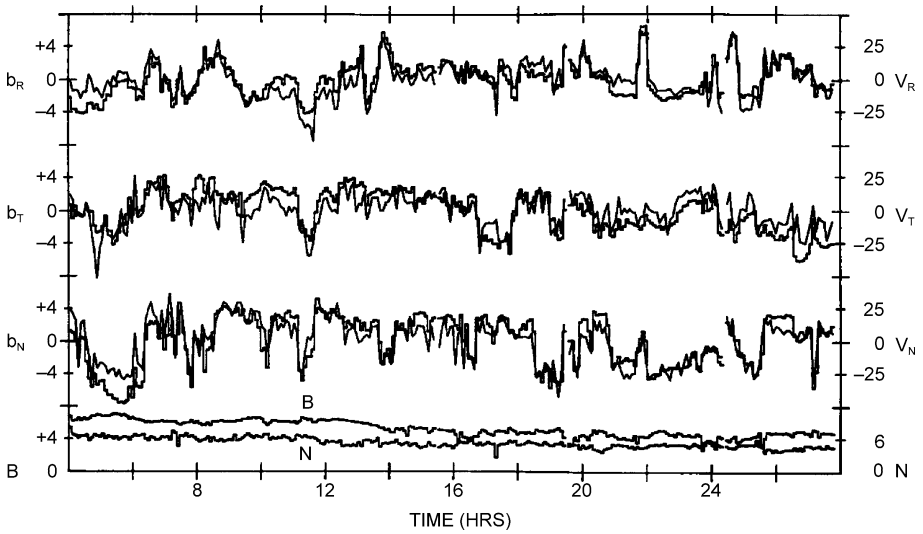
Self-Organization in Magnetohydrodynamic Turbulence, Figure 4
Typical profile of the magnetic field in an RFP device (reprinted from [48])

called *dynamical alignment* in that velocity and magnetic field tend to become aligned to each other.

It has been shown that also in numerical simulations of 2D and 3D MHD equations [17,38,39,46,53] there exists systematic behavior of turbulent flows. In particular, relaxation processes which bring the fluid toward those particular states like *Taylor solutions* or *Alfvénic solutions* have been found to emerge systematically from numerical simulations. In terms of ideal invariants, it has been shown that the Taylor's solution is obtained, when starting with almost zero cross helicity, while the *dynamical alignment* is found when the initial value of the cross helicity is substantially different from zero. In general when both effects are in competition, the situation is much more complicated. A systematic analysis of a wide number of 2D MHD simulations at different resolutions have been performed by Ting et al. [53]. Their results can be summarized by referring to the time behavior of the system projected on the plane (a, h), where $a = A/E$ and $h = 2H_c/E$. In fact, starting from whatever initial conditions, in that plane the system tends to approach the ellipse

$$\frac{1}{a} = 2\left(\frac{k_{\min}}{h}\right)^2 \left[1 \pm \sqrt{1-h^2}\right], \tag{20}$$

where $k_{\min}$ represents the minimum wave vector allowed in the simulations. Of course the points $(a, h) = (1/2, \pm 1)$ represents the *attractors* of dynamical alignment solutions, the point $(a, h) = (1, 0)$ represents the attractor of selective decay, and the point $(a, h) = (0, 0)$ represents the attractor when the system behaves like a fluid, with zero magnetic field. These extreme points can be recovered through the minimization of energy subject to the



Self-Organization in Magnetohydrodynamic Turbulence, Figure 5

Velocity (v_R, v_T, and v_N) and magnetic field (b_R, b_T, and b_N) fluctuations measured in solar wind in the RTN reference system (reprinted from [3]). The RTN system has the R unity vector along the radial direction, positive from the sun to the spacecraft, the T unity vector perpendicular to the plane formed by the rotation axis Ω of the sun and the radial direction, i. e., T = Ω × R, and the N component resulting from the vector product N = R × T. Particle density N and magnetic field intensity B are also shown in the lowest part of the plot. Units are 10^{-5} gauss for magnetic fields, km/s for velocities, and cm^{-3} for particle density

conservation of the other invariants. When the magnetic field is zero, the kinetic helicity is invariant and the point $(a, h) = (0, 0)$ can be recovered by allowing that the kinetic energy decays, subject to the constraint that the kinetic helicity remains constant. The entire curve, however, does not represent the locus of the extreme of anything over its entire range of definition.

A Minimal Triad-Interaction Model of Relaxation Processes in 2D MHD

As we have seen, 2D MHD plays a privileged role, in that it describes low-beta and low aspect ratio 3D plasmas. Here we briefly present how the basic non-chaotic triad-interaction works in 2D MHD, and how this represents a minimal basic model for relaxation processes. Choosing only three wave vectors, namely $\mathbf{k}_1 = (1, 1)$, $\mathbf{k}_2 = (2, -1)$, $\mathbf{k}_3 = (3, 0)$, we obtain a set of 12 ordinary differential equations [18] for the complex Fourier modes of both velocity $u_i(t) = u(\mathbf{k}_i, t)/2\sqrt{5}$ and magnetic field $b_i(t) = b(\mathbf{k}_i, t)/2\sqrt{5}$. The system can be further reduced through a projection of equations on a subset of the phase space, that is by considering only real fields. This can be seen by writing the fields in the form $u_j = |u_j|e^{i\alpha_j}$ and $b_j = |b_j|e^{i\beta_j}$, and by defining real fields through $V_j(t) = |u_j| \cos\alpha_j$ and $B_j(t) = |b_j| \cos\beta_j$ (subject to the conditions $\sin\alpha_j = 0$ and $\sin\beta_j = 0$). In this case we found a set of 6 ordinary differential equations for V_j and B_j, namely

$$
\begin{aligned}
(\mathrm{d}/\mathrm{d}t + 2\nu)V_1 &= 4(V_3V_2 - B_3B_2)\\
(\mathrm{d}/\mathrm{d}t + 5\nu)V_2 &= -7(V_3V_1 - B_3B_1)\\
(\mathrm{d}/\mathrm{d}t + 9\nu)V_3 &= 3(V_1V_2 - B_1B_2)\\
(\mathrm{d}/\mathrm{d}t + 2\mu)B_1 &= 2(B_3V_2 - V_3B_2)\\
(\mathrm{d}/\mathrm{d}t + 5\mu)B_2 &= 5(V_3B_1 - B_3V_1)\\
(\mathrm{d}/\mathrm{d}t + 9\mu)B_3 &= 9(V_1B_2 - B_1V_2)\,.
\end{aligned}
$$

For its simplicity the model represents the basic system to investigate the structure of nonlinear interactions in 2D MHD, and to study the role played by the rugged invariants during the dynamical evolution.

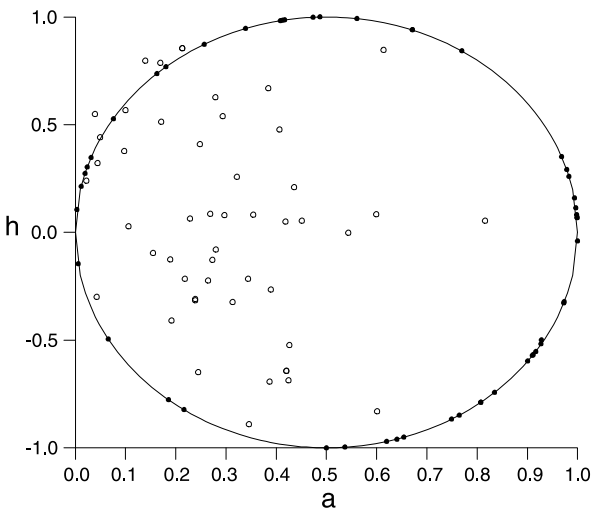
Let us introduce the phase space Ω, of dimension $\mathrm{Dim}[\Omega] = 6$, which can be built up by using as coordinates the Fourier amplitudes. A point $\Psi_i(t) \in \Omega$, defined as $\Psi_i(t) := \{[u_i(t), b_i(t)] \in \Omega\}$ represents the system at a certain time, and this point moves in Ω according to the flow $T_\tau[\Psi_i(t)] = \Psi_i(t + \tau)$ which represents the result of the equation of motion for the system. In absence of dissipative terms the phase space volume is conserved

$$
H = \sum_i \frac{\partial}{\partial\Psi_i}\left(\frac{\mathrm{d}\Psi_i}{\mathrm{d}t}\right) = 0\,, \tag{21}
$$

where H represents the rate of change of volumes in phase space.

If we define the ideal flow T_τ^{id} as the flow T_τ obtained when $\nu = \mu = 0$, the phase space volume conservation can be expressed as $T^{id}_{\tau\to\infty}[\Psi_i(t)] = \Psi_i(t+\tau)$. In this case, for a given set of initial values $\Psi_i(0)$, the point moves on a hyper-surface $S \subset \Omega$ defined by the initial value of the invariants. In presence of dissipative terms the quadratic invariants decay, and the rate of change of the volume in the phase space is $H = -16(\nu + \mu) \leq 0$. The condition $H \leq 0$ implies that the dissipative flow pushes the system toward the trivial asymptotic state where all the amplitudes of the fields are zero $T^{\mathrm{diss}}_{\tau\to\infty}[\phi(t)] = 0$.

The triad-interaction model is able to capture the dynamics of the quiescent states observed in MHD. We solved our system by starting from random initial conditions uniformly distributed $\Psi_i(0) \in [-1, 1]$. We used a fourth-order Runge–Kutta scheme, with a time step $\Delta t = 10^{-3}$ and $\nu = \mu = 0.01$. This value for the dissipative coefficients allows the nonlinear interactions to have sufficient time to set up the dynamical behavior. In Fig. 6 we report the curve ε along with two ensembles of points. The first ensemble (white circles) represents the set of points (a, h) obtained with some different initial conditions $\Psi_i(0)$ randomly chosen in the interval $[-1; 1]$. The second ensemble (black circles) represents the set of points (a, h) at time $t = 80$ (in unit of time steps), calculated from the set of fields $\Psi_i(t)$ which are obtained through the time evolution of the set $\Psi_i(0)$, that is $\Psi_i(t) = T_t^{\mathrm{diss}}[\Psi_i(0)]$. As it can be seen that all the initial conditions lead to the final state which belongs to the ellipse ε.



Self-Organization in Magnetohydrodynamic Turbulence, Figure 6

Numerical simulations of the triad-interaction model reported in the plane (a, h). We show two sets of different solutions at two different times. *White circles* refer to a set of 45 initial values $\Psi_i(0)$, *black circles* represent the set of point at a given time $t = 80$ (in unit of time steps) obtained from the time evolution of the above initial values $\Psi_i(t = 80) = T^{\mathrm{diss}}_{t=80}[\Psi_i(0)]$. The solutions represented on that plane are such that, after an initial transient, they fall on the ellipse represented as a *full line*

Since nonlinear interactions in the simple triad-interaction model have the same structure as in the 2D MHD equations, the model is able to capture relaxation properties. These properties are the fixed points and some invariant subspaces. The fixed points of the system can be classified as follows:

a) Two Alfvénic fixed points (say $A^{\pm}$) characterized by $u_i = \pm b_i$;
b) Three fluid fixed points $F(i)$ $(i = 1, 2, 3)$ where all variables but the velocity V_i are zero;
c) Three magnetic fixed points $M(i)$ $(i = 1, 2, 3)$ where all variables but the magnetic field B_i are zero.

A standard analysis of stability can be performed by linearizing the system around each fixed point. We find that $A^{\pm}$ are always stable, $M(1)$ is the only stable magnetic fixed point, while $F(1)$ and $F(3)$ are stable. The only stable magnetic fixed point $M(1)$ is such that the energy is localized on the minimum wave vector.

Apart from fixed points the system displays some other interesting properties. Looking at the equations, it can be easily shown that there exists some 3D subspaces of the 6D phase space which remain invariant under the ideal flow operator. Let us denote by I^α the αth invariant subspace, which is then characterized by the fact that if $\Psi_i(0) \in I^\alpha$ then $\Psi_i(t) = T_t^{id}[\Psi_i(0)] \in I^\alpha$ for each $t > 0$. In other words an ideal invariant subspace is a portion of the phase space where the system lies for all times. The most useful way to classify these structures is through the initial values the rugged invariants assume on them, since under the ideal flow they remain constant:

a) A fluid subspace F, characterized by $a = A/E = 0$ and $h = 2H_c/E = 0$. This can be recovered by imposing that the magnetic field is always zero, namely

$$\Psi_i[F] = (V_1, V_2, V_3, 0, 0, 0)$$

is the vector which describes this subspace.
b) Two Alfvénic subspaces $A^{\pm}$, characterized by $a = A/E = 0$ and $h = 2H_c/E = 1$, which can be recovered by imposing that $V_i = \pm B_i$ for each $i = 1, 2, 3$. These subspaces are also fixed points of the system.

c) Three magnetic subspaces H^A, H^B and H^C characterized by $a \neq 0$ and $h = 0$, i. e. a minimal value for the cross-helicity. These subspaces can be recovered by imposing that the cross-helicity is initially zero over all wave vectors, namely either $V_i = 0$ and $B_i \neq 0$ or vice versa; specifically,

$$\Psi_i[H^A] = (0, 0, V_3, B_1, B_2, 0)$$
$$\Psi_i[H^B] = (V_1, 0, 0, 0, B_2, B_3)$$
$$\Psi_i[H^C] = (0, V_2, 0, B_1, 0, B_3)\,.$$

The stability properties of each subspace can be investigated numerically [11]. Simulations start by putting the system on a given subspace, and by adding a small perturbation on the complementary manifold. The motion is not limited to the 3D subspace, and the problem of the stability of a particular I^α consists of determining if the perturbed solution remains close to I^α or goes away from it covering all the allowed 6D phase space. For each subspace we can define two pseudo-energies $E^{(\alpha)}_{\text{int}}(t)$ (built up with the fields which belong to the αth subspace) and $E^{(\alpha)}_{\text{ext}}(t)$ (built up with the fields which do not belong to the αth subspace). The external energy represents the distance $\|\Delta^{(\alpha)}[\Psi_i(t)]\|$ between the point $\Psi_i(t)$ and the αth invariant subspace. Since the total energy $E(t) = E^{(\alpha)}_{\text{ext}}(t) + E^{(\alpha)}_{\text{int}}(t)$ must remain constant under the ideal flow, two situations can arise, namely

1) During the time evolution both $E^{(\alpha)}_{\text{ext}}(t) \simeq E^{(\alpha)}_{\text{ext}}(0)$ and $E^{(\alpha)}_{\text{int}}(t) \simeq E^{(\alpha)}_{\text{int}}(0)$, which means that the solution remains trapped near the invariant subspace. In that case, the subspace is ideally stable.
2) During the time evolution energies become comparable $E^{(\alpha)}_{\text{ext}}(t) \simeq E^{(\alpha)}_{\text{int}}(t)$, which means that the solution is repelled from the invariant subspace. In that case the subspace is ideally unstable.

Numerical simulations show that the fluid subspace is always stable. As far as the magnetic subspaces are concerned, we find that H^A is always stable, while H^B and H^C are always unstable. More information about the behavior of the system near the invariant subspaces can be obtained by considering the characteristic of solutions when the dissipative coefficients are set different from zero. In that case, the energies decay in time, but the rate of decay is different, thus indicating a kind of selective dissipation. In particular, looking at the time evolution of their ratios $R^\alpha(t) = E^{(\alpha)}_{\text{ext}}(t)/E^{(\alpha)}_{\text{int}}(t)$, we find that it decays to zero for the ideally stable subspaces, while it settles to a constant value for the unstable subspaces. This means that ideally stable subspaces, namely the Alfvénic, fluid and magnetic H^A represent a kind of attractor for the system, while unstable subspaces, say H^B and H^C, repel solutions.

When we start numerical solutions near the stable subspaces, the system is attracted to the extreme points of the plane (a, h). When we start near the fluid attractor, the system reach the point $(a, h) = (0, 0)$, and when we start near one of the Alfvénic attractors the system is attracted to $(a, h) = (1/2, \pm 1)$. Finally, when we start near the stable magnetic subspace H^A the system is attracted to the extreme point $(a, h) = (1, 0)$. This last point represents a Taylor regime, say the magnetic field lies on the lowest allowed wave vector. On the contrary, when we start near the unstable magnetic attractors H^B or H^C, the system evolves in an erratic way towards any point (a, h) of the curve (20). Each point (a, h) is made by only one mode with wave vector $k_1 = k_{\min} = \sqrt{2}$, so that if $\eta = V_1/B_1$ we have

$$a = \frac{2}{\left(1 + \eta^2\right) k^2_{\min}}$$
$$h = \frac{\eta}{1 + \eta^2}\,. \tag{22}$$

This last equation is nothing but the parametric equation of the curve (20).

The behavior we have just described can be recovered also when we start the numerical computation with general initial conditions. In that case, we can associate to each initial condition an invariant subspace according to the rule that the αth subspace is associated with the initial condition $\Psi_i(0)$ when the distance $\|\Delta^{(\alpha)}[\Psi_i(0)]\|$ is the minimum one over α. In other words, we associate an initial condition to the nearest subspace. Numerical simulations show that when the initial conditions are associated to the unstable subspaces H^B and H^C, the final point reached by the system will be any point (a, h) on the curve (20). When the initial conditions are associated with one of the stable subspaces, the final state will be one of the extreme points of the curve (20).

Self-Organization at Small Scales and Intermittent Structures in MHD Turbulence

The self organization processes discussed in the previous sections are related to the spontaneous creation of large scale coherent structures during the MHD turbulence dynamics. However, also at small scales the nonlinear energy cascade process in fully developed MHD turbulence is characterized by the presence of self organization and coherent structures. These structures are strictly related to the intermittency phenomenon, that is, the breakdown of global self-similarity in the turbulent cascade.

The turbulence theory proposed by Kolmogorov in 1941 [30] is based on the conjecture that, within the inertial range, the statistical properties of the fluid motions depend only on the mean energy transfer rate ε and on the scale ℓ, and that the nonlinear energy cascade occurring in this range is a self similar (fractal) process. According to this idea, the velocity field increments $\delta \boldsymbol{v}_\ell(\boldsymbol{r}) = \boldsymbol{v}(\boldsymbol{r}+\boldsymbol{\ell}) - \boldsymbol{v}(\boldsymbol{r})$ scale as $\delta v_\ell \sim \ell^{1/3}$. The self similarity of the turbulent cascade implies that the probability density function (PDF) of field increments $P(\delta v_\ell)$ should be invariant under the scale change

$$P(\delta v_\ell) = \ell^{-1/3} F\left(\frac{\delta v_\ell}{\ell^{1/3}}\right) , \tag{23}$$

which means that PDFs of normalized velocity increments at different scales collapse onto the same curve. Another consequence of the global self similarity is that the structure functions $S_p(\ell) = \langle \delta v_\ell^p \rangle$, usually defined using the longitudinal field increments, that is, $\delta v_\ell(\boldsymbol{r}) = \left[\boldsymbol{v}(\boldsymbol{r}+\boldsymbol{\ell}) - \boldsymbol{v}(\boldsymbol{r})\right] \cdot \frac{\boldsymbol{\ell}}{\ell}$, should follow the scaling law $S_p(\ell) \sim \ell^{p/3}$.

When considering incompressible MHD turbulence, the scaling properties of the field increments $\delta \boldsymbol{z}_\ell^\sigma(\boldsymbol{r}) = \boldsymbol{z}^\sigma(\boldsymbol{r}+\boldsymbol{\ell}) - \boldsymbol{z}^\sigma(\boldsymbol{r})$, expressed here in terms of Elsässer variables $\boldsymbol{z}^\sigma = \boldsymbol{v} + \sigma \boldsymbol{B}/\sqrt{4\pi\rho}$ ($\boldsymbol{B}$ being the magnetic field and ρ the mass density) can be modified with respect to the fluid case by the so-called Alfvén effect, consisting of the fact that nonlinear interactions take place between eddies of different σ, which propagate in opposite directions at the Alfvén velocity along the large scale magnetic field reducing the efficiency of the non linear energy cascade [20,29,31].

In the fluid-like case a Kolmogorov-like scaling can be expected for the field increments, that is, $\delta z_\ell^\sigma \sim \ell^{1/3}$. On the other hand, when the Alfvén effect is at work the Kraichnan's scaling $\delta z_\ell^\sigma \sim \ell^{1/4}$ is recovered. However, in both cases the turbulent energy cascade is supposed to be a self similar process, resulting in a linear behavior of the structure function exponents, that is, $S_p(\ell) \sim \ell^{p/m}$, where $m = 3$ and $m = 4$ for the fluid-like and Alfvenic case respectively.

Several studies on the statistical properties of fields in fully developed turbulence have shown that the PDFs of field increments are not self similar, that is, their shape change with the scale ℓ. The PDFs at large scales are nearly Gaussian, while, as the scale decreases, the PDFs show sharper and sharper peaks and, correspondingly, higher and higher tails. As far as MHD turbulence is concerned, this behavior has been found both in the solar wind [36,50] and in laboratory plasmas [13]. This departure from self similarity has also been inferred from structure function analysis (see e. g. [8,12,13,35]) which have shown that for both velocity and magnetic field fluctuations the scaling exponents are nonlinear functions of p.

This behavior indicates thus that the turbulent, nonlinear energy cascade is not a fractal (self similar), but rather a multifractal process. In other words, small scale fluctuations much larger than their RMS are present in some spatial positions and the corresponding spatial fluctuations of the energy transfer rate $\varepsilon(\boldsymbol{r})$ play a primary role in the cascade process. These small scale fluctuations have been interpreted as the signature of localized coherent structures spontaneously produced by the nonlinear dynamics. The phenomenon described above is known as *intermittency* in fully developed turbulence.

Another important aspect of intermittency is the occurrence of impulsive bursts of energy dissipation which are observed in different systems characterized by the presence of MHD turbulence, e. g. the magnetic loops of the solar corona [1], and laboratory plasma devices such as reversed field pinches [5] and tokamaks [56]. This phenomenon is often referred to as *temporal intermittency*, to distinguish it from the *spatial intermittency* described above, as it consists in the observation of strong bursts of activity in the time series of a quantity which traces the energy dissipation in the system (e. g. radiation intensity). These bursts are separated from each other by intervals of low activity, denoted as laminar times or waiting times.

While temporal intermittency occurring in turbulent systems has been described, until the end of 90s, in the framework of theoretical paradigms other than turbulence (e. g. Self organized criticality [2]), the idea that spatially and temporal intermittency in fully developed turbulence are both due to the underlying nonlinear dynamics of the energy cascade, giving rise to small scale coherent structures, has now gained considerable standing. In the MHD turbulence context this picture has been supported by a number of results based on the analysis of solar wind and solar corona observations and laboratory plasma experimental data.

The intermittent, coherent events can be viewed as localized zones of fluid where phase correlations exist. These structures, which dominate the statistics of small scales, occur as isolated events. They continuously appear and disappear, apparently in a random fashion, at some locations in the fluid, and they carry most of the flow energy. Among the different techniques which can be used to identify such structures within a turbulent signal, a very effective one is based on the use of wavelet transforms [22]. The wavelet transform of a real square integrable signal $f(x)$ (where x is usually the time or a spatial coordinate) is de-

fined as

$$W(x,r) = C_\psi^{-1/2} r^{-1/2} \int_{-\infty}^{\infty} \psi\left(\frac{x'-x}{r}\right) f(x')\mathrm{d}x' \,, \quad (24)$$

where r is a scale dilation, x is a position (time) translation, $\psi(x)$ is the so called mother wavelet function, and C_ψ is a normalization constant, which must satisfy the admissibility condition $\int_{-\infty}^{+\infty} |k|^{-1}|\hat{\psi}(k)|^2 \mathrm{d}k < \infty$ where $\hat{\psi}(k)$ is the Fourier transform of $\psi(x)$. The wavelet coefficients $W(x,r)$ give a decomposition of $f(x)$ at the scale r as a function of the position x. For each scale, it is possible to identify the position (or the time occurrence) of strong, intermittent events through a method proposed by Farge [22] based on the so-called Local Intermittency Measure (LIM) $l(x,r)$ defined as

$$l(x,r) = \frac{|W(x,r)|^2}{\langle |W(x,r)|^2 \rangle} \,. \quad (25)$$

The identification method of intermittent structures is based on the idea that large values of $l(x,r)$ represent a signature of large fluctuations with respect to the background level. The wavelet coefficients can thus be classified as "passive" if $l(x,r) < F$ or "intermittent" when $l(x,r) > F$, where F is a threshold which can be chosen by using different suitable criteria [9,41,54].

Another possible method to select intermittent bursts [6] consists of defining bursts as the time intervals during which the condition $f(x) \geq f_{\mathrm{th}}$ is satisfied, where the threshold value f_{th} is calculated through the following iterative method. One starts by defining $f_{\mathrm{th}} = \langle f(t) \rangle + a\sigma$ (a is an arbitrary positive number, usually 2 or 3), where the average and the standard deviation are computed from the whole time series. After excluding the values which exceed f_{th}, the average and the standard deviation are calculated again using the remaining part of the time series. This process is repeated until convergence of f_{th} is reached. This procedure allows one to remove strong events and to evaluate the threshold taking into account only the background contribution.

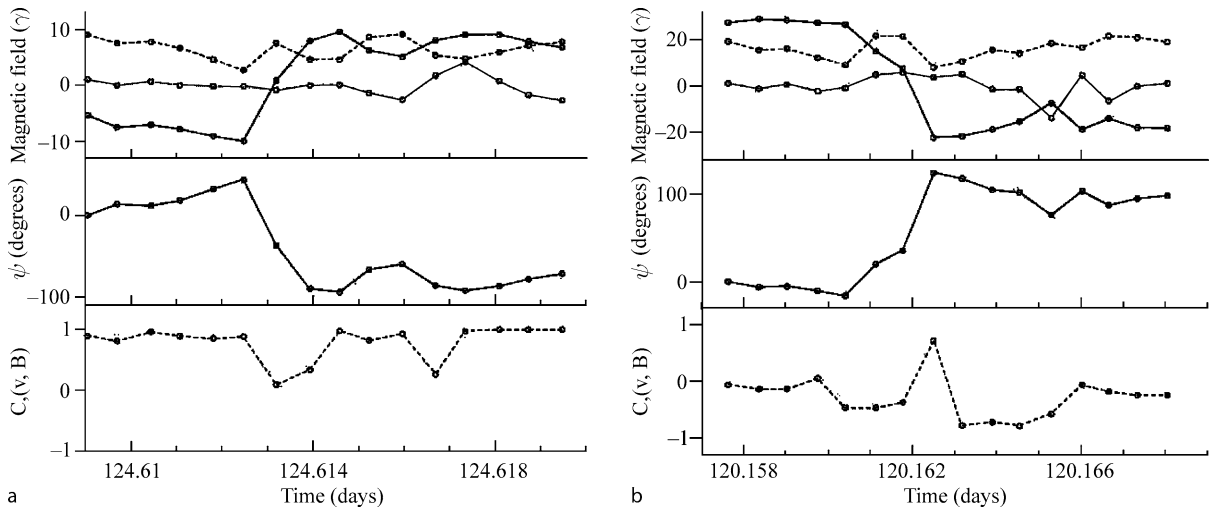
Once intermittent structures are identified, it becomes possible to study their typical profiles and some relevant statistical properties such as the distributions of event energy and of time intervals between successive bursts. Such investigations allow one to shed light on the physical mechanisms underlying the bursting process. We discuss now the properties of intermittent events observed in three different systems where MHD turbulence occurs, namely the solar wind, reversed field pinches devices, and magnetic loops of the solar corona. Related aspects of intermittency and multifractality are discussed in the review by Chang ▶ Space Plasmas, Dynamical Complexity in in this Encyclopedia.

Intermittent Structures in Solar Wind MHD Turbulence

Intermittent events in solar wind turbulence have been studied by Veltri and Mangeney [54] and Veltri et al. [55] by applying the LIM technique, with the Haar wavelet basis, to fluid velocity and magnetic field measurements performed during about 1 year in the space experiment ISEE. In this experiment only two components of the fluid velocity, namely, V_x and V_y, were measured, together with all the three components of the magnetic field. The reference frame used was the standard GSE frame and the sample was formed by data at a time resolution of $T = 1$ min, so that the sampling rate was $\Delta = V_{\mathrm{sw}} T \sim 24.000$ km (V_{sw} is the average solar wind velocity).

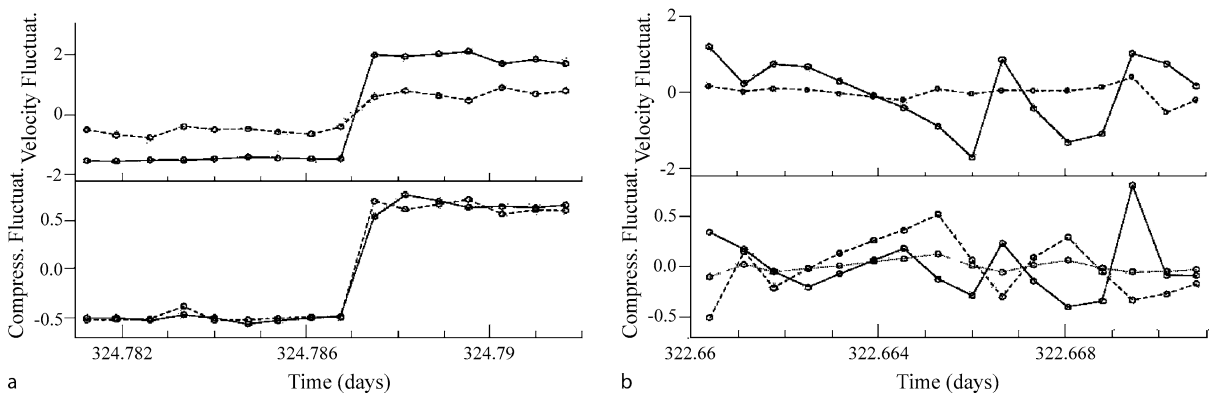
The classification of wavelet coefficients allows for an identification and a study of the most intermittent events in solar wind turbulence, which occur in those positions where the amplitude of the wavelet coefficients displays the largest values compared to the average. These events, which occur on time scale of the order of few minutes, exhibit a small number of typical profiles, summarized as follows:

a) "Tangential discontinuities": These structures are almost incompressible, pressure balanced one dimensional current sheets. A minimum variance analysis performed on the magnetic field around the singularity shows (Fig. 7) that the component of the magnetic field which varies most changes sign, and this component is perpendicular to the average magnetic field (the magnetic field component along the third axis being almost zero). The magnetic field rotates then in a plane by an angle which is about 120°–130°. There is one more interesting property: When these structures occur during an Alfvénic period (velocity and magnetic field fluctuations highly correlated), the Alfvénic correlations go from 1 to zero during the traversal of the current sheet (Fig. 7, left panel), when, on the contrary, these structures occur during a period of almost no Alfvénic correlation, the correlation increases to about 1 at the current sheet location (Fig. 7, right panel).
b) "Compressive discontinuities": These structures can be either parallel shocks, mainly observed on the radial component of the velocity field, but clearly seen also on the magnetic field intensity, proton temperature and density (Fig. 8, left panel); or slow mode wavetrains, characterized by a value of $\beta \sim 0$, a constantpressure,



Self-Organization in Magnetohydrodynamic Turbulence, Figure 7

Current sheet intermittent events in solar wind MHD turbulence: The three components of the magnetic field obtained through a minimum variance analysis (*upper panels*); the angle of rotation of the magnetic field in the plane perpendicular to the minimum variance direction (*middle panels*); the coefficient of correlation between velocity and magnetic field fluctuations (*lower panels*)



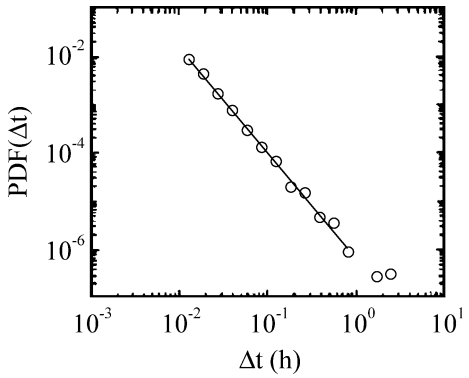
Self-Organization in Magnetohydrodynamic Turbulence, Figure 8

Intermittent events associated with compressive discontinuities in the solar wind. a A parallel shock intermittent event in solar wind, V_x (*full line*) and V_y (*dashed line*) fluctuations normalized to the local average sound velocity (*upper panel*); proton density (*full line*) and sound velocity (*dashed line*) fluctuations normalized to their local average values (*lower panel*). b A slow shock intermittent event in solar wind, velocity fluctuations parallel (*full line*) and perpendicular (*dashed line*) to the local average magnetic field normalized to the local average sound velocity (*upper panel*); proton density (*full line*), sound velocity (*dashed line*) and total pressure (*dotted line*) fluctuations normalized to their local average values (*lower panel*)

anticorrelated density and proton temperature fluctuations and with velocity fluctuations along the average magnetic field (Fig. 8, right panel).

Very interesting statistics can be studied on the time separation Δt (often denoted as *waiting time*) between the occurrence of two consecutive structures. Waiting times of solar wind intermittent structures are distributed according to a power law $P(\Delta t) \sim \Delta t^{-\gamma}$ extended over two decades at least, with an exponent $\gamma \simeq 2.18$ (Fig. 9) [14].

This property is very interesting, because it indicates that the energy cascade process is non-Poissonian. Waiting times occurring between isolated Poissonian events must be distributed according to an exponential function. The power law represents the asymptotic behavior of a Lévy function, which describes self-affine processes and is obtained from the central limit theorem by relaxing the hypothesis that the variance of variables is finite. The power law waiting time PDF is thus a clear evidence that long-range correlations (or in other words "memory") exist in the underlying cascade process.



Self-Organization in Magnetohydrodynamic Turbulence, Figure 9

Probability density functions of waiting times between consecutive intermittent structures in solar wind MHD turbulence identified by applying the LIM technique on magnetic field measurements acquired by the HELIOS II spacecraft. The *solid line* represents a power law with an exponent $\gamma = 2.18$

Intermittent Structures in Laboratory MHD Turbulence

The properties of intermittent events in the magnetic turbulence observed at the edge of a toroidal plasma device in reversed field pinch configuration has been studied by Carbone et al. [14] through the analysis of magnetic field fluctuations in the Reversed Field Pinch experiment RFX [48]. Events have been identified by applying the LIM technique on radial and toroidal magnetic field time series. The magnetic field fluctuations for some of the most intermittent events are shown in Fig. 10 by using the minimum variance reference frame.

One of the two components displays a peak when the other component changes sign. This behavior suggests that the intermittent structures identified in the RFX magnetic field are tangential discontinuities similar to those found in the solar wind, although measurements of all the magnetic field components would be needed to confirm this interpretation.

Also in RFP magnetic turbulence, the waiting time PDF of intermittent structures shows a power law behavior $P(\Delta t) \sim \Delta t^{-\gamma}$ extended over two decades as in the case of solar wind MHD turbulence, with an exponent $\beta \simeq 1.5$ (Fig. 11) [14].

Intermittent Energy Release Events in the Solar Corona

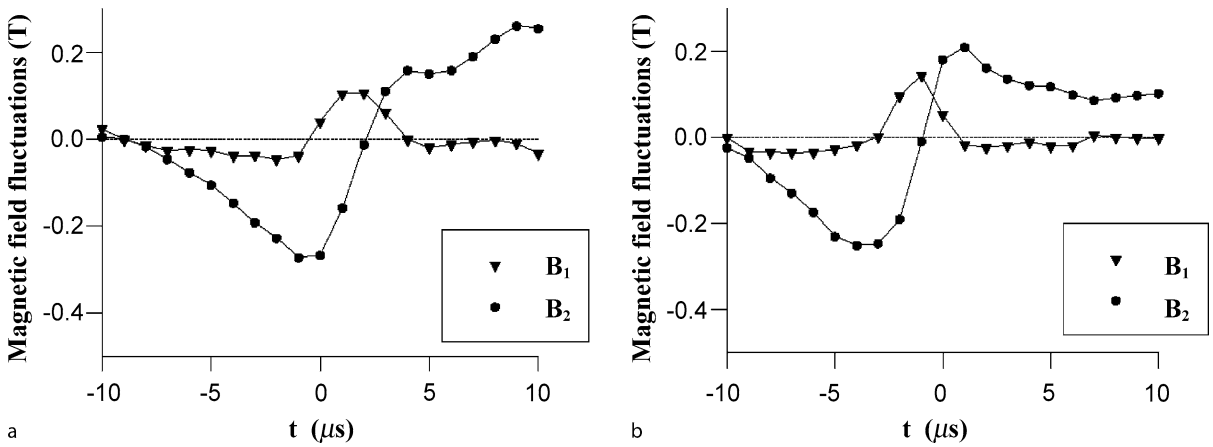
Another manifestation of intermittency phenomena in systems where MHD turbulence is present is the occurrence of impulsive energy release events, the so-called solar flares, which take place in the magnetic loops of the solar corona. Energy is released mainly through accelerated particles and emission of electromagnetic radiation in a wide range of wavelengths, from radio to γ-rays, and can vary between 10^{24} and 10^{33} erg. The smallest events, namely those between 10^{24} and 10^{27}, are usually denoted as nanoflares and microflares.

Parker [43] conjectured that nanoflares are produced by the dissipation of small current sheets, associated with tangential discontinuities, forming as a consequence of the continuous shuffling and intermixing of field footpoints in the photospheric convection. In the Parker's picture, flares of all sizes result from the superposition of small dissipation events (nanoflares) which can trigger energy releases in the neighboring discontinuities, originating a fragmented energy release process which can produce both small and large-scale events, depending on the details of the magnetic configuration.

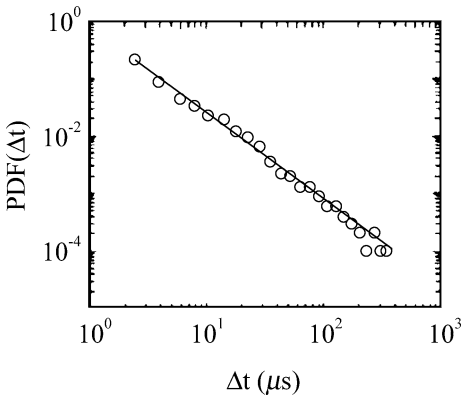
Starting from the idea that MHD turbulence is most probably a fundamental ingredient of coronal loop (Fig. 12) dynamics, it has been proposed [6] that nanoflares and flares can be identified with dissipation events of small-scale current sheets forming as a consequence of the nonlinear turbulent cascade which occurs inside coronal magnetic structures. The cascade would be driven by the energy input due to photospheric footpoint motions. In this framework, current sheets are coherent intermittent small scale structures of MHD turbulence. The energy injected at large scales by photospheric motions is transferred to small scales through the nonlinear cascade, which goes down to the dissipative scales. The intermittent nature of the energy release process is thus associated with the intermittency of energy dissipation in MHD turbulence.

This picture is supported by the statistical properties of solar flares inferred from the analysis of the associated radiation bursts. Nanoflares are detected as extreme ultraviolet (EUV) brightenings, while larger flares are observed through several types of emission (radio, microwaves, $H\alpha$, UV, X-rays) and most often studied analyzing Soft X-ray (SXR) and Hard X-ray (HXR) flare bursts. Probability distributions of the relevant quantities (peak flux, total energy, duration) characterizing the flare bursts have been found to be well represented by power laws $P(x) = Ax^{-\alpha}$ (see Fig. 13 for an example).

Using various SXR and HXR observations it has been found that, for energy and peak flux distributions, $\alpha \simeq 1.6$–2 (see e. g. [15,16,34]). For the EUV nanoflare brightenings the power law exponent is much more uncertain, $\alpha = 1.3$–2.8 (see e. g. [32,44]). The waiting time distribution (WTD) has been studied both for HXR and SXR bursts [6,45,57]. SXR observations acquired by the GOES



Self-Organization in Magnetohydrodynamic Turbulence, Figure 10
Two components of magnetic field fluctuations (denoted as B_1 and B_2), in the minimum variance reference frame, of two intermittent events observed in the RFX machine



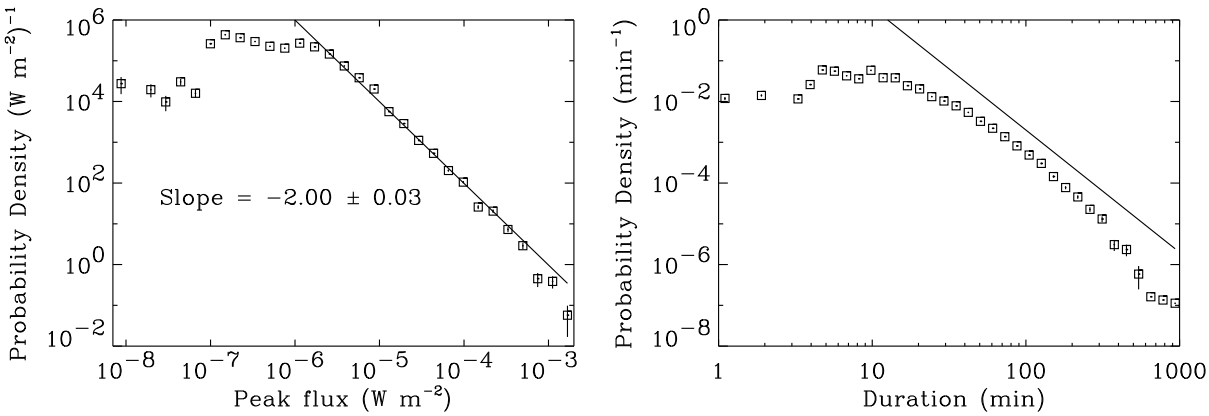
Self-Organization in Magnetohydrodynamic Turbulence, Figure 11
Probability density functions of waiting times between consecutive intermittent structures in RFP magnetic turbulence identified by applying the LIM technique on magnetic field measurements acquired at the RFX machine. The *solid line* represents a power law with an exponent $\gamma = 1.5$

Self-Organization in Magnetohydrodynamic Turbulence, Figure 12
An image of hot coronal loops which span 30 or more times the diameter of planet Earth obtained from Transition Region and Coronal Explore (TRACE) satellite. The image was taken in the Fe IX 171 Å emission line. (Transition Region and Coronal Explorer, TRACE, is a mission of the Stanford–Lockheed Institute for Space Research, and part of the NASA Small Explorer program)

satellites have the advantage to provide a long sequence of bursts (from 1975 to today) with few gaps and allow, thus, to analyze the WTD with a much better statistical accuracy than HXR data. The WTD of GOES SXR flares has been shown to display a clear power law tail $P(\Delta t) \sim \Delta t^{-\gamma}$ in the range 5 h $\lesssim \Delta t \lesssim$ 100 h, with an exponent $\gamma \simeq 2.4$ (Fig. 14) [6].

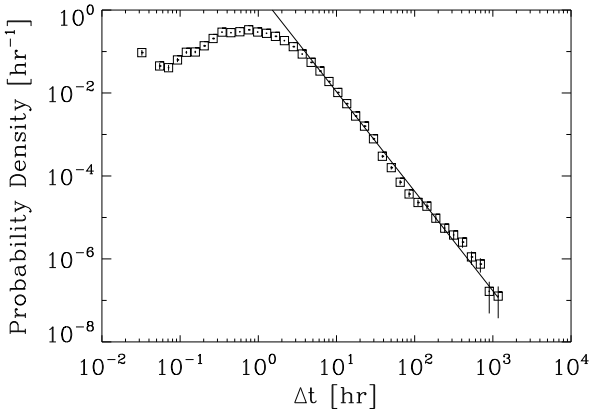
The power law behavior of the solar flare WTD represents an indication of the existence of correlations between successive bursts [6,33]. These correlations can be related to the nonlinear dynamics of the MHD turbulent energy cascade which generates intermittent bursts of chaoticity at small scales.

From the theoretical point of view, describing in a proper way the intermittency of the turbulent energy



Self-Organization in Magnetohydrodynamic Turbulence, Figure 13

Probability density functions of peak flux (*left panel*) and duration (*right panel*) of Soft *X*-ray solar flare bursts detected by the Geostationary Operational Environmental Satellites (GOES). The *solid line* in the *left panel* represents a power law with an exponent $\alpha = 2$ obtained from a least squares fit. The *solid line* in the *right panel*, shown as a comparison, represents a power law with an exponent $\alpha = 3$



Self-Organization in Magnetohydrodynamic Turbulence, Figure 14

Probability density function of waiting times for SXR solar flare bursts detected by the GOES satellites in the interval from September 1975 to December 2001. The solid line represents a power law with an exponent $\alpha = 2.4$ as found by Boffetta et al. [6]

dissipation process is one of the basic ingredients for the study of intermittent events occurring in the solar corona, as well as in other astrophysical systems such as the solar wind and accretion disks. These astrophysical plasmas are characterized by huge Reynolds numbers. The number $N_{\rm df}$ of degrees of freedom, i.e the number of grid points necessary to resolve the entire inertial range of turbulence in a direct numerical simulation of Navier–Stokes or MHD equations, grows with the Reynolds number Re as $N_{\rm df} \sim {\rm Re}^{9/4}$. Therefore performing direct numerical simulations at Reynolds number regimes of interest for space plasmas is out of the present computational possibilities. The so called shell models, a class of dynamical deterministic models of turbulence in which $N_{\rm df}$ grows logarithmically with Re, can represent an extremely helpful tool for the modeling of such physical systems, as these models are able to simulate the turbulent cascade and the related intermittency of the energy dissipation process in Reynolds number regimes which are not far from the real ones (at least with respect to direct numerical simulations). The statistical properties of intermittent events in MHD turbulence can thus be effectively investigated through the use of shell models [6,14,15].

Shell models [7] are dynamical systems designed in order to represent in a simplified way the spectral Navier–Stokes or MHD equations for turbulence. They were originally proposed by Obukhov [40], Desnyansky and Novikov [19], and Gledzer [26] in hydrodynamic turbulence.

In order to build up the evolution equations for a shell model, the wavevector space is divided into a discrete number of shells of radius $k_n = \rho^n k_0$, with $\rho > 1$, $n = 1, 2, \ldots, N$. Each shell is assigned a scalar dynamic variable (real or complex) $v_n(t)$ which represents the averaged effect of velocity modes with wavenumber between k_n and k_{n+1}. $v_n(t)$ can also be regarded as the velocity increments $|v(x + \ell) - v(x)|$ on an eddy of scale $\ell \sim k_n^{-1}$. For MHD shell models, besides $v_n(t)$, another variable $b_n(t)$, representing magnetic field modes between k_n and k_{n+1} (or, alternatively, magnetic field increments $|b(x + \ell) - b(x)|$), is assigned to each shell. The nonlinear terms of the equations are quadratic combinations of the

dynamic variables and are written under the assumption that the interactions among shells are local in k-space. This means that only nearest and, at most, next-nearest shells are introduced in the non-linear terms. The coupling coefficients are found by imposing conservation of the ideal invariants of Navier–Stokes or MHD equations.

The main advantage of shell models is that they can be studied through numerical simulations at very high Reynolds numbers. Moreover, they provide the possibility to obtain long time series with a fairly small computational effort, allowing thus a robust investigation of the statistical properties of intermittent events The weak point is that they are scalar models, that is, any information about spatial structures, such as vortices, sheets, and filaments, is lost.

Among the various shell models which have been proposed in the literature, we consider here the so called GOY (Gledzer–Ohkitani–Yamada) model, which has been used by several authors both in hydrodynamics and MHD. In the MHD context, the advantage of the GOY model with respect to previous shell models, where only the conservation of two quadratic invariants, i. e. total energy and cross helicity, can be imposed, is that it allows to conserve also magnetic helicity. The GOY model involves nearest and next-nearest interactions and was originally introduced in the framework of hydrodynamic turbulence by Gledzer [26], using real shell variables. The model was extended by Yamada and Ohkitani [60] to the case of complex variables. The MHD generalization of the GOY model has been considered by several authors (see [6,27,28] and references therein). The evolution equations for the dynamic variables $v_n(t)$ and $b_n(t)$ in the GOY shell model can be written as [6]

$$\frac{\mathrm{d}v_n}{\mathrm{d}t} = -\nu k_n^2 v_n + f_n + ik_n \Big\{ (v_{n+1}v_{n+2} - b_{n+1}b_{n+2}) - \frac{1}{4}(v_{n-1}v_{n+1} - b_{n-1}b_{n+1}) - \frac{1}{8}(v_{n-2}v_{n-1} - b_{n-2}b_{n-1}) \Big\}^* , \quad (26)$$

$$\frac{\mathrm{d}b_n}{\mathrm{d}t} = -\mu k_n^2 b_n + g_n + ik_n \frac{1}{6} \Big\{ (v_{n+1}b_{n+2} - b_{n+1}v_{n+2}) + (v_{n-1}b_{n+1} - b_{n-1}v_{n+1}) + (v_{n-2}b_{n-1} - b_{n-2}v_{n-1}) \Big\}^* . \quad (27)$$

f_n and g_n represent external forcing terms, usually acting on low wavenumber shells, while ν and μ appearing in the linear, dissipative terms, are the kinematic viscosity and the resistivity respectively, viscous and resistive terms provide a mechanism for energy dissipation at small scales. In turbulent space plasmas such as solar corona and solar wind dissipation mechanisms are different, involving kinetic plasma processes. Nevertheless, the properties of low frequency MHD turbulent cascade in the inertial range and of the related intermittency do not depend, in general, on the details of the dissipation mechanism, hence the use of viscous and resistive dissipation in the standard incompressible form does not represent a serious limitation in this context.

The GOY model described here has been designed to investigate 3D MHD turbulence. This means that it satisfies the conservation of the ideal 3D MHD quadratic invariants (4), (5) and (6). In terms of shell variables, the invariants read [27,28]

$$E = \frac{1}{2}\sum_{n=1}^{N}(|v_n|^2 + |b_n|^2) , \quad (28)$$

$$H_\mathrm{c} = \frac{1}{4}\sum_{n=1}^{N}\Re(v_n b_n^*) , \quad (29)$$

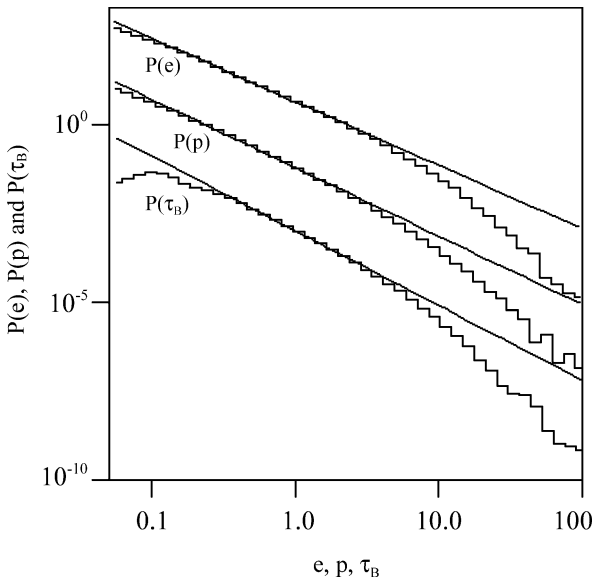
$$H_\mathrm{m} = \sum_{n=1}^{N}(-1)^n \frac{|b_n|^2}{k_n} . \quad (30)$$

The energy dissipation rate in the GOY MHD shell model can be defined as [6]

$$\varepsilon(t) = \nu \sum_{n=1}^{N} k_n^2 |v_n|^2 + \mu \sum_{n=1}^{N} k_n^2 |b_n|^2 . \quad (31)$$

A statistical analysis of intermittent events in the GOY MHD shell model shows that the PDFs of event energies, duration and waiting times display power law tails (Figs. 15 and 16) [6,14]. This behavior is consistent with the statistics of intermittent events observed in the solar wind, reversed field pinch devices and solar corona.

In closing this section, it can thus be remarked that the study of intermittent events in MHD turbulence, mainly based on wavelet transforms and PDFs of various event parameters, has provided a quite detailed picture of the coherent structures produced by the self-organization of the MHD turbulence cascade at small scales. The analysis of data from solar wind and RFP devices has demonstrated that the most intermittent structures in incompressible MHD are one dimensional current sheets associated with tangential discontinuities. The statistical analysis of intermittent events occurring in solar wind, laboratory plasma and magnetic structures of the solar corona has shown that these events are characterized by power law distributions of event energy and waiting times, which are naturally reproduced by MHD turbulence models. The power law behavior of the waiting time distribution indicates the



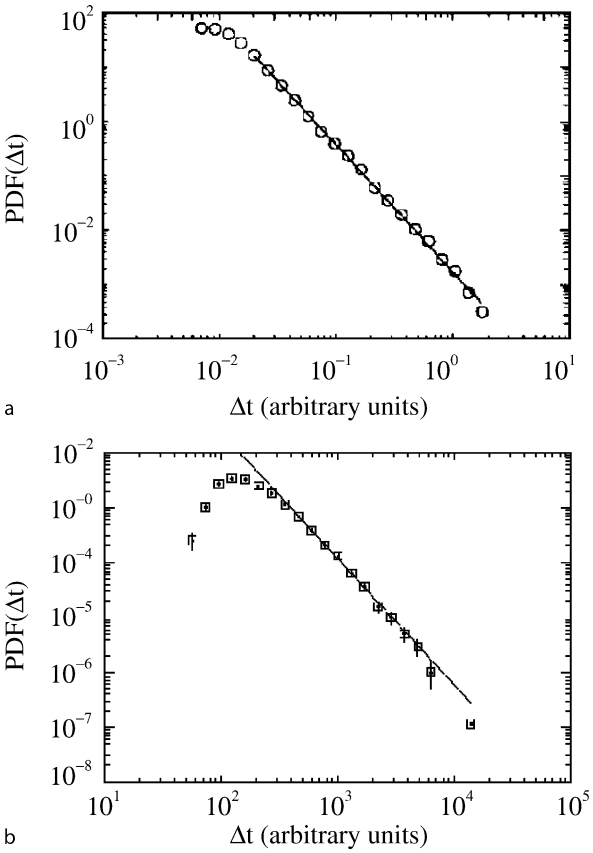
Self-Organization in Magnetohydrodynamic Turbulence, Figure 15

Total energy distribution $P(e)$, peak energy distribution $P(p)$ and burst duration distribution $P(\tau_B)$ for the GOY MHD shell model. The values of $P(e)$ and $P(\tau_B)$ are offset by a factor 10^2 and 10^{-2} respectively. The straight lines represent power law fits, with exponents $\alpha_e = 1.8$, $\alpha_p = 2.05$ and $\alpha_{\tau_B} = 2.2$ respectively

presence of long range time correlations which can be attributed to the non-linear dynamics of the turbulent energy cascade.

Future Directions

In the previous sections we have discussed how coherent dissipative structures are spontaneously formed by non-linear interactions both in laboratory and in natural plasmas. A further understanding of these processes requires an increase in the performance of computers, which will allow one to simulate 3D MHD turbulence at Reynolds numbers larger than 10^3, which represents the actual limit. On the other hand, an improved effort in the analysis of solar wind space data and laboratory plasma devices measurements should also be necessary. An increased knowledge of these phenomena is crucial in order to try to control the transport processes in fusion plasmas and then to improve the plasma confinement, since the presence of coherent structures inside turbulence greatly affect the plasma diffusion process. Moreover a better comprehension of the self organization in MHD turbulence can greatly help in explaining some basic phenomena in solar corona (flares and coronal heating), which occur in a region where in situ observations cannot be performed.



Self-Organization in Magnetohydrodynamic Turbulence, Figure 16

Probability density function of waiting times between successive intermittent events in the GOY MHD shell model. a Waiting time PDF for intermittent events identified in the time series of the magnetic field variables at the largest wavenumber shell. The *solid line* represents a power law with exponent $\beta = 2.13$. b Waiting time PDF for intermittent events identified in the time series of the energy dissipation rate ε (Eq. 31). The solid line represents a power law with exponent $\beta = 2.33$. In both cases, intermittent events were identified through the iterative procedure proposed in [6] and described in the text, using the threshold $\langle\cdot\rangle + 2\sigma$

Bibliography

Primary Literature

1. Aschwanden MJ (2004) Physics of the solar corona: An introduction. Springer, Berlin
2. Bak P, Tang C, Wiesenfeld K (1987) Self-organized criticality: An explanation of 1/*f* noise. Phys Rev Lett 59:381–384
3. Belcher JW, Davis L (1971) Large-amplitude Alfvèn waves in the interplanetary medium-II. J Geophys Res 76:3534–3563
4. Bellan PM (2000) Spheromaks: A practical application of magnetohydrodynamic dynamos and plasma self-organization. Imperial College Press, London

5. Biskamp D (1993) Non-linear magnetohydrodynamics. Cambridge University Press, Cambridge
6. Boffetta G, Carbone V, Giuliani P, Veltri P, Vulpiani A (1999) Power laws in solar flares: Self-organized criticality or turbulence? Phys Rev Lett 83:4662–4665
7. Bohr T, Jensen MH, Paladin G, Vulpiani A (1998) Dynamical systems approach to turbulence. Cambridge University Press, Cambridge
8. Burlaga LF (1991) Multifractal structure of the interplanetary magnetic field Voyager 2 observations near 25 AU, 1987 1988. Geophys Res Lett 18:69–72
9. Camussi R, Guj G (1997) Orthonormal wavelet decomposition of turbulent flows: Intermittency and coherent structures. J Fluid Mech 348:177–199
10. Carbone V, Veltri P (1987) A simplified cascade model for M.H.D. turbulence. Astron Astrophys 188:239–250
11. Carbone V, Veltri P (1992) Relaxation processes in MHD: A triad-interaction model. Astron Astrophys 259:359–372
12. Carbone V (1994) Scaling exponents of the velocity structure functions in the interplanetary medium. Ann Geophys 12:585–590
13. Carbone V, Sorriso-Valvo L, Martines E, Antoni V, Veltri P (2000) Intermittency and turbulence in a magnetically confined fusion plasma. Phys Rev E 62:49–56
14. Carbone V, Cavazzana R, Antoni V, Sorriso-Valvo L, Spada E, Regnoli G, Giuliani P, Vianello N, Lepreti F, Bruno R, Martines E, Veltri P (2002) To what extent can dynamical models describe statistical features of turbulent flows? Europhys Lett 58:349–355
15. Carbone V, Lepreti F, Sorriso-Valvo L, Veltri P, Antoni V, Bruno R (2004) Scaling laws in plasma turbulence. Riv Nuovo Cim 27(8-9):1–108
16. Crosby NB, Aschwanden MJ, Dennis BR (1993) Frequency distributions and correlations of solar *X*-ray flare parameters. Solar Phys 143:275–299
17. Dahlburg JP, Montgomery D, Doolen GD, Turner R (1986) Turbulent relaxation to a force-free field-reversed state. Phys Rev Lett 57:428–431
18. De Bartolo R, Carbone V (2006) The role of the basic three-modes interaction during the free decay of magnetohydrodynamic turbulence. Europhys Lett 73:547–552
19. Desnyansky VN, Novikov EA (1974) Evolution of turbulence spectra to self-similar regime. Atmos Oceanic Phys 10:127–136
20. Dobrowolny M, Mangeney A, Veltri P (1980) Fully developed anisotropic hydromagnetic turbulence in interplanetary space. Phys Rev Lett 45:144–147
21. Elsässer WM (1950) The hydromagnetic equations. Phys Rev 79:183–183
22. Farge M (1992) Wavelet transforms and their applications to turbulence. Ann Rev Fluid Mech 24:395–457
23. Frisch U, Pouquet A, Leorat J, Mazure A (1975) Possibility of an inverse cascade of magnetic helicity in magnetohydrodynamic turbulence. J Fluid Mech 68:769–778
24. Fyfe D, Montgomery D (1976) High-beta turbulence in two-dimensional magnetohydrodynamics. J Plasma Phys 16: 181–191
25. Heyvaerts J, Priest ER (1984) Coronal heating by reconnection in DC current systems – A theory based on Taylor's hypothesis. Astronom Astrophys 137:63–78
26. Gledzer EB (1973) System of hydrodynamic type admitting two quadratic integrals of motion. Sov Phys Dokl 18:216–217
27. Giuliani P, Carbone V (1998) A note on shell models for MHD turbulence. Europhys Lett 43:527–532
28. Giuliani P (1999) Non-linear MHD waves and turbulence, edited by Passot T, Sulem PL. Lect Notes Phys 536:331–345
29. Iroshnikov PS (1963) Turbulence of a conducting fluid in a strong magnetic field. Astron Zhur 40:742–750 (in Russian, translated in Sov Astron 7:566–571)
30. Kolmogorov AN (1941) The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. Dokl Akad Nauk SSSR 30:299–303 (in Russian translated in Proc Royal Soc London 434:9–13)
31. Kraichnan RH (1965) Inertial range spectrum of hydromagnetic turbulence. Phys Fluids 8:1385–1387
32. Krucker S, Benz AO (1998) Energy Distribution of Heating Processes in the Quiet Solar Corona. Astrophys J 501:L213–L216
33. Lepreti F, Carbone V, Veltri P (2001) Solar flare waiting time distribution: Varying rate Poisson or Lévy function? Astrophys J 555:L133–136
34. Lin RP, Schwartz RA, Kane SR, Pelling RM, Hurley KC (1984) Solar hard *X*-ray microflares. Astrophys J 283:421–425
35. Marsch E, Liu S (1993) Structure functions and intermittency of velocity fluctuations in the inner solar wind. Ann Geophys 11:227–238
36. Marsch E, Tu CY (1994) Non-Gaussian probability distributions of solar wind fluctuations. Ann Geophys 12:1127–1138
37. Tu CY, Marsch E (1995) MHD structures, waves and turbulence in the solar wind: Observations and theories. Space Sci Rev 73:1–210
38. Matthaeus WH, Montgomery D (1980) Selective decay hypothesis at high mechanical and magnetic Reynolds numbers. Ann NY Acad Sci 357:203–222
39. Matthaeus WH, Goldstein ML, Montgomery D (1982) Turbulent generation of outward-traveling interplanetary alfvnic fluctuations. Phys Rev Lett 51:1484–1487
40. Obukhov AM (1971) Some general characteristic equations of the dynamics of the atmosphere. Atmos Oceanic Phys 7:41–45
41. Onorato M, Camussi R, Iuso G (2000) Small scale intermittency and bursting in a turbulent channel flow. Phys Rev E 61:1447–1454
42. Parker EN (1983) Magnetic neutral sheets in evolving fields. II. Formation of the solar corona. Astrophys J 264:642–647
43. Parker EN (1988) Nanoflares and the solar *X*-ray corona. Astrophys J 330:474–479
44. Parnell CE, Jupp PE (2000) Statistical Analysis of the Energy Distribution of Nanoflares in the Quiet Sun. Astrophys J 529:554–569
45. Pearce G, Rowe A, Yeung J (1993) A statistical analysis of HRD *X*-ray solar flares. Astrophys Space Sci 208:99–111
46. Pouquet A, Meneguzzi M, Frish U (1986) Growth of correlations in magnetohydrodynamic turbulence. Phys Rev A 33:4266–4276
47. Rosenbluth MN, Monticello DA, Strauss HR (1976) Numerical studies of non-linear evolution of kink modes in tokamaks. Phys Fluids 19:1987–1996
48. Rostagni G (1995) RFX: An expected step in RFP research. Fusion Eng Design 25:301–313
49. Servidio S, Carbone V (2005) Non-linear dynamics of inviscid reduced MHD plasmas: The appearance of quasi-single-helicity states. Phys Rev Lett 95:045001
50. Sorriso-Valvo L, Carbone V, Veltri P, Consolini G, Bruno R (1999) Intermittency in the solar wind turbulence through proba-

bility distribution functions of fluctuations. Geophys Res Lett 26:1801–1804
51. Taylor JB (1974) Relaxation of toroidal plasma and generation of reverse magnetic fields. Phys Rev Lett 33:1139–1141
52. Taylor J (1986) Relaxation and magnetic reconnection in plasmas. Rev Mod Phys 58:741–763
53. Ting AC, Matthaeus WH, Montgomery D (1986) Turbulent relaxation processes in magnetohydrodynamics. Phys Fluids 29:3261–3274
54. Veltri P, Mangeney A (1999) Scaling laws and intermittent structures in solar wind MHD turbulence. In: Habbal SR, Esser R, Hollweg JV, Isenberg PA (eds) Solar wind nine, Proceedings of the ninth international solar wind conference, AIP Conference Proceedings vol 471 pp 543–546
55. Veltri P, Nigro G, Malara F, Carbone V, Mangeney A (2005) Intermittency in MHD turbulence and coronal nanoflares modelling. Non-lin Proc Geophys 12:245–255
56. Wesson J (2003) Tokamaks. Oxford Science, Oxford
57. Wheatland MS, Sturrock PA, McTiernan JM (1998) The waiting-time distribution of solar flare hard *X*-Ray bursts. Astrophys J 509:448–455
58. Woltjer L (1958) A theorem on force-free magnetic fields. Proc Natl Acad Sci 44:489–491
59. Yamada M (1999) Study of magnetic helicity and relaxation phenomena in laboratory plasmas. AGU Geophys Monogr 111:129–140
60. Yamada M, Ohkitani K (1987) Lyapunov spectrum of a chaotic model of three dimensional turbulence. J Phys Soc Japan 56:4210–4213
61. Zank GP, Matthaeus WH (1992) The equations of reduced magnetohydrodynamics. J Plasma Phys 48:85–100.

Books and Reviews

Bruno R, Carbone V (2005) The solar wind as a turbulence laboratory. Living Rev Solar Phys. 2: http://www.livingreviews.org/lrsp-2005-4
Biskamp D (2003) Magnetohydrodynamic turbulence. Cambridge University Press, Cambridge
Frisch U (1995) Turbulence: The legacy of A. N. Kolmogorov. Cambridge University Press, Cambridge
Moffat HK (1978) Magnetic field generation in electrically conducting fluids. Cambridge University Press, Cambridge

Self-organized Criticality and Cellular Automata

Michael Creutz
Physics Department,
Brookhaven National Laboratory, Upton, USA

Article Outline

Glossary
Definition of the Subject
Introduction
Self-Organized Criticality
Cellular Automata
The Sandpile Revisited
Future Directions
Acknowledgments
Bibliography

Glossary

Abelian group A mathematical group wherein all the elements commute.

Avalanche A possibly large disturbance induced in a system by a small perturbation.

Cellular automaton This refers to the dynamics of a collection of cells each of which can be in a finite set of states. The evolution is discrete, with the state of a cell at the next time step being dependent only on its previous state and that of its neighbors.

Chaos The tendency of a system of a few degrees of freedom to exhibit highly erratic behavior characterized by an infinite range of time scales.

Self-organized criticality The tendency of certain discrete and dissipative dynamical systems to evolve to a state where changes occur over all possible length scales.

Definition of the Subject

Self-organized criticality is a concept invoked to explain the frequent occurrence of fractal structures and multi-scale phenomena in nature. In contrast with the ideas of chaos, here simple common features appear in systems with many degrees of freedom. For modeling this phenomenon, cellular automata provide an elegant class of dynamical systems which are easily simulated numerically.

Introduction

Cellular automata provide a fascinating class of dynamical systems based on very simple rules of evolution yet capable of displaying highly complex behavior. These include simplified models for many phenomena seen in nature. Among other things, they provide insight into self-organized criticality, wherein dissipative systems naturally drive themselves to a critical state with important phenomena occurring over a wide range of length and time scales.

This article begins with an overview of self-organized criticality. This is followed by a discussion of a few examples of simple cellular automaton systems, some of which may exhibit critical behavior. Finally, some of the fascinating exact mathematical properties of the Bak–Tang–Wiesenfeld sandpile model [1] are discussed.

Self-Organized Criticality

Self-organized criticality refers to the tendency of many dynamical systems to naturally drive themselves to a state displaying fluctuations over a wide range of scales [1]. The concept is invoked as a possible "explanation" of the omnipresent multi-scale structures throughout the natural world, ranging from the fractal structure of mountains, to the power law spectra of earthquake sizes [2]. Recent applications include such diverse topics as punctuated evolution [3] and traffic flow [4]. The concept has even been invoked to explain the unpredictable nature of economic systems; i. e. why you can not beat the stock market [5].

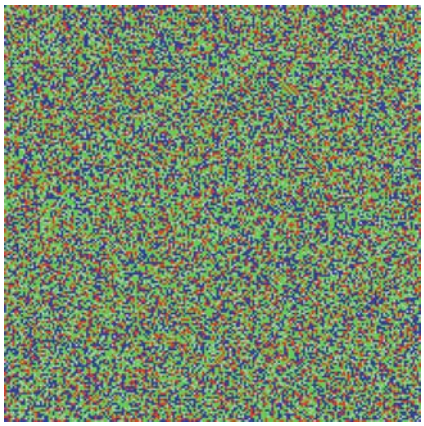
Self-organized criticality nicely compliments the concept of chaos. In the latter, dynamical systems with a few degrees of freedom, say as little as three, can display highly complex behavior, often generating beautiful fractal structures. With self-organized criticality, we start instead with systems of many degrees of freedom, and find a few general common features.

Another attractive feature of both self organized criticality and chaos is the ease with which computer models can be implemented and the elegance of the resulting graphics. Most of the figures in this chapter were produced using my publicly available set of programs "xtoys" [6]. Indeed, much of this presentation is based on my similar article in [7].

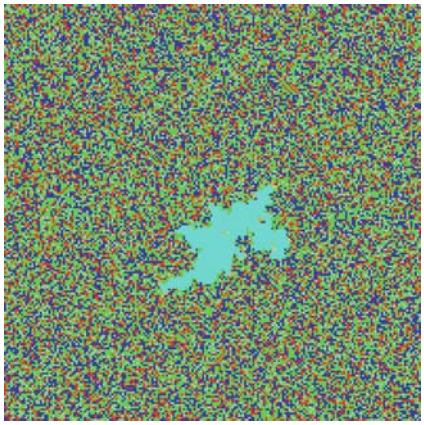
The paradigm for the phenomenon is the sandpile. On slowly adding grains of sand to an empty table, a pile will grow until its slope becomes critical and avalanches start spilling over the sides. If the slope becomes too large, a large catastrophic avalanche is likely, and the slope will reduce. If the slope is too small, then the sand will accumulate to make the pile steeper. Ultimately one should obtain avalanches of all sizes, with the prediction of the size for the next avalanche being impossible to determine without actually running the experiment.

The original Bak, Tang, Wiesenfeld paper [1] presented a particularly simple model to mimic the sandpile idea. For this, each site of a two dimensional lattice has a state represented by a positive integer z_i. This integer can be thought of as representing the amount of sand at that location, or in another sense it represents the slope of the sandpile at that point. Neither of these analogies is fully accurate, the model has aspects of each.

The dynamics follows by setting a threshold z_T above which any given z_i is unstable. Without loss of generality, I take this threshold to be $z_T = 3$. Time now proceeds in discrete steps. In one such step each unstable site with $z_i \geq 4$ "tumbles" or "topples," dropping by four and adding one grain to each of its four nearest neighbors. This



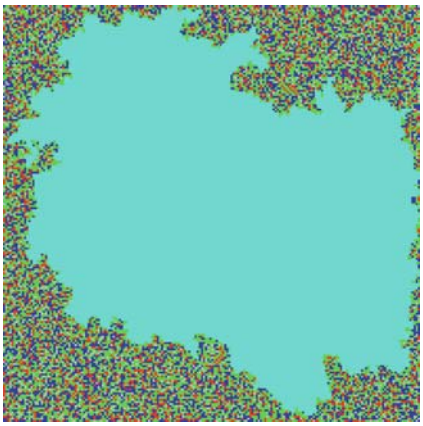
Self-organized Criticality and Cellular Automata, Figure 1
The sandpile model in the final stable state after adding lots of sand to random places. The lattice is 198 cells by 198 cells. The color code is *gray*, *red*, *blue*, and *green* for heights 0, 1, 2, and 3, respectively. Despite the lack of obvious patterns, subtle correlations are present; for example no two adjacent sites have height zero



Self-organized Criticality and Cellular Automata, Figure 2
An ongoing avalanche obtained by adding a small amount of sand to the configuration in Fig. 1. Stable sites which have tumbled during the avalanche are distinguished by being colored *light blue*. The still active sites are colored *yellowish brown*

may produce other unstable sites, and thus an avalanche can ensue. This proceeds for further time steps until all sites are stable. Figure 1 shows a typical configuration on a 198 by 198 lattice after lots of random sand addition followed by relaxation. Figure 2 shows an avalanche proceeding on this lattice, and Fig. 3 shows the final avalanch region after the system reaches stability.

A natural experiment consists of adding a grain of sand to a random site and measuring the number of topplings and the number of time steps for the resulting avalanche.



Self-organized Criticality and Cellular Automata, Figure 3
The final state after the avalanch in Fig. 2 has completed. The sites which tumbled during the avalanche are distinguished by being colored *light blue*. Note that the final avalanche region is simply connected. This is a general result proven later in the text

Repeating this many times to gain statistics, the distribution of avalanche sizes and lengths displays a power law behavior, with all sizes appearing. In [8] such experiments showed that the distribution of the number of tumbling events s in an avalanche empirically scales as

$$P(s) \sim s^{-1.07} \tag{1}$$

and the number of time steps τ for avalanches scales as

$$P(\tau) \sim \tau^{-1.14} . \tag{2}$$

This model has been extensively studied analytically. While as yet there is no exact calculation of these exponents, a lot is known. In particular, the critical ensemble is well characterized. I will return to these points later.

The extent to which laboratory experiments reproduce these phenomena is somewhat controversial. A study of avalanche dynamics [9] in rice piles showed power laws with long-grain rice, but more ambiguous results followed similar experiments with short-grain rice.

Cellular Automata

The sandpile model is a simple example of a system of cellular automata [10,11]. Each site or "cell" of our lattice follows a prescribed rule evolving in discrete time steps. At each step, the new value for a cell depends only on the current state of itself and its neighbors. These systems are fascinating in that deceptively simple rules can give rise to extremely complex behavior. Furthermore, slight changes in the rules can dramatically change their behavior.

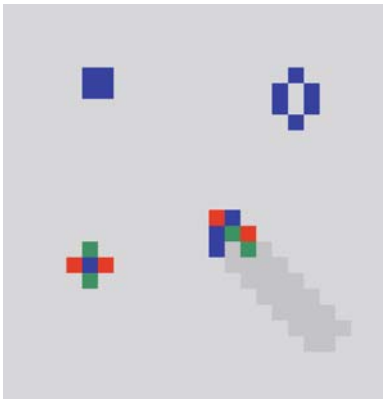
Even though the formulation of a cellular automaton may seem almost trivial, there are a huge number of possible rules. For example, suppose I consider two dimensional models where each cell can take only one of two possible states. These might be referred to as unset or set bits, or more figuratively as "dead" or "alive." Suppose furthermore that I restrict myself to rules where the evolution of a given cell to the next time step depends only on the current values of the cell and each of its eight neighbors. In this case there are $2^9 = 512$ possible arrangements for the cell and its neighbors. A general rule needs to specify the next state of the cell for each of these arrangements. This gives $2^{512} = 1.3 \times 10^{154}$ possible rules. Given that the universe is only of order 4×10^{17} seconds old, clearly only a vanishing fraction of these rules have a chance of being studied in any of our lifetimes.

A simple subset of rules called "totalistic" have the state of the updated cell only depend on the total number of living neighbors. With the eight cell neighborhood, there are nine possible values for this sum, and the new value for the cell requires specification of the new state for each of these as well as the current state of the cell. This gives $2^{18} = 262,144$ rules; still large, but not truly astronomical. If I restrict the rule to depending on the total of only the four nearest neighbors, I then have a modest $2^{10} = 1024$ cases to consider. Other than the sandpile model, most of the following will be restricted to such totalistic rules.

With a discrete set of states, cellular automata have the appealing feature of being easily implementable entirely by logical operations, the natural functions of computer circuitry. Also, the state of several cells can be stored and manipulated within a single computer word. Using such tricks, these models can often be implemented to run extremely fast, leading to hope that such models may supply simulation methods as good as or better than the conventional use of floating point fields on a discrete grid. With this motivation, considerable attention has been paid to cellular automata that may simulate fluid flow. Another advantage of this approach is the ability to work with arbitrary boundary conditions. These topics go beyond the scope of this article. A nice review can be found in [12].

Conway's Life

Perhaps the most famous cellular automaton model is Conway's "Game of Life" [13]. For this there exists a vast literature; so, I will only mention a couple of interesting features. The rule involves the eight cell neighborhood, and if a cell is initially "dead" it becomes alive if and only if it has exactly three live neighbors, or "parents". A living cell dies of loneliness if it has less than two live neighbors,



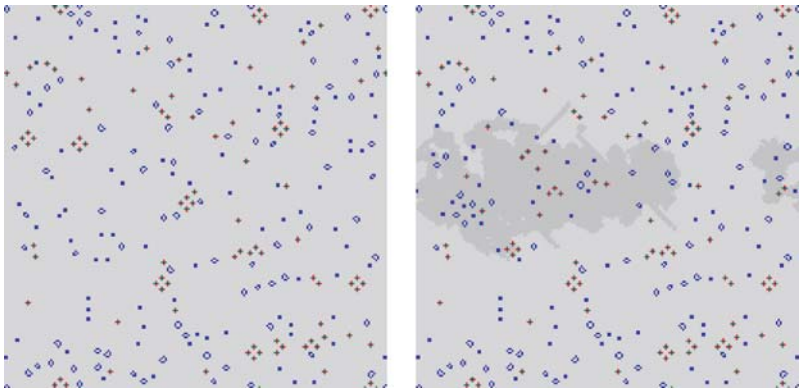
Self-organized Criticality and Cellular Automata, Figure 4
Some living configurations in life. The *top* two are stable patterns. The *lower left* shows a "blinker" or "traffic light" which oscillates with a period of two. On the *lower right* is a glider, which propagates diagonally through the lattice. *Blue* denotes a state that is and just was alive, *red* is newborn, and *green* represents just died. The track of the glider is *darkened* slightly over the remaining *gray background* to show its motion

and of overcrowding if it has more than three live neighbors. Only in the case of exactly two or three live neighbors does it survive.

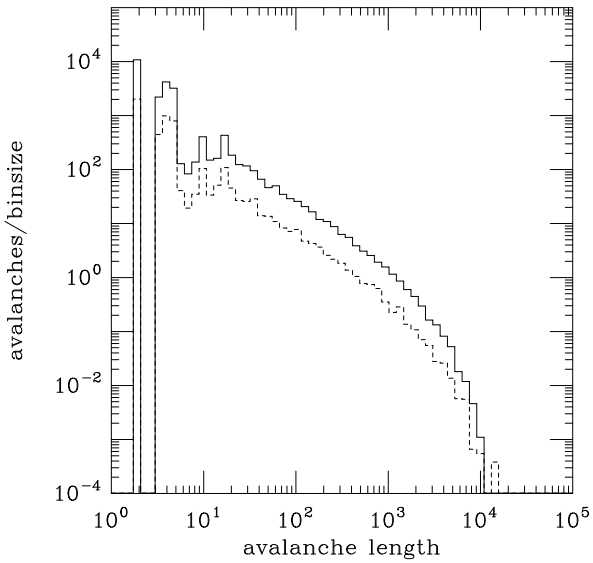
While simple to state, this model displays fascinating complexity. There are simple isolated sets of live cells that quietly survive, such as a block of four neighboring live cells forming a two by two square. Other configurations oscillate, such as three live cells in a row, which alternate between being vertically and horizontally oriented. A particularly amusing local configuration has five live cells; say starting with coordinates $\{(0,0),(0,1),(0,2),(1,2),(2,1)\}$. After four time steps this configuration returns to its original shape, but displaced by $(-1,1)$. On an otherwise empty board, this "glider" continues to propagate as a single entity. In an on-screen simulation, it appears much as a small insect crawling about. Some elementary configurations are shown in Fig. 4. A large collection of fascinating life configurations can be found in the Wikipedia [14].

Gliders allow information to be propagated over long distances, and it has been proven that with a complicated enough initial configuration, one can construct a computer out of live cells on a life board [13]. Special sub-configurations form the analog of electronic gates, which can control beams of gliders representing bits. Indeed, since life is capable of universal computation, one might imagine a life board programmed to simulate the game of life.

There is some limited evidence that the game of life also displays self-organized criticality [15,16]. One can repeatedly throw down gliders, which collide and create a background of static and oscillating clumps. While oscillators of arbitrarily long period are known to exist, those with period longer than two are extremely rare and almost never created from unorganized initialization. Once the system has settled into a loop, then another glider can be tossed on, giving a disturbance. An avalanche is defined to occur during the period until the system again goes into an oscillating state. Figure 5 shows the effect of such a disturbance. In Fig. 6 I show the distribution of such avalanches as measured on modest lattices. There is a hint of a power law superposed on additional structure from avalanches of only a few time steps, and a rounding at large times possibly due to finite size effects. The criticality of life remains controversial; [17] has looked unsuccessfully for a power



Self-organized Criticality and Cellular Automata, Figure 5
On the *left* is a configuration in life resulting from a random start and evolved until only stable and period two oscillators remain. On the *right* is the state after a small disturbance was introduced in the center and allowed to die out. Note the irregular shape of the disturbed region, which has been tinted a *darker gray*. The lattice here is 198 sites wide by 198 sites high, with periodic boundaries



Self-organized Criticality and Cellular Automata, Figure 6
The distribution of avalanches generated by adding gliders randomly to a system in the game of life consisting of stable and period two oscillators. An avalanche occupies the period until the system has relaxed again into such a periodic state. The *solid line* represents 25,000 avalanches on a 512 by 512 lattice, and the *dashed line* is for 6,000 avalanches on a 1024 by 1024 system. This figure is taken from [16]

law distribution of activity as one moves in from a source on the boundary. The relation between these two experiments is unclear.

Fredkin's Modulo-Two Rule

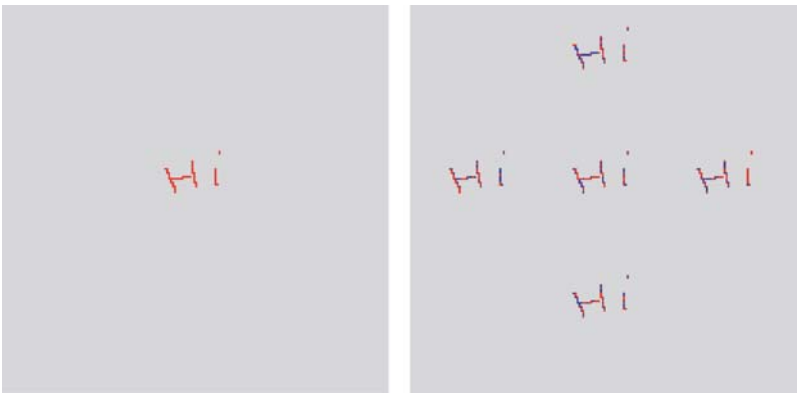
An extremely simple but highly amusing rule takes at each time step the "exclusive or" (XOR) operation between a site and its neighbors. This rule has the remarkable property of self replication [18]. Starting with any given initial pattern, after 2^n time steps copies of the original state occupy positions separated by 2^n spatial sites from the original in every direction as specified in the chosen neighborhood. In Fig. 7 I show an example of this with the four cell neighborhood.

In this rule, the pattern is generally rather complex just before returning to the replicated case, i. e. after $2^n - 1$ steps. Figure 8 shows the pattern obtained from a single set pixel after this rule has been applied for 63 time steps using the four nearest cells as the neighborhood. Note the fractal structure. In one more time step, all but five copies of the original set bit die.

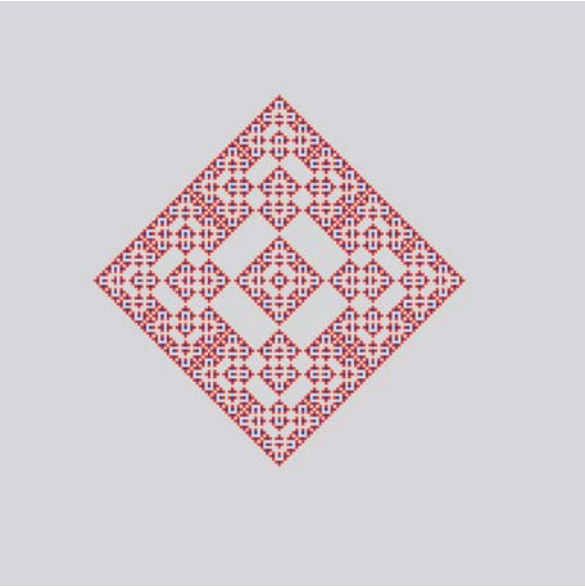
Unlike most cellular automaton rules, this gives a dynamics which in some sense is not really "complex". In most cases the simplest way to predict the evolution of a cellular automaton rule is to actually run it. Here, however, there is an easier way to predict what the final pattern will look like; it is always an XOR operation between several displaced copies of the configuration that appeared 2^n time steps in the past. Despite the lack of complexity, this rule shows rather dramatically that cellular automata are capable of "reproduction".

Reversible Rules

Reversibility is rather elusive among cellular automata. In the game of life, a single isolated cell immediately dies leaving no trace; thus it is impossible from the state at a given time to reconstruct what was there one time step back. A related difficult problem is to construct "garden of Eden" configurations which are impossible to arrive at from any previous state [19].



Self-organized Criticality and Cellular Automata, Figure 7
Starting from the initial configuration *on the left*, the modulo two rule is evolved for 64 steps using the four nearest neighbors. At a certain stage, five copies of the original image appear. The *blue pixels* indicate which sites were also alive one step before



Self-organized Criticality and Cellular Automata, Figure 8
The state after applying 63 steps of the modulo two rule using the four nearest neighbors to an initial state of a single set bit. After the next time step this fractal structure decays into only five remaining live cells

Fredkin pointed out an interesting class of reversible rules based on an analogy with molecular dynamics [11]. In the later one specifies both the position and the velocities of a set of particles and evolves the system under Newton's equations with some given inter-particle force law. Reversal can then be accomplished by merely changing the signs of all the velocities.

In a cellular automaton an analog of velocity requires the value of the cells at two successive time steps. Based on this, Fredkin presented a very simple scheme using the previous state to generate a wide class of reversible rules. He considered taking an arbitrary automaton rule at a given time, and then added an exclusive or (XOR) operation of the result with the state one step back in time. These combined operations could then be reversed by merely interchanging two successive time steps, the analogy of reversing the velocities.

To see this more mathematically, suppose the state at time i is s_i, and the underlying rule begins by taking some arbitrary function $f(s_i)$. Then the full rule takes for the next time step $s_{i+1} = f(s_i)$ XOR s_{i-1}. Here the exclusive or operation is taken site by site over the entire lattice. Elementary properties of the XOR operation then give $s_{i-1} = f(s_i)$ XOR s_{i+1}, which is the identical rule for the time reversed dynamics.

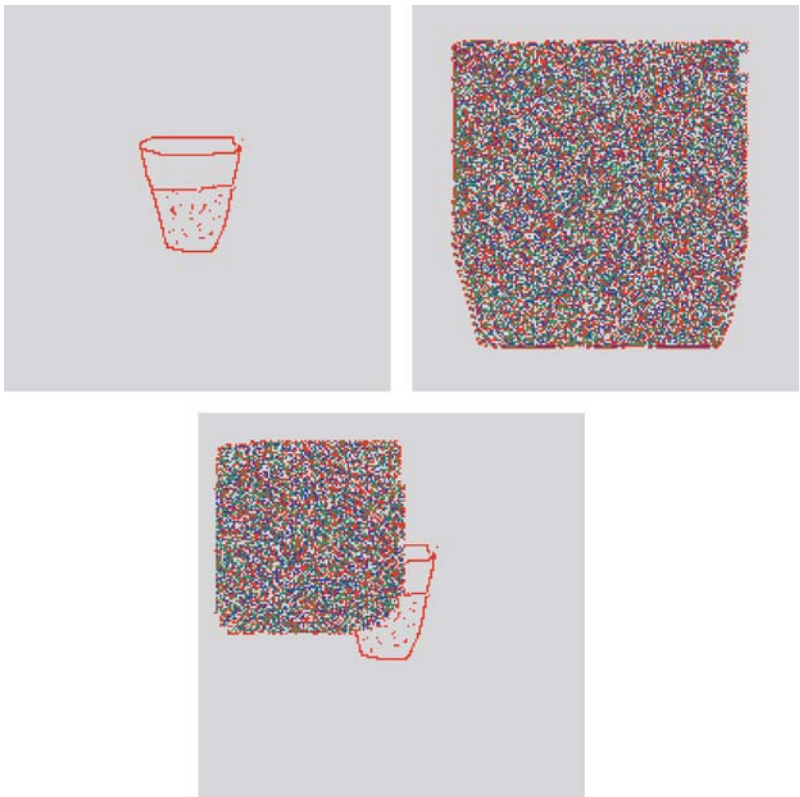
These rules provide a wonderful way to play with the concepts of entropy and reversibility. Indeed, an idealized universe of cellular automata enables experiments which would be impossible to carry out in the real world. In Fig. 9 I show the evolution of a simple image under such a rule. The experiment is a crude simulation of a beer glass shattering after being dropped on the floor. After a few steps it appears quite randomized. Reversal of the momenta of all relevant atoms in the beer glass would allow its reconstruction. In the model this is easily accomplished by swapping two time steps. After reversal, continuing with the same rule reconstructs the original image. At all stages the "information" contained in the system must be constant, even though the image may appear of drastically different complexity.

The reconstruction process is highly sensitive to the reversal being precise. The analog here is to the sensitivity to initial conditions in dynamical systems. In the bottom of Fig. 9c I try to reproduce the beer cup from its shards as in the above experiment, except that now at the time of reversal I modify the state of exactly one pixel. The reversal process recovers the original image only in regions outside the "light cone" for the modified pixel. As the disturbance can only propagate to neighbors in one time step, pixels outside n steps can not know of the change before an equal number of time steps. This use of an XOR operation to generate reversible complex mappings is an integral part of the Data Encryption Standard; see, for example, [20].

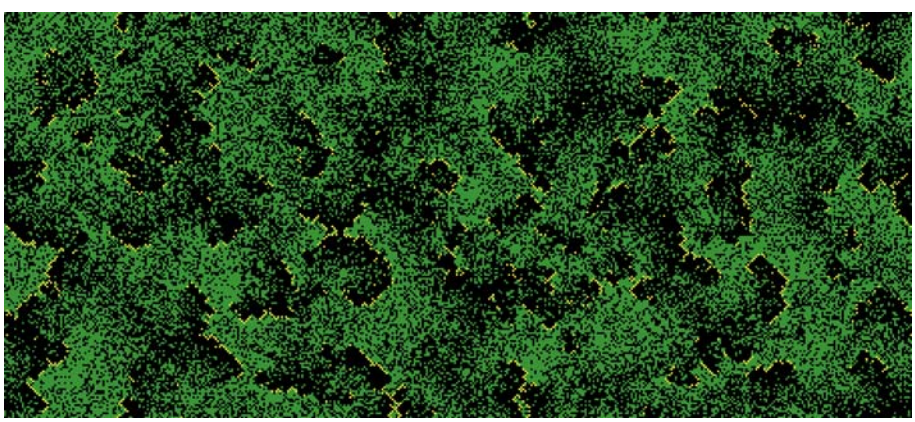
Forest Fires

An amusing model of forest fires has three possible states per cell, empty, a tree, or a fire. For the updating step, any empty site can have a tree born with a small probability. At the same time, any existing fire spreads to neighboring trees leaving its own cell empty. The rule here differs from those discussed previously in having a stochastic nature. As the system is made larger, the growth rate for the trees should decrease to just enough to keep the fires going.

If too many trees grow, one obtains a large fire reducing their density, while if there are too few trees, fires die out. On a finite system, one should light a fire somewhere to get the system started. On the other hand, as the system becomes larger, the growth rate for the trees can be reduced without the fire expiring. In a steady state the system has fire fronts continually passing through the system, as illustrated in Fig. 10. Perhaps there is a moral here that one should be careful about extinguishing all fires in the real world, for this may enhance the possibility for a catastrophic uncontrollable fire. It is not entirely clear whether this model is actually critical. What seems to happen on



Self-organized Criticality and Cellular Automata, Figure 9
The encryption of a glass of beer. The original rule uses the eight cell neighborhood with births on 1, 3, 5, and 7 neighbors and survivors on exactly 1 neighbors. The rule is modified at each step by XOR'ing the result with the history one time step back. Swapping two adjacent time steps will bring the glass back exactly. The first figure is the starting configuration, the second after 50 steps of evolution. At this point one bit in the *upper left hand quadrant* is flipped, and the dynamics is reversed. The glass is restored in all places beyond 50 steps from the flipped bit. Note the effect of a "speed of light" in the problem



Self-organized Criticality and Cellular Automata, Figure 10
A snapshot of the forest fire model on a 450 by 200 periodic lattice. Trees are continuously burning at a slow rate, while fires burn them down and spread to nearest neighbor trees. Here the four cell neighborhood is used

large systems is that stable spiral structures form and set up a steady rotation. For a review of this and several related models, see [21].

The Sandpile Revisited

Very little is known analytically about general cellular automata. However, in a series of papers, Deepak Dhar and co-workers have shown that the sandpile model has some rather remarkable mathematical properties [22,23,24,25]. In particular, the critical ensemble of the system has been well characterized in terms of an Abelian group. In the following I will generally follow the discussions given in [2,26].

Dhar introduced the useful toppling matrix $\Delta_{i,j}$ with integer elements representing the change in the height, z at site i resulting from a toppling at site j [22]. More precisely, under a toppling at site j, the height at any site i becomes $z_i - \Delta_{i,j}$. For the simple two dimensional sand model the toppling matrix is thus

$$\begin{array}{ll} \Delta_{i,j} = 4 & i = j \\ \Delta_{i,j} = -1 & i, j \text{ nearest neighbors} \\ \Delta_{i,j} = 0 & \text{otherwise.} \end{array} \tag{3}$$

For this discussion there is little special to the specific lattice geometry; indeed, the following results easily generalize to other lattices and dimensions. The analysis requires only that under a toppling of a single site i, that site has its slope decreased ($\Delta_{i,i} > 0$), the slope at any other site is either increased or unchanged ($\Delta_{i,j} \leq 0, j \neq i$), the total amount of sand in the system does not increase ($\sum_j \Delta_{i,j} \geq 0$), and, finally, that each site be connected through toppling events to some location where sand can be lost, such as at a boundary.

For the specific case in Eq. (3), the sum of slopes over all sites is conserved whenever a site away from the lattice edge undergoes a toppling. Only at the lattice boundaries can sand be lost. Thus the details of this model depend crucially on the boundaries, which we take to be open. A toppling at an edge loses one grain of sand and at a corner loses two.

The actual value of the maximum stable height z_T is unimportant to the dynamics. This can be changed by simply adding constants to all the z_i. Thus, as in Sect. "Self-Organized Criticality", I consider $z_T = 3$. With this convention, if all z_i are initially non-negative they will remain so, and I thus restrict myself to states C belonging to that set. The states where all z_i are non-negative and less than 4 are called stable; a state that has any z_i larger than or equal to 4 is called unstable. One conceptually useful configuration is the minimally stable state C^* which has all the heights at the critical value z_T. By construction, any addition of sand to C^* will give an unstable state leading to a large avalanche.

I now formally define various operators acting on the states C. First, the "sand addition" operator α_i acting on any C yields the state $\alpha_i C$ where $z_i = z_i + 1$ and all other z are unchanged. Next, the toppling operator t_i transforms C into the state with heights z'_j where $z'_j = z_j - \Delta_{i,j}$. The operator U which updates the lattice one time step is now simply the product of t_i over all sites where the slope is unstable,

$$UC = \prod_i t_i^{p_i} C \tag{4}$$

where $p_i = 1$ if $z_i \geq 4$; 0 otherwise. Using U repeatedly gives the relaxation operator R. Applied to any state C this corresponds to repeating U until no more z_i change. Neither U nor R have any effect on stable states. Finally, I define the avalanche operators a_i describing the action of adding a grain of sand followed by relaxation

$$a_i C = R\alpha_i C\,. \tag{5}$$

At this point it is not entirely clear that the operator R exists; in particular, it might be that the updating procedure enters a non-trivial cycle consisting of a never ending avalanche. I now prove that this is impossible. First note that a toppling in the interior of the lattice does not change the total amount of sand. A toppling on the boundary, however, decreases this sum due to sand falling off the edge. Thus, during an avalanche the total sand in the system is a non-increasing quantity. No closed cycle can have toppling at the boundary since this will decrease the sum. Next, the sand on the boundary will monotonically increase if there is any toppling one site further in. This also can not happen in a cycle; thus, there can be no topplings one site away from the edges. By induction there can be no toppling arbitrary distances in from the boundary; thus, there can be no cycle, and the relaxation operator exists. Note that for a general geometry this argument requires that every site be eventually connected to an edge where sand can be lost.

With a system lacking edges, such as under periodic boundaries, no sand would be lost and thus cycles are expected and easily observed. These models might be called "Escher models" after the artist constructing drawings of water flowing perpetually downhill and yet circulating in the system. While little is known about the dynamics of this variation on the sandpile model, some studies have been done under the nomenclature of "chip-firing games" [27]. It has been argued [28] that this lossless sand-

pile model on an appropriate lattice is capable of universal computation.

I now introduce the concept of recursive states. This set, denoted $\mathcal{R}$, includes those stable states which can be reached from any stable state by some addition of sand followed by relaxation. This set is not empty because it contains at least the minimally stable state C^*. Indeed, that state can be obtained from any other by carefully adding just enough sand to each site to make each z_i equal to three. Thus, one might alternatively define $\mathcal{R}$ as the set of states which can be obtained from C^* by acting with some product of the operators a_i.

It is easily shown that there exist non-recursive, transient states; for instance, no recursive state can have two adjacent heights both being zero. If you try to tumble one site to zero height, then it drops a grain of sand on its neighbors. If you then tumble a neighbor to zero, it dumps a grain back on the original site. One can also show that the self-organized critical ensemble, reached under random addition of sand to the system, has equal probability for each state in the recursive set. This is a consequence of the Abelian nature of this system, as discussed below.

The crucial results of [22,23,24,25] are that the operators a_i acting on stable states commute, and they generate an Abelian group when restricted to recursive states. I begin by showing that the operators commute, that is $a_i a_j C = a_j a_i C$ for all C. First I express the a's in terms of toppling and adding operators

$$a_i a_j C = \left(\prod_{k=1}^{n_1} t_{l_k} \right) \alpha_i \left(\prod_{k=n_1+1}^{n} t_{l_k} \right) \alpha_j C \tag{6}$$

where the specific number of topplings n_1 and n depend on i, j, and C. Acting on general states, the operators t and α all commute because they merely linearly add or subtract heights. Therefore I can shift α_i to the right in this expression:

$$a_i a_j C = \left(\prod_{k=1}^{n} t_{l_k} \right) \alpha_i \alpha_j C \,. \tag{7}$$

Now I rearrange the product of topplings. In the nontrivial case that the α-operators render either i or j (or both) unstable, the product must contain toppling operators corresponding to those unstable sites. I shift those operators to the right. Those operators constitute by definition the update operator, U, so I can write

$$a_i a_j C = \left(\prod t_{l_k} \right) U \alpha_i \alpha_j C \tag{8}$$

where the factors within the bracket are the remaining t's. Now, the update operator may leave some sites still unstable, and then the product must include further toppling operators; working on those sites, I can pull out another factor of the update operator. This procedure can be repeated until I have used all the toppling factors and the state is stable. Thus, I can identify the operator within the brackets in Eq. (8) as the relaxation operator R. But $\alpha_i \alpha_j C$ is the same state as $\alpha_j \alpha_i C$, so $a_i a_j C = a_j a_i C$.

A trivial consequence of this argument is that the total number of tumbling events occurring in the operations $a_i a_j C$ and $a_j a_i C$ are the same. Of course, if a particular site k tumbles it can be caused by either addition; the orders of the tumbling events may or may not be altered.

An intuitive argument that sand addition may be commutative uses an analogy with combining many digit numbers under long addition. The tumbling operation is much like carrying, except rather than transferring to the next digit, the overflow spreads to several neighbors. As addition is known to be Abelian, despite the confusing elementary-school rules, I might expect the sandpile addition rule also to be.

I now prove that the avalanche operators have unique inverses when restricted to recursive states; that is, there exists a unique operator a_i^{-1} such that $a_i(a_i^{-1} C) = C$ for all C in $\mathcal{R}$. This implies that the operators a_i acting on the recursive set generate an Abelian group. For any recursive state C I first find another recursive state such that a_i acting on it gives C, and I then show that this construction is unique.

I begin by adding a grain of sand at site i to the state C and then relax the system. This generates a new recursive state $a_i C$. Now since the state C is by assumption recursive, there is some way to add sand to regenerate C from any given state. In particular, there is some product P of addition operators a_j such that

$$C = P a_i C \,. \tag{9}$$

But the a's commute, so I have

$$C = a_i P C \tag{10}$$

and thus PC is a recursive state on which a_i gives C.

I must now show that this state is unique. Consider repeating the above process to find a series of states C_n satisfying

$$(a_i)^n C_n = C \,. \tag{11}$$

Because on a finite system the total number of stable states is finite, the sequence of states C_n must eventually enter a loop. I can run backwards around this loop by adding back the sand repeatedly to the given site. As the original

state C appears in resupplying the sand, C itself must itself belong to the loop. Calling the length of the loop m, I have $(a_i)^m C = C$. I now uniquely define $a_i^{-1} C = a_i^{m-1} C$.

I now have sufficient machinery to count the number of recursive states. As all such can be obtained by adding sand to C^*, I can write any state $C \in \mathcal{R}$ in the form

$$C = \left(\prod_i a_i^{n_i} \right) C^* \,. \tag{12}$$

Here the integers n_i represent the amount of sand to be added at the respective sites. However, in general there are several different ways to reach any given state. In particular, adding four grains of sand to any one site must force a toppling and is equivalent to adding a single grain to each of its neighbors. This can be expressed as the operator statement

$$a_i^4 = \prod_{j \in nn} a_j \tag{13}$$

where the product is over the nearest neighbors to site i. I can rewrite this equation by multiplying by the product of inverse avalanche operators on the nearest neighbors on both sides, thus obtaining for any site i

$$\prod_j a_j^{\Delta_{ij}} = E \tag{14}$$

where E is the identity operator. This allows me to shift the powers appearing in Eq. (12). Define N to be the number of sites in the system. If I label states by the vector $\mathbf{n} = (n_1, n_2, n_3, \ldots, n_N)$ I see that two states are equivalent if the difference of these vectors is of the form $\sum_j \beta_j \Delta_{ij}$ where the coefficients β_j are integers. These are the only constraints; if two states can not be related by toppling they are independent. Thus any vector $\mathbf{n}$ can be translated repeatedly until it lies in an N-dimensional hyper-parallelepiped whose base edges are the vectors Δ_{ji}, $j = 1, \ldots, N$. The vertices of this object have integer coordinates and its volume is the number of integer coordinate points inside it. This volume is just the absolute value of the determinant of Δ. Thus the number of recursive states equals the absolute value of the determinant of the toppling matrix Δ.

For large lattices this determinant can be found easily by Fourier transform. In particular, whereas there are $4N$ stable states, there are only

$$\exp\left(N \int_{(-\pi,-\pi)}^{(\pi,\pi)} \frac{\mathrm{d}^2 q}{(2\pi)^2} \ln(4 - 2q_x - 2q_y) \right) \simeq (3.2102\ldots)^N \tag{15}$$

recursive states. Thus starting from an arbitrary state and adding sand, the system "self-organizes" into an exponentially small subset of states forming the attractor of the dynamics.

An Isomorphism

Following [26], I now look into the consequences of stacking sand piles on top of one another. Given stable configurations C and C' with configurations z_i and z_i', I define the state $C \oplus C'$ to be that obtained by relaxing the configuration with heights $z_i + z_i'$. Clearly, if either C or C' is a recursive state, so is $C \oplus C'$.

Under the operation $\oplus$ the recursive states form an Abelian group isomorphic to the algebra generated by the a_i. First, the addition of a state C with heights z_i is equivalent to operating with a product of a_i raised to z_i, that is

$$B \oplus C = \left(\prod a_i^{z_i} \right) B \tag{16}$$

for any recursive state B. The operation $\oplus$ is associative and Abelian because the operators a_i are.

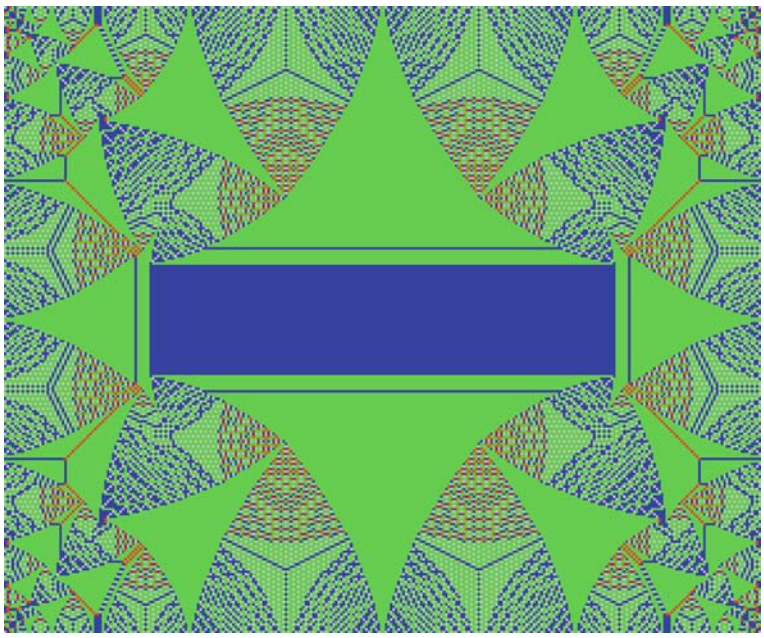
Since any element of a discrete group raised to the order of the group gives the identity, it follows that $a_i^{|\Delta|} = E$. This implies the simple formula $a_i^{-1} = a_i^{|\Delta|-1}$. The analog of this for the states is the existence of an inverse state, $-C$

$$-C = (|\Delta| - 1) \otimes C \,. \tag{17}$$

Here, $n \otimes C$ means adding n copies of C and relaxing. The state $-C$ has the property that for any state $B \oplus C \oplus (-C) = B$.

The state $I = C \oplus (-C)$ represents the identity and has the property $I \oplus B = B$ for every recursive state B. The state which is isomorphic to the operator a_i is simply $a_i I$. The identity state provides a simple way to check if a state, obtained for instance by a computer simulation, has reached the attractor, i. e. if a given state is a recursive state: A stable state is in $\mathcal{R}$ if and only if $C \oplus I = C$. The proof is simple. By construction, a recursive state has this property. On the other hand, since I is recursive, so is $C \oplus I$.

The identity state can be constructed by taking any recursive state, say C^* and repeatedly adding it to itself to use $|\Delta| \otimes C = I$. However, on any but the smallest lattices, $|\Delta|$ is a very large integer. A more economical scheme is to start with an empty table but use sandy boundary conditions which continually pour sand onto the table. Once it reaches a steady state, switch to open boundary condi-

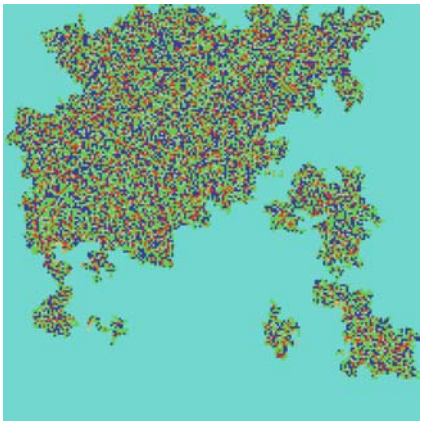


Self-organized Criticality and Cellular Automata, Figure 11
The identity state for the sandpile model on a 302 by 250 lattice. The color code is *gray*, *red*, *blue*, and *green* for heights 0, 1, 2, and 3, respectively

tions and let the sand run back off. This then relaxes to the desired identity. Figure 11 shows the identity state on a 302 by 250 lattice. Note the fractal structure, with features on many length scales.

Majumdar and Dhar [25] have constructed a simple "burning" algorithm to determine if a state belongs to the recursive set. For a given configuration, first add one particle to each of the edge sites and two particles to the corners. This again corresponds to imagining a large source of sand just outside the boundaries, which then tumbles one step onto the system. Then return to open boundaries and update according to the usual rules. If and only if the original state is recursive, this will generate an avalanche under which each site of the system tumbles exactly once. Also, the final state after the avalanche will be identical to the original. However, if the state is not recursive, some untumbled sites will remain. Figure 12 shows such a process underway on the configuration of Fig. 1. Here sites which have already burned are shown in cyan, while the remaining sites in the center have not yet tumbled. The small number of sites shown in orange are the still active sites, which eventually burn the entire remaining lattice.

The burning algorithm provides a simple way to prove that the avalanche regions are simply connected once one



Self-organized Criticality and Cellular Automata, Figure 12
The burning algorithm being applied to the state in Fig. 1. Burnt sites are *cyan*, burning sites are *orange*, and the remaining sites are colored as previously. This avalanche eventually tumbles every site exactly once

is in the critical state. In a burning process, any sub-lattice of the original will have all of its sites tumbled onto from outside. This is the condition for starting a burning on the sub-lattice. Thus, if a configuration is in the critical ensemble for the whole lattice, then any extracted piece of

this configuration on a subset of the original lattice is also in the critical ensemble of the extracted part. Now suppose that one constructs an avalanche with any initial addition to a state from the critical ensemble. In any subregion enclosed by this avalanche, sand will fall from the tumbling sites on its outside. Since the sub-lattice is itself in its own critical ensemble, this must induce an avalanche which, by the burning algorithm, will tumble all enclosed sites. Thus any avalanche on a state from the critical ensemble cannot leave untumbled any sites in a region isolated from the boundary, i. e. an untumbled island. This result that avalanches must be simply connected does not follow for states outside the recursive set, as can be easily demonstrated by considering a sandpile with a hole of empty sites in the middle.

The burning algorithm has several amusing consequences. One is that any configuration with only height 2 or 3 present is in the critical ensemble as long as the lattice has corners. For example, with all height 2, the burning will start at the four corners of a rectangular lattice and steadily work its way to the center of the system. Another consequence is that in addition to the tumbling region from an avalanche being simply connected, so will the smaller region where the number of tumblings exceed any fixed number; i. e. the region of sites that tumble twice or more is also simply connected.

Future Directions

Simple models as implemented by cellular automata provide a rich area for the study of complex phenomena. Some systems can self organize with physics at many scales, while others provide fascinating demonstrations of thermodynamic laws. I have only touched on a few issues here, leaving out many related topics such as lattice gasses, driven interfaces in random media, growth processes, and evolution. As the ease of programming and the speed of modern computers continue to rush forward, so will the fascination with such models.

Acknowledgments

I am thankful for discussions with many people, but most particularly P. Bak, D. Dar, E. Fredkin, N. Margolis, M. Paczuski, T. Toffoli, F. Van Scoy, and G. Vichniac. This manuscript has been authored under contract number DE-AC02-76CH00016 with the US Department of Energy. Accordingly, the US Government retains a non-exclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for US Government purposes.

Bibliography

Primary Literature

1. Bak P, Tang C, Wiesenfeld K (1987) Phys Rev Lett 59:381; (1988) Phys Rev A 38:3645
2. Bak P, Creutz M (1994) Fractals and self-organized criticality. In: Bunde A, Havlin S (eds) Fractals in Science. Springer, Berlin, pp 26–47
3. Paczuski M, Maslov S, Bak P (1996) Phys Rev E 53:414
4. Nagel K, Paczuski M (1995) Phys Rev E 51:2909
5. Levy M, Solomon S, Ram G (1996) Int J Mod Phys C 7:65
6. The latest version of the xtoys package is available at http://thy.phy.bnl.gov/www/xtoys/xtoys.html
7. Creutz M (1997) Cellular automata and self organized criticality. In: Bhanot G, Chen S, Seiden P (eds) Some new directions in science on computers. World Scientific, Singapore, pp 147–169
8. Christensen K (1992) Ph D Thesis, University of Aarhus
9. Frette V et al (1996) Nature 379:49
10. Wolfram S (1986) Theory and Applications of Cellular Automata. World Scientific, Singapore
11. Toffoli T, Margolus N (1987) Cellular Automata Machines. MIT Press, Cambridge
12. Bogosian B (1993) Nucl Phys B, Proc Suppl 30:204
13. Berlekamp E, Conway J, Guy R (1982) Winning Ways for your Mathematical Plays, vol 2. Academic Press, New York
14. Wikipedia (2007) Conway's Game of Life. http://en.wikipedia.org/wiki/Conway's_life. Accessed 6 Apr 2007
15. Bak P, Chen K, Creutz M (1989) Nature 342:780
16. Creutz M (1992) Nuclear Phys B, Proc Suppl 26:252
17. Bennett C, Bourzutschy M (1991) Nature 350:468
18. Gardner M (1983) Wheels, Life, and Other Mathematical Amusements. W.H. Freeman, New York
19. Wikipedia (2007) Garden of Eden pattern. http://en.wikipedia.org/wiki/Garden_of_Eden_pattern. Accessed 6 Apr 2007
20. Press W, Teukolsky S, Vetterling W, Flannery B (1988) Numerical Recipes in C. Cambridge University Press, Cambridge
21. Clar S, Drossel B, Schwabl F (1996) Phys J Cond Mat 8:6803
22. Dhar D (1990) Phys Rev Lett 64:1613
23. Dhar D, Ramaswamy R (1989) Phys Rev Lett 63:1659
24. Dhar D, Majumdar SN (1990) Phys J A 23:4333
25. Majumdar SN, Dhar D (1992) Physica A 185:129
26. Creutz M (1991) Comp Phys 5:198
27. Anderson R et al (1989) Amer Math Monthly 96:981; Björner A, Lovász L, Shor P (1991) Europ J Combinatorics 12:283; Eriksson K (1996) SIAM J Discret Math 9:118
28. Goles E, Margenstern M (1996) Int J Mod Phys C 7:113

Books and Reviews

Bak P (1996) How Nature Works: The Science of Self-Organised Criticality. Springer, Berlin

Gore A (1992) Earth in the Balance: Ecology and the Human Spirit. Plume, Boston

Jensen HJ (1998) Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological Systems. Cambridge, Cambridge

Toffoli T, Margolus N (1987) Cellular Automata Machines: A New Environment for Modeling. MIT Press, Cambridge

Wolfram S (1994) Cellular Automata and Complexity: Collected Papers. Westview Press, Boulder

Self-organizing Systems

WOLFGANG BANZHAF
Department of Computer Science, Memorial University of Newfoundland, St. John's, Canada

Article Outline

Glossary
Definition of the Subject
Introduction
History of the Concept of Self-Organization
Examples of Natural Self-Organizing Systems
Examples of Artificial Self-Organizing Systems
Explanatory Concepts of Self-Organization
Modeling Methods
The Role of Self-Organization in Science, the Social Sciences and Engineering
Open Issues and Future Directions
Cross References
Bibliography

Glossary

Attractor A special set of system states approached by a dynamical system after some time has passed when starting from a variety of initial states.

Autopoiesis The process by which systems maintain their identity and organization and regenerate their components in the course of their operation.

Competition and cooperation Types of interaction between two or more elements of a system. Competition refers to each element striving to maximize its use of a finite and/or non-renewable resource. Cooperation refers to the elements engaging in a mutually beneficial exchange.

Complexity Measure of number of elements and way of their interaction (structural c.); measure of variety of behavioral repertoire of a system (functional c.).

Constructive system A system whose later components are generated during the interaction of its earlier components.

Dynamics The quantitative development of a system's state variables over time.

Emergence The appearance of qualitatively new phenomena on higher levels of a hierarchical system.

Evolution A process of structural or qualitative change in some direction.

Instability Inability of a system to keep its state or structure.

Mode Macroscopic behavior of a system caused by the interaction of its microscopic parts via long-range correlations.

Non-equilibrium System state with inflow of matter, energy and/or information causing it to stay away from its most probable state under the hypothetical condition of isolation.

Phase transition A point at which the appearance or behavior, or qualitative nature of the steady state of a system changes suddenly.

Resilience Measure of a system's ability to remain within a domain of stability in response to fluctuations of the system by a perturbation, and the ability of the system to return to that stable domain having once left.

Self-organized criticality The ability of a system to evolve in such a way as to approach a critical point and then maintain itself at that point.

Definition of the Subject

Self-organization is a core concept of Systems Science. It refers to the ability of a class of systems (self-organizing systems (SOS)) to change their internal structure and/or their function in response to external circumstances. Elements of self-organizing systems are able to manipulate or organize other elements of the same system in a way that stabilizes either structure or function of the whole against external fluctuations. The process of self-organization is often achieved by growing the internal space-time complexity of a system and results in layered or hierarchical structures or behaviors. This process is understood not to be instructed from outside the system and is therefore called *self-organized*.

Modern ideas about self-organization start with the foundation of cybernetics in the 1940s. W. Ross Ashby, H. von Foerster and N. Wiener, among others, have contributed to an early understanding. Later, the concept was adopted in physics and nowadays pervades most of natural sciences. Many systems have been identified as possessing aspects of self-organization, though a clear definition is still lacking. As a result of this inaccuracy, the theory of self-organization is still in its infancy. While the concept has found applications in the social sciences and engineering as well, SOSs are an area of active research, with fundamental questions still being explored.

Introduction

Over the last decades a variety of features have been identified as typical for self-organizing systems. Not all of these features are present in all systems able to self-organize. Self-organizing systems are *dynamic*, often *non-determin-*

istic, *open*, exist *far from equilibrium* and sometimes employ *autocatalytic amplification of fluctuations*. Often, they are characterized by *multiple time-scales* of their internal and/or external interactions, they possess a *hierarchy of structural and/or functional levels* and they are able to *react* to external input *in a variety of ways*. Many self-organizing systems are *non-teleological*, i. e. they do not have a specific purpose except their own existence. As a consequence, *self-maintenance* is an important function of many self-organizing systems. Most of these systems are *complex* and use *redundancy* to achieve *resilience* against external pertubation tendencies.

Key aspects of self-organizing systems are:

- Growth of Complexity
- Emergence of new phenomena
- Positive and negative feedback loops of internal regulation.

The process of self-organization has been invoked to explain numerous phenomena in the natural sciences. From non-living systems like galaxies and stars down to nanoparticle aggregates, self-organizing systems have been observed. In the living world cells, organisms and ecosystems provide examples of systems classified as self-organizing. The concept has found applications in man-made systems like communication networks, societies, economies, and has been identified to be at work in the world of ideas in the development of world views, scientific beliefs and norm systems.

History of the Concept of Self-Organization

Early History

The concept of self-organization can be traced back to at least two sources: Western philosophy influenced heavily by Greek thinking; and eastern philosophy, centered around the process thinking of Bhuddism. The ideas derived from both sources resound with the modern way of thinking about self-organization although the word itself had never been used.

On wondering about the origin of the world, Greek atomists from Democritos of Abdera to Epicuros of Samos argued that world order arose from chance collisions of particles. First, the cosmos (from Greek *kosmos* = *the ordered*) did not exist but chaos instead (from Greek *chaos* = *the disordered*). In modern times chaos theory has taken up this topic again, with deep connections to ideas about self-organization and the origin of order in the universe.

In the Christian tradition, St Thomas Aquinas contributed through his interest in logical proofs for the existence of God. One of these proofs considered God to be the ultimate organizer or designer. The argument was that everything had to be organized and this called for an organizer. In turn, the organizer had to be organized and so on back to the original organizer: this was God. Since God is present without cause (otherwise he would have to be organized by another entity), he must have somehow organized himself.

The Bhuddist way of thinking, on the other hand, was fundamentally process-oriented. Things are considered not to be in static existence, but rather are thought to be generated and maintained by proper processes. The emphasis on processes is reminescent of self-organizing systems whose structure is determined by proper processes of internal and external interactions.

The First Use of the Term

Work on General Systems Theory (von Bertalanffy) [1] and Cybernetics (Wiener) [2] paved the way for the idea of self-organization.

The concept of a self-organizing system was introduced by Ashby in 1947 [3]. In the 1950s a self-organizing system was considered to be a system which changes its basic structure as a function of its experience and environment. The term was used by Farley and Clark in 1954 to describe learning and adaptation mechanisms [4]. Ashby [5], in 1960, redefined a self-organizing system to include the environment with the system proper. Von Foerster argued [6], also in 1960, that a self-organizing dynamical system possesses some stable structures (eigenvalues, attractor states) which he later termed eigenbehavior.

Further Developments

This notion was further developed by Haken [7] in 1977 who termed the global cooperation of elements of a dynamical system – resulting in it assuming an attractor state – *self-organization*. Both Haken and Kauffman (1993) [8] argued for a deep connection between self-organization and *selection*. Haken found that modes of collective behavior are competing against each other and considered this process to be Darwinian selection in the non-living world. Kauffman, on the other hand, emphasized the role of constraints on the direction of evolution (mostly of the living), caused by self-organization.

Already in the 1970s, however, ideas branched out into different directions. One branch of the development of the idea deepened the relation to studies of learning and adaptation (Conrad, Kohonen, [9,10]), another branch studied processes of self-organization in systems far from equilibrium (Prigogine, Haken) [11,12]. Chaos

Self-organizing Systems, Table 1
Examples of Self-Organizing Systems

System	Flow	Self -organizing entities	Emergence
Atmosphere	Solar energy	Gas molecules	Patterns of atmospheric circulation
Climate	Energy	Weather conditions (humidity, precipitation, temperature, …)	Distribution patterns
Liquid between plates	Heat	Particle circulation	Movement patterns
Laser	Excitation energy	Phase of light waves	Phase-locked mode
Reaction vessel	Chemicals for BZ reaction	Chemical reactions	Patterns of reaction fronts
Neural networks	Information	Synapses	Connectivity patterns
Living cells	Nutrients	Metabolic reactions	Metabolic pathways/network patterns
Food webs	Organisms of different species	Species rank relation	Networks of species
Highway traffic	Vehicles	Distance of vehicles	Density waves of traffic
City	Goods, information	Human housing density	Settling patterns
Internet	Computer nodes	Connections between nodes	Network connection pattern
Web	Information posted in websites	Links between websites	Patterns of web communities

theory (Thom, Mandelbrot) [13,14] was the line of inquiry into nonlinear systems in mathematics, whereas autopoiesis and self-maintenance where at center stage in biology (Eigen, Rosen) [15,16] neurophysiology (von der Mahlsburg, Linsker [17,18]) and cognitive science (von Foerster, Maturana and Varela) [19,20].

In recent years, self-organizing systems have assumed center stage in the natural sciences [21,22], and the social sciences [23,24,25]. Engineering is beginning to see the usability of the concept [26] in connection with the approach of nano-scale applications and the growing complexity of human artefacts.

Examples of Natural Self-Organizing Systems

Classical examples of natural self-organizing systems are the formation of Benard convection cells in non-equilibrium thermodynamics, the generation of laser light in non-linear optics and the Belousov–Zhabotinsky reaction in chemistry. These are examples from the non-living world, and the complexity of resulting macroscopic space-time patterns is restricted.

Nearly unrestricted complexity through self-organization can be achieved in the living world. For instance, the interaction of species in foodwebs could be looked at from this point of view [22]. Here, we shall briefly look at the self-organization of the Earth's biosphere known as the Gaia hypothesis [27]. This hypothesis states that Earth's living and non-living components self-organize into a single entity called *Gaia*. Gaia can be understood as the whole of the biosphere, that is able to self-stabilize. The model states, in other words, that the biomass of Earth self-regulates to make conditions on the planet habitable for life. In this way, a sort of homeostasis would be sought by the self-organizing geophysical/physiological system of Earth.

In recent years, the Gaia hypothesis has found its place in Earth Systems Science as the realization that there is just one global ecosystem, containing the entirety of resources and all living organisms, all interacting with each other in multiple regulatory cycles. These ideas have been connected to the Darwinian theory of evolution via natural selection [28,29], providing a mechanism by which such a stable state can be assumed to have emerged.

Other examples of natural self-organizing systems can be found in Table 1.

Examples of Artificial Self-Organizing Systems

There are numerous examples of man-made systems or systems which involve man that exhibit self-organization phenomena. Among them are traffic patterns, self-organizing neural networks, celular phone networks or the development of web communities.

The example we shall briefly discuss is that of traffic flow patterns. Macroscopic patterns of traffic jams on highways have been observed and experimentally examined [30]. Their appearance is closely related to traffic density, the model of behavior for drivers and the traffic flow that this allows [31]. Traffic flow is an open system, and it develops waves of traffic jams (solitons) excited by the density of traffic. Transitions between different traffic flow patterns have been considered as phase transitions, typical products of self-organization in the non-living world.

A number of examples of self-organizing systems from different fields is given in Table 1, lower section.

Explanatory Concepts of Self-Organization

Despite half a century of inquiry, the theory of self-organizing systems is still in its infancy. There is no "standard model" of SOS, only various aspects emphasized by different researchers. Here we shall discuss the most important of these.

Non-Equilibrium Thermodynamics

Thermodynamics has been concerned with the notion of order and disorder in physical systems for more than a century. The theory of self-organization has to address fundamental issues of this field. The most important question in this regard is, how order can arise through self-organization.

Classical thermodynamics has focused on closed systems, i. e. systems isolated from external influence in the form of matter and energy flow. This allowed to understand the processes involved when a system evolves undisturbed. A key result of this inquiry is the second law of thermodynamics, originally formulated by Carnot and later refined by Clausius in the 19th century. It states that "any physical or chemical process under way in a system will always degrade the energy". Clausius introduced a quantitative measure of this irreversibility by defining entropy:

$$S \equiv \int \mathrm{d}Q/T \tag{1}$$

with Q the heat energy at a given temperature T. In any process of a closed system, entropy always rises

$$\frac{\mathrm{d}S}{\mathrm{d}t} \geq 0\,. \tag{2}$$

According to Eddington, 1928 [32] this universal increase in entropy "draws the arrow of time" in nature.

Boltzmann had reformulated entropy earlier in terms of the energy microstates of matter. In his notion, entropy is a measure of the number of different combinations of microstates in order to form a specific macrostate.

$$S = k_\mathrm{B}\,\ln(W) \tag{3}$$

with k_B Boltzmann's constant and W the thermodynamic probability of a macrostate. He argued that the macrostate with most microstates (with maximum entropy) would be most probable and would therefore develop in a closed system. This is the central tenet of equilibrium thermodynamics.

More interesting phenomena occur if the restrictions for isolation of a system are removed. Nicolis and Prigogine [11] have examined these systems of non-equilibrium thermodynamics which allow energy and matter to flow across their boundary. Under those conditions, total entropy can be split into two terms, one characterizing internal processes of the system, $\mathrm{d_i}S$ and one characterizing entropy flux across the border $\mathrm{d_e}S$. In a generalization of the second law of thermodynamics, Prigogine and Nicolis postulated the validity of the second law for the internal processes,

$$\frac{\mathrm{d_i}S}{\mathrm{d}t} \geq 0 \tag{4}$$

but explicitly emphasized that nothing can be said about the sign of the entropy flux. Specifically, it could carry a negative sign and it could be larger in size than the internal entropy production. Since the total entropy is the sum of both parts, the sign of the total entropy change of an open system could be negative,

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{\mathrm{d_i}S}{\mathrm{d}t} + \frac{\mathrm{d_e}S}{\mathrm{d}t} < 0 \tag{5}$$

a situation impossible in equilibrium thermodynamics. Thus, increasing order of the system considered would be possible through export of entropy. Self-organization of a system, i. e. the increase of order, would not contradict the second law of thermodynamics. Specifically, the non-equilibrium status of the system could be considered a source of order.

Even in the distance from thermodynamic equilibrium, however, certain stable states will occur, the *stationary* states. These states assume the form of *dissipative structures* if the system is far enough from thermodynamic equilibrium and dominated by non-linear interactions. The preconditions for dissipative structures can be formulated as follows:

1. The system is open.
2. The inner dynamics is mainly non-linear.
3. There are cooperative microscopic processes.
4. A sufficient distance from equilibrium is assumed, e. g. through flows exceeding critical parameter values.
5. Appropriate fluctuations appear.

If those conditions are fullfilled, the classical thermodynamic branch of stationary solutions becomes unstable and dissipative structures become stable system solutions.

Synergetics

Prigogine's description of dissipative structures is formally limited to the neighborhood of equilibrium states. As

Haken pointed out, this is a severe restriction on its application and in particular precludes its formal application to living systems. Instead, Haken proposed *order parameters* and the *slaving principle* as key concepts for systems far from equilibrium. Let the time evolution of a continuous dynamical system is described by

$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} = \mathbf{N}(\mathbf{q}, \alpha) + \mathbf{F}(t) \tag{6}$$

where $\mathbf{q}(t) = [q_1(t), q_2(t), \ldots, q_N(t)]$ is the system's state vector and $\mathbf{N}$ is the deterministic part of the system's interaction whereas $\mathbf{F}$ represent fluctuating forces, and α are the so-called control parameters. Then the stable and unstable parts of the solution can be separated by linear stability analysis, as can the time dependent and time independent parts. As a result, the solution can be written as

$$\mathbf{q}(t) = \mathbf{q}_0 + \sum_u \xi_u(t)\mathbf{v}_u + \sum_s \xi_s(t)\mathbf{v}_s \tag{7}$$

$\mathbf{v}_u, \mathbf{v}_s$ are the unstable and stable modes, respectively, and $\xi_u(t), \xi_s(t)$ are their amplitudes. These amplitudes obey the following equations

$$\frac{\mathrm{d}\xi_u}{\mathrm{d}t} = \lambda_u \xi_u + N_u(\xi_u, \xi_s) + \mathbf{F}_u(t) \tag{8}$$

$$\frac{\mathrm{d}\xi_s}{\mathrm{d}t} = \lambda_s \xi_s + N_s(\xi_u, \xi_s) + \mathbf{F}_s(t) \tag{9}$$

with λ_u, λ_s characterizing the linear part of the equations and function N summarizing the non-linear deterministic components. The slaving principle formulated by Haken now allows to eliminate the stable mode development by expressing them as a function of unstable modes

$$\xi_s(t) = f_s[\xi_u(t), t]\,. \tag{10}$$

Thus, the unstable modes (order parameters) enslave the stable modes and determine the development of the system's dynamics. This result is useful both to describe *phase transitions* and *pattern formation* in systems far from equilibrium.

Synergetic concepts have been applied in a variety of disciplines [33].

Chaos and Complexity

The treatment of *chaotic systems* has been derived from non-linear system theory. Chaotic systems are usually low-dimensional systems which are unpredictable, despite being deterministic. The phenomenon was originally discovered by the meteorologist E. Lorenz in 1963 [34], although H. Poincare in 1909 was aware of the possibility of certain systems to be sensitive to initial conditions [35]. The reason for the difficulty to predict their behavior stems from the fact that initially infinitesimal differences in trajectories can be amplified by non-linear interactions in the system. These instabilities, together with the lack of methods for solving even one-dimensional non-linear equations analytically, produce the difficulties for predictions. Modern theory of deterministic chaos came into being with the publication of a seminal article by May in 1976 [36].

Complex systems, on the other hand, have many degrees of freedom, mostly interacting in complicated ways, i. e. they are high-dimensional. All the more astonishing is the fact that our world is not totally chaotic in the sense that nothing can be predicted with any degree of certainty. It became apparent, that chaotic behavior is but one of the ways non-linear dynamical systems behave, with other modes being complex attractors of a different kind.

Complexity itself can be measured, notably there exist a number of complexity measures in computer science, but describing or measuring complexity is not enough to understand complex systems.

Self-Organized Criticality

For particular high-dimensional systems, Bak et al. [43] have suggested a dynamic system approach toward the formation of fractal structures, which are found to be widespread both in natural and artificial environments. Their canonical example was a pile of sand. They examined the size and frequency of avalanches under certain well-prepared conditions, notably that grains of sand would fall on the pile one by one. This is an open system with the forces of gravity and friction acting on the possibly small fluctuations that are caused by deviations in the hitting position of each grain of sand. They observed how the grains would increase the slope of the sand pile until more or less catastrophic avalanches developed.

Bak suggested the notion of *self-organized criticality* (SOC) as a key concept which states that large dissipative systems drive themselves to a critical state with a wide range of length and time scales. This idea provided a unifying framework for the large-scale behavior in systems with many degrees of freedom. It has been applied to a diverse set of phenomena, e. g. in economic dynamics and biological evolution. SOC serves as an explanation for many power-law distributions observed in natural, social and technical systems, like earthquakes, forest fires, evolutionary extinction events, and wars. As opposed to the widely studied low-dimensional chaotic systems, SOC systems

have a large number of degrees of freedom, and still exhibit fractal structures as are found in the extended space-time systems in nature.

The Hypercycle

In a series of contributions since 1971, Eigen and Schuster have discussed particular chemical reaction systems responsible for the origin, self-organization and evolution of life [37,38,39,40]. By considering *autocatalytic sets* of reactions they arrived at the most simple form of organization, the *hypercycle*, which is able to explain certain aspects of the origin of life. They have considered a chemical reaction system composed of a variety of self-reproductive macromolecules and energy-rich monomers required to synthesize those macromolecules. The system is open and maintained in a non-equilibrium state by a continuous flux of energy-rich momomers. Under further assumptions they succeeded in deriving Darwinian selection processes at the molecular level. Eigen and Schuster have proposed rate equations to describe the system.

The simplest system realizing the above mentioned conditions can be described by the following rate equations

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = (A_i Q_i - D_i)x_i + \sum_{k\neq i} w_{ik}x_k + \Phi_i(\mathbf{x}) \qquad (11)$$

where i enumerates the individual self-reproducing units and x_i measures their respective concentrations. Metabolism is quantified by the formation and decomposition terms $A_i Q_i x_i$ and $D_i x_i$. The ability of the self-reproducing entities to mutate into each other is summarized by the quality factor for reproduction, Q_i, and the term $w_{ik}x_k$ which takes into account all catalytic productions of one sort using the other. A_i, D_i are rate constants for selfreproduction and decay respectively. The flow term Φ_i finally balances the production/destruction in this open system in order to achieve $\sum_k x_k = \text{const}$.

By introducing a new feature called excess production

$$E_i \equiv A_i - D_i \qquad (12)$$

and its weighted average

$$\bar{E}(t) = \sum_k E_k x_k / \sum_k x_k \qquad (13)$$

and symbolizing the "intrinsic selection value" of a sort i by

$$W_{ii} = A_i Q_i - D_i \qquad (14)$$

one arrives at reduced rate equations

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = (W_{ii} - \bar{E})x_i + \sum_{k\neq i} w_{ik}x_k \,. \qquad (15)$$

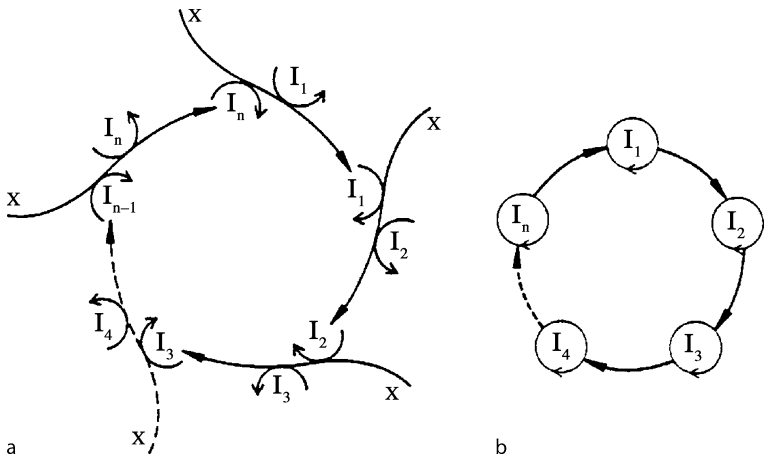
These equations can be solved under certain simplifying assumptions and notably yield the concept of a *quasi-species* and the following extremum principle: A quasi-species y_i is a transformed self-replicating entity with the feature that it can be considered as a cloud of sorts x_i whose average or consensus sequence it is. The extremum principle reads: Darwinian selection in the system of quasi-species will favor that quasi-species which possesses the largest eigen-value of the rate equation system above.

The Origin of Order

In the 1990s Kauffman [41] pointed out one of the weaknesses of Darwinian theory of evolution by natural selection: It cannot explain the '*origin* of species' but rather only their subsequent development. Kauffman instead emphasized the tendency of nature to constrain developments along certain paths, due to restrictions in the type of interaction and the constraints of limited resources available to evolution. In particular he held up the view that processes of spontaneous order formation conspire with the Darwinian selection process to create the diversity and richness of life on Earth.

Previously, Kauffman had formulated and extensively studied [42] the NK fitness landscapes formed by random networks of N Boolean logic elements with K inputs each. Kauffman observed the existence of cyclic attractor states whose emergence depended on the relation between N and K, and the absolute value of K. In the case of large K ($K \approx N$), the landscape is very rugged and behavior of the network appears stochastic. The state sequence is sensitive to minimal disturbances and to slight changes of the network. The attractor length is very large, $\approx N/2$, and there are many attractors. In the other extremal case, $K = 2$, the network is not very sensitive to disturbances. Changes of the network do not have strong and important consequences for the behavior of the system.

Kauffman proposed NK networks as a model of regulatory systems of living cells. He further developed the notion of a *canalizing function* that is a Boolean function in which at least one variable in at least one state can completely determine the output of the function. He proposed that canalizing functions are an essential part of regulatory genetic networks.



Self-organizing Systems, Figure 1
The Hypercycle (reproduced from [15])

Emergence and Top-Down Causation

The notion of *emergence* has been introduced in complex systems theory in order to explain the appearance of new qualitative features on the level of an entire system that could not be observed at the level of its components. Emergent behavior can be connected to the afore-mentioned complex attractors. It requires switching the level of description of behavior of a system, from local (component-centered) to global (system-centered), or at least to a meso-level (sub-system-centered). Emergent behavior happens when

a) the system shows qualitatively new behavior on a higher level of description which
b) could not have been easily predicted from the interactions of components at the lower level (obeys a non-linear relationship)
c) is the result of a self-organization process.

Emergence is strongly related to self-organization. It is often understood as a pattern formation process. While it essentially has to do with changing the perspective and looking at the system at a different level, it concerns itself with a change in behavior (e. g. the system is getting more organized, shows new coordinated modes of behavior). It has been further conjectured that there is top-down causation, i. e. the structures forming on the higher level of the system are able to affect the lower levels (system components) and influence them in a way that stabilizes the newly emergent behavior. Haken could show in the context of Synergetics that this phenomenon exists. *Top-down causation* is believed to be an important source of complexity, especially in living systems, because it stabilizes patterns.

Self-organization draws heavily from this source of qualitative innovation in complex systems.

Modeling Methods

A formal model is a simplified mathematical or algorithmic representation of a system. Often it has been simplified to the point of a carricature, and this has to be born in mind when making conclusions about the consequences of model predictions. No model can predict beyond the limits of its approximations.

Mean-Field Methods

One of the most important methods used to model complex systems is tied to the notion of dynamical systems. Dynamical systems are systems whose time development is accessible to a description by state changes. It entails the existence of a state space in which these changes can be traced and quantified.

Mean-field methods of description focus on average behavior. They abstract away from the local correlations between a system's elements and describe only long-range changes. For instance, the behavior of a planet could be described as a point on its trajectory around the star it circles. Detailed interactions of its atmosphere would not be part of that description.

Mean-field methods are formulated in the form of time-dependent differential or difference equations which can be solved under certain conditions and predict the behavior of a system in its state space.

Assuming that the state of a system can be subsumed in a vector of state variables $\mathbf{x}$ whose values are observable

and depend on time, we can generally formulate an equation for continuous time development as:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}) \,. \tag{16}$$

If time does not develop continuously, a discrete (iterative) equation can be used to describe system behavior:

$$\mathbf{x}_{t+1} = \mathbf{f}(\mathbf{x}_t) \,. \tag{17}$$

The notion is that the state of the system at time t, $\mathbf{x}(t)$ or $\mathbf{x}_t$, is everything that is of interest and can be known about the system. It turned out that with this extreme simplification many systems became treatable that would have otherwise resisted quantitative treatment. The non-linear nature of many of the state development equations, however, and the high-dimensionality of state space vectors often constitute prohibitive hurdles to exact or even approximate mathematical solution of these equations.

As a result, algorithmic approaches for modeling self-organizing systems have become more prevalent in recent years.

Agent-Based Models

A very general class of algorithmic systems is subsumed under the term agent-based models. In these systems, individual entities are modeled that interact with each other. Thus, the approximation of average behavior, and the interest for long-term behavior only is abandoned in favor of a microscopic description of the elements of a system and their interactions. The abstraction of features of a system is achieved through the assumption of rules of behavior of the agents, including their interaction behavior. Agent-based systems must be implemented as computational systems, and run on a computer to obtain results. Agents are assigned states, and transition rules between states, depending on interacting agents, and then these rules are executed in parallel over the set of agents under consideration.

Cellular Automata A particular subset of agent-based models is the class of cellular automata introduced by von Neumann [47] going back to lattice networks of Ulam. The agents of cellular automata are placed on a grid of cells and allowed to assume a finite number of states. Interactions are determined by state transition rules and the definition of a neighborhood, which determines the interactivity of the cellular automaton. Many variants of cellular automata exist, differing in the number of dimensions of the grid, the number of states, the sort and distribution of transition rules and the nature of the neighborhood.

A typical cellular automata model might, e. g. consist of digital cells (allowing only two states, "ON" and "OFF"), homogeneous and deterministic transition rules between states, a one-dimensional grid, and nearest-neighbor interactions. Cellular automata of this type have been thoroughly examined in [44,45] and show a surprising variety and richness in behavior.

In a cellular automaton like LIFE, for instance, one can observe how macroscopic and mesocopic structures appear through self-organization, that is, as a process determined solely by the local interaction of the CA's elements. some structures, e. g. spiral waves, are more resilient against perturbation than others, e. g. glider canons. A moving structure like a glider can be interpreted as an emergent phenomenon as it does not seem to be present on the microscopic scale (single CA cells do not move).

Graphs and Networks A more general class of automata can be formulated if the notion of cellular neighborhood is abandoned. Instead of a rigidly defined grid, a graph or network of automata connected through edges to other automata is introduced. Each node of the graph/network represents an automaton, with interactions allowed via edges.

The notion of a graph is, however, more general, and allows other agent-based systems to be simulated. For instance, the nodes of a graph might represent species of an ecosystem interacting with other species (connected by edges). Each species might be represented by a state counting the number of individuals of that species. Nodes might further hold information on particular features of individuals, and possibly their variants. Explicit simulations of such systems have been considered in the context of "Artificial Chemistries" [46].

In recent years, the structure and dynamics of networks has been a major focus of interest in the scientific community. Network science has become a converging point for different disciplines interested in modelling complex behavior.

Observables

Self-organizing phenomena rest on the appearance of particular sets of behaviors. If ever they are to be understood, a clear notion of observable quantities needs to be developed that allows a proper description of the behavior of such systems. At present, no such canonical set of observables exists, owing to the bewildering variety of systems that show signs of self-organization. However, one can discern a number of different measures and observables that might form the core of such a set [49].

Entropic and Information Theoretic Measures One class of observables can be considered entropic and information-theoretic measures. These measures have in common a statistical root, and seek to describe a self-organizing system in terms of the order (or disorder) that develops over time [48].

Stability Measures Another class of observables can be discussed as stability measures. In this class, systems are sought to be disturbed from their regular behavior in order to obtain a clearer idea of their resilience.

Scaling Measures A further class of observables can be attached to features of scaling. Both theoretical and experimental approaches can be used to vary the number of dimensions, number of equations/agents, number and complexity of interactions, etc, in these systems. Scaling behavior can then be observed for particular quantities and systems classified accordingly.

Patterns and Flows The defining observables of a self-organizing system are patterns. These refer to the collective behavior of the elements of a system, differentiating them from noise. If individual entities would not show such correlations in their behavior, self-organization could not be observed. Patterns can be described in a variety of ways, e. g. as multidimensional vectors, using spatial and temporal coordinates. If patterns change dynamically one can speak of flows.

The central tenet of self-organization is that systems exist whose pattern forming tendencies are determined by themselves, and not by an outside agency.

The Role of Self-Organization in Science, the Social Sciences and Engineering

Self-organization as a concept has assumed center stage in Science. With the advent of nonlinear systems and studies on complex systems in non-equilibrium situations, the explanatory power of self-organization now permeates every branch of scientific enquiry.

From structure formation at the level of super-galactic clusters, even starting from the development of the entire universe, down to microscopic particles and their interaction patterns, self-organizing phenomena have been postulated, theorized, observed and confirmed.

In particular the origin and evolution of life have been studied under the aspect of self-organization. Within Biology, the developmental process of organisms as well as their metabolisms, growth and learning have been identified as self-organizing processes.

In the humanities, the idea of self-organization has taken roots, although the paradigm is far from being fully recognized yet. Since the 1990s the origin and development of languages has been an object of study under the premise of self-organization. In social science the concept of self-organization has been studied since a number of years, due to the obvious fact that interaction between social actors generate a society. Even in psychology, self-organizing principles begin to appear.

Economy and Management Science have taken notice of the concept, and a growing number of enterprise concepts promote the idea of a form as a self-organizing entity.

Finally, Philosophy has embraced the concept of self-organization and connected it to earlier thoughts on the development of the scientific process and epistemology. Whitehead put forward his process philosophy, and Smuts, already in the 1920s, promoted the notion of holism which has strong connections to self-organization. Evolutionary epistemology was formulated as a response to traditional epistemology and emphasizes the aspect of natural selection affecting senses and cognitive abilities.

Engineering is beginning to grasp the ubiquity of self-organization in Nature. Specifically in the area of nanotechnology the concept is used extensively for the purpose of self-assembly of molecular entities. At nanoscales, it is very difficult to directly specify the structuring behavior of entities. As a result, self-organizing properties of matter are used to the advantage of the structural outcome.

Different kinds of infrastructure networks have been recognized as self-organizing, and Engineering begins to make use of the tendency of networked systems to self-organize.

In the area of adaptation, there exists a long tradition of making use of self-organization principles. The self-organizing feature map, introduced by Kohohen, has been a key step forward in the domain of unsupervised learning of artificial neural networks.

Open Issues and Future Directions

So far, there is no unique theory of self-organization. Over the course of many years different approaches have been used, but a coherent picture has not yet emerged.

An important open question in the area of the mathematical basis for self-organization is the formulation of a theory of *constructive (evolutionary) systems*, that is systems which, in the course of their development, generate new elements that subsequently interact with elements already created earlier.

Another question aims at the raison d'etre of *hierarchical systems*. Why do they form, how do they structure

themselves, and what would be possible to apply from these principles in Engineering? Notably, how would one build self-organizing systems such that they do something useful? How could they be controlled?

In Science, the build-up of complexity remains a controversial issue. Is it true that evolution of the universe tends to increase complexity, or is there no tendency of complexity increase at all? What are the mechanisms by which Nature increases complexity, if any? How could we apply this knowledge in planning and managing complexity in the human world?

A wealth of questions remains, and it is anticipated that the 21st century will shed light on at least a few of them.

Cross References

- ► Entropy in Ergodic Theory
- ► Mathematical Basis of Cellular Automata, Introduction to
- ► Swarm Intelligence
- ► Biological Development and Evolution, Complexity and Self-organization in
- ► Rough and Rough-Fuzzy Sets in Design of Information Systems

Bibliography

Primary Literature

1. von Bertalannfy L (1968) General System Theory: Essays on its Foundation and Development. George Braziller, New York
2. Wiener N (1965) Cybernetics: Or Control and Communication in the Animal and the Machine. MIT Press, Cambridge
3. Ashby WR (1947) Principles of the Self-Organizing Dynamic System. J Gen Psychol 37:125–128
4. Farley BG, Clark WA (1954) Simulation of self-organizing systems by digital computer. IRE Trans Inf Theor 4:76–84
5. Ashby WR (1960) Design for a Brain: The Origin of Adaptive Behavior. Chapman and Hall, London
6. von Foerster H (1960) On Self-organizing systems and their environments, In: Yovits MC, Cameron S (eds) Self-Organizing Systems. Pergamon Press, New York, pp 31–50
7. Haken H (1983) Synergetics – An Introduction, 3rd edn. Springer, Berlin
8. Kauffman S (1993) The Origins of Order. Oxford University Press, Oxford
9. Conrad M (1983) Adaptability. Plenum Press, New York
10. Kohonen T (2000) Self-Organizing Maps. Springer Series in Information Sciences, vol 30. Springer, Berlin
11. Nicolis G, Prigogine I (1977) Self-Organization in Nonequilibrium Systems. Wiley, New York
12. Haken H (2000) Information and Self-Organization: A Macroscopic Approach to Complex Systems. Springer Series in Synergetics, 2nd edn. Springer, Berlin
13. Thom R (1989) Structural Stability and Morphogenesis: An Outline of a General Theory of Models. Addison-Wesley, Reading
14. Mandelbrot B (1982) The fractal Geometry of Nature. W.H. Freeman, San Francisco
15. Eigen M, Schuster P (1979) The Hypercycle. Springer, Berlin
16. Rosen R (1991) Life Itself: A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life. Columbia University Press, New York
17. von der Malsburg C (1973) Self-organization of orientation sensitive cells in the striate cortex. Kybernetik 14:85–100
18. Linsker R (1988) Self-organization in a perceptual network. Computer 21:105–117
19. von Foerster H (2002) Understanding Understanding. Springer, New York
20. Maturana H, Varela F (1979) Autopoiesis and Cognition. Reidel, Dordrecht
21. Camazine S, Deneubourg J, Franks N, Sneyd J, Bonabeau E, Theraulaz G (2000) Self-Organization in Biological Systems. Princeton University Press, Princeton
22. Sole R, Bascompte J (2006) Self-Organization in Complex Ecosystems. Princeton University Press, Princeton
23. Luhmann N (1995) Social Systems. Stanford University Press, Palo Alto
24. Focardi S, Cincotti S, Marchesi M (2002) Self-organization and market crashes. J Econ Behav Organ 49:241–267
25. Portugali J (2000) Self-Organization and the City. Springer, Berlin
26. Bruckner S, di Marzo Serugendo G, Karageorgos A, Nagpal R (eds) (2005) Engineering Self-Organizing Systems. Springer, Berlin
27. Lovelock JE, Margulis L (1974) Atmospheric homoeostasis by and for the biosphere: The Gaia hypothesis. Tellus 26:2–9
28. Lenton TM (1998) Gaia and natural selection. Nature 394:439–447
29. Staley M (2002) Darwinian Selection Leads to Gaia. J Theor Biol 218:35–46
30. Kerner BS (1998) Experimental Features of Self-Organization in Traffic Flow. Phys Rev Lett 81:3797–3800
31. Treiber M, Helbing D (1999) Explanation of Observed Features of Self-Organization in Traffic Flow. Arxiv preprint cond-mat/9901239
32. Eddington AS (1928) The Nature of the Physical World. Gifford Lectures. Cambridge University Press, Cambridge
33. Haken H (2004) Synergetics. Introduction and Advanced Topics. Springer, Berlin
34. Lorenz E (1963) Deterministic Nonperiodic Flow. J Atmospheric Sci 20:130–141
35. Poincare H (1909) Science et Methode. Flammarion, Paris
36. May RM (1976) Simple Mathematical Models with very complicated Dynamics. Nature 261:459–467
37. Eigen M (1971) Selforganization of Matter and the Evolution of Biological Macromolecules. Naturwissenschaften 58:465–523
38. Eigen M, Schuster P (1977) The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle. Naturwissenschaften 64:541–565
39. Eigen M, Schuster P (1978) The hypercycle: a principle of natural self-organization, Part B. Naturwissenschaften 65:7–41
40. Eigen M, Schuster P (1978) The hypercycle: a principle of natural self-organization, Part C. Naturwissenschaften 65:341–369
41. Kauffman SA (1993) The Origins of Order. Oxford University Press, Oxford
42. Kauffman SA (1969) Metabilic Stability and epigenesis in randomly constructed nets. J Theor Biol 22:437–467

43. Bak P, Tang C, Wiesenfeld K (1988) Self-organized criticality. Phys Rev A 38:364–374
44. Wolfram S (1983) Statistical Physics of Cellular Automata. Rev Mod Phys 55:601–644
45. Wolfram S (1984) Cellular Automata as Models of Complexity. Nature 311:419–424
46. Dittrich P, Ziegler J, Banzhaf W (2001) Artificial Chemistries – A Review. Artif Life 7:225–275
47. von Neumann J (1966) Theory of Self-Reproducing Automata. Univ of Illinois Press, Chicago
48. Polani D (2003) Measuring Self-Organization via Observers, In: Banzhaf W et al (eds) Advances in Artificial Life. Lecture Notes in Artificial Intelligence, vol 2801. Springer, Berlin, pp 667–675
49. Shalizi CR, Shalizi KL, Haslinger R (2004) Quantifying Self-Organization with Optimal Predictors. Phys Rev Lett 93:118701

Books and Reviews

Bak P (1996) How Nature Works – The Science of Self-Organized Criticality. Springer, New York
Bar-Yam Y (2003) Dynamics of Complex Systems. Westview Press/Perseus Books, New York
Boccara N (2004) Modelling Complex Systems. Springer, New York
Edelman G (1987) Neural Darwinism. Basic Books, New York
Hemelrijk C (ed) (2005) Self-organization and Evolution of Social Systems. Cambridge University Press, Cambridge
Holland J (2000) Emergence: From Chaos to Order. Oxford University Press, Oxford
Jantsch E (1980) The Self Organizing Universe: Scientific and Human Implications. Pergamon Press, New York
Johnson S (2001) Emergence. Scribner, New York
Kelso S (1995) Dynamic Patterns: The Self-organization of Brain and Behavior. MIT Press, Cambridge
Lovelock JE (1995) The Ages of Gaia. W.W. Norton, New York
Mingers J (1995) Self-Producing Systems. Plenum Press, New York
Morowitz H (2002) The Emergence of Everything. Oxford University Press, Oxford
Rosen R (1999) Essays on Life Itself. Columbia University Press, New York
Sornette D (2003) Critical Phenomena in Natural Sciences. Springer, Berlin

Self-replicating Robotic Systems

JACKRIT SUTHAKORN[1], MATTHEW MOSES[2], GREGORY S. CHIRIKJIAN[2]
[1] Biomedical and Robotics Technology Lab, Faculty of Engineering, Mahidol University, Salaya, Thailand
[2] Robot and Protein Kinematics Lab, Department of Mechanical Engineering, Johns Hopkins University, Baltimore, USA

Article Outline

Glossary
Introduction
Our Previous Replicating Prototypes
Design and Descriptions of an Autonomous Self-Replicating Robot
Controls and Programming of the Autonomous Self-Replicating Robot
Experiments and Results of the Autonomous Self-Replicating Robot
Our Approach to Self-Replicating Control Circuitry
Towards a Universal Constructor
Discussion
Future Works
Acknowledgments
Bibliography

Glossary

Artificial life Artificial life refers to the study of artificial systems that exhibit lifelike properties, typically involving some form of self-replication and/or evolution. This is a very broad field of study and includes work involving computer simulations, cellular automata, chemistry, robotics, and synthetic biology.

Cellular automata A cellular automaton (CA) is a theoretical construct where a collection of cells are organized into regular grids or lattices. Many arrangements are possible, but typically one-dimensional CAs are composed of square cells arranged in a line, and two-dimensional CAs are composed of square cells arranged in a square grid. Each cell contains a finite state machine. The state of all cells in a CA are typically updated synchronously (at the same time) with each cell changing to a new state that is a function of its previous state and the previous states of its neighbors. The surrounding cells that affect a given cell's state transition are called the cell's neighborhood. CAs can simulate a wide variety of physical processes by designing an appropriate neighborhood and state transition function. CAs are most often studied by implementing them with computer simulations.

Finite state machine A finite state machine (FSM) is a conceptual computing machine with an input, output, and memory. At regular intervals in time the machine transitions (changes) to a new state, which is a function of the current state and the machine's current input. The output is a function of the input and the state. The function of the memory is to store the state information between each transition. FSMs are easily implemented with discrete digital electronic components, microcontrollers, or computer simulations. Implementations of FSMs are widely used to control industrial devices and consumer electronics.

Modular robot A modular robot is composed of distict "modules" which contain motors, mechanisms, electronics, and interconnections. Modules are typically designed so that they can easily be connected and disconnected from each other. In some modular robotic systems, many identical modules are assembled into larger robots. This allows the same set of modules to form robots optimized for different tasks. In other cases, the modules may be specialized for certain tasks, and when assembled they form a robot with additional functionality. Motivation for building modular robots includes increased versatility, ease of replacing damaged components, and potentially lower manufacturing costs.

Self-reconfigurable robot A self-replicating robotic system is a robotic device that exhibits some form of self-replication. This chapter is concerned primarily with directed robotic self-replication, in which a robotic device interprets some form of coded instructions in order to carry out the replication process. In nature deoxyribonucleic acid (DNA) typically encodes replication instructions. In the examples of robotic self-replication presented in this chapter, the instructions may be encoded in a computer program, an arrangement of modular components, or as a pattern of lines that guide the motion of a mobile robot.

Self-replication Self-replication is the process by which an entity creates a duplicate of itself. The most familiar example is the self-replication of living organisms, although other natural processes such as crystal growth can be classified as self-replication.

Universal constructor A universal constructor (UC) is a conceptual machine that reads instructions and executes them to construct an object. A key property of a UC is that it can construct any object which can be described to it via the instructions, including duplicates of itself. UCs have been demonstrated in computer simulations using cellular automata. Modern manufacturing tools, such as assembly robots and computer-controlled machine tools, are similar to UCs but practical limitations make these machines "somewhat less than universal" constructors. The ribosome, a complex molecule present in nearly all biological cells, performs a function very similar to that of a UC, assembling proteins according to instructions encoded in messenger ribonucleic acid (mRNA).

von Neumann universal constructor The mathematician John von Neumann proposed a cellular automata model of a universal constructor capable of self-replication. Many researchers have refined and improved von Neumann's original design since it was first presented in the 1950s. The original design used tens of thousands of cells in a two-dimensional CA with 29 states. The general structure of the model is a movable constructing arm controlled by instructions encoded in a long line of cells resembling a tape, conceptually similar to a computer-controlled robot arm.

Introduction

The concept of an artificial self-replicating system was introduced in the 1950s by John von Neumann [31]. Von Neumann introduced the theory of self-replicating automata and established a quantitative definition of self-replication. His early results on self-replicating machines have become useful in several diverse research areas such as: cellular automata, nanotechnology, macromolecular chemistry, and computing [8,23,26,27]. However, prior to the turn of the millennium, a fully autonomous self-replicating physical robot had never been implemented. In this chapter, a series of prototype designs from our laboratory and their physical implementation are described. We begin by discussing some motivation and history, then go on to describe a remote-controlled replicating robotic system and a semi-autonomous replicating robotic system. We then describe some fully autonomous self-replicating systems, and discuss how manufacturing work cells might be designed so as to reproduce.

Motivation

People have imagined for years a factory that could autonomously replicate itself for multiple generations, requiring neither people nor the monstrous machinery typically associated with a factory. Over recent decades, outer space has been mentioned as one potential application for such self-replicating robotic factories (see e. g., [3,10], and [7]). However, enormous technical barriers must be overcome before these systems can become feasible. The purpose of the current work is to take one small step toward realizing this goal.

In contrast to self-reconfigurable robotics [1,12,14,21,34], self-replication utilizes an original unit to actively assemble an exact copy of itself from passive components. This has the potential to result in exponential growth in the number of robots available to perform a job, thus drastically shortening the original unit's task time.

Descriptions of Self-Replication by Johns von Neumann

According to von Neumann there are four components required for a self-replicating machine: the builder, the

controller, the copier, and the blueprint [31]. The process of self-reproduction begins with the controller commanding the builder to fabricate an exact mechanical system by following the blueprint. Then the controller instructs the copier to replicate the blueprint, input the new blueprint in the replica, and start the new machine. The process described in the previous paragraph was written as a mathematical model called the "Von Neumann Kinematic Beast" by Rolf Pfeifer et al. of the University of Zurich [24], which goes as follows:

1. Let A represent the builder. If a machine G is desired, one can build this machine with the blueprint, *Blue(G)*.

   $$A + Blue(G) \rightarrow G\ .$$

   Where '+' indicates that the machine is composed of the left and right components ($A + Blue(G)$) and '$\rightarrow$' indicates construction.
2. Let B represent the copier. B would make a copy of the blue print, *Blue(G)*.

   $$B + Blue(G) \rightarrow Blue(G)\ .$$

3. Let C be the controller. With the combination of A and B, this would trigger them to follow in the correct order to generate a desired mechanical system G and the new blueprint *Blue(G)*, and then wrap them up together and split them from the original machine

   $$A + B + C + Blue(G) \rightarrow G + Blue(G)\ .$$

4. Let G be the machine $A + B + C$ then we get:

$$A+B+C+Blue(A+B+C) \rightarrow A+B+C+Blue(A+B+C).$$

Any system that observes this is a self-reproducing machine in Neumann's view. A short time after successfully presenting some of his theoretical work, von Neumann began working on the implementation of his theories. In the 1950's, with the help of Stanislaw Ulam, von Neumann invented the Cellular Automata concept (redrawn and shown in Fig. 1.)

Previous Efforts in Mechanical Self-Replicating System

Von Neumann [31] was the first to seriously study the idea of self-replicating machines from a theoretical perspective. In the late 1950's, Penrose performed the first recognized demonstration of a self-replicating mechanical system [22]. It consisted of passive elements that self-assembled under external agitation. This is similar in many ways to the modern work of Whitesides [32], only at a different length scale. Moore [19] was interested in von Neumann's concepts, but he commented that Neumann's Cellular Automata was only for demonstration, and that applications of self-replicating systems needed to be carried out. Moore described several conceptual designs of artificial living plants which could duplicate themselves not only from off-the-shelf artificial parts, but also from materials from nature. Jacobson [13] constructed a self-replicating machine using parts of toy trains. His replication was done on a round section of toy track.

In the 1980s, NASA established a series of studies on the topic of "Advanced Automation for Space Missions" [10]. These studies investigated the possibility of building a self-replicating factory on the moon. References [5,6,7,9,30] also outlined strategies for space utilization. Recently, research on robots that are capable of designing other machines with little help from humans has also been performed (see [17] and references therein). This uses rapid prototyping technologies.

Our Previous Replicating Prototypes

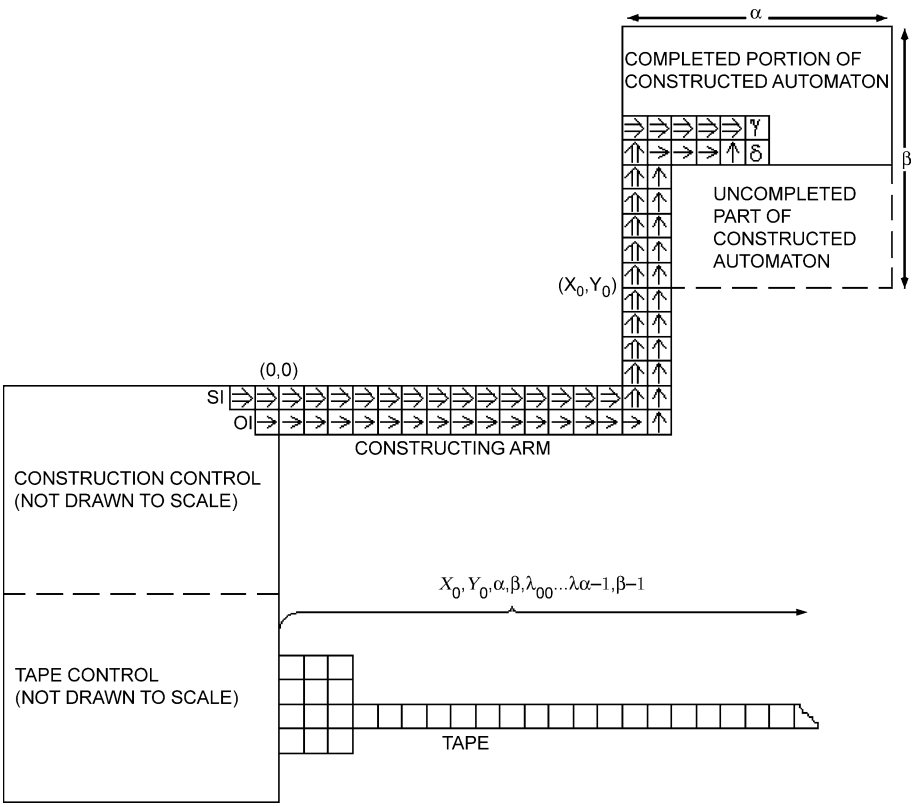
This section describes in chronological order the robotic prototypes capable of various levels of self-replication that have been designed and built in our laboratory. Our prototypes are constructed from modified LEGO Mindstorm kits with enhanced electrical connections because of their modularity, functionality and ease of use.

Demonstration I: Prototype 1

We built this robot as the first prototype to demonstrate that it is mechanically feasible for one simple robot to produce a copy of itself. This robot depends on external passive fixtures for self-replication. Figure 2 shows fixtures, the original prototype 1, and a set of the replica's subsystems. (Details of Prototype 1 can be found in [27].)

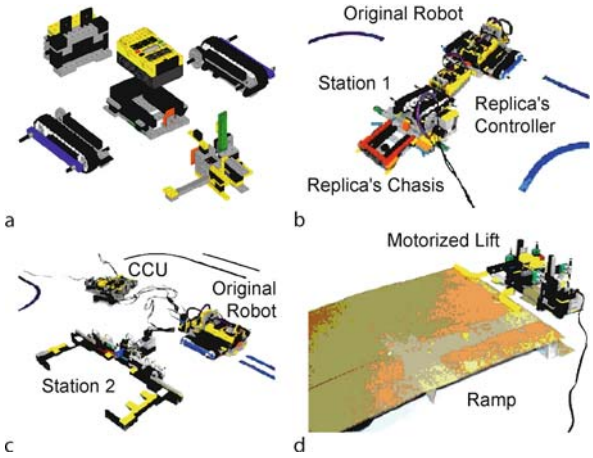
Demonstration II: Prototypes from the Spring 2002 Mechatronics Class at JHU.

From experience gained from Demonstration 1, we introduced the concept of self-replicating robots to students in a hands-on Mechatronics course (taught in the Department of Mechanical Engineering at the Johns Hopkins University) under the supervision of the last author, and TA'ed by the first author. We divided students in this course into eight groups to explore designs and implementations of the concept of self-replicating robotic systems. In order to focus on the mechanical issues involved in the design of the self-replicating systems, the robots



Self-replicating Robotic Systems, Figure 1
Schematic of von Neumann's self-reproducing cellular automaton (Redrawn from [31])

Self-replicating Robotic Systems, Figure 2
Fixtures, the original prototype 1, and a set of the replica's subsystems



Self-replicating Robotic Systems, Figure 3
a An assembly view of the semi-autonomous robotic system, b Assembling station 1, c Assembling station 2, and d Assembly station 3

were remote-controlled rather than autonomous. (Details of Demonstration II can be found in [27], and [2].)

Demonstration III: A Semi-Autonomous Replicating System

This work builds upon previous results in remote-controlled robotic replication with a new feature that many subtasks in the replication process are now autonomously performed by the robot. The use of feedback sensors was implemented. The robot is unable to directly build

its replicas. Therefore, several work-cells (intermediate robots) are required to assist the original robot in the replicating process. Each work-cell works as a station in this factory-liked replicating system. Figures 3a, b, c, and d illustrate an assembly view of the original robot, stations 1, 2, and 3, respectively. Full details of the semi-replicating robot can be found in [27,28].

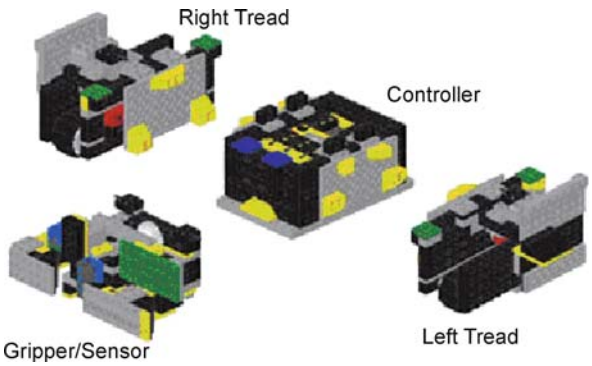
Design and Descriptions of an Autonomous Self-Replicating Robot

The robot and its replicas each consist of four subsystems: controller, left tread, right tread, and gripper/sensor subsystems. All subsystems are connected to others using magnets and shape constraints. Figure 4 shows an assembly view of the robot.

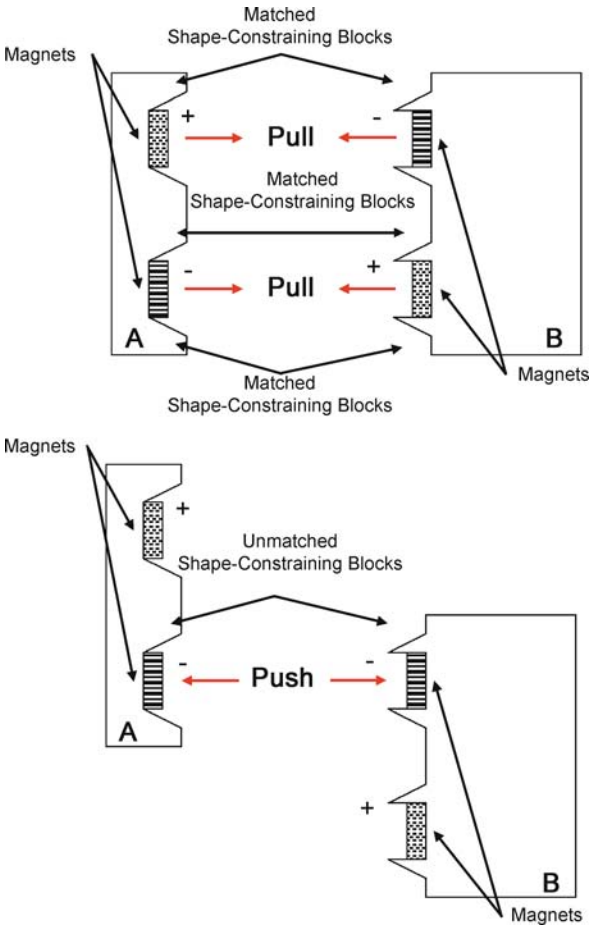
The controller subsystem is made up of a LEGO RCX programmable controller fit inside a chassis. The chassis's sides are used to connect to the left and right treads. Each side has a set of magnets, a set of shape-constraining blocks, and a set of electrical connections. The front end of the chassis is designed to attach with the gripper. The front end also has a set of magnets, a set of shape-constraining blocks, and a set of electrical connections, which transfer electrical signals and power from the controller to the gripper's motor and the navigating sensors installed on the gripper subsystem.

The magnets and the shape-constraining blocks are used in collaboration to aid aligning and interlocking subsystems. On each chassis side, the magnets are symmetrically placed in opposite polar directions to each other. This is to protect against incorrect positioning of the subsystems. The concept of using the magnets (with different polarizations) and shape-constraining blocks was influenced by the self-complementary molecules of Rebek [25]. Figure 5 illustrates the concepts of using the polar magnets and shape-constrained blocks to align and interlock subsystems. By design, it is very difficult for these connectors to misalign.

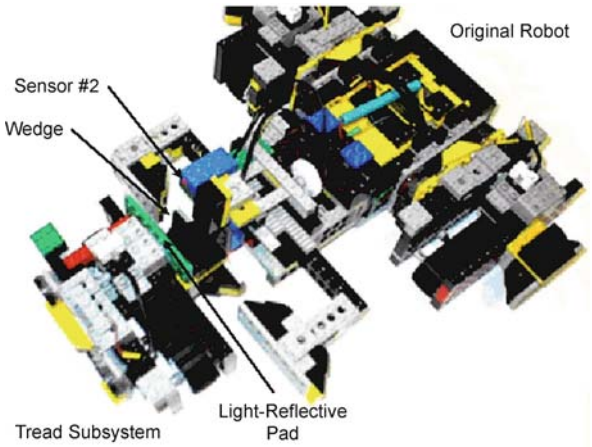
The left and right tread subsystems are designed to be identical to each other, with the purpose of reducing the system's design complexity. A tread subsystem hosts a rubber tread with a driving gear system, a 9 V LEGO DC motor, and a light-reflective pad which helps the original robot's navigation. One side of the tread has a set of magnets, a set of shape-constrained blocks, and a set of electrical connections, all of which correspond to the side of the controller subsystem. On the other side, the tread has a wedge which is fitted to the gripper. The wedge is used during the tread subsystem's transferring and assembling processes. Figure 6 shows how the original robot grasps



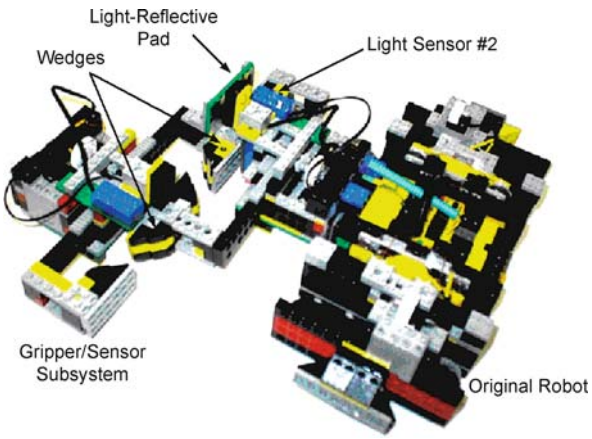
Self-replicating Robotic Systems, Figure 4
An assembly view of the self-replicating robot



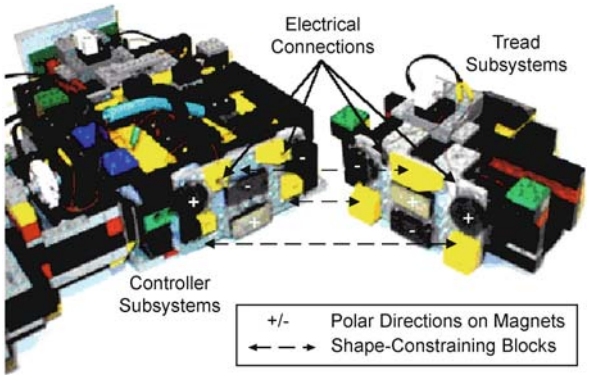
Self-replicating Robotic Systems, Figure 5
This diagram illustrates the concept of using polar magnets and shape-constraining blocks (*top*: correctly aligning, and *bottom*: incorrectly aligning)



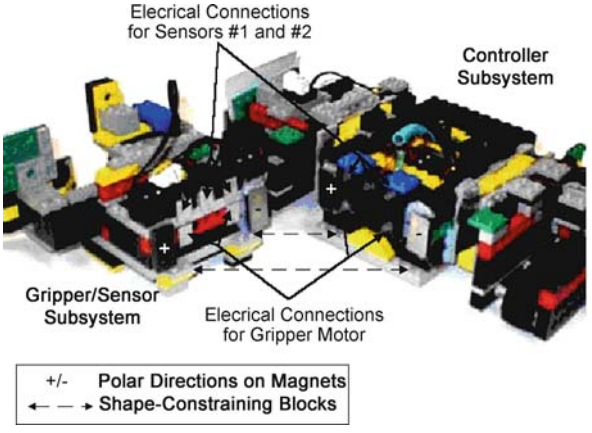
Self-replicating Robotic Systems, Figure 6
The original robot grasps the tread subsystem



Self-replicating Robotic Systems, Figure 8
The original robot grasps the gripper/sensor subsystem



Self-replicating Robotic Systems, Figure 7
The connections located between controller and tread subsystems



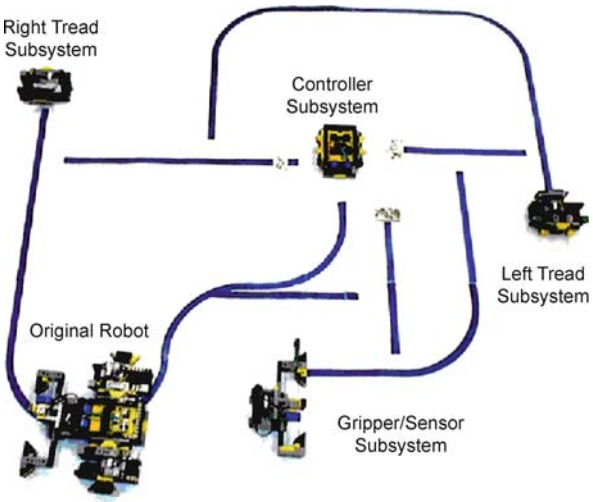
Self-replicating Robotic Systems, Figure 9
The connections located between gripper/sensor and controller subsystems

the tread subsystem, and Fig. 7 shows the connections located between the controller and tread subsystems.

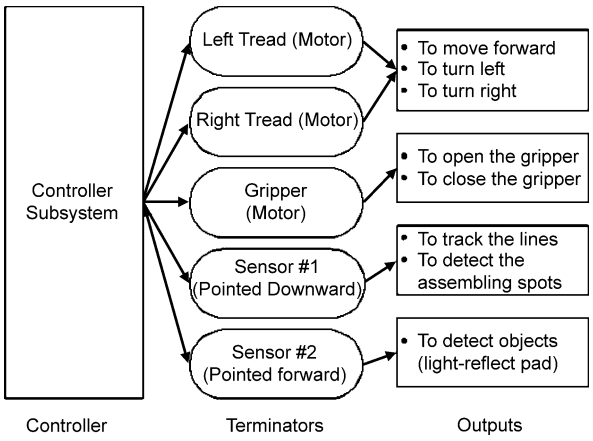
The gripper/sensor subsystem is comprised of a 9 V LEGO DC motor, a set of rack and pinion gears used to drive the left/right fingers of the gripper, a set of magnets, a set of shape-constrained blocks, a set of electrical connections, and two light sensors (one is pointed downward, and the other is pointed forward).

The set of magnets, shape-constraining blocks, and electrical connections are attached to their corresponding part, on the front side of the controller subsystem. The left finger of the gripper is designed in a wedge shape to be fitted with the gripper in any identical robot. This wedge is used in the same manner as in the tread subsystems during assembling processes. Figure 8 shows how the original robot grasps the gripper/sensor subsystem, and Fig. 9 shows the connections located between gripper/sensor and controller subsystems. The two LEGO light sensors are employed in the robot's navigation system. The first light sensor (pointed downward) is used to detect the blue painted lines and silver acrylic spots on the experiment surface. The second light sensor (pointed forward) is used to detect objects (the subsystems of the replica) which the robot runs into.

The experimental area is a 2 m x 3 m area made of white colored paper with lines and spots painted in blue and silver acrylic colors. The original robot starts at the initial position, and the replica's subsystems are at their locations. Figure 10 shows the experimental area with locations of the replica's subsystems and the initial position of the original robot.



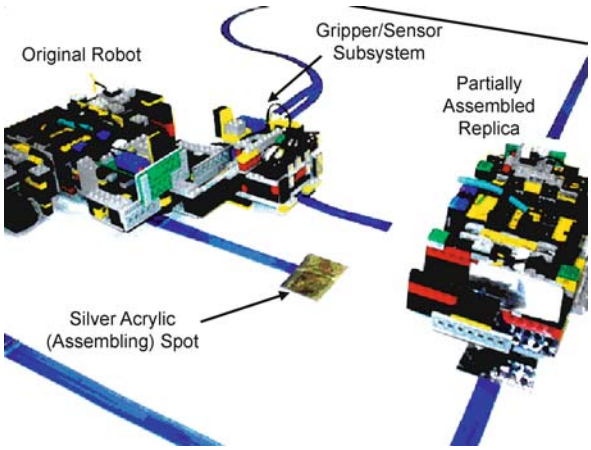
Self-replicating Robotic Systems, Figure 10
A map of the experimental area



Self-replicating Robotic Systems, Figure 11
The control architecture of the autonomous self-replicating robot

Controls and Programming of the Autonomous Self-Replicating Robot

The prototype robot is a fully autonomous system. Figure 11 shows the conceptual control architecture of the robot. The robot and its replicas, using the LEGO light sensor No. 1 (pointed downward), is capable of tracking the blue lines, and it can recognize the assembling spots, painted in a silver acrylic. The sensor detects and returns different analog values, corresponding to different colors. The robot tracks the painted lines to navigate between positions. Once the robot detects the assembling spot, the robot begins the assembling process. The LEGO light sensor No. 2 (pointed forward) returns an analog value once



Self-replicating Robotic Systems, Figure 12
The robot is searching for the assembly location while holding the replica's gripper/sensor subsystem

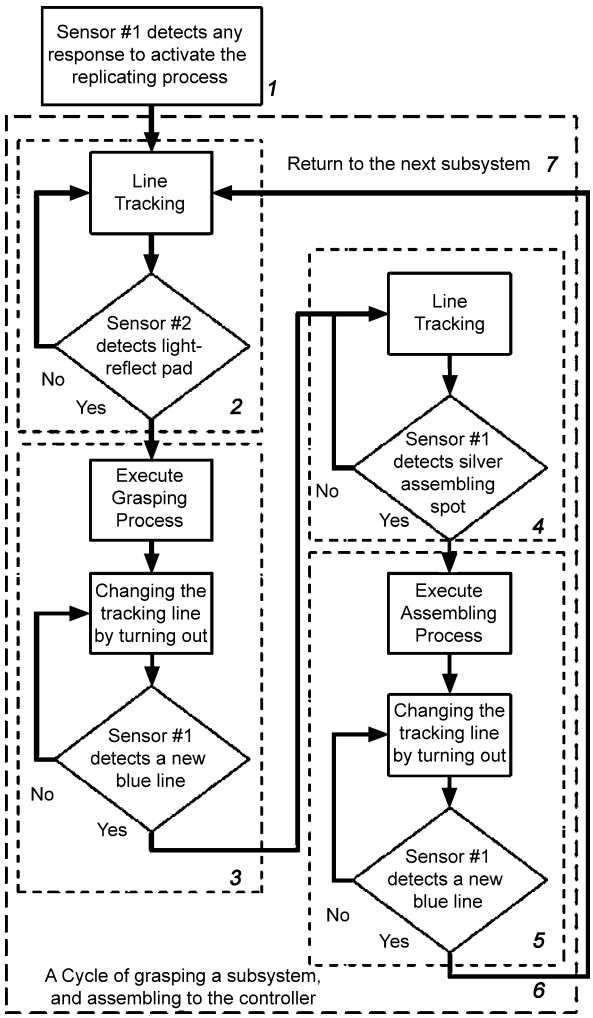
it detects a light-reflective pad attached to the tread and gripper/sensor subsystems. This notifies the robot to begin grasping the detected subsystem.

The grasping process consists of an aligning push toward the subsystem, and closing the gripper to grasp the subsystem. On the other hand, the assembly process consists of opening the gripper to release the subsystem, and an aligning push forward to snap the subsystem to the controller. Figure 12 shows the original robot grasping a subsystem, and moving toward an assembling spot, in silver acrylic, along the blue line.

The programming of the prototype is described here. The code is programmed on a PC and transferred through a LEGO infrared program-transferring tower. In the order in which events take place in the replication process, the programming is separated into seven stages: 1) replication process is activated, 2) line tracking and searching for a subsystem, 3) grasping the subsystem and changing to a new path which leads to the next step, 4) line tracking and searching for the assembly location, 5) assembling the subsystem to the controller and changing to a new path, 6) the new path leads to the next subsystem, 7) The final step loops back to steps 2 through 6 so the process repeats indefinitely. Figure 13 illustrates the programming flowchart of the self-replicating robot system.

Experiments and Results of the Autonomous Self-Replicating Robot

The following is a step-by-step outline of the procedure that our autonomous self-replicating robot system under-



Self-replicating Robotic Systems, Figure 13
The flowchart of the self-replication process

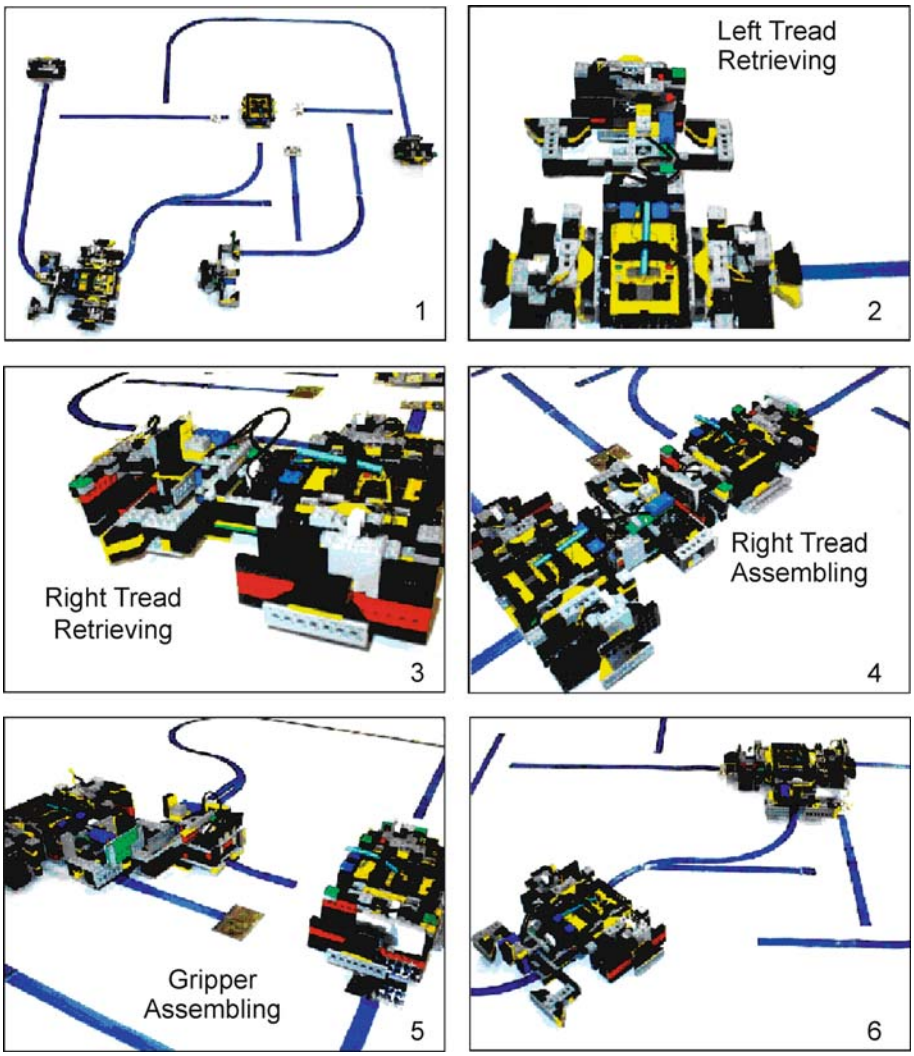
goes (Fig. 14 is a photographic representation of these step but they are not synchronous):

1. The original robot starts following the line from the starting point to the first subsystem using sensor No. 1.
2. Once light sensor No. 2 detects the first subsystem (right tread), the original robot begins the grasping process and grasps the right tread subsystem.
3. After the grippers are closed the original robot turns to the right until it detects a line.
4. The original robot follows the second line until it reaches the assembly location.
5. When light sensor No. 1 on the original robot detects the silver acrylic spot (the assembly location), the robot stops, and begins the attaching process.
6. The original robot opens the grippers, and gives a final push to secure the right tread subsystem to the controller subsystem.
7. The original robot then backs up and turns to the left until it detects a line value on sensor No. 1.
8. The original robot follows the line until it reaches the left tread subsystem.
9. Once light sensor No. 2 detects the second subsystem, the robot will stop, and begin the grasping process by closing its gripper around the left tread's wedge.
10. The original robot turns right until it detects the next line.
11. The original robot will follow the second line until it reaches the assembly location.
12. The original robot opens its gripper to release the left tread subsystem.
13. The original robot gives a final push on the left tread subsystem to help secure it.
14. The original robot then backs up and turns left until it detects the next line, using sensor No. 1.
15. The original robot follows the line to the final subsystem.
16. Once it reaches the gripper/sensor subsystem, it stops, and begins the grasping process.
17. The original robot closes its gripper, and turns right until it detects a line value with sensor No. 1.
18. The gripper/sensor subsystem is now transferred to the assembly location.
19. Once the original robot reaches the assembly location it stops, and opens the gripper.
20. The original robot backs up and turns left until sensor No. 1 is a line value.
21. The original robot then follows the line back to the starting point, and is ready to replicate again.
22. The completed replica self-activates (20 seconds after completion) and begins following the line to the starting point.
23. Once each robot reaches the starting point, it begins the replication procedure again.

The replication process takes two minutes and fifteen seconds per cycle. Although each subsystem is required to be placed in its starting location, errors in initial position and orientation are not very critical. We found slight errors during the grasping process in a few experiments caused by improper placement of the subsystems. Overall, the system is robust and very repeatable.

Our Approach to Self-Replicating Control Circuitry

Ideally, for a robotic system to be truly self-replicating, it would have to demonstrate the ability to assemble all of



Self-replicating Robotic Systems, Figure 14
Self-replication process: 1) the original robot begins at the initial position with every part placed in their position. 2) The robot detects the right-tread subsystem. 3) The robot grasps the right-tread subsystem, and attaches it to the controller. 4) In the same manner, the robot performs the assembly of the left-tread subsystem. 5) After the robot grasps the gripper/sensor subsystem the robot transfers the subsystem to the next assembly step at the controller. 6) After being fully assembled, the replica is self-activated, and ready to replicate just like the original

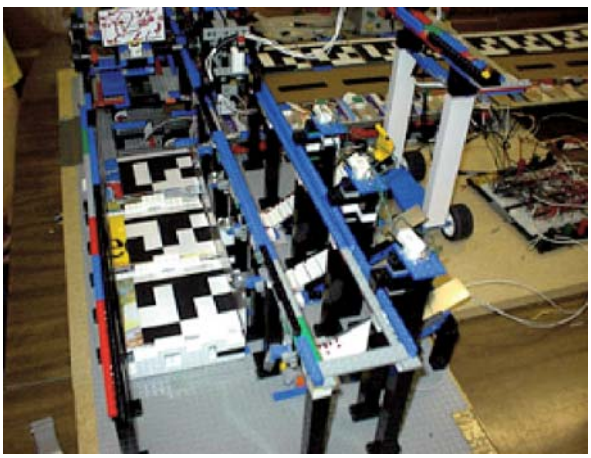
its own subsystems from the most fundamental components. In the case of the robot controller, we consider the most fundamental components to be transistors, resistors, capacitors, etc., whereas microcontrollers are too complex to be considered as basic elements.

Our approach is to build a circuit capable of controlling an electro-mechanical system to re-build replicas of the control circuit from the most fundamental electronic components. In the von Neumann universal constructor paradigm, an associated instruction code is also required. In contrast it is possible to replicate a particular system by self-inspection without invoking von Neumann's universal constructor. We illustrate both concepts in hardware designed and constructed by students in a Mechatronics course taught at Johns Hopkins University in 2003. Two prototypes illustrate replication by self-inspection, and one demonstrates the universal constructor. In all three cases, pre-built electro-mechanical systems (called the SRI-builders) use the transistorized circuit as its controller. While in the von Neumann paradigm, the controller follows instructions that are explicitly encoded (and hence must reproduce the code for the overall system to

be self-replicating), in the self-inspection paradigm, actions are taken implicitly as a result of observing the spatial layout of components in the original and feeding that information into the circuit itself. Clever electromechanical design ensures that observations obtained during self-inspection are translated directly into actions without requiring the interpretive step of consulting a long sequence of encoded construction commands.

A von Neumann Universal Constructor Prototype

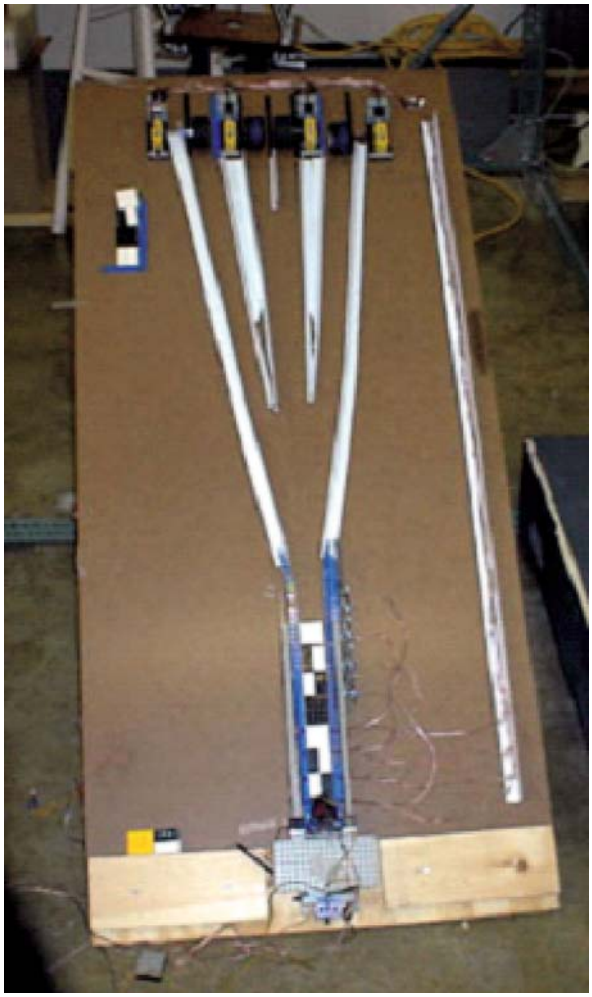
The prototype is a two-arm gantry-style robot with two degrees of freedom. The first degree of freedom moves along the whole system, includes the feeders and the assembly boards. The second degree of freedom moves vertically to pick and place codes and circuit pieces. A controller circuit is used to control motions of the robot. There are two boards being assembled at a certain amount of time. The first board is the replicated circuit, and the second is the replicated codes. The circuit is pre-wired before it is placed into the carrier. This is similar to a process of placing a chip into a circuit board. In the code part, each code consists of three lines of black and white strip representing 3 bits. A series of codes is set up on a code array where it is fed simultaneously to a reader array. The reader array is made of an array of photo-transistors and infrared LED emitters used to read black (0) and white (1) colors in each code. Infrared LED emitters are used instead of regular LED for reducing problem with ambient light distractions. Figure 15 shows the system in a side view. Once the acquisition part of the system reads the code, the assembly part of the system works by following the instruction code, replicating the code and circuit of that part.



Self-replicating Robotic Systems, Figure 15
Side view of the replicating system (von Neumann universal constructor prototype)

Non-Universal Self-Replication by Self-Inspection (Design 1)

This design is the first design of a non-universal self-replication by self-inspection. The self-replicating control circuit has the ability to identify the proper electronic components required, translate information about its own constituent parts obtained from self-inspection into mechanical tasks that create a replica, and transfer all functions to the replica. There is no list of instructions in the form of a code. Each electronic component has a black-and-white color code. Parts are loaded into feeders, and as a reading head traverses the control circuit, the information about which part of the control circuit is being observed is fed into the circuit itself. This actuates the solenoid in the ap-



Self-replicating Robotic Systems, Figure 16
Side view of the replicating system (non-universal self-replication by self-inspection – design 1. (See [11] for details))

propriate feeder to release the parts needed to form the replica. Parts then slide down an incline and form an orderly array. The reading head continues to move and creates replicas until resources are completely utilized or its track ends. The design is scalable and the components are modular, allowing many different levels of intelligence to be replicated. This concept is one of many which we are investigating to enable self-replicating robots to perform complex behaviors. See Fig. 16.

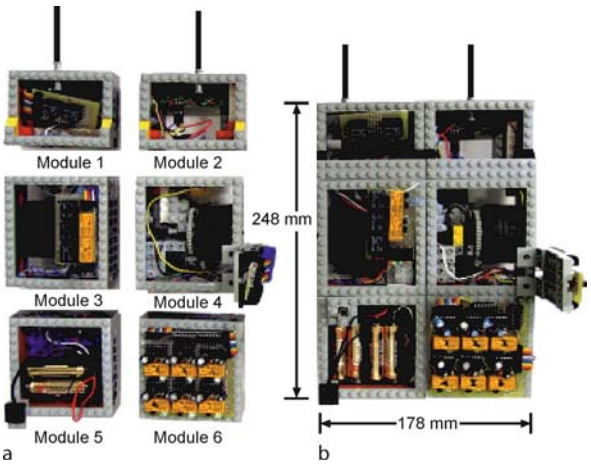
Non-Universal Self-Replication by Self-Inspection (Design 2)

This robotic system is an X-Y table constructed from modified LEGO components. A photo-transistor sensor system is attached to the end-effector of the X-Y system in order to inspect the control circuit (the components of which are each assigned a unique black and white code). On the top of the X-Y system, a set of component feeders is installed. The circuit converts the signal from the sensor system to control the component feeders to release the correct component to the parts assembler. The parts assembler then arranges all the components to create a new replica of the control circuit. See Fig. 17.

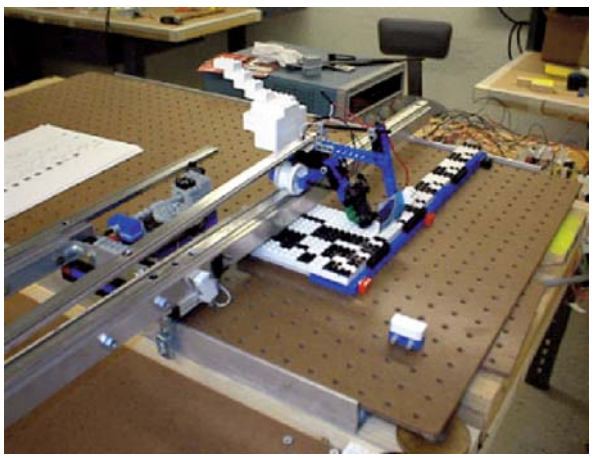
Merging the Self-Replicating Robot and Self-Replicating Circuit Concepts

In prior prototypes we considered two disjoint cases: (1) self-replicating robots constructed from modules where one module contained a computer; (2) a control circuit that commanded a mechanical device to assemble copies of the circuit. In this section we review very recent work with K. Lee and the JHU authors that merges these two concepts.

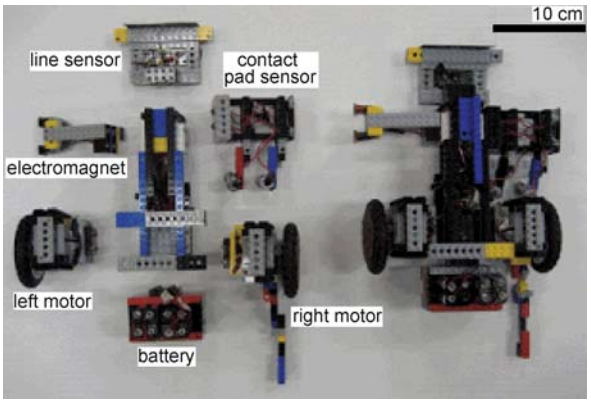
Basically, the autonomous self-replicating robot described previously in this paper functions as a finite-state machine in a highly structured environment. This same behavior can be implemented by a robot controlled by a simple circuit without using a microprocessor. Discrete electronics elements of this circuit such as transistors, resistors, etc., can be distributed over the modules from which the robot is constructed. This means that as the original robot assembles these modules to form a replica, the controller for the replica is assembled as the robot is. Prototypes are shown in Figs. 18 and 19.



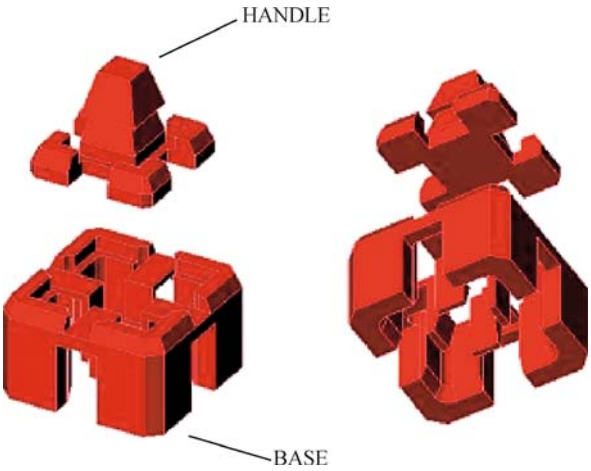
Self-replicating Robotic Systems, Figure 18
A modular self-replicating robot controlled by a finite state circuit (see [15,16])



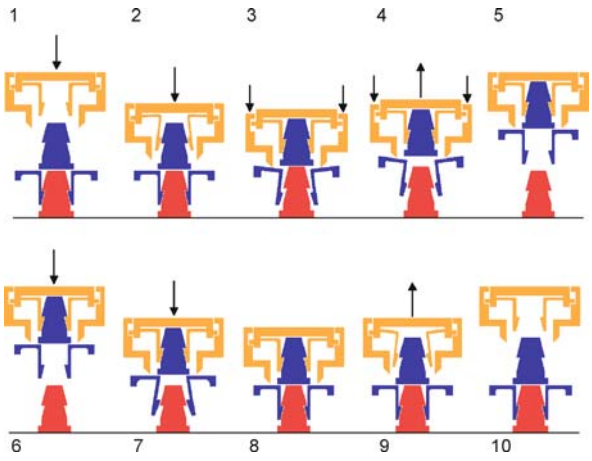
Self-replicating Robotic Systems, Figure 17
Side view of the replicating system (non-universal self-replication by self-inspection – design 2)



Self-replicating Robotic Systems, Figure 19
Example of another modular self-replicating robot controlled by a finite state machine (see [4,16,18] for details)



Self-replicating Robotic Systems, Figure 20
Two exploded views of the basic component. The handle and base are individually cast in polyurethane, then bonded together with epoxy



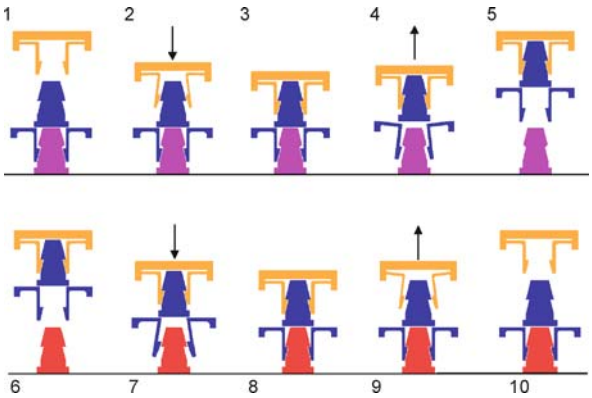
Self-replicating Robotic Systems, Figure 21
Steps in reversible assembly process. The tool (top part) grasps, retrieves, and reconnects one component to another

Towards a Universal Constructor

Whereas other prototypes of self-replicating robotic systems use relatively few subsystems as the initial parts, and connect these subsystems with magnets, the desire to create machines that can reproduce from a large number of basic parts requires thought about mechanical design, manufacturing, and assembly issues. Therefore, in this section the work of the second author in [20] is reviewed. In this work, a novel set of mechanical components that can be assembled into a wide variety of devices is presented. The components are specifically designed to be handled and assembled by devices made of the same type of components. A 3-axis Cartesian manipulator built with these components is presented. It is shown that in principle the manipulator can assemble duplicates of itself in addition to arbitrary devices when provided with properly oriented components and controlled by a human operator.

All components in the component set are variants of the "basic component". Figure 20 shows a cutaway view of two basic components. Each component consists of two parts: an upper "handle" and a lower "base". The base dimensions are 3.8 by 3.8 by 2.3 cm. The handle and base are individually cast from polyurethane resin in a silicone mold and then fixed together with epoxy adhesive. The base contains four compliant snap tangs that are complementary to the four tapered surfaces of the handle. Figure 21 shows how the snap tangs of the base grasp the undercut of the handle of another part. Additionally, the convex tapered surfaces of the base mate with concave tapered surfaces of the handle in order to provide rigidity with respect to shear and torsion between parts. Chamfers on the edges of the parts allow a certain amount of positioning error during assembly. The tangs and handle also contribute to error tolerance, since they are tapered such that two parts need not be exactly positioned before assembly. The slots cut in each side of the base allow the part to be assembled onto other parts that have reinforcing segments. The tangs are L-shaped, and the top part of each tang is exposed by a slot cut in each side of the handle.

The parts have a common handle, so they all can be manipulated by a single grasping tool. A part can be connected to another part by simply lowering it into place such that the tangs engage the handle of the other part (sliding components are slid sideways onto the handle of the other part). Depending on the grasping tool used, part connections can be reversible or irreversible. A section view of the steps in reversible assembly are shown in Fig. 21. The grasping tool (upper part) has thin tangs and a release mechanism. The process is initiated when the grasping tool is lowered onto the part to be disconnected. The release mechanism is engaged, depressing the tangs of the part to be disconnected. The upper part of the tangs deflect, disengaging the tangs of the middle part from the handle of the lower part (step 3). In step 4 the grasping tool and middle part are lifted from the lower part. To re-connect the part, the grasping tool and part are lowered onto the destination part (step 7). The grasping tool is removed simply by lifting it from the newly connected part. The snap tangs in the grasping tool are thinner than normal tangs, so they deflect first and disengage the grasping tangs from the handle on the newly connected part (step 9).

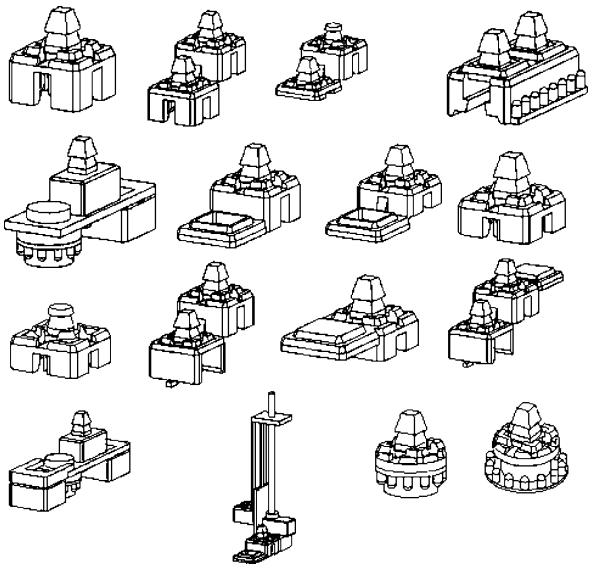


Self-replicating Robotic Systems, Figure 22
Steps in irreversible assembly process. The tool retrieves a part from a special storage site and attaches it to another component

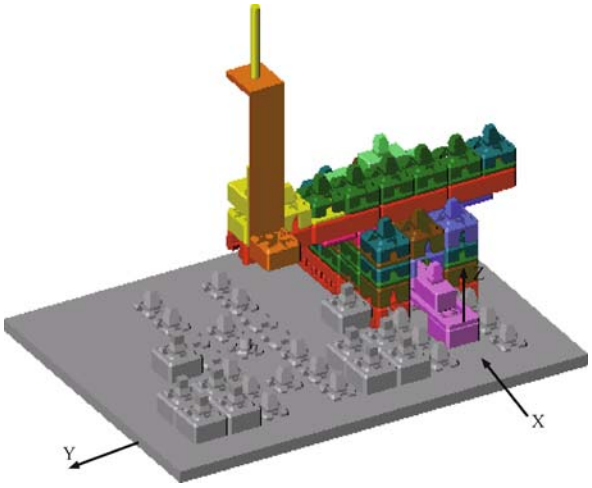
Irreversible assembly is also possible, as shown in Fig. 22. In this case, the grasping tool has thin tangs like the reversible tool, but there is no release mechanism. Parts must be stored at a special "storage site" before assembly. The storage site (lower part in steps 1–5) is a normal part with reduced undercut on the handles. The amount of undercut on the storage site handles and the thickness of the tangs on the tool are carefully chosen so that the force required to disassemble two components is greatest between two normal parts, and least between a normal part and the storage site. This allows the grasping tool to pick a part from a storage site and connect it to another part with two simple up/down vertical movements. The irreversible assembly method, due to its simplicity, is used in the manipulator described below.

These basic components can be used to form a constructing machine capable of assembling a wide variety of devices, including duplicates of itself. By "capable" is meant kinematically capable - the machine can grasp, manipulate, and place all of the components required for constructing a duplicate of its "mechanical self". The machine's "mechanical self" includes the hardware required to manipulate mechanical components, but excludes the portion responsible for control. The machine lacks any type of sensor or onboard control, and so must be controlled either by a human operator or a sophisticated external controller. By "assemble" is meant the retrieval of components from a storage site and subsequent connection of them to an assembly or operating plane. Components must arrive at the storage site through action of another entity – namely a human operator.

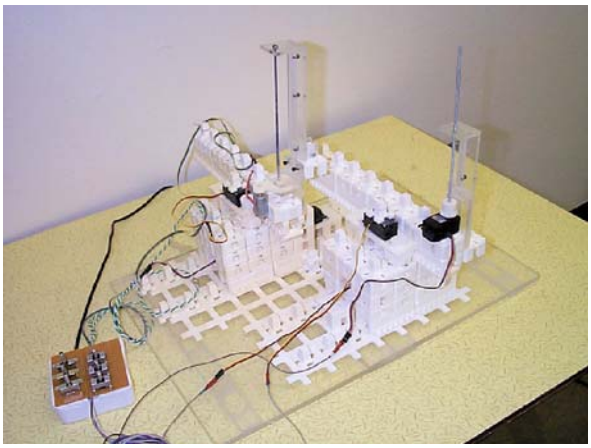
A total of 16 different variations on the basic component were built. Most of these are used in the manipulator, and the remainder are useful for constructing other types of devices. Figure 23 shows a diagram of these parts. Figure 24 shows a CAD model of a constructing machine made from parts in the component set. It is essentially a 3-axis Cartesian manipulator. The base of the machine is a platform that slides along the x-axis, and is actuated by a DC motor. A motor-driven boom rides on top of the platform, and slides along the y-axis. The end-effector is mounted to a vertically-linear motor on the end of the boom. The reachable space of the end-effector is a rectangular volume of size $3 \times 4 \times 6$ in units of component-widths. Note that the tool does not rotate. This requires



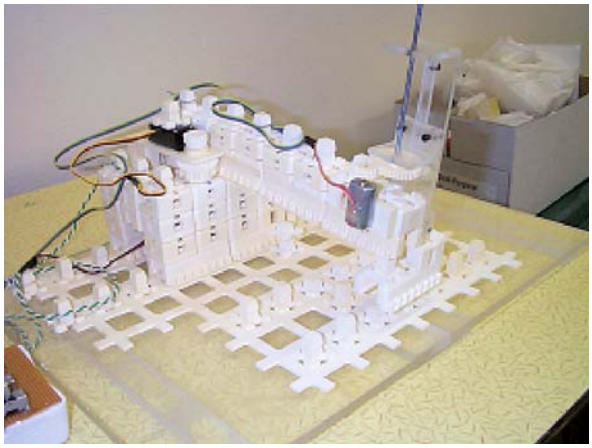
Self-replicating Robotic Systems, Figure 23
Sixteen variations on the basic component



Self-replicating Robotic Systems, Figure 24
A constructing machine made from parts in the component set



Self-replicating Robotic Systems, Figure 25
An initial physical implementation of a universal constructor [20]



Self-replicating Robotic Systems, Figure 26
Large downward deflection of extended boom

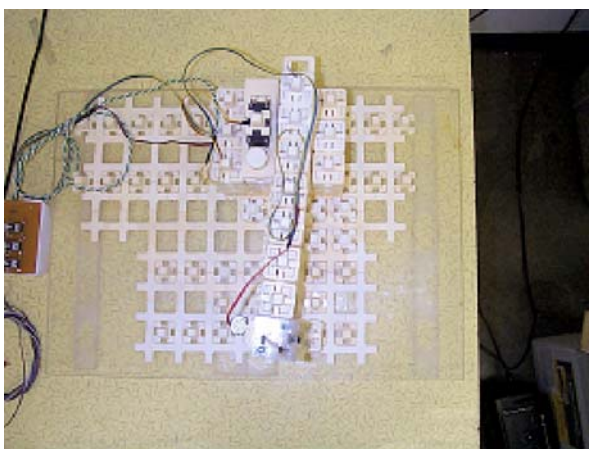
parts to be loaded at the storage site in their proper orientation. The constructor has a total of 45 components of eleven different types, not counting the operating plane. The details of the assembly process of this machine can be found in [20].

Figure 25 shows the experimental setup in its finished state. Both devices are operable. A human operator controls them through a switchbox that turns the various motors on or off. A topic that has been neglected so far is that of running wires to the motors. As seen from the figure, this is done rather haphazardly. This prototype had a number of problems that are worth mentioning which are helping us to design better systems.

The first problems that were encountered dealt with getting the constructor boom to move smoothly. When the boom is extended, it applies high forces to the parts that hold it to the platform. This results in more friction, and requires a higher torque from the motor. In these conditions the motor tended to separate from the platform, and the drive gear would lose contact with the racks on the boom. This problem was eventually "solved" by using a metal screw and adhesive between certain parts to hold them in place.

Most of the other problems occurred during assembly, and centered around the extended boom. Figure 26 shows the large downward deflection occurring in the extended boom. This made it difficult to align parts. A human operator could successfully place parts, using repeated effort and visual feedback, but a simple controller surely could not.

Deflection of the boom in the x-direction presented problems for sliding the far racks into place (Fig. 27). This would cause the sliding platform under the boom to jam.

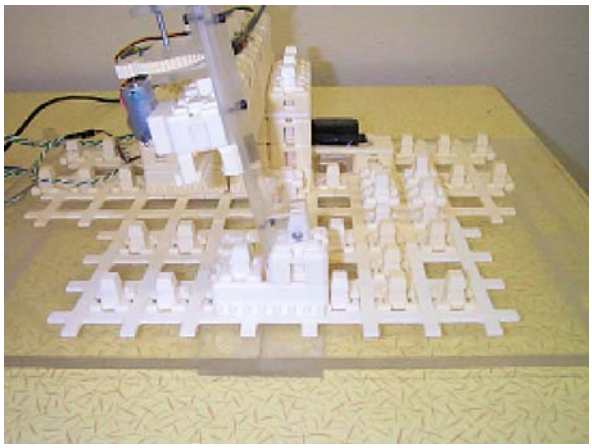


Self-replicating Robotic Systems, Figure 27
Sideways deflection of extended boom caused the sliding platform to jam

In addition to sideways deflection, the boom would also twist along the y-axis (Fig. 28). The constructor was able to place the near racks, but the far racks had to be placed by hand.

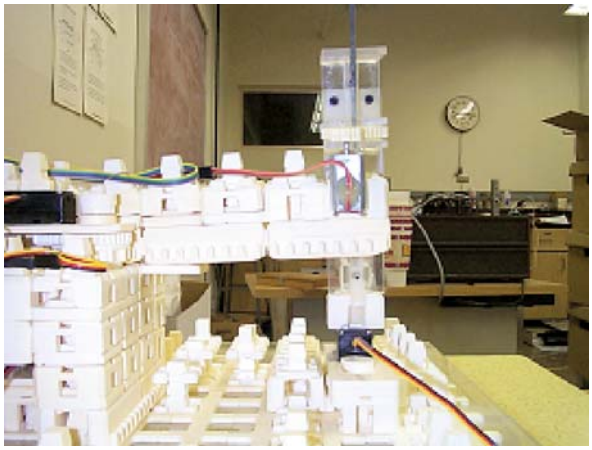
The torsion of the boom created problems in later assemblies steps as well. The twisting deflection of the boom would tend to misalign parts so that their snap-tangs would not engage with the parts below them. For this reason, the tool was mounted to the slide in a rotating bracket. This allowed the tool a small, passive amount of rotation about the y-axis. With this feature, the boom could still twist, but the parts would self-align and engage.

In many cases, the constructor could not develop the necessary force to assemble components. This was especially true when assembling components with multiple sets

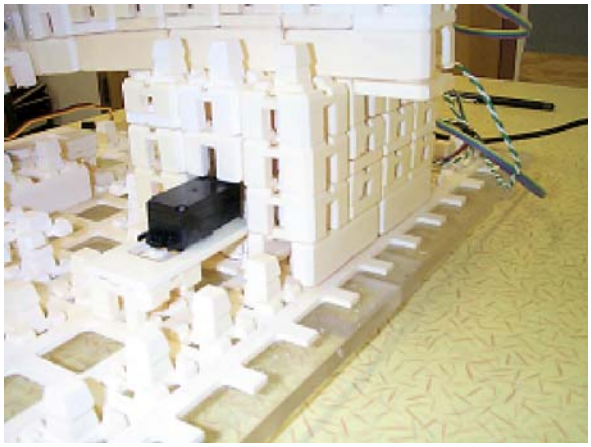


Self-replicating Robotic Systems, Figure 28
Torsion of boom

Self-replicating Robotic Systems, Figure 30
Separation of components within the boom



Self-replicating Robotic Systems, Figure 29
Upward deflection of boom



Self-replicating Robotic Systems, Figure 31
Separation of sliding platform and operating plane

of snap tangs, and in parts where the tangs were offset horizontally from the handle. Sometimes this occurred because the vertical linear actuator did not produce enough force, and sometimes it occurred because upward deflection of the boom caused the parts to misalign (Fig. 29). In a few cases, components in the boom would separate (Fig. 30). This only occurred when the boom was prevented from deflecting because it was engaged on the growing assembly.

The extended boom indirectly caused another problem. Recall that two of the handles in the tracks on the operating plane are modified as entry sites for racks. These sites do not help hold the platform to the plane. This, in combination with the high moment exerted on the platform by the extended boom, would cause the platform to separate from the plane and get stuck on adjacent, non-modified handles (Fig. 31).

Some aspects of the constructor worked well, such as the stage-by-stage assembly of the boom and the retrieval of parts from the storage site. The boom does not have to extend to reach the storage site, so deflections are not a problem and part retrieval was easy. The thickness of the tool tangs was found by trial and error – the tangs of a basic part were successively machined until the disassembly force fell within the correct range. Although no measurements were taken, the distribution of assembly and disassembly forces across various parts was fairly tight. The tool could lift all of the parts. Once parts were connected to an assembly, the tool could be separated from the connected parts in all cases.

In short, this initial concept of a desktop universal constructor provided us many valuable lessons about the importance of mechanical design, parts assembly and manufacturability issues that we are incorporating in future designs.

Discussion

Several self-replicating robot prototype have been constructed and tested. An autonomous prototype uses two light sensors in its navigation system to detect objects and also to track lines. Magnets and shape-constraining blocks are used to aid in aligning and interlocking the subsystems of the replica. As a result, the robot is capable of automatically assembling its replicas. All of the replicas are also capable of completing the same replicating process. The autonomous self-replicating system presented here has been recognized by authors of the book, "Kinematic Self-Replicating Machines" [25] that this prototype is the world's first fully functional autonomous self-replicating robot. Self-replicating circuits, as well as self-replicating robots whose controllers can be fully decomposed into basic parts have also been reviewed here.

Whereas von Neumann's architecture for self-replicating kinematic automata is the most widely known approach, it is not the only one. Self-reproduction by self-inspection in which a non-universal constructor "reads" an original device and "writes" a copy by executing a very small set of hardwired commands is an alternative. In our experience observing students attempting to build self-replicating devices, self-replication by self-inspection appears to be a more robust and less complicated alternative to the universal constructor.

Future Works

In previous sections we discussed a series of mechanical replicating prototypes in which programs are preloaded onto control computers which are then treated as one of several subsystems to be assembled. However, our ultimate goal is to develop a self-replicating robotic system capable of autonomously assembling its replicas from simple components using only electro-mechanical intelligence, i. e. a mechanical code, and transistor-based control circuits. This eliminates complicated electronic components, such as programmable micro-controllers, and makes the concept more appropriate for future space systems that can use in-situ resources for self-replication. The prototypes reviewed here are one step in this direction. Other related technologies are discussed in [33].

Acknowledgments

This work was made possible by support from the Royal Thai government, the Mellon Foundation, and the National Science Foundation under Grant ITS 0205466. The results and opinions expressed are solely those of the authors. We thank A. Cushing, Y. Kwon, K. Lee, M. Kutzer and all of the undergraduate students in the Mechatronics courses over the past five years who have designed several of the prototypes shown in this review. The work presented in Sect. "Towards a Universal Constructor" was performed at University of New Mexico with help from Professors G. Starr, H. Tran, and J. Wood.

Bibliography

1. Chirikjian GS, Pamecha A, Ebert-Uphoff I (1996) Evaluating Efficiency of Self-Reconfiguration in a Class of Modular Robots. J Robot Syst 13(5):317–338
2. Chirikjian GS, Suthakorn J (2002) Towards Self-Replicating Robots. In: Proceedings of the Eight International Symposium on Experimental Robotics (ISER), Italy, July 2002
3. Chirikjian GS, Zhou Y, Suthakorn J (2002) Self-Replicating Robots for Lunar Development. IEEE/ASME Trans Mechatron 7(4):462–472, December 2002
4. Eno S, Mace L, Liu J, Benson B, Raman K, Lee K, Moses M, Chirikjian GS (2007) Robotic Self-Replication in a Structured Environment without Computer Control. In: Proceedings of the 2007 IEEE CIRA. Piscataway
5. Freitas RA Jr (1980) A Self-Reproducing Interstellar Probe. J Br Interplanet Soc 33:251–264
6. Freitas RA Jr (1980) Report on the NASA/ASEE Summer Study on Advanced Automation for Space Missions. J Br Interplanet Soc 34:139–142
7. Freitas RA Jr (1983) Terraforming Mars and Venus Using Self-Replicating Systems. J Br Interplanet Soc 36:139–142
8. Freitas RA Jr, Merkle RC (2004) Kinematic Self-Replicating Machines. Landes Bioscience, Georgetown
9. Freitas RA Jr, Valdes F (1980) Comparison of Reproducing and Non-Reproducing Starprobe Strategies for Galactic Exploration. J Br Planet Soc 33:402–408
10. Freitas RA Jr, WP Gilbreath (eds) (1982) Advanced Automation for Space Missions. In: Proceedings of the 1980 NASA/ASEE summer study. Replicating Systems Concepts: Self-Replicating Lunar Factory and Demonstration, NASA, Scientific and Technical Information Branch (Conference Publication 2255), US Government Printing Office, Washington DC, chap 5
11. Hastings WA, Labarre M, Viswanathan A, Lee S, Sparks D, Tran T, Nolin J, Curry R, David M, Huang S, Shuthakorn J, Zhou Y, Chirikjian GS (2004) A minimalist parts manipulation system for a self replicating electromechanical circuit. IMG'04, Genoa, Italy, July 1–2, 2004
12. Hosokawa K, Fujii T, Kaetsu H, Asama H, Kuroda HY, Endo I (1999) Self-organizing collective robots with morphogenesis in a vertical plane. JSME Int J Ser C-Mechanical Syst Mach Elements Manuf 42(1):195–202
13. Jacobson H (1958) On Models of Reproduction. Am Scient 46:255–284

14. Kotay K, Rus D, Vona M, McGray C (1998) The Self-reconfiguring Molecule: Design and Control Algorithms. WAFR'98: Proceedings of the third workshop on the algorithmic foundations of robotics. Publisher A.K. Peters, Ltd. Wellesley
15. Lee K, Chirikjian GS (2007) Robotic Self-Replication. Low Complexity Parts. IEEE Robotics and Automation Magazine 14(4):34–43. Published by IEEE, Piscataway
16. Lee K, Moses M, Chirikjian GS (2008) Robotic Self-Replication in Structured Environments: Physical Demonstrations and Complexity Measures. Int J Robot Res 27(3-4):387–401. doi:10.1177/0278364907084982
17. Lipson H, Pollack B (2000) Automatic design and Manufacture of Robotic Lifeforms. Nature 406:974–978
18. Liu A, Sterling M, Kim D, Pierpont A, Schlothauer A, Moses M, Lee K, Chirikjian GS (2007) A Memoryless Robot that Assembles Seven Subsystems to Copy Itself. In: Proceedings of the 2007 IEEE ISAM, Piscataway
19. Moore EF (1956) Artificial Living Plants. Scient Am 195:118–126
20. Moses M (2001) A Physical Prototype of a Self-Replicating Universal Constructor. M.S. University of New Mexico
21. Murata S, Kurokawa H, Kokaji S (1994) Self-Assembling Machine. In: Proceedings of the 1994 IEEE International Conference on Robotics and Automation. Piscataway, San Diego, CA, pp 441–448
22. Penrose LS (1959) Self-Reproducing Machines. Sci Am 200(6):105–114
23. Pesavento U (1995) An implementation of von Neumann's self-reproducing machine. Artif Life J 2(4):337–354
24. Pfeifer R, Kunz H, Weber MM, Thomas D (2001) Lecture of the Artificial Life. http://www.ifi.unizh.ch/ailab/teaching/AL01/chap7.pdf. Dept of Information Technology, University of Zurich. Accessed 19 June 2008
25. Rebek J Jr (1994) Synthetic Self-Replicating Molecules. Sci Am 271(1):48–55
26. Sipper M (1998) Fifty Years of Research on Self-Replication: An Overview. Artif Life 4(3):237–257
27. Suthakorn J (2004) Paradigm for Service Robotics. Ph D Dissertation. Johns Hopkins University
28. Suthakorn J, Kwon YT, Chirikjian GS (2003) A Semi-Autonomous Replicating Robotic System. In: Proceedings of the 2003 IEEE/ASME International Symposium on Computational Intelligence for Robotics and Automation (CIRA), Kobe, Japan, Piscataway
29. Suthakorn J, Zhou Y, Chirikjian GS (2002) Self-Replicating Robots for Space Utilization. In: Proceedings of the 2002 Robosphere workshop on Self Sustaining Robotic Ecologies, NASA Ames Research Center, California, NASA Ames Research Center
30. Tiesenhausen GV, Darbro WA (1980) Self-Replicating Systems – A Systems Engineering Approach. In: Technical Memorandum: NASA TM-78304, Washington DC
31. Von Neumann J, Burks AW (1966) Theory of Self-Reproducing Automata. University of Illinois Press,Champaign
32. Whitesides GM (1995) Self-Assembling Materials. Sci Am 273(3):146–149
33. Yim M, Shen WM, Salemi B, Rus D, Moll M, Lipson H, Klavins E, Chirikjian GS (2007) Modular self-reconfigurable robot systems – Challenges and opportunities for the future. IEEE Robotics Autom Mag 14(1):43–52
34. Yim M, Zhang Y, Lamping J, Mao E (2001) Distributed Control for 3D Metamorphosis. Auton Robots 10:41–56

Self-Replication and Cellular Automata

GIANLUCA TEMPESTI[1], DANIEL MANGE[2], ANDRÉ STAUFFER[2]
[1] University of York, York, UK
[2] Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Article Outline

Glossary
Definition of the Subject
Introduction
Von Neumann's Universal Constructor
Self-Replication for Artificial Life
Other Approaches to Self-Replication
Future Directions
Bibliography

Glossary

Cellular automaton A cellular automaton (CA) is a mathematical framework modeling an array of *cells* that interact locally with their neighbors. In this *cellular space*, each cell has a set of *neighbors*, cells have *values* or *states*, all the cells update their values simultaneously at discrete *time steps* or *iterations*, and the new state of a cell is determined by the current state of its neighbors (including itself) according to a local *function* or *rule*, identical for all cells. In the article, the term is extended to account for systems that introduce variations to the basic definition (for example, systems where cells do not update simultaneously or do not have the same set of rules in every cell).
Following the historical pattern, in the article the same term is also used to refer to an object or structure built within the cellular space, i. e., a set of cells in a particular, usually active, state (overlapping with the definition of *Configuration*).

Configuration A set of cells in a given state at a given time. Usually, but not always, the term refers to the state of all the cells in the entire space. The *initial configuration* is the state of the cells at time $t = 0$.

Self-replication The process whereby a cellular automaton configuration creates a copy of itself in the cellular space. Incidentally, you will note that in the article we use the terms self-replication and self-reproduction interchangeably. In reality, the two terms are not really synonyms: self-reproduction is more properly applied to the reproduction of organisms, while self-replica-

tion concerns the cellular level. The more correct term to use in most cases would probably be self-replication, but since von Neumann favored self-reproduction, we will ignore the distinction.

Self-reproduction See *Self-Replication*

Construction The process that occurs when one or more cells, initially in the inactive or *quiescent* state are assigned an active state (in the context of this article, by the self-replicating structure).

Definition of the Subject

Machine self-replication, besides inspiring numerous fictional books and movies, has long been considered a powerful paradigm to allow artifacts, for example, to survive in hostile environments (such as other planets) or to operate more efficiently by creating *populations* of machines working together to achieve a given task. Where the self-replication of *computing* machines is concerned, other motivations can also come into play, related to concepts such as fault tolerance and self-organization.

Cellular automata have traditionally been the framework of choice for the study of self-replicating computing machines, ever since they were used by John von Neumann, who pioneered the field in the 1950s. In this context, self-replication is seen as the process whereby a configuration in the cellular space is capable of creating a copy of itself in a different location.

As a mathematical framework, CA allow researchers to study the mechanisms required to achieve self-replication in a simplified environment, in view of eventually applying this process to real-world systems, either to electronics or, more generally, to computing systems.

Introduction

The self-replication of computing systems is an idea that dates back to the very origins of electronics. One of the pioneers of the field, John von Neumann, was among the first to investigate the possibility of creating machines capable of self-replication [1] with the purpose of achieving reliability through the redundant operation of "populations" of computing machines.

Throughout the more than 50 years since von Neumann's seminal work, research on this topic has gone through several transformations. While interest in applying self-replication to electronic systems waned because of technological hurdles, the field of Artificial Life, starting with the pioneering work of Chris Langton [14], began studying this process in the more general context of achieving life-like properties in artificial systems.

Throughout its long history, cellular automata (CA) have remained one of the environments of choice to study how self-replication can be applied to computing systems. In general, researchers in the domain (including von Neumann) have never regarded CA as the environment in which self-replication would be ultimately applied. Rather, CA have traditionally provided a useful platform to test the complexity of self-replication at an early stage, in view of eventually applying this process to real-world systems, either to electronics or, more generally, to computing systems.

Of course, the concept of self-replication has been applied to artificial systems in contexts other than computing. A classic example is the 1980 NASA study by Robert Freitas Jr. and Ralph Merkle [10] (recently expanded in a remarkable book [11]), where self-replication is used as a paradigm for efficiently exploring other planets. However, this kind of self-replication, applied to physical machines rather then computing systems, does not commonly make use of cellular automata and is beyond the scope of this article.

Following the historical progress of self-replication in cellular automata (derived in part from [40]), we will first examine in some detail von Neumann's seminal work (Sect. "Von Neumann's Universal Constructor"). Then, the use of self-replication as an Artificial Life paradigm will be discussed (Sect. "Self-Replication for Artificial Life") before dealing with some of the latest advances in the field in Sect. "Other Approaches to Self-Replication".

Von Neumann's Universal Constructor

Many of the existing approaches to the self-replication of computing systems are essentially derived from the work of John von Neumann [1], who pioneered this field of research in the 1950s.

Von Neumann, confronted with the lack of reliability of computing systems, turned to nature to find inspiration in the design of fault-tolerant computing machines. Let us remember that the computers von Neumann was familiar with were based on vacuum-tube technology, and that vacuum tubes were much more prone to failure than modern transistors. Moreover, since the writing and the execution of complex programs on such systems represented many hours (if not many days) of work, the failure of a system had important consequences in wasted time and effort.

In particular, Von Neumann investigated self-replication as a way to design and implement digital logic devices. Unfortunately, the state of the art in the fifties restricted von Neumann's investigations to a purely theoretical level, and the work of his successors mirrored this constraint.

Indeed, it is not until fairly recently that some of the technological problems associated with the implementation of such a process in silicon have been resolved with the introduction of new kinds of electronic devices (see Sect. "Other Approaches to Self-Replication").

In this section, we will analyze von Neumann's research on the subject of self-replicating computing machines, and in particular his universal constructor, a self-replicating cellular automaton [44].

Von Neumann's Self-Replicating Machines

Natural systems are among the most reliable complex systems known to man, and their reliability is a consequence not of any particular robustness of the individual cells (or organisms), but rather of their extreme redundancy. The basic natural mechanism which provides such reliability is *self-reproduction*, both at the cellular level (where the survival of a single organism is concerned) and at the organism level (where the survival of the species is concerned).

Thus von Neumann, drawing inspiration from natural systems, attempted to develop an approach to the realization of self-replicating computing machines (which he called *artificial automata*, as opposed to natural automata, that is, biological organisms). In order to achieve his goal, he imagined a series of five distinct models for self-reproduction ([44], pp. 91–99):

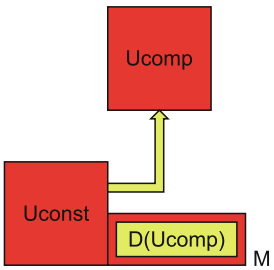
1. The *kinematic* model, introduced by von Neumann on the occasion of a series of five lectures given at the University of Illinois in December 1949, is the most general. It involves structural elements such as sensors, muscle-like components, joining and cutting tools, along with logic (switch) and memory elements. Concerning, as it does, physical as well as electronic components, its goal was to define the bases of self-replication, but was not designed to be implemented.
2. In order to find an approach to self-replication more amenable to a rigorous mathematical treatment, von Neumann, following the suggestion of the mathematician S. Ulam, developed a *cellular* model. This model, based on the use of cellular automata as a framework for study, was probably the closest to an actual realization. Even if it was never completed, it was further refined by von Neumann's successors and was the basis for most further research on self-replication.
3. The *excitation-threshold-fatigue* model was based on the cellular model, but each cell of the automaton was replaced by a neuron-like element. Von Neumann never defined the details of the neuron, but through a careful analysis of his work, we can deduce that it would have borne a fairly close relationship to today's simplest artificial neural networks, with the addition of some features which would have both increased the resemblance to biological neurons and introduced the possibility of self-replication.
4. For the *continuous* model, von Neumann planned to use differential equations to describe the process of self-reproduction. Again, we are not aware of the details of this model, but we can assume that von Neumann planned to define systems of differential equations to describe the excitation, threshold and fatigue properties of a neuron. At the implementation level, this would probably correspond to a transition from purely digital to analog circuits.
5. The *probabilistic* model is the least well-defined of all the models. We know that von Neumann intended to introduce a kind of automaton where the transitions between states were probabilistic rather than deterministic. Such an approach would allow the introduction of mechanisms such as mutation and thus of the phenomenon of evolution in artificial automata. Once again, we cannot be sure of how von Neumann would have realized such systems, but we can assume they would have exploited some of the same tools used today by genetic algorithms.

Of all these models, the only one von Neumann developed in some detail was the cellular model. Since it was the basis for the work of his successors, it deserves to be examined more closely.

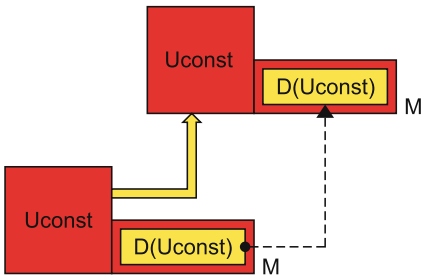
Von Neumann's Cellular Model

In von Neumann's work, self-reproduction is always presented as a special case of universal construction, that is, the capability of building any machine given its description (Fig. 1). This approach was maintained in the design of his cellular automaton, which is therefore much more than a self-replicating machine. The complexity of its purpose is reflected in the complexity of its structure, based on three separate components:

1. A *memory tape*, containing the description of the machine to be built, in the form of a one-dimensional string of elements. In the special case of self-reproduction, the memory contains a description of the universal constructor itself (Fig. 2). It is interesting to note that the memory of von Neumann's automaton bears a strong resemblance to the biological genome. This resemblance is even more remarkable when considering that the structure of the genome was not discovered until after the death of von Neumann.



Self-Replication and Cellular Automata, Figure 1
Von Neumann's universal constructor `Uconst` can build a specimen of any machine (e. g., a universal Turing machine `Ucomp`) given its description `D(Ucomp)`



Self-Replication and Cellular Automata, Figure 2
Given its own description `D(Uconst)`, von Neumann's universal constructor is capable of self-replication

2. The *constructor* itself, a very complex machine capable of reading the memory tape and interpreting its contents.
3. A *constructing arm*, directed by the constructor, used to build the offspring (i. e., the machine described in the memory tape). The arm moves across the space and sets the state of the elements of the offspring to the appropriate value.

The implementation as a cellular automaton is no less complex. Each element has 29 possible states, and thus, since the next state of an element depends on its current state and that of its four cardinal neighbors, $29^5 = 20,511,149$ transition rules are required to exhaustively define its behavior.

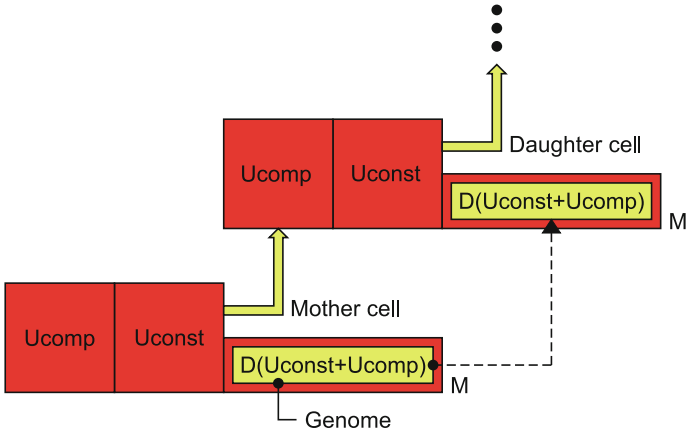
If we consider that the size of von Neumann's constructor is of the order of several hundred thousand elements, we can easily understand why a hardware realization of such a machine is not really feasible. In fact, as part of our research, we did realize a hardware implementation of a set of elements of von Neumann's automaton [3,35]. By carefully designing the hardware structure of each element, we were able to considerably reduce the amount of memory required to host the transition rules. Nevertheless, our system remains a demonstration unit, as it consists of a few elements only, barely enough to illustrate the behavior of a tiny subset of the entire machine.

It is also worth mentioning that von Neumann went one step further in the design of his universal constructor. If we consider the universal constructor from a biological viewpoint, we can associate the memory tape with the genome, and thus the entire constructor with a single cell (which would imply a parallel between the automaton's elements and molecules). However, the constructor, as we have described it so far, has no functionality outside of self-reproduction. Von Neumann recognized that a self-replicating machine would require some sort of functionality to be interesting from an engineering point of view, and postulated the presence of a *universal computer* (in practice, a universal Turing machine, an automaton capable of performing any computation) alongside the universal constructor (Fig. 3). Von Neumann's constructor can thus be regarded as a unicellular organism, containing a genome stored in the form of a memory tape, read and interpreted by the universal constructor (the mother cell) both to determine its operation and to direct the construction of a complete copy of itself (the daughter cell).

Von Neumann's Successors

The extreme size of von Neumann's universal constructor has so far prevented any kind of physical implementation (apart from the small demonstration unit we mentioned). But further, even the simulation of a cellular automaton of such complexity was far beyond the capability of early computer systems. Today, such a simulation is starting to be conceivable. Umberto Pesavento, a young Italian high school student, developed a simulation of von Neumann's entire universal constructor [27]. The computing power available did not allow him to simulate either the entire self-replication process (the length of the memory tape needed to describe the automaton would have required too large an array) or the Turing machine necessary to implement the universal computer, but he was able to demonstrate the full functionality of the constructor.

Considering the rapid advances in computing power of modern computer systems, we can assume that a complete simulation could be envisaged with today's technology. In fact, an effort is currently under way [4] to implement a complete specimen of the constructor. To achieve this goal, Buckley is revisiting and analyzing in detail the operation of the constructor. To give an idea of the scope of this work, Buckley's results indicate that the interpretercopier (without the tape) is bounded by a region of $751 \times 1048 = 787.048$ cells.



Self-Replication and Cellular Automata, Figure 3
By extension, von Neumann's universal constructor can include a universal computer and still be capable of self-replication

The impossibility of achieving a physical realization did not however deter some researchers from trying to continue and improve von Neumann's work [2,15,22]. Arthur Burks, for example, in addition to editing von Neumann's work on self-replication [5,44], also made several corrections and advances in the implementation of the cellular model. Codd [9], by altering the states and the transition rules, managed to simplify the constructor by a considerable degree. Vitanyi [43] studied the possibility of introducing sexual reproduction in von Neumann's approach. However, without in any way lessening these contributions, we can say that no major theoretical advance in the research on self-reproducing automata occurred until C. Langton, in 1984, opened a second stage in this field of research.

Self-Replication for Artificial Life

While the *implementation* of von Neumann's universal constructor faced insurmountable (at the time) technological hurdles, the same could not be said of the *theoretical* contribution that his approach represented as an attempt to study a biologically-inspired process within the world of computing systems.

In this context, the main drawback of von Neumann's work lay in the inability to achieve self-replication without resorting to an extremely complex simulation of a complete machine. Von Neumann's Universal Constructor was so complex because it tried to implement self-reproduction as a particular case of construction universality, i. e. the capability of constructing any other automaton, given its description. C. Langton approached the problem somewhat differently, by attempting to define the simplest cellular automaton, commonly known as *Langton's loop* [14], capable exclusively of self-reproduction.

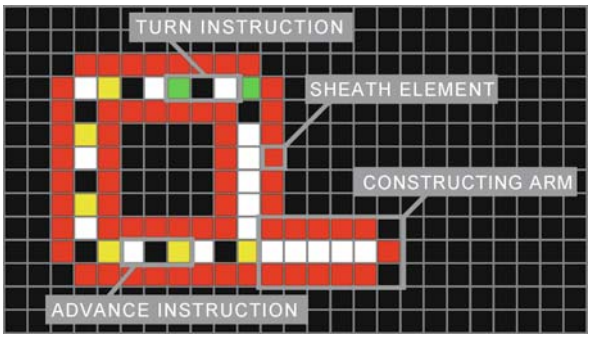
Langton's loop had a major impact on research in self-replication by introducing a new way to think about this process in more "abstract" terms as a study of the application of biologically-inspired mechanisms to computing, exemplifying the field known as *Artificial Life*. In this context, rather than the replicating machine, it is the process of self-replication itself that becomes the object of study.

This novel approach generated research that can be considered fundamentally different from that of von Neumann and started discussion on topics such as the analogy with cellular division and with the reproduction of individuals in a population (e.g, in [17], Section V.B), the difference between trivial and non-trivial self-replication in cellular automata (e. g., in automata such as those described in [16]), or the connections between evolution and self-replication (e.g, in the work of Sayama [17]).

Langton's Loop

As a consequence of his approach, Langton's Loop is orders of magnitude simpler than von Neumann's constructor. In fact, it is loosely based on one of the simplest organs (an organ in this context can be seen as a self-supporting structure capable of a single sub-task) in Codd's automaton: the *periodic emitter* (itself derived from von Neumann's periodic pulser), a relatively simple structure capable of generating a repeating string of a given sequence of pulses.

Langton's loop (Fig. 4) is named after the dynamic storage of data inside a square sheath (red in the figure). The data is stored as a sequence of instructions for directing the constructing arm, coded in the form of a set of three states. The data turns counterclockwise in permanence within the sheath, creating a loop.



Self-Replication and Cellular Automata, Figure 4
The initial configuration of Langton's Loop (iteration 0)

The two instructions in Langton's loop are extremely simple. One instruction (uniquely identified by the yellow element in the figure) tells the arm to advance by one position (Fig. 5), while the other (green in the figure) directs the arm to turn 90 degrees to the left (Fig. 6). Obviously, after three such turns, the arm has looped back on itself (Fig. 7), at which stage a messenger (the pink element) starts the process of severing the connection between the parent and the offspring, thus concluding the replication process.

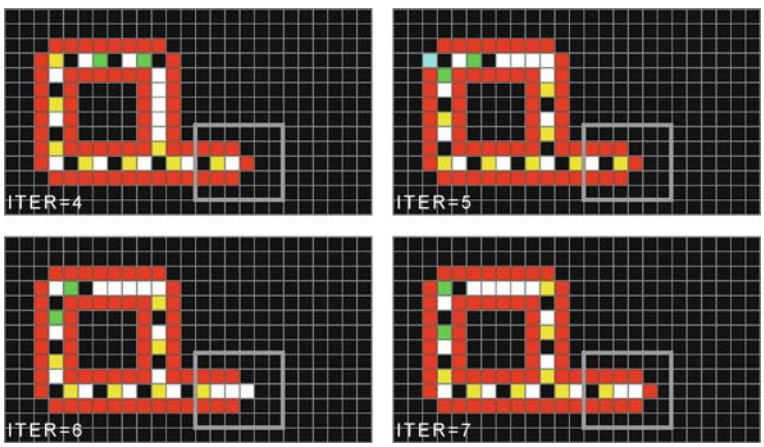
Once the copy is over, the parent loop proceeds to construct a second copy of itself in a different direction (to the north in this example), while the offspring itself starts to reproduce (to the east in this example). The sequential nature of the self-reproduction process generates a spiraling pattern in the propagation of the loop through space (Fig. 8): as each loop tries to reproduce in the four cardinal directions, it finds the place already occupied either by its parent or by the offspring of another loop, in which case it dies (the data within the loop is destroyed).

Langton's loop uses 8 states for each of the 86 non-quiescent cells making up its initial configuration, a 5-cell neighborhood, and a few hundred transition rules (the exact number depends on whether default rules are used and whether symmetric rules are included in the count). Further simplifications to Langton's automaton were introduced by Byl [6], who eliminated the internal sheath and reduced the number of states per cell, the number of transition rules, and the number of non-quiescent cells in the initial configuration. Reggia et al. [29] managed to remove also the external sheath, thus designing the smallest self-replicating loop known to date. Given their modest complexity, at least relative to von Neumann's automaton, all of the mentioned automata have been thoroughly simulated.

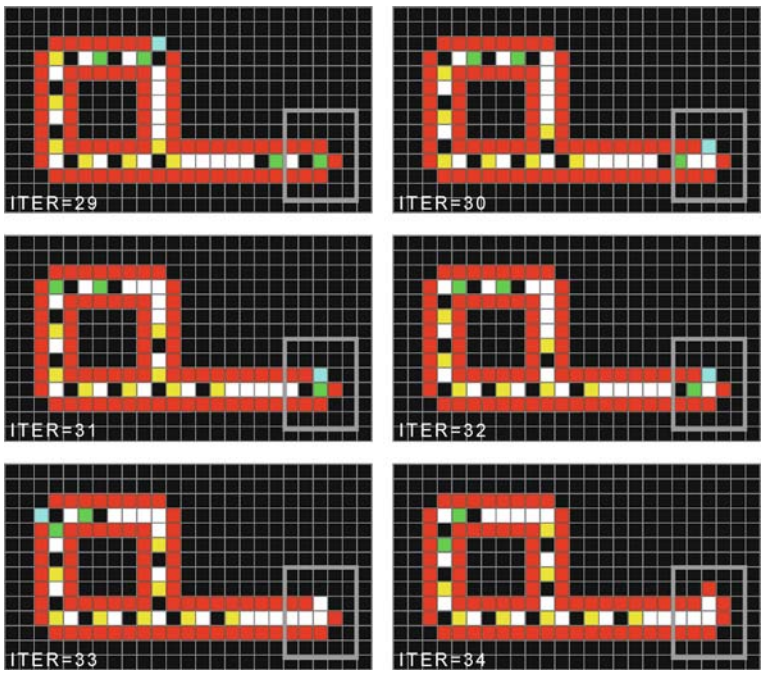
Langton's loop has been used as the basis for several approaches, mostly aimed at studying the properties of self-replication within a cellular system in the context of artificial life. Sayama [32] introduced structural dissolution (whereby a loop can destroy itself, in addition to replicating) to obtain colonies of loops that are dynamically stable and exhibit a potentially evolvable behavior. Nehaniv [21] extended Langton's approach to asynchronous cellular automata, while Sipper [34] developed a self-replicating loop in a non-uniform CA (i. e., a CA where the transition rules are not necessarily identical in all cells).

Perrier's Loop

In the context of applying self-reproduction to the replication of computing machines, and hence return to von Neumann's original goals, the main weakness of Langton's loop resides in the absence of any functionality beyond self-reproduction itself. To overcome this limitation, Per-



Self-Replication and Cellular Automata, Figure 5
The constructing arm advances by one space



Self-Replication and Cellular Automata, Figure 6
The constructing arm turns 90 degrees to the left

rier and Zahnd developed a relatively complex automaton (Fig. 9) in which a two-tape Turing machine was appended to Langton's loop [26].

This automaton exploits Langton's loop as a sort of "carrier" (Fig. 10): the first operation of Perrier's loop is to allow Langton's loop to build a copy of itself (iteration 128: note that the copy is limited to one dimension, since the second dimension is taken up by the Turing machine). The main function of the offspring is to determine the location of the copy of the Turing machine (iteration 134). Once the new loop is ready, a "messenger" runs back to the parent loop and starts to duplicate the Turing machine (iterations 158 and 194), a process completely disjoint from the operation of the loop. When the copy is finished (iteration 254), the same messenger activates the Turing machine in the parent loop (the machine had to be inert during the replication process in order to obtain a perfect copy). The process is then repeated in each offspring until the space is filled (iteration 720: as the automaton exploits Langton's loop for replication, meeting the boundary of the array causes the last machine to crash).

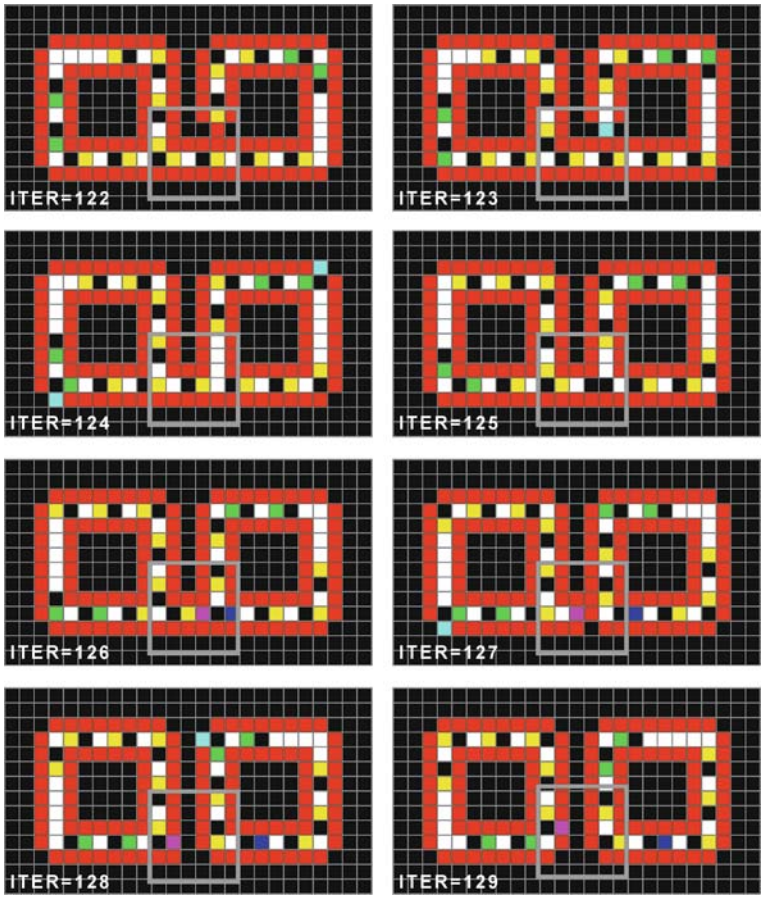
Perrier's automaton implements a self-replicating Turing machine, a powerful construct which is unfortunately handicapped by its complexity: in order to implement a Turing machine, the automaton requires a very considerable number of additional states (more than 60), as well as an important number of additional transition rules. This kind of complexity, while still relatively minor compared to von Neumann's universal constructor, is nevertheless too important to be really considered for an actual implementation.

Tempesti's Loop

Always in the context of achieving self-reproduction of computing machines, and beside the lack of functionality mentioned in Subsect. "Perrier's Loop", another problem of Langton's loop is that it is not well adapted to finite CA arrays. Its self-reproduction mechanism assumes that there is enough space for a copy of the loop, and the entire loop becomes nonfunctional otherwise (Fig. 8).

To overcome this limitation and move a step closer to the realization of self-replicating machines, we developed a self-replicating loop designed specifically to exist in a finite, but arbitrarily large, space, and at the same time capable, unlike Langton's loop, to have a functionality in addition to self-replication.

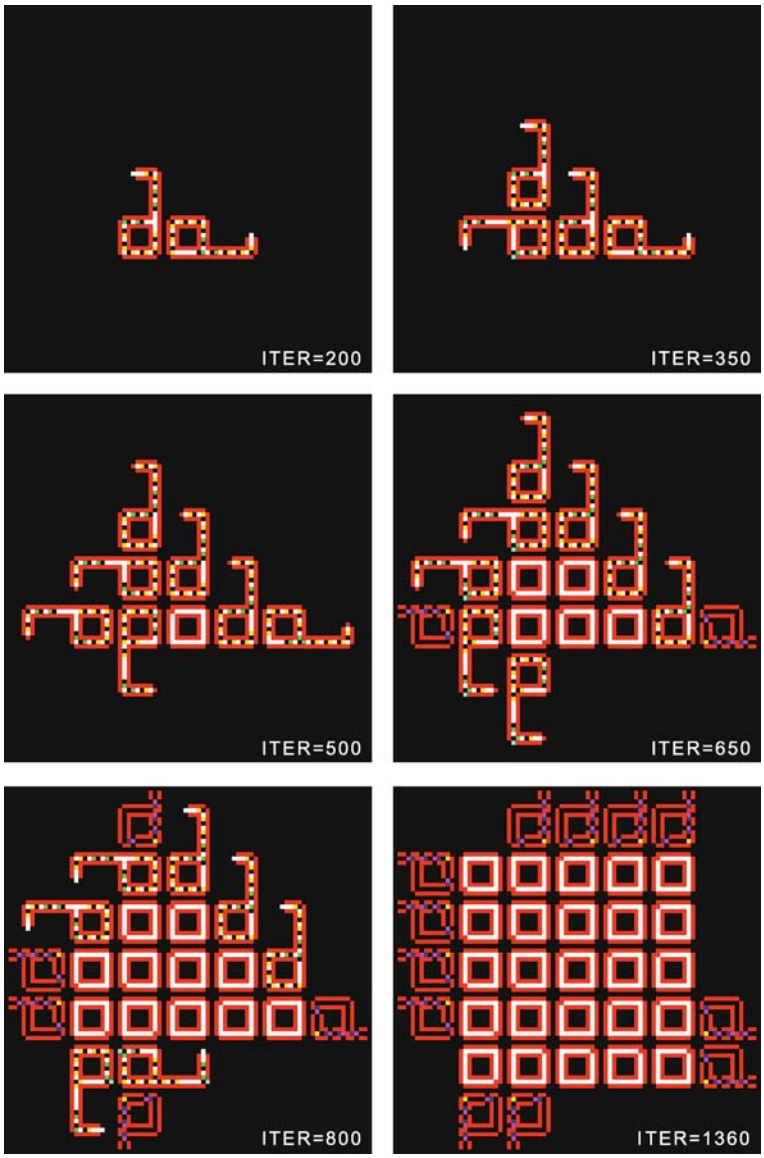
In designing our self-replicating automaton [39,40], we did maintain some of the more interesting features of Langton's loop. In particular, we preserved the structure based on a square loop to dynamically store information. Such storage is convenient in CA because of the locality of



Self-Replication and Cellular Automata, Figure 7
The copy is complete and the connection from parent to offspring is severed

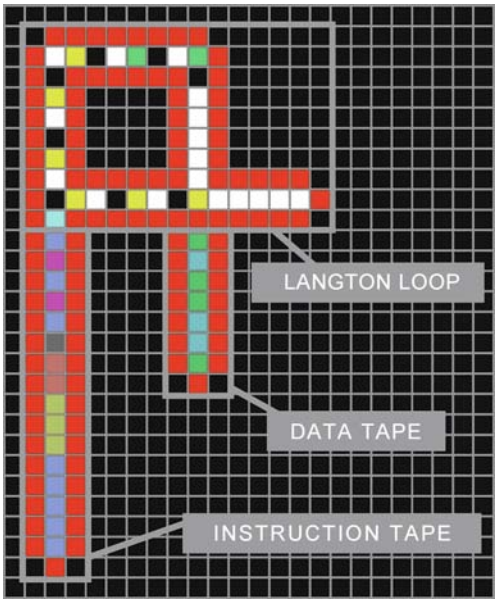
the rules. Also, we maintained the concept of constructing arm, in the tradition of von Neumann and his successors, even if we introduced considerable modifications to its structure and operation. While preserving some of the more interesting features of Langton's loop, we nevertheless introduced some basic structural alterations (Fig. 11):

- As in Byl's version of Langton's loop, we use only one sheath. In addition, four *gate elements* (in the same state as the sheath) at the four corners of the automaton enable or disable the replication process.
- We extend four constructing arms in the four cardinal directions at the same time, and thus create four copies of the original automaton in the four directions in parallel. When the arm meets an obstacle (either the boundary of the array or an existing copy of the loop), it simply retracts and puts the corresponding gate element in the closed position.
- The arm does not immediately construct the entire loop. Rather, it constructs a sheath of the same size as the original. Once the sheath is ready, the data circulating in the loop is duplicated and the copy is sent along the constructing arm to wrap around the new sheath. When the new loop is completed, the constructing arm retracts and closes the gate.
- As a consequence, we use only four of the elements circulating in the loop to control the self-replication process. The others can be used as a "program", i. e., a set of states with their own transition rules which will then be applied alongside the self-reproduction to execute some function.
- Unlike Langton's loop, our loop does not "die" once duplication is achieved, as the circulating data remains untouched by the self-reproduction process. This feature is a requirement for implementing functions which work after the copy has finished.



Self-Replication and Cellular Automata, Figure 8
Propagation pattern of Langton's loop

We use a 9-element neighborhood (the element itself plus its 8 neighbors) and the basic configuration of the loop (Fig. 11) requires five states plus at least one data state. State 0 (black) is the *quiescent* state: it represents the inactive background. State 1 (white) is the *sheath* state, that is the state of the elements making up the sheath and the four gates. State 2 (red) is the *activation* state or *control* state. The four gate elements are in state 2, as are the messengers which will be used to command the constructing arm and the tip of the constructing arm itself for the first phase of construction, after which the tip of the arm will switch to state 3 (light blue), the *construction* state. State 3 will construct the sheath that will host the offspring, signal the parent loop that the sheath is ready, and lead the duplicated data to the new loop. State 4 (green), the destruction state, will destroy the constructing arm once the copy is completed. In addition to these states, two additional *data* states (light and dark grey) represent the information stored in the loop. In this example, they are inactive, while the next section describes a loop where they are used to store an executable program.



Self-Replication and Cellular Automata, Figure 9
A two-tape Turing machine appended to Langton's loop (iteration 0)

The initial configuration is in the form of a square loop wrapped around a sheath. The size of the loop is variable, and for our example is set to 8×8. Once the iterations begin, the data starts turning counterclockwise around the loop. Nothing happens until the first control element (red) reaches a corner of the loop, where it checks the status of the gate. Since the gate is open, the control element splits into two identical elements: the first continues turning around the loop, while the second starts extending the arm (Fig. 12).

The arm advances automatically by one position every two iterations. Once the arm has started extending, each control element that arrives to a corner will again split and one of the copies will start running along the arm, advancing twice as fast. When the first of these messengers reaches the tip of the arm, the tip, which was until then in state 2, passes to state 3 and continues to advance at the same speed. This transformation tells the arm that it has reached the location of the offspring loop and to start constructing the new sheath. The next three messengers will force the tip of the arm to turn left, while the fourth causes the sheath to close upon itself (Fig. 13). It then runs back along the arm to signal to the original loop that the new sheath is ready. Once the return signal arrives at the corner of the original loop, it causes a copy of the data in the loop to run along the arm and wrap itself around the new sheath. Once the second copy has completed the loop (Fig. 14), it sends a destruction signal (green) back along the arm. The signal will destroy the arm until it reaches the corner of the original loop, where it closes the gate to avoid further copies.

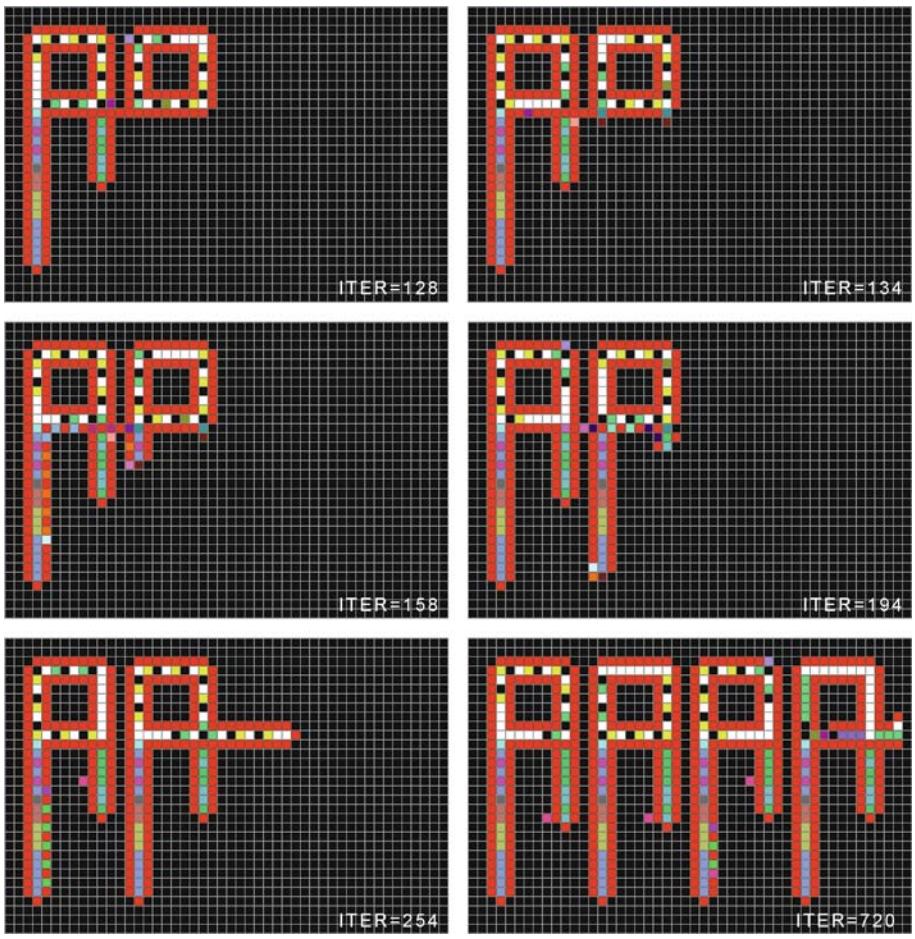
After 121 time periods the gates of the original automaton will be closed and it will enter an inactive state, with the understanding that it will be ready to reproduce itself again should the gates be opened. The main advantage of the new mechanism is that it becomes relatively simple to retract the arm if an obstacle (either the boundary of the array or another loop) is encountered, and therefore our loop is perfectly capable of operating in a finite space.

In Fig. 15, we illustrate an example of how the data states can be used to carry out operations alongside self-reproduction. The operation in question is the construction of three letters, LSL (the acronym of Logic Systems Laboratory, where the research was made), in the empty space inside the loop. Obviously this is not a very useful operation from a computational point of view, but it is a far from trivial construction task which should suffice to demonstrate the capabilities of the automaton.

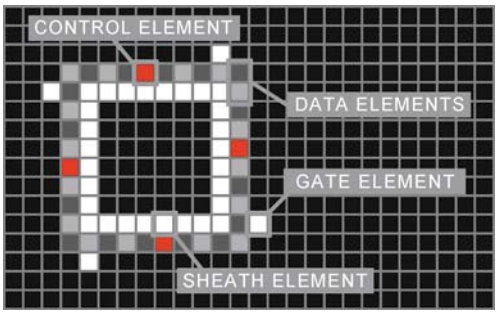
As should be obvious, while our loop owes to von Neumann the concept of constructing arm and to Langton (and/or Codd) the basic loop structure, it is in fact a very different automaton, endowed with some of the properties of both. We have seen that von Neumann's automaton is extremely complex, while Langton's loop is very simple. The complexity of our automaton is more difficult to estimate, as it depends on the data circulating in the loop. The number of non-quiescent elements making up the initial configuration depends directly on the size of the circulating program. The more complex (i. e. the longer) the program, the larger the automaton (it should be noted, however, that the complexity of the self-reproduction process does not depend on the size of the loop). The number of transition rules is obviously a function of the number of data states: in the basic configuration (i. e., one data state), the automaton needs 692 rules (173 rules rotated in the four directions), assuming that, by default, all elements remain in the same state. The complexity of the basic configuration is therefore in the same order as that of Langton's and Byl's loops, with the proviso that it is likely to increase drastically if the data in the loop is used to implement a complex function.

Other Approaches to Self-Replication

Von Neumann's and Langton's structures represent the main landmarks in the study of self-replication in com-



Self-Replication and Cellular Automata, Figure 10
Self-replication of the Turing machine

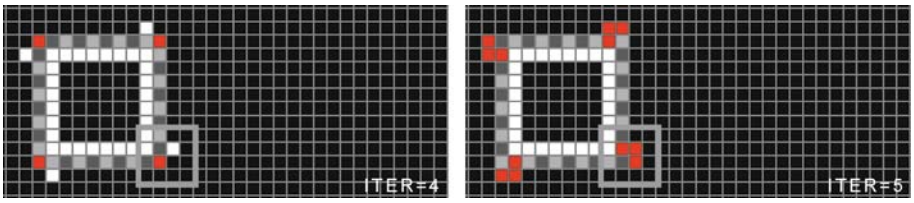


Self-Replication and Cellular Automata, Figure 11
The initial configuration of our loop (iteration 0)

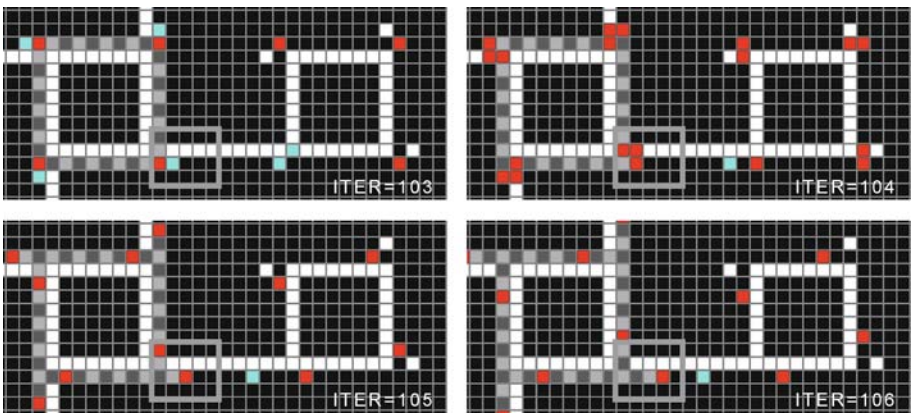
puting machines. It can safely be said that all other approaches refer, directly or indirectly, to these two systems. However, there exist some approaches to self-replication that cannot be easily reduced to simple variations on one of these two themes, either because they specifically take into consideration some issues that are not addressed by Langton and von Neumann, or because they occur in environments that are considerably different from the two original approaches.

In this section, we will deal in depth with one example, the *Tom Thumb algorithm* that, while referring back to von Neumann insofar as its goal is the implementation of self-replicating logic circuits, is specifically designed to operate efficiently in the kind of digital devices that are available today. The algorithm approaches cellular automata from a slightly unconventional angle [37], with the objective of a hardware realization of self-replication within a programmable logic device, or FPGA [41].

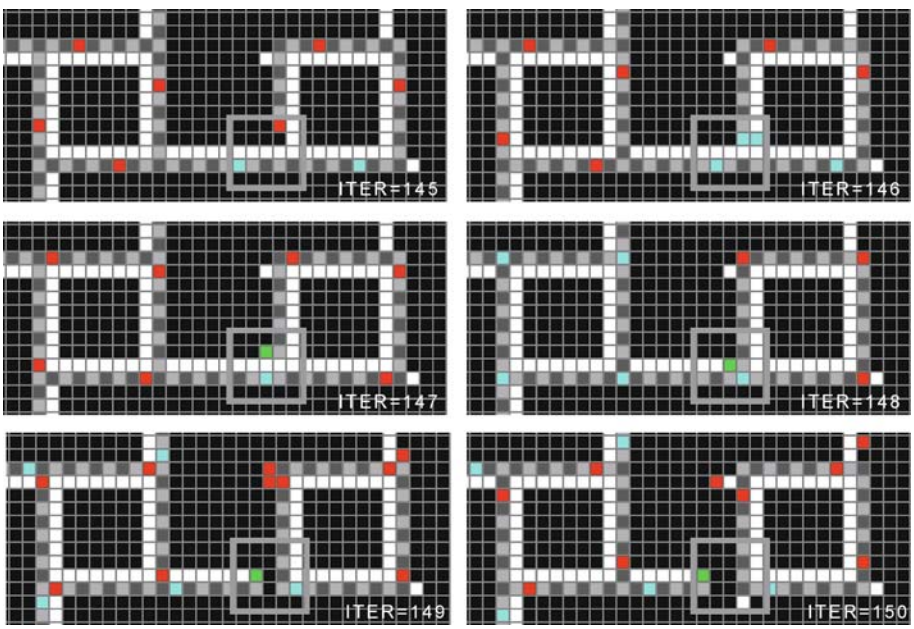
In the second part of the section, we will look at a set of approaches to self-replication that represent notable extensions to the approaches of von Neumann and Langton, because of different mechanisms (self-inspection), operat-



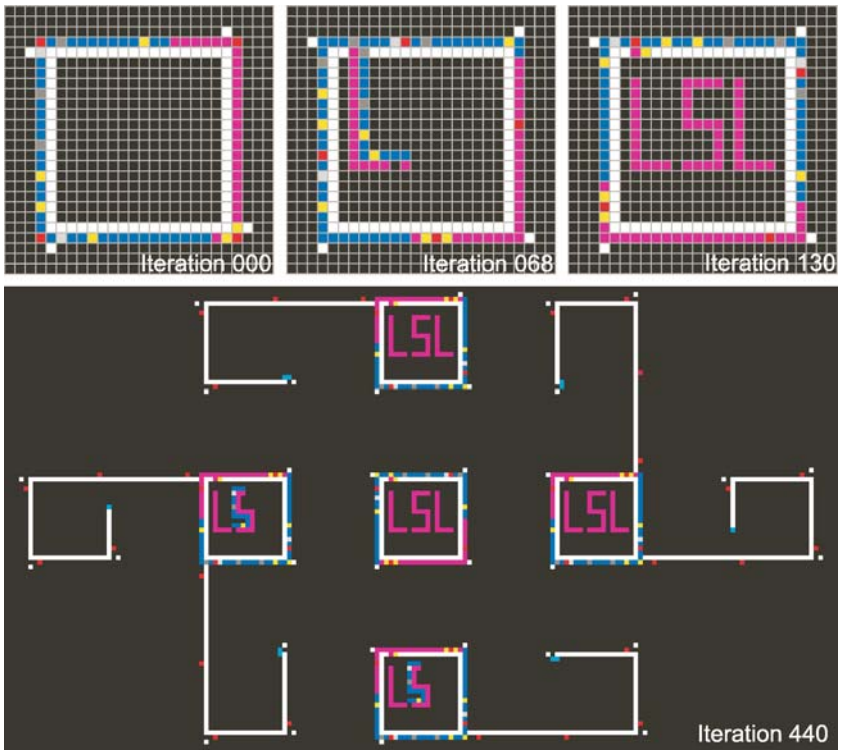
Self-Replication and Cellular Automata, Figure 12
The constructing arm begins to extend



Self-Replication and Cellular Automata, Figure 13
The new sheath has been fully constructed and a copy of the data is sent from the parent to the offspring



Self-Replication and Cellular Automata, Figure 14
The copy is complete and the constructing arm retracts



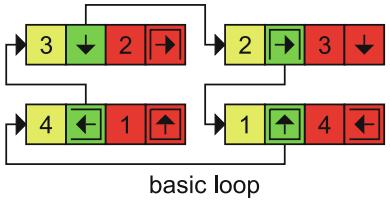
Self-Replication and Cellular Automata, Figure 15
The LSL automaton at different iterations

ing milieus (three-dimensional or self-timed CA), or design rules (evolutionary approaches).

Self-Replication in Hardware: The Tom Thumb Algorithm

In past years, we have devoted considerable effort to research on self-replication, studying this process from the point of view of the design of high-complexity multi-processor systems (with particular attention to next-generation technologies such as nanoelectronics). When considering self-replication in this context, Langton's loop and its successors share several weaknesses. Notably, besides the lack of functionality of Langton's loop (remedied only partially by its successors), which severely limits its usefulness for circuit design, each of these automata is characterized by a very loose utilization of the resources at its disposal: the majority of the elements in the cellular array remain in the quiescent state throughout the entire self-replication process.

A new loop was then developed specifically to address these very practical issues. In fact, the system is targeted to the implementation of self-replication within the context of digital circuits realized with programmable logic devices

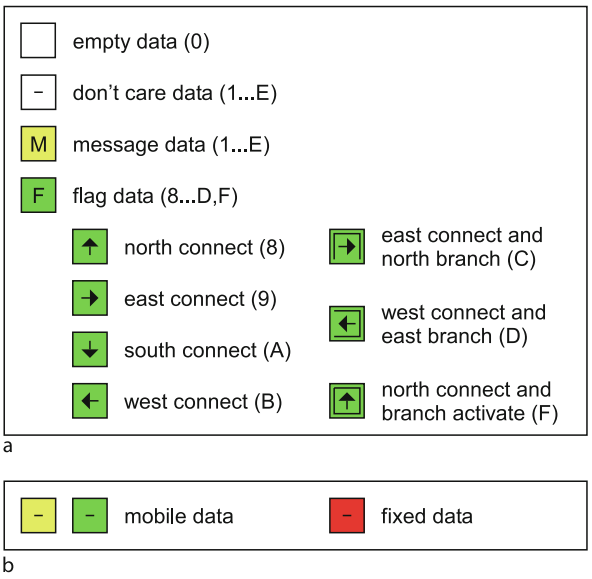


Self-Replication and Cellular Automata, Figure 16
Tom Thumb algorithm: basic loop

(the states of the cellular automaton can then be seen as the configuration bits of the elements of the device). The new loop is based on an original algorithm, the so-called *Tom Thumb algorithm* [18,19].

The minimal loop compatible with this algorithm is made up of four cells, organized as a square of two rows by two columns (Fig. 16). Each cell is able to store in its four memory positions four hexadecimal characters of an artificial *genome* (defined as the information required for the construction of the loop). The whole loop thus embeds 16 such characters.

The original genome for the minimal loop is organized as another loop, the *basic loop*, of eight hexadecimal characters, i. e. half the number of characters in the minimal



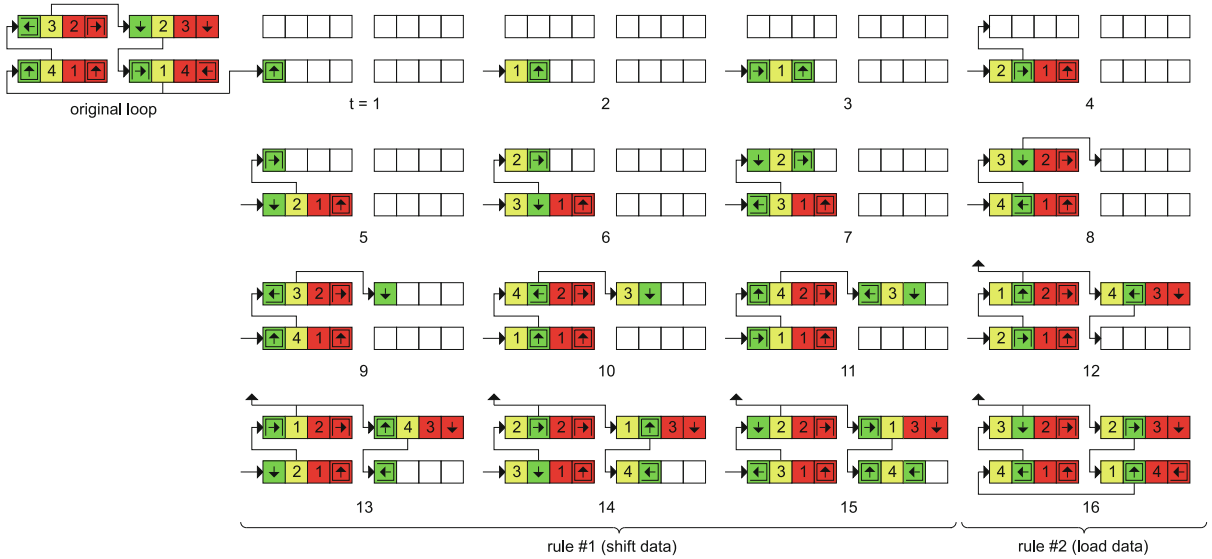
Self-Replication and Cellular Automata, Figure 17
a Graphical and hexadecimal representations of the 15 characters forming the alphabet of the artificial genome. b Graphical representation of the status of each character

loop, moving counterclockwise by one character at each time step.

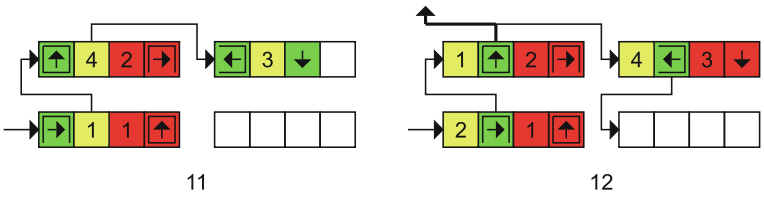
The 15 hexadecimal characters that compose the alphabet of the artificial genome are detailed in Fig. 17a. They are either *empty data* (0), *message data* ($M = 1 \ldots E$), or *flag data* ($F = 8 \ldots D, F$). Message data will be used to configure our final artificial organism, while flag data are indispensable for constructing the skeleton of the loop. Furthermore, each character is given a status and will eventually be *mobile data*, moving indefinitely around the loop, or *fixed data*, definitely trapped in a memory position of a cell (Fig. 17b). It is important to note that, while in this simple example the message data can take value from 1 to E, the Tom Thumb algorithm is perfectly scalable in this respect, that is, the size of the message data can be increased at will, while the flag data remain constant. This is a crucial observation in view of the exploitation of this algorithm in a programmable logic device, where the message data (the configuration data for the programmable elements of the circuit) are usually much more complex.

At each time step ($t = 1, 2, \ldots$), a character of the *original loop* is sent to the lower leftmost cell (Fig. 18). The construction of the loop, i. e., storing the fixed data and defining the paths for mobile data, depends on two rules:

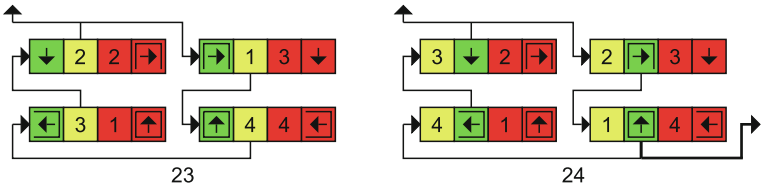
- If the four, three, or two rightmost memory positions of the cell are empty (blank squares), the characters are shifted by one position to the right (rule #1: shift data).
- If the rightmost memory position is empty, the characters are shifted by one position to the right (rule #2: load data). In this situation, the two rightmost characters are trapped in the cell (fixed data), and a new connection is established from the second leftmost position toward the northern, eastern, southern or western cell, depending on the fixed flag information (in Fig. 18, at



Self-Replication and Cellular Automata, Figure 18
Construction of a first specimen of the loop



Self-Replication and Cellular Automata, Figure 19
Creation of a new daughter loop to the north (rule #3)



Self-Replication and Cellular Automata, Figure 20
Creation of a new daughter loop to the east (rule #4)

time $t = 4$, the fixed flag F = F determines a northern connection).

At time $t = 16$, 16 characters, i. e., twice the contents of the basic loop, have been stored in the 16 memory positions of the loop (Fig. 18). Eight characters are fixed data, forming the *phenotype* of the final loop, and the eight remaining ones are mobile data, composing a copy of the original genome, i. e., the *genotype*. Both *interpretation* (the construction of the cell) and *copying* (the duplication of the genetic information) have been therefore achieved.

The fixed data trapped in the rightmost memory positions of each cell remind us of the pebbles left by Tom Thumb for memorizing his way in the famous children's story, an analogy that gives our algorithm its name.

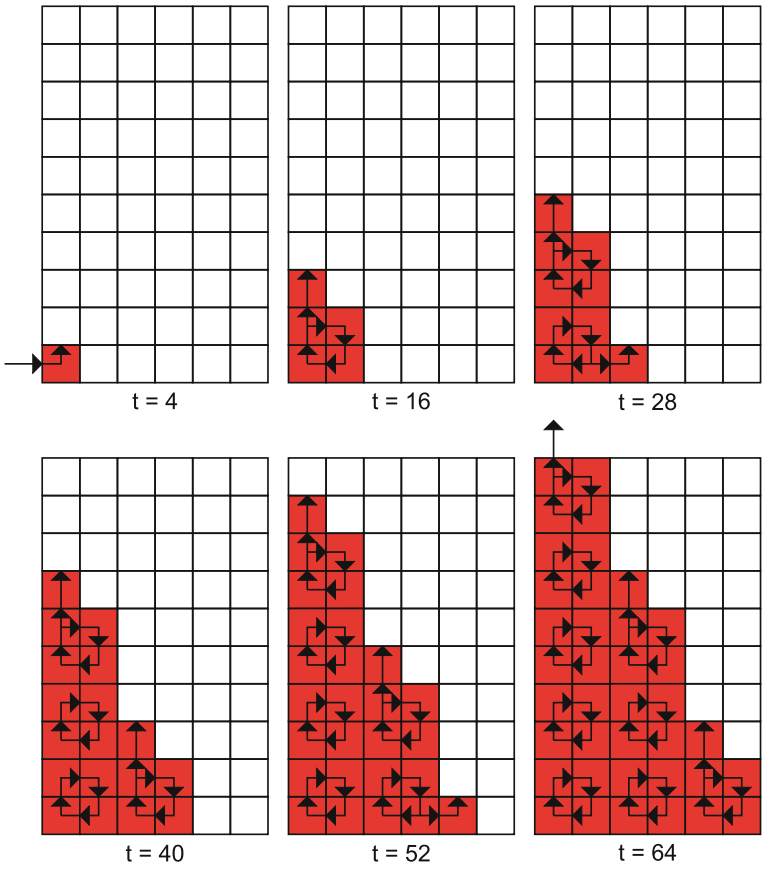
In order to grow loops in both horizontal and vertical directions, the mother loop should be able to trigger the construction of two daughter loops, northward and eastward. Two new rules are then necessary:

- At time $t = 11$ (Fig. 19), we observe a pattern of characters which is able to start the construction of the northward daughter loop; the upper leftmost cell is characterized by two specific flags, i. e., a fixed flag in the rightmost position, indicating a north branch (F = C) and the branch activation flag (F = F), in the leftmost position (rule #3: daughter loop to the north). The new path to the northward daughter loop will start from the second leftmost memory position ($t = 12$).
- At time $t = 23$ (Fig. 20), another particular pattern of characters starts the construction of the eastward daughter loop; the lower rightmost cell is characterized by two specific flags, i. e., a fixed flag indicating an east branch (F = D), in the rightmost position, and the branch activation flag (F = F), in the leftmost position (rule #4: daughter loop to the east). The new path to the eastward daughter loop starts from the second leftmost memory position ($t = 24$).

When two or more paths are activated simultaneously, a clear priority should be established between the different paths. Three growth patterns were chosen (Fig. 21), leading to four more rules:

- For loops in the lower row a collision occurs between the closing path, inside the loop, and the path entering the lower leftmost cell. The westward path has priority over the eastward path (rule #5).
- With the exception of the bottom loop, the inner path (i. e. the westward path) has priority over the northward path (rule #6) for the loops in the leftmost column.
- For all other loops, two types of collisions may occur: between the northward and eastward paths (2-signal collision) or between these two paths and a third one, the closing path (3-signal collision). In this case, the northward path will have priority over the eastward path (2-signal collision), and the westward path will have priority over the two other ones (3-signal collision)(rules #7 and #8).

We finally opted the following hierarchy: an east to west path has priority over a south to north path, which has priority over a west to east path.



Self-Replication and Cellular Automata, Figure 21
Growth of a colony of minimal loops represented at different time steps

The results of such a choice are as follows (Fig. 21): a closing loop has priority over all other outer paths, which makes the completed loop entirely independent of its neighbors, and the loops will grow bottom-up vertically. This choice is quite arbitrary and may be changed according to other specifications.

Unlike its predecessors, the Tom Thumb loop has been developed with a specific purpose beyond the theoretical study of self-replication. We believe that, in the not-so-distant future, circuits will reach densities such that conventional design techniques will become unwieldy. Should such an hypothesis be confirmed, self-replication could become an invaluable tool, allowing the engineer to design a single processing element, part of an extremely large array that would build itself through the duplication of the original element.

Current technology does not, of course, provide a level of complexity that would render this kind of process necessary. However, programmable logic devices (already among the densest circuits on the market) can be used as a first approximation of the kind of circuits that will become available in the future. Our loop is then targeted to the implementation of self-replication on this kind of device.

To this end, our loop introduces a number of features that are not present in any of the historical self-replicating loops we presented. Most notably, the structure of the loop (that is, the path used by the configuration data) is determined by the sequence of flags in the genome, implying that structures of almost any shape and size can be constructed and replicated using this algorithm, as long as the loop can be closed and that there is space for the daughter organisms. In practice, this implies that, if the Tom Thumb algorithm is used for the configuration logic of a programmable device, any of its configurations, and hence any digital circuit, can be made capable of self-replication.

In addition, particular care was given to develop a self-replication algorithm that is *efficient* (in the sense that it fully exploits the underlying medium, rather than leaving the vast majority of elements inert as past algorithms did), *scalable* (all the interactions between the elements are

purely local, implying that organisms of any size can be implemented), and amenable to a *systematic design process*. These features are important requirements for the design of highly-complex systems based on either silicon or molecular-scale components.

Different Techniques and Environments

Von Neumann's and Langton's automata share a common basic technique to obtain self-replication: the construction of the new machine is directed through the *interpretation* of a description, coded as a sequence of states. In the case of von Neumann, this description (which, in biological terms, is usually identified as the genome of the artificial organism) is stored within the memory tape, which is read and interpreted by the universal constructor to build a copy of the machine. In Langton's case, the description is stored in the mobile data that runs within the sheath of the loop.

This mechanism of interpretation, while standard in many approaches, is not however unique: some examples of self-replicating CA exploit a different mechanism, that of *self-inspection*. In these approaches, instead of reading and interpreting a description, the self-replicating automaton inspects itself and produces a copy of what it finds. While less general than the universal constructor (obviously, the machine can only build an exact copy of itself), the functionality of this approach is similar to that of Langton's loop. Indeed, the most representative example of self-inspection is that of a self-replicating loop [12]. A more recent example is a variation of the Tom Thumb algorithm, where self-inspection was used to self-replicate a small processor within a field-programmable gate array [30].

And while the Tom Thumb algorithm targets in priority silicon-based circuits, other approaches have tried to explore alternative environments that, in some way, might more closely resemble the kind of technologies that will be available in the future. An example is Morita and Imai's study of self-replication in the context of reversible cellular automata [20] (in a reversible CA, every configuration has at most one predecessor), inspired by reversible logic in digital circuits.

Similarly, Peper et al. [24,38] have developed self-replicating structures in Self-Timed Cellular Automata (STCA). This kind of automata do not rely on a global synchronization mechanism to update the states of the cells, but rather the state transitions only occur when triggered by transitions in neighboring cells. The basic assumption in this work is that STCA is a model that might more closely resemble molecular-scale nanoelectronic devices.

A final example in this context is the three-dimensional extension of self-replication, usually based on the assumption that silicon, with its rigidly two-dimensional structure, will one day be superseded by a technology that can exploit all three dimensions. In this context, Imai et al. [13] have extended their reversible approach, with the assumption that reversible logic is more amenable to an extension to three dimensions than conventional logic because of the greatly reduced power dissipation. Stauffer at al. [36] have also shown that the Data and Signal Cellular Automaton (DSCA) approach, designed to simplify the implementation of CA in digital hardware, can be extended to realize self-replication in three dimensions.

The study of self-replicating CA in the context of new technologies holds the promise of one day bringing a major contribution to computation. To determine how self-replication might be useful in this context, some attempts have been made at using self-replicating structures for computation. An example of this approach is the work of Chou and Reggia [8], who use self-replication as a mechanism to obtain massively parallel machines which can potentially be used to solve hard problems (the example used in the paper is the NP-complete problem of satisfiability).

An attempt was also made to perform computation using Tempesti's loops. In alternative to embedding a complex program, this kind of loops are used to perform computation by inter-loop communication. Using the *collision-based computing* paradigm, Petraglio et al. [28] have shown that it is possible to implement arithmetic operations by passing messages from one loop to another after building a network structure through self-replication. This approach, while valuable from a theoretical standpoint, shares however the same weakness of other loop-based computing approaches in that the poor utilization of resources makes a physical realization of such a system highly impractical.

Another problem to be solved for a practical implementation of self-replicating structures is their design: few approaches have attempted to define a precise methodology to define and create self-replicating structures. In this context, several researchers have attempted to use evolutionary techniques to find automatically self-replicating machines. In this context, the work of Sayama et al. has gone through several iterations [31,32,33] in an attempt to define loops that evolve through the self-replication process towards "fitter" individuals. Chou and Reggia [7], on the other hand, use evolution to find novel self-replicating structures within a CA, whereas Pan and Reggia [23] studied the conditions in which self-replicating structures might spontaneously emerge in a cellular space.

Future Directions

Historically, self-replication in cellular automata began as a paradigm to achieve fault tolerance in computing devices. In the following decades, much of the emphasis shifted towards a more "theoretical" approach where self-replication was considered as part of a more general investigation into the application of biologically-inspired mechanisms to computing. And while the latter approach remains an active research topic, the original paradigm was somewhat set aside because technology would not allow a practical implementation of self-replication in digital hardware.

More recently, however, advances in electronic devices (notably with the introduction of programmable devices, or FPGAs), together with emerging technological issues have rekindled interest in self-replication in a context similar to von Neumann's original work. In particular, the drastic device shrinking, low power supply levels, and increasing operating speeds, which accompany the technological evolution of silicon to deeper submicron levels, significantly reduce the noise margins and increase the soft-error rates [42]. In addition, the nascent field of nanoelectronics holds great promise for the future of computing devices, but introduces extremely high fault rates (e. g., [25]) and complex layout issues.

Thus, self-replication is currently attracting a considerable amount of attention for the same reasons that initially pushed von Neumann to investigate it as a possible solution to reliability and layout issues. Fault tolerance and self-organization are thus becoming the focal point of research in the field and the features of molecular-scale electronics seem to imply that self-replication at the device level will be an extremely useful paradigm in next-generation devices.

Bibliography

1. Asprey W (1992) John von Neumann and the Origins of Modern Computing. The MIT Press, Cambridge
2. Banks ER (1970) Universality in Cellular Automata. In: Proc. IEEE 11th Annual Symposium on Switching and Automata Theory, Santa Monica, CA, pp 194-215
3. Beuchat J-L, Haenni J-O (2000) Von Neumann's 29-State Cellular Automaton: A Hardware Implementation. IEEE Trans Education 43(3):300–308
4. Buckley WR, Mukherjee A (2005) Constructibility of Signal-Crossing Solutions in von Neumann 29-State Cellular Automata. Proc. 2005 Int. Conf. on Computational Science. LNCS, vol 3515. Springer, Berlin, pp 395–403
5. Burks A (ed) (1970) Essays on Cellular Automata. University of Illinois Press, Urbana
6. Byl J (1989) Self-Reproduction in Small Cellular Automata. Physica D 34:295–299
7. Chou H-H, Reggia JA (1997) Emergence of self-replicating structures in a cellular automata space. Physica D 110(3-4):252–276
8. Chou H-H, Reggia JA (1998) Problem solving during artificial selection of self-replicating loops. Physica D 115(3-4):293–312
9. Codd EF (1968) Cellular Automata. Academic Press, New York
10. Freitas RA Jr, Gilbreath WP (eds) (1980) Advanced Automation for Space Missions. In: Proc. 1980 NASA/ASEE Summer Study, Scientific and Technical Information Branch (available from U.S. G.P.O.), Washington, DC
11. Freitas RA Jr, Merkle RC (2004) Kinematic Self-Replicating Machines. Landes Bioscience, Georgetown
12. Ibanez J, Anabitarte D, Azpeitia I, Barrera O, Barrutieta A, Blanco H, Echarte F (1995) Self-inspection based reproduction in cellular automata. In: Proc. 3rd European Conf. on Artificial Life (ECAL95). LNCS, vol 929. Springer, Berlin, pp 564–576
13. Imai K, Hori T, Morita K (2002) Self-reproduction in three-dimensional reversible cellular space. Artif Life 8(2):155–174
14. Langton CG (1984) Self-Reproduction in Cellular Automata. Physica D 10:135–144
15. Lee C (1968) Synthesis of a Cellular Computer. In: Applied Automata Theory. Academic Press, London, pp 217–234
16. Lohn JD, Reggia JA (1997) Automatic Discovery of Self-Replicating Structures in Cellular Automata. IEEE Trans Evol Comput 1(3):165–178
17. Mange D, Sipper M, Stauffer A, Tempesti G (2000) Towards Robust Integrated Circuits: The Embryonics Approach. Proc IEEE 88(4):516–541
18. Mange D, Stauffer A, Petraglio E, Tempesti G (2004) Self-replicating loop with universal construction. Physica D 191:178–192
19. Mange D, Stauffer A, Peparolo L, Tempesti G (2004) A Macroscopic View of Self-Replication. Proc IEEE 92(12):1929–1945
20. Morita K, Imai K (1996) Self-reproduction in a reversible cellular space. Theor Comput Sci 168:337–366
21. Nehaniv CL (2002) Self-Reproduction in Asynchronous Cellular Automata. In: Proc. 2002 NASA/DoD Conf. on Evolvable Hardware (EH02). IEEE Computer Society, Washington, DC, pp 201–209
22. Nourai F, Kashef RS (1975) A Universal Four-State Cellular Computer. IEEE Trans Comput 24(8):766–776
23. Pan Z, Reggia J (2005) Evolutionary Discovery of Arbitrary Self-replicating Structures. In: Proc. 5th Int. Conf. on Computational Science (ICCS 2005) - Part II. LNCS, vol 3515. Springer, Berlin, pp 404–411
24. Peper F, Isokawa T, Kouda N, Matsui N (2002) Self-Timed Cellular Automata and their computational ability. Future Gener Comput Syst 18(7):893–904
25. Peper F, Lee J, Abo F, Isokawa T, Adaki S, Matsui N, Mashiko S (2004) Fault-Tolerance in Nanocomputers: a Cellular Array Approach. IEEE Trans Nanotechnol 3(1):187–201
26. Perrier J-Y, Sipper M, Zahnd J (1996) Toward a Viable, Self-Reproducing Universal Computer. Physica D 97:335–352
27. Pesavento U (1995) An Implementation of von Neumann's Self-Reproducing Machine. Artif Life 2(4):337–354
28. Petraglio E, Tempesti G, Henry J-M (2002) Arithmetic Operations with Self-Replicating Loops. In: Adamatsky A (ed) Collision-Based Computing. Springer-Verlag, London, pp 469-490
29. Reggia JA, Armentrout SA, Chou H-H, Peng Y (1993) Simple Systems That Exhibit Self-Directed Replication. Science 259:1282–1287

30. Rossier J, Thoma Y, Mudry PA, Tempesti G (2006) MOVE Processors that Self-Replicate and Differentiate. In: Proc. 2nd Int. Workshop on Biologically-Inspired Approaches to Advanced Information Technology (Bio-ADIT06). LNCS, vol 3853. Springer, Berlin, pp 328–343
31. Salzberg C, Antony A, Sayama H (2004) Evolutionary dynamics of cellular automata-based self-replicators in hostile environments. BioSystems 78:119–134
32. Sayama H (1998) Introduction of Structural Dissolution into Langton's Self-Reproducing Loop. In: Artificial Life VI, Proc. 6th Int. Conf. on Artificial Life, MIT Press, Cambridge, pp 114–122
33. Sayama H (2000) Self-replicating worms that increase structural complexity through gene transmission. In: Artificial Life VII: Proc. 7th Int. Conf. on Artificial Life, MIT Press, Cambridge, pp 21–30
34. Sipper M (1995) Studying artificial life using a simple, general cellular model. Artif Life 2(1):1–35
35. Sipper M, Mange D, Stauffer A (1997) Ontogenetic Hardware. BioSystems 44:193–207
36. Stauffer A, Mange D, Petraglio E, Vannel F (2004) DSCA Implementation of 3D Self-Replicating Structures. In: Proc. 6th Int. Conf. on Cellular Automata for Research and Industry (ACRI2004). LNCS, vol 3305. Springer, Berlin, pp 698–708
37. Stauffer A, Sipper M (2004) The Data-and-Signals Cellular Automaton and Its Application to Growing Structures. Artif Life 10(4):463–477
38. Takada Y, Isokawa T, Peper F, Matsui N (2006) Universal Construction and Self-Reproduction on Self-Timed Cellular Automata. Int J Mod Phys C 17(7):985–1007
39. Tempesti G (1995) A New Self-Reproducing Cellular Automaton Capable of Construction and Computation. In: Proc. 3rd European Conf. on Artificial Life. LNAI, vol 929. Springer, Berlin, pp 555–563
40. Tempesti G (1998) A Self-Repairing Multiplexer-Based FPGA Inspired by Biological Processes. Ph.D. Thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL)
41. Trimberger S (ed) (1994) Field-Programmable Gate Array Technology. Kluwer, Boston
42. Various (1999) A D&T Roundtable: Online Test. IEEE Design Test Comput 16(1):80–86
43. Vitanyi PMB (1973) Sexually reproducing cellular automata. Math Biosci 18:23–54
44. Burks AW (ed) von Neumann J (1966) The Theory of Self-Reproducing Automata. University of Illinois Press, Urbana

Semantic Web

Wendy Hall, Kieron O'Hara
Intelligence, Agents, Multimedia Group, School of Electronics and Computer Science, University of Southampton, Southampton, United Kingdom

Article Outline

Glossary
Definition of the Subject
Introduction
Linking Data
The Layered Model of the Semantic Web
Applications
Controversies
Future Directions
Bibliography

Glossary

Dereferencing Dereferencing a URI is using the URI to identify the object or resource it refers to.

Folksonomy 'Folksonomy' is a neologism applied to structures that emerge from the practice of 'tagging' Web content. In some Web 2.0 applications, users can apply a tag (a descriptive term) to content such as a photograph or video clip. The tags need only be meaningful to the individual tagger, but if a large enough number of users tag content, descriptive structures analogous to more formal ontologies can emerge that are meaningful to wide communities.

GRDDL Gleaning Resource Descriptions from Dialects of Languages (GRDDL, pronounced 'griddle') is a mechanism for helping bootstrap the Semantic Web, by extracting RDF from XML documents, using transformations expressed in XSLT. GRDDL became a W3C recommendation in 2007.

Metadata Metadata is data about data. In the context of the Semantic Web, metadata are also called 'markup' or 'annotations'. Because one aim of the Semantic Web is to support machine processing of information, metadata are helpful in describing the content of data. For example, metadata attached to a series of numerals could explain that it represents a zip code or a height or a population figure.

Ontology An ontology defines, describes and constrains the concepts and relationships that are used in some particular domain of knowledge. Ontologies may therefore have an important role in data sharing, for example by providing a means of expressing which concepts (in the ontology) are referred to by particular terms (in a set of databases, which may use widely differing vocabularies).

OWL The Web Ontology Language (OWL) is a language for describing and sharing ontologies on the Web. It is an extension of RDF, and has three variants. OWL Full is maximally expressive and compatible with RDF (any legal RDF document is a legal OWL Full document), but is undecidable. OWL DL is based on a series of restrictions of OWL Full which support efficient reasoning, at the cost of losing the strong connection with RDF (not all RDF documents are legal OWL DL

documents). OWL Lite is even more restricted in expressivity, but is easier to grasp. OWL became a W3C recommendation in 2004.

RDF The Resource Description Framework (RDF) is a standard framework for representing data on the Web, representing it in a three-place relation, of subject-relation-object form, called a *triple*. It uses URIs to refer not only to the two items related, but also to the relationship asserted between them. In this way it extends the linking structure of the Web by allowing the nature of the link asserted to be described. Its syntax is XML-based, to support syntactic interoperability between the two. RDF became a W3C recommendation in 1999.

RDF(S) RDF Schema (RDF(S) or RDFS) is a language for representing information on the Web. It is an extension of RDF to allow the description of the relationships which can be asserted between resources using RDF. So, for instance, RDF allows the assertion of the relationship 'author' between, say, 'Herman_Melville' and 'Moby_Dick', but RDF(S) is required to assert the properties of the 'author' relationship (e. g. that every document has at least one author, or that the 'author' relationship is the inverse of the 'written_by' relationship). RDF(S) became a W3C recommendation in 2004.

Rules and RIF Rules govern the transformation of, and inference from, data. In particular, when data are being shared, it may be useful to make basic inferences over the data (for example, to determine whether two names refer to the same object or different objects). Rule-based knowledge such as this cannot be expressed in individual data stores, and may be hard to express in an ontology. The Rule Interchange Format (RIF) is a language, not complete at the time of writing, to express the most common or basic types of rule.

SPARQL SPARQL is a special query language designed specifically for querying data stored in RDF ('SPARQL' is a recursive acronym standing for 'SPARQL Protocol And RDF Query Language', and is pronounced 'sparkle'). SPARQL became a W3C recommendation in January 2008.

Triples and triplestores A triple is a statement in RDF, consisting of a subject, an object, and a binary predicate that relates them. Each item of the triple is identified by a URI reference (or can be a string literal, such as a date or a number or name). A knowledge repository which contains RDF triples is called a triplestore. Triplestores may need to contain several millions of triples, and so need to be able to support fast querying at these potentially very large scales.

URI A Uniform Resource Identifier (URI) is a string of characters for identifying an abstract or physical object or resource. There are different URI schemes (i. e. different ways to restrict the syntax of a URI to make it meaningful). A Uniform Resource Locator (URL) is a particular type of URI that identifies its object by means of its access mechanism or 'location' in the network. URIs are important in that they act as a standard way of referring to objects on the Web.

W3C The World Wide Web Consortium (W3C) is an international non-profit consortium which coordinates the development of Web standards, founded in 1994 under the directorship of Sir Tim Berners-Lee. The W3C groups together all the bodies involved in specifying Semantic Web standards into the Semantic Web Activity. A W3C working group works on a standard for a particular formalism, and when the standard is judged by the working group to be in a final form, fit for purpose and properly interoperable with other W3C standards, it ratifies the formalism by recommending it. Hence a W3C recommendation is an important standard.

Web of data The Web of Data is another way of referring to or explaining the vision of the Semantic Web, which emphasises the idea of creating links between data rather than documents (as on the current World Wide Web). Linking data, as with linking documents, enables them to be reused in interesting or unexpected contexts (when the data are interpreted using explicit semantic theories of the sort provided by ontologies). RDF is the anticipated mechanism for creating the links between data; dereferencing one or more of the URIs in an RDF triple will lead to descriptions of the resources referred to, which in turn are likely to contain further triples, which can again lead to dereferencing and so on.

Web science Web Science is the multidisciplinary activity of trying to understand the two-way dynamic relationship between Web technology and wider society, in order to ensure that the technological changes made to the Web (including the Semantic Web) are generally beneficial rather than otherwise.

XML The eXtensible Markup Language (XML) is a language for allowing users to tag or mark up content using tags of their own devising (for instance based on a particular vocabulary used in a small community of practice). As such, XML can serve as a basic data exchange format between applications. It is an advance on other well-known markup languages (such as HTML, a standard language for marking up content for display on the Web) in that it separates instructions

to do with content and document structure from those to do with formatting. It is a basic language for representing and exchanging structured information.

Definition of the Subject

The Semantic Web is a proposed extension to the World Wide Web (WWW) that aims to provide a common framework for sharing and reusing data across applications. The most common interfaces to the World Wide Web present it as a Web of Documents, linked in various ways including hyperlinks. But from the data point of view, each document is a black box – the data are not given independently of their representation in the document. This reduces its power, and also (as most information needs to be extracted from documents by a human agent) inhibits the use of automatic information processing methods on the Web. The Semantic Web is an effort, steered by the World Wide Web Consortium, to develop a set of protocols, formalisms and standards to transform the Web into a *Web of Data*. Links would be between data, and data could be accessed independently of the applications that created them. This would allow both the sharing of data, and the amalgamation of data from different sources, using heterogeneous formats, in new contexts.

Introduction

The idea of the Semantic Web (SW), of exploiting the possibilities for serendipitous reuse of linked data, dates back at least to Sir Tim Berners-Lee's plenary talk at the first International World Wide Web Conference at CERN in Geneva in 1994 [33]. In that talk, Berners-Lee argued that there is too little machine-readable information on the WWW as was currently constituted. "The meaning of the documents is clear to those with a grasp of (normally) English, and the significance of the links is only evident from the context around the anchor. To a computer, then, the Web is a flat, boring world devoid of meaning. This is a pity, as in fact documents on the Web describe real objects and imaginary concepts, and give particular relationships between them. … Adding semantics to the Web involves two things: allowing documents which have information in machine-readable forms, and allowing links to be created with relationship values. Only when we have this extra level of semantics will we be able to use computer power to help us exploit the information to a greater extent than our own reading." Indeed, the original vision of the WWW was intended to support greater machine understanding of people's work and interactions; the 'flat' understanding of the world that the WWW produces in machines is a first step towards the richer vision.

The SW was from the beginning conceived as a set of layered standards and formalisms (see Sect. "The Layered Model of the Semantic Web" for details). The development of the Resource Description Framework (RDF) in the 1990s was a key technology [80]. RDF is an important formalism, as it allows expression not only *of* the link between two objects, but also about the *nature* of the link itself. Hence, one can follow a chain of links not only via the objects linked, but also the types of links involved. In the WWW of documents, links connect documents written in the Hypertext Markup Language HTML; in the SW, links in RDF connect not only documents but arbitrary things (objects and relationships) identified by the Uniform Resource Identifiers (URIs [38]) in the RDF triples representing the data.

The ability to move between data linked in such a way opens up the possibility of exposing data to the Web, and then being able to access such data from any application. As a simple example, consider personal information such as one's bank statements, information-based resources such as digital photographs, and an application such as a calendar or diary. Each of these depends on data which is controlled by the applications that use them. But in a genuine Web of Data, we could link these data in a productive way – something as simple as being able to present one's financial information in one's calendar. The metadata in one's photographs often includes information about the time of their creation; an application able to get at the photograph metadata and the data on one's calendar might be able to suggest where the photo was taken, and its possible location. The ability to use all these data in a constructive way is impossible without a Web of linked data to enable applications to move between the data sources.

In this way, the SW changes our model of the value of information. Currently, it is generally presumed that the value of information stems from its *scarcity* – people and organizations gain value from information they have gathered, and are given monopoly rights to exploit that information via such legal contrivances as copyright, intellectual property rights, licensing, and so on. Even when organizations do not resort to the law, they will make great investments in protecting trade secrets. However, this scarcity-based model seems inadequate for the digital age.

In the first place, as economist William Baumol has argued, the social benefits from unlicensed use of 'protected' knowledge and innovation, were already large in the pre-digital economy, and indeed account for much of our wealth today: "some 80 percent of the benefits [innovation] may – i. e. change round brackets to square, as this is an editorial interpolation by us plausibly have gone to per-

sons who made no direct contribution to innovation. The rather startling implication of all this is that the spillovers of innovation, both direct and indirect, can be estimated to constitute well over half of current [US] GDP – and it can even be argued that this is a very conservative figure" (see p. 135 in [32]). And secondly, the Internet and the Web have made it harder to preserve monopoly rights to information, as copying and distribution reduces the marginal cost to producers to close to zero. Although many media companies have taken rearguard action to protect their intellectual property, so simple is the distribution model on the Web that the basis of the value of information is rapidly switching from scarcity to *abundance*. It is the large quantity of data, that can be placed in novel and unintended contexts with little cost, that makes it increasingly valuable in the age of digital technologies – and it is this abundance that the SW is designed to foster.

One of the major drivers of the SW has been the transformation of science into *e-science*, a computer-enabled, data-heavy view of science as the analysis of the very large quantities of information that improved instrumentation, larger computer power, more prevalent sensor networks and greater memory storage have released. Several disciplines have seized on the opportunity to exploit such data, which are often available only in diverse and heterogeneous datasets. In particular, interdisciplinary research is growing in importance, requiring data developed in different disciplines, using a confusion of vocabularies and methods of collection. Methods for dealing with such large and heterogeneous datasets are required in many areas, including the life sciences, climate research, medicine and epidemiology, and genomics, to name but four, which explains the interest in many of these fields in the SW.

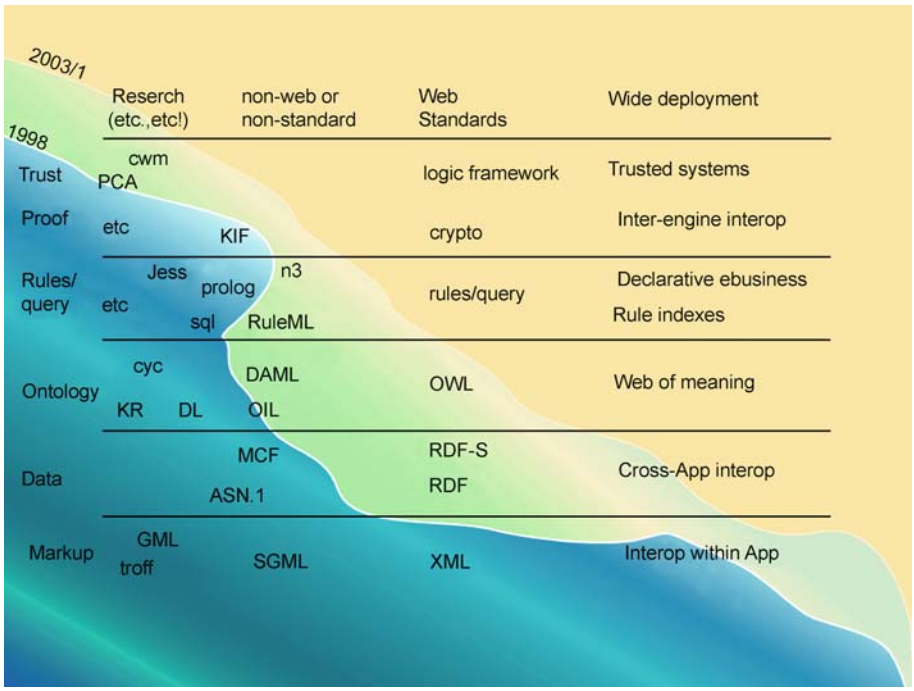
The use of the SW in such large, public projects was perhaps predictable, but much debate and discussion has focused more on the individual's interface to it. We discuss this in more detail below, but one reason for this was that the landmark publication for the public view of the SW, an article in *Scientific American* for 2001, written by Berners-Lee, James Hendler and Ora Lassila [37], developed the idea of a Web of Data with a number of household gadgets interfacing with it. The possibilities envisaged in their scenario included: a telephone that turned down the volume of all local devices with volume controls when it rang; an agent that could plan a program of medical care; and a calendar that could integrate this information to adjust a set of appointments. The point of the article was not the impressive set of agents, but rather the Web of Data that sat underneath them. However, many readers focused on the gadgets, and – given that such gadgets are not at the time of writing very common or effective – have concluded either that the SW has been a failure, or that it was an unrealistic vision from the beginning (see Sect. "Controversies").

In 2006, a further publication by Berners-Lee, together with Nigel Shadbolt and Wendy Hall, appeared in the publication *IEEE Intelligent Systems* de-emphasising the agents and focusing on the idea of the SW as a Web of Data or actionable information [101]. This paper argued that the agents described in 2001 could only flourish when standards for data sharing are well-established. The need for such standards, and for SW technologies in general, was growing, thanks to developments such as e-science, information-based medicine, and e-government.

Since the late 1990s, the World Wide Web Consortium (W3C), under the direction of Berners-Lee, has led the drive to create the standards. Figure 1 shows a diagrammatic representation of the progress of the SW in the development of its layered standards (the hierarchically-arranged layers are marked in Fig. 1 down the left hand column as markup/data/ontology etc.). The development of each layer goes through a long, sometimes tortuous, process of research and discussion. Initial stages of research, sometimes competitive, after some time produce a rough consensus about the general properties of a formalism to implement a layer. At that point, consideration is undertaken about creating a Web standard by the W3C.

To create a standard, a group, representative of stakeholders from academe and private enterprise, is assembled by the W3C, which first creates a *working draft*, which is released under that status for review by the community. Commentary is welcomed, and the draft may be changed dramatically. Once the group responsible for the standard is satisfied that it is fully capable of doing what it required, it is released as a *candidate recommendation*, when it is critiqued in terms of the practicability of its implementation. The next stage is to become a *proposed recommendation*, when it is submitted to the W3C advisory council. Finally, it is released as a W3C recommendation. It generally takes years to negotiate these various stages. Currently, the markup language XML, the data representation and interchange languages RDF and RDF(S), and the ontology language OWL are full recommendations. The next layer up from the ontology layer is that of rules and querying: the query language SPARQL became a W3C recommendation in January 2008 [96], while the rule expression language RIF is at the time of writing at the working draft stage [41].

Figure 1 represents the historical progress of the SW in terms of a 'wave' rolling over 'dry land'. The depth of the 'sea' indicates the extent to which SW standards have been accepted and been deployed widely. The SW wave is traveling from the bottom left of the diagram to the top



Semantic Web, Figure 1
The wave of development of the Semantic Web [51]

right; the net result of this is that the lower levels of the SW are coming into being and wide acceptance ahead of the higher levels. RDF is, at the time of writing, reasonably widely deployed; SPARQL has recently become a recommendation; trust (the highest level represented here) is still a research issue, and the consensus about what form a standard to represent and promote trust in data is still in the process of formation.

Linking Data

The underlying aim of the SW is to allow data to be explored and queried on the Web, analogously to the way that documents are currently investigated online. One precondition for this is obviously the publishing of data on the Web, but another is to create the links that allow data to be explored. RDF allows representation of data in such a way that anything referred to in the data can be linked to, and from. If URIs are used to name things, common naming schemes are allowed to emerge; one of the most important is the Hypertext Transfer Protocol (HTTP [59]), which affords a straightforward mechanism for people to look up the names. In a properly linked Web of Data, the URI, once looked up, should provide access to useful information about the resource named, as well as useful links out to other data.

Links can be made using various mechanisms, the simplest of which is to use a URI that points to another. For example (taken from [35]), someone might describe some relationships in RDF as follows:

```
<rdf:Description about="#albert"
   <fam:child rdf:Resource="#brian">
   <fam:child rdf:Resource="#carol">
</rdf:Description>
```

This RDF is about three resources given the local identifiers '#albert', '#brian' and '#carol', and might be placed in a file called '<http://example.org/smith>'. The architecture of the Web can use these names to provide a global identifier for the three resources; for instance "http://example.org/smith#albert" refers to #albert, and so on. And now there is a global identifier, links can be made. For instance, a document '<http://example.org/jones>' might contain the following RDF:

```
<rdf:Description about="#denise"
   <fam:child rdf:Resource="#edwin">
   <fam:child rdf:Resource=
                    "http://example.org/smith#carol">
</rdf:Description>
```

Here a series of relationships between resources #denise, #edwin and #carol have been asserted, but the datum

about #carol makes it possible to link to the data in the other file. Someone following the link dereferences the URI, i. e. decomposes 'http://example.org/smith#carol' into two parts: the part before the '#' which gives the name and location of the file; and '#carol' which is the local identifier in that file. Hence the information about #carol in the first file can be accessed thanks to the link included in the second file. The series of links between different resources can be represented, at least on a small scale, graphically, as in Fig. 3. This is the simplest way of linking data, though there are others [35]. And if the URI used for reference is created under a widely-supported system in a community, then the prospects for linking data are that much larger.

One of the main drivers of the SW is the vast quantity of data available in the form of relational databases (RDBs) (RDBs), which often exist in isolation from each other. Each database has its own value, but as argued above, the major source of informational value in the digital age is abundance, the possibilities for serendipitous reuse of data by placing it in fruitful contexts. To that end, a key aim of the SW is to harvest the large amount of data held in RDBs – a much larger quantity than is currently available in the document Web – and to support its amalgamation. The net result will be to facilitate the treatment of all the data as, in effect, sitting in a single queryable database.

Much of the current Web is supported by larger databases that sit below the level of what can be seen on a webpage; these databases are known as the *deep Web*, as opposed to the *shallow Web* on webpages. So, for instance, when one looks at one's bank statement on the shallow Web, the webpage that the bank's site creates uses a much larger quantity of data about one's bank account than is visible at any one time. Hence an alternative way of looking at this role for the SW is as a method for allowing users to query the whole of the deep Web, rather than simply the information released onto the shallow Web by the current set of implemented applications.

An RDB consists of a series of *tables*, which consist of *rows* or *records* of individual data items. Each record consists of a set of values of *fields* or *attributes*. The records (rows) and fields (columns) together can be conceived as a table or matrix (m records and n fields give us an $m \times n$ matrix) [50]. So, for instance, each record may represent an individual person, while a particular field may represent zip codes. The value placed in the zip code field of a person's record is therefore the value of that person's zip code. This tabular structure for RDBs can be mapped straightforwardly into an RDF representation. The record can be seen as the subject of an RDF triple, the field name is a link or property type, while the value of that field is the object of the triple. Hence the person who is represented by the record in the example RDB above would be represented by the first item of the triple; the zip code field would be represented by a zip_code property the second item of the triple; and the value of the zip code would be the third item. In that way, each cell of the RDB matrix is represented by an individual RDF triple, and the total set of triples would represent the entire database [34].

Having said that, there are additional factors about RDBs that are harder to capture in the RDF, and there are many open questions about the export of RDB. For instance, it may be that the database is definitive – that is, the institution holding the data has some responsibility for the data. It may be that the State of Texas holds a database of all Texas vehicle registration numbers; any car not on the database is not, as a matter of definition, registered in Texas. On a smaller scale, some RDBs allow a primary field for a unique identifier for the record, which also holds a significance beyond the particular piece of data. There are various ways of modeling these context-based properties of RDBs, perhaps most likely devolving the representation of such matters to the applications that use the data, via the ontologies, rules or query types that they use. See [34] for a worked example of methods to expose an RDB to the Web.

The process of exposing databases to the Web should not be too prescriptive – the whole point of the Web is as a decentralized collection of linked information, whether in the form of data or documents. The links are entirely democratic, and can be made between any pair of data items, or any pair of documents. This is where the power of the Web's ability to promote serendipitous reuse comes in. Attempting to fix the methods or languages used, or to 'police' the links made, will blur this vision, and create bottlenecks impeding information flow. Indeed, the widespread use of the Web is largely down to its non-prescriptive nature – prescription will simply drive users, who will not want particular information management strategies forced upon them, from the Web.

The result is a particularly untidy situation, unusual in the history of information management. If data, information or knowledge is generated within a single organization or affiliation (the usual situation for information managers before the growth of the Web), then information systems can trade on a number of simplifying assumptions. The size of such repositories can be assumed to be small or medium, and representation schemes would be planned and homogeneous. The quality of information would be likely to be high, and managers' trust in it correspondingly high. But on the Web scale, these assumptions fail. The amount of data available for query may be extremely high, and represented in highly heterogeneous ways – rarely in

the optimal way for the manager's task in hand. Information quality, and trust in that quality, would be very variable. Linking data using common URI schemes will never work perfectly, precisely because there are deliberately no enforcement mechanisms on the Web, and people cannot be forced to use any particular naming convention.

It is primarily for this reason that ontologies have always played a central role in the vision of the SW [57]. It is the ontology that puts the 'semantic' into 'Semantic Web'. Ontologies specify the vocabulary, concepts and relationships of a domain. They must be a rationalization of current practice, and managed and endorsed by a community. They act as a specification of the terms used in discussion. But they should not be too prescriptive – terms change over time, others are in constant dispute. Ontologies need to develop as a discipline or domain develops. Different areas will have different requirements of ontologies; some sciences will have large, publicly-managed ontologies which act as a public vocabulary standard, while others will make do with small, lightweight ontologies that only define the relationships between a few terms. The ontological requirements of any individual application may be quite small; the question for the application developer is whether to develop a small special-purpose ontology for her own individual purposes, or alternatively whether to reuse a larger, better-known ontology that may overspecify vocabulary for her purposes. Much will depend on the usual practices of her wider community. It should be noted that the SW project does *not* require a single overarching ontology, referring to and prescribing everything.

The ontology is key to being able to deal with heterogeneous datasets as described above; the data, and the terms used in it, must be mapped onto other terms held in common. Once this has been done, then databases can be understood in common terms – and, most crucially, the information they hold amalgamated and processed by machines. It is of course obvious to a human user that if one database of people has a field *ZC*, while another has a field *zip_code*, that there is at least a good chance that the two fields refer to the same attribute of people, viz., the zip codes of their addresses; a computer will merely try, and fail, to match the strings identifying each field. But if the computer is referred to an ontology and given mappings from the terms used in the two databases to the ontology's terms, then it can be told about the equivalence, and accordingly its inference space is opened up (for instance, it could make some inferences from the fact that $ZC(X) = zip_code(y)$). Machine processing of this heterogeneous, distributed data is made possible by ontologies.

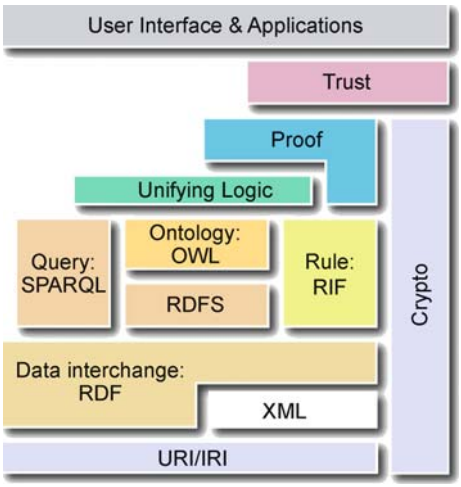
There are, of course, several issues pertaining to the use of ontologies in the SW, some of which will be discussed in greater detail in the Sect. "Controversies". The development of an ontology for an application is a (potentially large) initial cost to an application developer, and it is likely that ontologies, especially well-known ones, will be reused. To that end, searching for ontologies is likely to be a growth area in the future; there are already specialized search engines, such as Swoogle, dedicated to this task [8,54]. It may also be that one application might use several ontologies (for instance, an interdisciplinary scientific application may well reuse well-known ontologies from each discipline that it crosses). In that case, mappings *between* the ontologies will be important, and such mappings, as opposed to the ontologies themselves, could become the semantic basis for the application [78].

Building an ontology from scratch is always an option, especially if the application requires only relatively lightweight ontological apparatus [89]; again, special-purpose tools, such as Protégé are already available and well-used in the SW community [6,90]. Generating an ontology from an RDB can be done semi-automatically, and then mappings (which will be fairly straightforward, given the method of ontology generation) defined between the database and the ontology [105]. The problem is more complex if the aim is to map a legacy database onto an existing ontology. In particular, the mappings between the database and the ontology can be expected to be quite complex, and therefore very expressive languages will be required to describe them, such as the language R_2O [31].

The Layered Model of the Semantic Web

The Web, as a decentralized construction, cannot be created by fiat or prescription, which would limit its growth and create bottlenecks for information mobility. But to allow the Web of Data to reach fruition at the scale envisaged, several related tasks are required to be performed [101]. As discussed above, the W3C has devoted resources over the last few years to developing formalisms and standards to allow these tasks to be addressed, and the tasks themselves have also been arranged in a series of layers, depicted in a hierarchical diagram. The diagram has evolved with the vision of the SW, but is not dissimilar to its first incarnation, and at the time of writing is seen as in Fig. 2.

As noted above, and in Fig. 1, the development of these layers has been bottom-up, concentrating on the lower levels before work begins on those further up. The lower layers are now often the subject of W3C recommendations, while work on the upper layers remains generally more theoretical, and contains more open research questions.



Semantic Web, Figure 2
The layered view of the Semantic Web [39]

At the lowest level we have URIs (and IRIs, Internationalized Resource Identifiers, which are generalizations of URIs allowing non-ASCII characters to be used [56]). URIs identify resources in a global way – in other words, they are interpreted consistently across applications, unlike individual naming conventions, and therefore are central to the vision of a Web of Data [38]. Using a URI to identify a resource (whether that resource be a piece of information, a real-world object, an abstract concept, etc.) allows others to use the same identifier to link to the resource, refer to it, or retrieve a representation of it; this shifts the emphasis online from documents to data, and allows direct machine processing of data. If the URI scheme used is the Hypertext Transfer Protocol (http), then that is particularly helpful as http guides the user as to the location of the resource (although there are several other URI schemes, and indeed users can invent their own).

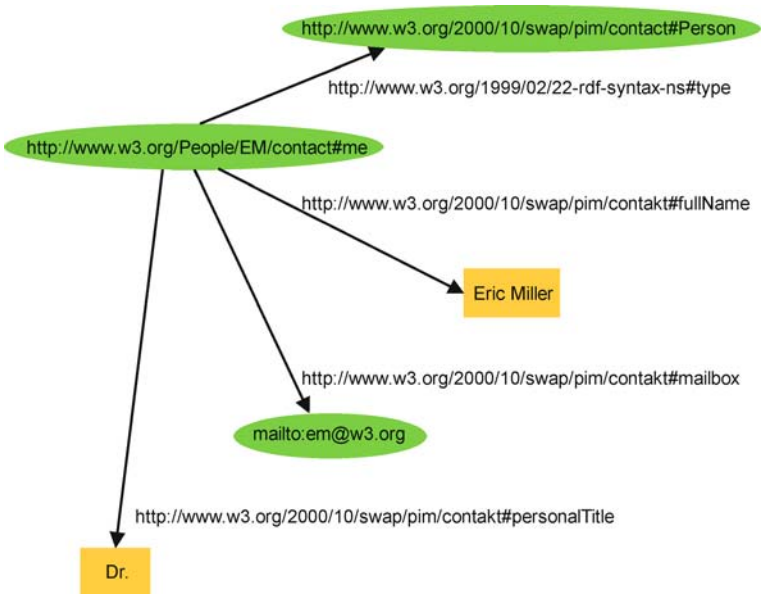
This ability to refer globally is why the export of data from RDBs to the SW should be facilitated, as noted above, by exporting database objects as first-class objects identified with URIs. A number of SW applications have diverged from this vision by not releasing their data onto the Web, but instead archiving them in inaccessible files (at least sometimes because of privacy concerns); Berners-Lee in particular has complained about this tendency [35]. However, this has happend at least sometimes because of privacy concerns.

The next layer up from names is that of markup and data interchange – the realm of XML and RDF. The eXtensible Markup Language (XML [44]) is a metalanguage for markup – in other words, a way of supporting communication and data interchange within communities by defining specialized vocabularies – commonly used in a number of sectors.

XML, like the Hypertext Markup Language (HTML) that underpins the current Web, is descended from the Standard Generalized Markup Language (SGML), an international standard for defining system-independent methods of representing information, and so has no conceptual connection to the SW. The main language for data interchange on the SW, on the other hand, RDF, is specially designed for the task, by assigning specific URIs to the fields in its triples. Figure 3 shows an RDF graph of nodes and arcs made up of several triples – each triple consists of two labeled nodes, from which is pointing a labeled directed arc. The two nodes are the first and third elements of the triple; the arc is the second (it points *from* the first element *to* the third). The use of URIs to refer to the properties as well as the objects is an important step to providing semantics – it enables us to reason about and link to relationships as well as objects.

Figure 3 shows four triples, all 'about' an individual called Eric Miller, identified by 'http://www.w3.org/People/EM/contact#me'. If we look at these triples clockwise from the right, the first represents a connection between Miller, the property of 'having the name . . . ' (which is referred to by the URI 'http://www.w3.org/2000/10/swap/pim/contact#fullName'), and a character string 'Eric Miller'. The second links Miller, by the property of having a mailbox, to the value of that property, which is his email address given using the common 'mailto:' URI scheme. The third again links Miller with a personal title, the property given as an http URI, and the title as a character string. The fourth triple provides some vocabulary in RDF – it refers to a namespace (an RDF document defining expressive resources which are imported by the RDF graph in Fig. 3) which defines some important relations – and in effect says that Miller (the first object in the triple) is an instance of (a relationship which is the second object in the triple) a person (a class which is the third object in the triple).

Based as it is on triples, RDF is simple yet powerful, exploiting the resources of the common subject/predicate/object structure, and its basis in URIs is very important for the SW. It is a minimalist knowledge representation language for the Web – there are some types of knowledge that cannot be represented in RDF, or only represented with difficulty. For instance, a predicate with more than two arguments has to be represented in a somewhat awkward way as a conjunction of two-argument predicates, while statements about hierarchical class relationships, say, need a further formalism. Furthermore, although the graph structure is quite intuitive, the actual



Semantic Web, Figure 3
An RDF graph representing Eric Miller [83]

syntax of RDF is based on XML (it is called RDF/XML) and, although it is well-suited to machine processing, it is not very easy for the human to read (see especially pp. 68–69 in [27]).

The growth in use of RDF has led to the need for special-purpose data stores for holding large quantities of RDF triples (often a set of data will be represented by millions of triples). Such data stores are known as *triplestores*, and need to provide not only storage, but efficient means of reasoning over and retrieving the data that will scale to the large sizes that will be needed. Examples of triplestores include JENA [4], 3store [7,67] and Oracle 11*g* [16].

Greater expressivity is required than is given in RDF, and to that end there another layer upwards which allows the expression of important information about the vocabularies used to express data. As we can see in Fig. 2, the layer here is relatively complex, and contains four boxes, RDFS, Ontology, Rules and Query. These between them provide representation and capabilities that are essential for putting the SW to use.

RDF Schema (RDFS, and sometimes RDF(S) [47]) provides a basic set of tools for producing structured vocabularies that allow different users to agree on particular uses of terms. An extension of RDF, it adds a few modelling primitives with a fixed meaning (such as class, subclass and property relations, and domain and range restriction). It is a basic ontology language that has been adopted fairly widely, and although fairly minimal it can express important constraints on vocabularies.

RDFS is deliberately minimal, and concentrates on expressing subclass and property hierarchies, with various restrictions on these, but the research community, including the Web Ontology Working Group of W3C, identified a number of requirements for greater expressivity for ontologies. As a couple of examples, RDFS allows the stating of subclass relationships, but not, say, that two classes are disjoint; neither does it allow class cardinality restrictions (e. g. a person has *exactly two* parents). As can be seen in Fig. 1, early research efforts into ontology languages led to two leading candidates being developed: DAML (DARPA Agent Markup Language [84]), and OIL (Ontology Interchange Language or Ontology Inference Layer [58]). These two, combined as DAML+OIL [94], became the seed for the W3C ontology language OWL. Unsubstantiated rumor suggests that the fact that it is called 'OWL' and not 'WOL' is an arcane joke: in A. A. Milne's *Winnie-the-Pooh* stories, Owl, who is wise and unusually literate for a forest-dweller, spells his name W-O-L. It is more likely that this is a post hoc rationalization of the naming decision.

The needs of ontology languages are great and potentially problematic. Expressivity is an issue. Even at the level of RDF, its reification mechanism allows the modeler to make statements about statements – an expressive possibility that can lead to logical problems. RDFS has even more powerful modelling primitives, including 'rdfs:Class', the class of all classes. OWL [85] is a strong language for representing concepts and their relations, and its creators needed to wrestle with the inevitable trade-

off between expressivity and efficient reasoning support, with two particular constraints demanded by the Semantic Web. First, there is the strong decentralization and lack of enforcement mechanisms on the Web, so that people cannot be forced to use a language they do not want to. Creating a language this powerful, for all purposes, may well have resulted in it not being used at all. Those who need great expressivity might be inclined to use their own favorite non-standard language, while those who want efficient reasoning might prefer to revert to RDFS. The second constraint is that the layered view of the SW (Fig. 2) makes it desirable that OWL should be an extension of RDFS – it should use the RDF interpretation of classes and properties and add primitives to provide richer expressivity. However, this cannot be the basis for OWL, because the addition of powerful reasoning to the expressive power of RDFS (to define such items as the class of all classes) would result in a language very hard to control.

To pick their way between these various pitfalls, the developers of OWL created three separate languages, OWL Full, OWL DL and OWL Lite [85]. OWL Full is the complete language, a full set of OWL primitives, which can be combined with RDF and RDFS in arbitrary ways. This includes the possibility that an OWL Full ontology could augment or alter the meaning of a pre-defined RDF, RDFS or OWL term (for instance, one could put a cardinality constraint on the size of the class of all classes, thereby limiting the number of possible classes that could be constructed). OWL Full is compatible with RDF, so that any legal RDF document is an OWL Full document. The downside of all this expressivity is that the language is undecidable, which rules out the possibility of complete reasoning support.

OWL DL is intended to be open to computational support, and is a sublanguage of OWL Full (so that any legal OWL DL ontology is a legal OWL Full ontology). OWL DL is so named because it is based on *description logic*, a type of knowledge representation language used to describe the knowledge definitive of a particular application domain [29], and is designed to be complete and decidable (in particular, application of OWL's constructors to each other is restricted). This does mean that full compatibility with RDF is lost: although every legal OWL DL document is a legal RDF document, the inverse is not true.

OWL Lite is a very lightweight language to support users requiring a classification hierarchy with simple constraints. Reasoning support should be relatively efficient, and it is intended to provide a straightforward migration route for bringing thesauri and taxonomies to the SW. A legal OWL Lite ontology is also a legal OWL DL ontology, but the cost of ease of reasoning is a lack of expressivity, so for instance enumerated classes, statements of disjointness and assignment of arbitrary cardinality constraints (i. e. restrictions of class cardinality to any number other than 0 and 1) are disallowed.

Even though the relationship between OWL and RDF is complex, its roots in RDF allow OWL to exploit RDF's linking capabilities to allow ontologies to be distributed across systems. When constructing an ontology in OWL, the developer can refer to terms in other ontologies, which then encourages the sharing of terminology across distributed data sources. Sharing ontologies is not always sufficient when it comes to data sharing – an organization may find that nearly all of an imported ontology is adequate, but it needs extra identifiers and descriptions, and in such a case it should be allowed to add them, rather than build a new ontology from scratch [70]. This ability to assemble distributed ontologies is central to the SW vision.

Expressing ontological relations vital, but having achieved a representation of the domain with semantics, one still needs to make inferences. OWL has some inferential support, such as subsumption and classification, but there are several inferential methods that will be required on the SW. Hence, work is currently ongoing on the Rule Interchange Format (RIF), which is intended to allow a variety of rule-based formalisms, ranging from Horn-clause logics, higher order logics and production systems, to be used [41]. Various insights from Artificial Intelligence (AI) have also been adapted for use for the SW for various purposes, including temporal (time-based) logic, causal logic and probabilistic logics [30,75,101,102].

And given the domain description at its desired level of expressivity and a means of making inferences, then the next important function at this level is the ability to query the data. Query language SPARQL works in effect by constructing a graph of RDF-like triples that may contain variables, which is then matched against the RDF graph to be queried; the query is successful if there is a subgraph of the RDF graph which matches the query when RDF terms are substituted for its variables [96].

Sitting on top of these layers are further layers with a unifying logic, proof systems and trust systems. As can be seen from Fig. 1, these upper layers remain topics of exploratory research.

Trust is perhaps key to widespread application of the SW. If information is being drawn from heterogeneous sources, then it is important that users are able to trust such sources if they are to act on the inferences that result. Trust will of course depend on the criticality of the inferences – trust entails risk, and a risk-averse user will naturally trust fewer sources [42,92]. Measuring trust, however, is a complex problem [62]. A key parameter is that of the

provenance of data, a statement of the conditions under which data were produced (including statements about the methods of production and the organization that carried them out). Methods are appearing to describe provenance [63], but more needs to be discovered about how information spreads across the Web, and therefore how it can be tracked and understood [39].

Related issues include respect for intellectual property, and the privacy of data subjects. In each case the reasoning abilities of the SW can be of value, and initiatives are currently under way to try to exploit them [93]. Protocols that allow users to express their own privacy preferences, and to enable those who wish to reuse information to reason about those preferences, are being created under the program of research into the Policy Aware Web [108]. Creative Commons is an initiative for representing copyright policies and preferences based on RDF to promote reuse where possible (current standard copyright assumptions are deliberately restrictive with respect to reuse) [2]. Cryptography protocols to protect information and privacy will also play an important role at all levels, as shown in Fig. 2.

Applications

The top layer of Fig. 2 is that of a user interface and applications. This recognizes the fact that if the SW cannot be used easily, and integrated into people's workflows in order to add value to their informational transactions, then it will not attract a large user base, without which the network effects already seen in the development of the World Wide Web will not transpire. Network effects are those positive benefits that increase in certain communication systems faster than its user base expands. In the same way as a telephone system is of limited value to a handful of people and enormous value to a large number, a few people exposing data to the SW is unlikely to make much of a difference, whereas if scalable SW technologies were applied to something like the quantity of data to be found currently in the Deep Web, the gains would be immense.

One not entirely frivolous way of expressing the need for the top layer of the SW is to say that its user base needs to grow quickly, and what is needed is a 'killer app', in other words an application that will meet a felt need and create a perception of the technology as 'essential'. Less ambitiously, the SW's spread depends not only on having an impressive set of formalisms, but also the tools to use the linked data [26].

Bootstrapping

One particular user issue is the importance of bootstrapping content for the SW. Even if RDF began to be published routinely, the amount of legacy content on the Web would dwarf new data for some time, and to make this legacy accessible to SW technology some automation of the process of creating RDF from other formats is required. CS AKTive Space [98], discussed in more detail below, amassed a large quantity of information about the state of computer science research in the United Kingdom through a relatively laborious process of harvesting information from the webpages of computer science departments in British universities without necessarily acting as a source, and using natural language processing and an ontology to interpret the data. The application was very successful, but the researchers were SW researchers, and the process is likely to be too onerous to be repeated on a large scale by non-experts. Assumptions can be made about webpage structure (for example, about regular layouts generated from a database by an individual website), and tools have been developed to exploit them [82].

An important development in this field is GRDDL (Gleaning Resource Descriptions from Dialects of Languages) which became a W3C recommendation in September 2007, and which allows the extraction of RDF from XML and XHTML (a further markup language) documents using transformations expressed in XSLT, an extensible stylesheet language based on XML. It is hoped that such extraction could allow bootstrapping of some of the hoped-for SW network effects, given the amount of XML and XHTML data in the Deep Web [52].

Annotating documents and data with metadata about their content, provenance and other useful dimensions (even including the emotional dimension to content – [100]) is also important for the effort to bring more content into the range of SW technologies [64]. Multimedia are a particular focus for research into annotation [106]. Manual annotation is a great burden for information holders, and a major initial cost for the SW, so methods of automating annotation have been investigated by a number of research teams in order to increase the quantity of annotated data available without excessive expenditure of resources [64,65,107].

In addition, as a large quantity of the Web is actually written in natural language, some have seen a role for natural language processing (NLP), and information extraction (IE), for analyzing this text statistically. So large is the Web's store of written language (two thousand billion words) that it can function as a corpus which dwarfs the most ambitious attempts of dedicated corpus builders in computational linguistics of only a few years ago [79]. And given this, and the need for automating or semi-automating annotation, NLP techniques, augmented by ontologies and training with humans, can be used to

extracted machine-readable structured information from plain text [43,48,49,76]. There have also been attempts to build ontologies using NLP techniques, another of the major anticipated bottlenecks for the SW [45]. See also the Subsect. "Commercial and Non-academic Applications" for some examples of SW applications using NLP.

Application Areas

Predicting which particular applications will succeed is unscientific and usually inaccurate. In fact, as it is the rich information contexts that the web of linked data provides that will increase the value of individual pieces of data, one way in which such growth can be encouraged is to focus on small communities with pressing information-processing requirements, and various more-or-less common goals; such communities can be the 'killer apps', or, more accurately, the early adopters of the technology, exactly as the high energy physics discipline played a vital role in the development of the WWW (cf. e. g. [20]). A series of case studies and use cases is maintained at [21].

The most promising of these communities is *e-science*, the data-driven, computationally-intensive pursuit of science in highly distributed computational environments [72]. Large quantities of data are created by analysis and experiments in disciplines such as particle physics, meteorology and the life sciences. Furthermore, in many contexts, different communities of scientists will come together to perform interdisciplinary work, so that data from various fields (e. g. genomics, clinical drug trials and epidemiology) varying not only in vocabulary, but also in the scale of description, need to be integrated. Many scientific disciplines have created large-scale and robust ontologies for this and other purposes. The most well-known of these is the Gene Ontology, a controlled vocabulary to describe gene and gene product attributes in organisms, and related vocabularies developed by Open Biomedical Ontologies. Others include the Protein Ontology, the Cell Cycle Ontology, MeSH (Medical Subject Headings, used to index life science publications), SNOMED (Systematized Nomenclature of Medicine) and AGROVOC (agriculture, forestry, fisheries and food). For more examples and references see [101].

E-government is another important application area, where heterogeneous information of varying quality is deployed widely. Government information varies in provenance, confidentiality and "shelf life" (some information will be good for decades or even centuries, while other information can be out of date within hours), while it can also have been created by various levels of government (national/federal, regional, state, city, parish). Privacy and security are also obviously important factors in this space. Integrating government information in a timely way is clearly an important challenge (see for instance a pilot study for the United Kingdom's Office of Public Sector Information, exploring the use of SW technologies for disseminating, sharing and reusing data held in the public sector [25]).

Academic Applications

Many applications for the SW have been developed with the specific purpose of bringing the SW to maturity. These are often written up in conferences such as the regular World Wide Web Conferences, the International Semantic Web Conferences (ISWC), the European Semantic Web Conferences (ESWC), as well as several one-off conferences and workshops, and can be found in the proceedings (usually online) of any of these. See also the *Journal of Web Semantics* [22].

One initiative of interest here is the Semantic Web Challenge [1], which runs annually alongside the ISWC. This is a good-natured competition to find applications that show SW technology in the best light and which can act as benchmarks for the research community. These applications, therefore, are to some extent an objective list of applications through the years that use semantic technologies to solve real-world problems involving heterogeneous real-world data. The winners of the SW Challenges from its inception, 2003, to the time of writing, 2007, are as follows.

2003: CS AKTive Space (University of Southampton) is an application to explore the UK Computer Science Research domain across multiple dimensions for multiple stakeholders, allowing the tracking of the activities of all agents from funding agencies to individual researchers, using information harvested from the Web, and mediated through an ontology [98].

2004: Flink (Vrije Universiteit Amsterdam) is a 'Who's Who' of the SW which allows the interrogation of information gathered automatically from Web-accessible resources about researchers who have participated in ISWC conferences [86].

2005: CONFOTO (appmosphere web applications, Germany) is a browsing and annotation service for conference photographs [88].

2006: MultimediaN E-Culture Demonstrator (Vrije Universiteit Amsterdam, Centre for Mathematics and

Computer Science, Universiteit van Amsterdam, Digital Heritage Netherlands and Technical University of Eindhoven) searches, navigates and annotates media collections interactively, using digital representations of items from the collections of several well-known museums and art repositories [99].

2007: Revyu.com (Open University) is a reviewing and rating site specifically designed for the SW, allowing reviews of any kind of resource, content or event to be integrated and interlinked with data from other sources (in particular, other reviews, which proliferate on the Web) [69].

A typical SW application will generate a new ontology for its application domain (e. g. art, as with MultimediaN or computer science, as with CS AKTive Space), and use it to interrogate large stores of data, whether legacy data or freshly harvested. This strand of research is tending to confirm the hypothesis that ontologies have an important role in mediating the integration of data from heterogeneous sources.

Commercial and Non-academic Applications

SW applications are generally presented using custom-built interfaces. This suggests a very important area for future research, the development of scalable visualizers capable of navigating the graph of connected information expressed in RDF. As can be seen, the importance of applications and user interfaces was made clear in the layered SW diagram (Fig. 2). However, we shouldn't expect to 'see' the SW in a special browser, in the way that we can see the Web of Documents through browsers such as Internet Explorer, Netscape or Mozilla Firefox. Rather, SW technologies, facilitating the exploration of data, may well work at the back end of websites to improve the user experience. Examples of such sites pointed to by the W3C include Sun's white paper collection site [18], Nokia's developers' discussion forum [11], Oracle's virtual press room [5], and Harper's online magazine [13].

There is an increasing number of applications supporting deeper querying of linked data. The DBpedia [28] is based on the collaborative encyclopedia Wikipedia created by volunteers, and is intended to extract structured information from Wikipedia allowing much more sophisticated querying. Sample queries given on the DPpedia website include a list of people influenced by Friedrich Nietzsche, and the set of images of American guitarists. DBpedia uses RDF, and is also interlinked with other data sources on the Web. When accessed in late 2007, the DBpedia dataset contained 103 million RDF triples. Other examples of linked data applications include the DBLP bibliography of scientific papers [23], and the GeoNames database which represents descriptions of millions of geographical features in RDF [12].

As well as existing organizations using semantic technologies to improve user experience, and applications exploiting linked data, commercial firms are beginning to appear whose business model is based on the possibilities of the SW. Garlik [24] aims to provide individual consumers with more power over their digital data. It reviews what is held about people, harvesting data from the open Web, and represents this in a people-centric structure. Natural Language Processing is used to find occurrences of people's names, sensitive information, and relations to other individuals and organisations (Declaration of interest: Wendy Hall is Chair of the Garlik Advisory Board). Twine [19] aims to facilitate knowledge and information sharing, and to organize that information using various SW technologies (also, like Garlik, using NLP). Twine's developer Nova Spivack coined the term 'knowledge networking' to describe the sharing process, analogous to the Web 2.0 idea of 'social networking'.

Controversies

The SW has been controversial during its history, with several commentators arguing that it is based upon unrealistic expectations, or repeats the mistakes of other initiatives. The arguments against the SW have tended to appear more in the blogosphere rather than the academic world, perhaps because people in the SW world are genuinely enthusiasts while those without confidence in the SW project are doing other things. The pro-SW website GetSemantic supports a wiki page of arguments against the SW, with references and responses [3]. In this section, we will examine three of the most prominent arguments raised against the SW.

The Semantic Web Repeats the Mistakes of "Good Old-Fashioned Artificial Intelligence"

It has been argued that the SW is basically a throwback to the project to program machine intelligence [77] which was jokingly christened by John Haugeland 'GOFAI' (Good Old-Fashioned AI). GOFAI proved impossible: so much of human intelligence is implicit, context-dependent and situated that writing down everything a computer needs to know to produce output that exhibits human-like intelligence is out of the question [68].

One attempt to work around this problem is the Cyc project, set up in 1984, which aims to produce a gigantic ontology that will encode all common-sense knowledge,

in order to support human-like reasoning by machines (i. e. GOFAI) [81]. The project has always aroused controversy, but it is fair to say that over two decades later, Cyc is not nearing completion and has not widely perceived as a solution for the GOFAI problem. The implicit nature of common-sense knowledge arguably makes it impossible to write it all down.

Many commentators have argued that the SW is basically a re-creation of the (misconceived) GOFAI idea, that the aim is to create machine intelligence over the Web, to allow machines to reason about Web content in such a way as to exhibit intelligence [77]. This, however, is a misconception, possibly abetted by the strong focus in the 2001 *Scientific American* article on an agent-based vision of the SW [37], although co-author of that article James Hendler has stated very firmly that he believes that the article was radically misinterpreted, and that no-one "can … say we're advocates of the big AI vision when we explicitly make it clear we're pushing for something else" [71]. The *Scientific American* article states that "Traditional knowledge-representation systems typically have been centralized, requiring everyone to share exactly the same definition of common concepts such as 'parent' or 'vehicle'. But central control is stifling, and increasing the size and scope of such a system rapidly becomes unmanageable" [37].

The SW is not GOFAI reheated, but rather an attempt to facilitate sharing of, and context-based machine reasoning over, content (and therefore the provision of machine-readable data on the Web). The aim is not to bring a single ontology, such as Cyc, to bear on all problems (implicitly defining or anticipating all problems and points of view), but to allow data to be interrogated in ways that were not anticipated by their creators. Different ontologies will be appropriate for different purposes; composite ontologies can be assembled from distributed parts [78,85] and it is frequently very basic ontologies (defining simple terms such as 'customer', 'account number' or 'account balance') that add most value to content. In this respect, the situation in the SW simply mirrors offline life where people from different communities and disciplines can and do interact without making any kind of common *global* ontological commitment [36,39,101]. The engineering challenge, as Berners-Lee et al. argue, is to allow independent consistent data systems to be connected locally without requiring global consistency [40].

Yorick Wilks, accepting that the SW is not an attempt to recreate GOFAI, argues that this is both a gain and a loss: a gain because the knowledge representation structures the SW proposes are computationally tractable, as opposed to the various GOFAI formalisms; a loss because DAML+OIL (and presumably by extension OWL) is less sophisticated than those formalisms, and may not have the representational power for the complexity of the world, whether common-sense or scientific [109]. Equally, as both Wilks and Berners-Lee point out, many in the SW world began their research careers in artificial intelligence, as Shadbolt et al. argue that "it will draw on some key insights, tools and techniques derived from 50 years of AI research" [101].

Ontologies

Ontologies, as we have seen, are vital for the SW vision of a Web of Data, but are perceived by many as expensive to develop and hard to maintain. The ideal conceptual apparatus is relative to the task in hand, and different ontologies are appropriate for different tasks. Classifications are also made relative to some background assumptions, and impose those assumptions onto the resulting ontology. To that extent, the expensive development of ontologies reflects the world view of the ontology builders, not necessarily the users. They are top-down and authoritarian, and therefore opposed to the Web ethos of decentralization and open conversation. They are fixed in advance, and so they don't work very well to represent knowledge in dynamic, situated contexts. [95] argues, for instance, that ontologies do not capture the situated processes of scientific research, the social construction of knowledge or the emergence and evolution of understanding over time, and presents an alternative way of representing this knowledge. [104] implicitly endorses this view, showing how there are issues in biology that OWL DL is not well-equipped to handle.

Other papers have made the point that some types of knowledge are more naturally modeled in ontologies than others, and, while not opposing the use of ontologies, warn against too strong a reliance on them for knowledge representation. [61] argues that ontologies cannot be too ambitious, and attempts to reify the context of an ontology (i. e. to provide context-independent accounts of knowledge) will be undermined by knowledge's situated nature. [91] argues that the social context of knowledge requires application builders to be maximally receptive to diverse types of heterogeneous reasoning, which might use knowledge that is hard to capture in hierarchical structures. See also [46] for a series of short essays debating this point.

A related critical point is that the Web as a decentralized, linked information structure must reflect the pragmatic needs of its large, heterogeneous user base which includes very many people who are naive in their understanding of computing issues. The infrastructure has to be usable widely, which argues for simplicity. The rich link-

ing structure of the Web of Documents, combined with statistically-based search engines such as Google, is much more responsive to the needs of unsophisticated users. The SW, in contrast, demands new information representation, markup and publishing practices, and corporations and information owners need to invest in new technologies. Not only that, but current statistical methods will scale up as the number of users and interactions grows, whereas logic-based methods such as those advocated by the SW, on the other hand, scale less well (cf. e. g. [111]).

The dispute has been fueled by the flowering since 2005 or so of the so-called 'Web 2.0' paradigm (of systems, communities and services facilitating collaboration and information-sharing among users). In particular, it has been argued that the meaningful structures that emerge when sufficiently large numbers of users 'tag' content with key words, structures which have been called 'folksonomies', resulting in a structure of connections and classifications emerging without central control, 'really' express the assumptions of the users, and furthermore in such a way as to respect their familiar patterns of communication and workflow. Meanwhile, ontologies 'really' express the needs of the ontology developers and their sponsors [103].

However, folksonomies are much less expressive than ontologies; they are basically variants on keyword searches. A tag 'SF' may refer to science fiction or San Francisco, even if we make the unrealistic assumption of a monoglot English user community. In a multilingual environment such as the Web, further ambiguity is possible – for instance, 'SF' might refer to the Swiss television station Schweizer Fernsehen. Furthermore, the semantics of Web 2.0 are relatively shallow, with few links and very sparse hierarchies.

When a community is large enough and the benefits clear enough to provide incentives to work together, then a large-scale ontology building and maintenance program is justified. It is true that large fixed costs will tend to skew the effort involved towards authorities who may be unrepresentative [91], but Shadbolt et al. argue explicitly that "the ontologies that will furnish the Semantics for the Semantic Web must be developed, managed, and endorsed by committed practice communities. Whether the subject is meteorology or bank transactions, proteins or engine parts, we need concept definitions we can use" [101].

It is of course an undecided question as to whether this community involvement will transpire, but in a recent note, Berners-Lee argues that such conditions will be perhaps more frequently encountered than sceptics believe. On the broad assumption that the size of an ontology-building team increases on the order of the log of the size of the ontology's user community, and that the resources needed to build an ontology increase on the order of the square of community size, the cost per individual of ontology building will diminish rapidly as community size increases. These assumptions are explicitly intended to be indicative rather than realistic [36].

More to the point, not all ontologies need be of great size and expressive depth. It is certainly not the case that the SW requires a single ontology of all discourse on the model of Cyc. Such an ontology, even if possible, would not scale, and in a decentralized structure like the Web its use could not be enforced. Even in complex scientific domains, [74] argues, using a case study from the field of medical informatics, that ontologies should be firmly based on work practices in the domain. In more mundane applications, we should expect a lot of use of small-scale, *shallow* ontologies defining just a few terms that nevertheless are widely applicable [36].

For example, the machine-readable Friend-of-a-Friend (FOAF) ontology is intended to describe people, their activities and their relations to other people. It is not complex, and publishing a FOAF profile is a fairly simple matter for which there are dedicated tools [15]. The resulting network of people has become very large indeed. A survey performed in 2004 discovered over 1.5 million documents using the FOAF ontology [55].

In any case, ontologies and folksonomies serve different purposes. Folksonomies are based on word tags, whereas the basis for ontology reference is via a URI. One of the main aims of ontology definition is to *remove* ambiguity – not globally, for this may well be impossible, but rather within the particular context of the application. Folksonomies will necessarily inherit the ambiguity from the natural language upon which they are based. Nevertheless, a strong possibility that has been considered is to use cheaply-gathered folksonomies as starting points for ontology development, gradually morphing the Web 2.0 structures into something with greater precision and less ambiguity [73,87].

Symbol Grounding

An important aspect of the SW is that URIs must be interpreted consistently. However, terms and symbols are highly variable in their definitions and use through time and space. The SW project will be boosted by processes whereby URIs are given to objects by communities and individuals, endorsed by the user community, who ensure consistency. Responsible URI 'ownership' is critical to the smooth functioning of the SW [101].

But the process of ensuring a fixed and known link between a symbol and its referent, which has been called *symbol grounding*, is at best hard [66], and at worst impossible [110]. Meanings do not stay fixed, but alter, often imperceptibly. They are delineated not only by logical definitions in terms of necessary and sufficient conditions, but also by procedures, technologies and instrumentation, and alter subtly as practice alters.

Any attempt to fix the reference of URIs is a special case of symbol grounding, and is consequently hard to do globally. Attempting to resist the alteration in community practices and norms, and reformulation of meanings of terms, would be doomed. This is understood by leading developers of the SW, who agree that "communities and practice will change norms, conceptualizations, and terminologies in complex and sociologically subtle ways. We shouldn't be surprised or attempt to resist these reformulations" [101]. But there is an important issue, as the same authors concede. "The issue for a Semantic Web built (in a community-driven way) is to know when parts need revision" [101].

Yorick Wilks has argued that Natural Language Processing techniques are essential for grounding the SW, because of the preponderance of text-based content on the Web. NLP is a vital procedural bridge from texts to knowledge representation, usually via automatic information extraction [109]. Berners-Lee has argued in response to Wilks, at a Web Science Workshop in 2005 that the SW was necessarily based on logic and firm definitions (even if those definitions were imperfect, or highly situated and task-relative), not words, use patterns or statistics. Though meanings are not fully stable, they can be stable *enough* relative to individual applications and in particular contexts to allow the SW approach to work [10]. In the case of large-scale, deep ontologies describing sciences, that perhaps will be where the SW is likely to add most value, the Berners-Lee view is reminiscent of that of Hilary Putnam that scientists are 'guardians' of meaning, who determine the 'true' referent of a word like 'water' [97]. But Berners-Lee agrees that ontologies will need to evolve – some quite quickly, and that such meanings cannot be fixed irrevocably; nevertheless, for the purposes of particular applications, this is unlikely to be a problem in practice [101].

Future Directions

The SW is a work in progress, though Shadbolt et al. argue that the need for shared semantics and a Web of Data have increased, and furthermore that the SW is "attainable" [101]. This final section will sketch some of the anticipated directions of future SW work.

Standards

The most obvious future direction is to continue the research as planned. The development of the SW has been conceived as a tide rolling over a beach, covering some areas fully, enveloping other areas more slowly (Fig. 1). As has been noted, the upper layers of the SW, looking at trust, logic and proof, are relatively underdeveloped, and are the focus for exploratory research at the cutting edge. The lower layers of the SW are in place and deployed widely. The middle layers are more or less in place; OWL and SPARQL are complete, while RIF is progressing, and should become a W3C recommendation in the fullness of time.

The Semantic Grid

Grid computing is a type of distributed computing designed to apply computational power from a number of different distributed, complete computers working in parallel, and in cooperation, on a single problem. For some extremely data-heavy problems requiring a lot of computation (particularly in e-science), grid computing is an important time-saving solution. Particular issues in grid computing include the problems of coordinated resource sharing, distributed problem-solving and the creation of 'virtual organisations' to pool data and share outcomes. The SW, of course, is another distributed computing paradigm where data sharing is a key issue – with the SW, a Web of Data, sharing is the whole point. A third distributed paradigm – software agents – is also a relevant factor.

This synergy has led to a research strand to apply semantic technologies to the problems of grid computing, adding meaning via ontologies and RDF metadata annotations to the grid. Information and services for the grid are thereby given well-defined meaning, which enables the interaction between humans and computers to be better coordinated. In particular, all the components, services and resources are adequately described for machine processing. The use of semantics to describe grid resources is known as the *Semantic Grid*, and research is ongoing [17,53,60,72].

The Policy-Aware Web

As is clear in Fig. 2, trusted systems are very important to the development of the SW. There are two reasons for this. First, if someone is reasoning with heterogeneous data harvested from the Web, then they will need to trust the data they have harvested and are using. As noted above, research is ongoing into methods for specifying the provenance of such data [63]. The second reason is that peo-

ple will not release their data if they thought it would be misused; the importance of data privacy in our digital age is easily underestimated [93]. The *policy-aware Web* is an initiative designed to rectify this problem.

The assumption behind the policy-aware Web is that inflexible and simplistic security systems and access control for the decentralized environment of the Web has hampered its development. Insufficiently sophisticated controls have made people reluctant to share data, particularly with other parties with which they do not have pre-existing information-sharing policies. Furthermore, the Web of Documents is rather coarse-grained for detailed security: the security decision to be made is to grant access to an entire website or page, or not, because policy control mechanisms for access at a finer-grained level aren't available. Thus, despite increasing amounts of useful information residing on the Web in a machine-retrieval form, reluctance to share that information remains.

The aim of policy-aware Web technology is to provide for the publication of access policies in a way that allows significant transparency for sharing among partners without requiring pre-agreement. In addition, greater control over information release can be placed in the hands of the information owner, allowing discretionary (rather than mandatory) access control to flourish. Policies would be another kind of metadata attached to information, and those wishing to use that information would be able to reason about them. For instance, one should be able to specify that the information can only be used by the agent gaining access, and that that agent should not pass the information on. Or it may be specified that the information should be deleted after a certain period of time. Or if it is to be used in a certain manner, then data should be anonymized.

Enforcement of these policies is another matter, but at present the research effort is focused on how to express such policies, and on creating theorem provers to reason about them. The result should be a much more fine-grained security picture, with greater transparency and accountability of information use [108].

Web Science

Although since its inception the Web has revolutionized communication, collaboration and education (particularly within science), relatively little is known about the way it develops. There is a growing feeling among researchers across a number of disciplines that a clear research agenda aimed at understanding the current, evolving and potential Web is needed to assure its continued growth. Such researchers want to model the Web, understand the architectural principles that have provided for its growth, and be as sure as possible that it supports the basic social values of trustworthiness, privacy, and respect for social boundaries, and their solution is to chart out a research agenda that targets the Web as a primary focus of attention [39,40].

This agenda has been dubbed *Web Science*, a combination of analysis of the Web and its dynamics, and synthesis of new languages and protocols. The Web is an engineered space created via formally specified languages but, as humans are the creators of Web pages and links between them, their interactions form emergent patterns in the Web at a macroscopic scale. These human interactions are, in turn, governed by social conventions and laws. Web Science is, therefore, inherently interdisciplinary; its goal is to both understand the growth of the Web and to create approaches allowing new powerful and more beneficial patterns to occur.

Such a research area does not yet exist in a coherent form. Within computer science Web-related research has largely focused on information retrieval algorithms and the algorithms for the routing of information through the underlying Internet. Outside of computing, researchers grow ever more dependent on the Web, but there is no concerted agenda for exploring emerging trends on the Web. Nor are those outside computer science fully engaged with the emerging Web research community to focus more specifically on the needs of science and of society as a whole, while preserving the essential invariants of the Web experience: decentralization to avoid social and technical bottlenecks, openness to the reuse of information in unexpected ways, and freedom and equality of information as it passes across the Web.

Despite excitement about the SW, the majority of the world's data is locked in large data stores and is not published as an open web of inter-referring resources. As a result, the reuse of information has been limited. Substantial research challenges arise in changing this situation. We have already discussed the need for policy controls, and for tools to allow scientists to exploit data when it emerges. But on top of that, releasing data is both a technical and a social problem, and understanding how to free data to the SW is a matter of understanding society in relation to the Web (in social, legal and economic terms) and the Web in relation to society. This is the foundation of the emerging Web Science agenda which it is hoped will inform the development of the SW [101]. The recent foundation of the Web Science Research Initiative (WSRI [9]), a joint venture between the Massachusetts Institute of Technology and the University of Southampton, is intended to drive the agenda on, acting as a focus (e. g. advising in particular on curricula to support it).

Bibliography

Primary Literature

1. http://challenge.semanticweb.org/. Accessed Aug 2008
2. http://creativecommons.org/about/. Accessed Dec 2007
3. http://getsemantic.com/wiki/Arguments_against_the_Semantic_Web. Accessed Dec 2007
4. http://jena.sourceforge.net/. Accessed Dec 2007
5. http://pressroom.oracle.com/. Accessed Dec 2007
6. http://protege.stanford.edu/. Accessed Dec 2007
7. http://sourceforge.net/projects/threestore. Accessed Dec 2007
8. http://swoogle.umbc.edu/. Accessed Dec 2007
9. http://webscience.org/. Accessed Dec 2007
10. http://www.cs.umd.edu/~hendler/2005/WebSciReport.pdf. Accessed Dec 2007
11. http://www.forum.nokia.com/. Accessed Dec 2007
12. http://www.geonames.org/. Accessed Dec 2007
13. http://www.harpers.org/. Accessed Dec 2007
14. http://www.informatik.uni-bremen.de/agki/www/swc/index.html. Accessed Dec 2007
15. http://www.ldodds.com/foaf/foaf-a-matic. Accessed Dec 2007
16. http://www.oracle.com/technology/tech/semantic_technologies/index.html. Accessed Dec 2007
17. http://www.semanticgrid.org/. Accessed Dec 2007
18. http://www.sun.com/servers/wp.jsp. Accessed Dec 2007
19. http://www.twine.com/. Accessed Dec 2007
20. http://www.w3.org/2001/sw/SW-FAQ. Accessed Dec 2007
21. http://www.w3.org/2001/sw/sweo/public/UseCases/. Accessed Dec 2007
22. http://www.websemanticsjournal.org/. Accessed Dec 2007
23. http://www4.wiwiss.fu-berlin.de/dblp/. Accessed Dec 2007
24. https://www.garlik.com/index.php. Accessed Dec 2007
25. Alani H, Dupplaw D, Sheridan J, O'Hara K, Darlington J, Shadbolt N, Tullo C (2007) Unlocking the potential of public sector information with Semantic Web technology. In: Proceedings of the 6th international Semantic Web conference, Busan, 2007. http://iswc2007.semanticweb.org/papers/701.pdf. Accessed Dec 2007
26. Alani H, Kalfoglou Y, O'Hara K, Shadbolt N (2005) Towards a killer app for the Semantic Web. In: Gil Y, Motta E, Benjamins VR, Musen MA (eds) The Semantic Web, proceedings of the international Semantic Web conference, Hiroshima, 2005. Springer, Berlin, pp 829–843
27. Antoniou G, van Harmelen F (2004) A Semantic Web primer. MIT Press, Cambridge
28. Auer S, Bizer C, Kobilarov G, Lehmann J, Cyganiak R, Ives Z (2007) DBpedia: a nucleus for a Web of open data. In: Proceedings of the 6th international Semantic Web conference, Busan, South Korea, 2007. http://iswc2007.semanticweb.org/papers/715.pdf. Accessed Dec 2007
29. Baader F, Calvanese D, McGuinness D, Nardi D, Patel-Schneider P (eds) (2003) The description logic handbook: theory, implementation and applications. Cambridge University Press, Cambridge
30. Baclawski K, Niu T (2005) Ontologies for bioinformatics. MIT Press, Cambridge
31. Barrasa J, Corcho O, Gómez-Pérez A (2004) R_2O, an extensible and semantically based database-to-ontology mapping language. In: 2nd Workshop on Semantic Web and Databases (SWDB2004), Toronto, 2004. http://www.cs.man.ac.uk/~ocorcho/documents/SWDB2004_BarrasaEtAl.pdf. Accessed Dec 2007
32. Baumol WJ (2002) The free-market innovation machine: analyzing the growth miracle of capitalism. Princeton University Press, Princeton
33. Berners-Lee T (1994) Plenary at WWW Geneva 94. http://www.w3.org/Talks/WWW94Tim/. Accessed Dec 2007
34. Berners-Lee T (1998) Relational databases on the Semantic Web. http://www.w3.org/DesignIssues/RDB-RDFhtml. Accessed Dec 2007
35. Berners-Lee T (2006/2007) Linked data. http://www.w3.org/DesignIssues/LinkedData.html. Accessed Dec 2007
36. Berners-Lee T (2007) The fractal nature of the Web. http://www.w3.org/DesignIssues/Fractal.html. Accessed Dec 2007
37. Berners-Lee T, Hendler J, Lassila O (2001) The Semantic Web. Scientific American 284(5):34–43. http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21. Accessed Dec 2007
38. Berners-Lee T, Fielding R, Masinter L (2005) Uniform Resource Identifier (URI): generic syntax. http://gbiv.com/protocols/uri/rfc/rfc3986.html. Accessed Dec 2007
39. Berners-Lee T, Hall W, Hendler JA, O'Hara K, Shadbolt N, Weitzner DJ (2006) A framework for Web Science. Found Trends Web Sci 1(1):1–134 http://www.nowpublishers.com/product.aspx?product=WEB&doi=1800000001. Accessed Dec 2007
40. Berners-Lee T, Hall W, Hendler J, Shadbolt N, Weitzner D (2006) Creating a science of the Web. Science 313(5788):769–771
41. Boley H, Kifer M (2007) RIF basic logic dialect. http://www.w3.org/TR/rif-bld/. Accessed Dec 2007
42. Bonatti PA, Duma C, Fuchs N, Nejdl W, Olmedilla D, Peer J, Shahmehri N (2006) Semantic Web policies – a discussion of requirements and research issues. In: Sure Y, Domingue J (eds) The Semantic Web: research and applications, 3rd European Semantic Web Conference 2006 (ESWC-06), Budva. Springer, Berlin
43. Bontcheva K, Tablan V, Maynard D, Cunningham H (2004) Evolving GATE to meet new challenges in language engineering. Nat Lang Eng 10(3/4):349–373
44. Bray T, Paoli J, Sperberg-McQueen CM, Maler E, Yergeau F (2006) Extensible Markup Language (XML) 1.0 (Fourth Edition). http://www.w3.org/TR/xml/. Accessed Dec 2007
45. Brewster C, Ciravegna F, Wilks Y (2002) User-centred ontology learning for knowledge management. In: Andersson B, Bergholtz M, Johannesson P (eds) Proceedings of the 6th international conference on applications of natural language to information systems. Springer, Berlin, pp 203–207
46. Brewster C, O'Hara K (2004) Knowledge representation with ontologies: the present and future. IEEE Intell Syst 19(1):72–81
47. Brickley D, Guha RV, McBride B (2004) RDF vocabulary description language 1.0: RDF Schema. http://www.w3.org/TR/rdf-schema/. Accessed Dec 2007
48. Ciravegna F, Dingli A, Guthrie L, Wilks Y (2003) Integrating information to bootstrap information extraction from Web sites. In: IJCAI 2003 Workshop on Information Integration on the Web, in conjunction with the 18th International Joint

Conference on Artificial Intelligence (IJCAI 2003), Acapulco, 2003
49. Ciravegna F, Wilks Y (2003) Designing adaptive information extraction for the Semantic Web in Amilcare. In: Handschuh S, Staab S (eds) (2003) Annotation for the Semantic Web. IOS Press, Amsterdam
50. Codd EF (1970) A relational model of data for large shared data banks. Commun ACM 13(6):377–387
51. Connolly D (2003) Semantic Web update: OWL and beyond. http://www.w3.org/2003/Talks/1017-swup/all.htm. Accessed Dec 2007
52. Connolly D (ed) (2007) Gleaning Resource Descriptions from Dialects of Langages (GRDDL). http://www.w3.org/TR/grddl/. Accessed Dec 2007
53. De Roure D, Sure Y (eds) (2006) The Semantic Grid. J Web Semant 4(2):81–123. http://www.websemanticsjournal.org/navigation.html#4. Accessed Dec 2007
54. Ding L, Finin T, Joshi A, Pan R, Cost RS, Peng Y, Reddivari P, Doshi VC, Sachs J (2004) Swoogle: a search and metadata engine for the Semantic Web. In: Proceedings of the Thirteenth ACM Conference on Information and Knowledge Management. ACM Press, pp 652–659. http://ebiquity.umbc.edu/_file_directory_/papers/115.pdf. Accessed Dec 2007
55. Ding L, Zhou L, Finin T, Joshi A (2005) How the Semantic Web is being used: an analysis of FOAF documents. In: Proceedings of the 38th international conference on system sciences. http://ebiquity.umbc.edu/_file_directory_/papers/120.pdf. Accessed Dec 2007
56. Duerst M, Suignard M (2005) Internationalized Resource Identifiers (IRIs). http://www.ietf.org/rfc/rfc3987.txt. Accessed Dec 2007
57. Fensel D (2004) Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce, 2nd edn. Springer, Berlin
58. Fensel D, van Harmelen F, Horrocks I, McGuinness DL, Patel-Schneider PF (2001) OIL: an ontology infrastructure for the Semantic Web. IEEE Intell Syst 16(2):38–45 http://www.cs.man.ac.uk/~horrocks/Publications/download/2001/IEEE-IS01.pdf. Accessed Dec 2007
59. Fielding R, Irvine UC, Gettys J, Mogul JC, Frystyk H, Masinter L, Leach P, Berners-Lee T (1999) Hypertext Transfer Protocol – HTTP/1.1. http://www.w3.org/Protocols/HTTP/1.1/rfc2616.pdf. Accessed Dec 2007
60. Goble C, Kesselman C, Sure Y (eds) (2005) Proceedings of the Dagstuhl seminar on Semantic Grid: the convergence of technologies. http://drops.dagstuhl.de/portals/index.php?semnr=05271. Accessed Dec 2007
61. Goguen JA (2004) Ontology, ontotheology and society. In: International conference on formal ontology in information systems (FOIS 2004). http://charlotte.ucsd.edu/users/goguen/pps/fois04.pdf. Accessed Dec 2007
62. Golbeck J, Hendler J (2004) Accuracy of metrics for inferring trust and reputation in Semantic Web-based social networks. In: Motta E, Shadbolt N, Stutt A, Gibbins N (eds) Engineering Knowledge in the Age of the Semantic Web, Proceedings of 14th International Conference, EKAW 2004, Whittlebury Hall. Springer, Berlin, pp 116–131
63. Groth P, Jiang S, Miles S, Munroe S, Tan V, Tsasakou S, Moreau L (2006) An architecture for provenance systems. http://eprints.ecs.soton.ac.uk/13216/1/provenanceArchitecture10.pdf. Accessed Dec 2007
64. Handschuh S, Staab S (eds) (2003) Annotation for the Semantic Web. IOS Press, Amsterdam
65. Handschuh S, Staab S, Ciravegna F (2002) S-CREAM – Semi-automatic CREAtion of Metadata. In: Gómez-Pérez A, Benjamins VR (eds) Knowledge engineering and knowledge management: ontologies and the Semantic Web, proceedings of 13th international conference, EKAW 2002, Siguënza. Springer, Berlin, pp 358–372
66. Harnad S (1990) The symbol grounding problem. Physica D 42:335–346 http://users.ecs.soton.ac.uk/harnad/Papers/Harnad/harnad90.sgproblem.html. Accessed Dec 2007
67. Harris S, Gibbins N (2003) 3store: efficient bulk RDF storage. In: Proceedings of the 1st International Workshop on Practical and Scalable Systems, Sanibel Island, Florida. http://km.aifb.uni-karlsruhe.de/ws/psss03/proceedings/harris-et-al.pdf. Accessed Dec 2007
68. Haugeland J (1979) Understanding natural language. J Philos 76:619–632
69. Heath T, Motta E (2007) Revyu.com: a reviewing and rating site for the Web of data. In: Proceedings of the 6th international Semantic Web conference 2007, Busan. http://iswc2007.semanticweb.org/papers/889.pdf. Accessed Dec 2007
70. Heflin J (2004) OWL Web Ontology Language use cases and requirements. http://www.w3.org/TR/webont-req/. Accessed Dec 2007
71. Hendler J (2007) Shirkyng my responsibility. http://www.mindswap.org/blog/2007/11/21/shirkyng-my-responsibility. Accessed Dec 2007
72. Hendler J, de Roure D (2004) E-science: the grid and the Semantic Web. IEEE Intell Syst 19(1):65–71
73. Hendler J, Golbeck J (2008) Metcalfe's law, Web 2.0 and the Semantic Web. J Web Semant 6(1):14–20 http://www.websemanticsjournal.org/papers/2007119/MetcalfsLawGolbeckV6I1.pdf. Accessed Dec 2007
74. Hu B, Dasmahapatra S, Dupplaw D, Lewis P, Shadbolt N (2007) Reflections on a medical ontology. Int J Human-comput Stud 65(7):569–582
75. Huang Z, Stuckenschmidt H (2005) Reasoning with multi-version ontologies: a temporal logic approach. In: Proceedings of the 4th International Semantic Web Workshop, http://www.cs.vu.nl/~heiner/public/ISWC05a.pdf. Accessed Dec 2007
76. Iria J, Ciravegna F (2005) Relation extraction for mining the Semantic Web. In: Dagstuhl seminar on machine learning for the Semantic Web. http://tyne.shef.ac.uk/t-rex/pdocs/dagstuhl.pdf. Accessed Dec 2007
77. Jones KS (2004) What's new about the Semantic Web? Some questions. SIGIR forum 38(2) http://www.sigir.org/forum/2004D/sparck_jones_sigirforum_2004d.pdf. Accessed Dec 2007
78. Kalfoglou Y, Schorlemmer M (2003) Ontology mapping: the state of the art. Knowl Eng Rev 18(1):1–31
79. Kilgarrif A, Grefenstette G (2003) Introduction to the special issue on the Web as corpus. Comput Linguist 29(3):333–348 http://www.kilgarriff.co.uk/Publications/2003-KilgGrefenstette-WACIntro.pdf. Accessed Dec 2007
80. Klyne G, Carroll JJ McBride B (2004) Resource Description Framework (RDF): concepts and abstract syntax. http://www.w3.org/TR/rdf-concepts/. Accessed Dec 2007
81. Lenat DB (1995) Cyc: a large-scale investment in knowledge infrastructure. Commun ACM 38(11):33–38

82. Leonard T, Glaser H (2001) Large scale acquisition and maintenance from the Web without source access. In: proceedings of workshop on knowledge markup and semantic annotation, K-CAP2001. http://eprints.ecs.soton.ac.uk/6185/1/Paper.pdf. Accessed Dec 2007
83. Manola F, Miller E, McBride B (2004) RDF primer. http://www.w3.org/TR/rdf-primer/. Accessed Dec 2007
84. McGuinness DL, Fikes R, Stein LA, Hendler J (2003) DAML-ONT: an ontology language for the Semantic Web. In: Fensel D, Hendler J, Lieberman H, Wahlster W (eds) Spinning the Semantic Web: bringing the World Wide Web to its full potential. MIT Press, Cambridge, pp 65–93
85. McGuinness DL, van Harmelen F (2004) OWL Web Ontology Language overview. http://www.w3.org/TR/owl-features/. Accessed Dec 2007
86. Mika P (2005) Flink: Semantic Web technology for the extraction and analysis of social networks. J Web Semant 3(2):211–223 http://www.websemanticsjournal.org/papers/20050719/document7.pdf. Accessed Dec 2007
87. Mika P (2007) Ontologies are us: a unified model of social networks and semantics. J Web Semant 5(1):5–15
88. Nowack B (2005) CONFOTO: A semantic browsing and annotation service for conference photos. In: Gil Y, Motta E, Benjamins VR, Musen MA (eds) The Semantic Web, proceedings of the international Semantic Web conference 2005, Hiroshima. Springer, Berlin, pp 1067–1070
89. Noy NF, McGuinness DL (2001) Ontology development 101: a guide to creating your first ontology. http://smi.stanford.edu/smi-web/reports/SMI-2001-0880.pdf. Accessed Dec 2007
90. Noy NF, Sintek M, Decker S, Crubezy M, Fergerson RW, Musen MA (2001) Creating Semantic Web contents with Protégé-2000. IEEE Intell Syst 16(2):60–71
91. O'Hara K (2004) Ontologies and technologies: knowledge representation or misrepresentation. SIGIR forum 38(2) http://sigir.org/forum/2004D/ohara_sigirforum_2004d.pdf. Accessed Dec 2007
92. O'Hara K, Alani H, Kalfoglou Y, Shadbolt N (2004) Trust strategies for the Semantic Web. In: Workshop on trust, security and reputation on the Semantic Web, 3rd international Semantic Web conference (ISWC 04), Hiroshima, 2004. http://eprints.ecs.soton.ac.uk/10029/. Accessed Dec 2007
93. O'Hara K, Shadbolt N (2008) The spy in the coffee machine: the end of privacy as we know it. Oneworld, Oxford
94. Patel-Schneider P, Horrocks I, van Harmelen F (2002) Reviewing the design of DAML+OIL: an ontology language for the Semantic Web. In: Proceedings of the 18th National Conference on Artificial Intelligence (AAAI02), Edmonton, 2002. http://www.cs.vu.nl/~frankh/postscript/AAAI02.pdf. Accessed Dec 2007
95. Pike W, Gahegan M (2007) Beyond ontologies: toward situated representations of scientific knowledge. Int J Human-comput Stud 65(7):674–688
96. Prud'hommeaux E, Seaborne A (2008) SPARQL query language for RDF. http://www.w3.org/TR/rdf-sparql-query/. Accessed Aug 2008
97. Putnam H (1975) The meaning of 'meaning'. Philosophical papers vol 2: mind, language and reality. Cambridge University Press, Cambridge
98. Schraefel MMC, Shadbolt NR, Gibbins N, Glaser H, Harris S (2004) CS AKTive Space: representing computer science on the Semantic Web. In: Proceedings of WWW 2004, New York, 2004. http://eprints.ecs.soton.ac.uk/9084/. Accessed Dec 2007
99. Schreiber G, Amin A, van Assem M, de Boer V, Hardman L, Hildebrand M, Hollink L, Huang Z, van Kersen J, de Niet M, Omelayenko B, van Ossenbruggen J, Siebes R, Taekema J, Wielemaker J, Wielinga B (2006) MultimediaN e-culture demonstrator. http://www.cs.vu.nl/~guus/papers/Schreiber06a.pdf. Accessed Dec 2007
100. Schröder M, Zovato E, Pirker H, Peter C, Burkhardt F (2007) W3C emotion incubator group report. http://www.w3.org/2005/Incubator/emotion/XGR-emotion/. Accessed Dec 2007
101. Shadbolt N, Hall W, Berners-Lee T (2006) The Semantic Web revisited. IEEE Intell Syst 21(3):96–101
102. Shafer G (1998) 'Causal logic'. In: Proceedings of IJCAI-98. http://www.glennshafer.com/assets/downloads/articles/article62.pdf. Accessed Dec 2007
103. Shirky C (2005) Ontology is overrated: categories, links and tags. http://www.shirky.com/writings/ontology_overrated.html. Accessed Dec 2007
104. Stevens R, Egaña Aranguren M, Wolstencroft K, Sattler U, Drummond N, Horridge M, Rector A (2007) Using OWL to model biological knowledge. Int J Human-comput Stud 65(7):583–594
105. Stojanovic L, Stojanovic N, Volz R (2002) Migrating data-intensive Web sites into the Semantic Web. Symposium on Applied Computing, Madrid
106. Troncy R, van Ossenbruggen J, Pan JZ, Stamou G, Halaschek-Wiener C, Simou N, Tsouvaras V (2007) Image annotation on the Semantic Web. http://www.w3.org/2005/Incubator/mmsem/XGR-image-annotation/. Accessed Dec 2007
107. Vargas-Vera M, Motta E, Domingue J, Lanzoni M, Stutt A, Ciravegna F (2002) MnM: ontology-driven semi-automatic and automatic support for semantic markup. In: Gómez-Pérez A, Benjamins VR (eds) Knowledge engineering and knowledge management: ontologies and the Semantic Web, proceedings of 13th international conference, EKAW 2002, Siguënza, 2002. Springer, Berlin, pp 379–391
108. Weitzner DJ, Hendler J, Berners-Lee T, Connolly D (2005) Creating a policy-aware Web: discretionary, rule-based access for the World Wide Web. In: Ferrari E, Thuraisingham B (eds) Web and information security. Idea Group Inc, Hershey
109. Wilks Y (2008) The Semantic Web as the apotheosis of annotation, but what are its semantics? IEE Intell Syst 23(3):41–49
110. Wittgenstein L (1953) Philosophical investigations. Basil Blackwell, Oxford
111. Zambonini D (2006) The 7 (f)laws of the Semantic Web. http://www.oreillynet.com/xml/blog/2006/06/the_7_flaws_of_the_semantic_we.html [sic]. Accessed Dec 2007

Books and Reviews

Antoniou G, van Harmelen F (2004) A Semantic Web primer. MIT Press, Cambridge

Berners-Lee T (1999) Weaving the Web: the past, present and future of the World Wide Web by its inventor. Texere Publishing, London

Berners-Lee T, Hendler J, Lassila O (2001) The Semantic Web. Scientific American 284(5):34–43. http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21. Accessed Dec 2007

Berners-Lee T, Hall W, Hendler JA, O'Hara K, Shadbolt N, Weitzner DJ (2006) A framework for Web Science. Found Trends Web Sci 1(1):1–134 http://www.nowpublishers.com/product.aspx?product=WEB&doi=1800000001. Accessed Dec 2007
Berners-Lee T, Hall W, Hendler J, Shadbolt N, Weitzner D (2006) Creating a science of the Web. Science 313(5788):769–771
Fensel D (2004) Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce, 2nd edn. Springer, Berlin
Fensel D, Hendler J, Lieberman H, Wahlster W (eds) (2003) Spinning the Semantic Web: bringing the World Wide Web to its full potential. MIT Press, Cambridge
Shadbolt N, Hall W, Berners-Lee T (2006) The Semantic Web revisited. IEEE Intell Syst 21(3):96-101

Semiclassical Spin Transport in Spin-Orbit Coupled Systems

Dimitrie Culcer[1,2]
[1] Advanced Photon Source, Argonne National Laboratory, Argonne, USA
[2] Northern Illinois University, De Kalb, USA

Article Outline

Abstract
Glossary
Definition of the Subject
Introduction
Spin Currents in Electric Fields
Future Directions
Bibliography

Abstract

This article discusses spin transport in systems with spin-orbit interactions and how it can be understood in a semiclassical picture. I will first present a semiclassical wavepacket description of spin transport, which explains how the microscopic motion of carriers gives rise to a spin current. Due to spin non-conservation the definition of the spin current has some arbitrariness. In the second part I will briefly review the physics from a density matrix point of view, which makes clear the relationship between spin transport and spin precession and the important role of scattering.

Glossary

Extrinsic effect An effect which has an explicit dependence on the form or strength of the disorder potential.

Intrinsic effect An effect which does not depend explicitly on the form and strength of the disorder potential.

Semiclassical theory A theory in which a particle's position and momentum are considered simultaneously.

Spin-orbit interaction A relativistic interaction between the spin of a particle and its momentum (which is associated with its orbital motion.)

Steady-state spin current A flow of spins induced by an electric field.

Steady-state spin density A net spin density induced by an electric field.

Definition of the Subject

Spin transport refers to the physical movement of spins across a sample and, if spin were a conserved quantity, one could make a straightforward distinction between spin-up and spin-down charge currents. The recent upsurge of interest in spin transport is, however, motivated by systems in which spin is not conserved due to the presence of spin-orbit interactions, which give rise to spin precession. Here, due to non-conservation of spin the spin current is not well defined [1,2,3]. Spin transport in these cases usually does not involve charge transport as the charge currents in the direction of spin flow cancel out. Finally, in certain materials, spin currents are accompanied by steady-state spin densities. The appearance of a spin density is not a transport phenomenon, but it is a steady-state process and is intimately connected to spin transport.

The word *semiclassical* as used in this work refers to theories which consider the position and momentum of a particle simultaneously. Semiclassical pictures are intuitive and useful in descriptions of transport, particularly in inhomogeneous systems and in spatially dependent fields, which typically vary on length scales much larger than atomic size.

In recent years, steady progress has been made towards realization of convenient semiconducting ferromagnets and spin injection into semiconductors from ferromagnetic metals [4,5,6,7] yet spin injection from a ferromagnetic metal into a semiconductor is hampered by the resistivity mismatch between the two [8]. This is one factor, in addition to basic science, motivating the search for an understanding of the way spins are manipulated electrically. The last few years have seen many experimental advances in spin transport, and spin currents have been measured directly [9,10] and indirectly [11,12,13,14,15].

Introduction

Novel physical phenomena that may lead to improved memory devices and advances in quantum information processing are closely related to spin-orbit interactions [16]. Spin-orbit interactions are present in the band

structure and in potentials due to impurity distributions. Spin-orbit coupling is in principle always present in impurity potentials and gives rise to skew scattering. Band structure spin-orbit coupling may arise from the inversion asymmetry of the underlying crystal lattice [17] (bulk inversion asymmetry), from the inversion asymmetry of the confining potential in two dimensions [18] (structure inversion asymmetry), and may be present also in inversion symmetric systems [19].

Although many observations in this entry are general, the discussion will focus on non-interacting spin-1/2 electron systems, which are pedagogically easier. The Hamiltonian of these systems typically contains a kinetic energy term and a spin-orbit coupling term, $H_{\boldsymbol{k}} = (\hbar^2 k^2)/(2m^*) + H^{\rm so}_{\boldsymbol{k}}$, where m^* is the electron effective mass. In spin-1/2 electron systems, band structure spin-orbit coupling can always be represented as a Zeeman-like interaction of the spin with a wave vector-dependent effective magnetic field $\boldsymbol{\Omega}_{\boldsymbol{k}}$, thus $H^{\rm so}_{\boldsymbol{k}} = (1/2)\boldsymbol{\sigma}\cdot\boldsymbol{\Omega}_{\boldsymbol{k}}$. Common examples of effective fields are the Rashba spin-orbit interaction, [18] which is often dominant in quantum wells with inversion asymmetry, and the Dresselhaus spin-orbit interaction, [17] which is due to bulk inversion asymmetry. The spin operator is given by $s^\sigma = (\hbar/2)\sigma^\sigma$, where σ^σ is a Pauli spin matrix. The spin current operator in these systems will be taken to be $\hat{J}^\sigma_i = (1/2)\{s^\sigma, v^i\}$, where the velocity operator is $v^i = (1/\hbar)\partial H_{\boldsymbol{k}}/\partial k_i$.

An electron spin at wave vector $\mathbf{k}$ precesses about the effective field $\boldsymbol{\Omega}_{\boldsymbol{k}}$ with frequency $\Omega_{\boldsymbol{k}}/\hbar \equiv |\boldsymbol{\Omega}_{\boldsymbol{k}}|/\hbar$ and is scattered to a different wave vector within a characteristic momentum scattering time τ_p. I will assume in this work that $\varepsilon_F\tau_p/\hbar \gg 1$, where ε_F is the Fermi energy, which is equivalent to the assumption that the carrier mean free path is much larger that the de Broglie wavelength. Within this range, the relative magnitude of the spin precession frequency $\Omega_{\boldsymbol{k}}$ and inverse scattering time $1/\tau_p$ define three qualitatively different regimes. In the ballistic (clean) regime no scattering occurs and the temperature tends to absolute zero, so that $\varepsilon_F\tau_p \to \infty$ and $\Omega_{\boldsymbol{k}}\tau_p/\hbar \to \infty$. The weak scattering regime is characterized by fast spin precession and little momentum scattering due to, e. g., a slight increase in temperature, yielding $\varepsilon_F\tau_p/\hbar \gg \Omega_{\boldsymbol{k}}\tau_p/\hbar \gg 1$. In the strong momentum scattering regime $\varepsilon_F\tau_p/\hbar \gg 1 \gg \Omega_{\boldsymbol{k}}\tau_p/\hbar$. I will concentrate on effects originating in the band structure, the observation of which requires the assumption that the materials under study are in the weak momentum scattering regime. Electric fields will be assumed uniform.

The first part of this article will present a semiclassical theory of spin transport, identifying the terms responsible for spin currents in the microscopic dynamics of carriers. Spin non-conservation as a result of spin precession leads to several possible definitions of the spin current, which emerge out of the spin equation of continuity. The second part presents a different point of view, which explains aspects not easily captured in the semiclassical approach. The steady-state density matrix is shown to contain a contribution due to precessing spins and one due to conserved spins. Steady state corrections $\propto \tau_p$ are associated with the *absence* of spin precession and give rise to spin densities in external fields [20,21,22,23,24,25,26,27,28]. Steady state corrections independent of τ_p are associated with spin precession and give rise to spin currents in external fields [1,2,3,9,10,11,12,13,14,15,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46]. Scattering between these two distributions induces significant corrections to steady-state spin currents.

Spin Currents in Electric Fields

Wave-Packet Picture of Spin Transport

This section presents a semiclassical theory of spin transport valid for a general spin-orbit system. The semiclassical method is a suitable approach to the study of transport, because, typically, in the relevant systems the external fields vary smoothly on atomic length scales. All information about the system is taken to be contained in the band structure, thus allowing a description of spin transport which does not make reference to the detailed form of the spin-orbit interaction.

The system under study is regarded as as a collection of carriers, whose semiclassical dynamics in a non-degenerate band i are described by a wave packet [47], with its charge centroid having coordinates $(\mathbf{r}_c, \mathbf{k}_c)$

$$|w_i\rangle = \int \mathrm{d}^3k\; a(\mathbf{k}, t)\mathrm{e}^{i\mathbf{k}\cdot\hat{\mathbf{r}}}|u_i(\mathbf{r}_c, \mathbf{k}, t)\rangle\;. \tag{1}$$

In the above, the function $a(\mathbf{k}, t)$ is a narrow distribution sharply peaked at $\mathbf{k}_c$, the phase of which specifies the center of charge position $\mathbf{r}_c$, while $|u_i(\mathbf{r}_c, \mathbf{k}, t)\rangle$ are lattice-periodic Bloch wave functions. The size of the wave packet in momentum space must be considerably smaller than that of the Brillouin zone. In real space, this implies that the wave packet must stretch over many unit cells.

The external electric field drives the center of the wave packet in k-space according to the semiclassical equations of motion

$$\begin{aligned}\hbar\dot{\mathbf{r}}_c &= \frac{\partial\varepsilon_i}{\partial\mathbf{k}_c} - q\mathbf{E}\times\boldsymbol{\Omega}_i\\ \hbar\dot{\mathbf{k}}_c &= q\mathbf{E}\,,\end{aligned} \tag{2}$$

with q the charge of the carriers, ε_i the band energy, and the Berry curvature

$$\boldsymbol{\Omega}_i = i\left\langle \frac{\partial u_i}{\partial \mathbf{k}} \right| \times \left| \frac{\partial u_i}{\partial \mathbf{k}} \right\rangle . \tag{3}$$

The electric field also gives rise to an adiabatic correction to the wave functions, which mixes the states making up the wave packet. The wave functions $|u_i\rangle$ therefore have the following form:

$$|u_i\rangle = |\phi_i\rangle - \sum_j \frac{\langle \phi_j | i\hbar \frac{\mathrm{d}}{\mathrm{d}t} | \phi_i \rangle}{\varepsilon_i - \varepsilon_j} |\phi_j\rangle , \tag{4}$$

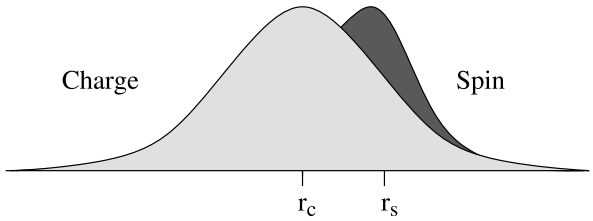
where the ϕ_i are the unperturbed Bloch eigenstates. The $|u_i\rangle$ form a complete set and retain the Bloch periodicity.

The distribution of carriers is described by a function f. When scattering is present, the distribution function satisfies the following equation:

$$\frac{\partial f}{\partial t} + \dot{\mathbf{r}}_c \cdot \frac{\partial f}{\partial \mathbf{r}_c} + \dot{\mathbf{k}}_c \cdot \frac{\partial f}{\partial \mathbf{k}_c} = \left(\frac{\mathrm{d}f}{\mathrm{d}t}\right)_{\text{coll}} , \tag{5}$$

where $(\frac{\mathrm{d}f}{\mathrm{d}t})_{\text{coll}}$ is the usual collision term. In independent bands, in the relaxation time approximation, the collision term takes the form $(f_0 - f)/\tau_p$, with f_0 the equilibrium distribution and τ_p the momentum relaxation time. In the Boltzmann theory, the change in the distribution function with time arises through the drift terms, which are determined from the semiclassical equations of motion, as well as through scattering with other carriers, with localized impurities or with phonons. For transport in a nondegenerate band, it is consistent to ignore interband scattering effects in the weak scattering limit. In this case the relaxation time is a scalar quantity. The effects of interband coherence due to scattering will be explored in the next section.

In order to obtain expressions for macroscopic quantities of interest, such as densities and currents, one needs to carry out a coarse graining by averaging over microscopic fluctuations. In classical dynamics this coarse graining is performed by means of a sampling function, which is smooth and has a significant magnitude only in a finite range [48]. This range is large compared to atomic dimensions, but small compared to the scale of variation of the distribution function. Moreover, it has a rapidly converging Taylor expansion over distances of atomic dimensions, and its form does not need to be specified. This method has a close analog in wavepacket dynamics, where the sampling function is replaced by a δ-function.



Semiclassical Spin Transport in Spin-Orbit Coupled Systems, Figure 1

For a particle of finite extent the charge and spin distributions in real space in general do not coincide. The same is true of the charge and spin distributions in reciprocal space

It is crucial to recognize that, in general, the center of spin and the center of charge are distinct (Fig. 1), since the wave packet samples a range of wave vectors and the spin is usually a function of $\mathbf{k}$. Following the line of thought outlined above, the spin density is defined to be (henceforth $\mathbf{k}_c$ will be abbreviated to $\mathbf{k}$)

$$S^\sigma(\mathbf{R}, t) = \iint \mathrm{d}^3k\mathrm{d}^3r_c f(\mathbf{r}_c, \mathbf{k}, t) \langle \delta(\mathbf{R} - \hat{\boldsymbol{r}}) \hat{s}^\sigma \rangle , \tag{6}$$

where the bracket indicates quantum mechanical averaging over the wave packet with charge centroid $(\mathbf{r}_c, \mathbf{k})$. As the δ-function has operator arguments, it will be regarded as a *sampling operator*, whose expectation value yields a spatial average, evaluated at position $\mathbf{r}$. To account for the fact that spin is not conserved (Fig. 2), a new quantity is introduced, which will be referred to as the torque density, defined by

$$\mathcal{T}^\sigma(\mathbf{R}, t) = \iint \mathrm{d}^3k\mathrm{d}^3r_c f(\mathbf{r}_c, \mathbf{k}, t) \langle \delta(\mathbf{R} - \hat{\boldsymbol{r}}) \hat{\tau}^\sigma \rangle . \tag{7}$$

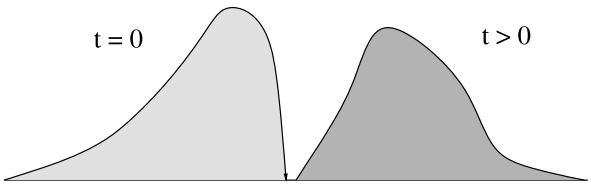
$\hat{\tau}^\sigma$ in the above stands for the rate of change of the spin operator, given by $i/\hbar[\hat{H}, \hat{s}^\sigma]$, and symmetrization of products of non-commuting operators has been assumed. Finally, the microscopic spin current density is defined as:

$$\boldsymbol{\mathcal{J}}^\sigma(\mathbf{R}, t) = \iint \mathrm{d}^3k\mathrm{d}^3r_c f(\mathbf{r}_c, \mathbf{k}, t) \langle \delta(\mathbf{R} - \hat{\boldsymbol{r}}) \hat{s}^\sigma \hat{\mathbf{v}} \rangle . \tag{8}$$

We obtain the following continuity equation for the spin density and current:

$$\frac{\partial S^\sigma}{\partial t} + \nabla \cdot \boldsymbol{\mathcal{J}}^\sigma = \mathcal{T}^\sigma + \mathcal{F}^\sigma . \tag{9}$$

The equation of continuity contains a bulk source term, which coincides with the torque density and acts as a mechanism for spin generation. Similar source terms are associated with nonconserved quantities, for example,



Semiclassical Spin Transport in Spin-Orbit Coupled Systems, Figure 2

In the presence of spin-orbit interactions the spin distribution of a particle changes in time. The horizontal axis may represent position or wave vector

in quantum electrodynamics and in Maxwell's equations. The last term in (9) represents the scattering contribution, which will be discussed further below.

Let us discuss the terms in the equation of continuity, beginning with the spin density. The argument of the sampling operator can be expressed as $[\mathbf{r} - \mathbf{r}_c - (\hat{\mathbf{r}} - \mathbf{r}_c)]$, and, as the second term is of atomic dimensions, the sampling operator can be written as a Taylor expansion about $(\hat{\mathbf{r}} - \mathbf{r}_c)$. The density can therefore be re-expressed, in terms of macroscopic quantities, as

$$S^\sigma(\mathbf{R}, t) = \rho^{s\sigma}(\mathbf{R}, t) - \nabla_{\mathbf{R}} \cdot \mathbf{P}^{s\sigma}(\mathbf{R}, t) \,, \tag{10}$$

where summation over repeated indices has been assumed. In the above, the monopole density is given by

$$\begin{aligned} \rho^{s\sigma}(\mathbf{R}, t) &= \iint \mathrm{d}^3k\mathrm{d}^3r_c f(\mathbf{r}_c, \mathbf{k}, t)\langle\hat{s}^\sigma\rangle\delta(\mathbf{R} - \mathbf{r}_c) \\ &= \int \mathrm{d}^3k f\langle\hat{s}^\sigma\rangle|_{\mathbf{r}_c=\mathbf{R}} \,, \end{aligned} \tag{11}$$

where f in the second line, and henceforth, is to be understood as $f(\mathbf{R}, \mathbf{k}, t)$, and the dipole density is

$$\begin{aligned} \mathbf{P}^{s\sigma}(\mathbf{R}, t) &= \iint \mathrm{d}^3k\mathrm{d}^3r_c f\langle(\hat{\mathbf{r}} - \mathbf{r}_c)\hat{s}^\sigma\rangle\delta(\mathbf{R} - \mathbf{r}_c) \\ &= \int \mathrm{d}^3k f\mathbf{p}^{s\sigma}|_{\mathbf{r}_c=\mathbf{R}} \,. \end{aligned} \tag{12}$$

The average spin of the wave packet has been denoted by $\langle\hat{s}^\sigma\rangle$, and the spin-dipole is defined to be $\mathbf{p}^{s\sigma} = \langle(\hat{\mathbf{r}} - \mathbf{r}_c)\hat{s}^\sigma\rangle|_{\mathbf{r}_c=\mathbf{R}}$. It will be seen that the first term in the density is the average of a monopole density located at $\mathbf{r}_c$, while the dipole term is the average of a point dipole density located at $\mathbf{r}_c$, and similarly for higher orders. The dipole must be understood as the average of the quantum mechanical dipole operator, as an exact analogy with the electric dipole of classical electrodynamics cannot be made. The density can thus be viewed as a collection of point multipoles, located at the centroid of each wave packet. The microscopic distribution of spin is important at the molecular level, but at the macroscopic level the effect of this molecular distribution is replaced by a sum of multipoles. Since the center of spin is different from the center of charge, in principle all multipoles are present.

Following a similar manipulation and using the Boltzmann equation, the torque density is re-expressed as:

$$\mathcal{T}^\sigma(\mathbf{R}, t) = \rho^{\tau\sigma}(\mathbf{R}, t) - \nabla \cdot \mathbf{P}^{\tau\sigma}(\mathbf{R}, t) \tag{13}$$

with the torque monopole density

$$\begin{aligned} \rho^{\tau\sigma}(\mathbf{R}, t) &= \iint \mathrm{d}^3k\mathrm{d}^3r_c f(\mathbf{r}_c, \mathbf{k}, t)\langle\hat{\tau}^\sigma\rangle\delta(\mathbf{R} - \mathbf{r}_c) \\ &= \int \mathrm{d}^3k f\langle\hat{\tau}^\sigma\rangle|_{\mathbf{r}_c=\mathbf{R}} \,, \end{aligned} \tag{14}$$

and the torque dipole density

$$\begin{aligned} \mathbf{P}^{\tau\sigma}(\mathbf{R}, t) &= \iint \mathrm{d}^3k\mathrm{d}^3r_c f(\mathbf{r}_c, \mathbf{k}, t)\langle(\hat{\boldsymbol{R}} - \mathbf{r}_c)\hat{\tau}^\sigma\rangle|_{\mathbf{r}_c=\mathbf{R}} \\ &= \int \mathrm{d}^3k f\mathbf{p}^{\tau\sigma}|_{\mathbf{r}_c=\mathbf{R}} \,. \end{aligned} \tag{15}$$

In analogy with the spin dipole, the torque dipole has been defined as $\mathbf{p}^{\tau\sigma} = \langle(\hat{\mathbf{r}} - \mathbf{r}_c)\hat{\tau}^\sigma\rangle|_{\mathbf{r}_c=\mathbf{R}}$. The torque density is therefore also a sum of multipole moments, that is, the moments of a point spin source located at $\mathbf{r}_c$. Even in the case when the center of $\langle\hat{s}^\sigma\rangle$ coincides with the center of charge, $\langle\hat{\tau}^\sigma\rangle$ may not be centered at $\mathbf{r}_c$, with the result that the higher order terms in the torque density are in general present. The second and higher terms of $\mathcal{T}^\sigma$ cancel exactly the analogous terms in the continuity equation which come from the current.

Since only the gradient of the spin current appears in the equation of continuity, in the expansion of the sampling operator we keep the leading term

$$\mathcal{J}^s(\mathbf{R}, t) = \int \mathrm{d}^3k f\langle\hat{\mathbf{v}}\hat{s}^\sigma\rangle|_{\mathbf{r}_c=\mathbf{R}} \,. \tag{16}$$

Keeping terms to first order in $(\hat{\boldsymbol{r}} - \mathbf{r}_c)$, the current can be decomposed into the following:

$$\mathcal{J}^{s\sigma} = \mathbf{C}^{s\sigma} + \mathbf{D}^{s\sigma} - \mathbf{P}^{\tau\sigma} \,. \tag{17}$$

The convective term $\mathbf{C}^{s\sigma}$ represents the spin being transported along with the wave packet

$$\begin{aligned} \mathbf{C}^{s\sigma}(\mathbf{R}, t) &= \iint \mathrm{d}^3k\mathrm{d}^3r_c f(\mathbf{r}_c, \mathbf{k}, t)\dot{\mathbf{r}}_c\langle\hat{s}^\sigma\rangle\delta(\mathbf{R} - \mathbf{r}_c) \\ &= \int \mathrm{d}^3k f\mathbf{c}^{s\sigma}|_{\mathbf{r}_c=\mathbf{R}} \,, \end{aligned} \tag{18}$$

while $\mathbf{D}^{s\sigma}$ comes from the rate of change of the spin dipole, which has already been introduced. It has the form:

$$\mathbf{D}^{s\sigma}(\mathbf{R}, t) = \int \mathrm{d}^3k f \frac{\mathrm{d}\mathbf{p}^{s\sigma}}{\mathrm{d}t}|_{\mathbf{r}_c=\mathbf{R}} . \tag{19}$$

$\mathbf{P}^{\tau\sigma}$ is the torque dipole introduced above. The corresponding monopole term appears in the source term of the continuity equation, as will emerge below. The presence of the torque dipole here is to be contrasted with the absence of an analogous term in classical electrodynamics. There, an electric dipole arises from the placement of two charges a small distance from each other, but the charge itself is conserved.

Finally, we come to the source term in (9). The first part, composed of the torque density, has already been discussed. The second term, denoted by $\mathcal{F}^{\sigma}$, becomes, in the relaxation time approximation

$$\mathcal{F}^{\sigma} = \frac{S_0^{\sigma} - S^{\sigma}}{\tau_p} = \frac{1}{\tau_p}\int \mathrm{d}^3k (f_0 - f)\langle \hat{s}^{\sigma}\rangle , \tag{20}$$

where τ_p is the momentum relaxation time and f_0 the equilibrium distribution, which is usually the Fermi–Dirac distribution function.

Based on the continuity equation alone, there is some flexibility in defining the current and the source. In systems in which spin is conserved, the torque density becomes, to first order in $(\hat{\boldsymbol{r}} - \mathbf{r}_c)$, a pure divergence, which can be incorporated into a redefinition of the spin current. This current, henceforth referred to as the spin transport current, is only due to the convective and spin dipole contributions:

$$\mathbf{J}^{t\sigma}(\mathbf{R}, t) = \mathbf{C}^{s\sigma}(\mathbf{R}, t) + \mathbf{D}^{s\sigma}(\mathbf{R}, t) . \tag{21}$$

With respect to this spin transport current, the continuity equation takes the following form:

$$\frac{\partial S^{\sigma}}{\partial t} + \nabla \cdot \mathbf{J}^{t\sigma} = \frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}^3k f \langle \hat{s}^{\sigma}\rangle . \tag{22}$$

In the steady state under a constant electric field, the distribution function is composed of an equilibrium part, independent of the field, and a non-equilibrium part, which is first order in the field. Henceforth, terms in the spin current and source which depend on the equilibrium distribution function will be referred to as intrinsic, whereas the terms depending on the nonequilibrium shift in the distribution will be referred to as extrinsic. For example, the integrand in Eq. (16) can be decomposed into a zero order spin-velocity, $\mathbf{v}\langle \hat{s}^{\sigma}\rangle$, where $\mathbf{v}$ is the usual group velocity of the band, and a first order correction. Therefore, there will be a contribution to the current from the non-equilibrium part of the distribution and the zero order spin-velocity, which has been discussed extensively in previous work [29,30,31,32,33]. There will also be a contribution from the equilibrium distribution and the first order correction to the spin-velocity, which is referred to as the intrinsic contribution. In the wave packet formalism presented here this effect arises from the change in wave functions induced by the electric field, rather than from the change in distribution functions that is responsible for most conventional transport effects. The intrinsic spin current is calculated from (16) using the equilibrium distribution and the expectation values of the spin and spin dipole operators in a Bloch state perturbed to first order in $\mathbf{E}$.

In its turn, the source in (9) can be decomposed into intrinsic and extrinsic contributions. The present entry considers homogeneous systems, so that all the gradient terms vanish, and the torque density is simply $f\langle \hat{\tau}^{\sigma}\rangle$. The zeroth order contribution to this term is null, as the Bloch wave functions are eigenstates of the Hamiltonian. Thus, to first order in the electric field, we find that $\langle \hat{\tau}^{\sigma}\rangle$ is simply given by $(e\mathbf{E}/\hbar)\cdot(\partial\langle \hat{s}^{\sigma}\rangle/\partial\mathbf{k})$. One is thus justified in replacing f by its equilibrium value f_0, in which case this term is purely intrinsic. The second term in the source, $\mathcal{F}^{\sigma}$, which depends on the nonequilibrium shift in the distribution function, is entirely extrinsic.

The extrinsic source term takes into account the effect of scattering, and is a term which usually appears in the equation of continuity. The intrinsic source accounts for the effect of all spin-nonconserving terms, and must be present even in a clean system, if the Hamiltonian contains spin-dependent contributions. In general, in addition to the rate of change of spin arising from the spin-dependent terms in the Hamiltonian, scattering processes may alter the orientation of the spin, with the result that any one spin component is not conserved, and the orientation of spins is randomized over a longer time period. For a uniform steady-state system, the current is constant and the intrinsic source term must vanish. However, near the boundary of the system, or at an interface with a different semiconductor with (for example) weaker spin-orbit interactions, the spin current driven by an electric field will vary spatially and $\mathcal{T}$ must reach a non-zero value.

Let us take a closer look at the spin dipole and torque dipole, which are seen to be the main mechanisms responsible for generating the spin current. Because of its narrow distribution in $\mathbf{k}$, the mean spin of the wave packet is $\langle w_i|\hat{s}^{\sigma}|w_i\rangle = \langle u_i|\hat{s}^{\sigma}|u_i\rangle$, where it is understood that the wave vector of the Bloch function is set at $\mathbf{k}_c$ and $\hat{s}^{\sigma}$ is an arbitrary projection of the spin vector operator. The spin

dipole of the wave packet, defined relative to the charge center of the wave packet is given, in terms of Bloch functions, by the expression:

$$\mathbf{p}_i^{s\sigma} = \frac{i}{2}\left[\left\langle u_i|\hat{s}^\sigma|\frac{\partial u_i}{\partial \mathbf{k}}\right\rangle - \left\langle \frac{\partial u_i}{\partial \mathbf{k}}|\hat{s}^\sigma|u_i\right\rangle\right] - \left\langle u_i|i\frac{\partial u_i}{\partial \mathbf{k}}\right\rangle\langle u_i|\hat{s}^\sigma|u_i\rangle \ . \quad (23)$$

Interestingly, the spin dipole is independent of the wave packet width. The expression is also invariant under a local gauge transformation, in the sense that if $|u_i\rangle$ is modified by a phase factor $e^{i\alpha(\mathbf{k})}$ the spin dipole is unchanged.

The torque dipole term has a special interpretation in the case of spin transport. The rate of change of spin is equivalent to a torque, and the torque dipole represents the moment exerted by this torque about the center of the wave packet. The semiclassical expression for the torque moment is

$$\mathbf{p}_i^{\tau\sigma} = \frac{i}{2}\left[\left\langle u_i|\dot{\hat{s}}^\sigma|\frac{\partial u_i}{\partial \mathbf{k}}\right\rangle - \left\langle \frac{\partial u_i}{\partial \mathbf{k}}|\dot{\hat{s}}^\sigma|u_i\right\rangle\right] - \left\langle u_i|i\frac{\partial u_i}{\partial \mathbf{k}}\right\rangle\left\langle u_i|\dot{\hat{s}}^\sigma|u_i\right\rangle \ . \quad (24)$$

The torque moment has the same gauge invariance properties as the spin dipole, and like the spin dipole it also does not depend on the wave packet width.

It is important to note that the spin current, $\boldsymbol{\mathcal{J}}^\sigma$, can be simplified to:

$$\boldsymbol{\mathcal{J}}^\sigma(\mathbf{R}, t) = \int \mathrm{d}^3k f \,\mathrm{tr}\langle u_i|\hat{s}^\sigma\hat{\mathbf{v}}|u_i\rangle \ , \quad (25)$$

which is the semiclassical equivalent of the Kubo formula for spin currents.

Density Matrix Picture of Spin Transport

Semiclassical theory provides a straightforward, intuitive picture of the way spin currents arise in the course of carrier dynamics in an electric field. The theory was developed for independent bands. It turns out that interband coherence arising from scattering is crucial in spin transport, and is difficult to treat semiclassically. Although the semiclassical theory can be generalized to multiple bands [49,50], it is more instructive to examine spin transport from a different point of view that is closer in outlook to the philosophy underlying the Kubo formula (with which the semiclassical theory agrees.) This will shed some light on additional issues, such as the relationships between spin currents and spin precession, between spin currents and spin densities, the complex effect of disorder and the vanishing of spin current in certain systems.

A large, uniform system of non-interacting spin-1/2 electrons is represented by a one-particle density operator $\hat{\rho}$. The expectation value of an observable represented by a Hermitian operator $\hat{O}$ is given by $\mathrm{tr}(\hat{\rho}\hat{O})$ and $\hat{\rho}$ satisfies the quantum Liouville equation

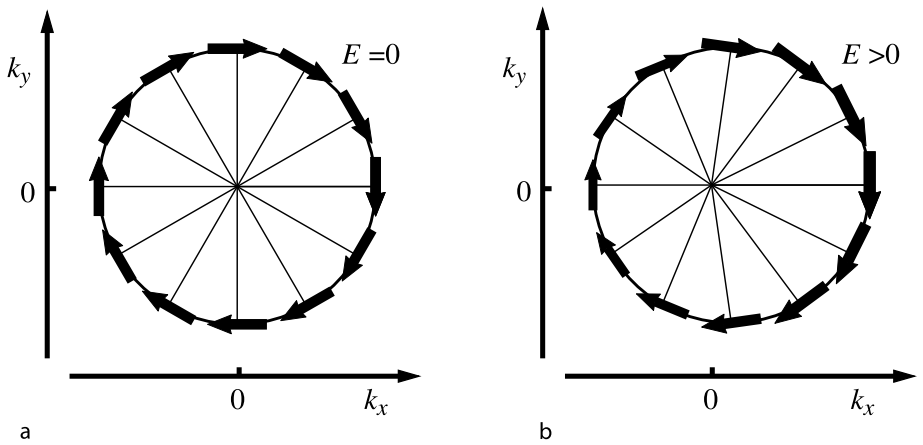
$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} + \frac{i}{\hbar}[\hat{H} + \hat{U}, \hat{\rho}] = 0 \ . \quad (26)$$

The Liouville equation is projected onto a set of time-independent states of definite wave vector $\{|\mathbf{k}\rangle\mathbf{s}\}$, which are not assumed to be eigenstates of the Hamiltonian $\hat{H}$. The matrix elements of $\hat{\rho}$ in this basis will be written as $\rho_{\mathbf{k}\mathbf{k}'} \equiv \rho_{\mathbf{k}\mathbf{k}'}^{ss'} = \langle \mathbf{k}s\hat{\rho}|\mathbf{k}'s'\rangle$. Spin indices are not shown explicitly, and $\rho_{\mathbf{k}\mathbf{k}'}$ is a matrix in spin space, referred to as the density matrix. In this work we require the expectation values of operators which are diagonal in wave vector, and will thus require the part of the density matrix diagonal in wave vector, $\rho_{\mathbf{k}\mathbf{k}} \equiv f_\mathbf{k} = n_\mathbf{k}\mathbb{1} + S_\mathbf{k}$. In the presence of a constant uniform electric field $\mathbf{E}$, $f_\mathbf{k} = f_{0\mathbf{k}} + f_{E\mathbf{k}}$, where the equilibrium density matrix $f_{0\mathbf{k}}$ is given by the Fermi–Dirac distribution, and the correction $f_{E\mathbf{k}}$ is due to the $\mathbf{E}$. We subdivide $f_{0\mathbf{k}} = n_{0\mathbf{k}}\mathbb{1} + S_{0\mathbf{k}}$ and $f_{E\mathbf{k}} = n_{E\mathbf{k}}\mathbb{1} + S_{E\mathbf{k}}$. The spin-dependent part of the nonequilibrium correction to the density matrix $S_{E\mathbf{k}}$ is interpreted as the spin density induced by $\mathbf{E}$. The equations governing the time evolution of $n_{E\mathbf{k}}$ and $S_{E\mathbf{k}}$ is [51]

$$\begin{aligned} \frac{\partial n_{E\mathbf{k}}}{\partial t} + \hat{J}_0(n_{E\mathbf{k}}) &= \frac{e\mathbf{E}}{\hbar}\cdot\frac{\partial n_{0\mathbf{k}}}{\partial \mathbf{k}} \\ \frac{\partial S_{E\mathbf{k}}}{\partial t} + \frac{i}{\hbar}[H_\mathbf{k}, S_{E\mathbf{k}}] + \hat{J}_0(S_{E\mathbf{k}}) &= \frac{e\mathbf{E}}{\hbar}\cdot\frac{\partial S_{0\mathbf{k}}}{\partial \mathbf{k}} - \hat{J}_s(n_{E\mathbf{k}}) \\ &\equiv \Sigma_{E\mathbf{k}} \ , \end{aligned} \quad (27)$$

where the scalar part of the scattering operator $\hat{J}_0$ and its spin-dependent part $\hat{J}_s$ have been defined in [51]. The equation for $n_{E\mathbf{k}}$ has the well-known solution $n_{E\mathbf{k}} = (e\mathbf{E}\tau_p/\hbar)\cdot(\partial n_{0\mathbf{k}}/\partial\mathbf{k})$, in other words, $n_{E\mathbf{k}}$ describes the shift of the Fermi sphere in the presence of the electric field $\mathbf{E}$, with the momentum relaxation time τ_p. It is seen from Eq. (27) that spin-dependent scattering gives rise to a renormalization of the driving term in the equation for $S_{E\mathbf{k}}$. This renormalization has no analog in charge transport.

We need to find the expectation value of the spin current operator $\hat{J}_i^\sigma$ defined in the introduction. In the systems under study the spin current operator can be written as $\hat{J}_i^\sigma = \hbar k_i s^\sigma/m^* + (1/4\hbar)\partial\Omega^\sigma/\partial k_i\mathbb{1}$. We need to determine $S_{E\mathbf{k}}$. To this end we remember that an electron spin at wave vector $\mathbf{k}$ precesses about an effective magnetic field $\boldsymbol{\Omega}_\mathbf{k}$. The spin can be resolved into components parallel



Semiclassical Spin Transport in Spin-Orbit Coupled Systems, Figure 3
Effective field $\boldsymbol{\Omega}_k$ at the Fermi energy in the Rashba model [18] a without ($E = 0$) and b with an external electric field ($E > 0$)

and perpendicular to $\boldsymbol{\Omega}_{\mathbf{k}}$. In the course of spin precession the component of the spin parallel to $\boldsymbol{\Omega}_{\mathbf{k}}$ is conserved, while the perpendicular component is continually changing. Corresponding to this decomposition of the spin is an analogous decomposition of the spin distribution $S_{E\mathbf{k}}$ into a part representing conserved spin and a part representing precessing spin, denoted by $S_{E\mathbf{k}\parallel}$ and $S_{E\mathbf{k}\perp}$ respectively. There is an analogous decomposition of the source on the RHS of Eq. (27) into $\Sigma_{E\mathbf{k}\parallel}$ and $\Sigma_{E\mathbf{k}\perp}$. This decomposition is carried out by introducing projection operators $P_\parallel$ and $P_\perp$ as described in [51], giving for $S_{E\mathbf{k}\parallel}$ and $S_{E\mathbf{k}\perp}$ in the weak momentum scattering limit

$$\frac{\partial S_{E\mathbf{k}\parallel}}{\partial t} + P_\parallel \hat{J}_0(S_{E\mathbf{k}}) = \Sigma_{E\mathbf{k}\parallel}\,, \tag{28a}$$

$$\frac{\partial S_{E\mathbf{k}\perp}}{\partial t} + \frac{i}{\hbar}[H_{\mathbf{k}}, S_{E\mathbf{k}\perp}] = \Sigma_{E\mathbf{k}\perp} - P_\perp \hat{J}_0(S_{E\mathbf{k}})\,. \tag{28b}$$

Equation (28b) shows that scattering mixes the distributions of conserved and precessing spins. This is so because when one spin at wave vector $\mathbf{k}$ and precessing about $\boldsymbol{\Omega}_{\mathbf{k}}$ is scattered to wave vector $\mathbf{k}'$ and precesses about $\boldsymbol{\Omega}_{\mathbf{k}'}$, its conserved component changes, a process which alters the distributions of conserved and precessing spin. Equations (28a,28b) can be solved straightforwardly if one assumes the impurity potential to be short-ranged, obtaining [51]

$$S_{E\mathbf{k}\parallel} = \Sigma_{E\mathbf{k}\parallel}\tau_p + P_\parallel(1-\bar{P}_\parallel)^{-1}\bar{\Sigma}_{E\mathbf{k}\parallel}\tau_p\,, \tag{29a}$$

$$S_{E\mathbf{k}\perp} = \frac{\boldsymbol{\Omega}_{\mathbf{k}} \times (\boldsymbol{\Sigma}_{E\mathbf{k}\perp}\tau_p + P_\perp \bar{\mathbf{S}}_{E\mathbf{k}\parallel}) \cdot \boldsymbol{\sigma}\,\tau_p}{2\hbar(1+\Omega_{\mathbf{k}}^2\tau_p^2/\hbar^2)} - \frac{(\Sigma_{E\mathbf{k}\perp}\tau_p + P_\perp \bar{S}_{E\mathbf{k}\parallel})}{1+\Omega_{\mathbf{k}}^2\tau_p^2/\hbar^2}\,. \tag{29b}$$

The correction $S_{E\mathbf{k}\parallel}$ does not give rise to a spin current. Inspection of Eq. (29a) shows that integrals of the form $\int \mathrm{d}\theta\, \hat{J}_i^\sigma S_{E\mathbf{k}\parallel}$ contain an odd number of powers of $\mathbf{k}$ and are therefore zero. It can, however, give rise to a nonequilibrium spin density, since integrals of the form $\int \mathrm{d}\theta \hat{s}^\sigma S_{E\mathbf{k}\parallel}$ contain an even number of powers of $\mathbf{k}$ and may be nonzero. Similarly $S_{E\mathbf{k}\perp}$ does not lead to a nonequilibrium spin density. The expectation value of the spin operator yields integrals of the form $\int \mathrm{d}\theta \hat{s}^\sigma S^{(0)}_{E\mathbf{k}\perp}$, which involve odd numbers of powers of $\mathbf{k}$ and are therefore zero. This term does, however, give rise to nonzero spin currents, since integrals if the form $\int \mathrm{d}\theta \hat{J}_i^\sigma S_{E\mathbf{k}\perp}$ contain an even numbers of powers of $\mathbf{k}$ and may be nonzero. Therefore, in the absence of spin-orbit coupling in the scattering potential, nonequilibrium spin currents arise from spin precession (as outlined by Sinova et al. [35]), and nonequilibrium spin densities from the absence of spin precession. The dominant contribution to the nonequilibrium spin density in an electric field exists because in the course of spin precession a component of each individual spin is preserved. For an electron with wave vector $\mathbf{k}$, this spin component is parallel to $\boldsymbol{\Omega}_k$. In equilibrium the average of these conserved components is zero. When an electric field is applied the Fermi surface is shifted and the average of the conserved spin components may be nonzero, as illustrated in Fig. 1. This argument explains why the nonequilibrium spin density $\propto \tau_p^{-1}$ and *requires* scattering to balance the drift of the Fermi surface. Although spin densities in electric fields require band structure spin-orbit interactions and therefore spin precession, the dominant contribution arises as a result of the absence of spin precession.

Systems in which $\boldsymbol{\Omega}_{\mathbf{k}}$ is linear in $\mathbf{k}$ are special, in that the spin current as defined in this section vanishes [39,40,41,42,43,44,45,46]. This is because of the renormalization of the driving term on the RHS of Eq. (28b) for $S_{E\mathbf{k}\perp}$, in other words because of scattering from the conserved spin distribution to the precessing spin distribution. In Eq. (29b) it is also clear that if $\Sigma_{E\mathbf{k}\perp}\tau_p + P_\perp \bar{S}_{E\mathbf{k}\parallel}$ vanishes, then all the contributions to $S_{E\mathbf{k}\perp}$ also vanish. Since $\bar{S}_{E\mathbf{k}\parallel}$ effectively represents a steady-state spin density, we see that the presence of this spin density tends to diminish the spin current. In systems with energy dispersion linear in $\mathbf{k}$ it cancels the spin current completely.

Future Directions

Whereas the community appears to be in agreement that spin currents exist and are measurable, many questions remain unanswered. Theoretically, intrinsic and extrinsic effects (such as due to skew scattering and side jump) have not been studied on the same footing for an arbitrary form of band structure spin-orbit interactions. The relative magnitude of intrinsic and extrinsic spin currents in such a general system remains to be determined. Also, different definitions of the spin current give results that often differ by a sign [1,2]. The relationship between spin current and spin accumulation at the boundary is not clear, again thanks to the non-conservation of spin. It appears that what happens at the boundary is sensitive to the type of boundary conditions assumed. Thus so far as quantitative interpretation of experimental data is concerned, theory has some way to go.

Despite tremendous progress, experiment is still searching for a reliable way to *measure*, as opposed to *detect*, spin currents directly. Practically, the question of what to do with spin once it has been transported/generated remains. The revolutionary electronic device that harnesses spin currents for a practical purpose remains to be made, and the challenge of its design confronts experimentalists and theorists alike.

The research at Argonne National Laboratory was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.

Bibliography

1. Shi J, Zhang P, Xiao D, Niu Q (2006) Proper definition of spin current in spin-orbit coupled systems. Phys Rev Lett 96:076604
2. Sugimoto N, Onoda S, Murakami S, Nagaosa N (2006) Spin hall effect of a conserved current: Conditions for a nonzero spin hall current. Phys Rev B 73:113305
3. Wang Y, Xia K, Su Z-B, Ma Z (2006) Consistency in formulation of spin current and torque associated with a variance of angular momentum. Phys Rev Lett 96:066601
4. Fiederling R et al (1999) Injection and detection of a spin-polarized current in a light-emitting diode. Nature 402:787
5. Ohno Y et al (1999) Electrical spin injection in a ferromagnetic semiconductor heterostructure. Nature 402:790
6. Jiang X et al (2003) Optical detection of hot-electron spin injection into GaAs from a magnetic tunnel transistor source. Phys Rev Lett 90:256603
7. Jonker BT (2003) Progress toward electrical injection of spin-polarized electrons into semiconductors. Proceedings IEEE 91:727
8. Schmidt G et al (2000) Fundamental obstacle for electrical spin injection from a ferromagnetic metal into a diffusive semiconductor. Phys Rev B 62:R4790
9. Valenzuela SO, Tinkham M (2006) Direct electronic measurement of the spin Hall effect. Nature 442:176
10. Liu B, Shi J, Wang W, Zhao H, Li D, Zhang S-C, Xue Q, Chen D (2006) Experimental observation of the inverse spin hall effect at room temperature. cond-mat/0610150
11. Kato YK, Myers RC, Gossard AC, Awschalom DD (2004) Observation of the spin hall effect in semiconductors. Science 306:1910
12. Wunderlich J, Kaestner B, Sinova J, Jungwirth T (2005) Experimental observation of the Spin-Hall Effect in a two-dimensional Spin-Orbit coupled semiconductor system. Phys Rev Lett 94:047204
13. Sih V, Myers RC, Kato YK, Lau WH, Gossard AC, Awschalom DD (2005) Spatial imaging of the spin Hall effect and current-induced polarization in two-dimensional electron gases. Nature Phys 1:31–35
14. Stern NP, Ghosh S, Xiang G, Zhu M, Samarth N, Awschalom DD (2006) Current-induced polarization and the Spin Hall Effect at room temperature. Phys Rev Lett 97:126603
15. Ganichev SD, Ivchenko EL, Danilov SN, Eroms J, Wegscheider W, Weiss D, Prettl W (2001) Conversion of spin into directed electric current in quantum wells. Phys Rev Lett 86:4358
16. Žutić I, Fabian J, Das Sarma S (2004) Spintronics: Fundamentals and applications. Rev Mod Phys 76:323
17. Dresselhaus G (1955) Spin-Orbit coupling effects in zinc blende structures. Phys Rev 100:580
18. Bychkov YA, Rashba EI (1984) Properties of a 2D electron gas with lifted spectral degeneracy. JETP Lett 39:78
19. Luttinger JM (1956) Quantum theory of cyclotron resonance in semiconductors: General theory. Phys Rev 102:1030
20. Ivchenko EL, Pikus GE (1978) New photogalvanic effect in gyrotropic crystals. JETP Lett 27:604
21. Levitov LS, Nazarov YV, Eliashberg GM (1985) Magnetoelectric effects in conductors with mirror isomer symmetry. Sov Phys JETP 61(1):133
22. Edelstein VM (1990) Spin polarization of conduction electrons induced by electric current in two-dimensional asymmetric electron systems. Solid State Comm 73:233
23. Aronov AG, Lyanda-Geller YB, Pikus GE (1991) Spin polarization of electrons by an electric current. Sov Phys JETP 73:537
24. Magarill LI, Chaplik AV, Entin MV (2001) Spin response of 2D electrons to a lateral electric field. Semiconductors 35(9):1081
25. Vorob'ev LE, Ivchenko EL, Pikus GE, Farbstein II, Shalygin VA, Sturbin AV (1979) Optical activity in tellurium induced by a current. JETP Lett 29:441

26. Kato YK, Myers RC, Gossard AC, Awschalom DD (2004) Current-induced spin polarization in strained semiconductors. Phys Rev Lett 93:176601
27. Ganichev SD, Danilov SN, Schneider P, Bel'kov VV, Golub LE, Wegscheider W, Weiss D, Prettl W (2004) Can an electric current orient spins in quantum wells? cond-mat/0403641
28. Silov AY, Blajnov PA, Wolter JH, Hey R, Ploog KH, Averkiev NS (2004) Current-induced spin polarization at a single heterojunction. Appl Phys Lett 85:5929
29. D'yakonov MI, Perel' VI (1971) Spin orientation of electrons associated with the interband absorption of light in semiconductors. Sov Phys JETP 33:1053–1059
30. Hirsch JE (1999) Spin Hall Effect. Phys Rev Lett 83:1834
31. Zhang S (2000) Spin Hall Effect in the presence of spin diffusion. Phys Rev Lett 85:393
32. Qi Y, Zhang S (2002) Crossover from diffusive to ballistic transport properties in magnetic multilayers. Phys Rev B 65:214407 (ibid 67:052407 (2003))
33. Engel H-A, Halperin BI, Rashba EI (2005) Theory of Spin Hall Conductivity in n-doped GaAs. Phys Rev Lett 95:166605
34. Murakami S, Nagaosa N, Zhang S-C (2003) Dissipationless quantum spin current at room temperature. Science 301:1348
35. Sinova J, Culcer D, Niu Q, Sinitsyn NA, Jungwirth T, MacDonald AH (2004) Universal intrinsic Spin Hall Effect. Phys Rev Lett 92:126603
36. Adagideli I, Bauer GEW (2005) Intrinsic Spin Hall edges. Phys Rev Lett 95:256602
37. Rashba EI (2004) Sum rules for Spin Hall conductivity cancellation. Phys Rev B 70:201309
38. Zhang S, Yang Z (2005) Intrinsic spin and orbital angular momentum hall effect. Phys Rev Lett 94:066602
39. Inoue J, Bauer GEW, Molenkamp LW (2004) Suppression of the persistent spin Hall current by defect scattering. Phys Rev B 70:041303
40. Dimitrova OV (2005) Spin-Hall conductivity in a two-dimensional Rashba electron gas. Phys Rev B 71:245327
41. Schwab P, Raimondi R (2002) Magnetoconductance of a two-dimensional metal in the presence of spin-orbit coupling. Eur Phys J B 25:483
42. Mishchenko EG, Shytov AV, Halperin BI (2004) Spin current and polarization in impure two-dimensional electron systems with spin-orbit coupling. Phys Rev Lett 93:226602
43. Khaetskii A (2006) Nonexistence of intrinsic Spin Currents. Phys Rev Lett 96:056602
44. Shytov AV, Mishchenko EG, Engel H-A, Halperin BI (2006) Small-angle impurity scattering and the spin hall conductivity in two-dimensional semiconductor systems. Phys Rev B 73:075316
45. Liu SY, Lei XL, Horing NJM (2006) Vanishing spin-hall current in a diffusive Rashba two-dimensional electron system: A quantum Boltzmann equation approach. Phys Rev B 73:035323
46. Mal'shukov AG, Chao KA (2005) Spin Hall conductivity of a disordered two-dimensional electron gas with Dresselhaus spin-orbit interaction. Phys Rev B 71:121308
47. Sundaram G, Niu Q (1999) Wave-packet dynamics in slowly perturbed crystals: Gradient corrections and Berry-phase effects. Phys Rev B 59:14915
48. Jackson JD (1999) Classical Electrodynamics, 3rd edn. Wiley, New York, section 6.6, pp 248–258
49. Culcer D, Yao Y, Niu Q (2005) Coherent wave-packet evolution in degenerate bands. Phys Rev B 72:085110
50. Culcer D, Niu Q (2006) Geometrical phase effects in the Wigner distribution of Bloch electrons. Phys Rev B 74:035209
51. Culcer D, Winkler R (2007) Steady states of spin distributions in the presence of spin-orbit interactions. Phys Rev B 76:245322

Shallow Water Waves and Solitary Waves

Willy Hereman
Department of Mathematical and Computer Sciences, Colorado School of Mines, Golden, USA

Article Outline

Glossary
Definition of the Subject
Introduction
Completely Integrable Shallow Water Wave Equations
Shallow Water Wave Equations of Geophysical Fluid Dynamics
Computation of Solitary Wave Solutions
Water Wave Experiments and Observations
Future Directions
Acknowledgments
Bibliography

Glossary

Deep water A surface wave is said to be in deep water if its wavelength is much shorter than the local water depth.

Internal wave A internal wave travels within the interior of a fluid. The maximum velocity and maximum amplitude occur within the fluid or at an internal boundary (interface). Internal waves depend on the density-stratification of the fluid.

Shallow water A surface wave is said to be in shallow water if its wavelength is much larger than the local water depth.

Shallow water waves Shallow water waves correspond to the flow at the free surface of a body of shallow water under the force of gravity, or to the flow below a horizontal pressure surface in a fluid.

Shallow water wave equations Shallow water wave equations are a set of partial differential equations that describe shallow water waves.

Solitary wave A solitary wave is a localized gravity wave that maintains its coherence and, hence, its visibility through properties of nonlinear hydrodynamics. Solitary waves have finite amplitude and propagate with constant speed and constant shape.

Soliton Solitons are solitary waves that have an elastic scattering property: they retain their shape and speed after colliding with each other.

Surface wave A surface wave travels at the free surface of a fluid. The maximum velocity of the wave and the maximum displacement of fluid particles occur at the free surface of the fluid.

Tsunami A tsunami is a very long ocean wave caused by an underwater earthquake, a submarine volcanic eruption, or by a landslide.

Wave dispersion Wave dispersion in water waves refers to the property that longer waves have lower frequencies and travel faster.

Definition of the Subject

The most familiar water waves are waves at the beach caused by wind or tides, waves created by throwing a stone in a pond, by the wake of a ship, or by raindrops in a river (see Fig. 1). Despite their familiarity, these are all different types of water waves. This article only addresses shallow water waves, where the depth of the water is much smaller than the wavelength of the disturbance of the free surface. Furthermore, the discussion is focused on gravity waves in which buoyancy acts as the restoring force. Little attention will we paid to capillary effects, and capillary waves for which the primary restoring force is surface tension are not covered.

Although the history of shallow water waves [15,20,22] goes back to French and British mathematicians of the eighteenth and early nineteenth century, Stokes [71] is considered one of the pioneers of hydrodynamics (see [21]). He carefully derived the equations for the motion of incompressible, inviscid fluid, subject to a constant vertical gravitational force, where the fluid is bounded below by an impermeable bottom and above by a free surface. Starting from these fundamental equations and by making further simplifying assumptions, various shallow water wave models can be derived. These shallow water models are widely used in oceanography and atmospheric science.

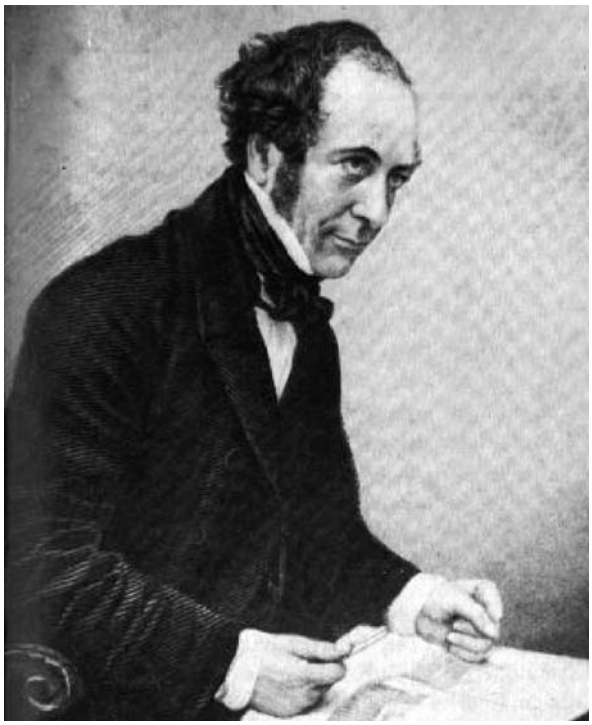
Shallow Water Waves and Solitary Waves, Figure 1
Capillary surface waves from raindrops. Photograph courtesy of E. Scheller and K. Socha

This article discusses shallow water wave equations commonly used in oceanography and atmospheric science. They fall into two major categories: Shallow water wave models with wave dispersion are discussed in Sect. "Completely Integrable Shallow Water Wave Equations" Most of these are completely integrable equations that admit smooth solitary and cnoidal wave solutions for which computational procedures are outlined in Sect. "Computation of Solitary Wave Solutions". Sect. "Shallow Water Wave Equations of Geophysical Fluid Dynamics" covers classical shallow water wave models without dispersion. Such hyperbolic systems can admit shocks. Sect. "Water Wave Experiments and Observations" addresses a few experiments and observations. The article concludes with future directions in Sect. "Future Directions".

Introduction

The initial observation of a solitary wave in shallow water was made by John Scott Russell, shown in Fig. 2. Russell was a Scottish engineer and naval architect who was conducting experiments for the Union Canal Company to design a more efficient canal boat.

In Russell's [63] own words: "I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses, when the boat suddenly stopped – not so the mass of water in the channel which it had put in motion; it accumulated round the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth and well-defined heap of water, which continued its course along the channel apparently without change of form or diminution of speed. I followed it on horseback, and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet long and a foot to a foot and a half in height. Its height gradually diminished, and after a chase of one or two miles I lost it in the windings of the channel. Such, in the month of August 1834, was my first chance interview with that singular and beautiful phenomenon which I have called the Wave of Translation."



Shallow Water Waves and Solitary Waves, Figure 2
John Scott Russell. Source: [27]. Courtesy of John Murray Publishers

Russell built a water tank to replicate the phenomenon and research the properties of the solitary wave he had observed. Details can be found in a biography of John Scott Russell (1808–1882) by Emmerson [27], and in review articles by Bullough [15], Craik [20], and Darrigol [22], who pay tribute to Russell's research of water waves.

In 1895, the Dutch professor Diederik Korteweg and his doctoral student Gustav de Vries [47] derived a partial differential equation (PDE) which models the solitary wave that Russell had observed. Parenthetically, the equation which now bears their name had already appeared in seminal work on water waves published by Boussinesq [13,14] and Rayleigh [59]. The solitary wave was considered a relatively unimportant curiosity in the field of nonlinear waves. That all changed in 1965, when Zabusky and Kruskal realized that the KdV equation arises as the continuum limit of a one dimensional anharmonic lattice used by Fermi, Pasta, and Ulam [29] to investigate "thermalization" – or how energy is distributed among the many possible oscillations in the lattice. Zabusky and Kruskal [78] simulated the collision of solitary waves in a nonlinear crystal lattice and observed that they retain their shapes and speed after collision. Interacting solitary waves merely experience a phase shift, advancing the faster and retarding the slower. In analogy with colliding particles, they coined the word "solitons" to describe these elastically colliding waves. A narrative of the discovery of solitons can be found in [77].

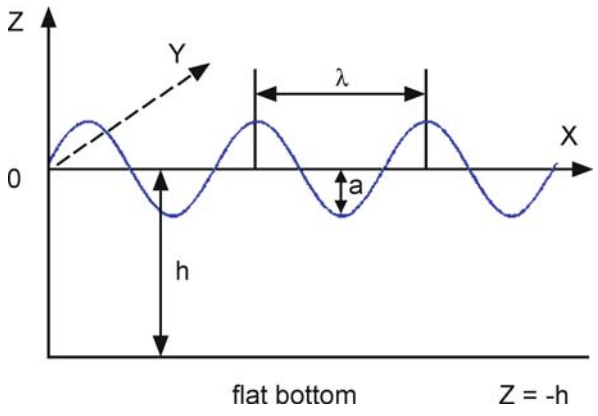
Shallow Water Waves and Solitary Waves, Figure 3
Recreation of a solitary wave on the Scott Russell Aqueduct on the Union Canal. Photograph courtesy of Heriot–Watt University

Since the 1970s, the KdV equation and other equations that admit solitary wave and soliton solutions have been the subject of intense study (see, e. g., [23,30,60]). Indeed, scientists remain intrigued by the physical properties and elegant mathematical theory of the shallow water wave models. In particular, the so-called completely integrable models which can be solved with the Inverse Inverse scattering transform (IST). For details about the IST method the reader is referred to Ablowitz et al. [3], Ablowitz and Segur [2], and Ablowitz and Clarkson [1]. The completely integrable models discussed in the next section are infinite-dimensional Hamiltonian systems, with infinitely many conservation laws and higher-order symmetries, and admit soliton solutions of any order.

As an aside, in 1995, scientists gathered at Heriot–Watt University for a conference and successfully recreated a solitary wave but of smaller dimensions than the one observed by Russell 161 years earlier (see Fig. 3).

Completely Integrable Shallow Water Wave Equations

Starting from Stokes' [71] governing equations for water waves, completely integrable PDEs arise at various levels of approximation in shallow water wave theory. Four length scales play a crucial role in their derivation. As shown in Fig. 4, the wavelength λ of the wave measures the distance between two successive peaks. The amplitude a measures the height of the wave, which is the varying distance between the undisturbed water to the peak of the wave. The



Shallow Water Waves and Solitary Waves, Figure 4
Coordinate frame and periodic wave on the surface of water

depth of the water h is measured from the (flat) bottom of the water up to the quiescent free surface. The fourth length scale is along the Y-axis which is along the crest of the wave and perpendicular to the (X, Z)-plane.

Assuming wave propagation in water of uniform (shallow) depth, i. e. h is constant, and ignoring dissipation, the model equations discussed in this section have a set of common features and limitations which make them mathematically tractable [68]. They describe (i) long waves (or shallow water), i. e. $h \ll \lambda$, (ii) with relatively small amplitude, i. e. $a \ll h$, (iii) traveling in one direction (along the X-axis) or weakly two-dimensional (with a small component in the Y-direction). Furthermore, the small effects must be comparable in size. For example, in the derivation of the KdV and Boussinesq equations one assumes that $\varepsilon = a/h = O(h^2/\lambda^2)$, where ε is a small parameter ($\varepsilon \ll 1$), and O indicates the order of magnitude.

The Korteweg–de Vries Equation

The KdV equation was originally derived to describe shallow water waves of long wavelength and small amplitude. In the derivation, Korteweg and de Vries assumed that all motion is uniform in the Y-direction, along the crest of the wave. In that case, the surface elevation (above the equilibrium level h) of the wave, propagating in the X-direction, is a function only of the horizontal position X (along the canal) and of time T, i. e. $Z = \eta(X, T)$.

In terms of the physical parameters, the KdV equation reads

$$\frac{\partial \eta}{\partial T} + \sqrt{gh}\frac{\partial \eta}{\partial X} + \frac{3}{2}\frac{\sqrt{gh}}{h}\eta\frac{\partial \eta}{\partial X} + \frac{1}{2}h^2\sqrt{gh}\left(\frac{1}{3} - \frac{\mathcal{T}}{\rho g h^2}\right)\frac{\partial^3 \eta}{\partial X^3} = 0\,, \quad (1)$$

where h is the uniform water depth, g is the gravitational acceleration (about 9.81m/sec^2 at sea level), ρ is the density, and $\mathcal{T}$ stands for the surface tension. The dimensionless parameter $\mathcal{T}/\rho g h^2$ is called the *Bond number* which measures the relative strength of surface tension and the gravitational force.

Keeping only the first two terms in (1), the speed of the associated linear (long) wave is $c = \sqrt{gh}$. This is indeed the maximum attainable speed of propagation of gravity-induced water waves of infinitesimal amplitude. The speed of propagation of the small-amplitude solitary waves described by (1) is slightly higher. According to Russell's empirical formula the speed equals $\sqrt{g(h+k)}$, where k is the height of the peak of the solitary wave above the surface of undisturbed water. As Bullough [15] has shown, Russell's approximate speed and the true speed of solitary waves only differ by a term of $O(k^2/h^2)$.

The KdV equation can be recast in dimensionless variables as

$$u_t + \alpha u u_x + u_{xxx} = 0\,, \quad (2)$$

where subscripts denote partial derivatives. The parameter α can be scaled to any real number. Commonly used values are $\alpha = \pm 1$ or $\alpha = \pm 6$.

The term u_t describes the time evolution of the wave propagating in one direction. Therefore, (2) is called an *evolution* equation. The nonlinear term $\alpha u u_x$ accounts for steepening of the wave, and the linear dispersive term u_{xxx} describes spreading of the wave. The linear first-order term $\sqrt{gh}\frac{\partial \eta}{\partial X}$ in (1) can be removed by an elementary transformation. Conversely, a linear term in u_x can be added to (2).

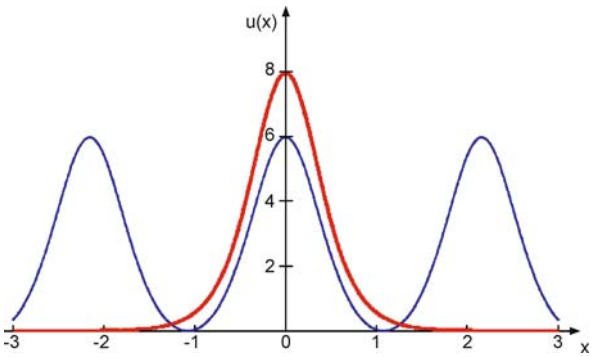
The nonlinear steepening of the water wave can be balanced by dispersion. If so, the result of these counteracting effects is a stable solitary wave with particle-like properties. A solitary wave has a finite amplitude and propagates at constant speed and without change in shape over a fairly long distance. This is in contrast to the concentric group of small-amplitude capillary waves, shown in Fig. 1, which disperse as they propagate.

The closed-form expression of a solitary wave solution is given by

$$u(x,t) = \frac{\omega - 4k^3}{\alpha k} + \frac{12k^2}{\alpha}\operatorname{sech}^2(kx - \omega t + \delta) \quad (3)$$

$$= \frac{\omega + 8k^3}{\alpha k} - \frac{12k^2}{\alpha}\tanh^2(kx - \omega t + \delta)\,, \quad (4)$$

where the wave number k, the angular frequency ω, and δ are arbitrary constants.



Shallow Water Waves and Solitary Waves, Figure 5
Solitary wave (red) and periodic cnoidal (blue) wave profiles

Requiring that $\lim_{x\to\pm\infty} u(x,t) = 0$ for all time leads to $\omega = 4k^3$. Then (3) and (4) reduce to

$$u(x,t) = \frac{12k^2}{\alpha}\operatorname{sech}^2(kx - 4k^3t + \delta) = \frac{12k^2}{\alpha}[1 - \tanh^2(kx - 4k^3t + \delta)]\,. \quad (5)$$

The position of the hump-type wave at $t = 0$ is depicted in Fig. 5 for $\alpha = 6$, $k = 2$, and $\delta = 0$. As time changes, the solitary wave with amplitude $2k^2 = 8$ travels to the right at speed $v = \omega/k = 4k^2 = 16$. The speed is exactly twice the peak amplitude. So, the higher the wave the faster it travels, but it does so without change in shape. The reciprocal of the wavenumber k is a measure of the width of the sech-square pulse.

As shown by Korteweg and de Vries [47], Eq. (2) also has a simple periodic solution,

$$u(x,t) = \frac{\omega - 4k^3(2m-1)}{\alpha k} + \frac{12k^2m}{\alpha}\operatorname{cn}^2(kx - \omega t + \delta; m), \quad (6)$$

which they called the *cnoidal wave* solution since they involve the Jacobi elliptic cosine function, cn, with modulus m $(0 < m < 1)$. The wavenumber k gives the characteristic width of each oscillation in the "cnoid."

Three cycles of the cnoidal wave are depicted in Fig. 5 at $t = 0$. The graph corresponds to $\alpha = 6, k = 2, m = 3/4, \omega = 16$, and $\delta = 0$. Using the property $\lim_{m\to 1} \operatorname{cn}(\xi; m) = \operatorname{sech}(\xi)$, one readily verifies that (6) reduces to (3) as m tends to 1. Pictorially, the individual oscillations then stretch infinitely far apart leaving a single-pulse solitary wave.

The celebrated KdV equation appears in all books and reviews about soliton theory. In addition, the equation has been featured in, e. g., Miura [52] and Miles [51].

Regularized Long-Wave Equations

A couple of alternatives to the KdV equation have been proposed. A first alternative,

$$u_t + u_x + \alpha u u_x - u_{xxt} = 0\,, \quad (7)$$

was proposed by Benjamin, Bona, and Mahony [7]. Hence, (7) is referred to as the BBM or regularized long-wave (RLW) equation.

Equation (7), which has a solitary wave solution,

$$u(x,t) = \frac{\omega - k - 4k^2\omega}{\alpha k} + \frac{12k\omega}{\alpha}\operatorname{sech}^2(kx - \omega t + \delta), \quad (8)$$

was also derived by Peregrine [57] to describe the behavior of an undular bore (in water), which comprises a smooth wavefront followed by a train of solitary waves. An undular bore can be interpreted as the dispersive analog of a shock wave in classical non-dispersive, dissipative hydrodynamics [26].

The linear dispersion relation for the KdV equation, $\omega = k(1 - k^2)$, can be obtained by substituting $u(x,t) = \exp[i(kx - \omega t + \delta)]$ into $u_t + u_x + u_{xxx} = 0$. The linear phase velocity, $v_p = \omega/k = 1 - k^2$, becomes negative for $|k| > 1$, thereby contradicting the assumption of unidirectional propagation. Furthermore, the group velocity $v_g = d\omega/dk = 1 - 3k^2$ has no lower bound which implies that there is no limit to the rate at which shorter ripples propagate in the negative x-direction.

The BBM equation, where $\omega = k/(1 + k^2), v_p = 1/(1 + k^2)$, and $v_g = (1 - k^2)/(1 + k^2)^2$, was proposed to get around these objections and to address issues related to proving the existence of solutions of the KdV equation. The dispersion relation of (7) has more desirable properties for high wave numbers, but the group velocity becomes negative for $|k| > 1$. In addition, the KdV and BBM equations are first order in time making it impossible to specify both u and u_t as initial data.

To circumvent these limitations, a second alternative,

$$u_t + u_x + \alpha u u_x + u_{xtt} = 0\,, \quad (9)$$

was proposed by Joseph and Egri [45] and Jeffrey [43]. It is called the time regularized long-wave (TRLW) equation and its solitary wave solution is given by

$$u(x,t) = \frac{\omega - k - 4k\omega^2}{\alpha k} + \frac{12\omega^2}{\alpha}\operatorname{sech}^2(kx - \omega t + \delta). \quad (10)$$

The TRLW equation shares many of the properties of both the KdV and BBM equations, at the cost of a more complicated dispersion relation, $\omega = (-1 \pm \sqrt{1 + 4k^2})/2k$ with two branches. Only one of these branches is valid because

the derivation of the TRLW equation shows that (9) is uni-directional, despite the two time derivatives in u_{xtt}.

Bona and Chen [8] have shown that the initial value problem for the TRLW equation is well-posed, and that for small-amplitude, long waves, solutions of (9) agree with solutions of either (2) or (7). As a matter of fact, all three equations agree to the neglected order of approximation over a long time scale, provided the initial data is properly imposed (see also [9]).

Fine-tuning the dispersion relation of the KdV equation comes at a cost. In contrast to (2), the RLW and TRLW equations are no longer completely integrable. Perhaps that is why these equations never became as popular as the KdV equation.

The Boussinesq Equation

The classical Boussinesq equation,

$$\eta_{TT} - c^2\eta_{XX} - \frac{3c^2}{h}\left(\eta_X^2 + \eta\eta_{XX}\right) - \frac{c^2h^2}{3}\eta_{XXXX} = 0, \quad (11)$$

was derived by Boussinesq [12] to describe gravity-induced surface waves as they propagate at constant (linear) speed $c = \sqrt{gh}$ in a canal of uniform depth h.

In contrast to the KdV equation, (11) has a second-order time-derivative term. Ignoring all but the first two terms in (11), one obtains the linear *wave* equation, $\eta_{TT} - c^2\eta_{XX} = 0$, which describes both left-running and right-running waves. However, (11) is *not* bi-directional because in the derivation Boussinesq used the constraint $\eta_T = -c\eta_X$, which limits (11) to waves traveling to the right.

In dimensionless form, the Boussinesq equation reads

$$u_{tt} - c^2u_{xx} - \alpha u_x^2 - \alpha u u_{xx} - \beta u_{xxxx} = 0\,. \quad (12)$$

The values of the parameters $c, \alpha > 0$, and β do not matter, but the sign of β matters. Typically, one sets $c = 1, \alpha = 3$, and $\beta = \pm 1$.

A simple solitary wave solution of (12) is given by

$$u(x,t) = \frac{\omega^2 - c^2k^2 - 4\beta k^4}{\alpha k^2} + \frac{12\beta k^2}{\alpha}\operatorname{sech}^2(kx - \omega t + \delta)\,. \quad (13)$$

The equation with $\beta = 1$ is a scaled version of (11) and thus most relevant to shallow water wave theory. Mathematically, (12) with $\beta = 1$ is ill-posed, (even without the nonlinear terms), which means that the initial value problem cannot be solved for arbitrary data. This shortcoming does not happen for (12) with $\beta = -1$, which is therefore nicknamed the "good" Boussinesq equation [50]. Nonetheless, the classical and good Boussinesq equations are completely integrable.

The "improved" or "regularized" Boussinesq equation (see, e. g., [10]) has $\beta = 1$ but u_{xxtt} instead of u_{xxxx}, which improves the properties of the dispersion relation. Like (12), the regularized version describes uni-directional waves. The regularized Boussinesq equation and other alternative equations listed in the literature (see, e. g., Madsen and Schäffer [49]) are not completely integrable.

Bona et al. [10,11] analyzed a family of Boussinesq systems of the form

$$\begin{aligned} w_t + u_x + (uw)_x + \alpha u_{xxx} - \beta w_{xxt} &= 0\,, \\ u_t + w_x + uu_x + \gamma w_{xxx} - \delta u_{xxt} &= 0\,, \end{aligned} \quad (14)$$

which follow from the Euler equations as first-order approximations in the parameters $\varepsilon_1 = a/h \ll 1$, $\varepsilon_2 = h^2/\lambda^2 \ll 1$, where the Stokes number, $S = \varepsilon_1/\varepsilon_2 = a\lambda^2/h^3 \approx 1$.

In (14) $w(x,t)$ is the non-dimensional deviation of the water surface from its undisturbed position; $u(x,t)$ is the non-dimensional horizontal velocity field at a height θh (with $0 \le \theta \le 1$) above the flat bottom of the water. The constant parameters α through δ in (14) satisfy the following consistency conditions: $\alpha + \beta = \frac{1}{2}(\theta^2 - \frac{1}{3})$ and $\gamma + \delta = \frac{1}{2}(1 - \theta^2) \ge 0$. Solitary wave solutions of various special cases of (14) have been computed by Chen [18].

Boussinesq systems arise when modeling the propagation of long-crested waves on large bodies of water (such as large lakes or the ocean). The Boussinesq family (14) encompasses many systems that appeared in the literature. Special cases and properties of well-posedness of (14) are addressed by Bona et al. [10,11].

1D Shallow Water Wave Equation

The so-called one-dimensional (1D) shallow water wave equation,

$$v_{xxt} + \alpha v v_t - v_t - v_x + \beta v_x \int_\infty^x v_t(y,t)\,dy = 0\,, \quad (15)$$

can be derived from the classical shallow water wave theory (see Sect. "Shallow Water Wave Equations of Geophysical Fluid Dynamics") in the Boussinesq approximation. In that approximation one assumes that vertical variations of the static density, ρ_0, are neglected, except the buoyancy term proportional to $d\rho_0/dz$, which is, in fact, responsible for the existence of solitary waves. The integral term in (15) can be removed by introducing the potential u. Indeed, setting $v = u_x$, Eq. (15) can be written as

$$u_{xxxt} + \alpha u_x u_{xt} - u_{xt} - u_{xx} + \beta u_{xx} u_t = 0\,. \quad (16)$$

The equation is completely integrable and can be solved with the IST if and only if either $\alpha = \beta$ [41] or $\alpha = 2\beta$ [3]. When $\alpha = \beta$, Eq. (16) can be integrated with respect to x and thus replaced by

$$u_{xxt} + \alpha u_x u_t - u_t - u_x = 0\,. \tag{17}$$

Closed-form solutions of (15), and in particular of (17), have been computed by Clarkson and Mansfield [19].

The Camassa–Holm Equation

The CH equation, named after Camassa and Holm [16,17],

$$u_t + 2\kappa u_x + 3uu_x - \alpha^2 u_{xxt} + \gamma u_{xxx} - 2\alpha^2 u_x u_{xx} - \alpha^2 u u_{xxx} = 0\,, \tag{18}$$

also models waves in shallow water. In (18), u is the fluid velocity in the x-direction or, equivalently, the height of the water's free surface above a flat bottom, and κ, γ and α are constants. Retaining only the first four terms in (18) gives the BBM Eq. (7). Setting $\alpha = 0$ reduces (18) to the KdV equation.

The CH equation admits solitary wave solutions, but in contrast to the hump-type solutions of the KdV and Boussinesq equations, they are implicit in nature (see, e. g., [44]). In the limit $\kappa \to 0$, Eq. (18) with $\gamma = 0, \alpha = 1$ has a cusp-type solution of the form $u(x,t) = c\exp(-|x - ct - x_0|)$. The solution is called a peakon since it has a peak (or corner) where the first derivatives are discontinuous. The solution travels at speed $c > 0$ which equals the height of the peakon.

The Kadomtsev–Petviashvili Equation

In their 1970 study [46] of the stability of line solitons, Kadomtsev and Petviashvili (KP) derived a 2D-generalization of the KdV equation which now bears their name. In dimensionless variables, the KP equation is

$$(u_t + \alpha u u_x + u_{xxx})_x + \sigma^2 u_{yy} = 0\,, \tag{19}$$

where y is the transverse direction. In the derivation of the KP equation, one assumes that the scale of variation in the y-direction (along the crest of the wave as shown in Fig. 4) is much longer than the wavelength along the x-direction.

The solitary wave and periodic (cnoidal) solutions of (19) are, respectively, given by

$$u(x,t) = \frac{k\omega - 4k^4 + \sigma^2 l^2}{\alpha k^2} + \frac{12k^2}{\alpha}\operatorname{sech}^2(kx + ly - \omega t + \delta), \tag{20}$$

Shallow Water Waves and Solitary Waves, Figure 6
Periodic plane waves in shallow water, off the coast of Lima, Peru. Photograph courtesy of A. Segur

and

$$u(x,t) = \frac{k\omega - 4k^4(2m-1) - \sigma^2 l^2}{\alpha k^2} + \frac{12k^2 m}{\alpha}\operatorname{cn}^2(kx + ly - \omega t + \delta; m)\,. \tag{21}$$

As shown in Fig. 6, near a flat beach the periodic waves appear as long, quasilinear stripes with a cn-squared cross section. Such waves are typically generated by wind and tides.

The equation with $\sigma^2 = -1$ is referred to as KP1, whereas (19) with $\sigma^2 = 1$ is called KP2, which describes shallow water waves [67]. Both KP1 and KP2 are completely integrable equations but their solution structures are fundamentally different (see, e. g., [65], pp. 489–490).

Shallow Water Wave Equations of Geophysical Fluid Dynamics

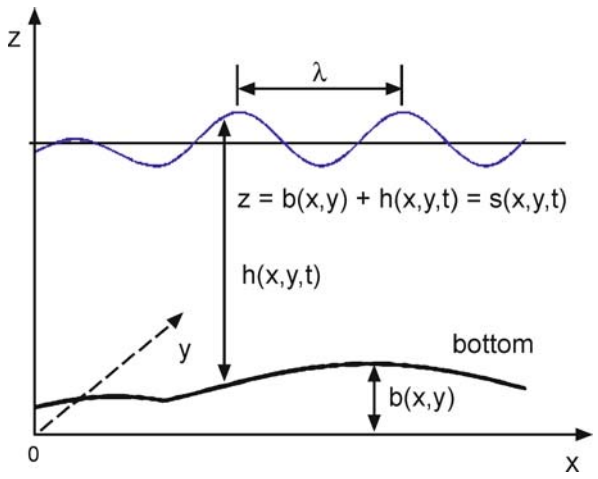
The shallow water equations used in geophysical fluid dynamics are based on the assumption $D/L \ll 1$, where D and L are characteristic values for the vertical and horizontal length scales of motion. For example, D could be the average depth of a layer of fluid (or the entire fluid) and L could be the wavelength of the wave.

The geophysical fluid dynamics community (see, e. g., [55,72,73]) uses the following 2D shallow water equations,

$$u_t + uu_x + vu_y + gh_x - 2\Omega v = -gb_x\,, \tag{22}$$

$$v_t + uv_x + vv_y + gh_y + 2\Omega u = -gb_y, \tag{23}$$

$$h_t + (hu)_x + (hv)_y = 0, \tag{24}$$



Shallow Water Waves and Solitary Waves, Figure 7
Setup for the geophysical shallow water wave model

to describe water flows with a free surface under the influence of gravity (with gravitational acceleration g) and the Coriolis force due to the earth's rotation (with angular velocity Ω.) As usual, $\mathbf{u} = (u, v)$ denotes the horizontal velocity of the fluid and $h(x, y, t)$ is its depth. As shown in Fig. 7, $h(x, y, t)$ is the distance between the free surface $z = s(x, y, t)$ and the variable bottom $b(x, y)$. Hence, $s(x, y, t) = b(x, y) + h(x, y, t)$. Eqs. (22) and (23) express the horizontal momentum-balance; (24) expresses conservation of mass. Note that the vertical component of the fluid velocity has been eliminated from the dynamics and that the number of independent variables has been reduced by one. Indeed, z no longer explicitly appears in (22)–(24), where u, v, and h only depend on x, y, and t.

A shortcoming of the model is that it does not take into account the density stratification which is present in the atmosphere (as well as in the ocean). Nonetheless, (22)–(24) are commonly used by atmospheric scientists to model flow of air at low speed.

More sophisticated models treat the ocean or atmosphere as a stack of layers with variable thickness. Within each layer, the density is either assumed to be uniform or may vary horizontally due to temperature gradients. For example, Lavoie's rotating shallow water wave equations (see [24]),

$$\boldsymbol{u}_t + (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} + 2\boldsymbol{\Omega} \times \boldsymbol{u} = -\nabla(h\theta) + \tfrac{1}{2} h \nabla\theta \,, \qquad (25)$$

$$h_t + \nabla \cdot (h\boldsymbol{u}) = 0 \,, \qquad (26)$$

$$\theta_t + \boldsymbol{u} \cdot (\nabla\theta) = 0 \,, \qquad (27)$$

consider only one active layer with layer depth $h(x, y, t)$, but take into account the forcing due to a horizontally varying potential temperature field $\theta(x, y, t)$. Vector $\boldsymbol{u} = u(x, y, t)\boldsymbol{i} + v(x, y, t)\boldsymbol{j}$ denotes the fluid velocity and $\boldsymbol{\Omega} = \Omega\boldsymbol{k}$ is the angular velocity vector of the Earth's rotation. $\nabla = \frac{\partial}{\partial x}\boldsymbol{i} + \frac{\partial}{\partial y}\boldsymbol{j}$ is the gradient operator, and $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$ are unit vectors along the x, y, and z-axes.

Lavoie's equations are part of a family of multi-layer models proposed by Ripa [61] to study, for example, the effects of solar heating, fresh water fluxes, and wind stresses on the upper equatorial ocean. A study of the validity of various layered models has been presented by Ripa [62]. The more sophisticated the models become the harder they become to treat with analytic methods so one has to apply numerical methods. Numerical aspects of various shallow water models in atmospheric research and beyond are discussed in e. g. [48,75,76].

Computation of Solitary Wave Solutions

As shown in Sect. "Completely Integrable Shallow Water Wave Equations", solitary wave solutions of the KdV and Boussinesq equations (and like PDEs), can be expressed as polynomials of the hyperbolic secant (sech) or tangent (tanh) functions, whereas their simplest period solutions involve the Jacobi elliptic cosine (cn) function.

There are several methods to compute exact, analytic expressions for solitary and periodic wave solutions of nonlinear PDEs. Two straightforward methods, namely the direct integration method and the tanh-method, will be discussed. Both methods seek traveling wave solutions. By working in a traveling frame of reference the PDE is replaced by an ordinary differential equation (ODE) for which one seeks closed-form solutions in terms of special functions.

In the terminology of dynamical systems, the solitary wave solutions correspond to heteroclinic or homoclinic trajectories in the phase plane of a first-order dynamical system corresponding to the underlying ODE (see, e. g., [6]). The periodic solutions are bounded in the phase plane by these special trajectories, which correspond to the limit of infinite period and modulus one.

Other more powerful methods, such as the aforementioned IST and Hirota's method (see, e. g., [40]) deal with the PDE directly. These methods allow one to compute closed-form expressions of soliton solutions (in particular, solitary wave solutions) addressed elsewhere in the encyclopedia.

Direct Integration Method

Exact expressions for solitary wave solutions can be obtained by direct integration. The steps are illustrated for

the KdV equation given in (2). Assuming that the wave travels to the right at speed $v = \omega/k$, Eq. (2) can be put into a traveling frame of reference with independent variable $\xi = k(x - vt - x_0)$. This reduces (2) to an ODE, $-v\phi' + \alpha\phi\phi' + k^2\phi''' = 0$, for $\phi(\xi) = u(x, t)$. A first integration with respect to ξ yields

$$-v\phi + \frac{\alpha}{2}\phi^2 + k^2\phi'' = A\,, \tag{28}$$

where A is a constant of integration. Multiplication of (28) by ϕ', followed by a second integration with respect to ξ, yields

$$-\frac{v}{2}\phi^2 + \frac{\alpha}{6}\phi^3 + \frac{k^2}{2}\phi'^2 = A\phi + B\,, \tag{29}$$

where B is an integration constant. Separation of variables and integration then leads to

$$\int_{\phi_0}^{\phi} \frac{d\phi}{\sqrt{a\phi^2 - b\phi^3 + \tilde{A}\phi + \tilde{B}}} = \pm\int_{\xi_0}^{\xi} d\xi\,, \tag{30}$$

where $a = v/k^2$, $b = \alpha/3k^2$, $\tilde{A} = 2A/k^2$, and $\tilde{B} = 2B/k^2$.

The evaluation of the elliptic integral in (30) depends on the relationship between the roots of the function $f(\phi) = a\phi^2 - b\phi^3 + \tilde{A}\phi + \tilde{B}$. In turn, the nature of the roots depends on the choice of $\tilde{A}$ and $\tilde{B}$. Two cases lead to physically relevant solutions.

Case 1: If the three roots are real and distinct, then the integral can be expressed in terms of the inverse of the cn function (see, e. g., [25] for details). This leads to the cnoidal wave solution given in (6).

Case 2: If the three roots are real and (only) two of them coincide, then the tanh-squared solution follows. This happens when $\tilde{A} = \tilde{B} = 0$. Integrating both sides of (30) then gives

$$\begin{aligned}\int_{\phi_0}^{\phi} \frac{d\phi}{\phi\sqrt{a - b\phi}} &= -\frac{2}{\sqrt{a}}\,\mathrm{Arctanh}\left[\frac{\sqrt{a - b\phi}}{\sqrt{a}}\right] + C \\ &= \pm(\xi - \xi_0)\,.\end{aligned} \tag{31}$$

where, without loss of generality, C and ξ_0 can be set to zero. Solving (31) for ϕ yields

$$\phi(\xi) = \frac{a}{b}\left(1 - \tanh^2\left(\frac{\sqrt{a}}{2}\xi\right)\right) = \frac{a}{b}\,\mathrm{sech}^2\left(\frac{\sqrt{a}}{2}\xi\right). \tag{32}$$

Returning to the original variables, one gets

$$\phi(\xi) = \frac{3v}{\alpha}\,\mathrm{sech}^2\left(\frac{\sqrt{v}}{2k}\xi\right), \tag{33}$$

or

$$u(x, t) = \frac{3v}{\alpha}\,\mathrm{sech}^2\left(\frac{\sqrt{v}}{2}(x - vt - x_0)\right), \tag{34}$$

where v is arbitrary. Setting $v = \omega/k = 4k^2$, where k is arbitrary, and $\delta = -kx_0$, one can verify that (34) matches (5).

The Tanh Method

If one is only interested in tanh- or sech-type solutions, one can circumvent explicit integration (often involving elliptic integrals) and apply the so-called tanh-method. A detailed description of the method has been given by Baldwin et al. [5]. The method has been fully implemented in *Mathematica*, a popular symbolic manipulation program, and successfully applied to many nonlinear differential equations from soliton theory and beyond.

The tanh-method is based on the following observation: all derivatives of the tanh function can be expressed as polynomials in tanh. Indeed, using the identity $\cosh^2\xi - \sinh^2\xi = 1$ one computes $\tanh'\xi = \mathrm{sech}^2\xi = 1 - \tanh^2\xi$, $\tanh''\xi = -2\tanh\xi + 2\tanh^3\xi$, etc. Therefore, all derivatives of $T(\xi) = \tanh\xi$ are polynomials in T. For example, $T' = 1 - T^2$, $T'' = -2T + 2T^3$, and $T''' = -2T + 8T^2 - 6T^4$.

By applying the chain rule, the PDE in $u(x, t)$ is then transformed into an ODE for $U(T)$ where $T = \tanh\xi = \tanh(kx - \omega t + \delta)$ is the new independent variable. Since all derivatives of T are polynomials of T, the resulting ODE has polynomial coefficients in T. It is therefore natural to seek a polynomial solution of the ODE. The problem thus becomes algebraic. Indeed, after computing the degree of the polynomial solution, one finds its unknown coefficients by solving a nonlinear algebraic system.

The method is illustrated using (2). Applying the chain rule (repeatedly), the terms of (2) become $u_t = -\omega(1 - T^2)U'$, $u_x = k(1 - T^2)U'$, and

$$\begin{aligned}u_{xxx} = k^3(1 - T^2)\big[&-2(1 - 3T^2)U' \\ &- 6T(1 - T^2)U'' + (1 - T^2)^2U'''\big]\,,\end{aligned} \tag{35}$$

where $U(T) = U(\tanh(kx - \omega t + \delta)) = u(x, t)$, $U' = dU/dT$, etc.

Substitution into (2) and cancellation of a common $1 - T^2$ factor yields

$$\begin{aligned}&-\omega U' + \alpha k U U' - 2k^3(1 - 3T^2)U' \\ &\quad - 6k^3T(1 - T^2)U'' + k^3(1 - T^2)^2U''' = 0\,.\end{aligned} \tag{36}$$

This ODE for $U(T)$ has polynomial coefficients in T. One therefore seeks a polynomial solution

$$U(T) = \sum_{n=0}^{N} a_n T^n \,, \tag{37}$$

where the integer exponent N and the coefficients a_n must be computed.

Substituting T^N into (36) and balancing the highest powers in T gives $N = 2$. Then, substituting

$$U(T) = a_0 + a_1 T + a_2 T^2 \tag{38}$$

into (36), and equating to zero the coefficients of the various power terms in T, yields

$$\begin{aligned} a_1(\alpha a_2 + 2k^2) &= 0 \,, \\ \alpha a_2 + 12k^2 &= 0 \,, \\ a_1(\alpha k a_0 - 2k^3 - \omega) &= 0 \,, \\ \alpha k a_1^2 + 2\alpha k a_0 a_2 - 16k^3 a_2 - 2\omega a_2 &= 0 \,. \end{aligned} \tag{39}$$

The unique solution of this nonlinear system for the unknowns a_0, a_1 and a_2 is

$$a_0 = \frac{8k^3 + \omega}{\alpha k}, \;\; a_1 = 0, \;\; a_2 = -\frac{12k^2}{\alpha} \,. \tag{40}$$

Finally, substituting (40) into (38) and using $T = \tanh(kx - \omega t + \delta)$ yields (4).

The solitary wave solutions and cnoidal wave solutions presented in Sect. "Completely Integrable Shallow Water Wave Equations" have been automatically computed with a *Mathematica* package [5] that implements the tanh-method and variants.

A review of numerical methods to compute solitary waves of arbitrary amplitude can be found in Vanden-Broeck [74].

Water Wave Experiments and Observations

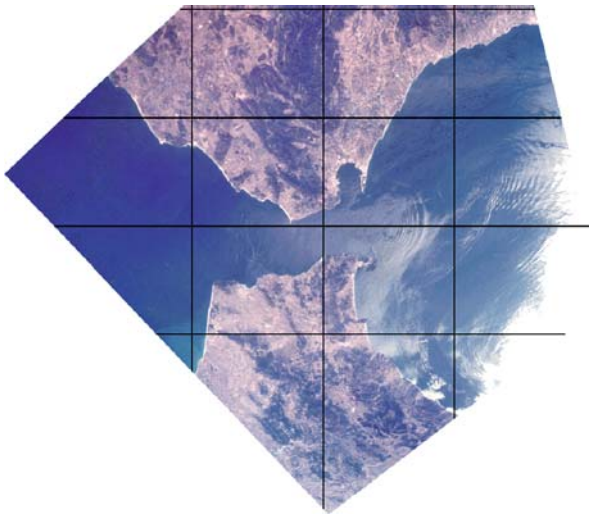
Through a series of experiments in a hydrodynamic tank, Hammack investigated the validity of the BBM equation [35] and KdV equation as models for long waves in shallow water [36,37,38] and long internal waves [69]. Their research addressed the question: Would an initial displacement of water, as it propagates forward, eventually evolve in a train of localized solitary waves (solitons) and an oscillatory tail as predicted by the KdV equation? Based on the experimental data, they concluded that (i) the KdV dynamics only occurs if the waves travel over a long distance, (ii) a substantial amount of water must be initially displaced (by a piston) to produce a soliton train, (iii) the water volume of the initial wave determines the shape of the leading wave in the wave train, and (iv) the initial direction of displacement (upward or downward piston motion) determines what happens later. Quickly raising the piston causes a train of solitons to emerge; quickly lowering it causes all wave energy to distribute into the oscillatory tail, as predicted by the theory.

Several other researchers have tested the validity of the KdV equation and variants in laboratory experiments (see, e. g., [39,60]). Bona et al. [9] give an in-depth evaluation of the BBM Eq. (7) with and without dissipative term u_{xx}. Their study includes (i) a numerical scheme with error estimates, (ii) a convergence test of the computer code, (iii) a comparison between the predictions of the theoretical model and the results of laboratory experiments. The authors note that it is important to include dissipative effects to obtain reasonable agreement between the forecast of the model and the empirical results.

Water tank experiments in conjunction with the analysis of actual data, allows researchers to judge whether or not the KdV equation can be used to model the dynamics of tsunamis (see [67]). Tsunami research intensified after the December 2004 tsunami devastated large coastal regions of India, Indonesia, Sri Lanka, and Thailand, and killed nearly 300,000 people.

Apart from shallow water waves near beaches, the KdV equation and its solitary wave solution also apply to internal waves in the ocean. Internal solitary waves in the open ocean are slow waves of large amplitude that travel at the interface of stratified layers of different density. Stratification based on density differences is primarily due to variations in temperature or concentration (e. g. due to salinity gradients). For example, absorption of solar radiation creates a near surface thin layer of warmer water (of lower density) above a thicker layer or colder, denser water. The smaller the density change, the lower the wave frequency, and the slower the propagation speed. If the upper layer is thinner than the lower one, then the internal wave is a wave of depression causing a downward displacement of the fluid interface.

Internal solitary waves are ubiquitous in stratified waters, in particular, whenever strong tidal currents occur near irregular topography. Such waves have been studied since the 1960s. An early, well-documented case deals with internal waves in the Andaman Sea, where Perry and Schimke [58] found groups of internal waves up to 80 m high and 2000 m long on the main thermocline at 500 m in 1500 m deep water. Their measurements were confirmed by Osborne and Burch [53] who showed that internal waves in the Andaman Sea are generated by tidal flows and can travel over hundreds of kilometers.



Shallow Water Waves and Solitary Waves, Figure 8
Three solitary wave packets generated by internal waves from sills in the Strait of Gibraltar. Original image STS41G-34-81 courtesy of the Earth Sciences and Image Analysis Laboratory, NASA Johnson Space Center (http://eol.jsc.nasa.gov). Ortho-rectified, color adjusted photograph courtesy of Global Ocean Associates

Strong internal waves can affect biological life and interfere with underwater navigation. Understanding the behavior of internal waves can aid in the design of offshore production facilities for oil and natural gas.

The near-surface current associated with the internal wave locally modulates the height of the water surface. Hence, the internal wave leaves a "signature" or "footprint" at the sea surface in the form of a packet of solitary waves (sometimes called current rips or tide rips). These visual manifestations appear as long, quasilinear stripes in satellite imagery or photographs taken during space flights. Over 50 case studies and hundreds of images of oceanic internal waves can be found in "An Atlas of Internal Solitary-like Waves and Their Properties" [42].

Figure 8 shows a photograph of three solitary waves packets which are the surface signature of internal waves in the Strait of Gibraltar. The photograph was taken from the Space Shuttle on October 11, 1984. Spain is to the North, Morocco to the South. Alternate solitary wave packets move toward the northeast or the southeast. The amplitude of these waves is of the order of 50 m; their wavelength is in the range of 500–2000 m. The separation between the packets is approximately 30 km. Waves of longer wavelengths and higher amplitudes have traveled the furthest. The number of oscillations within each packet increases as time goes on. Solitary wave packets can reach 200 km into the Western Mediterranean sea and live for more than two days before dissipating. A in-depth study of solitary waves in the Strait of Gibraltar can be found in Farmer and Armi [28].

The KdV model is applicable to stratified fluids with two layers and internal solitary waves if (i) the ratio of the amplitude a to the upper layer depth h is small, and (ii) the wavelength λ is long compared with the upper layer depth. More precisely, $a/h = O(h^2/\lambda^2) \ll 1$. A detailed discussion of internal solitary waves and additional references can be found in Garrett and Munk [31], Grimshaw [32,33,34], Helfrich and Melville [39], Apel et al. [4], and Pelinovsky et al. [56]. The last three papers discuss a variety of other theoretical models including the extended KdV equation (also known as Gardner's equation or combined KdV-modified KdV equation) which contains both quadratic and cubic nonlinearities. Solitary wave solutions of the extended KdV equation can be found in Scott ([65], p. 856), Helfrich and Melville [39], and Apel et al. [4]. A review of laboratory experiments with internal solitary waves was published by Ostrovsky and Stepanyants [54].

As discussed in the review paper by Staquet and Sommeria [70], internal gravity waves also occur in the atmosphere, where they are often caused by wind blowing over topography and cumulus convective clouds. Internal gravity waves reveal themselves as unusual cloud patterns, which are the counterpart of the solitary wave packets on the ocean's surface.

Future Directions

For many shallow water wave applications, the full Euler equations are too complex to work with. Instead, various approximate models have been proposed. Arguably, the most famous shallow water wave equations are the KdV and Boussinesq equations.

The KdV equation was originally derived to describe shallow water waves in a rectangular channel. Surprisingly, the equation also models ion-acoustic waves and magneto-hydrodynamic waves in plasmas, waves in elastic rods, mid-latitude and equatorial planetary waves, acoustic waves on a crystal lattice, and more (see, e. g., [64,65,66]). The KdV equation has played a pivotal role in the development of the Inverse Scattering Transform and soliton theory, both of which had a lasting impact on twentieth-century mathematical physics.

Historically, the classical Boussinesq equation was derived to describe the propagation of shallow water waves in a canal. Boussinesq systems arise when modeling the propagation of long-crested waves on large bodies of water (e. g. large lakes or the ocean). As Bona et al. [10] point

out, a plethora of formally-equivalent Boussinesq systems can be derived. Yet, such systems may have vastly different mathematical properties. The study of the well-posedness of the nonlinear models is of paramount importance and is the subject of ongoing research.

Shallow water wave theory allows one to adequately model waves in canals, surface waves near beaches, and internal waves in the ocean (see [4]). Due to their widespread occurrence in the ocean (see [42]), solitary waves and "solitary wave packets" (solitons) are of interest to oceanographers and geophysicists. The (periodic) cnoidal wave solutions are used by coastal engineers in studies of sediment movement, erosion of sandy beaches, interaction of waves with piers and other coastal structures.

Apart from their physical relevance, the knowledge of solitary and cnoidal wave solutions of nonlinear PDEs facilitates the testing of numerical solvers and aids in stability analysis.

Shallow water wave models are widely used in atmospheric science as a paradigm for geophysical fluid motions. They model, for example, inertia-gravity waves with fast time scale dynamics, and wave-vortex interactions and Rossby waves associated with slow advective-timescale dynamics.

This article has reviewed commonly used shallow water wave models, with the hope of bridging two research communities: one that focuses on nonlinear equations with dispersive effects; the other on nonlinear hyperbolic equations without dispersive terms. Of common concern are the testing of the theoretical models on measured data and further validation of the equations with numerical simulations and laboratory experiments. A fusion of the expertise of both communities might advance research on water waves and help to answer open questions about wave breaking, instability, vorticity, and turbulence. Of paramount importance is the prevention of natural disasters, ecological ravage, and damage to man-made structures due to a better understanding of the dynamics of tsunamis, steep waves, strong internal waves, rips, tidal currents, and storm surges.

Acknowledgments

This material is based upon work supported by the National Science Foundation (NSF) of the USA under Award No. CCF-0830783. The author is grateful to Harvey Segur, Jerry Bona, Michel Grundland, and Douglas Poole for advise and comments. William Navidi, Benoit Huard, Anna Segur, Katherine Socha, Christopher Jackson, Michel Peyrard, and Christopher Eilbeck are thanked for help with figures and photographs.

Bibliography

Primary Literature

1. Ablowitz MJ, Clarkson PA (1991) Solitons, nonlinear evolution equations and inverse scattering. Cambridge University Press, Cambridge
2. Ablowitz MJ, Segur H (1981) Solitons and the inverse scattering transform. SIAM, Philadelphia, Pennsylvania
3. Ablowitz MJ et al (1974) The inverse scattering transform: Fourier analysis for nonlinear problems. Stud Appl Math 53:249–315
4. Apel JR et al (2007) Internal solitons in the ocean and their effect on underwater sound. J Acoust Soc Am 121:695–722
5. Baldwin D et al (2004) Symbolic computation of exact solutions expressible in hyperbolic and elliptic functions for nonlinear PDEs. J Symb Comp 37:669–705
6. Balmforth NJ (1995) Solitary waves and homoclinic orbits. Annu Rev Fluid Mech 27:335–373
7. Benjamin TB et al (1972) Model equations for long waves in nonlinear dispersive systems. Phil Trans Roy Soc London Ser A 272:47–78
8. Bona JL, Chen H (1999) Comparison of model equations for small-amplitude long waves. Nonl Anal 38:625–647
9. Bona JL et al (1981) An evaluation of a model equation for water waves. Phil Trans Roy Soc London Ser A 302:457–510
10. Bona JL et al (2002) Boussinesq equations and other systems for small-amplitude long waves in nonlinear dispersive media. I: Derivation and linear theory. J Nonl Sci 12:283–318
11. Bona JL et al (2004) Boussinesq equations and other systems for small-amplitude long waves in nonlinear dispersive media. II: The nonlinear theory. Nonlinearity 17:925–952
12. Boussinesq J (1871) Théorie de l'intumescence liquide appelée onde solitaire ou de translation, se propageant dans un canal rectangulaire. C R Acad Sci Paris 72:755–759
13. Boussinesq J (1872) Théorie des ondes et des remous qui se propagent le long d'un canal rectangulaire horizontal, en communiquant au liquide contenu dans ce canal des vitesses sensiblement pareilles de la surface au fond. J Math Pures Appl 17:55–108
14. Boussinesq J (1877) Essai sur la théorie des eaux courantes. Académie des Sciences de l'Institut de France, Mémoires présentés par divers savants (ser 2) 23:1–680
15. Bullough RK (1988) "The wave" "par excellence", the solitary, progressive great wave of equilibrium of the fluid–an early history of the solitary wave. In: Lakshmanan M (ed) Solitons: introduction and applications. Springer, Berlin, pp 7–42
16. Camassa R, Holm D (1993) An integrable shallow water equation with peakon solitons. Phys Rev Lett 71:1661–1664
17. Camassa R, Holm D (1994) An new integrable shallow water equation. Adv Appl Mech 31:1–33
18. Chen M (1998) Exact solutions of various Boussinesq systems. Appl Math Lett 11:45–49
19. Clarkson PA, Mansfield EL (1994) On a shallow water wave equation. Nonlinearity 7:975–1000
20. Craik ADD (2004) The origins of water wave theory. Annu Rev Fluid Mech 36:1–28

21. Craik ADD (2005) George Gabriel Stokes and water wave theory. Annu Rev Fluid Mech 37:23–42
22. Darrigol O (2003) The spirited horse, the engineer, and the mathematician: Water waves in nineteenth-century hydrodynamics. Arch Hist Exact Sci 58:21–95
23. Dauxois T, Peyrard M (2004) Physics of solitons. Cambridge University Press, Cambridge, UK. Transl of Peyrard M, Dauxois T (2004) Physique des solitons. Savoirs Actuels, EDP Sciences, CNRS Editions
24. Dellar P (2003) Common Hamiltonian structure of the shallow water equations with horizontal temperature gradients and magnetic fields. Phys Fluids 15:292–297
25. Drazin PG, Johnson RS (1989) Solitons: an introduction. Cambridge University Press, Cambridge, UK
26. El GA (2007) Korteweg–de Vries equation: solitons and undular bores. In: Grimshaw RHJ (ed) Solitary waves in fluids. WIT Press, Boston, pp 19–53
27. Emmerson GS (1977) John Scott Russell: A great Victorian engineer and naval architect. John Murray Publishers, London
28. Farmer DM, Armi L (1988) The flow of Mediterranean water through the Strait of Gibraltar. Prog Oceanogr 21:1–105
29. Fermi F et al (1955) Studies of nonlinear problems I. Los Alamos Sci Lab Rep LA-1940. Reproduced in: AC Newell (ed) (1974) Nonlinear wave motion. AMS, Providence, RI
30. Filippov AT (2000) The versatile soliton. Birkhäuser-Verlag, Basel
31. Garrett C, Munk W (1979) Internal waves in the ocean. Annu Rev Fluid Mech 11:339–369
32. Grimshaw R (1997) Internal solitary waves. In: Liu PLF (ed) Advances in coastal and ocean engineering, vol III. World Scientific, Singapore, pp 1–30
33. Grimshaw R (2001) Internal solitary waves. In: Grimshaw R (ed) Environmental stratified flows. Kluwer, Boston, pp 1–28
34. Grimshaw R (2005) Korteweg–de Vries equation. In: Grimshaw R (ed) Nonlinear waves in fluids: Recent advances and modern applications. Springer, Berlin, pp 1–28
35. Hammack JL (1973) A note on tsunamis: their generation and propagation in an ocean of uniform depth. J Fluid Mech 60:769–800
36. Hammack JL, Segur H (1974) The Korteweg–de Vries equation and water waves, part 2. Comparison with experiments. J Fluid Mech 65:289–314
37. Hammack JL, Segur H (1978a) The Korteweg–de Vries equation and water waves, part 3. Oscillatory waves. J Fluid Mech 84:337–358
38. Hammack JL, Segur H (1978b) Modelling criteria for long water waves. J Fluid Mech 84:359–373
39. Helfrich KR, Melville WK (2006) Long nonlinear internal waves. Annu Rev Fluid Mech 38:395–425
40. Hirota R (2004) The direct method in soliton theory. Cambridge University Press, Cambridge, UK
41. Hirota R, Satsuma J (1976) N-soliton solutions of model equations for shallow water waves. J Phys Soc Jpn 40:611–612
42. Jackson CR (2004) An atlas of internal solitary-like waves and their properties, 2nd edn. Global Ocean Associates, Alexandria, VA
43. Jeffrey A (1978) Nonlinear wave propagation. Z Ang Math Mech (ZAMM) 58:T38–T56
44. Johnson RS (2003) On solutions of the Camassa–Holm equation. Proc Roy Soc London Ser A 459:1687–1708
45. Joseph RI, Egri R (1977) Another possible model equation for long waves in nonlinear dispersive systems. Phys Lett A 61:429–432
46. Kadomtsev BB, Petviashvili VI (1970) On the stability of solitary waves in weakly dispersive media. Sov Phys Doklady 15:539–541
47. Korteweg DJ, de Vries G (1895) On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. Philos Mag (Ser 5) 39:422–443
48. LeVeque RJ (2002) Finite volume methods for hyperbolic problems. Cambridge University Press, Cambridge, UK
49. Madsen PA, Schäffer HA (1999) A review of Boussinesq-type equations for surface gravity waves. In: Liu PLF (ed) Advances in coastal and ocean engineering, vol V. World Scientific, Singapore, pp 1–95
50. McKean HP (1981) Boussinesq's equation on the circle. Comm Pure Appl Math 34:599–691
51. Miles JW (1981) The Korteweg–de Vries equation: a historical essay. J Fluid Mech 106:131–147
52. Miura RM (1976) The Korteweg–de Vries equation: a survey of results. SIAM Rev 18:412–559
53. Osborne AR, Burch TL (1980) Internal solitons in the Andaman sea. Science 208:451–460
54. Ostrovsky LA, Stepanyants YA (2005) Internal solitons in laboratory experiments: Comparison with theoretical models. Chaos 15:037111
55. Pedlosky J (1987) Geophysical fluid dynamics, 2nd edn. Springer, Berlin
56. Pelinovsky E et al (2007) In: Grimshaw RHJ (ed) Solitary waves in fluids. WIT Press, Boston, pp 85–110
57. Peregrine DH (1966) Calculations of the development of an undular bore. J Fluid Mech 25:321–330
58. Perry RB, Schimke GR (1965) Large amplitude internal waves observed off the north-west coast of Sumatra. J Geophys Res 70:2319–2324
59. Rayleigh L (1876) On waves. Philos Mag 1:257–279
60. Remoissenet M (1999) Waves called solitons: Concepts and experiments, 3rd edn. Springer, Berlin
61. Ripa P (1993) Conservation laws for primitive equations models with inhomogeneous layers. Geophys Astrophys Fluid Dyn 70:85–111
62. Ripa P (1999) On the validity of layered models of ocean dynamics and thermodynamics with reduced vertical resolution. Dyn Atmos Oceans Fluid Dyn 29:1–40
63. Russell JS (1844) Report on waves, 14th meeting of the British Association for the Advancement of Science. John Murray, London, pp 311–390
64. Scott AC (2003) Nonlinear science: Emergence and dynamics of coherent structures. Oxford University Press, Oxford, UK
65. Scott AC (ed) (2005) Encyclopedia of Nonlinear science. Routledge, New York
66. Scott AC et al (1973) The soliton: a new concept in applied science. Proc IEEE 61:1443–1483
67. Segur H (2007a) Waves in shallow water, with emphasis on the tsunami of 2004. In: Kundu A (ed) Tsunami and nonlinear waves. Springer, Berlin
68. Segur H (2007b) Integrable models of waves in shallow water. In: Pinski M, Birnir B (eds) Probability, geometry and integrable systems. Cambridge University Press, Cambridge, UK
69. Segur H, Hammack JL (1982) Soliton models of long internal waves. J Fluid Mech 118:285–304

70. Staquet C, Sommeria J (2002) Internal gravity waves: From instabilities to turbulence. Annu Rev Fluid Mech 34:559–593
71. Stokes GG (1847) On the theory of oscillatory waves. Trans Camb Phil Soc 8:441–455
72. Toro EF (2001) Shock-capturing methods for free-surface shallow flows. Wiley, New York
73. Vallis GK (2006) Atmospheric and oceanic fluid dynamics: fundamentals and large scale circulation. Cambridge University Press, Cambridge, UK
74. Vanden-Broeck J-M (2007) Solitary waves in water: numerical methods and results. In: Grimshaw RHJ (ed) Solitary waves in fluids. WIT Press, Boston, pp 55–84
75. Vreugdenhil CB (1994) Numerical methods for shallow-Water flow. Springer, Berlin
76. Weiyan T (1992) Shallow water hydrodynamics: mathematical theory and numerical solution for two-dimensional systems of shallow water equations. Elsevier, Amsterdam
77. Zabusky NJ (2005) Fermi–Pasta–Ulam, solitons and the fabric of nonlinear and computational science: History, synergetics, and visiometrics. Chaos 15:015102
78. Zabusky NJ, Kruskal MD (1965) Interaction of 'solitons' in a collisionless plasma and the recurrence of initial states. Phys Rev Lett 15:240–243

Books and Reviews

Boyd JP (1998) Weakly nonlinear solitary waves and beyond-all-order asymtotics. Kluwer, Dordrecht
Calogero F, Degasperis A (1982) Spectral transform and solitons I. North Holland, Amsterdam
Dickey LA (2003) Soliton equations and Hamiltonian systems, 2nd edn. World Scientific, Singapore
Dodd RK et al (1982) Solitons and nonlinear wave equations. Academic Press, London
Eilenberger G (1981) Solitons: Mathematical methods for physicists. Springer, Berlin
Faddeev LD, Takhtajan LA (1987) Hamiltonian methods in the theory of solitons. Springer, Berlin
Fordy A (ed) (1990) Soliton theory: A survey of results. Manchester University Press, Manchester
Grimshaw RHJ (ed) (2007) Solitary waves in fluids. WIT Press, Boston
Infeld E, Rowlands G (2000) Nonlinear waves, solitons, and chaos. Cambridge University Press, New York
Johnson RS (1977) A modern introduction to the mathematical theory of water waves. Cambridge University Press, Cambridge, UK
Lamb GL (1980) Elements of soliton theory. Wiley Interscience, New York
Miles JW (1980) Solitary waves. Annu Rev Fluid Mech 12:11–43
Mei CC et al (2005) Theory and applications of ocean surface waves. World Scientific, Singapore
Mitropol'sky YZ (2001) Dynamics of internal gravity waves in the ocean. Kluwer, Dordrecht
Newell AC (1983) The history of the soliton. J Appl Mech 50:1127–1137
Newell AC (1985) Solitons in mathematics and physics. SIAM, Philadelphia, PA
Makhankov VG (1990) Soliton phenomenology. Kluwer, Dordrecht, The Netherlands
Novikov SP et al (1984) Theory of solitons: The inverse scattering method. Consultants Bureau(Plenum Press), New York
Pedlosky J (2003) Waves in the ocean and atmosphere: introduction to wave dynamics. Springer, Berlin
Russell JS (1885) The wave of translation in the oceans of water, air and ether. Trübner, London
Stoker JJ (1957) Water waves. Wiley Interscience, New York
Whitham GB (1974) Linear and nonlinear waves. Wiley Interscience, New York

Signaling Games

Joel Sobel
Department of Economics, University of California, San Diego, USA

Article Outline

Glossary
Definition of the Subject
Introduction
The Model
Equilibrium
The Basic Model
Cheap Talk
Verifiable Information
Communication About Intentions
Applications
Future Directions
Acknowledgments
Bibliography

Glossary

Babbling equilibrium An equilibrium in which the sender's strategy is independent of type and the receiver's strategy is independent of signal.

Behavior strategy A strategy for an extensive-form game that specifies the probability of taking each action at each information set.

Behavioral type A player in a game who is constrained to follow a given strategy.

Cheap-talk game A signaling game in which players' preferences do not depend directly on signals.

Condition D1 An equilibrium refinement that requires out-of-equilibrium beliefs to be supported on types that have the most to gain from deviating from a fixed equilibrium.

Divinity An equilibrium refinement that requires out-of-equilibrium beliefs to place relatively more weight on types that gain more from deviating from a fixed equilibrium.

Equilibrium outcome The probability distribution over terminal nodes in a game determined by equilibrium strategy.

Handicap principle The idea that animals communicate fitness through observable characteristics that reduce fitness.

Incomplete information game A game in which players lack information about the strategy sets or payoff functions of their opponents.

Intuitive criterion An equilibrium refinement that requires out-of-equilibrium beliefs to place zero weight on types that can never gain from deviating from a fixed equilibrium outcome.

Nash equilibrium A strategy profile in a game in which each player's strategy is a best response to the equilibrium strategies of the other players.

Neologism-proof equilibrium An equilibrium that admits no self-signaling set.

Pooling equilibrium A signaling-game equilibrium in which each all sender types send the same signal with probability one.

Receiver In a signaling game, the uninformed player.

Self-signaling set A set of types C with the property that precisely types in the set C gain from inducing the best response to C relative to a fixed equilibrium.

Sender In a signaling game, the informed agent.

Separating equilibrium A signaling-game equilibrium in which sender types sent signals from disjoint subsets of the set of available signals.

Signaling game A two-player game of incomplete information in which one player is informed and the other in not. The informed player's strategy is a type-contingent message and the uninformed player's strategy is a message-continent action.

Single-crossing condition A condition that guarantees that indifferent curves from a given family of preferences cross at most one.

Spence-Mirrlees condition A differential condition that orders the slopes of level sets of a function.

Standard signaling game A signaling game in which strategy sets and payoff functions satisfy monotonicity properties.

Type In an incomplete information game, a variable that summarizes private information.

Verifiable information game A signaling game with the property that each type has a signal that can only be sent by that type.

Definition of the Subject

Signaling games refer narrowly to a class of two-player games of incomplete information in which one player is informed and the other is not. The informed player's strategy set consists of signals contingent on information and the uninformed player's strategy set consists of actions contingent on signals. More generally, a signaling game includes any strategic setting in which players can use the actions of their opponents to make inferences about hidden information. The earliest work on signaling games was Spence [72]'s model of educational signaling and Zahari [76]'s model of signaling by animals. During the 1980s researchers developed the formal model and identified conditions that permitted the selection of unique equilibrium outcomes in leading models.

Introduction

The framed degree in your doctor's office, the celebrity endorsement of a popular cosmetic, and the telephone message from an old friend are all signals. The signals are potentially valuable because they allow you to infer useful information. These signals are indirect and require interpretation. They may be subject to manipulation. The doctor's diploma tells you something about the doctor's qualifications, but knowing where and when the doctor studied does not prove that she is a good doctor. The endorsement identifies the product with a particular lifestyle, but what works for the celebrity may not work for you. Besides, the celebrity was probably paid to endorse the product and may not even use it. The phone message may tell you how to get in touch with your friend, but is unlikely to contain all of the information you need to find him – or to evaluate whether you'll meet to discuss old times or to be asked a favor. While the examples all involve signaling, the nature of the signaling is different. The doctor faces large penalties for misrepresenting her credentials. She is not required to display all of her diplomas, but it is reasonable to assume that degrees are not forged. The celebrity endorsement is costly – certainly to the manufacturer who pays for the celebrity's services and possibly to the celebrity himself, whose reputation may suffer if the product works badly. It is reasonable to assume that it is easier to obtain an endorsement of a good product, but there are also good reasons to be skeptical about the claims. In contrast, although a dishonest or misleading message may lead to a bad outcome, leaving a message is not expensive and the content of the message is not constrained by your friend's information. The theory of signaling games is a useful way to describe the essential features of all three examples.

Opportunities to send and evaluate signals arise in many common natural and economic settings. In the canonical example (due to Spence [72]), a high-ability worker invests in education to distinguish herself from less skilled workers. The potential employer observers educational attainment, but not innate skill, and infers that

a better educated worker is more highly skilled and pays a higher wage. To make this story work, there must be a reason that low-ability workers do not get the education expected of a more highly skilled worker and hence obtain a higher wage. This property follows from an assumption that the higher the ability the worker, the easier it is for her to produce a higher signal.

The same argument appears in many applications. For example, a risk-averse driver will purchase a lower cost, partial insurance contract, leaving the riskier driver to pay a higher rate for full insurance (Rothschild and Stiglitz [65] or Wilson [75]). A firm that is able to produce high-quality goods signals this by offering a warranty for the goods sold (Grossman [37]) or advertising extensively. A strong deer grows extra large antlers to show that it can survive with this handicap and to signal its fitness to potential mates [76].

Game theory provides a formal language to study how one should send and interpret signals in strategic environments. This article reviews the basic theory of signaling and discusses some applications. It does not discuss related models of screening. Kreps and Sobel [43] and Riley [64] review both signaling and screening.

Section "The Model" describes the basic model. Section "Equilibrium" defines equilibrium for the basic model. Section "The Basic Model" limits attention to a special class of signaling game. I give conditions sufficient for the existence of equilibria in which the informed agent's signal fully reveals her private information and argue that one equilibrium of this kind is prominent. The next three sections study different signaling games. Section "Cheap Talk" discusses models of costless communication. Section "Verifiable Information" discusses the implications of the assumptions that some information is verifiable. Section "Communication about Intentions" briefly discusses the possibility of signaling intentions rather than private information. Section "Applications" describes some applications and extensions of the basic model. Section "Future Directions" speculates on directions for future research.

The Model

This section describes the basic signaling game. There are two players, called S (for sender) and R (for receiver). S knows the value of some random variable t whose support is a given set T. t is called the **type** of S. The prior beliefs of R are given by a probability distribution $\pi(\cdot)$ over T; these beliefs are common knowledge. When T is finite, $\pi(t)$ is the prior probability that the sender's type is t. When T is uncountably infinite, $\pi(\cdot)$ is a density function. Player S learns t and sends to R a signal s drawn from some set M. Player R receives this signal, and then takes an action a drawn from a set A. (It is possible to allow A to depend on s and S to depend on t.) This ends the game: The payoff to i is given by a function $u^i: T \times M \times A \to \mathbb{R}$.

This canonical game captures the essential features of the classic applications of market signaling. In the labor-market signaling story due to Spence [72] a worker wishes to signal his ability to a potential employer. The worker has information about ability that the employer lacks. Direct communication about ability is not possible, but the worker can acquire education. The employer can observe the worker's level of education and use this to form a judgment about the worker's true level of ability. In this application, S is a worker; R represents a potential employer (or a competitive labor market); t is the student's productivity; s is her level of education; and a is her wage.

Equilibrium

Defining Nash equilibrium for the basic signaling game is completely straightforward when T, S, and A are finite sets. In this case a behavior strategy for S is a function $\mu: T \times M \to [0,1]$ such that $\sum_{s \in M} \mu(t,s) = 1$ for all t. $\mu(t,s)$ is the probability that sender-type t sends the signal s. A behavior strategy for R is a function $\alpha: M \times A \to [0,1]$ where $\sum_{a \in A} \alpha(s,a) = 1$ for all s. $\alpha(s,a)$ is the probability that R takes action a following the signal s.

Proposition 1 *Behavior strategies (α^*, μ^*) form a Nash Equilibrium if and only if for all $t \in T$*

$$\mu(t,s) > 0 \text{ implies } \sum_{a \in A} U^S(t,s,a)\alpha(s,a) = \max_{s' \in S} \sum_{a \in A} U^S(t,s',a)\alpha(s',a) \quad (1)$$

and, for each $s \in S$ such that $\sum_{t \in T} \mu(t,s)\pi(t) > 0$, if $\sum_{t \in T} \mu(t,s)\pi(t) > 0$, then

$$\alpha(s,a) > 0 \text{ implies } \sum_{t \in T} U^R(t,s,a)\beta(t,a) = \max_{a' \in A} \sum_{t \in T} U^R(t,s,a')\beta(t,a'), \quad (2)$$

where

$$\beta(t,s) = \frac{\mu(t,s)\pi(t)}{\sum_{t' \in T} \mu(t',s)\pi(t')}. \quad (3)$$

Condition (1) states that the S places positive probability only on signals that maximize expected utility. This condition guarantees that S responds optimally to R's strategy.

Condition (2) states that R places positive probability only on actions that maximize expected utility, where is taken with respect to the distribution $\beta(\cdot, s)$ following the signal s. Condition (3) states that $\beta(\cdot, s)$ accurately reflects the pattern of play. It requires that R's beliefs be determined using S's strategy and the prior distribution whenever possible. Equilibrium refinements also require that R has beliefs following signals s that satisfy

$$\sum_{t \in T} \mu(t, s)\pi(t) = 0\,, \tag{4}$$

that is are sent with probability zero in equilibrium. Specifically, sequential equilibrium permits $\beta(\cdot, m)$ to be an arbitrary distribution when Eq. (4) holds, but requires that Eq. (2) holds even for these values of s. This restriction rules out equilibria in which certain signals are not sent because the receiver responds to the signal with an action that is dominated.

The ability to signal creates the possibility that R will be able to draw inferences about S's type from the signal. Whether he is able to do so is a property of the equilibrium. It is useful to define two extreme cases.

Definition 1 An equilibrium (α^*, μ^*) is called a separating equilibrium if each type t sends different signals. That is, M can be partitioned into sets M_t such that for each t, $\sum_{s \in M_t} \mu(t, s) = 1$. An equilibrium (α^*, μ^*) is called a pooling equilibrium if there is a single signal s^* that is sent by all types with probability one.

In a separating equilibrium, R can infer S's private information completely. In a pooling equilibrium, R learns nothing from the sender's signal. This definition excludes other possible situations. For example, all sender types can randomize uniformly over a set of two or more signals. In this case, the receiver will be able to draw no inference beyond the prior from a signal received in equilibrium. More interesting is the possibility that the equilibrium will be partially revealing, with some, but not all of the sender types sending common signals.

It is not difficult to construct pooling equilibria for the basic signaling game. Take the labor market model and assume S sends the message s^* with probability one and that the receiver responds to s^* with his best response to the prior distribution and to all other messages with the best response to the belief that t is the least skilled agent. Provided that the least skilled agent prefers to send s^* to sending the cheapest alternative signal, this is a Nash Equilibrium outcome.

The Basic Model

The separating equilibrium is a benchmark outcome for signaling games. When a separating equilibrium exists, then it is possible for the sender to share her information fully with the receiver in spite of having a potential conflict of interest.

Existence of separating equilibria typically requires a systematic relationship between types and signals. An appropriate condition, commonly referred to as the single-crossing condition, plays a prominent role in signaling games and in models of asymmetric information more generally.

In this section I limit attention to a special class of signaling game in which there is a monotonic relationship between types and signals. In these models, separating equilibria typically exist.

I begin by stating the assumption in the environment most commonly seen in applications. Assume that the sets T, S, and A are all real intervals.

Definition 2 $U^S(\cdot)$ satisfies the single-crossing condition if $U^S(t, s, a) \leq U^S(t, s', a')$ for $s' > s$ implies that $U^S(t', s, a) < U^S(t', s', a')$ for all $t' > t$.

In a typical application, $U^S(\cdot)$ is strictly decreasing in its second argument (the signal) and increasing in its third argument (R's response) for all types. Consequently indifference curves are well defined in $M \times A$ for all t. The single-crossing condition states that indifference curves of different sender types cross once. If a lower type is indifferent between two signal-action pairs, then a higher type strictly prefers to send the higher signal. In this way, the single-crossing condition links signals to types in such a way as to guarantee that higher types send weakly higher signals in equilibrium.

Note two generalizations of Definition 2. First, the assumption that the domain of $U^S(\cdot)$ is the product of intervals can be replaced by the assumption that these sets are partially ordered. In this case, weak and strict order replace the weak and strict inequalities comparing types and actions in the statement of the definition. Second, it is sometimes necessary to extend the definition to mixed strategies. In this case, the ordering of A induces a partial ordering of distributions of A through first-order stochastic dominance.

When one thinks of the single-crossing condition geometrically, it is apparent that it implies a ranking of the slopes of the indifference curves of the sender. Suppose that $U^S(\cdot)$ is smooth, strictly increasing in actions and strictly decreasing in signals so that indifference curves are well defined for each t. Writing the indifference curve as $\{(s, \tilde{a}(s; t))\}$, it must be that $U^S(t, s, \tilde{a}(s; t)) \equiv 0$, so that

the slope of the indifference curve of a type t sender is

$$\bar{a}_1(s;t) = -\frac{U_2^S(t,s,a)}{U_3^S(t,s,a)}, \tag{5}$$

where $\bar{a}_1(s;t)$ is the partial derivative of $\bar{a}(s;t)$ with respect to the first argument, and $U_k^S(\cdot)$ denotes the partial derivative of $U^S(\cdot)$ with respect to its kth argument. Under these conditions, the single-crossing condition is implied by the requirement that the right-hand side of Eq. (5) is decreasing in t. The differentiable version of the single-crossing condition is often referred to as the Spence–Mirrlees condition. Milgrom and Shannon [57] contains general definitions of the single-crossing and Spence–Mirrlees conditions and Edlin and Shannon [26] provides a precise statement of when the conditions are equivalent.

To provide a simple construction of a separating equilibrium, limit attention to a standard signaling game in which the following conditions hold.

1. $T = \{0, \ldots, K\}$ is finite.
2. A and M are real intervals.
3. Utility functions are continuous in action and signal.
4. $U^S(\cdot)$ is strictly increasing in action and strictly decreasing in signal.
5. The single-crossing property holds.
6. The receiver's best-response function is uniquely defined, independent of the signal, and strictly increasing in t so that it can be written $BR(t)$.
7. There exists $\bar{s} \in S$ such that $U^S(K, \bar{s}, BR(K)) < U^S(K, s_0^*, BR(0))$.

Conditions 1 and 2 simplify exposition, but otherwise are not necessary. It is important that T, M, and A be partially ordered so that some kind of single-crossing condition applies. Conditions 4–6 impose a monotone structure on the problem so that higher types are more able to send high signals, and that higher types induce higher (and uniformly more attractive) actions. These conditions imply that in equilibrium higher types will necessarily send weakly higher signals. Condition 7 is a boundary condition that makes sending high signals unattractive. It states that the highest type of sender would prefer to be treated like the lowest type rather than use the signal $\bar{s}$. These properties hold in many standard applications. Condition 6 would be satisfied if $U^R(t,s,a) = -(a-t)^2$.

Separating Equilibrium

To illustrate these ideas, consider a construction of a separating equilibrium.

Proposition 2 *The standard signaling game has a separating equilibrium.*

One can prove the proposition by constructing a possible equilibrium path and confirming that the path can be part of a separating equilibrium.

Step 1. t_0 selects the signal s_0^* that maximizes $U^S(t_0, s, BR(t_0))$.

Step 2. Suppose that s_i^* have been specified for $i = 0, \ldots, k-1$ and let $U^*(t_i) = U^S(t_i, s_i^*, BR(t_i))$. Define s_k^* to solve:

$$\max U^S(t_k, s, BR(t_k)) \text{ subject to } U^S(t_{k-1}, s, BR(t_k)) \le U^*(t_{k-1}).$$

Provided that the optimization problems in Steps 1 and 2 have solutions, the process inductively produces a signaling strategy for the sender and a response rule for the receiver defined on $\{s_0^*, \ldots, s_K^*\}$. When $BR(\cdot)$ is strictly increasing, the single-crossing condition implies that the signaling strategy is strictly increasing. To complete the description of strategies, assume that the receiver takes the action $BR(t_k)$ in response to signals in the interval $[s_k, s_{k+1})$, $BR(t_0)$ for $s < s_0^*$, and $BR(t_K)$ for $s > s_K^*$. By the definition of the best-response function, the receiver is best responding to the sender's strategy. When the boundary condition fails, a fully separating equilibrium need not exist, but when M is compact, one can follow the construction above to obtain an equilibrium in which the lowest types separate and higher types pool at the maximum signal in M (see Cho and Sobel [22] for details).

In the construction, the equilibrium involves inefficient levels of signaling. When $U^S(\cdot)$ is decreasing in the signal, all but the lowest type of sender must make a wasteful expenditure in the signal in order to avoid being treating as having a lower quality. The result that expenditures on signals are greater than the levels optimal in a full-information model continue to hold when $U^S(\cdot)$ is not monotonic in the signal. The sender inevitably does no better in a separating equilibrium than she would do if R had full information about t. Indeed, all but the lowest type will do strictly worse in standard signaling games. On the other hand, the equilibrium constructed above has a constrained efficiency property: Of all separating equilibria, it is Pareto dominant from the standpoint of S. To confirm this claim argue inductively that in any separating equilibrium if t_j sends the signal s_j, then $s_j \ge s_j^*$, with equality only if all types $i < j$ send s_i^* with probability one.

Mailath [49] provides a similar construction when T is a real interval. In this case, the Spence–Mirrlees formulation of the single-crossing condition plays an important role and the equilibrium is a solution to a differential equation.

Multiple Equilibria and Selection

Section "Equilibrium" ended with the construction of a pooling equilibrium. A careful reconsideration of the argument reveals that there typically are many pooling equilibrium outcomes. One can construct a potential pooling outcome by assuming that all sender types send the same signal, that the receiver best responds to this common signal, and responds to all other signals with the least attractive action. Under the standard monotonicity assumptions, this strategy profile will be an equilibrium if the lowest sender type prefers pooling to sending the cheapest available out-of-equilibrium message. Section "Separating Equilibrium" ended with the construction of a separating equilibrium. There are also typically many separating equilibrium outcomes. Assume that types $t = 0, \ldots, r-1$ send signals $s^*(t)$, type r sends $\tilde{s}(k) > s^*(k)$, and subsequent signals $\tilde{s}^*(t)$ for $t > r$ solve:

$$\max U^S(t_k, s, BR(t_k)) \text{ subject to}$$
$$U^S(t_{k-1}, s, BR(t_k)) \leq U(t_{k-1}, \tilde{s}, BR(t_{k-1})) .$$

In both of these cases, the multiplicity is typically profound, with a continuum of distinct equilibrium outcomes (when M is an interval). The multiplicity of equilibria means that, without refinement, equilibrium theory provides few clear predictions beyond the observation that the lowest type of sender receives at least $U^*(t_0)$, the payoff it would receive under complete information, and the fact that the equilibrium signaling function is weakly increasing in the sender's type. The first property is a consequence of the monotonicity of S's payoff in a and of R's best response function. The second property is a consequence of the single-crossing condition.

This section describes techniques that refine the set of equilibria. Refinement arguments that guarantee existence and select unique outcomes for standard signaling games rely on the Kohlberg–Mertens [42] notion of strategic stability. The complete theory of strategic stability is only available for finite games. Consequently the literature applies weaker versions of strategic stability that are defined more easily for large games. Banks and Sobel [8], Cho and Kreps [21], and Cho and Sobel [22] introduce these arguments.

Multiple equilibria arise in signaling games because Nash equilibrium does not constrain the receiver's response to signals sent with zero probability in equilibrium. Specifying that R's response to these unsent signals is unattractive leads to the largest set of equilibrium outcomes. (In standard signaling games, S's preferences over actions do not depend on type, so the least attractive action is well defined.) The equilibrium set shrinks if one restricts the meaning of unsent signals. An effective restriction is condition D1, introduced in Cho and Kreps [21]. This condition is less restrictive than the notion of universal divinity introduced by Banks and Sobel [8], which in finite games is less restrictive than Kohlberg and Mertens's notion of strategic stability.

Given an equilibrium (α^*, μ^*), let $U^*(t)$ be the equilibrium expected payoff of a type t sender and let $D(s, t) = \{a\colon u(t, s, a) \geq U^*(t)\}$ be the set of pure-strategy responses to s that lead to payoffs at least as great as the equilibrium payoff for player t. Given a collection of sets, $X(t)$, $t \in T$, $X(t^*)$ is maximal if it not a proper subset of any $X(t)$.

Definition 3 Behavior strategies (α^*, μ^*) together with beliefs β^* satisfy D1 if for any unsent message s, $\beta(\cdot, s)$ is supported on those t for which $D(s, t)$ is maximal.

In standard signaling games, $D(s, t)$ is an interval: all actions greater than or equal to a particular action will be attractive relative to the equilibrium. Hence these sets are nested. If $D(s, t)$ is not maximal, then there is another type t' that is "more likely to deviate" in the sense that there exists out-of-equilibrium responses that are attractive to t' but not t. Condition D1 requires that the receiver place no weight on type t making a deviation in this case. Notice if $D(s, t)$ is empty for all t, then D1 does not restrict beliefs given s (and any choice of action will support the putative equilibrium). Condition D1 is strong. One can imagine weaker restrictions. The intuitive condition (Cho and Kreps [21]) requires that $\beta(t, s) = 0$ when $D(t, s) = \phi$ and at least one other $D(t', s)$ is non empty. Divinity (Banks and Sobel [8]) requires that if $D(t, s)$ is strictly contained in $D(t', s)$, then $\beta(t', s)/\beta(t, s) \geq \pi(t')/\pi(t)$, so that the relative probability of the types more likely to deviate increases.

Proposition 3 *The standard signaling game has a unique separating equilibrium outcome that satisfies Condition D1.*

In standard signaling games, the only equilibrium outcome that satisfies Condition D1 is the separating outcome described in the previous section. Details of the argument appear in Cho and Sobel. The argument relies on two insights. First, types cannot be pooled in equilibrium because slightly higher signals will be interpreted as coming from the highest type in the pool. Second, in any separating equilibrium in which a sender type fails to solve Step 2, deviation to a slightly lower signal will not lower R's beliefs.

The refinement argument is powerful and the separating outcome selected receives prominent attention in the literature. It is worth pointing out that the outcome has

one unreasonable property. The separating outcome described above depends only on the support of types, and not on the details of the distribution. Further, all types but the lowest type must make inefficient (compared to the full-information case) investments in signal in order to distinguish themselves from lower types. The efficient separating equilibrium for a sequence of games in which the probability of the lowest type converges to zero does not converge to the separating equilibrium of the game in which the probability of the lowest type is zero. In the special case of only two types, the (efficient) pooling outcome may be a more plausible outcome when the probability of the lower type shrinks to zero. Grossman and Perry [38] and Mailath, Okuno-Fujiwara, and Postlewaite [50] introduce equilibrium refinements that select the pooling equilibrium in this setting. These concepts share many of the same motivations of the refinements introduced by Banks and Sobel and Cho and Kreps. They are qualitatively different from the intuitive criterion, divinity, and Condition D1, because they are not based on dominance arguments and lack general existence properties.

Cheap Talk

Models in which preferences satisfy the single-crossing property are central in the literature, but the assumption is not appropriate in some interesting settings. This section describes an extreme case in which there is no direct cost of signaling.

In general, a cheap-talk model is a signaling model in which $u^i(t, s, a)$ is independent of s for all (t, a). Two facts about this model are immediate. First, if equilibrium exists, then there always exists an equilibrium in which no information is communicated. To construct this "babbling" equilibrium, assume that $\beta(t, s)$ is equal to the prior independent of the signal s. R's best response will be to take an action that is optimal conditional only on his prior information. Hence R's action can be taken to be constant. In this case, it is also a best response for S to send a signal that is independent of type, which makes $\beta(t, s)$ the appropriate beliefs. Hence, even if the interests of S and R are identical, so that it there are strong incentives to communicate, there is a possibility of complete communication break down.

Second, it is clear that non-trivial communication requires that different types of S have different preferences over R's actions. If it is the case that whenever some type t prefers action a to action a' then so do all other types, then (ruling out indifference), it must be the case that in equilibrium the receiver takes only one action with positive probability. To see this, note that otherwise one type of sender is not selecting a best response. The second observation shows that cheap talk is not effective in games, like the standard labor-market story, in which the sender's preferences are monotonic in the action of the receiver. With cheap communication, the potential employee in the labor market will always select a signal that leads to the higher possible wage and consequently, in equilibrium, all types of workers will receive the same wage.

A Simple Cheap-Talk Game

There are natural settings in which cheap talk is meaningful in equilibrium. To describe examples, I follow the development of Crawford and Sobel [24] (Green and Stokey [35] independently introduced a similar game in an article circulated in 1981). In this paper, A and T are the unit interval and M can be taken to be the unit interval without loss of generality. The sender's private information or type, t, is drawn from a differentiable probability distribution function, $F(\cdot)$, with density $f(\cdot)$, supported on $[0, 1]$. S and R have twice continuously differentiable von Neumann–Morgenstern utility functions $U^i(a, t)$ that are strictly concave in a and have a strictly positive mixed partial derivative. Let $i = R, S$, $a^i(t)$ denotes the unique solution to $\max_a U^i(a, t)$ and further assume that $a^S(t) > a^R(t)$ for all t. (The assumptions on $U^i(\cdot)$ guarantee that $U^i(\cdot)$ is well defined and strictly increasing.)

In this model, the interests of the sender and receiver are partially aligned because both would like to take a higher action with a higher t. The interests are different because S would always like the action to be a bit higher than R's ideal action. In a typical application, t represents the idea action for R, such as the appropriate expenditure on a public project. Both R and S want actual expenditure to be close to the target value, but S has a bias in favor of additional expenditure.

For $0 \leq t' < t'' \leq 1$, let $\tilde{a}(t', t'')$ be the unique solution to $\max_a \int_{t'}^{t''} U^R(a, t)\mathrm{d}F(t)$. By convention, $\tilde{a}(t, t) = a^R(t)$.

Without loss of generality, limit attention to pure-strategy equilibria. The concavity assumption guarantees that R's best responses will be unique, so R will not randomize in equilibrium. An equilibrium with strategies (μ^*, α^*) **induces** action a if $\{t\colon \alpha^*(\mu^*(t)) = a\}$ has positive prior probability. Crawford and Sobel [24] characterize equilibrium outcomes.

Proposition 4 *There exists a positive integer N^* such that for every integer N with $1 \leq N \leq N^*$, there exists at least one equilibrium in which the set of induced actions has cardinality N, and moreover, there is no equilibrium*

which induces more than N^* actions. An equilibrium can be characterized by a partition of the set of types, $t(N) = (t_0(N), \ldots, t_N(N))$ with $0 = t_0(N) < t_1(N) < \ldots < t_N(N) = 1$, and signals m_i, $i = 1, \ldots, N$, such that for all $i = 1, \ldots, N-1$

$$U^S(\bar{a}(t_i, t_{i+1}), t_i)) - U^S(\bar{a}(t_{i-1}, t_i), t_i)) = 0\,, \tag{6}$$

$$\mu(t) = m_i \;\; for \;\; t \in (t_{i-1}, t_i]\,, \tag{7}$$

and

$$\alpha(m_i) = \bar{a}(t_{i-1}, t_i)\,. \tag{8}$$

Furthermore, essentially all equilibrium outcomes can be described in this way.

In an equilibrium, adjacent types pool together and send a common message. Condition (6) states that sender types on the boundary of a partition element are indifferent between pooling with types immediately below or immediately above. Condition (7) states that types in a common element of the partition send the same message. Condition (8) states that R best responds to the information in S's message.

Crawford and Sobel make another monotonicity assumption, which they call condition (M). (M) is satisfied in leading examples and implies that there is a unique equilibrium partition for each $N = 1, \ldots, N^*$, the ex-ante equilibrium expected utility for both S and R is increasing in N, and N^* increases if the preferences of S and R become more aligned. These conclusions provide justification for the view that with fixed preferences "more" communication (in the sense of more actions induced) is better for both players and that the closer are the interests of the players the greater the possibilities for communication.

As in the case of models with costly signaling, there are multiple equilibria in the cheap-talk model. The multiplicity is qualitatively different. Costly signaling models have a continuum of Nash Equilibrium outcomes. Cheap-talk models have only finitely many. Refinements that impose restrictions on off-the-equilibrium path signals work well to identify a single outcome in costly signaling models. These refinements have no cutting power in cheap-talk models because any equilibrium distribution on type-action pairs can arise from signaling strategies in which all messages are sent with positive probability. To prove this claim, observe that if message m' is unused in equilibrium, while message m is used, then one can construct a new equilibrium in which R interprets m' the same way as m and sender types previously sending m randomize equally between m and m'.

In the basic model messages take on meaning only through their use in an equilibrium. Unlike natural language, they have no external meaning. There have been several attempts to formalize the notion that messages have meanings that, if consistent with strategic aspects of the interaction, should be their interpretation inside the game. The first formulation of this idea is due to Farrell [28].

Definition 4 Given an equilibrium (α^*, σ^*) with sender expected payoffs $u^*(\cdot)$, the subset $G \subset T$ is self signaling if $G = \{t\colon U^S(t, BR(G)) > u^*(t)\}$.

That is, G is self signaling if precisely the types in G gain by making a statement that induces the action that is a best response to the information that $t \in G$. (When $BR(t)$ is not single valued it is necessary to refine the definition somewhat and permit the possibility that $U^S(t, BR(G)) = u^*(t)$ for some t. See Matthews, Okuno-Fujiwara, and Postlewaite [51].) Farrell argues that the existence of a self-signaling set would destroy an equilibrium. If a subset G had available a message that meant "my type is in G," then relative to the equilibrium R could infer that if he were to interpret the message literally, then it would be sent only by those types in G (and hence the literal meaning would be accurate). With this motivation, Farrell proposes a refinement.

Definition 5 An equilibrium (α^*, σ^*) is neologism proof if there exist no self-signaling sets relative to the equilibrium.

Rabin [62] argues convincingly that Farrell's definition rules out too many equilibrium outcomes. Indeed, for leading examples of the basic cheap-talk game, there are no neologism-proof equilibria. Specifically, in the Crawford–Sobel model in which S has a bias towards higher actions, there exist self signaling sets of the form $[t, 1]$. On the other hand, Chen, Kartik, and Sobel [20] demonstrate that the N^*-step equilibrium always satisfies the no incentive to separate (NITS) condition:

$$U^S(\alpha^*(\mu^*(0)), 0) \geq U^S(a^R(0), 0)\,, \tag{9}$$

and that under condition (M) this is the only equilibrium that satisfies Condition (9).

NITS states that the lowest type of sender prefers her equilibrium payoff to the payoff she would receive if the receiver knew her type (and responded optimally). [41] introduced and named this condition. The NITS condition can be shown to rule out equilibria that admit if self-signaling sets of the form $[0, t]$. Chen [19] and Kartik [41] show that the condition holds in the limits of perturbed versions of the basic cheap-talk game.

Inequality (9) holds automatically in any perfect Bayesian equilibrium of the standard signaling model. This follows because when R's actions are monotonic in type and S's preferences are monotonic in action, the worst outcome for S is to be viewed as the lowest type. This observation would not be true in Nash Equilibrium, where it is possible for R to respond to an out-of-equilibirum message with an action $a < BR(0)$.

Variations on Cheap Talk

In standard signaling models, there is typically an equilibrium that is fully revealing. This is not the case in the basic cheap-talk model. This leads to the question of whether it is possible to obtain more revelation in different environments.

One possibility is to consider the possibility of signaling over many dimensions. Chakraborty and Harbaugh [18] consider a model in with T and A are multidimensional. A special case of their model is one in which the components of T are independent draws from the same distribution and A involves taking a real-valued action for each component of T. If preferences are additively separable across types and actions, Chakraborty and Harbaugh provide conditions under which categorical information transmission, in which the S transmits the order of the components of T, is credible in equilibrium even when it would not be possible to transmit information if the dimensions were treated in isolation. It may be credible for S to say "$t_1 > t_2$," even if she could not credibly provide information about the absolute value of either component of t.

Communication is non-trivial if some Sender type strictly prefers to induce one equilibrium action over another. Non-trivial communication requires that different types have different preferences over outcomes. In standard signaling models, the heterogeneity arises because different sender types have different costs of sending messages. In cheap-talk models, the heterogeneity arises with one-dimensional actions if different sender types have different ideal actions. With multi-dimensional actions, heterogeneity could come simply from different sender types having different preferences over the relative importance of the different issues. Another simple variation is to assume the existence of more than one sender. In the two-sender game, nature picks t as before, both senders learn t and simultaneously send a message to the receiver, who makes a decision based on the two messages. The second sender has preferences that depend on type and the receiver's action, but not directly on the message sent. In this environment, assume that $M = T$, so that the set of available messages (this is essentially without loss of generality). One can look for equilibria in which the senders report honestly. Denote by $a^*(t, t')$ R's response to the pair of messages (t, t'). If an equilibrium in which both senders report honestly exists, then R's response to identical messages, $a^*(t, t) = a^R(t)$, and it must be the case that there exists a specification of $a(t, t')$ for $t \neq t'$ such that for all $i = 1$ and 2 and $t \neq t'$,

$$U^{S_i}(t, a^*(t, t)) \geq U^{S_i}(t, a^*(t, t')) \,. \tag{10}$$

It is possible to satisfy Condition (10) if the biases of the senders are small relative to the set of possible best responses. Krishna and Morgan [45] studies a one-dimensional model of information transmission with two informed players. Ambrus and Takahashi [1] and Battiglini [9] provide conditions under which full revelation is possible when there are two informed players and possibly multiple dimensions of information.

In many circumstances, enriching the communication structure either by allowing more rounds of communication [2,29], mediation [10], or exogenous uncertainty [16] enlarges the set of equilibrium outcomes.

Verifiable Information

Until now, the focus has been on situations in which the set of signals available does not depend on the true state. There are situations in which this assumption is not appropriate. There may be laws that ban false advertisement. The sender may be able to document details about the value of t. Models of this kind were first studied by Grossman [37] and Milgrom [56]. For example, if t is the sender's skill at playing the piano, then if there is a piano available t could demonstrate that she has skill at least as great as t (by performing at her true ability), but she may not be able to prove that her skill is no more than t (the receiver may think that she deliberately played the piano badly).

To model these possibilities, suppose that the set of possible messages is the set of all subsets of T. In this case, messages have "literal" meanings: When the sender uses the message $s = C \in T$, this can be interpreted as a statement of the form: "my type is in C." If senders cannot lie, then $M(t)$ must be the set of subsets of T that contain t. If type t is verifiable, then $\{t\} \in M(t')$ if and only if $t' = t$. If there are no additional costs of sending signals, this model can be viewed as a variation of cheap talk models in which the message space depends on t. In general, one can treat verifiable information models as a special case of the general signaling game in which the cost of sending certain signals is so large that these signals can be ruled out. Lying

is impossible if $M(t) = \{C \subset 2^T : t \in C\}$. In this setting, it is appropriate to require equilibria to be consistent with the signaling structure.

Definition 6 The equilibrium (σ^*, α^*) is rationalizable if

$$\alpha(C, a) > 0 \text{ implies } \sum_{t \in T} U^R(t, s, a)\beta(t, a) = \max_{a' \in A} \sum_{t \in T} U^R(t, s, a')\beta(t, a') , \quad (11)$$

where $\beta(t, a) = 0$ if $t \notin C$.

Compared to (2), (11) requires that beliefs place positive probability only on types capable of sending the message "my type is an element of C."

Proposition 5 *Suppose that A and T are linearly ordered, that the receiver's best response function is increasing in type, and that all sender types prefer higher actions. If lying is not possible, then in any rationalizable equilibrium (α^*, μ^*), $\alpha^*(s, BR(t)) = 1$ whenever $\mu^*(t, s > 0)$.*

Grossman [37] and Milgrom [56] present versions of this proposition. Seidman and Winter [68] generalize the result.

Provided that the receiver responds to the signal $\{t\}$ with $BR(t)$, each type can guarantee a payoff of $BR(t)$. On the other hand, if any type receives a payoff greater than $BR(t)$, then some higher type must be doing worse. Another way to make the same point is to notice that the highest type $\{\bar{t}\}$ has a weakly dominant strategy to reveal her type by announcing $\{\bar{t}\}$. Once this type is revealed, the next highest type will want to reveal herself and so on. Hence verifiable information will be revealed voluntarily in an environment where cheap talk leads to no revealing and costly signaling will be compatible with full revelation, but at the cost of dissipative signaling.

The full-revelation result depends on the assumption that the sender and receiver share a linear ranking over the quality of information. Giovannini and Seidmann [31] discuss more general settings in which the ability to provide verifiable information need not lead to full revelation.

Communication About Intentions

In a simple signaling game, signals potentially provide information about private information. Another possibility is to add a round of pre-play communication to a given game. Even if the game has complete information, there is the possibility that communication would serve to select equilibria or permit correlation that would otherwise be infeasible. Farrell and Rabin [63]'s review article discusses this literature in more detail.

Aumann [3] argues that one cannot rely on pre-play communication to select a Pareto-efficient equilibrium. He considers a simple two-player game with Pareto-ranked equilibria and argues that no "cheap" pre-play signal would be credible.

Ben-Porath and Dekel [11] show that adding a stage of "money burning" (a signal that reduces all future payoffs by the same amount) when combined with an equilibrium refinement can select equilibria in a complete information game. Although no money is burned in the selected equilibrium outcome, the potential to send costly signals creates dominance relationships that lead to a selection.

Vida [74] synthesizes a literature that compares the set of equilibrium outcomes available when communication possibilities are added to a game to the theoretically larger set available if there is a reliable mediator available to collect information and recommend actions to the players.

Applications

Economic Applications

There is an enormous literature that uses signaling models in applications. Riley's [64] survey contains extended discussion of some of the most important applications. What follows is a brief discussion of some central ideas.

In a simple signaling game, one informed agent sends a single signal to one uninformed decision maker. This setting is reach enough to illustrate many important aspects of signaling, but it plainly limited. Interesting new issues arise if there are many informed agents, if there are many decision makers, and if the interaction is repeated. Several of the models below add some or all of these novel features to the basic model.

Advertising Advertisements are signals. Models similar to the standard model can explain situations in which higher levels of advertisement can lead consumers to believe the quality of the good is higher. In a separating equilibrium, advertising expenditures fully reveal quality. As in all costly signaling models, it is not important that there be a direct relationship between quality and signal, it is only necessary that firms with higher quality have lower marginal costs of advertising. Hence simply "burning money" or sending a signal that lowers utility by an amount independent of quality and response can be informative. The consumer may obtain full information in equilibrium, but someone must pay the cost of advertis-

ing. There are other situations where it is natural for the signal to be linked to the quality of the item. Models of verifiable information are appropriate in this case. When the assumptions of Proposition 5 hold, one would expect consumers to obtain all relevant information through disclosures without wasteful expenditures on signaling. Finally, cheap talk plays a role in some markets. One would expect costless communication to be informative in environments where heterogeneous consumers would like to identify the best product. Cheap talk can create more efficient matching of product to consumer. Here communication is free although in leading models separating equilibria do not exist.

Limit Pricing Signaling models offer one explanation for the phenomenon of limit pricing. An incumbent firm have private information about its cost. Potential entrants use the pricing behavior of the firm to draw inferences about the incumbent's cost, which determines profitability of entry. Milgrom and Roberts [54] construct an equilibrium in which the existence of incomplete information distorts prices: Relative to the full information model, the incumbent charges lower prices in order to signal that the market is relatively unprofitable. This behavior has the flavor of classical models of limit pricing, with one important qualification. In a separating equilibrium the entrant can infer the true cost of the incumbent and therefore the low price charged by the incumbent firm fails to change the entry decision.

Bargaining Several authors have proposed bargaining models with incomplete information to study the existence and duration of strikes [30,71]. If a firm with private information about its profitability makes a take-it-or-leave it offer to a union, then the strategic interaction is a simple signaling model in which the magnitude of the offer may serve as a signal of the firm's profitability. Firms with low profits are better able to make low wage offers to the union because the threat of a strike is less costly to a firm with low profits than one with high profits. Consequently settlement offers may reveal information. Natural extensions of this model permit counter offers. The variation of the model in which the uninformed agent makes offers and the uninformed agent accepts and rejects is formally almost identical to the canonical model of price discrimination by a durable-goods monopolist [4,39].

Finance Simple signaling arguments provide potential explanations for firms' choices of financial structure. Classic arguments due to Modigliani and Miller [58] and imply that firms' profitability should not depend on their choice of capital structure. Hence this theory cannot organize empirical regularities about firm's capital structure. The Modigliani–Miller theorem assumes that the firm's managers, shareholders, and potential shareholders all have access to the same information. An enormous literature assumes instead that the firm's managers have superior information and use corporate structure to signal profitability.

Leland and Pyle [47] assume that insiders are risk averse so they would prefer to diversify their personal holdings rather than maintain large investments in their firm. The value of diversification is greater the lower the quality of the firm. Hence when insiders have superior information than investors, there will be an incentive for the insiders of highly profitable firms to maintain inefficiently large investments in their firm in order to signal profitability to investors.

Dividends are taxed twice under the United States tax code, which raises the question of why firms would issue dividends when capital gains are taxed at a lower rate. A potential explanation for this behavior comes from a model in which investors have imperfect information about the future profitability of the firm and profitable firms are more able than less profitable firms to distribute profits in the form of dividends (see [14]).

Reputation Dynamic models of incomplete information create the opportunity for the receiver to draw inferences about the sender's private information while engaging in an extended interaction. Kreps and Wilson [44] and Milgrom and Roberts [55] provided the original treatments of reputation formation in games of incomplete information. Motivated by the limit pricing, their models examined the interaction of a single long-lived incumbent facing a sequence of potential entrants. The entrants lack information about the willingness of the incumbent to tolerate entry. Pricing decisions of the incumbent provide information to the entrants about the profitability of the market.

In these models, signals have implications for both current and future utility. The current cost is determined by the effect the signal has on current payoffs. In Kreps–Wilson and Milgrom–Roberts, this cost is the decrease in current profits associated with charging a low price. In other models (for example [59,70]) the actual signal is costless, but it has immediate payoff implications because of the response it induces. Signals also have implications for future utility because inferences about the sender's private information will influence the behavior of the opponents in future periods. Adding concern for reputation to a signaling game will influence behavior, but whether it

leads to more or less informative signaling depends on the application.

Signaling in Biology

Signaling is important in biology. In independent and almost contemporaneous work, Zahavi [76] proposed a signaling model that shared the essential features of Spence [72]'s model of labor-market signaling. Zahavi observed that there are many examples in nature of animals apparently excessive physical displays. It takes energy to produce colorful plumage, large antlers, or loud cries. Having a large tail may actually make it harder for peacocks to flea predators. If a baby bird makes a loud sound to get his mother's attention, he may attract a dangerous predator. Zahavi argued that costly signals could play a role in sexual selection. In Zahavi's basic model, the sender is a male and the receiver is a female of the same species. Females who are able to mate with healthier males are more likely to have stronger children, but often the quality of a potential mate cannot be observed directly. Zahavi argued that if healthier males could produce visible displays more cheaply than less healthy males, then females would be induced to use the signals when deciding upon a mate. Displays may impose costs that "handicap" a signaler, but displays would persist when additional reproductive success compensates for their costs. Zahavi identifies a single-crossing condition as a necessary condition for the existence of costly signals.

The development of signaling in biology parallels that in economics, but there are important differences. Biology replaces the assumption of utility maximization and equilibrium with fitness maximization and evolutionary stability. That is, their models do not assume that animals consciously select their signal to maximize a payoff. Instead, the biological models assume that the process of natural selection will lead to strategy profiles in which mutant behavior has lower reproductive fitness than equilibrium behavior. This notion leads to static and dynamic solution concepts similar to Nash Equilibrium and its refinements. Fitness in biological models depends on contributions from both parents. Consequently, a full treatment of signaling must take into account population genetics. Grafen [34] discusses these issues and Grafen [33] and Siller [69] provide further theoretical development of the handicap theory. Finally, one must be careful in interpreting heterogeneous quality in biological models. Natural selection should operate to eliminate the least fit genes in a population. To the extent that this arises, there is pressure for quality variation within a population to decrease over time. The existence of unobserved quality variations needed for signaling may be the result of relatively small variations about a population norm.

While most of the literature on signaling in biology focuses on the use of costly signals, there are also situations in which cheap talk is effective. A leading example is the "Sir Philip Sidney Game," originally developed by John Maynard Smith [53] to illustrate the value of costly communication between a mother and child. The child has private information about its level of hunger and the mother must decide to feed the child or keep the food for itself. Since the players are related, survival of one positively influences the fitness of the other. This creates a common interest needed for cheap-talk communication. There are two ways to model communication in this environment. The first is to assume that signaling is costly, with hungrier babies better able to communicate their hunger. This could be because the sound of a hungry baby is hard for sated babies to imitate or it could be that crying for food increases the risk of predation and that this risk is relatively more dangerous to well fed chicks than to starving ones (because the starving chicks have nothing to lose). This game has multiple equilibria in which signals fully reveal the state of the baby over a range of values (see [46,53]). These papers look at a model in which both mother and child have private information. Alternatively, Bergstrom and Lachmann [13] study a cheap-talk version of the game. Here there may be an equilibrium outcome in which the baby bird credibly signals whether or not he is hungry. Those who signal hunger get fed. The others do not. Well fed baby birds may wish to signal that they are not hungry in order to permit the mother to keep food for herself. Such an equilibrium exists if the fraction of genes that mother and child share is large and the baby is already well fed.

Political Science

Signaling games have played an important role in formal models of political science. Banks [7] reviews models of agenda control, political rhetoric, voting, and electoral competition. Several important models in this area are formally interesting because they violate the standard assumptions frequently satisfied in economic models. I describe two such models in this subsection.

Banks [6] studies a model of agenda setting in which the informed sender proposes a policy to a receiver (decision-maker), who can either accept or reject the proposal. If the proposal is accepted, it becomes the outcome. If not, then the outcome is a fall-back policy. The fall-back policy is known only to the sender. In this environment, the

sender's strategy may convey information to the decision maker. Signaling is costly, but, because the receiver's set of actions in binary, fully revealing equilibria need not exist. Refinements limit the set of predictions in this model to a class of outcomes in which only one proposal is accepted in equilibrium (and that this proposal is accepted with probability one), but there are typically a continuum of possible equilibrium outcomes.

Matthews [52] develops a cheap-talk model of veto threats. There are two players, a Chooser (C), who plays the role of receiver, and a Proposer (P), who plays the role of sender. The players have preferences that are represented by single-peaked utility functions which depend on the real-valued outcome of the game and an ideal point. P's ideal point is common knowledge. C's ideal point is her private information, drawn from a prior distribution that has a smooth positive density on a compact interval, $[\underline{t}, \bar{t}]$. The game form is simple: C learns her type, then sends a cheap-talk signal to P, who responds with a proposal. C then either accepts or rejects the proposal. Accepted proposals become the outcome of the game. If C rejects the proposal, then the outcome is the status quo point.

As usual in cheap-talk games, this game has a babbling outcome in which C's message contains no information and P makes a single, take-it-or-leave-it offer that is accepted with probability strictly between 0 and 1. Matthews shows there may be equilibria in which two outcomes are induced with positive probability (size-two equilibria), but size $n > 2$ (perfect Bayesian) equilibria never exist. In a size-two equilibrium, P offers his ideal outcome to those types of C whose message indicates that their ideal point is low; this offer is always accepted in equilibrium. If C indicated that his ideal point is high, P makes a compromise offer that is sometimes accepted and sometimes rejected.

Future Directions

The most exciting developments in signaling games in the future are likely to come from interaction between economics and other disciplines.

Over the last ten years the influence of behavioral economists have led the profession to rethink many of its fundamental models. An explosion of experimental studies have already influenced the interpretation of signaling models and have led to a re-examination of basic assumptions. There is evidence that economic actors lack the strategic sophistication assumed in equilibrium models. Further, economic agents may be motivated by more than their material well being. Existing experimental evidence provides broad support for many of the qualitative predictions of the theory (Banks, Camerer, and Porter [5] and Brandts and Holt [17]), but also suggests ways in which the theory may be inadequate.

The driving assumption of signaling models is that when informational asymmetries exist, senders will attempt to lie for strategic advantage and that sophisticated receivers will discount statements. These assumptions may be reconsidered in light of experimental evidence that some agents will behave honestly in spite of strategic incentives to lie. For example, Gneezy [32] and Hurkens and Kartik [40] present experimental evidence that some agents are reluctant to lie even when there is a financial gain from doing so. There is evidence from other disciplines that some agents are unwilling or unable to manipulate information for strategic advantage and that people may be well equipped to detect these manipulations in ways that are not captured in standard models (see, for example, Ekman [27] or Trivers [73]). Experimental evidence and, possibly, results from neuroscience may demonstrate that the standard assumption that some agents cannot manipulate information for their strategic advantage (or that other agents have ability to see through deception) will inform the development of novel models of communication that include behavioral types. Several papers study the implications of including behavioral types into the standard paradigm. The reputation models of Kreps and Wilson [44] and Milgrom and Roberts [54] are two early examples. Papers on communication by Chen [19], Crawford [23], Kartik [41], and Olszewski [61] are more recent examples. New developments in behavioral economics will inform future theoretical studies.

There is substantial interest in signaling in philosophy. Indeed, the philosopher David Lewis [48] (first published in 1969) introduced signaling games prior to the contributions of Spence and Zahavi. Recently linguists have been paying more attention to game-theoretic ideas. Benz, Jäger and Van Rooij [12] collects recent work that attempts to formalize ideas from linguistic philosophy due to Grice [36]. While there have been a small number of contributions by economists in this area (Rubinstein [66] and Sally [67] are examples), there is likely to be more active interaction in the future.

Finally, future work may connect strategic aspects of communication to the actual structure of language. Blume [15], Cucker, Smale, Zhou [25], and Nowak and Krakauer [60] present dramatically different models on how structured communication may result from learning processes. Synthesizing these approaches may lead to fundamental insights on how the ability to send and receive signals develops.

Acknowledgments

I thank the Guggenheim Foundation, NSF, and the Secretaría de Estado de Universidades e Investigación del Ministerio de Educación y Ciencia (Spain) for financial support and Jose Penalva for his comments. I am grateful to the Departament d'Economia i d'Història Econòmica and Institut d'Anàlisi Econòmica of the Universitat Autònoma de Barcelona for hospitality and administrative support.

Bibliography

Primary Literature

1. Ambrus A, Takahashi S (2008) Multi-sender cheap talk with restricted state space. Theor Econ 3(1):1–27
2. Aumann R, Hart S (2003) Long cheap talk. Econometrica 71(6):1619–1660
3. Aumann RJ (1990) Nash equilibria and not self-enforcing. In: Gabszewicz JJ, Richard J-F, Wolsey LA (eds) Economic Decision Making: Games, Econometrics and Optimisation. Elsevier, Amsterdam, pp 201–206
4. Ausubel LM, Deneckere RJ (1989) Reputation in bargaining and durable goods monopoly. Econometrica 57(3):511–531
5. Banks J, Camerer C, Porter D (1994) An experimental analysis of nash refinements in signaling games. Games Econ Behav 6(1):1–31
6. Banks JS (1990) Monopoly agenda control and asymmetric information. Q J Econ 105(2):445–464
7. Banks JS (1991) Signaling Games in Political Science. Routledge, Langhorne
8. Banks JS, Sobel J (1987) Equilibrium selection in signaling games. Econometrica 55(3):647–661
9. Battaglini M (2002) Multiple referrals and multidimensional cheap talk. Econometrica 70(4):1379–1401
10. Ben-Porath E (2003) Cheap talk in games with incomplete information. J Econ Theory 108:45–71
11. Ben-Porath E, Dekel E (1992) Signaling future actions and the potential for sacrifice. J Econ Theory 57:36–51
12. Benz A, Jäger G, Van Rooij R (eds) (2005) Game Theory and Pragmatics. Palgrave MacMillan, Houndmills, Basingstoke
13. Bergstrom CT, Lachmann M (1998) Signalling among relatives. III. talk is cheap. Proc Natl Acad Sci USA 95:5100–5015
14. Bhattacharya S (1979) Imperfect information, dividend policy, and the bird in the hand fallacy. Bell J Econ 10(1):259–270
15. Blume A (2000) Coordination and learning with a partial language. J Econ Theory 95:1–36
16. Blume A, Board O, Kawamura K (2007) Noisy talk. Theor Econ 2(4):396–440
17. Brandts J, Holt CA (1992) An experimental test of equilibrium dominance in signaling games. Am Econ Rev 82(5):1350–1365
18. Chakraborty A, Harbaugh R (2007) Comparative cheap talk. J Econ Theory 132(1):70–94
19. Chen Y (2005) Perturbed communication games with honest senders and naive receivers. Technical report. Arizona State University, Tempe
20. Chen Y, Kartik N, Sobel J (2007) On the robustness of informative cheap talk. Technical report, UCSD. In: Econometrica 76(1):117–136, Blackwell Synergy
21. Cho I-K, Kreps DM (1987) Signaling games and stable equilibria. Q J Econ 102(2):179–221
22. Cho I-K, Sobel J (1990) Strategic stability and uniqueness in signaling games. J Econ Theory 50(2):381–413
23. Crawford VP (2003) Lying for strategic advantage: Rational and boundedly rational misrepresentation of intentions. Am Econ Rev 93(1):133–149
24. Crawford VP, Sobel J (1982) Strategic information transmission. Econometrica 50(6):1431–1451
25. Cucker F, Smale S, Zhou D-X (2004) Modeling language evolution. Found Comput Math 4(3):315–343
26. Edlin AS, Shannon C (1998) Strict single-crossing and the strict spence-mirrlees condition: A comment on monotone comparative statics. Econometrica 66(6):1417–1425
27. Ekman P (2001) Telling Lies. W.W. Norton, New York
28. Farrell J (1993) Meaning and credibility in cheap-talk games. Games Econ Behav 5(4):514–531
29. Forges F (1990) Equilibria with communication in a job market example. Q J Econ 105(2):375–398
30. Fudenberg D, Tirole J (1983) Sequential bargaining with incomplete information. Rev Econ Stud 50(2):221–247
31. Giovannoni F, Seidmann DJ (2007) Secrecy, two-sided bias and the value of evidence. Games Econ Behav 59(2):296–315
32. Gneezy U (2005) Deception: The role of consequences. Am Econ Rev 95(1):384–394
33. Grafen A (1990) Biological signals as handicaps. J Theor Biol 144:517–546
34. Grafen A (1990) Sexual selection unhandicapped by the fisher process. J Theor Biol 144:473–516
35. Green JR, Stokey NL (2007) A two-person game of information transmission. J Econ Theory 135(1):90–104
36. Grice HP (1991) Studies in the Way of Words. Harvard University Press, Cambridge
37. Grossman S (1981) The role of warranties and private disclosure about product quality. J Law Econ 24:461–483
38. Grossman S, Perry M (1987) Perfect sequential equilibria. J Econ Theory 39:97–119
39. Gul F, Sonnenchein H, Wilson R (1986) Foundations of dynamic monopoly and the coase conjecture. J Econ Theory 39:155–190
40. Hurkens S, Kartik N (2006) (When) Would I lie to you? Comment on Deception: The role of consequences. Technical report, UCSD
41. Kartik N (2005) Information transmission with almost-cheap talk. Technical report, UCSD
42. Kohlberg E, Mertens J-F (1986) On the strategic stability of equilibria. Econometrica 54(5):1003–1037
43. Kreps DM, Sobel J (1994) Signalling. In: Aumann RJ, Hart S (eds) Handbook of game theory: with economics applications. Handbooks in Economics, no 11, vol 2. Elsevier, Amsterdam, chap 25, pp 849–868
44. Kreps DM, Wilson R (1982) Reputation and imperfect information. J Econ Theory 27:253–277
45. Krishna V, Morgan J (2001) A model of expertise. Q J Econ 116(2):747–775
46. Lachmann M, Bergstrom CT (1998) Signalling among relatives. II. beyond the tower of babel. Theor Popul Biol 54:146–160
47. Leland HE, Pyle DH (1977) Informational asymmetries, financial structure, and financial intermediation informational asymmetries, financial structure, and financial intermediation. J Finance 32(2):371–387

48. Lewis D (2002) Convention: A Philosophical Study. Blackwell, Oxford
49. Mailath GJ (1987) Incentive compatibility in signaling games with a continuum of types. Econometrica 55(6):1349–1365
50. Mailath GJ, Okuno-Fujiwara M, Postlewaite A (1993) On belief based refinements in signaling games. J Econ Theory 60(2):241–276
51. Mathews SA, Okuno-Fujiwara M, Postlewaite A (1991) Refining cheap-talk equilibria. J Econ Theory 55(2):247–273
52. Matthews SA (1989) Veto threats: Rhetoric in a bargaining game. Q J Econ 104(2):347–369
53. Maynard Smith J (1991) Honest signalling: the Philip Sidney game. Anim Behav 42:1034–1035
54. Milgrom P, Roberts J (1982) Limit pricing and entry under incomplete information: An equilibrium analysis. Econometrica 50(2):443–459
55. Milgrom P, Roberts J (1982) Predation, reputation, and entry deterence. J Econ Theory 27:280–312
56. Milgrom PR (1981) Good news and bad news: Representation theorems. Bell J Econ 21:380–391
57. Milgrom PR, Shannon C (1994) Monotone comparative statics. Econometrica 62(1):157–180
58. Modigliani F, Miller MH (1958) The cost of capital, corporation finance and the theory of investment. Am Econ Rev 48(3):261–297
59. Morris S (2001) Political correctness. J Political Econ 109:231–265
60. Nowak MA, Krakauer DC (1999) The evolution of language. Proc Natl Acad Sci 96(14):8028–8033
61. Olszewski W (2004) Informal communication. J Econ Theory 117:180–200
62. Rabin M (1990) Communication between rational agents. J Econ Theory 51:144–170
63. Rabin M, Farrell J (1996) Cheap talk. J Econ Perspectives 10(3):103–118
64. Riley JG (2001) Silver signals: Twenty-five years of screening and signaling. J Econ Lit 39(2):432–478
65. Rothschild M, Stiglitz J (1976) Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. Q J Econ 90(4):629–649
66. Rubinstein A (2000) Economics and language. Cambridge University Press, New York
67. Sally D (2005) Can I say, bobobo and mean, there's no such thing as cheap talk? J Econ Behav Organ 57(3):245–266
68. Seidmann DJ, Winter E (1997) Strategic information transmission with verifiable messages. Econometrica 65(1):163–169
69. Siller S (1998) A note on errors in grafen's strategic handicap models. J Theor Biol 195:413–417
70. Sobel J (1985) A theory of credibility. Rev Econ Stud 52(4): 557–573
71. Sobel J, Takahashi I (1983) A multistage model of bargaining. Rev Econ Stud 50(3):411–426
72. Spence AM (1974) Market Signaling. Harvard University Press, Cambridge
73. Trivers RL (1971) The evolution of reciprocal altruism. Q Rev Biol 46:35–58
74. Vida P (2006) Long Pre-Play Communication in Games. Ph D thesis, Autonomous University of Barcelona
75. Wilson C (1977) A model of insurance markets with incomplete information. J Econ Theory 16:167–207
76. Zahavi A (1975) Mate selection- a selection for a handicap. J Theor Biol 53:205–214

Books and Reviews

Admati A, Perry M (1987) Strategic delay in bargaining. Rev Econ Stud 54:345–364
Austen-Smith D (1990) Information transmission in debate. Am J Political Sci 34(1):124–152
Battigalli P, Siniscalchi M (2002) Strong belief and forward induction reasoning. J Econ Theory 106(2):356–391
Bernheim BD (1994) A theory of conformity. J Political Econ 102(5):841–877
Blume A, Kim Y-G, Sobel J (1993) Evolutionary stability in games of communication. Games Econ Behav 5:547–575
Fudenberg D, Tirole J (1991) Game Theory. MIT Press, Cambridge
Gibbons R (1992) Game Theory for Applied Economists. Princeton University Press, Princeton
Kartik N, Ottaviani M, Squintani F (2007) Credulity, lies, and costly talk. J Econ Theory 134(1):93–116
Krishna V, Morgan J (2001) A model of expertise. Q J Econ 116(2): 747–775
Lo P-Y (2006) Common knowledge of language and iterative admissibility in a sender-receiver game. Technical report. Brown University
Manelli AM (1996) Cheap talk and sequential equilibria in signaling games. Econometrica 64(4):917–942
Milgrom P, Roberts J (1986) Price and advertising signals of product quality. J Political Econ 94(4):796–821
Noldeke G, Samuelson L (1997) A dynamic model of equilibrium selection in signaling markets. J Econ Theory 73(1):118–156
Noldeke G, Van Damme E (1990) Signalling in a dynamic labour market. Rev Econ Stud 57(1):1–23
Ottaviani M, Sørensen PN (2006) Professional advice. J Econ Theory 126(1):120–142
Rabin M (1994) A model of pre-game communication. J Econ Theory 63(2):370–391
Ramey G (1996) D1 signaling equilibria with multiple signals and a continuum of types. J Econ Theory 69(2):508–531
Rasmusen EB (2006) Games and Information: An Introduction to Game Theory, 4th edn. Blackwell, New York
Riley JG (1979) Informational equilibrium. Econometrica 47:331–359
Sobel J, Stole L, Zapater I (1990) Fixed-equilibrium rationalizability in signaling games. J Econ Theory 52(2):304–331
Spence AM (1973) Job market signaling. Q J Econ 90:225–243
Swinkels JM (1999) Education signallling with preemptive offers. Rev Econ Stud 66(4):949–970

Single Column Modeling of Atmospheric Boundary Layers and the Complex Interactions with the Land Surface

Albert A. M. Holtslag, Gert-Jan Steeneveld
Department of Meteorology and Air Quality, Wageningen University, Wageningen, The Netherlands

Article Outline

Glossary
Definition of the Subject
Introduction
Background
Atmospheric Boundary-Layer Structure
Modeling Basics
Local Mixing Parameterization
More Advanced Mixing Parameterizations
Intercomparison of Single Column Models for Stable Conditions
Modeling Boundary Layers over Land
Impact of Land Surface Conditions on Model Results
Summary
Acknowledgments
Bibliography

Glossary

Atmospheric boundary layer The Atmospheric Boundary Layer (ABL) is the lower part of the atmosphere which is directly influenced by the presence of the earth's surface. As such its major characteristics are turbulence and the diurnal cycle.

Diurnal cycle The depth of the dry atmospheric boundary layer (ABL) can vary over land between tens of meters during night up to kilometers during daytime (see Fig. 2). Over sea the depth is often typical a few hundred meters and rather constant on the time scale of a day.

Turbulence Turbulence in the atmospheric boundary layer is the three-dimensional, chaotic flow of air with time scales typically between a second and an hour. The corresponding length scales are from a millimeter up to the depth of the boundary layer (or more in the case of clouds). Turbulence in the ABL originates due to friction of the flow and heating (convection) at the surface.

Definition of the Subject

In this article we deal with the single column modeling of the Atmospheric Boundary layer (ABL) and the the complex interactions which may occur with the land surface. As such we review the major characteristics of the ABL over land, and summarize the basic parameterizations for the represesentation of atmospheric turbulence and the surface fluxes. The modeling principles are illustrated with the outcome of single-column models for a variety of conditions using field data and fine-scale model results. Our emphasis is on stable conditions which occur over land at night-time under clear skies. For readers not familiar with atmospheric turbulence and meteorological definitions, some background and basic definitions are also given.

Introduction

The Atmospheric Boundary Layer (ABL) is generally characterized by turbulence. Because of its capability to mix air with different properties efficiently, the representation of turbulence is directly relevant for atmospheric and environmental modeling. For instance, turbulence directly impacts on the transfer of momentum, sensible heat, water vapor, ozone, and methane, among many other quantities, between the earth's surface and the atmosphere. Turbulence also defines the mixing of properties inside the atmospheric boundary layer, the transfer of quantities between the boundary layer and the clear or cloudy atmosphere aloft, and the mixing inside clouds.

Turbulence in the ABL is mainly due to the mechanical turbulence by vertical wind shear and turbulence by convection. Most of the atmosphere above the ABL is not turbulent, although turbulence can occur throughout the whole atmosphere. For instance, cumulus-type clouds, which may grow into thunderstorms, are always turbulent through convection produced by the heat released due to the condensation of water vapor. Turbulence can also occur in clear air above the ABL; most of this is produced in layers of strong vertical wind shear at the boundary between air masses (so-called 'Clear-Air Turbulence').

Because of the mixing capacity of turbulence, modeling atmospheric boundary layers is also relevant for many practical applications. For instance, chimney plumes are diluted and spread over larger volumes than they would be without turbulence. As such, strong local peaks of pollution are prevented and otherwise clean air is polluted. In practice turbulence may also cause engineering problems, because it shakes structures such as bridges, towers, and airplanes, causing failure of such systems in extreme cases. Turbulent fluctuations in the horizontal motions during severe storms can be fatal to tall buildings or bridges, particularly if resonance (e. g., forcing of a system at its natural frequency) occurs.

The correct formulation of the overall effects by turbulence, either inside or outside the atmospheric boundary layer, is an essential part of atmospheric models dealing with the prediction and study of weather, climate and air quality. These models are based on solving the equations dealing with atmosphere behavior. With state-of-the-art computers, the number of grid points in atmospheric models is limited to a number of typically 10^8 or

so. This implies that on the regional and global scale the atmospheric model equations are usually applied too fairly large 'air boxes'. Such boxes are often in the order of ten to hundred kilometers wide and ten to a few hundred meters thick. In these large boxes, smaller scale motions make air parcels interact and mix. For example, if a hot parcel is located next to a cold parcel, turbulent motion at their boundaries will heat the cool and cool the hot parcel. Thus, a closure formulation is needed to reproduce mixing by the turbulent motions into the model-resolved scales using the equations for the larger-scale 'mean' motions. It is important to realize that the closure formulation needs to be expressed in terms of variables available in the modeling context. This is called a 'parameterization'.

In this contribution we provide an overview of the modeling principles, the turbulent closures and parameterizations in use for of the atmospheric boundary layer, where we emphasize the modeling and parameterization of turbulence in the atmospheric boundary layer without clouds. Additionally we discuss the performance of models in current use [7,28], and we study the impact of the surface boundary condition over land [18].

Background

Atmospheric models for the forecasting and study of weather, climate, and air quality are typically based on integration of the basic equations governing atmospheric behavior. These equations are the gas law, the equation of continuity (mass), the first law of thermodynamics (heat), the conservation equations for momentum (the so-called "Navier–Stokes equations"), and usually equations expressing the conservation of moisture, trace gases and air pollutants. At one extreme, atmospheric models may deal with the world's climate and climate change; at the other, they may account for the behavior of local flows at coasts, in mountain-valley areas, or even deal with individual clouds. This all depends on the selected horizontal modeling domain and the available computing resources.

Since there is an enormous range of scales in atmospheric motion and turbulence, there is a need to separate the scales of atmospheric turbulence from larger-scale motions. Let C denote an atmospheric variable, such as specific humidity. Then $\overline{C}$ represents a mean or "smoothed" value of C, typically taken on a horizontal scale of order 10 (or more) km and a corresponding time scale in the order of 10 min to one hour. A local or instantaneous value of C would differ from $\overline{C}$. Thus, we have

$$C = \overline{C} + c\,. \tag{1}$$

Here c represents the smaller-scale fluctuations. Note that we use lower case for the latter (often primes are used as well to indicate fluctuations). In principle, the fluctuations around the mean motion also reflect gravity waves and other smaller scale motions, in addition to turbulence. Gravity waves often co-exist with turbulence or are generated by turbulence. If the wind at the same time is weak, there may be no turbulence at all. Anyhow, if turbulence exists, it is usually more important for most atmospheric applications, because it mixes more efficiently than the other small-scale motions.

To make the mathematical handling of c tractable, it must satisfy the so-called "Reynolds postulates". These require, for example, that $\overline{c} = 0$ and that small- and larger-scale values must not be correlated. After a quantity has been averaged to create a larger-scale quantity, further averaging should produce no further changes, in order for this postulate to apply. The mean of the summation of two variables A and C will produce $\overline{A \pm C} = \overline{A} \pm \overline{C}$. A further condition is that a mean variable $\overline{C}$ must be differentiable, since differentials show up in the atmospheric equations (see below). In practice, not all these conditions are rigorously satisfied. If the Reynolds postulates are fulfilled, then the averaging for the product of two variables provides

$$\overline{AC} = \overline{A}\,\overline{C} + \overline{ac}\,. \tag{2}$$

The second term at the right hand side of Eq. (2) is known as the turbulent covariance. Similarly, the turbulence variance of a quantity is given by $\overline{C^2} - (\overline{C})^2$ (which is the square of the standard deviation).

If in Eq. (2), the variable A represents one of the velocity components (U, V, W in the x, y, z direction, respectively), then $\overline{AC}$ is the total flux of C and the second term at the right hand side of Eq. (2) represents a turbulent flux of C. For instance, $\overline{uc}$ and $\overline{wc}$ are the horizontal and vertical turbulent fluxes of the variable C, respectively. Here u and w are the turbulent fluctuations of the horizontal and vertical velocities. Near the surface, the mean vertical wind $\overline{W}$ is usually small, and thus the total vertical fluxes are normally dominated by the turbulent contributions.

Atmospheric Boundary-Layer Structure

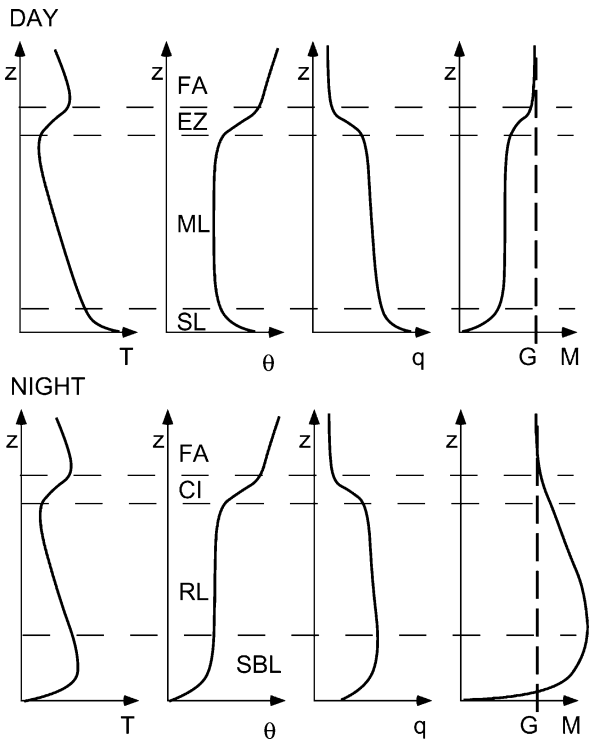
Turbulent fluctuations, variances and fluxes of variables are influenced by the vertical boundary-layer structure. Here the variation of temperature in the atmospheric boundary-layer plays an important role. Since pressure decreases with altitude, air parcels, which are forced to rise (sink), do expand (compress). According to the first law of thermodynamics, a rising (sinking) parcel will cool

(warm) if there is no additional energy source such as condensation of water vapor. Then this is called a dry adiabatic process.

It can be shown that in the atmospheric boundary layer, the temperature (T) variation with height for a dry adiabatic process is $\mathrm{d}T/\mathrm{d}z = -g/C_\mathrm{p}$ (here g is gravity constant and C_p is specific heat at constant pressure). The value for g/C_p is approximately 1 K per 100 m. An atmospheric layer which has such a temperature variation with height, is called neutral for dry air (at least when there is no convection arising from other levels). In that case $\Theta = T + (g/C_\mathrm{p})z$ is constant, where Θ is called the potential temperature (Note that the previous definition for potential temperature is not accurate above the boundary layer). Since air normally contains water vapor and because moist air is lighter than dry air, we have to correct for the influence of this on vertical motions. Consequently, a virtual potential temperature is defined as $\Theta_\mathrm{v} = \Theta(1 + 0.61q)$, where q is the specific humidity (defined as the mass of water vapor per unit mass of moist air).

In a neutral layer with constant Θ_v, vertical motions of moist (not saturated) air can maintain themselves. If the virtual potential temperature of the atmospheric layer increases with height, vertical displacements are suppressed. This is called a stable condition (or 'inversion'). At the other hand, when the virtual potential temperature decreases with height, vertical fluctuations may be accelerated. Consequently this is called an unstable condition. Thus in considerations with turbulent fluctuations and atmospheric stability, we have to deal with the virtual potential temperature and not with the actual temperature. Similarly, the vertical flux of sensible heat is connected to turbulent fluctuations of (virtual) potential temperature; e. g. it reads as $\overline{w\theta}_\mathrm{v}$ (in m K/s). The latter relates directly to the energy per time and unit area H by $H = \rho C_\mathrm{p}\overline{w\theta}_\mathrm{v}$ (in W/m^2), where ρ is density of the air (in kg/m^3).

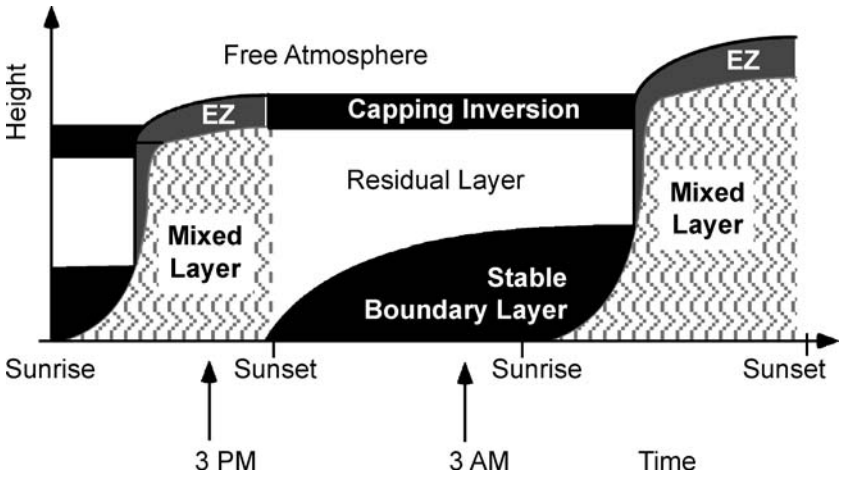
Figure 1 (after [30]), provides the typical, idealized, mean vertical profiles for temperature T, potential temperature Θ, specific humidity q, in addition to the horizontal wind M (defined by $M^2 = U^2 + V^2$). These profiles apply for an atmospheric boundary layer over land in clear sky conditions in the afternoon and around midnight. Note that in the free atmosphere the horizontal wind is mostly a result of the acting of the larger scale pressure differences and the Coriolis force due to the rotation of the earth (but other effects may play a role as well). The resulting wind is known as the 'geostrophic' wind and indicated with G in Fig. 1 (see dashed line). In the daytime boundary layer the actual wind is smaller due to surface friction, while at clear nights the actual wind away from the surface may be substantially stronger than G due to inertial effects (resulting in the so-called 'low level jet').



Single Column Modeling of Atmospheric Boundary Layers and the Complex Interactions with the Land Surface, Figure 1
Idealized vertical profiles of mean variables in the Atmospheric Boundary Layer over land in fair weather (after [30]). See text for additional information

The temporal variation of the mean boundary-layer profiles over land can be quite substantial due to the strong diurnal variation of solar incoming radiation and the nighttime cooling at the land surface. During daytime the turbulent boundary layer may grow to several kilometers into the non-turbulent 'free atmosphere' (indicted as FA in Fig. 1). At night the turbulent part of the stable boundary layer (SBL) may only extend up to a few hundred meters or less (the lowest dashed line in the lower figure). An idealized picture for the temporal variation of the boundary layer over land is given in Fig. 2 (after [30]). Here the arrows with local time indications refer to the day and nighttime figures of Fig. 1.

Figure 1 also indicates that the boundary layer during daytime shows a three-layer structure: an unstable 'surface layer (SL)', a 'well-mixed layer (ML)' with rather uniform (virtual) potential temperature, and a stably-stratified 'entrainment zone (EZ)'. In the latter zone, turbulence acts to exchange heat, momentum, water-vapor and trace gasses



Single Column Modeling of Atmospheric Boundary Layers and the Complex Interactions with the Land Surface, Figure 2
Idealized diurnal evolution of the Atmospheric Boundary Layer over land in fair weather (after [30])

between the boundary layer and the free atmosphere. During nighttime, often the vertical structure of the previous day persists above the SBL. As such a 'residual layer (RL)' with sporadic turbulence (remaining from the previous day) can be identified as well as a 'capping inversion (CI)'.

Modeling Basics

The challenge of modeling the atmospheric boundary layer is the prediction of the temporal variation of the vertical and horizontal structures in response to the influence of the major processes acting in the atmosphere and at the earth's surface. As such the governing equations have to be integrated. In practice, the variables are split into 'mean' larger-scale motions and smaller-scale fluctuations as in Eq. (1). Inserting this into the basic equations and after averaging this provides a set of equations for the behavior of the larger-scale (mean) variables. The larger-scale variables are then used explicitly in atmospheric models. This can be demonstrated as follows below.

The general character of any of the budget equations dealing with atmospheric motions is

$$\frac{\mathrm{D}C}{\mathrm{D}t} = S_i \,. \tag{3a}$$

Here S_i represents the subsequent sources and sinks for the variable C (such as radiation or chemistry effects). The notation $\mathrm{D}C/\mathrm{D}t$ represents the total rate of change for the variable C by local changes $(\partial/\partial t)$, and changes transported with the fluid motion in the three directions. As such, we have

$$\frac{\partial C}{\partial t} + U\frac{\partial C}{\partial x} + V\frac{\partial C}{\partial y} + W\frac{\partial C}{\partial z} = S_i \,. \tag{3b}$$

Here U, V, W are the wind speeds in the three directions x, y, z, respectively.

If in the atmospheric motion each variable is split into a mean component and a fluctuation then (3b) provides after Reynolds-averaging, some algebraic manipulations and simplifying assumptions, a budget equation for the mean variable $\overline{C}$. This reads as

$$\begin{aligned}\frac{\mathrm{D}\overline{C}}{\mathrm{D}t} &= \frac{\partial \overline{C}}{\partial t} + \overline{U}\frac{\partial \overline{C}}{\partial x} + \overline{V}\frac{\partial \overline{C}}{\partial y} + \overline{W}\frac{\partial \overline{C}}{\partial z} \\ &= \overline{S_i} - \frac{\partial \overline{uc}}{\partial x} - \frac{\partial \overline{vc}}{\partial y} - \frac{\partial \overline{wc}}{\partial z} \,.\end{aligned} \tag{4}$$

We may note that in the derivation of (4), single terms representing fluctuations have disappeared (as above in Eq. (2)). However, terms involving the product of two fluctuations did remain.

Thus because the basic equations are nonlinear, the budget equations for the mean variables contain terms involving smaller-scale motions. The latter terms are of the form of a divergence of fluxes produced by such motions in the three directions and appear as the last three terms in Eq. (4). These motions are said to be sub-grid and consequently, closure formulations or parameterizations are needed to introduce mixing by the smaller-scale, sub-grid, motions into the equations for the larger-scale motions (as resolved by the model). Note that additional terms may also appear in (4) when the source or sink term S_i incorporates nonlinear effects (such as in the case of chemistry).

The atmospheric model equations can also be applied on much smaller spatial and temporal scales then discussed here, for instance by using vertical and horizontal grid elements of 10 to 100 m, and time steps of seconds

only. It is important to realize that in such cases a significant part of the turbulent fluctuations are resolved by the model equations. This type of modeling is nowadays known as 'Large-eddy simulation (LES)'. This has become a powerful and popular tool in the last decade to study turbulence in clear and cloudy boundary layers under well-defined conditions. It is important to realize that in the case of LES the simplifying assumptions leading to Eq. (2) are normally not valid.

A special and simple form of Eq. (4) arises for horizontally homogeneous conditions. In such cases the terms including horizontal derivatives are negligible. If in addition the mean vertical wind is small and if there are no other sources and sinks, then (4) provides

$$\frac{\partial \overline{C}}{\partial t} = -\frac{\partial \overline{wc}}{\partial z} \,. \tag{5}$$

This equation is known as the one-dimensional, vertical diffusion equation. It shows that the local time rate of change for the mean of a variable (such as temperature or wind) at a certain height, is given by the divergence of the turbulent (corresponding heat or momentum) flux in the vertical direction. As such, information on the turbulent flux may produce a local forecast of the variation of a mean variable (but only under the simplifications mentioned).

Equation (5) can be seen as the basis of a single-column model where only local information of the atmosphere is relevant. However, normally the other terms in (4) are also relevant, in particular the terms with mean wind speed (the so-called "advection terms"). This means that in general the budget equations for momentum, heat, and the various scalars are closely coupled in any atmospheric model. Still one can solve for the local time rate of change in a single column model once the advection terms are known form observations or other means. This approach is widely adopted to study atmospheric boundary layers in comparison with observations on a local scale.

Before we proceed with more detailed parameterizations for the fluxes in the boundary layer, let us deal with the derivation of the surface fluxes. These fluxes enter as boundary conditions when solving the budget equations for all the relevant mean variables (in any approach). It is important to realize that near the surface, the average wind must vanish because the mean wind is zero at the earth's surface. At the other hand, we know from observations that the fluxes of heat, momentum and trace gasses are nonzero. Consequently, it is convenient to model an 'effective' surface flux $\overline{wc_0}$ of a conserved variable due to the combined effect of molecular diffusion and turbulence at the surface. This can be achieved by writing

$$\overline{wc_0} = \beta_t w_t (C_0 - C_a) \,. \tag{6}$$

Here C_0, and C_a are the values of the transported variable at the surface and in the air, respectively; β_t is a transfer coefficient, and w_t is an effective transport velocity representing the turbulence. For example, in near-neutral conditions the effective transport velocity is well represented by the well-known surface friction velocity u_{*0}. Then it can be shown that $\beta_t = \kappa / \ln(z/z_0)$, where κ is the 'Von Karman' constant (often specified as $\kappa \cong 0.4$), z is the corresponding height of C_a in the lowest part of the boundary layer and z_0 is the so-called surface roughness length for the variable C. We refer to the literature for a more detailed treatment (e. g., Beljaars and Holtslag [4]).

Local Mixing Parameterization

To solve the budget Eq. (4) for all the mean atmospheric variables involved, the terms involving turbulent fluxes need to be parametrized. As mentioned before, this means that the fluxes need to be expressed in terms of available mean model quantities, both in the atmosphere and at the surface. Once this has been achieved, the atmospheric model equations can be integrated. Thus, starting with proper initial values, new values can be calculated for the following time step and so on.

The most frequently used parameterization for environmental and atmospheric models, is known as first-order closure or often also called K-theory. In this theory it is assumed that the flux $\overline{wc}$ of a variable C in the vertical direction z, is down the vertical gradient of the mean concentration of C per unit mass. Thus

$$\overline{wc} = -K_c \frac{\partial \overline{C}}{\partial z} \,. \tag{7}$$

Here, K_c is known as the 'eddy-diffusivity' or mixing coefficient for the variable C. Similarly, the horizontal fluxes can be represented in terms of horizontal gradients. Note that the corresponding eddy-diffusivities typically are not constant, but that they generally depend on properties of the flow and the variable of interest. This also means that normally no analytic solutions are possible, not even for the simple case in which Eqs. (5) and (8) are combined.

We may note that the dimension of an eddy-diffusivity is a length scale ℓ times a velocity scale. These are proportional to the products of effective eddy sizes and eddy velocities in the corresponding directions. Often a diagnostic expression is used for the eddy-diffusivity, on basis of what is called 'mixing length theory' (in analogy with molecular

diffusion). The result reads as

$$K_c = \ell^2 S f(\mathrm{Ri})\,. \tag{8}$$

Here S is vertical wind shear (that is the variation of mean horizontal wind with height). Note that the combination ℓS in Eq. (9) has units of velocity. In Eq. (8), $f(\mathrm{Ri})$ denotes a functional dependence on local stability as represented by the gradient Richardson-number Ri defined by

$$\mathrm{Ri} = \frac{g}{\overline{\Theta_v}} \frac{\partial \overline{\Theta_v}/\partial z}{(\partial \overline{U}/\partial z)^2 + (\partial \overline{V}/\partial z)^2}\,. \tag{9}$$

Here g is the acceleration due to gravity, and $\overline{\Theta_v}$ is the mean 'virtual potential temperature'.

The specification of the length scale ℓ is not at all straightforward, except near the surface where so-called 'surface-layer similarity theory' (see cited literature) provides that $\ell \propto z$. A frequently used form for ℓ is:

$$\frac{1}{\ell} = \frac{1}{\kappa z} + \frac{1}{\lambda}\,. \tag{10}$$

Here λ is a turbulent length scale, which should be valid for the turbulence far above the surface. We note that the latter has a rather empirical nature and consequently there is no agreement on the specification of λ in the literature.

Equation (8) is a diagnostic equation, which indicates that the eddy-diffusivity may variy with height, wind speed, stability, et cetera. In combination with the flux parameterization of Eq. (7), it follows that the flux at a certain height depends on the local gradient of the mean variable involved. Consequently the approach is referred to as a 'diagnostic local mixing approach'. Such an approach is mostly suitable for relatively homogeneous conditions with neutral and stable stratification, and is not so suitable for cases with convection (see non-local mixing parameterizations below).

More Advanced Mixing Parameterizations

A physically realistic alternative to the diagnostic approach is to relate the eddy-diffusivity of Eq. (7) to the actual turbulent kinetic energy of the flow, by using the prognostic turbulent kinetic energy equation and an appropriate choice for the turbulent length scale. It is important to realize that the kinetic energy of atmospheric motion per unit of mass E is given by the half of the sum of the velocities squared in the three directions (as in classic mechanics), e. g. $E = (U^2 + V^2 + W^2)/2$. Similar as with respect to Eq. (2), we can separate between the Mean Kinetic Energy $\overline{E}$ of the mean atmospheric motions and the Turbulent Kinetic Energy (TKE or e) of the smaller-scale fluctuating motions by turbulence. Thus e is given by $e = (\overline{u^2} + \overline{v^2} + \overline{w^2})/2$.

The prognostic equation for e reads in its basic form as:

$$\frac{\mathrm{D}e}{\mathrm{D}t} = -\overline{uw}\frac{\partial \overline{U}}{\partial z} - \overline{vw}\frac{\partial \overline{V}}{\partial z} + \frac{g}{\overline{\Theta_v}}\overline{w\theta_v} + D - \varepsilon\,. \tag{11}$$

Here $\mathrm{D}e/\mathrm{D}t$ is the total variation of e with time (the sum of local variations and those transported with the mean air motion). The two terms at the immediate right hand side of (11) represent the shear production of turbulence. These depend primarily on vertical variations of wind or, near the ground, on wind speed and surface roughness. The terms are almost always positive. The third term in Eq. (11) represents the rate of production or breakdown of turbulence by buoyancy effects (such as heat convection). It depends directly on density effects, which can be written in terms of the virtual potential temperature $\overline{\Theta_v}$, and its turbulent flux $\overline{w\theta_v}$. The term D in Eq. (11) represents divergence and pressure redistribution terms. These have a tendency to cancel near the surface. Finally, the term ε reflects the molecular dissipation of turbulence into heat and this term is always positive. In fact ε is typically proportional to e/τ, where τ is the characteristic time scale for the turbulent mixing process.

Using Eq. (11), turbulent kinetic energy can be calculated for given mean profiles when the corresponding fluxes are calculated using Eq. (7) for all fluxes involved. In this approach the diffusivities are typically calculated with equations of the form

$$K_c = \alpha_c \ell \sqrt{e}\,. \tag{12}$$

Here α_c is a constant depending on the variable of interest. The length scale is typically calculated with a similar type of diagnostic equation as (10) provides. This approach is known as the 'TKE-length scale approach' and it is an example of so-called 1.5 order closure. Sometimes a prognostic equation is used for the length scale as well, but such an approach is more popular in engineering applications then in the atmospheric sciences.

It can be shown that Eq. (8) is a solution of (11) and (12) in stationary conditions and when other simplifications are made such as the neglect of the influences by advection and turbulence divergence in the TKE equation. A more advanced turbulence scheme is known as 'second-order closure'. In such an approach, prognostic equations are developed for the fluxes and variances themselves. Such equations have a very similar structure as Eq. (12) for kinetic energy. Unfortunately, new unknowns are present in these equations. These must be related to the

other variables in the model equations, always involving assumptions. Thus, second-order closure involves many more than the original equations and is therefore computationally more time consuming ('expensive') than first-order and 1.5 order closure.

One may expect that a model with 1.5 or second-order closure would produce more realistic results then a model with a first order closure. However, in practice this is often not the case, because of complex model interactions and the difficulty of representing all the relevant details with sufficient accuracy (see also below). That is the reason why diagnostic approaches remain popular. Nevertheless, second order equations are useful to gain insight in the governing physics, and after simplification useful extensions of the basic parameterizations may be achieved.

We continue our discussion with mixing parameterizations which have been proposed for boundary-layers with strong atmospheric convection. In such cases, the turbulent flux of a conserved quantity is typically not proportional to the local gradient alone as predicted by Eq. (7). In fact, in a large part of the ABL the mean gradients are small in conditions with dry convection, in particular for potential temperature (see Fig. 1). Then the fluxes depend mostly on the mixing characteristics of the large eddies across the ABL. Theories are available, which have modified K-theory to allow for the influence of convection, for example by including additional terms at the right hand side of Eq. (8) For details we refer to the literature (e. g., Holtslag and Moeng [15]).

In the next sections we apply the modeling concepts above and compare their results with field observations. In addition we present results from model intercomparison studies, and illustrate the role of boundary conditions.

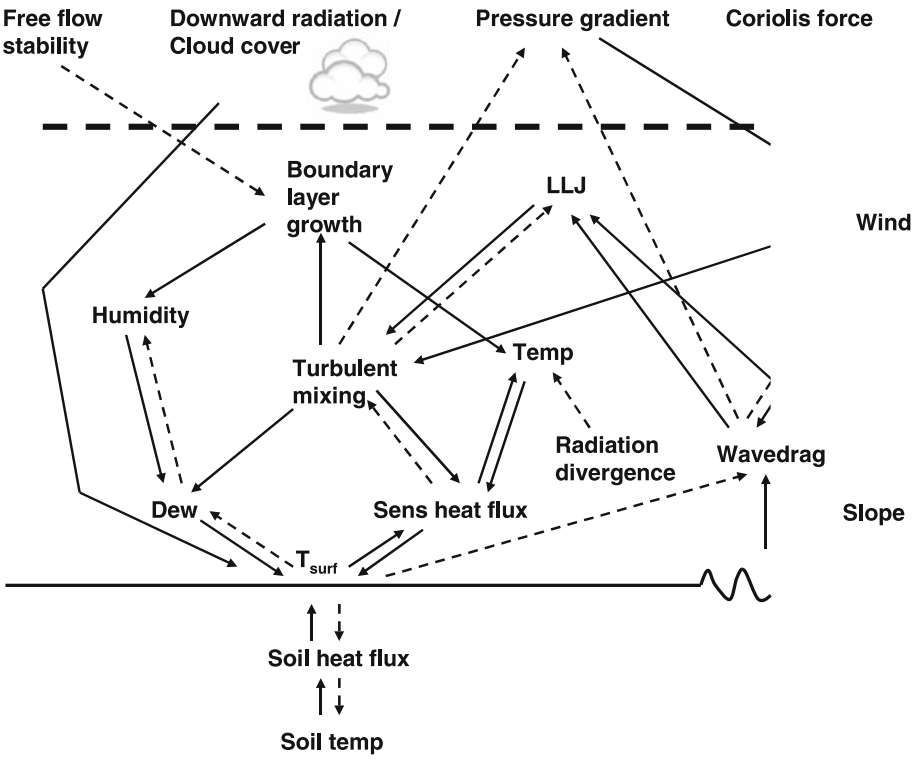
Intercomparison of Single Column Models for Stable Conditions

Atmospheric models for weather and climate need to make an overall representation of the smaller-scale boundary-layer and near surface processes. This appears to be more successful during daytime (e. g. Ek and Holtslag [10]; Holtslag and Ek [17]) then during nighttime stable conditions over land. The modeling of the stable boundary layer over land is rather complex because of the many different physical processes which are "at work" in stable conditions [21]. These small-scale processes are: clear air radiation divergence, drainage flow, generation of gravity waves and shear instabilities, fog and dew formation, the occurrence of a low-level jet and generation of discontinuous or intermittent turbulence [33]. In addition, the phenomenology of stable atmospheric boundary layers is quite divers, e. g. shallow and deep boundary layers with continuous turbulence through most of their depth, and on the other hand boundary layers with intermittent turbulence or even laminar flow.

The small-scale processes influence the vertical and horizontal exchange of quantities between the surface and the atmosphere as well as the mixing in the atmosphere on a variety of scales. In addition, it is known that turbulent mixing in stratified flow has an inherent non-linear character and may, as such, trigger positive feedbacks. These positive feedbacks, in turn, may cause unexpected transitions between totally different SBL regimes (e. g. van de Wiel et al. [34]).

Figure 3 depicts the interactions between relevant processes in the stable boundary layer. The non-linear behavior of the system is seen in e. g. the surface sensible heat-flux (H). A sudden change of the surface temperature can result in 2 different impacts on H. First, in weakly stable conditions (with strong wind and sufficient turbulence), a surface temperature decrease will provide a larger heat flux since H is proportional to the temperature difference between the surface and the atmosphere. The larger heat-flux from the atmosphere to the surface compsensates for the stronger cooling. In contrast for stronger stably stratified conditions, a surface temperature decrease will provide a stronger stratification and inhibits turbulent mixing, and consequently H will decrease. This allows for even stronger surface cooling (positive feedback). Note that a similar diagram for the daytime boundary layer can be found in Ek and Holtslag [10].

Having in mind the above mentioned complexity, one should not be surprised that atmospheric models encounter large forecast errors for stable conditions [20,24]. One strategy to improve model performance is to provide different models the same forecasting task, and analyze which model descriptions are in favor for which atmospheric stability. Recently such an intercomparison of boundary-layer schemes for stable conditions was made within the GEWEX Atmospheric Boundary Layer Study ('GABLS'). This GEWEX project aims to improve the understanding and the representation of the atmospheric boundary layer in regional and large-scale climate models [14]. A rather simple case was selected as a benchmark to review the state of the art and to compare the skills of single column (1D) models [7] and Large–Eddy Simulation models [3]. In this case a stable boundary layer is driven by an imposed, uniform geostrophic wind, with a specified constant surface-cooling rate over (homogeneous) ice. The case is initialized with $\theta = 265\,\mathrm{K}$ for $0 < z < 100$ and a lapse rate of 1 K/100 m aloft.



Single Column Modeling of Atmospheric Boundary Layers and the Complex Interactions with the Land Surface, Figure 3
Interaction diagram for the processes and variables in the stable atmospheric boundary layer over land (after [26])

It turns out that with the same initial conditions and model forcings, the models indicate a large range of results for the mean temperature and wind profiles. Figure 4 shows the mean profiles for several models after nine hours of constant surface cooling (sufficient to achieve a quasi-steady state). The variable results achieved are strongly related to the details of the boundary-layer mixing schemes [7]. An important finding is that the models in use at operational weather forecast and climate centers (as depicted at the left hand side of Fig. 4) typically allow for enhanced mixing resulting in too deep boundary layers, while the typical research models (at the right hand sides) show less mixing in more in agreement with the 'Large Eddy Simulation' results for this case [3].

Because of the enhanced mixing in weather and climate models, these models tend to show a too strong surface drag, too deep boundary layers, and an underestimation of the wind turning in the lower atmosphere [19]. At the other hand, by decreasing the mixing and surface drag, a direct impact on the atmospheric dynamics ('Ekman pumping') is noted (e. g. Beljaars and Viterbo [5]). Consequently, cyclones may become too active, corresponding in too high extremes for wind and precipitation, etc.

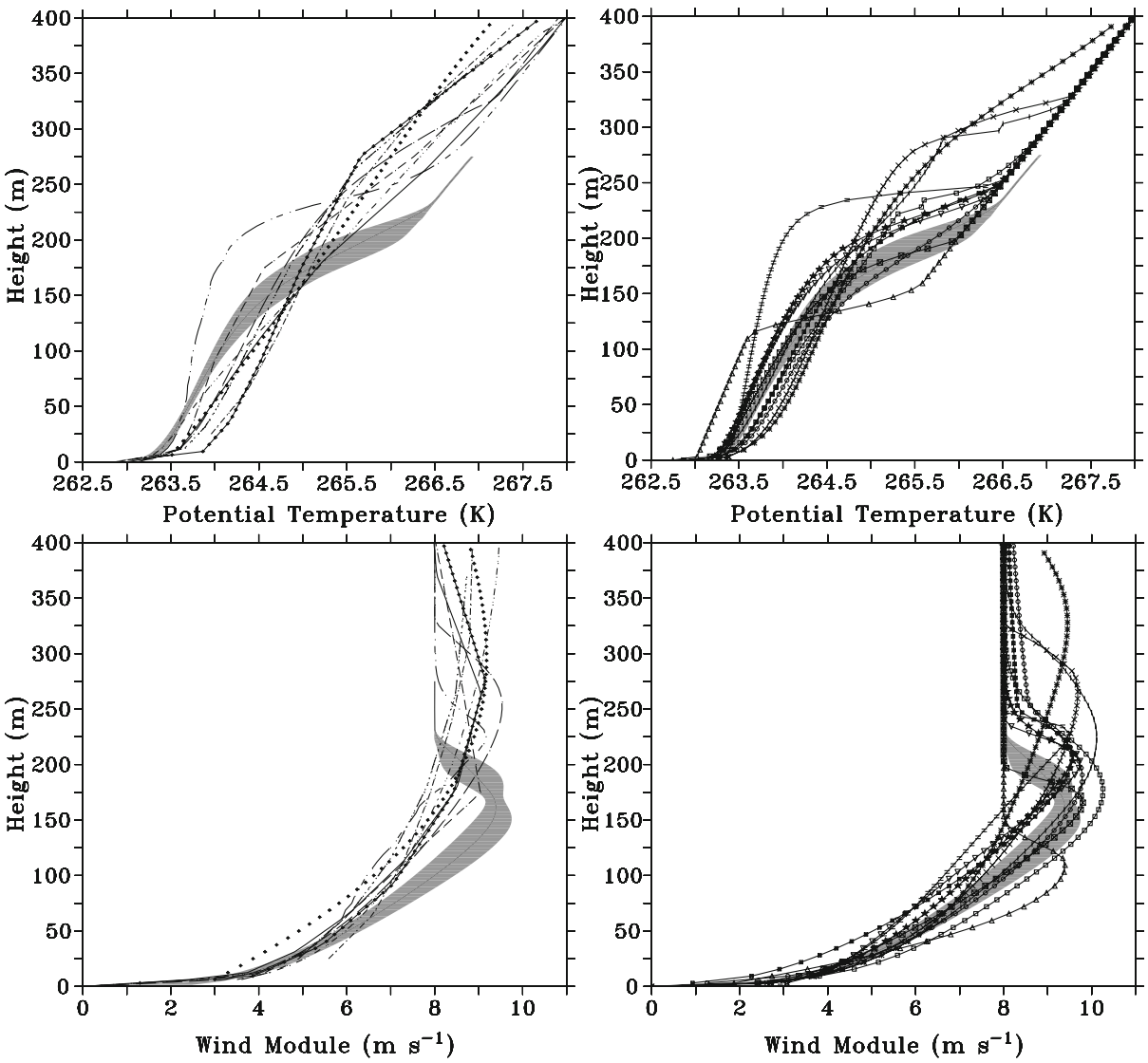
Modeling Boundary Layers over Land

To study the interactions of the ABL with the land surface we utilize the model by Duynkerke [9] with the extensions and modifications by Steeneveld et al. [28]. This model has been validated against tower observations at Cabauw, The Netherlands and later with CASES-99 field observations [1,28], and also participated in the GABLS model comparison in the previous section.

Instead of prescribing the surface temperature, and to enable interaction between the ABL and the land surface, the model is extended with a soil and a vegetation layer The soil temperature evolution is calculated by solving the diffusion equation (using a grid spacing of 1 cm) and the heat flux G_h from the soil to vegetation is calculated by:

$$G_h - (1 - f_{veg})K^{\downarrow} = r_g(T_{veg} - T_{s0}) \,. \tag{13}$$

In Eq. (13) $K^{\downarrow}$ is the incoming shortwave (solar) radiation, T_{veg} represents the vegetation surface temperature, and T_{s0} the soil temperature just below the vegetation. We use a vegetation fraction $f_{veg} = 0.9$ and conductance $r_g = 5.9\,\mathrm{Wm}^{-2}K^{-1}$, which are consistent with the obser-



Single Column Modeling of Atmospheric Boundary Layers and the Complex Interactions with the Land Surface, Figure 4
Modeled potential temperature and wind profiles by an ensemble of column models (after 9 h). *Grey areas* indicate the ensemble of Large Eddy Simulation results [3]. *Left panel* shows the results for first order closure models and the *right panel* for higher order closure models (after [7])

vations of CASES99 [28]. Initial soil and surface temperatures are also taken from the CASES99 observations.

Subsequently, the evolution of T_{veg} is computed by solving the surface energy budget for the vegetation layer:

$$C_{\mathrm{v}} \frac{\partial T_{\mathrm{veg}}}{\partial t} = Q^* - G_{\mathrm{h}} - H - L_{\mathrm{v}} E \,. \qquad (14)$$

Here C_{v} is the heat capacity of the vegetation layer per unit of area ($C_{\mathrm{v}} = 2000\,\mathrm{Jm^{-2}K^{-1}}$, van de Wiel [34]), Q^* is the net radiation, H is the sensible heat flux and $L_{\mathrm{v}}E$ the latent heat flux. The turbulent fluxes are calculated basically with the format of Eq. (6) above. Finally, Q^* is calculated by adopting the Garratt and Brost [12] radiation scheme. Note that Eqs. (13) and (14) provide a rather strong coupling of the atmosphere to the vegetated land surface for the current parameter setting which is found to be important [28].

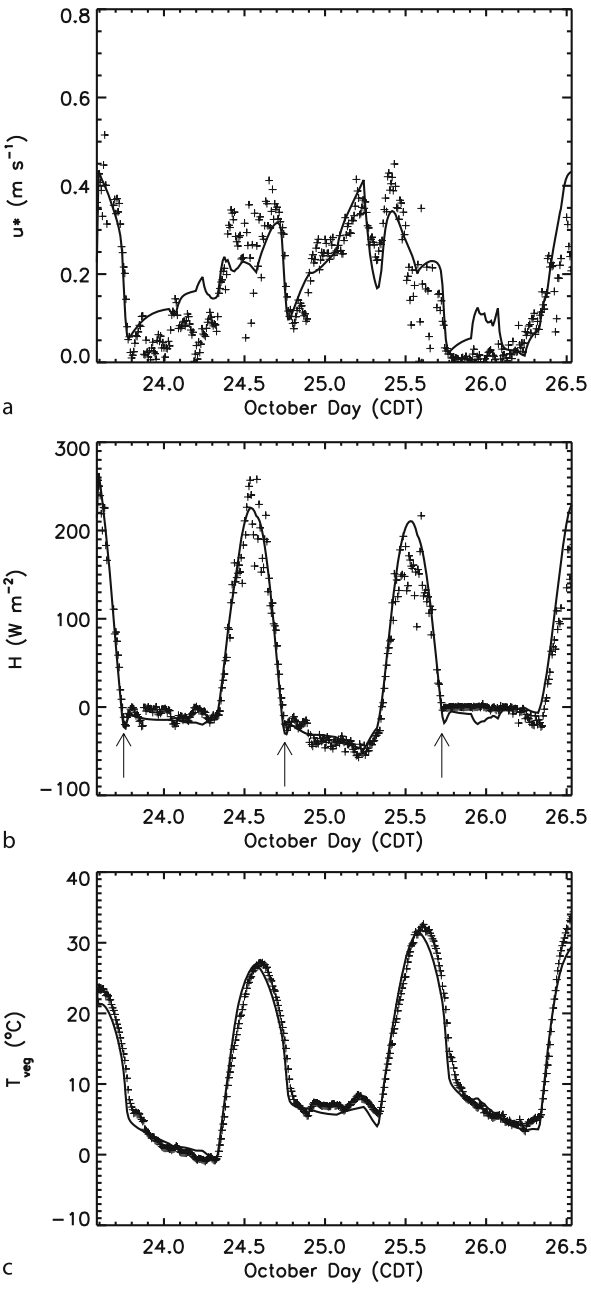
Model forecasts have been compared with CASES-99 observations for contrasting diurnal cycles (i. e. for different wind speeds) for 23–26 Oct. 1999. Note that the model

is only initialized once and that the total run comprises three full days. We distinguish between "radiative nights" with weak winds and small turbulent mixing, when radiative cooling dominates the SBL development. On the contrary, so-called "continuous turbulent" nights are characterized by strong winds. The final archetype is the so called "intermittently turbulent" night where turbulent episodes alternate with calm periods, when the radiative and turbulent forcings are of similar order of magnitude. Here we restrict ourselves to the results for surface fluxes, surface vegetation temperature, and vertical profiles of temperature and wind speed during nighttime. The diurnal cycle of the modeled net radiation Q^* (the balance of all incoming and outgoing short- and longwave radiative fluxes) agrees with the observations (not shown). Net radiation amounts typically $Q^* = 400\,\mathrm{Wm^{-2}}$ during daytime and $Q^* = -70\,\mathrm{Wm^{-2}}$ during nighttime.

The friction velocity (u_*) shows a clear diurnal cycle: u_* is large during the day and small at night, which is in general well captured by the model (Fig. 5a). Looking in more detail we find that during weak winds (1st and 3rd night) the model tends to overestimate u_*. The model lacks a clear turbulence collapse as observed during the first (intermittent) night. In the period 24 Oct., 700 CDT – 25 Oct., 1700 CDT the model performs well, while during the last (radiative) night u_* is slightly too high until midnight but follows the collapse at the end of the night. The overall bias amounts to $0.03\,\mathrm{ms^{-1}}$ for the last night. Sodar observations show much smaller wind speeds at 200 m AGL (which is above the SBL during this night) than the imposed **G**. This may suggest that **G** was overestimated, and this consequently may explain the bias in u_*. In general it is known that models correctly predict u_* for strong winds, but overestimate u_* for weak winds [25,32].

The sensible heat flux differs substantially between day ($\sim 250\,\mathrm{Wm^{-2}}$ here) and night (between 0 and $-60\,\mathrm{Wm^{-2}}$ depending on the wind speed). In the first (intermittent) night, the modeled $H = -14.1\,\mathrm{Wm^{-2}}$ on average, while $-9.1\,\mathrm{Wm^{-2}}$ was observed. However, the model does not simulate the observed intermittent character of the surface fluxes (Fig. 5b). Some models with more resolution [23,36] were also able to reproduce intermittent turbulence. On the other hand, the models by Sharan and Gopalakrishnan [25] and Derbyshire [8] did not show any intermittency. This subject needs further investigation.

Just after the day-night transition to the intermittent night, the observed magnitude of H shows a clear maximum (see arrows in Fig. 5b), which is well reproduced by the model. This maximum is caused by a sudden reversal of the stratification near the surface due to longwave radiation emission during the day-night transition,



Single Column Modeling of Atmospheric Boundary Layers and the Complex Interactions with the Land Surface, Figure 5
Modeled and observed (+) friction velocity (a), surface sensible heat flux (b), and surface vegetation temperature (c) for three diurnal cycles in CASES-99 (after [28])

and is maintained by residual turbulence of the convective boundary layer. This is an often observed *realistic feature* (e. g. during 11 of the 30 nights for CASES-99 and in FIFE observations shown by [6]). However, most model-

ing studies rarely show these minima. The reproduction of this detailed feature emphasizes the realism of the model outcome.

During the turbulent night (24–25 Oct.), the predicted H follows the observations. The specific minimum during the day-night transition (24 Oct., 1900 CDT) is present here as well. Radiative flux divergence dominates the last (radiative) night and the observed H is approximately zero. The model slightly overestimates the magnitude of H ($-2.9\,\mathrm{Wm^{-2}}$), mainly caused by an overestimation of u_*. This causes a weaker stratification and thus a larger magnitude of H. The second half of this night the model gives good results.

Reliable prediction of T_{veg} is a common problem for large-scale models. Some models show unphysical decoupling of the atmosphere from the surface resulting in so-called "runaway cooling" of T_{veg}. On the other hand, the pragmatic enhanced mixing approach which is commonly used for very stable conditions, leads to overestimation of T_{veg}. For both day- and nighttime T_{veg} *is simulated in very good agreement with the data, despite the fact that we cover a broad range of stability* (Fig. 5c).

We conclude that the present model generates surface fluxes which are in good agreement with observations, because of the detail in the description of the surface scheme, the soil heat flux and radiation physics (with high resolution). In general, the model is also able to estimate temperature and wind profiles (see results and discussion in Steeneveld et al. [28]). To examine the robustness of the results, we performed some sensitivity analysis on the initial conditions and model parameters. Disturbing the initial temperature (by 1 K), wind profiles (by 5%), soil temperature and vegetation temperature (both by 1 K) do not affect the results seriously. Also model re-initialization every 24 h (1400 CDT) with observed radiosonde information showed hardly any impact on the results (not shown).

Impact of Land Surface Conditions on Model Results

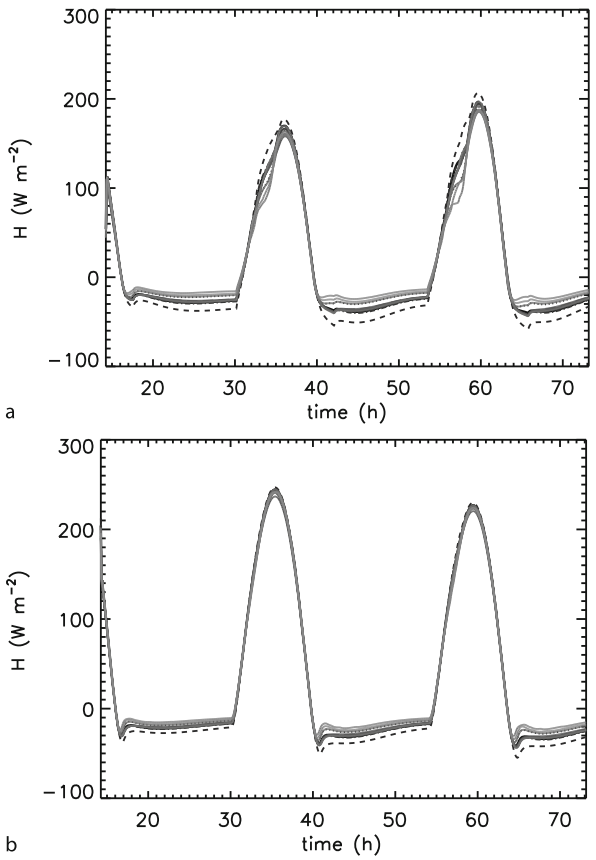
Inspired by the result in the previous section that a coupling with the land surface is necessary to obtain satisfactory model results, we now analyze the difference in variability of model results during model intercomparisons, as function of the chosen boundary condition. At first we use the first-order closure model and vary the parameters in the turbulence scheme for stable conditions in a reasonable range to mimic the apparent variability among boundary-layer models. As such first model runs are performed with a prescribed surface temperature as inspired by (but not identical to) the observations in CASES99 [22] and as described in the GABLS2 case description [31]. Subsequently, the model runs are repeated, but then using an interactive prognostic heat budget equation for the surface temperature (Eq. (14)).

To study the impacts of parameter values on the model results, reference runs are made for coupled and uncoupled cases with alternative permutations in some of the parameter settings for stable conditions. The parameter modifications are chosen such that they cover a realistic range in comparison with existing models of the stable boundary layer [7]. The local starting time in the model runs is 14.00 LT on October 22, 1999 (rather than 16.00 LT in the GABLS2 runs). The duration of all runs is 59 hours (so that the axis of all the figures indicates 14.00 until 73.00 h, covering a period of 2.5 diurnal cycles). In all model runs the roughness length for heat z_{oh} and momentum z_{om} (3 mm and 3 cm respectively), and the canopy resistance are constant, and the geostrophic wind is taken at a reference value of $9.5\,\mathrm{ms^{-1}}$ (as in Svensson and Holtslag [31]). The reference model set up has 50 logarithmically distributed layers and the first atmospheric model level is at 2 m.

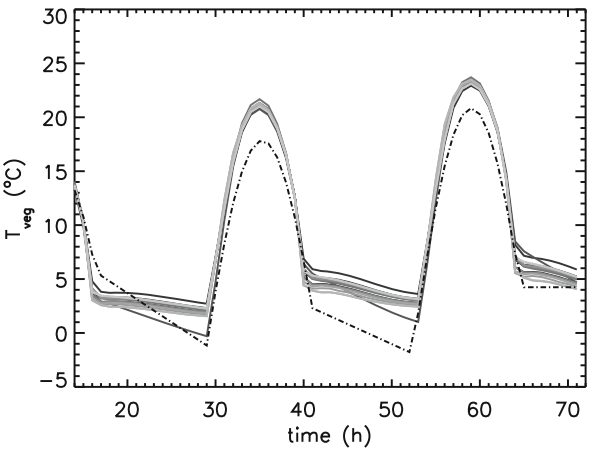
The model results for all parameter permutations are presented for the sensible heat flux (Fig. 6), In the upper sub-frame of the figure (labeled a) the results achieved with the uncoupled model are given (using prescribed surface temperature). Overall the variety of results in the upper frame is comparable to the variety within the GABLS2 intercomparison study in stable conditions for the uncoupled models (see Svensson and Holtslag [31]). Thus we have a range of -15 to $-50\,\mathrm{Wm^{-2}}$ for the sensible heat flux (at the end of the first night e. g. at the time of 30 h). The variability is a result of the range of parameters chosen above and the impact is apparently sufficient to mimic the different parameterizations for stable conditions in the models used within GABLS2.

Next we repeat all model runs and allow for surface feedback using Eqs. (13) and (14). The results for the sensible heat flux with the coupled model are given in the lower frame (Fig. 6b). Now we have a range of -10 to $-25\,\mathrm{Wm^{-2}}$ (again the values apply for the end of the first night e. g. at the time of 30 h). Thus it appears that the variety of model results is *smaller* for the sensible heat flux in the coupled case. At the same time it appears that the variability appears to be somewhat larger for friction velocity and boundary-layer depth, which seems to be related to the larger variability in the near surface air temperature and wind speed.

During daytime the sensible heat fluxes are rather similar for all model runs within one category (either coupled or uncoupled), but the maximum values differ. In addition, due to the coupling the sensible heat fluxes show



Single Column Modeling of Atmospheric Boundary Layers and the Complex Interactions with the Land Surface, Figure 6
Time series of an ensemble of model results for the sensible heat flux in a model intercomparison study with a prescribed surface temperatures, and b by solving the surface energy budget (after [18])



Single Column Modeling of Atmospheric Boundary Layers and the Complex Interactions with the Land Surface, Figure 7
Time series of an ensemble of modeled surface temperature for coupled runs. *Dash-dotted line*: prescribed surface temperature in the uncoupled case (after [18])

a more smooth behavior in the morning hours as compared with the uncoupled results. Thus, surface feedback is influencing the model results and is also able to compensate for some variation in the model parameter values. Note also that the variability in the friction velocities of the first night remains during the morning hours in the uncoupled runs, but not so much in the coupled case.

In Fig. 7 the surface temperatures are given as specified for the uncoupled case (the dashed line), and the temperatures as calculated in the various interactive model runs (various gray lines). It is seen that the latter ones are quite different from each other (in particular at night). It is also important to note that the surface temperature by the ensemble of coupled model runs is clearly different from the specified temperature in the uncoupled case. This impacts also on the absolute values and the range of air temperatures and the wind speeds.

Thus it is apparent that the treatment of the surface temperature impacts strongly on the outcome of the boundary-layer model results and their variety (see also Basu et al. [2]). By repeating the uncoupled model runs with a specified surface temperature as given by the ensemble mean value of the interactive runs, we achieve basically the same variety of model outputs for the potential temperature and wind as for the coupled cases. This confirms that in model evaluation studies the surface temperature should be taken consistent with the value of the geostrophic wind (although this may be model dependent).

Summary

In this paper a summary is given of the basic approaches for the modeling and parameterization of turbulence in the atmospheric boundary layer. The treated approaches are in current use in regional and global-scale models for the forecasting and study of weather, climate and air quality. Here we have shown the results of such approaches by using single column models in comparison with field data and fine-scale model results. We have also studied the impact of the surface temperature condition on the variability of results by an atmospheric boundary-layer model. From the coupled model results we achieve that surface feedback can compensate for some of the variety introduced by changing model parameters. Generally much work needs still to be done before we have a full understanding of the complexity of atmospheric turbulence and the interactions with the land surface. A better understanding of atmo-

spheric turbulence hopefully also contributes to our capability in refining and unifying the turbulence parameterizations for modeling of the atmospheric boundary layer in response to the different type of surfaces which are found in reality.

Acknowledgments

This contribution is a compilation of earlier works by the authors, in particular the works by Holtslag [13], Steeneveld et al. [28] and Holtslag et al. [18].

Bibliography

1. Baas P, Steeneveld GJ, van de Wiel BJH, Holtslag AAM (2006) Exploring Self-correlation in flux-gradient relationships for stably stratified conditions. J Atmos Sci 63:3045–3054
2. Basu S, Holtslag AAM, van de Wiel BJH, Moene AF, Steeneveld GJ (2007) An inconvenienth 'truth' about using the sensible heatflux as a surface boundary condition in models under stably stratified regimes. Acta Geophys 56:88–99. doi:10.2478/s11600-007-0038-y
3. Beare R, MacVean M, Holtslag AAM, Cuxart J, Esau I, Golaz J-C, Jimenez M, Khairoutdinov M, Kosovic B, Lewellen D, Lund T, Lundquist J, McCabe A, Moene A, Noh Y, Raasch S, Sullivan P (2006) An intercomparison of Large–Eddy Simulations of the stable boundary layer. Bound-Layer Meteorol 118:247–272
4. Beljaars ACM, Holtslag AAM (1991) Flux parameterization over land surfaces for atmospheric models. J Appl Meteor 30:327–341
5. Beljaars ACM, Viterbo P (1998) Role of the boundary layer in a numerical weather prediction model. In: Holtslag AAM, Duynkerke PG (eds) Clear and Cloudy boundary layers. Royal Netherlands Academy of Arts and Sciences, Amsterdam, 372 pp
6. Chen F, Dudhia J (2001) Coupling an Advanced Land Surface-Hydrology Model with the Penn State-NCAR MM5 Modeling System. Part II: Preliminary Model Validation. Mon Wea Rev 129:587–604
7. Cuxart J, Holtslag AAM, Beare RJ, Bazile E, Beljaars A, Cheng A, Conangla L, Ek MB, Freedman F, Hamdi R, Kerstein A, Kitagawa H, Lenderink G, Lewellen D, Mailhot J, Mauritsen T, Perov V, Schayes G, Steeneveld GJ, Svensson G, Taylor P, Weng W, Wunsch S, Xu K-M (2006) Single-column model intercomparison for a stably stratified atmospheric boundary layer. Bound-Layer Meteorol 118: 273–303
8. Derbyshire SH (1999) Boundary layer decoupling over cold surfaces as a physical boundary instability. Bound-Layer Meteorol 90:297–325
9. Duynkerke PG (1991) Radiation fog: A comparison of model simulation with detailed observations. Mon Wea Rev 119:324–341
10. Ek MB, Holtslag AAM (2004) Influence of Soil Moisture on Boundary Layer Cloud Development. J Hydrometeor 5:86–99
11. Garratt JR (1992) The Atmospheric Boundary Layer. Cambridge University Press, New York, 316 pp
12. Garratt JR, Brost RA (1981) Radiative Cooling effects within and above the nocturnal boundary layer. J Atmos Sci 38:2730–2746
13. Holtslag AAM (2002) Atmospheric Boundary Layers: Modeling and Parameterization. In: Holton JR, Pyle J, Curry JA (eds) Encyclopedia of Atmospheric Sciences, vol 1. Academic Press, pp 253–261
14. Holtslag AAM (2006) GEWEX Atmospheric Boundary Layer Study (GABLS) on Stable Boundary Layers. Bound-Layer Meteorol 118:243–246
15. Holtslag AAM, Moeng C-H (1991) Eddy diffusivity and countergradient transport in the convective boundary layer. J Atmos Sci 48:1690–1698
16. Holtslag AAM, de Bruin HAR (1988) Applied Modeling of the Nighttime Surface Energy Balance over Land. J Clim Appl Meteor 27:689–704
17. Holtslag AAM, Ek MB (2005) Atmospheric Boundary Layer Climates and Interactions with the Land Surface. In: Encyclopedia of Hydrological Sciences. Wiley
18. Holtslag AAM, Steeneveld GJ, van de Wiel BJH (2007) Role of land-surface feedback on model performance for the stable boundary layer. Bound-Layer Meteorol 125:361–376
19. Lenderink G, Van den Hurk B, van Meijgaard E, van Ulden A, Cuijpers H (2003) Simulation of present day climate in RACMO2: First results and model developments. KNMI Technical report TR-252, 24 p
20. Mahrt L (1998) Stratified atmospheric boundary layers and breakdown of models. Theor Comp Fluid Phys 11:263–279
21. Mahrt L (1999) Stratified atmospheric boundary layers, Bound-Layer Meteorol 90:375–396
22. Poulos GS et al (2002) CASES-99: A comprehensive investigation of the stable nocturnal boundary layer. Bull Am Meteor Soc 83:555–581
23. ReVelle DO (1993) Chaos and "bursting" in the planetary boundary layer. J Appl Meteor 32:1169–1180
24. Salmond JA, McKendry IG (2005) A review of turbulence in the very stable boundary layer and its implications for air quality. Prog Phys Geogr 29:171–188
25. Sharan M, Gopalakrishnan SG (1997) Comparative Evaluation of Eddy Exchange coefficients for strong and weak wind stable boundary layer modelling. J Appl Meteor 36:545–559
26. Steeneveld GJ (2007) Understanding and prediction of stable boundary layers over land. Ph D thesis, Wageningen University, 199 pp
27. Steeneveld GJ, van de Wiel BJH, Holtslag AAM (2006) Modeling the arctic stable boundary layer and its coupling to the surface. Bound-Layer Meteorol 118:357–378
28. Steeneveld GJ, van de Wiel BJH, Holtslag AAM (2006) Modeling the Evolution of the Atmospheric Boundary Layer Coupled to the Land Surface for Three Contrasting Nights in CASES-99. J Atmos Sci 63:920–935
29. Steeneveld GJ, Mauritsen T, de Bruijn EIF, Vila-Guerau de Arellano J, Svensson G, Holtslag AAM (2008) Evaluation of limited area models for the representation of the diurnal cycle and contrasting nights in CASES99. J Appl Meteor Clim 47:869–887
30. Stull RB (1988) An introduction to Boundary-Layer Meteorology. Kluwer, Dordrecht, 666 pp. (reprinted 1999)
31. Svensson G, Holtslag AAM (2006) Single column modeling of the diurnal cycle based on CASES99 data -GABLS second intercomparison project. 17th Symposium on Boundary layers and Turbulence, 22–25 May, San Diego. American Meteorol Soc, Boston, Paper 8.1 (available at http://ams.confex.com/ams/pdfpapers)

32. Tjemkes SA, Duynkerke PG (1989) The nocturnal boundary layer: model calculations compared with observations. J Appl Meteor 28:161–175
33. van de Wiel BJH (2002) Intermittency and Oscillations in the Stable Boundary Layer over Land. Ph D thesis, Wageningen University, 129 pp
34. van de Wiel BJH, Moene AF, Hartogensis OK, de Bruin HAR, Holtslag AAM (2003) Intermittent turbulence and oscillations in the stable boundary layer over land, Part III: a classification for observations during CASES99. J Atmos Sci 60:2509–2522
35. van de Wiel BJH, Moene AF, Steeneveld GJ, Hartogensis OK, Holtslag AAM (2007) Predicting the Collapse of Turbulence in Stably Stratified Boundary Layers. Turbul Flow Combust 79:251–274
36. Welch RM, Ravichandran MG, Cox SK (1986) Prediction of Quasi-Periodic Oscillations in Radiation Fogs. Part I: Comparison of Simple Similarity Approaches. J Atmos Sci 43:633–651

Slug Flow: Modeling in a Conduit and Associated Elastic Radiation

Luca D'Auria, Marcello Martini
Osservatorio Vesuviano, Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Napoli, Naples, Italy

Article Outline

Glossary
Definition of the Subject
Introduction
Slug Flow and Strombolian Activity
Numerical Modeling
Bubble Ascent
Slug Ascent in a Vertical Pipe
Slug Ascent in a Pipe with a Flare
Seismological Constraints on Numerical Models
Conclusions
Future Directions
Appendix A – Fluid Dynamics of a Two-Phase System
Appendix B – Numerical Implementation
Bibliography

Glossary

Strombolian activity Is a kind of volcanic activity consisting of discrete, intermittent ejections of gas and magma fragments. The height reached by the ejecta seldom exceeds a few hundred meters above the volcanic crater. They consist mostly of molten magma fragments that partially quench as they fly. Strombolian explosions are intermediate between small Hawaiian eruptions and stronger Plinian eruptions. This kind of activity takes its name from the Stromboli volcano (Southern Italy) famous for its perpetual activity consisting of repeated moderate explosions (usually about 10 every hours).

Gas slug A bubble whose diameter is close to the diameter of the conduit where it is flowing. The shape and the motion of the slug is largely controlled by the conduit walls.

Modeling A simplified representation of a system aimed at reproducing some of its features. Mathematical models describe a system through a set of variables related by analytical relations. When these relations are too complex to be solved exactly, they can be solved using approximate numerical methods. Actually these numerical techniques often involve the massive use of computers. An alternative kind of modeling (analogue modeling) involves the use of versions of the system under study rescaled to fit spatial and temporal laboratory scales. Different materials are used to simulate original ones. For instance silicon oil is often used to simulate magma. The scaling relationships are rigorously stated in order to get physically meaningful results.

Computational fluid dynamics A set of mathematical, numerical and computational tools aimed at simulating complex fluid flows on computers. It developed simultaneously with computer science starting from the 1950s mostly with the aim of solving engineering problems. Today CFD spans a wider range of fields, from aeronautics to chemical engineering, to astrophysics to geophysics and much more. The most recent developments of CFD are related with the increase in computer performances and with the wider use of parallel computers.

Diffuse interface theory A molecular theory for describing the variation of the chemical composition across the interface on the basis of a rigorous thermodynamical approach. Starting from the definition of a free-energy function dependent on the chemical composition and its gradient, it is possible to compute all the chemical-physical properties of the interface. Beyond its physical meaning, this theory can be used also as a numerical tool for modeling of multiphase flows.

Very-long-period events (VLP) Seismic events recorded on active volcanoes and geothermal systems having a typical period of 10^2–10^0 s. Their observation and study began during the 1990s, with the spreading of seismic broadband sensors and have shown to be one of the most powerful tools for investigating the geometries of volcanic conduits and the dynamics of volcanic eruptions. Until now they have been observed

in tens of volcanoes with different eruptive styles such as: Aso (Japan), Erebus (Antarctica), Kilauea (Hawaii), Popocatepetl (Mexico), Sakurajima (Japan), Stromboli (Italy).

Moment-tensor Tensor representation of the force systems acting on seismic sources. It can be applied to common earthquake sources (rupturing faults) as well as to volcanic sources (fluid filled conduits). In the former case its trace is null, which from a physical point of view, means that there is no netvolume change during common earthquakes. On the other hand in volcanic sources volumetric variations are very common and furthermore they are accompanied also by a single force component, related to the net acceleration of center of mass of the fluid filling the conduit.

Definition of the Subject

Among the eruptive styles, the Strombolian activity is one of the more easy to study because of its repetitive behavior. For this reason large amount of data can be comfortably collected. Strombolian volcanoes are like natural laboratories repeating the same experiment (individual explosions) many times each day.

The development of quantitative models of eruptive dynamics is driven by the comparison of experimental observations and synthetic data obtained through mathematical, numerical or analogue modeling.

Since Strombolian activity offers a profuse amount of interesting seismic signals, during the last decades there has been growing attention on seismological techniques aimed at retrieving the conduit geometry and the eruption dynamics from the seismological recordings. One of these techniques, the source function inversion, is able to retrieve a summary of the forces acting on the volcanic conduit during the VLP event generation [5]. The comparison of observed source functions with synthetic ones, obtained through numerical modeling, allow us to put constraints on the proposed models.

Quantitative models, able to fit seismological observations, are a powerful tool for interpreting seismic recordings and therefor the seismological monitoring of active volcanoes.

Introduction

In this paper we discuss the mechanism of generation of Very-Long-Period events related to Strombolian explosions. This eruptive style, occurring in many basaltic volcanoes worldwide, is characterized by the ascent and the bursting of large gas slugs. The mechanism of formation, ascent and explosion of bubbles and slugs and their relation to eruptive activity has been studied from a theoretical point of view and by means of analogue simulations. Here we introduce results from numerical simulations, focusing on the pressure variations induced on the conduit walls and responsible for the generation of seismic signals.

We will first illustrate the main features of the fluid dynamics related to Strombolian eruptive activity (Sect. "Slug Flow and Strombolian Activity") and an overview of the numerical modeling (Sect. "Numerical Modeling") Then we will show results obtained using simple conduit model (Sect. "Bubble Ascent", Sect. "Slug Ascent in a Vertical Pipe" and Sect. "Slug Ascent in a Pipe with a Flare") and we will compare the synthetic source functions with actual observations (Sect. "Seismological Constraints on Numerical Models").

Slug Flow and Strombolian Activity

A fundamental distinction can be made between eruptive regimes on the basis of the magma viscosity. In silicic systems, the magma viscosity ($>10^5$ Pa s) is too high to allow an independent motion of gas bubbles [18]. They can only grow by diffusion processes and expand under the effects of pressure variations until fragmentation occurs. In basaltic magmas the viscosity ($<10^3$ Pa s) allows independent motion of the gas bubbles leading to a different behavior with the possibility of bubble coalescence, splitting and turbulent fluid flow [11]. Theoretical laboratory and numerical studies have been published in order to understand the dynamics of Strombolian eruption in terms of gas/magma interaction [12,24].

The distinction between free bubble ascent and slug flow can be made using the ratio between the average bubble diameter d and the conduit diameter D [7]: the parameter λ (Table 1). Values of λ higher than 0.6 are related to slug flow. The range of behavior exhibited by slug flow depends on the physical properties of the fluids (density, viscosity, surface tension) and on the geometry of the conduit (diameter, shape, inclination). It is possible to define adimensional parameters for describing the particular flow regime (see Table 1). The Reynolds number is the ratio between inertial and viscous forces. Low values of Re are typical of laminar flows, while higher values are related to turbulent regimes. The Froude number Fr is the ratio between inertial and gravitational forces. The Eotvos number Eo is the ratio between buoyancy and surface tension effects. The values of Eo determines the bubble shape. High values of Eo are related to distorted bubble shapes, while lower values to sub-spherical shapes. The Morton number Mo has a similar meaning. The dimensionless inverse viscosity N_f assess the relative importance of viscous effects.

Slug Flow: Modeling in a Conduit and Associated Elastic Radiation, Table 1
Symbols used in the text

Symbol	Meaning
d	Average bubble diameter
D	Conduit diameter
L	Distance between the top of the bubble and the magma/air interface in the initial conditions.
U	Terminal slug velocity
λ	$\lambda = \frac{d}{D}$
g	Gravity
ρ	Liquid density
$\Delta\rho$	Difference between gas and liquid densities
μ	Dynamic viscosity
ν	Kinematic viscosity $\nu = \frac{\mu}{\rho}$
σ	Surface tension
γ	Isothermal expansion ratio
Re	Reynolds number $Re = \frac{Ud}{\nu}$
Fr	Froude number $Fr = \frac{U}{\sqrt{gD}}\sqrt{\frac{\rho}{\Delta\rho}}$
Mo	Morton number $Mo = \frac{g\mu^4\Delta\rho}{\rho^2\sigma^3}$
Eo	Eotvos number $Eo = \frac{\Delta\rho g D^2}{\sigma}$
N_f	Dimensionless inverse viscosity $N_f = \left(\frac{Eo^3}{Mo}\right)^{\frac{1}{4}}$

Viscous flows are characterized by $N_f < 2$ while inertial flows by $N_f > 200$ [9].

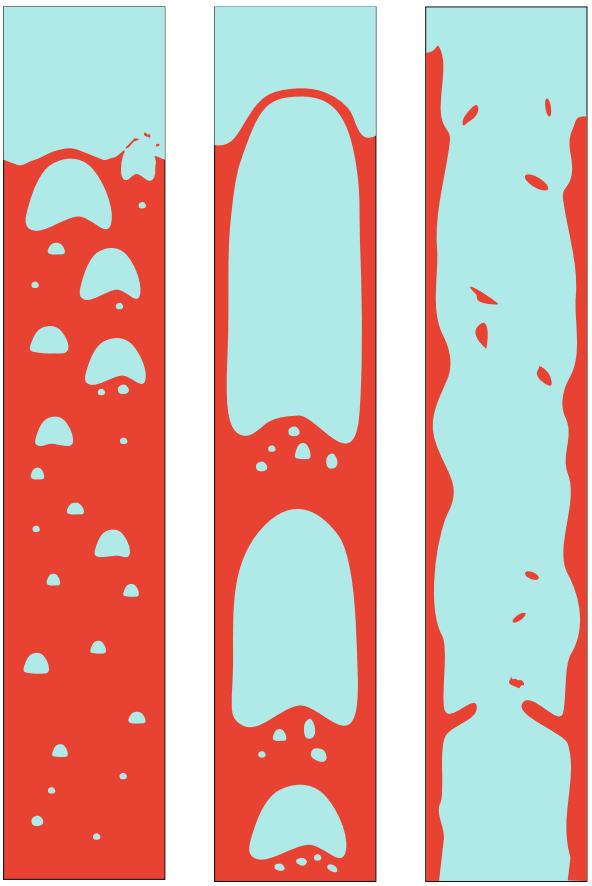
The ascent of gas in a basaltic system has been studied both from a theoretical and an experimental point of view. Observations of basaltic systems have shown that usually the flow conditions are transitional between viscous dominated and inertia dominated systems [17]. This puts constraints on the parameter range to explore both in numerical and analogue modeling. The range of adimensional numbers for basaltic systems, summarized from [17] and [9] is reported in Table 2.

For a basaltic system with given physical properties another variable plays a fundamental role in determining the physics of the flow: the gas/magma volumetric ratio. Flows are characterized by a low ratio consist of isolated bubbles rising with minor interaction between them and the conduit walls (Fig. 1a). In this regime bubbles rise assuming a shape that depends on their size and on the magma viscosity. Surface tension effect depends on the size of the bubbles and the ascent velocity depends on both the bubble size and the magma viscosity. In the range allowed for basaltic systems, larger bubbles have shapes ranging from dimpled ellipsoidal cap to spherical cap, while smaller bubbles have shapes ranging from spherical to ellipsoidal [7]. For higher gas/magma ratios bubbles begin to interact and to coalesce forming gas slugs occupying most of the conduit diameter (Fig. 1b). The ascent of gas slugs is strongly controlled by the conduit walls. During their passage, the magma is completely mixed and the motion of smaller bubbles is dominated by the fluid vorticity released (in a transitional regime) in their wake. For even higher gas/magma ratios (Fig. 1c) the slugs coalesce forming an almost continuous (turbulent) gas flow at the center of the conduit, while walls remains wet with a magma film. The high velocity of the flow produces instabilities on the surface of the magma film pulling out small fragments carried away from the gas.

Bubbly flows are related to effusive and to moderate Hawaiian explosive activity. Each single bubble explodes at the magma/air interface ejecting small magma fragments resulting from the shattering of the liquid film around the bubble. The bubble is usually overpressurized because of surface tension, viscous and inertial effects [23]. Slug flows are related instead to intermittent Strombolian activity.

Slug Flow: Modeling in a Conduit and Associated Elastic Radiation, Table 2
Range of adimensional numbers for flows related to basaltic systems (from [17] and [9])

Parameter	Range
Fr	$0.1 \div 0.345$
Re	$5 \div 31^3$
Eo	$51^5 \div 71^7$
Mo	$51^5 \div 101^{10}$
N_f	$16 \div 5000$



Slug Flow: Modeling in a Conduit and Associated Elastic Radiation, Figure 1
Schematic representation of gas/magma flows in basaltic systems. a represents bubbly flow, b slug flow and c annular flow

Each slug exploding at the magma/air interface generates a jet of gas of short duration (from a few seconds to some tens of seconds) [24]. The main Strombolian explosion can be followed by minor bursts caused by the smaller bubbles trapped behind the slug wake. Annular flows are related to continuous lava fountains with the central gas jet carrying molten magma fragments [11].

The nature of the flow can change along the conduit because of the gas expansion due to the magmastatic pressure decrease. This expansion makes the gas/magma volumetric ratio increase along the conduit. At the base of the conduit, where the gas starts to exolve from the magma there is always a bubbly flow. As the pressure decreases the flow can change into slug flow. This can be made easier by an inclined conduit [4] and by constrictions [12] that force the bubbles to coalesce even at lower gas/magma ratios than those required in a straight vertical conduit. The total gas flow can change in the conduit because of variations in the deep feeding system or because of non-linear instabilities due to the complex shape of the conduit system [11,12].

The expansion of bubbles and slugs is a fundamental factor to take into account when studying Strombolian activity [17]. Assuming a simple ideal isothermal state equation for the gas we get the relative volumetric change for a bubble rising in a magma with density ρ:

$$\frac{V_0}{V} = \frac{P_{\text{atm}} + \rho g h}{P_{\text{atm}}} . \tag{1}$$

It is easy to show that the expansion ratio in the upper few hundreds of meters is more than one order of magnitude.

The flow regime can change also horizontally, for instance in a subvertical dike the flow rate can be higher on one side, leading to a slug flow, and lower on the other side, leading to a bubbly flow. This explains the coexistence of effusive and Strombolian activity observed during some basaltic eruptions.

Numerical Modeling

The aim of this work is to investigate the pressure variations induced by a gas bubble rising in a magma-filled volcanic conduit. This phenomenon has been also investigated by means of analogue laboratory models [9,10,12,17].

The major drawback of analogue modeling, in this context, is that it provides only a limited number of pressure time series: one for each sensor. Numerical modeling provides a different point of view giving the full set of scalar (density), vector (velocity) and tensor (pressure) quantities over the whole computational domain. This allows quantitative inferences on the flow regimes and on the elastodynamic wavefield generated (see Sect. "Seismological Constraints on Numerical Models").

The modeling of two-phase systems as gas magma is not a simple task in computational fluid dynamics (CFD) [6]. Taking into account surface tension effects at the liquid-gas interface can be done in two different ways. The first consists in explicitly tracking the time evolution of the gas-liquid interface [22]. In situations involving complex flows with extensive occurrences of bubble coalescence and splitting, these methods show a rapid increase in computational effort. Another category of methods model the gas-liquid systems using the diffuse-interface theory. These methods consider two scalar fields, defining the relative local concentrations of the two components and a modified advection-diffusion equation for

Slug Flow: Modeling in a Conduit and Associated Elastic Radiation, Table 3
Description and adimensional numbers for the six simulations presented in the text

Sim.	Description	Comp. domain (pts.)	λ	γ	Re	log(Mo)	Eo	Fr	N_f
1A	Bubble ascent (low viscosity)	60 × 600	0.5	5	1.93×10^3	0	9×10^4	0.37	5.19×10^3
1B	Bubble ascent (high viscosity)	60 × 600	0.5	5	1.90×10^2	4	9×10^4	0.36	5.19×10^2
2A	Slug ascent in vertical pipe (low viscosity)	30 × 600	0.8	5	1.82×10^3	0	9×10^4	0.35	5.19×10^3
2B	Slug ascent in vertical pipe (high viscosity)	30 × 600	0.8	5	1.54×10^2	4	9×10^4	0.30	5.19×10^2
3A	Slug ascent in vertical pipe with a flare (low viscosity)	60 × 600	0.8	5.3	1.83×10^3	0	9×10^4	0.35	5.19×10^3
3B	Slug ascent in vertical pipe with a flare (high viscosity)	60 × 600	0.8	5.3	1.68×10^2	4	9×10^4	0.32	5.19×10^2

modeling the physical-chemical interaction between them. This is done using a thermodynamically consistent definition of a free-energy function that takes into account the phase equilibria. This approach leads to smooth interfaces where the concentration of one phase gradually decreases while the other increases. The thickness of these interfaces depends on the numerical method and on the surface tension value [25]. From a numerical point of view this problem can be faced using Lattice Boltzmann Methods (LBM) [20] or by classic CFD [6]. Here we use the latter. Some details about the numerical method are reported in the Appendices.

In our models we have not considered the mechanical interaction between the fluid phases and the elastic conduit walls. The exchange of linear momentum between them is a significant factor in the generation of seismic waves [3] in seismo-volcanic sources. However in conduits having ratios between the linear dimension and thickness lower than 10^1–10^2 the effect of the motion of the conduit walls does not affect significantly the fluid flow. Furthermore we show in Sect. "Seismological Constraints on Numerical Models" that we will not compare the result of the simulation directly with seismograms, but with an equivalent system of forces acting on the conduit walls.

In the following we present results of some elementary two-dimensional conduit models focusing on the slug flow regime which occurs during Strombolian activity. The numerical simulations generate snapshots of the physical quantities (composition, density, pressure and velocity) along all the point of the discretized computational domain. The effective computational domains for each simulation are showed in Table 3. In all the simulations we start from a static fluid with a bubble in the lower part of the conduit and a gas/magma interface in the upper part. The boundary conditions keeps a constant pressure value at the bottom and the top of the model. The no-slip boundary conditions are implemented along the conduit walls.

The adimensional numbers for each simulation are reported in Table 3. Velocities for determining Re and Fr are computed when the bubble/slug is moving in a steady state after a transient due to the initial conditions. These velocity values U are also used for normalizing times:

$$t^* = t\,\frac{U}{L}\,, \tag{2}$$

where L is the distance between the top of the bubble and the magma/air interface. So (2) represents a normalization for the virtual time the bubble needs to reach the surface in an ideal steady motion. We can also normalize distances:

$$x^* = x\,\frac{1}{D}\,, \tag{3}$$

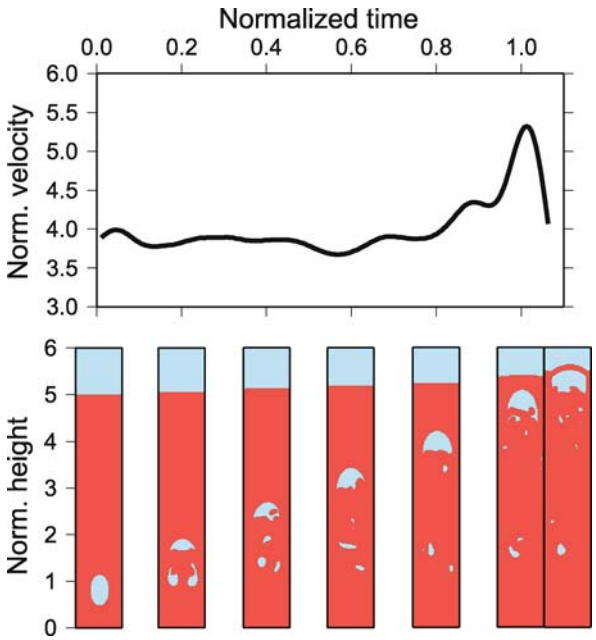
where D is the conduit diameter. We can define then normalized velocities as:

$$v^* = \frac{\partial x^*}{\partial t^*}\,. \tag{4}$$

In our simulations we have also defined an isothermal expansion ratio γ which is the ratio between the bubble volume at the initial conditions and its volume at (simulated) atmospheric pressure.

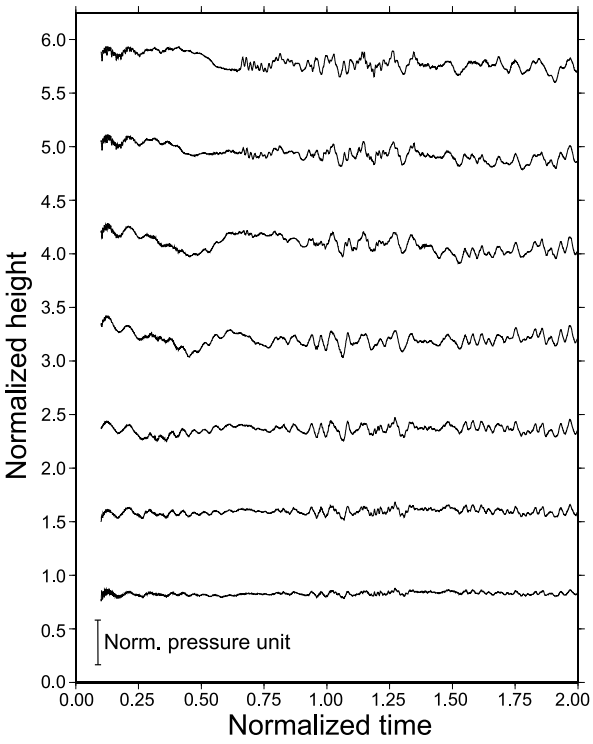
Bubble Ascent

We first consider the ascent of a single bubble in a vertical conduit. In the first simulation (1A in Table 3) the initial shape of the bubble is an ellipse (see Fig. 2). As the simulation evolves the bubble becomes unstable splitting into three smaller bubbles. The two smaller lateral bubbles are embedded in a symmetric vortex structure. As the flow



Slug Flow: Modeling in a Conduit and Associated Elastic Radiation, Figure 2

Ascent and bursting of a gas bubble (simulation 1A in Table 3). Time, height and velocity are normalized according to (2), (3) and (4)



Slug Flow: Modeling in a Conduit and Associated Elastic Radiation, Figure 3

Time series of pressure variations for simulation 1A. The represented pressure values are normalized by the external atmospheric pressure P_0. Each time series represents the recording of a virtual sensor located at the center of the conduit at the height of its first point (see *scale on the left*)

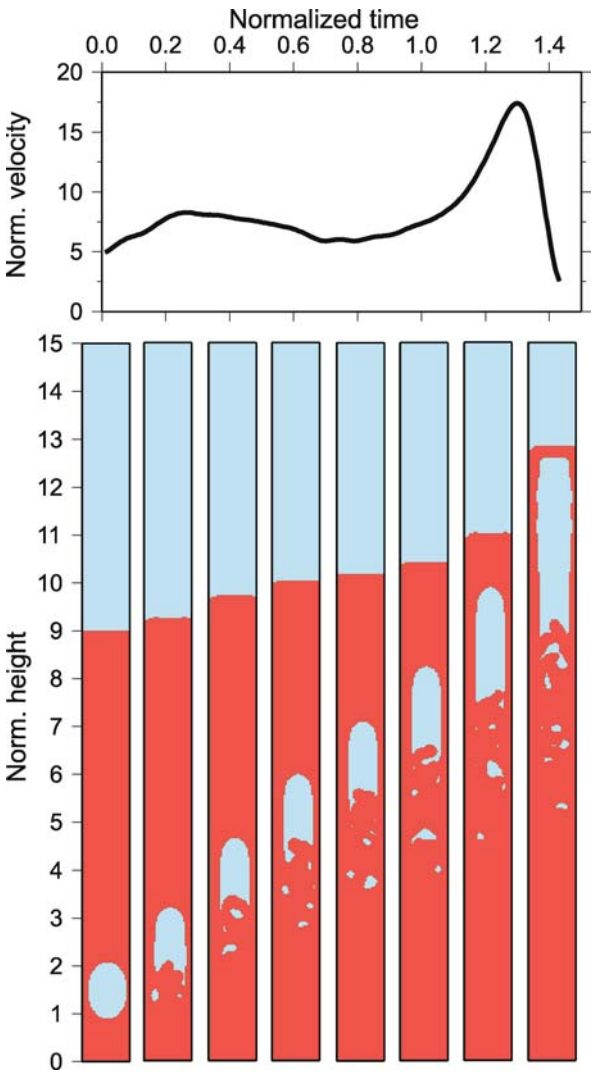
evolves the symmetry is broken and a turbulent wake develops behind the largest bubble. This one rises with an almost steady velocity (Fig. 2 top) with a spherical cap shape. In the final part of the simulation, the bubble suffers an expansion that made the bubble accelerate ($t > 0.8$) and causes the flow to become transitional toward the slug flow regime. Then the bubble reaches the surface and a curved liquid film develops above the bubble just before bursting. This phase is accompanied by a sudden deceleration of the bubble ascent ($t > 1.0$). The pressure variations in the conduit are modest, with a lack of low frequency oscillations. Also the pressure transients generated by the bubble burst are limited (Fig. 3).

The adimensional number for this simulation (Table 3) is in a range related to transitional flow regimes typical of basaltic systems [9,17]. An increase of an order of magnitude of the viscosity (1B in Table 3) makes the flow still in the transitional regime (Table 3). The high values of the Eotvos number indicates that the surface tension effect is not relevant in the flow dynamics [7] in both cases.

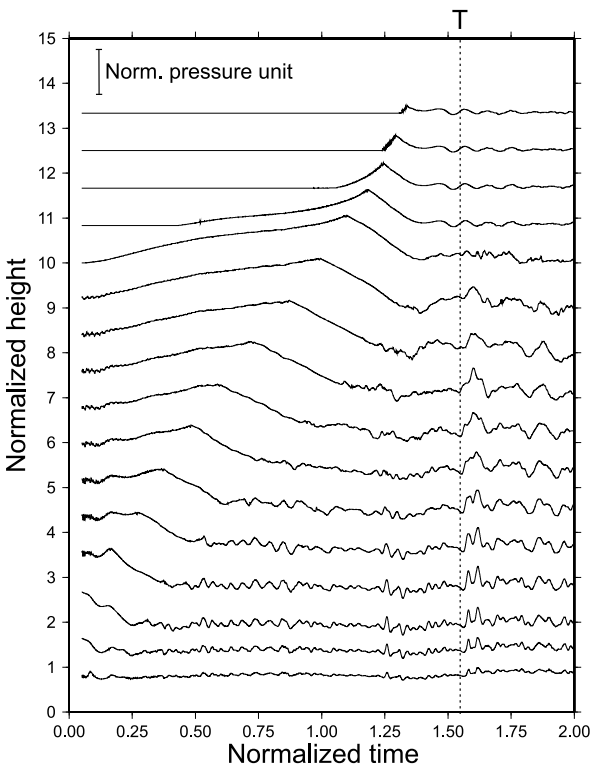
Slug Ascent in a Vertical Pipe

In these simulations (2A and 2B in Table 3) we model the ascent of a single slug in a vertical pipe. Again we start from an elliptical shape of the bubble. After an initial transient the typical slug shape develops (Fig. 4) and a turbulent wake appears behind the slug. As the slug rises it suffers a volume expansion. In the slug flow regime the liquid is pushed upward causing an overall increase of the hydrostatic pressure in the conduit. After the bursting of the slug the liquid film on the wall falls down to the original hydrostatic level. The slug expansion causes a significant acceleration of the liquid for $t > 1.0$ (Fig. 4). The velocity drops as the slug reaches the bursting point and the liquid film develops.

The major feature in the pressure pattern is the slow ramp-like increase followed by a drop (Fig. 5). The increase reflects the rise of the hydrostatic head above the slug while the gradual decrease is related to the passage of the slug. The bursting of the slug generates a moderate pressure transient in the conduit (marker "T" in Fig. 5) and damped resonant oscillations in the uppermost part of the pipe, filled with gas (above norm. height 11 in Fig. 5).



Slug Flow: Modeling in a Conduit and Associated Elastic Radiation, Figure 4
Ascent and bursting of a single gas slug in a straight pipe (simulation 2A in Table 3)



Slug Flow: Modeling in a Conduit and Associated Elastic Radiation, Figure 5
Time series of pressure variations for simulation 2A. *T* marks the most significant transient in the pressure patterns

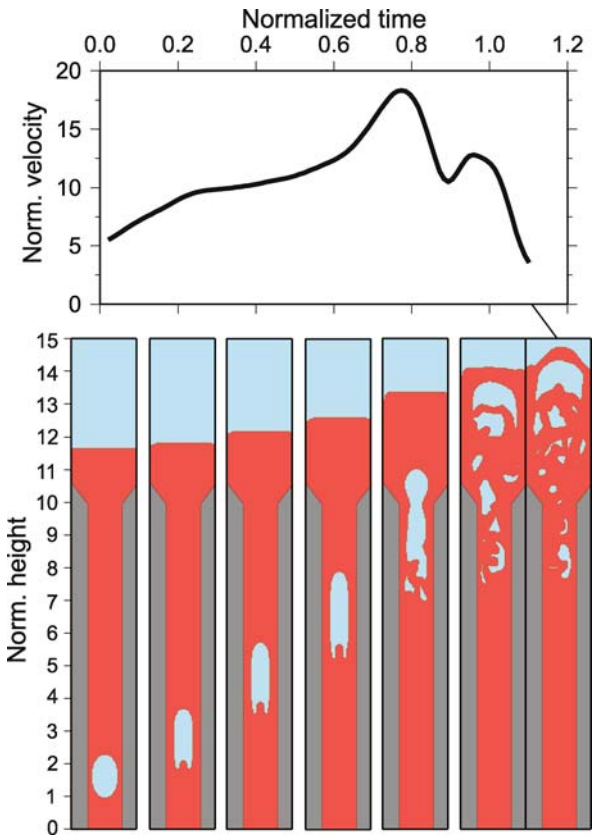
The simulated pressure pattern fits very well the observations of analogue modeling [9]. Both the pressure increase linked to the rise of the hydrostatic head and the pressure drop related to the passage of the slug are in good agreement. Together with these long period variations they observe also oscillatory pressure transients just before the slug approaches the surface and after the bursting. The simulations presented here lack this feature because of an intrinsic limit in the numerical method that does not model explicitly the interface (see Appendix).

The overall behavior of simulation 2B, where a tenfold increase of the liquid viscosity (Table 3) is very similar to 2A. The only remarkable difference is the lack of high frequency oscillations in the pressure (as the transient "T" in Fig. 5) due to the greater viscous damping effect.

Slug Ascent in a Pipe with a Flare

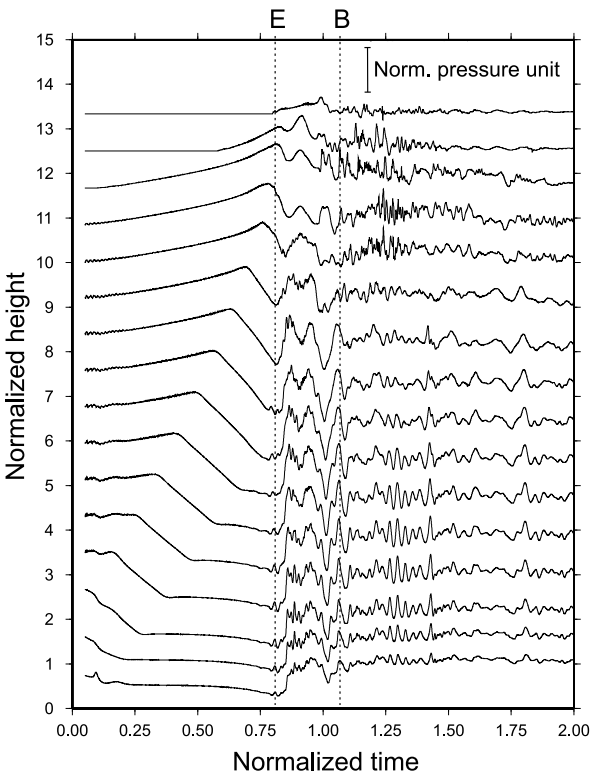
In this set of simulations (3A and 3B in Table 3) we model again the ascent of a single slug in a pipe. However in this case the width of the pipe doubles after a norm. height of 10. In the first part of the simulation, the behavior is similar to the one illustrated in Sect. "Slug Ascent in a Vertical Pipe". As the slug nose enters in the flare (Fig. 6) is starts to expand rapidly making the fluid accelerate upward. The expansion of the slug in the flare is followed by its breakup because of the development of strong turbulence. The velocity of the top of the slug increases until the slug passes through the flare then it drops because of the change in the conduit diameter.

The pressure pattern shows a major difference, compared with simulation 2A (Fig. 7). As the slug passes through the flare there is a sudden pressure increase



Slug Flow: Modeling in a Conduit and Associated Elastic Radiation, Figure 6

Ascent and bursting of a single gas slug in a pipe with a flare (simulation 3A in Table 3)



Slug Flow: Modeling in a Conduit and Associated Elastic Radiation, Figure 7

Time series of pressure variations for simulation 3A. *E* marks the entering of the slug in the upper conduit while *B* marks the bursting of the main bubble at the magma surface

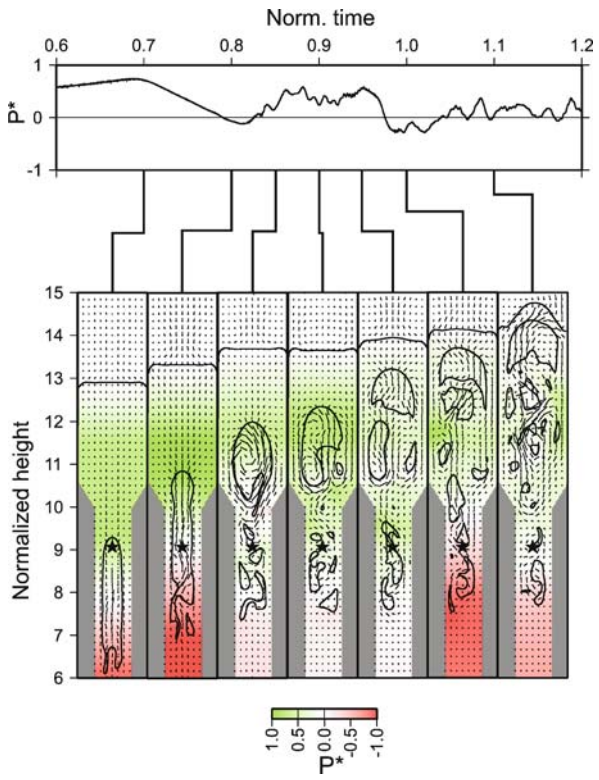
recorded along the whole conduit. The pressure then drops when the slug expansion has terminated. This strong pressure pulse is clearly related to the varying fluid flow regime as the slug enters the upper conduit and it deserves a closer analysis. In Fig. 8 we have represented the fluid dynamic regime during the slug expansion together with the related pressure variations. We observe that the pressure rises when most of the slug has passed through the flare. The strong acceleration induces turbulence both in the lower and the upper sections of the conduit. The upward acceleration first causes the disruption of the lower part of the slug leaving behind the main bubble a set of smaller ones, whose motion is driven by the fluid turbulence. When the slug has entered in the upper conduit it suffers another splitting due both to the induced vorticity and to the change in the boundary conditions. The behavior of the slug for $t^* > 0.8$ is characterized by a flow regime with a smaller value of λ and so it is similar to the initial part ($t^* < 0.2$) of simulation 1A.

Similar fluid dynamics behavior and pressure patterns have been observed in analogue simulations by [10]. In particular, the breakup of the slug and the generation of a positive pressure pulse as it passes through the flare are observed in analogue models within a wide range of fluid viscosities and conduit widenings.

As in the previous case the simulation with higher liquid viscosity (3B in Table 3) shows a similar behavior. Higher viscosities reduce the fluid vorticity leading to less fragmented slugs.

Seismological Constraints on Numerical Models

Seismological data analysis is a powerful tool for putting constraints on the geometry and the dynamics of volcanic systems. Non-stationary fluid flow in volcanic conduits generates pressure variations on the conduit walls and then seismic waves propagating toward the Earth's surface where they are recorded by seismometers. The frequency band of seismic signals recorded in volcanic ar-



Slug Flow: Modeling in a Conduit and Associated Elastic Radiation, Figure 8

Detail of the conduit during the slug expansion phase. Each snapshot represents the normalized pressure variations (see the *color scale* on the bottom) and the local fluid velocity vectors. The *thick contours* are the bubble interfaces. The pressure values represented above are measured at the virtual sensor indicated by a *black star* in the snapshots

eas spans a wide range: from the Ultra-Long-Period (ULP) band (>10^2)s, the Very-Long-Period (VLP) band (10^2–10^0 s) and the Long-Period (LP) band (10^0–10^1 s) [5].

From a formal point of view this can be expressed considering the elastodynamic field generated by a general extended source, whose external surface is Σ:

$$u_n(\mathbf{x}, t) = \iint_{\Sigma} [f_q * G_{np} + m_{pq} * G_{np,q}] \mathrm{d}\Sigma \,, \qquad (5)$$

where $u_n(\mathbf{x}, t)$ is the nth component of the ground displacement recorded at the position $\mathbf{x}$, f_q is the body-force distribution over Σ, m_{pq} is the moment density tensor and G_{np} are the Green's functions. In (5) we take into account all the details of the source dynamics. However the density of actual seismic networks is not sufficient for retrieving every minor feature of the seismic sources. Volcano monitoring networks usually have a limited extension (10^3–10^4 m) and the wavelengths associated with ULP and VLP signals are higher. Therefore the analysis of such signals the seismic source (i. e. the part of the volcanic conduit responsible of seismic wave generation) can be represented as a point [5]. Under the assumption of a point source we can express (5) as [5]:

$$u_n(\mathbf{x}, t) = F_q * G_{np} + M_{pq} * G_{np,q} \,, \qquad (6)$$

where:

$$F_q = \iint_{\Sigma} f_q \mathrm{d}\Sigma \qquad (7)$$

and:

$$M_{pq} = \iint_{\Sigma} m_{pq} \mathrm{d}\Sigma \,. \qquad (8)$$

In common earthquake sources the single force component F_q is null. In volcanic sources, the acceleration of the center of mass of the fluid makes this component noteworthy.

The inversion of the recorder waveforms $\mathbf{u}(\mathbf{x}, t)$, after the numerical computation of the Green's functions $\mathbf{G}$ allows the retrieval of the single force and moment tensor components of (6) [14]. $\mathbf{F}$ and $\mathbf{M}$ are a synthesis of the force systems acting on volcanic conduits and then can be used for discriminating among numerical models on the basis of their fit with observations.

Numerical simulations provide the pressure tensor field $\mathbf{P}$ over the whole computational domain. The pressure tensor can be used for computing the forces acting on each point of the conduit walls multiplying it for the unit vector normal to the wall $\hat{\mathbf{n}}$:

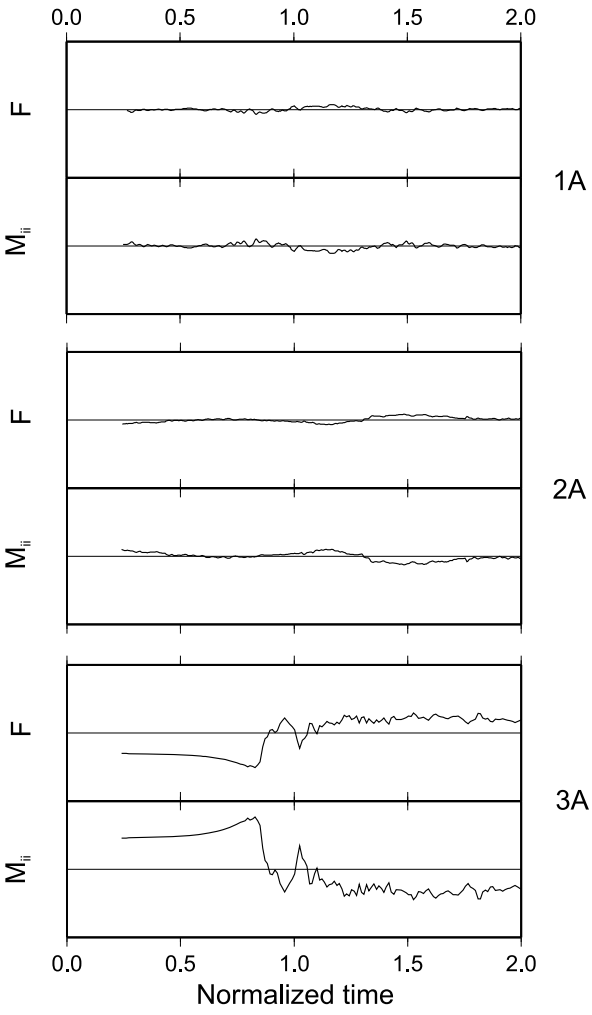
$$\mathbf{f} = \mathbf{P}\hat{\mathbf{n}} \,. \qquad (9)$$

These values can be integrated numerically using an expression similar to (7) to get the single force component $\mathbf{F}$. Then using the definition of moment they can be used for retrieving also the equivalent moment tensor:

$$M_{pq} = \iint_{\Sigma} \left(f_q - \frac{F_q}{\Sigma} \right) r_p \mathrm{d}\Sigma \,, \qquad (10)$$

where $\mathbf{r}$ is the arm formed by the force respect to an arbitrary origin $\mathbf{O}$.

In Fig. 9 we have represented the vertical force component and the moment tensor isotropic component (that is the trace $M_{xx} + M_{zz}$) for the three simulations with lower liquid viscosity. The represented values are normalized.



Slug Flow: Modeling in a Conduit and Associated Elastic Radiation, Figure 9

Vertical force and isotropic component for simulations 1A, 2A and 3A of the moment tensor for simulations 1A, 2A and 3A (see Table 3)

In the first case (1A) we note the absence of significant signals. There are only minor oscillations related to fluid dynamics instability during the bubble ascent. The rise of a bubble smaller than the conduit dimension seems not to be able to generate VLP signals. Therefore a repetitive bursting of small bubbles at the magma/air interface can be responsible for the generation of continuous infrasonic and seismic tremor recorded in some basaltic volcanoes [16].

In the second case, again, the signal amplitudes are quite low. It is evident that there is an anticorrelation between the vertical force and the isotropic moment. The vertical acceleration of the slug ($1.0 < t^* < 1.3$ in Fig. 4) causes a downward reaction force as well as an increase in the conduit pressure. The bursting ($t^* > 1.3$) causes a downward acceleration of the liquid and a pressure decrease. A similar effect, although with a minor magnitude, is also evident in 1A (Fig. 9).

In the third case the signal amplitude increases dramatically, about five times higher than 2A. The slug entering in the upper, wider portion of the conduit is marked by a downward force and a positive moment. At the end of the expansion phase ($t^* > 0.8$) there is a sudden inversion of the trends in both quantities, followed again by a positive downward peak of the force and positive of the momentum ($1.0 < t^* < 1.1$). This peak occurs when two pieces of the original slug coalesce again (Fig. 8) leading to a sudden, but limited in time, upward acceleration. In real cases the dynamics can be even more complex with multiple slugs or slug fragments entering in the upper conduit and interacting with each other. In this case during the main expansion phase, the lower conduit is filled almost exclusively with liquid (Fig. 6). If we suppose the simultaneous presence of many vertically aligned slugs in the lower conduit, then the pressure variations induced by the expansion of the topmost one would influence the lower ones in a complex non-linear mechanism still to investigate.

The signals presented in Fig. 9 need to be scaled to an actual time scale to be compared with real seismic signals source functions. In basaltic volcanoes the dimensions of the upper conduit is of the order of 10^2 m while the ascent velocities of the slugs are of the order of 10^0–10^1 m/s. This gives a scaling factor of about 10^1–10^2 *s*. Since the simulated transients have a normalized duration of about 10^{-1} the actual simulated signals should have a characteristic period of 10^0–10^1 *s*. These values are within the range of VLP and LP signals [5].

The patterns observed in simulation 3A (Fig. 9) closely matches the results obtained by [4] for the source function of VLP events at the Stromboli volcano. Thus this result, together with analogue simulations [10], strongly supports the hypothesis that VLP events at Stromboli are generated by the passage of a gas slug through a conduit widening and its subsequent expansion and bursting. The long ramp-like signals in both force and moment for $t^* < 0.8$ in Fig. 9 has a characteristic period longer than the VLP signals recorded and analyzed by [4]. Recordings at Stromboli, using a seismic sensor with a wider frequency range [13] interestingly suggested in some seismic signals a similar ramp having a length of more than 60 *s*, compatible with our scaling. This long ramp is related to the slow and gradual increase in the magmastatic head due to the continuous expansion of the ascending slug.

Conclusions

The rapid expansion of the gas slugs in the uppermost part of the conduit plays a fundamental role, both in the eruptive dynamics and the seismic wave generation process. We have focused on the role of conduit geometry in the fluid dynamics of gas slug ascent and its implication in the generation of seismic signals. In one of the simulations (3A) we have shown that the system of forces acting on the conduit is able to generate seismic waves with a higher efficiency compared with other cases. This observation can be generalized to more complex geometry such as an alternating of widening and narrowing [10]. The passage of slugs can occur also in very complex conduits [17]. In this case it is possible that pressure transients are generated in different positions along the conduit.

Future Directions

A deeper understanding of the relationship between slug ascent dynamics and seismic signals generation would require more advanced modeling techniques in various three-dimensional geometries, testing how the effect of changes in slug volumes and magma properties can affect the generation of seismic signals.

These studies are an important step toward more advanced seismic monitoring techniques of active basaltic volcanoes aimed at assigning in real time a volcanological meaning to variations in observed LP and VLP seismic signals.

Appendix A – Fluid Dynamics of a Two-Phase System

Definition of a Two-Phase System

A fluid two-phase system can be described using two scalar fields n_a and n_b, representing the local molar densities of the two components a and b. The actual local density ρ is then:

$$\rho = m_a n_a + m_b n_b \,, \tag{11}$$

where m_a and m_b are the molecular weight of the two phases. As it will be shown in the following, it is convenient to describe the system using an alternative representation based on the variables:

$$n = n_a + n_b \,, \tag{12}$$

and

$$\phi = n_a - n_b \,. \tag{13}$$

ϕ defines the local composition of the fluid. It can span the range $[-n, +n]$ with the value $-n$ representing a pure b composition and $+n$ a pure a composition. The relation between n, ϕ and ρ is:

$$\rho = \tfrac{1}{2}\left[m_a\,(n+\phi) + m_b\,(n-\phi)\right] \,. \tag{14}$$

The values of m_a and m_b are set so that in reference conditions the value of n is always equal to 1. So for instance, if we consider pure water at ambient conditions $m_w = 1000$ kg/mol.

State Equation

We assume that pure phases obey to a simple isothermal ideal gas state equation like:

$$P = \rho c^2 \,, \tag{15}$$

where P is the pressure and c is the sound speed. We assume that the pressure is zero in the reference state, so the state equation for the component a is:

$$P = m_a\,(n-1)\,c_a^2 \,. \tag{16}$$

A similar relation holds for b. Since along interfaces the composition varies continuously, we should define a state equation for a mixture that satisfies two requirements: the state equation for pure phases must match (16) and the isobaric contours in the (ϕ, n) plane must be straight lines. This second requirement follows from the consideration that diffusion processes between the components follows straight paths in this domain. In most of the actual Lattice Boltzmann literature, this question is not issued because the proposed implementations are almost isobaric ($n \simeq 1$) [21,25,26]. A suitable choice for the state equation is then:

$$P(n,\phi) = \frac{1}{4}\left[\Lambda + \sqrt{16 m_a m_b c_a^2 c_b^2\,(n-1) + \Lambda^2}\right] \tag{17}$$

with:

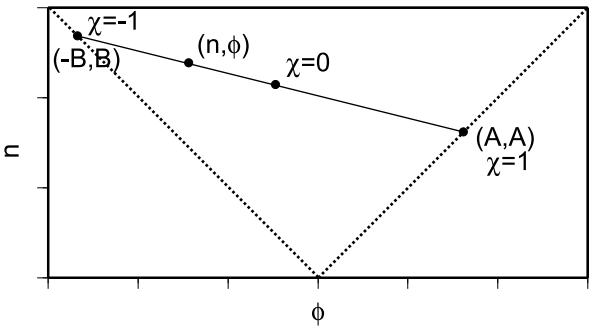
$$\Lambda = m_a c_a^2\,(n+\phi-2) + m_b c_b^2\,(n-\phi-2) \,. \tag{18}$$

In Fig. 11 an example contour plot of isobaric lines is shown. We can define another parameter χ as:

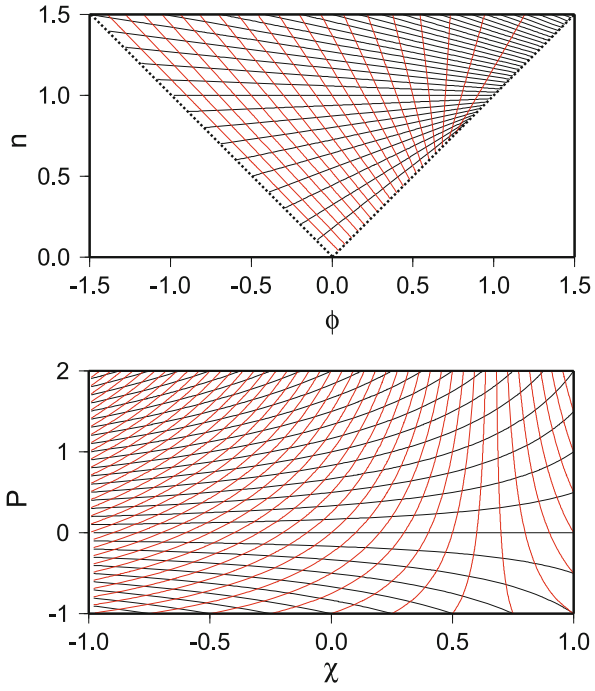
$$\chi(n,\phi) = -1 + 2\sqrt{\frac{(\phi+B)^2 + (n-B)^2}{(A+B)^2 + (A-B)^2}} \,, \tag{19}$$

where:

$$A = \frac{P(n,\phi)}{m_a c_a^2} + 1 \,, \tag{20}$$



Slug Flow: Modeling in a Conduit and Associated Elastic Radiation, Figure 10
Definition of χ. See text for details



Slug Flow: Modeling in a Conduit and Associated Elastic Radiation, Figure 11
In the upper panel, isolines of χ (*red lines*) and P (*black lines*) are represented in the (n, ϕ) domain. In the lower panel isolines of n (*black lines*) and ϕ (*red lines*) are represented in the (χ, P) domain. These plots represent a two-phase system where $m_a c_a^2 = 5$ and $m_b c_b^2 = 1$

and

$$B = \frac{P(n,\phi)}{m_b c_b^2} + 1 \,. \tag{21}$$

The actual meaning of χ is shown in Fig. 10. It is the rescaled ratio of the distance between the points (ϕ, n) and $(-B, B)$ and the distance between the points (A, A) and $(-B, B)$. So $\chi = -1$ along the line $\phi = -n$ and $\chi = 1$ along $\phi = n$. In practice χ is related to the local composition of the fluid. In Fig. 10 the relation between ϕ, n, P and χ is represented. This relation can be inverted and the system (χ, P) can be used as well as the (ϕ, n) coordinate system. In the following we show that the former is the best choice for describing chemical equilibria.

Chemical Equilibria and the Diffuse Interface Theory

The chemical equilibrium between the two phases can be described using a physically consistent thermodynamic approach [2,8,26]. We define a Ginzburg–Landau free energy functional for an heterogeneous mixture as [15]:

$$\mathcal{F} = \int \left[\psi(n,\phi) + \frac{\kappa_n}{2} (\nabla n)^2 + \frac{\kappa_\phi}{2} (\nabla \phi)^2 \right] \mathrm{d}V \,, \tag{22}$$

where κ_n and κ_ϕ are parameters related to the surface tension. The system reaches its final thermodynamical equilibrium making the free energy minimum.

We can rewrite (22) in the (χ, P) representation:

$$\mathcal{F} = \int \left[\psi(\chi,P) + \frac{\kappa_\chi}{2} (\nabla \chi)^2 \right] \mathrm{d}V \,. \tag{23}$$

Note that we have set $\kappa_P = 0$. This implies that surface tension effects are independent of the absolute pressure. In this paper we set:

$$\psi(\chi,P) = \alpha \left(\frac{\chi^4}{4} - \frac{\chi^2}{2} \right) + P \ln \frac{P}{P_0} \,, \tag{24}$$

where P_0 is an arbitrary reference pressure value. Using this definition of ψ we note that this function has two minima for $\chi = \pm 1$, corresponding to the two equilibrium compositions. Starting from an heterogeneous mixture, the system evolves by creating domains having a quite homogeneous composition separated by interfaces where the composition varies smoothly. The equilibrium is reached when the chemical potential:

$$\mu = \frac{\delta \mathcal{F}}{\delta \chi} \,, \tag{25}$$

is everywhere equal to zero [2,26]. So, in our case, using (23) and (24) we obtain the condition of chemical equilibrium for a plane interface orthogonal to the x-direction:

$$\mu = \frac{\partial \psi}{\partial \chi} - \kappa_\chi \frac{\partial^2 \chi}{\partial x^2} \,. \tag{26}$$

Assuming the boundary conditions $\psi(x) = 0$ and $\frac{\partial \chi}{\partial x} = 0$ for $x = \pm\infty$ and $\chi(0) = 0$, the previous ODE can be integrated giving the expression of the spatial variation of the

composition through an equilibrium interface:

$$\chi(x) = \tanh\left(\frac{2x}{\xi}\right), \tag{27}$$

where the interface thickness ξ is the width where the 96% of variation occurs. Its value is:

$$\xi = 2\sqrt{-\frac{2\kappa_\chi}{\alpha}}. \tag{28}$$

On the basis of the definition of surface tension σ [25]:

$$\sigma = \int_{-\infty}^{+\infty} \mathcal{F}(x)\,\mathrm{d}x, \tag{29}$$

we can write:

$$\sigma = \frac{2}{3}\sqrt{-2\alpha\kappa_\chi}. \tag{30}$$

Using (28) and (30) we can also retrieve useful inverse relations:

$$\alpha = 3\frac{\sigma}{\xi}, \tag{31}$$

and

$$\kappa_\chi = \frac{3}{8}\sigma\xi. \tag{32}$$

We should emphasize that in real physical systems expression (27) describes an actual interface having a thickness whose order of magnitude is of molecular scale. In this work we use the diffuse interface theory as a numerical tool for modeling of two-phase systems. In other words, we use unphysical interfaces having a macroscopic thickness (usually $10^{-3} \div 10^{-1} m$).

Cahn–Hilliard and Mass Conservation Equations

The time evolution under non-equilibrium conditions can be expressed by two Cahn–Hilliard equations [2,26]:

$$\frac{\mathrm{D}n}{\mathrm{D}t} = \Gamma\nabla^2\mu_n, \tag{33}$$

$$\frac{\mathrm{D}\phi}{\mathrm{D}t} = \Gamma\nabla^2\mu_\phi. \tag{34}$$

These equations are similar to common advection-diffusion equations with Γ being a diffusion coefficient and the operator $\mathrm{D}/\mathrm{D}t$ the substantial derivative. The explicit expressions for the chemical potentials μ_n and μ_ϕ are:

$$\mu_n = \frac{\delta\mathcal{F}}{\delta n} = \frac{\delta\mathcal{F}}{\delta\chi}\frac{\partial\chi}{\partial n} = \mu_\chi\frac{\partial\chi}{\partial n}, \tag{35}$$

$$\mu_\phi = \frac{\delta\mathcal{F}}{\delta\phi} = \frac{\delta\mathcal{F}}{\delta\chi}\frac{\partial\chi}{\partial\phi} = \mu_\chi\frac{\partial\chi}{\partial\phi}. \tag{36}$$

with:

$$\mu_\chi = \frac{\delta\mathcal{F}}{\delta\chi} = \alpha\chi(\chi^2-1) - \kappa_\chi\nabla^2\chi. \tag{37}$$

Then the Cahn–Hilliard Eqs. (33) and (34) can be rewritten as:

$$\frac{\mathrm{D}}{\mathrm{D}t}\begin{pmatrix} n \\ \phi \end{pmatrix} = \Gamma\nabla^2\mu_\chi\begin{pmatrix} \frac{\partial\chi}{\partial n} \\ \frac{\partial\chi}{\partial\phi} \end{pmatrix} \tag{38}$$

On the basis of the definitions of n (12) and ϕ (13), we can state that the previous equation expresses the mass transfer of components a and b because of chemical disequilibria. Since we are dealing with a compressible flow, we should obviously account for this in the mass balances. Then, following basic fluid dynamics [1], we can rewrite (38) in explicit form as:

$$\frac{\partial n}{\partial t} = -\mathbf{v}\nabla n - n\nabla\cdot\mathbf{v} + \frac{\partial\chi}{\partial n}\Gamma\nabla^2\mu_\chi \tag{39}$$

$$\frac{\partial\phi}{\partial t} = -\mathbf{v}\nabla\phi - \phi\nabla\cdot\mathbf{v} + \frac{\partial\chi}{\partial\phi}\Gamma\nabla^2\mu_\chi. \tag{40}$$

Thermodynamic Pressure Tensor

It can be shown, from statistical mechanics that the pressure tensor can be obtained from (23) [8,15]:

$$P_{ij}^{th} = p_0\delta_{ij} + \kappa_\chi\frac{\partial\chi}{\partial x_i}\frac{\partial\chi}{\partial x_j}, \tag{41}$$

where the isotropic component is:

$$p_0 = P\frac{\delta\mathcal{F}}{\delta P} + \chi\frac{\delta\mathcal{F}}{\delta\chi} - \left(\psi(\chi,P) + \frac{\kappa_\chi}{2}(\nabla\chi)^2\right). \tag{42}$$

so:

$$p_0 = P + \alpha\left(\frac{3}{4}\chi^4 - \frac{1}{2}\chi^2\right) - \kappa\chi(\nabla^2\chi) - \frac{\kappa}{2}(\nabla\chi)^2 \tag{43}$$

The tensor $\mathbf{P}^{th}$ describes the stresses induced by spatial variation in density and composition and has to be added to the viscous stress tensor $\boldsymbol{\tau}$.

Conservation Equations

Let us now apply the results of the previous section for building a system of fluid dynamics equations suitable for a numerical implementation. Together with the mass

conservation Eqs. (39) and (40) we need the conservation equation for the linear momentum [1]:

$$\rho \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \nabla \cdot \mathbf{P} + \rho \mathbf{g} \,, \tag{44}$$

where $\mathbf{P} = -\mathbf{P}^{th} + \boldsymbol{\tau}$ is the full pressure tensor of (41), $\mathbf{g}$ is the gravity and $\boldsymbol{\tau}$ is the viscous stress tensor that, for a Newtonian fluid is:

$$\tau_{ij} = \lambda e_{kk} \delta_{ij} + 2\mu e_{ij} \,, \tag{45}$$

with λ and μ the bulk and shear viscosities and e_{ij} is the strain rate tensor expressed by:

$$e_{ij} = \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right) . \tag{46}$$

Appendix B – Numerical Implementation

Transformed Equations

A common problem in the numerical solution of the conservation Eqs. (33) and (34) is the output of values having no physical meaning (for instance negative densities for one of the components) with bad effects on the numerical stability of the code. This is due to the fact that both variables are defined over a limited range of values that are $n \in [0, +\infty[$ and $\phi \in [-n, +n]$. These conditions are not explicit in the conservation equations, but this problem can be overcome by making simple changes of variable:

$$n = \mathrm{e}^{l} \,, \tag{47}$$

$$\phi = \mathrm{e}^{l} \tanh q \,. \tag{48}$$

Using the new variable l instead of n the conservation Eq. (39) become:

$$\frac{\partial l}{\partial t} = -\mathbf{v} \cdot \nabla l - \nabla \cdot \mathbf{v} + \frac{1}{\mathrm{e}^{l}} \left(\frac{\partial \chi}{\partial n}\right) \Gamma \nabla^2 \mu \,. \tag{49}$$

Substituting (48) in (40) we first obtain:

$$\frac{\partial q}{\partial t} = -\mathbf{v} \cdot \nabla q - \frac{1}{2} \sinh 2q \left(\frac{\partial l}{\partial t} + \mathbf{v} \cdot \nabla l + \nabla \cdot \mathbf{v}\right) + \frac{1}{\mathrm{e}^{l}} \cosh^2 q \left(\frac{\partial \chi}{\partial \phi}\right) \Gamma \nabla^2 \mu \,. \tag{50}$$

Using (49) and some algebraic manipulation the previous expression become:

$$\frac{\partial q}{\partial t} = -\mathbf{v} \cdot \nabla q - \frac{1}{\mathrm{e}^{l}} \cosh q \left[\frac{\partial \chi}{\partial n} \sinh q + \frac{\partial \chi}{\partial \phi} \cosh q\right] \cdot \Gamma \nabla^2 \mu \tag{51}$$

Boundary and Initial Conditions

Equations (49) and (51) have to be solved setting proper boundary and initial conditions.

The no-slip boundary condition, implemented along the conduit walls is:

$$\mathbf{v} = 0 \,. \tag{52}$$

Another boundary condition is set along the walls:

$$\bigtriangledown \phi \cdot \mathbf{n} = 0 \,. \tag{53}$$

This is the neutral wetting condition [25] and it is needed in order to avoid the wall to behave sticky respect to one of the phases.

At the conduit top and bottom conditions of constant pressure are implementer. On the top the pressure is kept to a reference atmospheric value, while at the bottom it is kept at the hydrostatic pressure value, computed on the initial conditions.

In the initial conditions we set $\mathbf{v} = 0$ and the computational domain is composed of domains having a homogeneous composition separated by smooth interfaces, close to the equilibrium solution of (27). At the top of the model a gas domain represents the atmosphere, while the remaining part of the conduit is filled with magma. A small elliptical gas bubble in hydrostatic equilibrium is placed in the lower part of the conduit. The details are specified for each set of simulations.

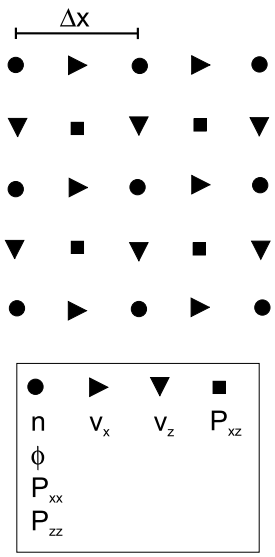
Finite Difference Implementation

Equations (49), (51) and (44) are discretized on a regular grid and the differential equations are transformed in finite-difference equations [19]. Scalar, vector and tensor quantities are discretized on staggered grids [6] (Fig. 12). Scalar quantities (l, q, μ and P_{ii}) are defined at integer grid steps (x_i, z_j). Velocities are staggered half grid step along x and z directions, that is v_x is defined at $(x_{i+1/2}, z_j)$ and v_z at $(x_i, z_{j+1/2})$. The deviatoric part of the pressure tensor P_{xz} is defined at grid points $(x_{i+1/2}, z_{j+1/2})$. This allows a second-order accuracy in the computation of spatial derivative.

We apply also a staggering in time. At the time step k we first solve (49) and (51). Then the current values of l and q are updated:

$$l_{i,j}^{(k)} = l_{i,j}^{(k-1)} + \frac{\mathrm{d}l}{\mathrm{d}t}^{(k)} \Delta t \,, \tag{54}$$

$$q_{i,j}^{(k)} = q_{i,j}^{(k-1)} + \frac{\mathrm{d}q}{\mathrm{d}t}^{(k)} \Delta t \,. \tag{55}$$



Slug Flow: Modeling in a Conduit and Associated Elastic Radiation, Figure 12
Spatial scheme of the staggered grid used in the computation. *Circle* represents the isotropic grid (density, composition and isotropic pressure), *rightward triangles* are the grids for v_x while *downward triangles* are for v_z. *Squares* are the grids used for discretizing P_{xz}

Using these new values the full pressure tensor $\mathbf{P}^{(k)}$ is computed using the discretized version of (41), (42) and (43). At the time step $k + 1/2$ (44) is solved and the values of the local velocities are updated.

The computation of $\frac{\mathrm{d}l}{\mathrm{d}t}^{(k)}$ and $\frac{\mathrm{d}q}{\mathrm{d}t}^{(k)}$ requires the values of $l^{(k-1)}$, $q^{(k-1)}$ and $v^{(k-1/2)}$. On the other hand the computation of $v^{(k+1/2)}$ requires the values of $l^{(k)}$, $q^{(k)}$ and $v^{(k-1/2)}$.

The time step Δt is chosen in order to satisfy a stability condition. Since we are dealing with isothermal flows with a very low Mach number, the condition can be expressed by:

$$\Delta x < \Delta t \; c_{\max} \,, \tag{56}$$

where $c_{\max}$ is the highest sound speed among the two components. In our simulations we have set:

$$\Delta x = 6 \; \Delta t \; c_{\max} \,, \tag{57}$$

Bibliography

Primary Literature

1. Aris R (1962) Vectors, tensors and the basic equation of fluid mechanics. Dover Publications
2. Cahn JW, Hilliard JE (1958) Free energy of a nonuniform system. i. interfacial free energy. J Chem Phys 28(2):258–267
3. Chouet B (1986) Dynamics of a fluid-driven crack in three dimensions by the finite difference method. J Geophys Res 91:13967–13992
4. Chouet B, Dawson P, Ohminato T, Martini M, Saccorotti G, Giudicepietro F, De Luca G, Milana G, Scarpa R (2003) Source mechanisms of explosions at stromboli volcano, italy, determined from moment-tensor inversions of very-long-period data. J Geophys Res 108(B1)
5. Chouet BA (1996) New methods and future trends in seismological volcano monitoring. In: Scarpa R, Tilling RI (eds) Monitoring and mitigation of volcano hazards. Springer
6. Chung TJ (2002) Computational fluid dynamics. Cambridge University Press, Cambridge
7. Clift R, Grace JR, Weber ME (1978) Bubbles, drops and particles. Dover Publications
8. Evans R (1979) The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids. Adv Phys 28(2):143–200
9. James MR, Lane SJ, Chouet B, Gilbert JS (2004) Pressure changes associated with the ascent and bursting of gas slugs in liquid-filled vertical and inclined conduits. J Volc Geotherm Res 129:61–82
10. James MR, Lane SJ, Chouet BA (2006) Gas slug ascent through changes in conduit diameter: Laboratory insight into a volcano-seismic source process in low-viscosity magmas. J Geophys Res 111
11. Jaupart C (2000) Magma ascent at shallow levels. In: Sigurdsson H (ed) Encyclopedia of Volcanoes. Academic Press
12. Jaupart C, Vergniolle S (1988) Laboratory models of hawaiian and strombolian eruptions. Nature 331:58–60
13. Kirchdorfer M (1999) Analysis and quasistatic fe modeling of long period impulsive events associated with explosions at stromboli volcano (italy). Annali di Geofisica 42(3):379–390
14. Ohminato T, Chouet BA, Dawson P, Kedar S (1998) Waveform inversion of very long period impulsive signals associated with magmatic injection beneath kilauea volcano, hawaii. J Geophys Res 103(B10):23839–23862
15. Pooleay CM, Kuksenok O, Balazs AC (2005) Convection-driven pattern in phase-separating binary fluids. Phys Rev E 71:030501(R)
16. Ripepe M, Poggi P, Braun T, Gordeev E (1996) Infrasonic waves and volcanic tremor at stromboli. Geophys Res Lett 23(2):181–184
17. Seyfried R, Freundt A (2000) Experiments on conduit flow and eruption behaviour of basaltic volcanic eruptions. J Geophys Res B10(B10):23727–23740
18. Sparks RSJ (1978) The dynamics of bubble formation and growth in magmas: a review and analysis. J Volc Geotherm Res 3:137–186
19. Strikwerda JC (2004) Finite difference schemes and partial differential equations. SIAM, Philadelphia
20. Succi S (2001) The Lattice Boltzmann Equation for fluid dynamics and beyond. In: Numerical mathematics and scientific computation. Oxford University Press, Oxford
21. Swift MR, Orlandini E, Osborn WR, Yeomans JM (1996) Lattice Boltzmann simulations of liquid-gas and binary fluid system. Phys Rev E 54(5):5041–5052
22. Tomiyama A, Takagi S, Matsumoto Y (1999) Numerical simulation of bubble flows using interface tracking and bubble tracking method. Trans Model Simul 23

23. Vergniolle S, Brandeis G (1996) Strombolian explosions 1. a large bubble breaking at the surface of a lava column as a source of sound. J Geophys Res 101(B9):20433–20447
24. Vergniolle S, Mangan M (2000) Hawaiian and strombolian eruptions. In: Sigurdsson H (ed) Encyclopedia of Volcanoes. Academic Press
25. Xu A, Gonnella G, Lamura A (2003) Phase-separating binary fluids under oscillatory shear. Phys Rev E 67
26. Yue P, Feng JJ, Liu C, Shen J (2004) A diffuse-interface method for simulating two-phase flows of complex fluids. J Fluid Mech 515:293–317

Books and Reviews

Brennen CE (2005) Fundamentals of multiphase flow. Cambridge University Press, Cambridge
Scarpa R, Tilling RI (eds) (1996) Monitoring and mitigation of volcano hazards. Springer
Sigurdsson H, Bruce Houghton BF, McNutt SR, Rymer H, Stix J (eds) (2000) In: Encyclopedia of Volcanoes. Academic Press
Tannehill JC, Anderson DA, Pletcher RH (1997) Computational fluid mechanics and heat transfer, second edn. Taylor and Francis, London
Zobin V (2003) Introduction to volcanic seismology. Elsevier Science

Smooth Ergodic Theory

Amie Wilkinson
Northwestern University, Evanston, USA

Article Outline

Glossary
Definition of the Subject
Introduction
The Volume Class
The Fundamental Questions
Lebesgue Measure and Local Properties of Volume
Ergodicity of the Basic Examples
Hyperbolic Systems
Beyond Uniform Hyperbolicity
The Presence of Critical Points and Other Singularities
Future Directions
Bibliography

Glossary

Conservative, dissipative Conservative dynamical systems (on a compact phase space) are those that preserve a finite measure equivalent to volume. Hamiltonian dynamical systems are important examples of conservative systems. Systems that are not conservative are called dissipative. Finding physically meaningful invariant measures for dissipative maps is a central object of study in smooth ergodic theory.

Distortion estimate A key technique in smooth ergodic theory, a distortion estimate for a smooth map f gives a bound on the variation of the jacobian of f^n in a given region, for n arbitrarily large. The jacobian of a smooth map at a point x is the absolute value of the determinant of derivative at x, measured in a fixed Riemannian metric. The jacobian measures the distortion of volume under f in that metric.

Hopf argument A technique developed by Eberhard Hopf for proving that a conservative diffeomorphism or flow is ergodic. The argument relies on the Ergodic Theorem for invertible transformations, the density of continuous functions among integrable functions, and the existence of stable and unstable foliations for the system. The argument has been used, with various modifications, to establish ergodicity for hyperbolic, partially hyperbolic and nonuniformly hyperbolic systems.

Hyperbolic A compact invariant set $\Lambda \subset M$ for a diffeomorphism $f\colon M \to M$ is hyperbolic if, at every point in Λ, the tangent space splits into two subspaces, one that is uniformly contracted by the derivative of f, and another that is uniformly expanded. Expanding maps and Anosov diffeomorphisms are examples of globally hyperbolic maps. Hyperbolic diffeomorphisms and flows are the archetypical smooth systems displaying chaotic behavior, and their dynamical properties are well-understood. Nonuniform hyperbolicity and partial hyperbolicity are two generalizations of hyperbolicity that encompass a broader class of systems and display many of the chaotic features of hyperbolic systems.

Sinai–Ruelle–Bowen (SRB) measure The concept of SRB measure is a rigorous formulation of what it means for an invariant measure to be "physically meaningful". An SRB measure attracts a large set of orbits into its support, and its statistical features are reflected in the behavior of these attracted orbits.

Definition of the Subject

Smooth ergodic theory is the study of the statistical and geometric properties of measures invariant under a smooth transformation or flow. The study of smooth ergodic theory is as old as the study of abstract ergodic theory, having its origins in Bolzmann's Ergodic Hypothesis in the late 19th Century. As a response to Boltzmann's hypothesis, which was formulated in the context of Hamiltonian Mechanics, Birkhoff and von Neumann defined er-

godicity in the 1930s and proved their foundational ergodic theorems. The study of ergodic properties of smooth systems saw an advance in the work of Hadamard and E. Hopf in the 1930s their study of geodesic flows for negatively curved surfaces. Beginning in the 1950s, Kolmogorov, Arnold and Moser developed a perturbative theory producing obstructions to ergodicity in Hamiltonian systems, known as Kolmogorov–Arnold–Moser (KAM) Theory. Beginning in the 1960s with the work of Anosov and Sinai on hyperbolic systems, the study of smooth ergodic theory has seen intense activity. This activity continues today, as the ergodic properties of systems displaying weak forms of hyperbolicity are further understood, and KAM theory is applied in increasingly broader contexts.

Introduction

This entry focuses on the basic arguments and principles in smooth ergodic theory, illustrating with simple and straightforward examples. The classic texts [1,2] are a good supplement.

The discussion here sidesteps the topic of ▶ Kolmogorov–Arnold–Moser (KAM) Theory, which has played an important role in the development of smooth ergodic theory. For reasons of space, detailed discussion of several active areas in smooth ergodic theory is omitted, including: higher mixing properties (Kolmogorov, Bernoulli, etc.), finer statistical properties (fast decay of correlations, Central Limit Theorem, large deviations), smooth thermodynamic formalism (transfer operators, pressure, dynamical zeta functions, etc.), the smooth ergodic theory of random dynamical systems, as well as any mention of infinite invariant measures. The text [3] covers many of these topics, and the texts [4,5,6] treat random smooth ergodic theory in depth. An excellent discussion of many of the recent developments in the field of smooth ergodic theory is [7].

This entry assumes knowledge of the basic concepts in ergodic theory and of basic differential topology. The texts [8] and [9] contain the necessary background.

The Volume Class

For simplicity, assume that M is a compact, boundaryless C^∞ Riemannian manifold, and that $f: M \to M$ is an orientation-preserving, C^1 map satisfying $m(D_x f) > 0$, for all $x \in M$, where

$$m(D_x f) = \inf_{v \in T_x M, \|v\|=1} \|D_x f(v)\| .$$

If f is a diffeomorphism, then this assumption is automatically satisfied, since in that case $m(D_x f) = \|D_{f(x)} f^{-1}\|^{-1} > 0$. For non-invertible maps, this assumption is essential in much of the following discussion. The Inverse Function Theorem implies that any map f satisfying these hypotheses is a covering map of positive degree $d \geq 1$.

These assumptions will avoid the issues of infinite measures and the behavior of f near critical points and singularities of the derivative. For most results discussed in this entry, this assumption is not too restrictive. The existence of critical points and other singularities is, however, a complication that cannot be avoided in many important applications. The ergodic-theoretic analysis of such examples can be considerably more involved, but contains many of the elements discussed in this entry. The discussion in Sect. "Beyond Uniform Hyperbolicity" indicates how some of these additional technicalities arise and can be overcome. For simplicity, the discussion here is confined almost exclusively to discrete time evolution. Many, though not all, of the the results mentioned here carry over to flows and semiflows using, for example, a cross-section construction (see Chap. 1 in [2]).

Every smooth map $f: M \to M$ satisfying these hypotheses preserves a natural measure *class*, the measure class of a finite, smooth Riemannian volume on M. Fix such a volume ν on M. Then there exists a continuous, positive *jacobian* function $x \mapsto \mathrm{jac}_x f$ on M, with the property that for every sufficiently small ball $B \subset M$, and every measurable set $A \subset B$ one has:

$$\nu(f(A)) = \int_B \mathrm{jac}_x f \, d\nu(x) .$$

The jacobian of f at x is none other than the absolute value of the determinant of the derivative $D_x f$ (measured in the given Riemannian metric). To see that the measure class of ν is preserved by f, observe that the Radon–Nikodym derivative $\frac{df_*\nu}{d\nu}(x)$ at x is equal to $\sum_{y \in f^{-1}(x)} (\mathrm{jac}_y f)^{-1} > 0$. Hence $f_*\nu$ is equivalent to ν, and f preserves the measure class of ν.

In many contexts, the map f has a natural *invariant* measure in the measure class of volume. In this case, f is said to be *conservative*. One setting in which a natural invariant smooth measure appears is Hamiltonian dynamics. Any solution to Hamilton's equations preserves a smooth volume called the *Liouville measure*. Furthermore, along the invariant, constant energy hypermanifolds of a Hamiltonian flow, the Liouville measure decomposes smoothly into invariant measures, each of which is equivalent to the induced Riemannian volume. In this way, many systems of physical or geometric origin, such as billiards, geodesic flows, hard sphere gases, and evolution of the n-body problem give rise to smooth conserva-

tive dynamical systems. See ► Dynamics of Hamiltonian Systems.

Note that even though f preserves a smooth measure class, it might not preserve any measure in that measure class. Consider, for example, a diffeomorphism $f\colon S^1 \to S^1$ of the circle with exactly two fixed points, p and q, $f'(p) > 1 > f'(q) > 0$. Let μ be an f-invariant probability measure. Let I be a neighborhood of p. Then $\bigcap_{n=1}^{\infty} f^{-n}(I) = \{p\}$, but on the other hand, $\mu(f^{-n}(I)) = \mu(I) > 0$, for all n. This implies that $\mu(\{p\}) > 0$, and so μ does not lie in the measure class of volume. This is an example of a dissipative map. A map f is called *dissipative* if every f-invariant measure with full support has a singular part with respect to volume. As was just seen, if a diffeomorphism f has a periodic sink, then f is dissipative; more generally, if a diffeomorphism f has a periodic point p of period k such that $\mathrm{jac}_p f^k \neq 1$, then f is dissipative.

The Fundamental Questions

For a given smooth map $f\colon M \to M$, there are the following fundamental questions.

1. Is f conservative? That is, does there exist an invariant measure in the class of volume? If so, is it unique?
2. When f is conservative, what are its statistical properties? Is it ergodic, mixing, a K-system, Bernoulli, etc.? Does it obey a Central Limit Theorem, fast decay of correlations, large deviations estimates, etc.?
3. If f is dissipative, does there exist an invariant measure, not in the class of volume, but (in some sense) natural with respect to volume? What are the statistical properties of such a measure, if it exists?

There are several plausible ways to "answer" these questions. One might fix a given map f of interest and ask these questions for that specific f. What tends to happen in the analysis of a single map f is that either:

- the question can be answered using "soft" methods, and so the answer applies not only to f but to perturbations of f, or even to *generic* or *typical* f inside a class of maps; or
- the proof requires "hard" analysis or precise asymptotic information and cannot possibly be answered for a specific f, but can be answered for a large set of f_t in a typical (or given) parametrized family $\{f_t\}_{t\in(-1,1)}$ of smooth maps containing $f = f_0$.

Both types of results appear in the discussion that follows.

Lebesgue Measure and Local Properties of Volume

Locally, any measure in the measure class of volume is, after a smooth change of coordinates, equivalent to Lebesgue measure in $\mathbb{R}^n$. In fact, more is true: Moser's Theorem implies that locally any Riemannian volume is, after a smooth change of coordinates, *equal* to Lebesgue measure in $\mathbb{R}^n$. Hence to study many of the local properties of volume, it suffices to study the same properties for Lebesgue measure.

One of the basic properties of Lebesgue measure is that every set of positive Lebesgue measure can be approximated arbitrarily well in measure from the outside by an open set, and from the inside by a compact set. A consequence of this property, of fundamental importance in smooth ergodic theory, is the following statement.

Fundamental Principle #1: Two disjoint, positive Lebesgue measure sets cannot mix together uniformly at all scales.

As an illustration of this principle, consider the following elementary exercise in measure theory. First, some notation. If ν is a measure and A and B are ν-measurable sets with $\nu(B) > 0$, the *density of A in B* is defined by:

$$\nu(A\colon B) = \frac{\nu(A \cap B)}{\nu(B)}\,.$$

Proposition 1 *Let $\mathcal{P}_1, \mathcal{P}_2, \ldots$ be sequence of (mod 0) finite partitions of the circle S^1 into open intervals, with the properties: a) any element of $\mathcal{P}_n$ is a (mod 0) union of elements of $\mathcal{P}_{n+1}$, and b) the maximum diameter of elements of $\mathcal{P}_n$ tends to 0 as $n \to \infty$.*

Let A be any set of positive Lebesgue measure in S^1. Then there exists a sequence of intervals $I_1, I_2, \ldots$, with $I_n \in \mathcal{P}_n$ such that:

$$\lim_{n\to\infty} \lambda(A\colon I_n) = 1\,.$$

Proof Assume that Lebesgue measure has been normalized so that $\lambda(S^1) = 1$. Fix a (mod 0) cover of $S^1 \setminus A$ by pairwise disjoint elements $\{J_i\}$ of the union $\bigcup_{n=1}^{\infty} \mathcal{P}_n$ with the properties:

$$\lambda(J_1) \geq \lambda(J_2) \geq \cdots, \text{ and}$$

$$\lambda\left(\bigcup_{i=1}^{\infty} J_i\right) = \sum_{i=1}^{\infty} \lambda(J_i) < 1\,.$$

For $n \in \mathbb{N}$, let U_n be the union of all the intervals J_i that are contained in $\mathcal{P}_n$, and let $V_n = \bigcup_{i=1}^{n} U_n$. This defines an increasing sequence of natural numbers $i_1 = 1 < i_2 < i_3 < \cdots$ such that $U_n = \bigcup_{i=i_n}^{i_{n+1}-1} J_n$ and $V_n = \bigcup_{i=1}^{i_{n+1}-1} J_n$.

For each n, the interval I_n will be chosen to be an element $\mathcal{P}_n$, disjoint from V_n, in which the density of $\bigcup_{i=n+1}^{\infty} U_i$ is very small (approaching 0 as $n \to \infty$). Since $(S^1 \setminus A) \cap I_n$ is contained in $\bigcup_{i=n+1}^{\infty} U_i$, this choice of I_n will ensure that the density of A in I_n is large (approaching 1 as $n \to \infty$).

To make this choice of I_n, note first that the density of $\bigcup_{i=n+1}^{\infty} U_i$ inside of $S^1 \setminus V_n$ is:

$$\frac{\lambda(\bigcup_{i=n+1}^{\infty} U_i)}{\lambda(S^1 \setminus V_n)} = \frac{\sum_{i=i_{n+1}}^{\infty} \lambda(J_i)}{1 - \sum_{i=1}^{i_{n+1}-1} \lambda(J_i)} = a_n \,.$$

Note that, since $\sum_{i=1}^{\infty} \lambda(J_i)1$, one has $a_n \to 0$ as $n \to \infty$. Since the density of $\bigcup_{i=n+1}^{\infty} U_i$ inside of $S^1 \setminus V_n$ is at most a_n, there is an interval I_n in $\mathcal{P}_n$, disjoint from V_n, such that the density of $\bigcup_{i=n+1}^{\infty} U_i$ inside of I_n is at most a_n. Then

$$\lim_{n\to\infty} \lambda(A\colon I_n) \geq \lim_{n\to\infty} 1 - a_n = 1 \,.$$

□

In smooth ergodic theory, it is often useful to use a variation on Proposition 1 (generally, in higher dimensions) in which the partitions $\mathcal{P}_n$ are nested, dynamically-defined partitions. A simple application of this method can be used to prove that the doubling map on the circle is ergodic with respect to Lebesgue measure, which is done in Sect. "Lebesgue Measure and Local Properties of Volume".

Notice that this proposition does not claim that the intervals I_n are nested. If one imposes stronger conditions on the partitions $\mathcal{P}_n$, then one can draw stronger conclusions.

A very useful theorem in this respect is the Lebesgue Density Theorem. A point $x \in M$ is a *Lebesgue density point* of a measurable set $X \subseteq M$ if

$$\lim_{r\to 0} m(X\colon B_r(x)) = 1 \,,$$

where $B_r(x)$ is the Riemannian ball of radius r centered at x. Notice that the notion of Lebesgue density point depends only on the smooth structure of M, because any two Riemannian metrics have the same Lebesgue density points. The Lebesgue Density Theorem states that if A is a measurable set and $\widehat{A}$ is the set of Lebesgue density points of A, then $m(A \,\Delta\, \widehat{A}) = 0$.

Ergodicity of the Basic Examples

This section contains proofs of the ergodicity of two basic examples of conservative smooth maps: irrational rotations on the circle and the doubling map on the circle. See ► Ergodic Theory: Basic Examples and Constructions for a more detailed description of these maps. These proofs serve as an elementary illustration of some of the fundamental techniques and principles in smooth ergodic theory.

Rotations on the circle. Denote by S^1 the circle $\mathbb{R}/\mathbb{Z}$, which is an additive group, and by λ normalized Lebesgue-Haar measure on S^1. Fix a real number $\alpha \in \mathbb{R}$. The rotation $R_\alpha\colon S^1 \to S^1$ is the translation defined by $R_\alpha(x) = x + \alpha$. Since translations preserve Lebesgue-Haar measure, the map R_α is conservative. Note that R_α is a diffeomorphism and an isometry with respect to the canonical flat metric (length) on S^1.

Proposition 2 *If $\alpha \notin \mathbb{Q}$, then the rotation $R_\alpha\colon S^1 \to S^1$ is ergodic with respect to Lebesgue measure.*

Proof Let A be an R_α-invariant set in S^1, and suppose that $0 < \lambda(A) < 1$. Denote by A^c the complement of A in S^1. Fix $\varepsilon > 0$. Proposition 1 implies that there exists an interval $I \subset S^1$ such that the density of A in I is large: $\lambda(A\colon I) > 1 - \varepsilon$. Similarly, one may choose an interval J such that $\lambda(A^c\colon J) > 1 - \varepsilon$. Without loss of generality, one may choose I and J to have the same length. Since α is irrational, R_α has a dense orbit, which meets the interval I. Since R_α is an isometry, this implies that there is an integer n such that $\lambda(R_\alpha^n(I) \,\Delta\, J) < \varepsilon\lambda(I)$. Since $\lambda(I) = \lambda(J)$, this readily implies that $|\lambda(A\colon R_\alpha^n(I)) - \lambda(A\colon J)| < \varepsilon$. Also, since A is invariant, and R_α is invertible and preserves measure, one has:

$$\lambda(A\colon R_\alpha^n(I)) = \lambda(R_\alpha^n(A)\colon R_\alpha^n(I)) = \lambda(A\colon I) > 1 - \varepsilon \,.$$

But for ε sufficiently small, this contradicts the facts that $\lambda(A\colon J) = 1 - \lambda(A^c\colon J) < \varepsilon$ and $|\lambda(A\colon R_\alpha^n(I)) - \lambda(A\colon J)| < \varepsilon$. □

Note that this is not a proof of the strongest possible statement about R_α (namely, minimality and unique ergodicity). The point here is to show how "soft" arguments are often sufficient to establish ergodicity; this proof uses no more about R_α than the fact that it is a transitive isometry. Hence the same argument shows:

Theorem 1 *Let $f\colon M \to M$ be a transitive isometry of a Riemannian manifold M. Then f is ergodic with respect to Riemannian volume.*

One can isolate from this proof a useful principle:

Fundamental Principle #2: Isometries preserve Lebesgue density at all scales, for arbitrarily many iterates.

This principle implies, for example, that a smooth action by a compact Lie group on M is ergodic along typical (nonsingular) orbits. This principle is also useful in studying area-preserving flows on surfaces and, in a refined form, unipotent flows on homogeneous spaces. In

the case of surface flows, ergodicity questions can be reduced to a study of interval exchange transformations. See the entry ► Ergodic Theory: Basic Examples and Constructions for a detailed discussion of interval exchange transformations and flows on surfaces. ► Ergodic Theory, Introduction to contains detailed information on unipotent flows.

Doubling map on the circle. Let $T_2 : S^1 \to S^1$ be the doubling map defined by $T_2(x) = 2x$. Then T_2 is a degree-2 covering map and endomorphism of S^1 with constant jacobian $\mathrm{jac}_x T_2 \equiv 2$. Since $\frac{d(T_2)_*\lambda}{d\lambda} = \frac{1}{2} + \frac{1}{2} = 1$, T_2 preserves Lebesgue-Haar measure. The doubling map is the simplest example of a hyperbolic dynamical system, a topic treated in depth in the next section.

As with the previous example, the focus here is on the property of ergodicity. It is again possible to prove much stronger results about T_2, such as Bernoullicity, by other methods. Instead, here is a soft proof of ergodicity that will generalize readily to other contexts.

Proposition 3 *The doubling map $T_2 : S^1 \to S^1$ is ergodic with respect to Lebesgue measure.*

Proof Let A be a T_2-invariant set in S^1 with $\lambda(A) > 0$. Let $p \in S^1$ be the fixed point of T_2, so that $T_2(p) = p$. For each $n \in \mathbb{N}$, the preimages of p under T_2^{-n} define a (mod 0) partition $\mathcal{P}_n$ into 2^n open intervals of length 2^{-n}; the elements of $\mathcal{P}_n$ are the connected components of $S^1 \setminus T_2^{-n}(\{p\})$. Note that the sequence of partitions $\mathcal{P}_1, \mathcal{P}_2, \ldots$ is nested, in the sense of Proposition 1. Restricted to any interval $J \in \mathcal{P}_n$, the map $T_2{}^n$ is a diffeomorphism onto $S^1 \setminus \{p\}$ with constant jacobian $\mathrm{jac}_x(T_2^n) = (T_2^n)'(x) = 2^n$.

Since A is invariant, it follows that $T_2^{-n}(A) = A$. Fix $\varepsilon > 0$. Proposition 1 implies that there exists an $n \in \mathbb{N}$ and an interval $J \in \mathcal{P}_n$ such that $\lambda(A : J) > 1 - \varepsilon$. Note that $T_2^n(A \cap J) \subset A$. But then

$$\begin{aligned}
\lambda(A) &\geq \lambda(T_2^n(A \cap J)) \\
&= \int_{A \cap J} \mathrm{jac}_x(T_2^n)\, d\lambda(x) \\
&= 2^n \lambda(A \cap J) \\
&= 2^n \lambda(A : J)\lambda(J) \\
&> 2^n(1-\varepsilon)\lambda(J) = 1 - \varepsilon\,.
\end{aligned}$$

Since ε was arbitrary, one obtains that $\lambda(A) = 1$. □

In this proof, the facts that the intervals in $\mathcal{P}_n$ have constant length 2^{-n} and that the jacobian of $T_2{}^n$ restricted to such an interval is constant and equal to 2^n are not essential. The key fact really used in this proof is the assertion that the ratio:

$$\frac{\lambda(T_2^n(A \cap J) : T_2^n(J))}{\lambda(A : J)}$$

is bounded, *independently of n*. In this case, the ratio is 1 for all n because T_2 has constant jacobian.

It is tempting to try to extend this proof to other expanding maps on the circle, for example, a C^1, λ-preserving map $f : S^1 \to S^1$ with $d_{C^1}(f, T_2)$ small. Many of the aspects of this proof carry through *mutatis mutandis* for such an f, save for one. A C^1-small perturbation of T_2 will in general no longer have constant jacobian, and the *variation* of the jacobian of f^n on a small interval can be (and often is) unbounded. The reason for this unboundedness is a lack of control of the modulus of continuity of f'. Hence this argument can fail for C^1 perturbations of T_2. On the other hand, the argument still works for C^2 perturbations of T_2, even when the jacobian is not constant.

The principle behind this fact can be loosely summarized:

Fundamental Principle #3: On controlled scales, iterates of C^2 expanding maps distort Lebesgue density in a controlled way.

This principle requires further explanation and justification, which will come in the following section. The C^2 hypothesis in this principle accounts for the fact that almost all results in smooth ergodic theory assume a C^2 hypothesis (or something slightly weaker).

Hyperbolic Systems

One of the most developed areas of smooth ergodic theory is in the study of hyperbolic maps and attractors. This section defines hyperbolic maps and attractors, provides examples, and investigates their ergodic properties. See [10,11] and ► Hyperbolic Dynamical Systems for a thorough discussion of the topological and smooth properties of hyperbolic systems.

A *hyperbolic structure* on a compact f-invariant set $\Lambda \subset M$ is given by a Df-invariant splitting $T_\Lambda M = E^u \oplus E^s$ of the tangent bundle over Λ and constants $C, \mu > 1$ such that, for every $x \in \Lambda$ and $n \in \mathbb{N}$:

$$\begin{aligned}
v \in E^u(x) &\implies \|D_x f^n(v)\| \geq C^{-1}\mu^n \|v\|\,, \\
\text{and} \quad v \in E^s(x) &\implies \|D_x f^n(v)\| \leq C\mu^{-n}\|v\|\,.
\end{aligned}$$

A hyperbolic attractor for a map $f : M \to M$ is given by an open set $U \subset M$ such that: $f(U) \subset \overline{U}$, and such that the set $\Lambda = \bigcap_{n \geq 0} f^n(U)$ carries a hyperbolic structure. The set Λ is called the *attractor*, and U is an *attracting region*. A map $f : M \to M$ is *hyperbolic* if M decomposes (mod 0)

into a finite union of attracting regions for hyperbolic attractors. Typically one assumes as well that the restriction of f to each attractor Λ_i is topologically transitive.

Every point p in a hyperbolic set Λ has smooth *stable manifold* $\mathcal{W}^{\mathrm{s}}(p)$ and *unstable manifold* $\mathcal{W}^{\mathrm{u}}(p)$, tangent, respectively, to the subspaces $E^s(p)$ and $E^{\mathrm{u}}(p)$. The set $\mathcal{W}^{\mathrm{s}}(p)$ is precisely the set of $q \in M$ such that $d(f^n(p), f^n(q))$ tends to 0 as $n \to \infty$, and it follows that $f(\mathcal{W}\mathrm{s}(p)) = \mathcal{W}^s(f(p))$. When f is a diffeomorphism, the unstable manifold $\mathcal{W}^{\mathrm{u}}(p)$ is uniquely defined and is the stable manifold of f^{-1}. When f is not invertible, local unstable manifolds exist, but generally are not unique. If Λ is a transitive hyperbolic attractor, then every unstable manifold of every point $p \in \Lambda$ is dense in Λ.

Examples of Hyperbolic Maps and Attractors

Expanding Maps The previous section mentioned briefly the C^r perturbations of the doubling map T_2. Such perturbations (as well as T_2 itself) are examples of *expanding maps*. A map $f\colon M \to M$ is *expanding* if there exist constants $\mu > 1$ and $C > 0$ such that, for every $x \in M$, and every nonzero vector $v \in T_x M$:

$$\|D_x f^n(v)\| \geq C\mu^n \|v\| ,$$

with respect to some (any) Riemannian metric on M. An expanding map is clearly hyperbolic, with $U = M$, E^s the trivial bundle, and $E^u = TM$. Any disk in M is a local unstable manifold for f.

Anosov Diffeomorphisms A diffeomorphism $f\colon M \to M$ is called *Anosov* if the tangent bundle splits as a direct sum $TM = E^u \oplus E^s$ of two Df-invariant subbundles, such that E^u is uniformly expanded and E^s is uniformly contracted by Df. Similarly, a flow $\varphi_t\colon M \to M$ is called Anosov if the tangent bundle splits as a direct sum $TM = E^u \oplus E^0 \oplus E^s$ of three $D\varphi_t$-invariant subbundles, such that E^0 is generated by $\dot{\varphi}$, E^u is uniformly expanded and E^s is uniformly contracted by $D\varphi_t$. Like expanding maps, an Anosov diffeomorphism is an Anosov attractor with $\Lambda = U = M$.

A simple example of a conservative Anosov diffeomorphism is a hyperbolic linear automorphism of the torus. Any matrix $A \in SL(n, \mathbb{Z})$ induces an automorphism of $\mathbb{R}^n$ preserving the integer lattice $\mathbb{Z}^n$, and so descends to an automorphism $f_A\colon T^n \to T^n$ of the n-torus $T^n = \mathbb{R}^n/\mathbb{Z}^n$. Since the determinant of A is 1, the diffeomorphism f_A preserves Lebesgue-Haar measure on T^n. In the case where none of the eigenvalues of A have modulus 1, the resulting diffeomorphism f_A is Anosov. The stablebundle E^s at $x \in T^n$ is the parallel translate to x of the sum of the contracted generalized eigenspaces of A, and the unstable bundle E^u at x is the translated sum of expanded eigenspaces.

In general, the invariant subbundles E^u and E^s of an Anosov diffeomorphism are integrable and tangent to a transverse pair of foliations $\mathcal{W}^u$ and $\mathcal{W}^s$, respectively (see, e. g [12]. for a proof of this). The leaves of $\mathcal{W}^s$ are uniformly contracted by f, and the leaves of $\mathcal{W}^u$ are uniformly contracted by f^{-1}. The leaves of these foliations are as smooth as f, but the tangent bundles to the leaves do not vary smoothly in the manifold. The regularity properties of these foliations play an important role in the ergodic properties of Anosov diffeomorphisms.

The first Anosov flows to be studied extensively were the geodesic flows for manifolds of negative sectional curvatures. As these flows are Hamiltonian, they are conservative. Eberhard Hopf showed in the 1930s that such geodesic flows for surfaces are ergodic with respect to Liouville measure [14]; it was not until the 1960s that ergodicity of all such flows was proved by Anosov [15]. The next section describes, in the context of Anosov diffeomorphisms, Hopf's method and important refinements due to Anosov and Sinai.

DA Attractors A simple way to produce a non-Anosov hyperbolic attractor on the torus is to start with an Anosov diffeomorphism, such as a linear hyperbolic automorphism, and deform it in a neighborhood of a fixed point, turning a saddle fixed point into a source, while preserving the stable foliation. If this procedure is carried out carefully enough, the resulting diffeomorphism is a dissipative hyperbolic diffeomorphism, called a *derived from Anosov (DA)* attractor. Other examples of hyperbolic attractors are the Plykin attractor and the solenoid. See [10].

Distortion Estimates

Before describing the ergodic properties of hyperbolic systems, it is useful to pause for a brief discussion of distortion estimates. Distortion estimates are behind almost every result in smooth ergodic theory. In the hyperbolic setting, distortion estimates are applied to the action of f on unstable manifolds to show that the volume distortion of f along unstable manifolds can be controlled for arbitrarily many iterates.

The example mentioned at the end of the previous section illustrates the ideas in a distortion estimate. Suppose that $f\colon S^1 \to S^1$ is a C^2 expanding map, such as a C^2 small perturbation of T_2. Then there exist constants $\mu > 1$ and $C > 0$ such that $(f^n)'(x) > C\mu^n$ for all x and n.

Let d be the degree of f. If I is a sufficiently small open interval in S^1, then for each n, $f^{-n}(I)$ is a union of d disjoint intervals. Furthermore, each of these intervals has diameter at most $C^{-1}\mu^{-n}$ times the diameter of I. It is now possible to justify the assertion in Fundamental Principle #3 in this context.

Lemma *There exists a constant $K \geq 1$ such that, for all $n \in \mathbb{N}$, and for all $x, y \in f^{-n}(I)$, one has:*

$$K^{-1} \leq \frac{(f^n)'(x)}{(f^n)'(y)} \leq K\,.$$

Proof Since f is C^2 and f' is bounded away from 0, the function $\alpha(x) = \log(f'(x))$ is C^1. In particular, α is Lipschitz continuous: there exists a constant $L > 0$ such that, for all $x, y \in S^1$, $|\alpha(x) - \alpha(y)| < Ld(x, y)$. For $n \geq 0$, let $\alpha_n(x) = \log((f^n)'(x))$. The Chain Rule implies that $\alpha_n(x) = \sum_{i=0}^{n-1} \alpha(f^i(x))$.

The expanding hypothesis on f implies that for all $x, y \in f^n(I)$ and for $i = 0, \ldots, n$, one has $d(f^i(x), f^i(y)) \leq C^{-1}\mu^{i-n} d(f^n(x), f^n(y)) \leq C^{-1}\mu^{i-n}$. Hence

$$\begin{aligned}
|\alpha_n(x) - \alpha_n(y)| &\leq \sum_{i=0}^{n-1} |\alpha(f^i(x)) - \alpha(f^i(y))| \\
&\leq L \sum_{i=0}^{n-1} d(f^i(x), f^i(y)) \\
&\leq L \sum_{i=0}^{n-1} C^{-1}\mu^{i-n} \\
&< LC^{-1}\mu^{-1}(1 - \mu^{-1})^{-1}\,.
\end{aligned}$$

Setting $K = \exp(LC^{-1}\mu^{-1}(1 - \mu^{-1})^{-1})$, one now sees that $(f^n)'(x)/(f^n)'(y)$ lies in the interval $[K^{-1}.K]$, proving the claim. □

In this distortion estimate, the function $\alpha\colon M \to \mathbb{R}$ is called a *cocycle*. The same argument applies to any Lipschitz continuous (or even Hölder continuous) cocycle.

Ergodicity of Expanding Maps

The ergodic properties of C^2 expanding maps are completely understood. In particular, every conservative expanding map is ergodic, and every expanding map is conservative. The proofs of these facts use Fundamental Principles #1 and 3 in a fairly direct way.

Every C^2 conservative expanding map is ergodic with respect to volume. The proof is a straightforward adaptation of the proof of Proposition 3 (see, e. g [2].). Here is a description of the proof for $M = S^1$. As remarked earlier, the proof of Proposition 3 adapts easily to a general expanding map $f\colon S^1 \to S^1$ once one shows that for every f-invariant set A, and every connected component J of $f^{-n}(S^1 \setminus \{p\})$, the quantity

$$\frac{\lambda(f^n(A \cap J)\colon f^n(J))}{\lambda(A\colon J)}$$

is bounded independently of n. This is a fairly direct consequence of the distortion estimate in Lemma 6.1 and is left as an exercise.

The same distortion estimates show that every C^2 expanding map is conservative, preserving a probability measure ν in the measure class of volume. Here is a sketch of the proof for the case $M = S^1$. To prove this, consider the push-forward $\lambda_n = f^n_*\lambda$. Then λ_n is equivalent to Lebesgue, and its Radon–Nikodym derivative $d\lambda_n \setminus d\lambda$ is the density function

$$\rho_n(x) = \sum_{y \in f^{-n}(x)} \frac{1}{\mathrm{jac}_y f^n}\,.$$

Since $f^n_*\lambda$ is a probability measure, it follows that $\int_{S^1} \rho_n \, \mathrm{d}\lambda = 1$. A simple argument using the distortion estimate above (and summing up over all d^n branches of f^{-n} at x) shows that there exists a constant $c \geq 1$ such that for all $x, y \in S^1$,

$$c^{-1} \leq \frac{\rho_n(x)}{\rho_n(y)} \leq c\,.$$

Since the integral of ρ_n is 1, the functions ρ_n are uniformly bounded away from 0 and ∞. It is easy to see that the measure $\nu_n = \frac{1}{n}\sum_{i=1}^n f^i_*\lambda$ has density $\frac{1}{n}\sum_{i=1}^n \rho_i$. Let ν be any subsequential weak* limit of ν_n; then ν is absolutely continuous, with density ρ bounded away from 0 and ∞. With a little more care, one can show that ρ is actually Lipschitz continuous.

As a passing comment, the ergodicity of ν and positivity of ρ imply that ν is the unique f-invariant measure absolutely continuous with respect to λ. With more work, one can show that ν is exact. See [2] for details.

Ergodicity of Conservative Anosov Diffeomorphisms

Like conservative C^2 expanding maps, conservative C^2 Anosov diffeomorphisms are ergodic. This subsection outlines a proof of this fact. Unlike expanding maps, however, Anosov diffeomorphisms need not be conservative. The subsection following this one describe a type of invariant measure that is "natural" with respect to volume, called a Sinai-Ruelle-Bowen (or SRB) measure. The central result for hyperbolic systems states that every hyperbolic attractor carries an SRB measure.

The Hopf Argument In the 1930s Hopf [14] proved that the geodesic flow for a compact, negatively-curved surface is ergodic. His method was to study the Birkhoff averages of continuous functions along leaves of the stable and unstable foliations of the flow. This type of argument has been used since then in increasingly general contexts, and has come to be known as the Hopf Argument.

The core of the Hopf Argument is very simple. To any $f: M \to M$ one can associate the *stable equivalence relation* $\sim_s$, where $x \sim_s y$ iff $\lim_{n\to\infty} d(f^n(x), f^n(y)) = 0$. Denote by $W^s(x)$ the stable equivalence class containing x. When f is invertible, one defines the *unstable equivalence relation* to be the stable equivalence relation for f^{-1}, and one denotes by $W^u(x)$ the unstable equivalence class containing x.

The first step in the Hopf Argument is to show that Birkhoff averages for continuous functions are constant along stable and unstable equivalence classes. Let $\phi: M \to \mathbb{R}$ be an integrable function, and let

$$\overline{\phi} = \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \phi \circ f^i \,. \tag{1}$$

Observe that if ϕ is continuous, then for every $x \in M$ and $x' \in W^s(x)$, $\lim_{n\to\infty} |\phi(f^n(x)) - \phi(f^i(x'))| = 0$. It follows immediately that $\overline{\phi}_f(x) = \overline{\phi}_f(x')$. In particular, if the limit in (1) exists at x, then it exists and is constant on $W^s(x)$.

Fundamental Principle #4: Birkhoff averages of continuous functions are constant along stable equivalence classes.

The next step of Hopf's argument confines itself to the situation where f is conservative and Anosov. In this case, f is invertible, the stable equivalence classes are precisely the leaves of the stable foliation $\mathcal{W}^s$, and the unstable equivalence classes are the leaves of the unstable foliation $\mathcal{W}^u$. Since f is conservative, the Ergodic Theorem implies that for every L^2 function ϕ, the function $\overline{\phi}_f$ is equal (mod 0) to the projection of ϕ onto the f-invariant functions in L^2. Since this projection is continuous, and the continuous functions are dense in L^2, to prove that f is ergodic, it suffices to show that the projection of any continuous function is trivial. That is, it suffices to show that for every continuous ϕ, the function $\overline{\phi}_f$ is constant (a.e.).

To this end, let $\phi: M \to \mathbb{R}$ be continuous. Since the f-invariant functions coincide with the f^{-1}-invariant functions, one obtains that $\overline{\phi}_f = \overline{\phi}_{f^{-1}}$ a.e. The previous argument shows $\overline{\phi}_f$ is constant along $\mathcal{W}^s$-leaves and $\overline{\phi}_{f^{-1}}$ is constant along $\mathcal{W}^u$-leaves. The desired conclusion is that $\overline{\phi}_f$ is a.e. constant. It suffices to show this in a local chart, since the manifold M is connected. In a local chart, after a smooth change of coordinates, one obtains a pair of transverse foliations $\mathcal{F}_1, \mathcal{F}_2$ of the cube $[-1, 1]^n$ by disks, and a measurable function $\psi: [-1, 1]^n \to \mathbb{R}$ that is constant along the leaves of $\mathcal{F}_1$ and constant along the leaves of $\mathcal{F}_2$.

When the foliations $\mathcal{F}_1$ and $\mathcal{F}_2$ are smooth (at least C^1), one can perform a further smooth change of coordinates so that $\mathcal{F}_1$ and $\mathcal{F}_2$ are transverse coordinate subspace foliations. In this case, Fubini's theorem implies that any measurable function that is constant along two transverse coordinate foliations is a.e. constant. This completes the proof in the case that the foliations $\mathcal{W}^s$ and $\mathcal{W}^u$ are smooth. In Hopf's original argument, the stable and unstable foliations were assumed to be C^1 foliations (a hypotheses satisfied in the examples he considered, due to low-dimensionality. See also [16], where a pinching condition on the curvature, rather than low dimensionality, implies this C^1 condition on the foliations.)

Absolute Continuity For a general Anosov diffeomorphism or flow, the stable and unstable foliations are not C^1, and so the final step in Hopf's orginal argument does not apply. The fundamental advance of Anosov and Anosov-Sinai was to prove that the stable and unstable foliations of an Anosov diffeomorphism (conservative or not) satisfy a weaker condition than smoothness, called *absolute continuity*. For conservative systems, absolute continuity is enough to finish Hopf's argument, proving that every C^2 conservative Anosov diffeomorphism is ergodic [15,17].

For a definition and careful discussion of absolute continuity of a foliation $\mathcal{F}$, see [13]. Two consequences of the absolute continuity of $\mathcal{F}$ are:

1. (AC1) If $A \subset M$ is any measurable set, then
$$\lambda(A) = 0 \iff \lambda_{\mathcal{F}(x)}(A) = 0 \,, \quad \text{for} \quad \lambda - \text{a.e. } x \in M \,,$$
where $\lambda_{\mathcal{F}(x)}$ denotes the induced Riemannian volume on the leaf of $\mathcal{F}$ through x.
2. (AC2) If τ is any small, smooth disk transverse to a local leaf of $\mathcal{F}$, and $T \subset \tau$ is a 0-set in τ (with respect to the induced Riemannian volume on τ), then the union of the $\mathcal{F}$ leaves through points in T has Lebesgue measure 0 in M.

The proof that $\mathcal{W}^s$ and $\mathcal{W}^u$ are absolutely continuous has a similar flavor to the proof that an expanding map has a unique absolutely continuous invariant measure (although the cocycles involved are Hölder continuous, rather than Lipschitz), and the facts are intimately related.

With absolute continuity of the stable and unstable foliations in hand, one can now prove:

Theorem 2 (Anosov) *Let f be a C^2, conservative Anosov diffeomorphism. Then f is ergodic.*

Proof By the Hopf Argument, it suffices to show that if ψ^s and ψ^u are L^2 functions with the following properties:

1. ψ^s is constant along leaves of $\mathcal{W}^s$,
2. ψ^u is constant along leaves of $\mathcal{W}^u$, and
3. $\psi^s = \psi^u$ a.e.,

then ψ^s (and so ψ^u as well) is constant a.e.

This is proved using the absolute continuity of $\mathcal{W}^u$ and $\mathcal{W}^s$. Since M is connected, one may argue this locally. Let G be the full measure set of $p \in M$ such that $\psi^s = \psi^u$. Absolute continuity of $\mathcal{W}^s$ (more precisely, consequence (AC1) of absolute continuity described above) implies that for almost every $p \in M$, G has full measure in $\mathcal{W}^s(p)$. Pick such a p. Then for almost every $q \in \mathcal{W}^s(p)$, $\psi^s(q) = \psi^u(p)$; defining G' to be the union over all $q \in \mathcal{W}^s(p) \cap G$ of $\mathcal{W}^u(q)$, one obtains that ψ^s is constant on $G \cap G'$. But now, since $\mathcal{W}^s(p) \cap G$ has full measure in $\mathcal{W}^s(p)$, the absolute continuity of $\mathcal{W}^u$ (consequence (AC2) above) implies that G' has full measure in a neighborhood of p. Hence ψ^s is a.e. constant in a neighborhood of p, completing the proof. □

SRB Measures

In the absence of a smooth invariant measure, it is still possible for a map to have an invariant measure that behaves naturally with respect to volume. In computer simulations one observes such measures when one picks a point x at random and plots many iterates of x; in many systems, the resulting picture is surprisingly insensitive to the initial choice of x. What appears to be happening in these systems is that the trajectory of almost every x in an open set U is converging to the support of a singular invariant probability measure μ. Furthermore, for any open set V, the proportion of forward iterates of x spent in V appears to converge to $\mu(V)$ as the number of iterates tends to ∞.

In the 1960s and 70s, Sinai, Ruelle and Bowen rigorously established the existence of these physically observable measures for hyperbolic attractors [18,19,20]. Such measures are now known as Sinai-Ruelle-Bowen (SRB) measures, and have been shown to exist for non-hyperbolic maps with some hyperbolic features. This subsection describes the construction of SRB measures for hyperbolic attractors.

An f-invariant probability measure μ is called an *SRB (or physical) measure* if there exists an open set $U \subset M$ containing the support of μ such that, for every continuous function $\phi : M \to \mathbb{R}$ and λ-a.e. $x \in U$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \phi(f^i(x)) = \int_M \phi \, d\mu \, .$$

The maximal open set U with this property is called the *basin* of f. To exclude the possibility that the SRB measure is supported on a periodic sink, one often adds the condition that at least one of the Lyapunov exponents of f with respect μ is positive. Other definitions of SRB measure have been proposed (see [21]). Note that every ergodic absolutely continuous invariant measure with positive density in an open set is an SRB measure. Note also that an SRB measure for f is not in general an SRB measure for f^{-1}, unless f preserves an ergodic absolutely continuous invariant measure.

Every transitive Anosov diffeomorphism carries a unique SRB measure. To prove this, one defines a sequence of probability measures ν_n on M as follows. Fix a point $p \in M$, and define ν_0 to be the normalized restriction of Riemannian volume to a ball B^u in $\mathcal{W}^u(p)$. Set $\nu_n = \frac{1}{n} \sum_{i=1}^{n} f^i_* \nu_0$. Distortion estimates show that the density of ν_n on its support inside $\mathcal{W}^u$ is bounded, above and below, independently of n. Passing to a subsequential weak* limit, one obtains a probability measure ν on M with bounded densities on $\mathcal{W}^u$-leaves.

To show that ν is an SRB measure, choose a point $q \in M$ in the support of ν. Since ν has positive density on unstable manifolds, almost every point in a neighborhood of q in $\mathcal{W}^u(q)$ is a regular point for f (that is, a point where the forward Birkhoff averages of every continuous function exist). A variation on the Hopf Argument, using the absolute continuity of $\mathcal{W}^s$, shows that ν is an ergodic SRB measure.

A similar argument shows that every transitive hyperbolic attractor admits an ergodic SRB measure. In fact this SRB measure has much stronger mixing properties, namely, it is Bernoulli. To prove this, one first constructs a Markov partition [22] conjugating the action of f to a Bernoulli shift. This map sends the SRB measure to a Gibbs state for a mixing Markov shift (see ► Pressure and Equilibrium States in Ergodic Theory). A result that subsumes all of the results in this section is:

Theorem 3 (Sinai, Ruelle, Bowen) *Let $\Lambda \subset M$ be a transitive hyperbolic attractor for a C^2 map $f : M \to M$. Then f has an ergodic SRB measure μ supported on Λ. Moreover: the disintegration of μ along unstable manifolds of Λ is equivalent to the induced Riemannian volume, the Lyapunov exponents of μ are all positive, and μ is Bernoulli.*

Beyond Uniform Hyperbolicity

The methods developed in the smooth ergodic theory of hyperbolic maps have been extended beyond the hyperbolic context. Two natural generalizations of hyperbolicity are:

- partial hyperbolicity, which requires uniform expansion of E^u and uniform contraction of E^s, but allows central directions at each point, in which the expansion and contraction is dominated by the behavior in the hyperbolic directions; and
- nonuniform hyperbolicity, which requires hyperbolicity along almost every orbit, but allows the expansion of E^u and the contraction of E^s to weaken near the exceptional set where there is no hyperbolicity.

This section discusses both generalizations.

Partial Hyperbolicity

Brin and Pesin [23] and independently Pugh and Shub [24] first examined the ergodic properties of partially hyperbolic systems soon after the work of Anosov and Sinai on hyperbolic systems. One says that a diffeomorphism $f: M \to M$ of a compact manifold M is *partially hyperbolic* if there is a nontrivial, continuous splitting of the tangent bundle, $TM = E^{\mathrm{s}} \oplus E^{\mathrm{c}} \oplus E^{\mathrm{u}}$, invariant under Df, such that E^{s} is uniformly contracted, E^{u} is uniformly expanded, and E^c is dominated, meaning that for some $n \geq 1$ and for all $x \in M$:

$$\|D_x f^n|_{E^s}\| < m(D_x f^n|_{E^c}) \leq \|D_x f^n|_{E^c}\| < m(D_x f^n|_{E^u}).$$

Partial hyperbolicity is a C^1-open condition: any diffeomorphism sufficiently C^1-close to a partially hyperbolic diffeomorphism is itself partially hyperbolic. For an extensive discussion of examples of partially hyperbolic dynamical systems, see the survey article [25] and the book [26]. Among these examples are: the time-1 map of an Anosov flow, the frame flow for a compact manifold of negative sectional curvature, and many affine transformations of compact homogeneous spaces. All of these examples preserve the volume induced by a Riemannian metric on M.

As in the Anosov case, the stable and unstable bundles E^s and E^u of a partially hyperbolic diffeomorphism are tangent to foliations, again denoted by $\mathcal{W}^s$ and $\mathcal{W}^u$ respectively [23]. Brin-Pesin and Pugh-Shub proved that these foliations are absolutely continuous.

A partially hyperbolic diffeomorphism $f: M \to M$ is *accessible* if any point in M can be reached from any other along an *su-path*, which is a concatenation of finitely many subpaths, each of which lies entirely in a single leaf of $\mathcal{W}^s$ or a single leaf of $\mathcal{W}^u$. Accessibility is a global, topological property of the foliations $\mathcal{W}^u$ and $\mathcal{W}^s$ that is the analogue of transversality of $\mathcal{W}^u$ and $\mathcal{W}^s$ for Anosov diffeomorphisms. In fact, the transversality of these foliations in the Anosov case immediately implies that every Anosov diffeomorphism is accessible. Fundamental Principle #4 suggests that accessibility might be related to ergodicity for conservative systems.

Conservative Partially Hyperbolic Diffeomorphisms Motivated by a breakthrough result with Grayson [27], Pugh and Shub conjectured that accessibility implies ergodicity, for a C^2, partially hyperbolic conservative diffeomorphism [28]. This conjecture has been proved under the hypothesis of center bunching [29], which is a mild spectral condition on the restriction of Df to the center bundle E^{c}. Center bunching is satisfied by most examples of interest, including all partially hyperbolic diffeomorphisms with $\dim(E^c) = 1$. The proof in [29] is a modification of the Hopf Argument using Lebesgue density points and a delicate analysis of the geometric and measure-theoretic properties of the stable and unstable foliations.

In the same article, Pugh and Shub also conjectured that accessibility is a widespread phenomenon, holding for an open and dense set (in the C^{r} topology) of partially hyperbolic diffeomorphisms. This conjecture has been proved completely for $r = 1$ [30], and for all r, with the additional assumption that the central bundle E^{c} is one dimensional [31].

Together, these two conjectures imply the third, central conjecture: in [28]:

Conjecture 1 (Pugh–Shub) *For any $r \geq 2$, the C^r, conservative partially hyperbolic diffeomorphisms contain a C^r open and dense set of ergodic diffeomorphisms.*

The validity of this conjecture in the absence of center bunching is currently an open question.

Dissipative Partially Hyperbolic Diffeomorphisms There has been some progress in constructing SRB-type measures for dissipative partially hyperbolic diffeomorphisms, but the theory is less developed than in the conservative case. Using the same construction as for Anosov diffeomorphisms, one can construct invariant probability measures that are smooth along the $\mathcal{W}^u$ foliation [32]. Such measures are referred to as u-Gibbs measures. Since the stable bundle E^{s} is not transverse to the unstable bundle E^{u}, the Anosov argument cannot be carried through to show that u-Gibbs measures are SRB measures.

Nonetheless, there are conditions that imply that a u-Gibbs measure is an SRB measure: for example,

a u-Gibbs measure is SRB if it is the unique u-Gibbs measure [33], if the bundle $E^s \oplus E^c$ is nonuniformly contracted [34], or if the bundle $E^u \oplus E^c$ is nonuniformly expanded [35]. The proofs of the latter two results use Pesin Theory, which is explained in the next subsection.

SRB measures have also been constructed in systems where E^c is nonuniformly hyperbolic [36], and in (noninvertible) partially hyperbolic covering maps where E^c is 1-dimensional [37]. It is not known whether accessibility plays a role in the existence of SRB measures for dissipative, non-Anosov partially hyperbolic diffeomorphisms.

Nonuniform Hyperbolicity

The concept of Lyapunov exponents gives a natural way to extend the notion of hyperbolicity to systems that behave hyperbolically, but in a nonuniform manner. The fundamental principles described above, suitably modified, apply to these nonuniformly hyperbolic systems and allow for the development of a smooth ergodic theory for these systems. This program was initially proposed and carried out by Yakov Pesin in the 1970s [38] and has come to be known as *Pesin theory*.

Oseledec's Theorem implies that if a smooth map f satisfying the condition $m(D_x f) > 0$ preserves a probability measure ν, then for ν-a.e. $x \in M$ and every nonzero vector $v \in T_x M$, the limit

$$\lambda(x, v) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \log \|D_x f^i(v)\|$$

exists. The number $\lambda(x, v)$ is called the *Lyapunov exponent at x in the direction of v*. For each such x, there are finitely many possible values for the exponent $\lambda(x, v)$, and the function $x \mapsto \lambda(x, \cdot)$ is measurable. See the discussion of Oseledec's Theorem in ► Ergodic Theorems.

Let f be a smooth map. An f-invariant probability measure μ is *hyperbolic* if the Lyapunov exponents of μ-a.e. point are all nonzero. Observe that any invariant measure of a hyperbolic map is a hyperbolic measure.

A conservative diffeomorphism $f : M \to M$ is *nonuniformly hyperbolic* if the invariant measure equivalent to volume is hyperbolic. The term "nonuniform" is a bit misleading, as uniformly hyperbolic conservative systems are also nonuniformly hyperbolic. Unlike uniform hyperbolicity, however, nonuniform hyperbolicity allows for the *possibility* of different strengths of hyperbolicity along different orbits.

Nonuniformly hyperbolic diffeomorphisms exist on all manifolds [39,40], and there are C^1-open sets of nonuniformly hyperbolic diffeomorphisms that are not Anosov diffeomorphisms [41]. In general, it is a very difficult problem to establish whether a given map carries a hyperbolic measure that is nonsingular with respect to volume.

Hyperbolic Blocks As mentioned above, the derivative of f along almost every orbit of a nonuniformly hyperbolic system looks like the derivative down the orbit of a uniformly hyperbolic system; the nonuniformity can be detected only by examining a positive measure set of orbits. Recall that Lusin's Theorem in measure theory states that every Borel measurable function is continuous on the complement of an arbitrarily small measure set. A sort of analogue of Lusin's theorem holds for nonuniformly hyperbolic maps: every C^2, nonuniformly hyperbolic diffeomorphism is uniformly hyperbolic on a (noninvariant) compact set whose complement has arbitrarily small measure. The precise formulation of this statement is omitted, but here are some of its salient features.

If μ is a hyperbolic measure for a C^2 diffeomorphism, then attached to μ-a.e. point $x \in M$ are transverse, smooth stable and unstable manifolds for f. The collection of all stable manifolds is called the *stable lamination* for f, and the collection of all unstable manifolds is called the *unstable lamination* for f. The stable lamination is invariant under f, meaning that f sends the stable manifold at x into the stable manifold for $f(x)$. The stable manifold through x is contracted uniformly by all positive iterates of f in a neighborhood of x. Analogous statements hold for the unstable manifold of x, with f replaced by f^{-1}.

The following quantites vary measurably in $x \in M$:

- the (inner) radii of the stable and unstable manifolds through x,
- the angle between stable and unstable manifolds at x, and
- the rates of contraction in these manifolds.

The stable and unstable laminations of a nonuniformly hyperbolic system are absolutely continuous. The precise definition of absolute continuity here is slightly different than in the uniformly and partially hyperbolic setting, but the consequences (AC1) and (AC2) of absolute continuity continue to hold.

Ergodic Properties of Nonuniformly Hyperbolic Diffeomorphisms Since the stable and unstable laminations are absolutely continuous, the Hopf Argument can be applied in this setting to show:

Theorem 4 (Pesin) *Let f be C^2, conservative and nonuniformly hyperbolic. Then there exists a (mod 0) partition $\mathcal{P}$ of*

M into countably many f-invariant sets of positive volume such that the restriction of f to each $P \in \mathcal{P}$ is ergodic.

The proof of this theorem is also exposited in [42]. The countable partition can in examples be countably infinite; nonuniform hyperbolicity alone does not imply ergodicity.

The Dissipative Case As mentioned above, establishing the existence of a nonsingular hyperbolic measure is a difficult problem in general. In systems with some global form of hyperbolicity, such as partial hyperbolicity, it is sometimes possible to "borrow" the expansion from the unstable direction and lend it to the central direction, via a small perturbation. Nonuniformly hyperbolic attractors have been constructed in this way [43]. This method is also behind the construction of a C^1 open set of nonuniformly hyperbolic diffeomorphisms in [41].

For a given system of interest, it is sometimes possible to prove that a given invariant measure is hyperbolic by establishing an *approximate* form of hyperbolicity. The idea, due to Wojtkowski and called the *cone method*, is to isolate a measurable bundle of cones in TM defined over the support of the measure, such that the cone at a point x is mapped by $D_x f$ into the cone at $f(x)$. Intersecting the images of these cones under all iterates of Df, one obtains an invariant subbundle of TM over the support of f that is nonuniformly expanded.

Lai-Sang Young has developed a very general method [44] for proving the existence of SRB measures with strong mixing properties in systems that display "some hyperbolicity". The idea is to isolate a region X in the manifold where the first return map is hyperbolic and distortion estimates hold. If this can be done, then the map carries a mixing, hyperbolic SRB measure. The precise rate of mixing is determined by the properties of the return-time function to X; the longer the return times, the slower the rate of mixing.

More results on the existence of hyperbolic measures are discussed in the next section.

An important subject in smooth ergodic theory is the relationship between entropy, Lyapunov exponents, and dimension of invariant measures of a smooth map. Significant results in this area include the Pesin entropy formula [45], the Ruelle entropy inequality [46], the entropy-exponents-dimension formula of Ledrappier–Young [47,48], and the proof by Barreira-Pesin-Schmeling that hyperbolic measures have a well-defined dimension [49]. ► Hyperbolic Dynamical Systems contains a discussion of these results; see this entry there for further information.

The Presence of Critical Points and Other Singularities

Now for a discussion of the aforementioned technical difficulties that arise in the presence of singularities and critical points for the derivative.

Singularities, that is, points where Df (or even f) fails to be defined, arise naturally in the study of billiards and hard sphere gases. The first subsection discusses some progress made in smooth ergodic theory in the presence of singularities.

Critical points, that is, points where Df fails to be invertible, appear inescapably in the study of noninvertible maps. This type of complication already shows up for noninvertible maps in dimension 1, in the study of unimodal maps of the interval. The second subsection discusses the technique of parameter exclusion, developed by Jakobson, which allows for an ergodic analysis of a parametrized family of maps with criticalities.

The technical advances used to overcome these issues in the interval have turned out to have applications to dissipative, nonhyperbolic, diffeomorphisms in higher dimension, where the derivative is "nearly critical" in places. The last subsection describes extensions of the parameter exclusion technique to these near-critical maps.

Hyperbolic Billiards and Hard Sphere Gases

In the 1870s the physicist Ludwig Boltzmann hypothesized that in a mechanical system with many interacting particles, physical measurements (observables), averaged over time, will converge to their expected value as time approaches infinity. The underlying dynamical system in this statement is a Hamiltonian system with many degrees of freedom, and the "expected value" is with respect to Liouville measure. Loosely phrased in modern terms, Boltzmann's hypothesis states that a generic Hamiltonian system of this form will be ergodic on constant energy submanifolds. Reasoning that the time scales involved in measurement of an observable in such a system are much larger than the rate of evolution of the system, Boltzmann's hypothesis allowed him to assume that physical quantities associated to such a system behave like constants.

In 1963, Sinai revived and formalized this ergodic hypothesis, stating it in a concrete formulation known as the Boltzmann-Sinai Ergodic Hypothesis. In Sinai's formulation, the particles were replaced by N hard, elastic spheres, and to compactify the problem, he situated the spheres on a k-torus, $k = 2, 3$. The Boltzmann-Sinai Ergodic Hypothesis is the conjecture that the induced Hamiltonian system

on the $2kN$-dimensional configuration space is ergodic on constant energy manifolds, for any $N \geq 2$.

Sinai verified this conjecture for $N = 2$ by reducing the problem to a billiard map in the plane. As background for Sinai's result, a brief discussion of planar billiard maps follows.

Let $D \subset \mathbb{R}^s$ be a connected region whose boundary ∂D is a collection of closed, piecewise smooth simple curves the plane. The *billiard map* is a map defined (almost everywhere) on $\partial D \times [-\pi, \pi]$. To define this map, one identifies each point $(x, \theta) \in \partial D \times [-\pi, \pi]$ with an inward-pointing tangent vector at x in the plane, so that the normal vector to ∂D at x corresponds to the pair $(x, \pi/2)$. This can be done in a unique way on the smooth components of ∂D. Then $f(x, \theta)$ is obtained by following the ray originating at (x, θ) until it strikes the boundary ∂D for the first time at (x', θ'). Reflecting this vector about the normal at x', define $f(x, \theta) = (x', \pi - \theta')$.

It is not hard to see that the billiard map is conservative. The billiard map is piecewise smooth, but not in general smooth: the degree of smoothness of f is one less than the degree of smoothness of ∂D. In addition to singularities arising from the corners of the table, there are singularities arising in the *second* derivative of f at the tangent vectors to the boundary.

In the billiards studied studied by Sinai, the boundary ∂D consists of a union of concave circular arcs and straight line segments. Similar billiards, but with convex circular arcs, were first studied by Bunimovich [51]. Sinai and Bunimovich proved that these billiards are ergodic and nonuniformly hyperbolic. For the Boltzmann-Sinai problem with $N \geq 3$, the relevant associated dynamical system is a higher dimensional billiard table in euclidean space, with circular arcs replaced by cylindrical boundary components.

In a planar billiard table with circular/flat boundary, the behavior of vectors encountering a flat segment of boundary is easily understood, as is the behavior of vectors meeting a circular segment in a neighborhood of the normal vector. If the billiard map is ergodic, however, every open set of vectors will meet the singularities in the table infinitely many times. To establish the nonuniform hyperbolicity of such billiard tables via conefields, it is therefore necessary to understand precisely the fraction of time orbits spend near these singularities. Furthermore, to use the Hopf argument to establish ergodicity, one must avoid the singularities in the second derivative, where distortion estimates break down. The techniques for overcoming these obstacles involve imposing restrictions on the geometry of the table (even more so for higher dimensional tables), and are well beyond the scope of this paper.

The study of hyperbolic billiards and hard sphere gases has a long and involved history. See the articles [50] and [52] for a survey of some of the results and techniques in the area. A discussion of methods in singular smooth ergodic theory, with particular applications to the Lorentz attractor, can be found in [53]. Another, more classical, reference is [54], which contains a formulation of properties on a critical set, due to Katok–Strelcyn, that are useful in establishing ergodicity of systems with singularities.

Interval Maps and Parameter Exclusion

The logistic family of maps $f_t\colon x \mapsto tx(1-x)$ defined on the interval $[0, 1]$ is very simple to define but exhibits an astonishing variety of dynamical features as the parameter t varies. For small positive values of t, almost every point in I is attracted under the map f_t to the sink at -1. For values of $t > 4$, the map has a repelling hyperbolic Cantor set. As the value of t increases between 0 and 4, the map f_t undergoes a cascade of period-doubling bifurcations, in which a periodic sink of period 2^n becomes repelling and a new sink of period 2^{n+1} is born. At the accumulation of period doubling at $t \approx 3.57$, a periodic point of period 3 appears, forcing the existence of periodic points of all periods. The dynamics of f_t for t close to 4 has been the subject of intense inquiry in the last 20 years.

The map f_t, for t close to 4, shares some of the features of the doubling map T_2; it is 2-to-1, except at the critical point $\frac{1}{2}$, and it is uniformly expanding in the complement of a neighborhood of this critical point. Because this neighborhood of the critical point is not invariant, however, the only invariant sets on which f_t is uniformly hyperbolic have measure zero. Furthermore, the second derivative of f_t vanishes at the critical point, making it impossible to control distortion for orbits that spend too much time near the critical point.

Despite these serious obstacles, Michael Jakobson [55] found a method for constructing absolutely continuous invariant measures for maps in the logistic family. The method has come to be known as *parameter exclusion* and has seen application far beyond the logistic family. As with billiards, it is possible to formulate geometric conditions on the map f_t that control both expansion (hyperbolicity) and distortion on a positive measure set. As these conditions involve understanding infinitely many iterates of f_t, they are impossible to verify for a given parameter value t.

Using an inductive formulation of this condition, Jakobson showed that the set of parameters t near 4 that fail to satisfy the condition at iterate n have exponentially small measure (in n). He thereby showed that for a posi-

tive Lebesgue measure set of parameter values t, the map f_t has an absolutely continuous invariant measure [55]. This measure is ergodic (mixing) and has a positive Lyapunov exponent. The delicacy of Jakobson's approach is confirmed by the fact that for an open and dense set of parameter values, almost every orbit is attracted to a periodic sink, and so f_t has no absolutely continuous invariant measure [56,57]. Jakobson's method applies not only to the logistic family but to a very general class of C^3 one-parameter families of maps on the interval.

Near-Critical Diffeomorphisms

Jakobson's method in one dimension proved to extend to certain highly dissipative diffeomorphisms. The seminal paper in this extension is due to Benedicks and Carleson; the method has since been extended in a series of papers [59,60,61] and has been formulated in an abstract setting [62].

This extension turns out to be highly nontrivial, but it is possible to describe informally the similarities between the logistic family and higher-dimensional "near critical" diffeomorphisms. The diffeomorphisms to which this method applies are crudely hyperbolic with a one dimensional unstable direction. Roughly this means that in some invariant region of the manifold, the image of a small ball under f will be stretched significantly in one direction and shrunk in all other directions. The directions of stretching and contraction are transverse in a large proportion of the invariant region, but there are isolated "near critical" subregions where expanding and contracting directions are nearly tangent.

The dynamics of such a diffeomorphism are very close to 1-dimensional if the contraction is strong enough, and the diffeomorphism resembles an interval map with isolated critical points, the critical points corresponding to the critical regions where stable and unstable directions are tangent.

An illustration of this type of dynamics is the Hénon family of maps $f_{a,b}\colon (x, y) \mapsto (1 - ax^2 + by, x)$, the original object of study in Benedicks-Carlesson's work. When the parameter b is set to 0, the map $f_{a,b}$ is no longer a diffeomorphism, and indeed is precisely a projection composed with the logistic map. For small values of b and appropriate values of a, the Hénon map is strongly dissipative and displays the near critical behavior described in the previous paragraph. In analogy to Jakobson's result, there is a positive measure set of parameters near $b = 0$ where $f_{a,b}$ has a mixing, hyperbolic SRB measure.

See [63] for a detailed exposition of the parameter exclusion method for Hénon-like maps.

Future Directions

In addition to the open problems discussed in the previous sections, there are several general questions and problems worth mentioning:

- What can be said about systems with everywhere vanishing Lyapunov exponents? Open sets of such systems exist in arbitrary dimension. Pesin theory carries into the nonuniformly hyperbolic setting the basic principles from uniformly hyperbolic theory (in particular, Fundamental Principles #3 and 4 above). To what extent do properties of isometric and unipotent systems (for example, Fundamental Principle #2) extend to conservative systems all of whose Lyaponov exponents vanish?
- Can one establish the existence of and analyze in general the conservative systems on surfaces that have two positive measure regimes: one where Lyapunov exponents vanish, and the other where they are nonzero? Such systems are conjectured exist in the presence of KAM phenomena surrounding elliptic periodic points.
- On a related note, how common are conservative systems whose Lyapunov exponents are nonvanishing on a positive measure set? See [64] for a discussion.
- Find a broad description of those dissipative systems that admit finitely (or countably) many physical measures. Are such systems dense among all dissipative systems, or possibly generic among a restricted class of systems? See [41,65] for several questions and conjectures related to this problem.
- Extend the methods in the study of systems with singularities to other specific systems of interest, including the infinite dimensional systems that arise in the study of partial differential equations.
- Carry the methods of smooth ergodic theory further into the study of smooth actions of discrete groups (other than the integers) on manifolds. When do such actions admit (possibly non-invariant) "physical" measures?

There are other interesting open areas of future inquiry, but this gives a good sample of the range of possibilities.

Bibliography

1. Arnol'd VI, Avez A (1986) Ergodic problems of classical mechanics. Translated from the French by Avez, A. Benjamin WA, Inc., New York Amsterdam
2. Mañé R (1987) Ergodic theory and differentiable dynamics. Translated from the Portuguese by Silvio L. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol 8. Springer, Berlin

3. Baladi V (2000) Positive transfer operators and decay of correlations. Advanced Series in Nonlinear Dynamics, 16. World Scientific Publishing Co., Inc., River Edge
4. Kifer Y (1986) Ergodic theory of random transformations. Progress in Probability and Statistics, vol 10. Birkhäuser, Boston
5. Kifer Y (1988) Random perturbations of dynamical systems. Progress in Probability and Statistics, 16. Birkhäuser, Boston
6. Liu PD, Qian M (1995) Smooth ergodic theory of random dynamical systems. Lecture Notes in Mathematics, vol 1606. Springer, Berlin
7. Bonatti C, Dìaz LJ, Viana M (2005) Dynamics beyond uniform hyperbolicity. A global geometric and probabilistic perspective. Encyclopaedia of Mathematical Sciences, 102. Mathematical Physics, III. Springer, Berlin
8. Cornfeld IP, Fomin SV, Sinaĭ YG (1982) Ergodic theory. Translated from the Russian by A. B. Sosinskiĭ. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol 245. Springer, New York
9. Hirsch MW (1979) Differential topology. Graduate Texts in Mathematics, No. 33. Springer, New York Heidelberg
10. Robinson C (1995) Dynamical systems. Stability, symbolic dynamics, and chaos. Studies in Advanced Mathematics. CRC Press, Boca Raton
11. Katok A, Hasselblatt B (1995) Introduction to the modern theory of dynamical systems. With a supplementary chapter by Katok and Leonardo Mendoza. Encyclopedia of Mathematics and its Applications, vol 54. Cambridge University Press, Cambridge
12. Hirsch MW, Pugh CC, Shub M (1977) Invariant manifolds. Lecture Notes in Mathematics, vol 583. Springer, Berlin New York
13. Brin M, Stuck G (2002) Introduction to dynamical systems. Cambridge University Press, Cambridge
14. Hopf E (1939) Statistik der geodätischen Linien in Mannigfaltigkeiten negativer Krümmung. Ber Verh Sachs Akad Wiss Leipzig 91:261–304 (German)
15. Anosov DV (1967) Geodesic flows on closed Riemann manifolds with negative curvature. Proceedings of the Steklov Institute of Mathematics, No. 90 1967. Translated from the Russian by S. Feder. American Mathematical Society, Providence
16. Sinaĭ JG (1961) Geodesic flows on compact surfaces of negative curvature. Dokl Akad Nauk SSSR 136:549–552 (Russian); translated as Soviet Math Dokl 2:106–109
17. Anosov DV, Sinaĭ JG (1967) Certain smooth ergodic systems. (Russian) Uspehi Mat Nauk 22 no. 5(137):107–172
18. Sinaĭ JG (1972) Gibbs measures in ergodic theory. (Russian) Uspehi Mat Nauk 27 no. 4(166):21–64
19. Ruelle D (1976) A measure associated with axiom-A attractors. Amer J Math 98(3):619–654
20. Bowen R (1975) Equilibrium states and the ergodic theory of Anosov diffeomorphisms. Lecture Notes in Mathematics, vol 470. Springer, Berlin New York
21. Young LS (2002) What are SRB measures, and which dynamical systems have them? Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays. J Statist Phys 108(5–6):733–754
22. Bowen R (1970) Markov partitions for Axiom A diffeomorphisms. Amer J Math 92:725–747
23. Brin MI, Pesin JB (1974) Partially hyperbolic dynamical systems. Izv Akad Nauk SSSR Ser Mat 38:170–212 (Russian)
24. Pugh C, Shub M (1972) Ergodicity of Anosov actions. Invent Math 15:1–23
25. Burns K, Pugh C, Shub M, Wilkinson A (2001) Recent results about stable ergodicity. Smooth ergodic theory and its applications. (Seattle, 1999), 327–366. Proc Sympos Pure Math, 69, Amer Math Soc, Providence
26. Pesin YB (2004) Lectures on partial hyperbolicity and stable ergodicity. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich
27. Grayson M, Pugh C, Shub M (1994) Stably ergodic diffeomorphisms. Ann of Math 140(2):295–329
28. Pugh C, Shub M (1996) Stable ergodicity and partial hyperbolicity. International Conference on Dynamical Systems (Montevideo, 1995), 182–187, Pitman Res Notes Math Ser, 362, Longman, Harlow
29. Burns K, Wilkinson A, On the ergodicity of partially hyperbolic systems. Ann of Math. To appear
30. Dolgopyat D, Wilkinson A (2003) Stable accessibility is C^1 dense. Geometric methods in dynamics. II. Astérisque No. 287
31. Rodríguez HA, Rodríguez HF, Ures R (2008) Partially hyperbolic systems with 1D-center bundle. Invent Math 172(2)
32. Pesin YB, Sinaĭ YG (1983) Gibbs measures for partially hyperbolic attractors. Ergod Theor Dynam Syst 2(3–4):417–438
33. Dolgopyat D (2004) On differentiability of SRB states for partially hyperbolic systems. Invent Math 155(2):389–449
34. Bonatti C, Viana M (2000) SRB measures for partially hyperbolic systems whose central direction is mostly contracting. Israel J Math 115:157–193
35. Alves JF, Bonatti C, Viana M (2000) SRB measures for partially hyperbolic systems whose central direction is mostly expanding. Invent Math 140(2):351–398
36. Burns K, Dolgopyat D, Pesin Y, Pollicott Mark Stable ergodicity for partially hyperbolic attractors with negative central exponents. Preprint
37. Tsujii M (2005) Physical measures for partially hyperbolic surface endomorphisms. Acta Math 194(1):37–132
38. Pesin JB (1977) Characteristic Ljapunov exponents, and smooth ergodic theory. Uspehi Mat Nauk 32(196):55–112, 287 (Russian)
39. Katok A (1979) Bernoulli diffeomorphisms on surfaces. Ann Math 110(2):529–547
40. Dolgopyat D, Pesin Y (2002) Every compact manifold carries a completely hyperbolic diffeomorphism. Ergod Theor Dynam Syst 22(2):409–435
41. Shub M, Wilkinson A (2000) Pathological foliations and removable zero exponents. Invent Math 139(3):495–508
42. Pugh C, Shub M (1989) Ergodic attractors. Trans Amer Math Soc 312(1):1–54
43. Viana M (1997) Multidimensional nonhyperbolic attractors. Inst Hautes Études Sci Publ Math No 85:63–96
44. Young LS (1998) Statistical properties of dynamical systems with some hyperbolicity. Ann Math 147(2):585–650
45. Pesin JB (1976) Characteristic Ljapunov exponents, and ergodic properties of smooth dynamical systems with invariant measure. Dokl Akad Nauk SSSR 226 (Russian)
46. Ruelle D (1978) An inequality for the entropy of differentiable maps. Bol Soc Brasil Mat 9(1):83–87
47. Ledrappier F, Young LS (1985) The metric entropy of diffeomorphisms. I. Characterization of measures satisfying Pesin's entropy formula. Ann Math 122(2):509–539
48. Ledrappier F, Young LS (1985) The metric entropy of diffeomorphisms. II. Relations between entropy, exponents and dimension. Ann Math 122(2):540–574

49. Barreira L, Pesin Y, Schmeling J (1999) Dimension and product structure of hyperbolic measures. Ann Math 149(2):755–783
50. Szász D (2000) Boltzmann's ergodic hypothesis, a conjecture for centuries? Hard ball systems and the Lorentz gas, Encycl Math Sci, vol 101. Springer, Berlin, pp 421–448
51. Bunimovič LA (1974) The ergodic properties of certain billiards. (Russian) Funkcional. Anal i Priložen 8(3):73–74
52. Chernov N, Markarian R (2003) Introduction to the ergodic theory of chaotic billiards. Second edition. Publicações Matemáticas do IMPA. IMPA Mathematical Publications 24º Colóquio Brasileiro de Matemática. 26th Brazilian Mathematics Colloquium Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro
53. Araújo V, Pacifico MJ (2007) Three Dimensional Flows. Publicações Matemáticas do IMPA. [IMPA Mathematical Publications] 26º Colóquioquio Brasileiro de Matemática. [26th Brazilian Mathematics Colloquium] Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro
54. Katok A, Strelcyn JM, Ledrappier F, Przytycki F (1986) Invariant manifolds, entropy and billiards; smooth maps with singularities. Lecture Notes in Mathematics, 1222. Springer, Berlin
55. Jakobson MV (1981) Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. Comm Math Phys 81(1):39–88
56. Graczyk J, Świątek G (1997) Generic hyperbolicity in the logistic family. Ann Math 146(2):1–52
57. Lyubich M (1999) Feigenbaum–Coullet–Tresser universality and Milnor's hairiness conjecture. Ann Math 149(2):319–420
58. Benedicks M, Carleson L (1991) The dynamics of the Hénon map. Ann Math 133(2):73–169
59. Mora L. Viana M (1993) Abundance of strange attractors. Acta Math 171(1):1–71
60. Benedicks M, Young LS (1993) Sinaĭ–Bowen–Ruelle measures for certain Hénon maps. Invent Math 112(3):541–576
61. Benedicks M, Viana M (2001) Solution of the basin problem for Hénon-like attractors. Invent Math 143(2):375–434
62. Wang QD, Young LS (2008) Toward a theory of rank one attractors. Ann Math 167(2)
63. Viana M, Lutstsatto S (2003) Exclusions of parameter values in Hénon-type systems. (Russian) Uspekhi Mat Nauk 58 (2003), no. 6(354), 3–44; translation in Russian Math. Surveys 58(6):1053–1092
64. Bochi J, Viana M (2004) Lyapunov exponents: how frequently are dynamical systems hyperbolic? Modern dynamical systems and applications. Cambridge Univ Press, Cambridge, pp 271–297
65. Palis J (2005) A global perspective for non-conservative dynamics. Ann Inst H Poincaré Anal Non Linéaire 22(4):485–507

Social Cognitive Complexity

JÜRGEN KLÜVER
Duisburg-Essen University, Essen, Germany

Article Outline

Glossary
Definition of the Subject
Introduction
Social-Cognitive Complexity, Algorithms, and Bottom Up Models
Theoretical Foundations
A Universal Modeling Schema
Examples
Conclusions and Future Directions
Bibliography

Glossary

Attractor A sub space of the state space of a certain system that the system does not leave anymore. If the sub space consists of only one point it is a point attractor.

Dialectical relation A relation between two persons or generally two sides representing a tension and a mutual dependency. In more contemporary terms a dialectical relation is a certain form of feed back relation.

Homo oeconomicus An actor who acts according to the assumptions of rational choice.

Homo sociologicus The assumption that man is basically a being following social rules in contrast to *homo oeconomicus*.

Meta rules Rules that operate on rules of interaction and change these rules.

Rational choice The theoretical assumption that man is a rational and egoistical actor, who tries to maximize his profit and selects in a certain situation the best strategies.

Socialization The biographical process by which an individual learns and internalizes the cultural norms and social rules of his specific society. The result is a social actor.

Social-cognitive systems Complex dynamical systems that consist of social actors and contain at least two different levels, namely a social level and a cognitive one.

Socio-cultural evolution The development of societies in the mutually interdependent dimensions of social structure and culture.

Universal modeling schema A schema that consists of different interdependent levels and each level is modeled as a complex dynamical system.

Definition of the Subject

Social-cognitive systems certainly belong to the most complex systems that we know. The reason for this complexity is due to the fact that social-cognitive systems at least consist of two different levels, namely a social and a cognitive one that permanently interact and mutually influence each other. In terms of the social and cognitive sciences, social-

cognitive systems consist of social actors who interact and whose actions are determined by certain social rules. Yet these actors are also complex dynamical systems, i. e., cognitive ones. The cognitive processes of the individual actors also influence the specific actions. The social rules determine in some part the cognitive processes that in turn generate the according actions. But also in turn the cognitive processes may determine the actions in such a way that the social rules are influenced and even changed. That is for example the case in times of revolutions and/or political reforms. In this sense social-cognitive systems are determined by a permanent interdependent dynamics between these levels and that is why a modeling and precise analysis of these systems is rather difficult. In addition, social-cognitive systems are able to adapt to changing environmental conditions. This means that these systems are able to change their rules and structure.

Introduction

Social-cognitive complexity, although of course not in this term, has been the subject of social thinking since the beginning of reflections on society and social actions. Already one of the earliest works on society in human history, namely Plato's *Politeia* (The Republic), emphasizes the importance of the interdependency of mind and social structure: A certain desired social structure, i. e. the differentiation of society into different social classes, can be generated and reproduced only by influencing the minds of the social actors via education. A suited education cares for the adequate thinking of the individuals, which in turn guarantees the production and reproduction of the class structure by the social actions of the educated actors. The class structure finally determines the minds of the individuals and produces the wished forms of socially determined thinking. Plato's ideal society is certainly not a paradigm for democratic societies but the logic of Plato's arguments already demonstrates the necessity to take account of both social and cognitive levels if one wants to understand society and the socially determined actions of human beings. It is not by chance that all authors who thought about utopian societies also stressed the importance of education, i. e. the influencing of minds to produce and reproduce social structures, for example Huxley and Orwell.

More than two millennia later, Karl Marx in his theory of Historical Materialism [27] dealt in a similar way with the subject of social-cognitive complexity. He postulated a "dialectical" relation between consciousness or mind respectively and social structure – the social conditions (*soziale Verhältnisse*) in his terms. The social conditions determine the minds of social actors and produce that way an ideological consciousness; accordingly the social actors reproduce the social conditions by their ideologically determined actions. That is the essence of Marx's famous base-superstructure theorem: the social conditions determine the consciousness, which in turn produces the ideological parts of society (In the German original: *Das gesellschaftliche oder materielle Sein bestimmt das Bewusstsein*). Yet in times of societal crisis the consciousness of certain social actors, namely the members of revolutionary classes, the consciousness is able to overcome the ideological constraints and produces by the social actions of revolution a new social structure. In such times historical progress is generated by the determination of social structures by the social minds of the revolutionary actors. That is why Marx called the relationship between mind and social structure a dialectical one: both sides of the relationship form a dynamical tension, influence the other and are influenced in turn by the other. If one substitutes the philosophical term of dialectic by the more precise contemporary term of feed back, one may say that Marx postulated social-cognitive systems as systems whose dynamics is characterized by permanent feed back loops between a social and a cognitive level. To be sure, although Marx frequently stressed the importance of mathematically formulated theories he never could think of transforming the venerable concept of dialectic into a mathematical framework.

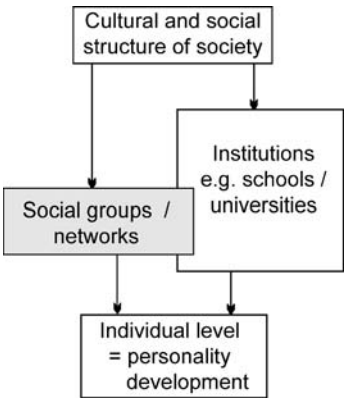
Marx had in principle captured the whole problem of social-cognitive complexity although only in an informal manner and formulated in classical philosophical terms. Yet his monumental approach was certainly not suited as a common foundation for the development of the social and cognitive sciences. Nearly at the same time the French sociologist Emile Durkheim, one of the founding fathers of modern sociology, postulated that social phenomena must only be explained by other *social* phenomena, which is one of the most important theoretical premises of the contemporary social sciences. As a consequence social scientists only deal with *social* complexity and consciously often neglect the influence of cognitive factors on social processes. Accordingly cognitive scientists only deal with *cognitive* complexity and neglect the impact of social factors on cognitive processes. Paradigmatic in this sense is the great work of Jean Piaget [34,35] who defined cognitive development as a principally autonomous process that can only be slowed down or accelerated by its social environment. With a famous expression of the German social theorist Niklas Luhmann one can say that the different disciplines of social-cognitive complexity reduced the complexity of their subject by concentrating on only one level of the two-level problem [26].

Matters became even more complicated when social theorists discovered that society is not a homogenous system but must itself be differentiated into different levels. The most basic differentiation is of course that of a micro-level and a macro-level [1,22]. In addition it is possible and sometimes necessary to introduce a meso-level, that is a middle range level of social reality. For example, if the micro-level is constituted by individual social actors or their actions respectively, if the macro-level is defined as a whole society or even an aggregation of societies like, e. g. the European Union, then a meso-level may consist of certain institutions, political parties and the like.

An additional complication and hence another increase of complexity must be accepted by taking account of the fundamental difference between social structure and culture. It is not possible for this article to mention all the numerous definitions of culture social theorists have developed in the last century. Following Habermas [15] and Giddens [12] it is sufficient to define the social structure of a society as the set of all social rules that are valid in a certain society and culture as the sum of all accepted knowledge, beliefs, world views, etc. in that society. By considering that fundamental distinction and by taking into account the different levels of society we apparently obtain a multi-dimension system, namely a system consisting of different levels that contain two dimensions – the socio-structural one and the cultural one. It is no wonder that social theorists concentrated on the social level of social-cognitive complexity, because this level is complex enough.

The main problem with such multi-level approaches is of course that the relations between the different levels and hence their interactions could usually be described only in an informal and frequently metaphorical manner. When a social theorist draws a graphical schema with different levels and additionally two dimensions the schema itself explains nothing but just states a problem. Somehow the levels interact and somehow individual actions are determined by the levels above, but in general this *somehow* cannot be precisely explained. To be sure, there are a lot of empirical studies that try, e. g., to analyze the impact of certain social structure on specific actions. But there is no general precise theory of links between the different levels of society [1].

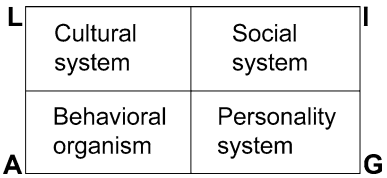
Yet despite these difficulties in several fields of the social and cognitive sciences not only the different levels of society but also the cognitive level of individual actors is taken into account. That is especially the case in socio-psychology and in particular in theories of socialization. The analysis of processes of socialization indeed demands that cognitive levels and social levels alike must be considered, if one understands socialization as the process by



Social Cognitive Complexity, Figure 1
A multi-level model of socialization

which the individual mind is formed according to the social structure and the cultural norms and values of a particular society. Although in the general models of processes of socialization the relation between the social levels and the cognitive one is expressed only one-sided, i. e., only the determination of the personality by social structure is taken into account in contrast to Marx, at least it is suited to speak of attempts to capture the whole social-cognitive complexity of social reality. Such a model of socialization is shown in Fig. 1; it is a shortened version of a more complex model of Geulen and Hurrelmann [11].

The well-known classical study of Berger and Luckmann "The Social Construction of Reality" is a paradigmatic example of such attempts [6] that in contrast to most one-sided models assumes a mutual interdependency of social and cognitive levels: According to them there is, in the tradition of Marx, a dialectical relation between the subjective social reality and the objective one. The subjective social reality consists of the internalized social rules and cultural values that constitute the personal identity, i. e. the individual personality. The objective social reality is generated by the according social actions and interactions of the socialized individual actors. Because the structure of this reality mirrors that of the individual personality the individual actions reproduce the objective reality, which was the determining factor of the constitution of the subjective reality. As in the case of Marx we have again a permanent feedback between the social and the cognitive level. Younger individuals participate in this process by taking over the structure of objective reality from their elders, by reproducing this structure in their own actions, and so forth. Although the theory of Berger and Luckmann does not deal with social progress, i. e., changes of social structure and culture, their model demonstrates how in principle it is possible to describe the interdependency



Social Cognitive Complexity, Figure 2
The general action system

of social and cognitive levels, even if only in an informal manner.

The most famous attempt to capture social-cognitive complexity in a whole model is certainly the well-known model of Parsons [8,29,32] of the differentiation of the action system. A simplified version is shown in Fig. 2.

The general assumption of Parsons is that each system is differentiated into four subsystems according to the famous AGIL-schema; the same logic of differentiation is valid for the subsystems and so forth. A means Adaptation, G stands for Goal Orientation, I is Integration, and L means Latent Pattern Maintenance. (The explanation of the AGIL-schema cannot be done in this article [29,33]). The picture shows that the social subsystems (social and cultural systems) constitute together with the cognitive subsystems (personality and behavioral system) the whole action system. The specific form of interaction between the four subsystems is called interpenetration, which roughly means that the four systems influence each other. For example, personalities are formed via socialization by the social and cultural subsystems, social and cultural systems are reproduced by personalities and according behavior and so on. Yet again we only have an informal theory, although the general logic of the AGIL-differentiation is certainly a great achievement.

To put it into a nutshell, the problem of dealing with social-cognitive complexity and not reducing the problem to one level is seen by many theorists. Yet the theoretical models are always just informal descriptions of the problem and programmatic foundations, if one looks at them from a precise, i. e. mathematical point of view.

Social-Cognitive Complexity, Algorithms, and Bottom Up Models

Despite many attempts to formulate the complexity of social-cognitive systems in a mathematical way for a long time it was not possible to transfer the mathematical tools of, e. g., mathematical physics to the problems of social reality. The use of statistical methods in empirical social research and experimental psychology is of course a well tested method since many decades. The problem was and frequently still is the mathematical formulation of *theoretical* models. Apparently the use of in particular differential equations is not suited for the problems one has to deal with in the social and cognitive realms. To be sure, simple cases can be analyzed in a mathematical way e .g. [5], but the complexity of social-cognitive systems in the sense of the great theorists mentioned in the preceding section could not be captured by models founded on the classical methods of the calculus.

The reasons for this are the peculiarities of social-cognitive systems. They are constituted, as was mentioned in the definition, by several levels that permanently interact and thus must be characterized by a multi-dimensional dynamics. In addition, social-cognitive systems are able to change their specific rules and structure by adapting to varying environmental conditions. It is of course possible to undertake a time series analysis in order to find some regularity in the historical development of such systems, but a time series analysis just describes a system's behavior in a phenomenological manner and gives no explanation. To make matters worse, even if one has found some regularity by a careful time series analysis the system may change its behavior via the varying of its rules; accordingly the time series will change and the analysis becomes worthless for the next time span. It is no wonder that more than once social and cognitive scientists alike believed that it would never be possible to develop a mathematical science of social and cognitive systems.

Fortunately things have changed by the development of new mathematical tools for the analysis of complex systems and in particular by the emergence of complex computer programs. These new mathematical modeling techniques, as cellular automata, Boolean networks, evolutionary algorithms, and artificial neural nets are on a first sight rather different methods but on a second sight they have much in common:

a) All modeling techniques are heuristically oriented to some natural processes as the reproductive capacity of living systems in the case of cellular automata and Boolean nets, the logic of biological evolution in the case of evolutionary algorithms, and the cognitive processes of the brain in the case of artificial neural nets. Sometimes these new mathematical tools are put together under the a bit misleading name of Soft Computing [48] but it would be more appropriate to call them nature oriented algorithms.
b) All these algorithms are very well suited for the construction of so-called bottom up models [24]. This term is used in contrast with top down models and means that the model construction is done by starting at the level of the elements of the system that shall be mod-

eled. Accordingly the units of the models represent the elements of the respective systems. The dynamics of the empirical system is represented by certain rules of interaction and the thus caused behavior of the whole model. The model's behavior is generated as an emergent consequence of the local interactions of the units of the model. This procedure allows constructing the respective model by *immediately* transferring the empirical observations of the system's elements and their rules of local interaction into the model. It is not necessary to make an abstract representation of the whole system by an aggregation of the elements.

For example, the modeling of a social group via a cellular automaton [21] can be done the following quite natural way: The cells of the cellular automaton represent the different individual members of the group and their specific rules of interaction determined by the group's social hierarchy are represented by the rules of transitions of the cellular automaton. It is then possible, at least in some cases, to predict the group's behavior by according simulation runs of the cellular automaton, as we did in several successful cases. The work of many researchers demonstrated that in particular cellular automata are very well suited for the analysis of social groups [13,16,25,30,39,41,43].

c) The capability of social-cognitive systems to adapt, i. e. to change their specific rules of interaction in order to fulfill certain environmental conditions can be modeled by the hybridization of the respective models, i. e., by the coupling of different modeling tools and such the constructing of a hybrid model [14]. In order to manage a successful adaptation adaptive systems must have not only rules of interaction but also meta rules, namely additional rules that operate on (*meta*) the rules of local interaction and change them according to the environmental demands. The best-known example for such adaptive capabilities is of course the changing of genomes by the genetic operators of mutation and recombination and by the force of natural selection in biological evolution. In this example the rules are those of individual epigenesis and ontogenesis; the meta rules are the genetic operators, steered by selection, that change the genome by variation.

A mathematical model of these processes of adaptation can be quite easily done by the construction of a hybrid system, consisting of several Boolean or logical nets respectively, representing the individual genotypes [18], and a genetic algorithm that operates on the rules of the Boolean nets and their specific topology.

Another example for adaptation is the case of political reforms. In nearly all advanced societies the *official* social rules of interaction are defined by laws and bureaucratic prescriptions. In parliamentary societies the meta rules are defined by the competence of the parliaments to change the laws and prescriptions by discussions and votes of majority. The meta rules then are given by the rules of procedure by which the parliaments have to operate. A variation of the rules of interaction by parliamentary decisions is done when the internal or external conditions of the society demand a changing of social structure. It can be demonstrated that democratic societies are more adaptive than dictatorships because dictatorial societies are both unwilling and unable to change their structure because of the resulting unrest and the danger for the ruling elite to lose its power [19]. An according model for such phases of political changing could, e. g. be constructed by the coupling of a cellular automaton (the model for the population) and an evolutionary strategy (the meta rules).

d) In contrast with a pure time series analysis bottom up models are able to explain, i. e. they can reduce observed phenomena to causes, namely to previous events that are causally linked to the observed event. In cases of complex systems such previous events can be either a previous state of the system. In that case the explanation consists of naming this particular previous state and the rules that generated the present state from the previous one. Or the previous event is an external disturbance of the system. Then the explanation consists of naming the disturbance and the reaction of the system's rules to the disturbance, i. e. to the externally changed state of the system. The third and most difficult case is that of an operation of the system's meta rules, their variation of the rules of interaction, and the generation of the present state via the changed rules of interaction. To be sure, the analysis and explanation of all different cases by the construction of an according valid model needs a lot of information about the empirical system. Yet a thorough explanation of such complex systems as socio-cognitive ones can only be done by the construction of models and simulation runs that use these new mathematical tools. To be sure, explanations of this kind can principally also be done by top down models and have been successfully done in physics and chemistry. Only the third case needs the distinction between rules and meta rules and hence demands models constructed the way sketched above.

Theoretical Foundations

The social sciences are, unfortunately, characterized by a multitude of different theoretical approaches that can-

not be enumerated in this article. Yet for attempts to construct mathematical models of social or even social-cognitive processes there are two main approaches, namely the Rational Choice approach (abbreviated RC) e .g. [17] and a social systems theory approach, founded on social rules and structure. Both are based on certain anthropological foundations about the nature of man or social actors respectively.

The RC-approach is based on the assumption that man is basically an egoistical being who always tries to maximize his own profit without caring for others. This assumption has a venerable tradition and it goes at least back to the Leviathan of Hobbes, namely the question how social order is possible in a world of egoists. The term "rational" means that a social actor tries to find the optimal strategies in each action situation, i. e., he analyzes, which action as means could serve him best for his maximum welfare as goal. Hence each social action situation is considered by an actor as a situation in a strategic game like chess or poker: he rationally analyzes the different possible strategies and chooses the best – therefore rational choice.

The RC-approach has a great advantage for the purpose of mathematical model construction in the social sciences, namely the famous Theory of Games by von Neumann and Morgenstern [45]. Game theory was originally developed for the mathematical analysis of economical decision problems and already Max Weber stated that economical actions are a paradigm of rational actions, oriented to the most favorable proportion between different means and the goal of profit maximizing e .g. [46] (Oscar Morgenstern was an economist). But soon the partisans of RC-approaches generalized this idea and tried to capture the logic of social action in general by game theoretical models.

Usually game theoretical models are based on a so-called pay-off matrix. Because most RC-theorists distinguish only between two forms of behavior, namely cooperation and defection, a pay-off matrix is a two-dimensional one. A characteristic a pay-off matrix that is frequently used for the analysis of social dilemmas, in particular the famous Prisoner's Dilemma (PD), is the following one:

	C	D
C	3	0
D	5	1

C of course stands for cooperation and D for defection. As usual the matrix-values show the result of the combination of the two different action strategies. For example, the combination DC = 5 means that a defective player D gets five points while the cooperative player gets zero points (CD = 0). Accordingly a cooperative player gets three points if the other is also cooperative; the other gets three points too (CC = 3). The absolute values of course are arbitrary; the only condition for strategic games of the PD-type is that DC > CC > DD > CD.

In a famous study about *The Evolution of Cooperation* based on a tournament of different computer programs, i. e. programs that simulated winning strategies, Axelrod [2] showed that the best strategy was *Tit for Tat* (TFT). TFT can roughly be translated as "I do the same to you as you to me". TFT is a very simple strategy because the first player *A* always starts with a cooperative action and waits for the response of the other player *B*. If *B* is also cooperative then *A* keeps his cooperative actions. If *B* is defective then *A* switches his strategy and becomes defective too as long as *B* acts in a defective manner. The first cooperative action of *B* then is answered by *A* with an also cooperative action and so forth. Axelrod used for his own analysis a cellular automaton and confirmed his results later by using a genetic algorithm [3].

TFT is basically a cooperative strategy because each player who acts according to TFT always takes account of the strategies of his opponent and becomes defective only if his opponent does the same. Hence Axelrod concluded that the emergence of cooperation in a world of rational egoists is not only possible but also even probable because it can be mathematically shown that a cooperative strategy like TFT is the most favorable one for an egoistic rational player. In this sense the old question of Hobbes could finally be answered by the methods of mathematical game theory.

Despite the fact that the results of Axelrod are not so generally valid as Axelrod believed [9,19,31] it is no wonder that many partisans of the RC-approach saw these results as confirmation for the RC-approach and for the necessity to study social dilemmas by game theoretical methods. Sometimes it seemed that the social sciences had to be reduced to game theoretical analysis of decision problems. Yet these studies neglect several important difficulties.

Some time ago Lord Dahrendorf [7] formulated the important distinction between *homo oeconomicus* and *homo sociologicus*. *Homo oeconomicus* is the rational and egoistical actor, i. e. the basis of RC. *Homo sociologicus* in contrast is the social actor who orientates his actions to social rules and norms. He may be or not be an egoistical being but the important point is that his actions must be understood as the consequences of certain social rules that are valid in a specific situation. *Homo oeconomicus* understands a social action situation as a decision problem for finding the optimal strategy; his rationality, therefore, is the calculation for the optimal proportion between

means and ends. The rationality of *homo sociologicus* on the other hand is an interpretative one: he must understand the situation in order to know, which rules determine the situation and which according actions he has to perform.

It is important to see that the question of Hobbes, which Axelrod and other partisans of RC tried to answer with a game theoretical analysis of social dilemmas, does not arise with the assumption of *homo sociologicus* as the basic nature of *homo sapiens*. Man is, according to this position, a social being by nature, i. e., his human nature is defined not only by characteristics like language or consciousness but also by his sociality. The basic question is then not how social order is possible at all – the question Hobbes answered by introducing the law enforcing state – but *how the construction* of social order is performed and how and why certain forms of social order were changed and others were not. Dahrendorf of course named this pair of concepts in this manner because he was convinced that *homo sociologicus* is the normal aspect of man as a social being; *homo oeconomicus* is a rather special case in situations where either no valid social rules exist, or where some actors deliberately decide *not to* follow the respective rules, or where social rules leave some space of freedom, which must be filled out by individual decisions.

It is not an arbitrary decision if one chooses the option for one anthropological option or the other. There are many reasons for the validity of the *homo sociologicus* position:

a) Studies in behavioral biology teach us that nearly all biological species consist of social beings, whether predators, prey, or plants; there are exceptions as, e. g., otters that live as loners but these are indeed exceptions. From anthropological studies it is to be learned that all our hominid ancestors lived in social communities and that socio-cultural evolution would not have been possible without human sociality [37]. Therefore, one may safely assume that human sociality is part of our biological heritage and not something humans had to create [47].
b) Social rules, and in particular social institutions, are important tools to reduce the complexity of the world and in particular that of the social world. Without accepted rules each new situation would be a new decision problem that has to be calculated anew. In contrast to that, social institutions offer a relief of the burden of new decisions because they enable actions by just following the rules [6,10,44]. That is particularly favorable if the rules have been proven in past situations. Hence, the evolution of social rules is more probable than Hobbes' famous *bellum omnium contra omnes* ("The war of all against all").
c) By following social rules it is possible to gain a social reputation as being reliable. Such a reputation frequently is favorable if others are undecided if they should trust an actor or not.
d) Finally, countless social studies have demonstrated that social actors mostly are rule-obeying beings. In a strict sense the social sciences could not have been as successful as they are if they would not have made the basic assumption of *homo sociologicus*.

These considerations do not mean that RC is wrong but that it is incomplete. There are certainly cases where social actors indeed act as egoistical rational players in a strategic game as for example in cases of social deviance or political revolutions. In particular, it is obviously a characteristic of modern democratic societies that the respective social rules nearly always leave a space of freedom for individual decisions. The social role of a teacher, for example, defines the necessity to act in a pedagogical way and to evaluate the efforts of the pupils. But it is an individual decision of the teacher if he mainly fosters the pupils or mainly acts as a critical evaluator of the pupils' performances. In this sense RC is an apt theory for some peculiarities of modern societies but it is not a general theory of human social action.

The consequence of these meta theoretical considerations is the following: The construction of mathematical models of social systems and of course of cognitive systems too must be founded on the basic logical unit of *rule*, either social rules or cognitive ones. To be sure, the elements of social systems are individual actors if one follows the bottom up modeling approach. Yet these actors are determined by the respective rules of their social systems; their actions, hence, must be understood as the result of the application of the rules valid in the respective situation. Certainly actors can be in error about the specific rules and/or they can decide to be socially deviant. In these cases it is particularly important to analyze the cognitive processes that lead the actor to social deviance. Yet the center of model construction and model interpretation must be the according rules.

We shall see in one example below that also in cases of social conformity the respective cognitive processes must be taken into account if one wants to understand how certain conformity emerged. Therefore, the interdependency of social and cognitive processes must be always be *principally* taken into account, although frequently it is possible to understand and even predict social behavior without considering the according cognitive processes.

A Universal Modeling Schema

The considerations of the preceding sections can be summarized in form of a universal modeling schema that captures the interdependency of different levels, the bottom up principle, and the logical dominance of the rule concept.

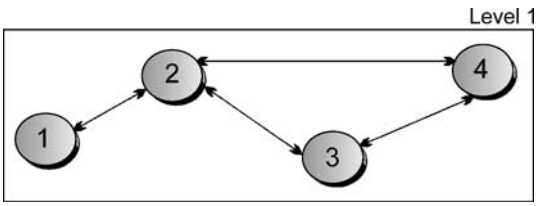
The basic *social* level of the schema consists of individual social actors who interact according to certain social rules. One may visualize this level, e. g., as the grid of a two-dimensional cellular automaton or the structure of a Boolean net. The level constitutes a dynamical complex system whose behavior depends, as usual, on the specific interaction rules, the initial states, and eventually operations of certain meta rules. Figure 3 shows such a level, which we may call level 1:

The concentration on this level and the neglect of a cognitive one apparently follows the mentioned methodical demand of Durkheim – social phenomena must only be explained by other social phenomena. Social actions are exclusively understood by the reconstruction of the respective rules that determined the actions. It is well known that many social processes can be understood and explained by taking into account only this level.

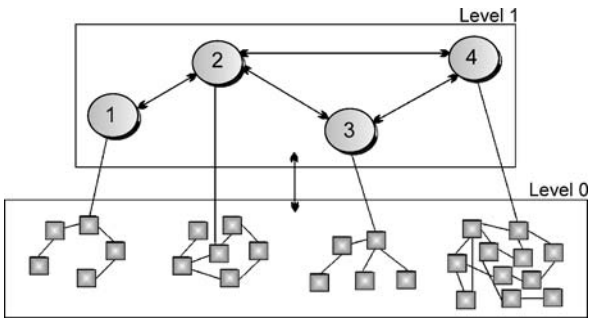
However, particularly in cases of social deviance, but not only there, the knowledge about the respective rules is not sufficient. Because social actors act with respect to certain rules *and* their particular thinking, worldviews etc. in many cases an according cognitive level must considered too. For example, the deeds of religious fanatics can only be understood if one reconstructs their fundamentalist worldviews and the according individual legitimization of criminal acts. In such cases social actors must also be considered as complex dynamical systems, i. e. cognitive ones with according cognitive rules and elements. The schema has to be extended as is shown in Fig. 4; the cognitive level may be called level 0.

Level 0 and level 1 already allow the modeling of many social processes, in particular those where an interplay of social and cognitive dynamics must be taken into account. Yet often even these two levels are not enough, as was mentioned in the introduction, because it is frequently necessary to model aggregations of individual actors. If, for example, one wants to analyze the political situation in a certain country with respect to the different political parties, then the concept of party of course means a certain institution consisting of specific rules and all party members. The interaction between such parties – coalitions, conflicts and so on – should be modeled on another level, namely level 2, the level of aggregations of actors (Fig. 5).

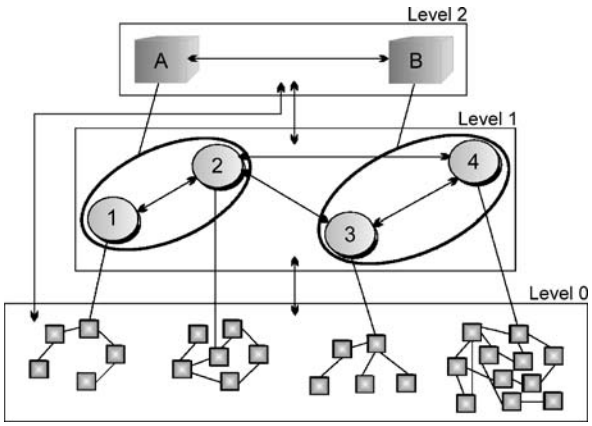
Note that level 2 must also be understood as a complex dynamical system with its own specific rules and elements. The elements there are usually called collective actors. All the three levels, hence, constitute certain complex systems whose dynamics is determined a) by the respective rules, elements, and initial states, b) by the eventual operation of meta rules that change the rules and in consequence the generation of states, and c) the interdependency between the different levels. This interdependency is visualized in



Social Cognitive Complexity, Figure 3
The social level as a complex dynamical system



Social Cognitive Complexity, Figure 4
Social actors modeled as cognitive complex systems



Social Cognitive Complexity, Figure 5
Social actors (level 1) modeled as cognitive complex systems (level 0) and as aggregations (level 2)

the pictures by the arrows that represent the mutual influences.

In contrast to many schemas of multi level models, e. g. the schema shown in the introduction, it must be noted that there is *in principle* a permanent feedback between all different levels. The action of individual actors (level 1) generate and reproduce the collective actors on level 2, yet collective actors in turn influence and determine the individual actors. A member of a political party acts *as* a party member. Accordingly the cognitive processes of individual actors determine their actions but these cognitive processes in turn are influenced by the result of the actions on the one hand and the collective actors the individual actor is a part of on the other. A party member not only acts but also thinks in concepts of the respective party. In addition, individual cognitive processes may directly influence collective actors, particularly in times of social and/or political change. History is full of the impact of individual cognitive processes on whole institutions, if the individual actors operate as reformers or revolutionaries.

It is certainly possible to extend even the schema shown in Fig. 5 by either adding a third social level, consisting of aggregations of collective actors, or by adding a second cognitive level by distinguishing between, e. g. conscious and unconscious thinking. Such extensions would lead to four- or even five-level models. Yet in most cases of social-cognitive modeling it is sufficient to operate with one, two or three levels. To be sure, it is always a question of the specific research interest how many levels and which ones one must take account of. The schema, after all, is just a *schema*, that is a frame work to guide the modeling processes when dealing with socio-cognitive processes – and not only these. The construction of a certain concrete model for a specific problem might be still a difficult task, in particular if one has to evaluate the validity of the specific model.

It is rather evident that this modeling schema can practically be applied to all complex social-cognitive processes. Yet the schema is in an even more ambitious sense a *universal* schema. This is a consequence from the mathematical characteristics of the algorithms mentioned in Sect. "Social-Cognitive Complexity, Algorithms, and Bottom Up Models". Imagine for example that level 1 is modeled by a cellular automaton, level 0 by one or several types of artificial neural networks, and level 3 again by a cellular automaton or a Boolean net. It is known from mathematical investigations that cellular automata are potentially equivalent to Universal Turing machines [36], which roughly means that no formal system is more general than such systems. This is the so-called Church–Turing thesis, which no mathematician seriously doubts. A corollary, the *physical* Church–Turing thesis, states that each physical, i. e. empirical system can be modeled via the use of a suited Universal Turing machine or an equivalent system. Although the physical Church–Turing thesis has not been proven in an exact mathematical manner nobody doubts this thesis either.

Without great difficulties it can be shown that each neural net and each Boolean net is equivalent to an according cellular automaton – the mappings between cellular automaton and the two classes of nets are injective. Hence principally each complex system can be modeled by using this schema *and* by using one or several of the mentioned algorithms. An important consequence from these considerations is the fact that a mathematical science of social-cognitive systems is not only possible but also that there is for each problem a general way how to treat it in a mathematical manner. The schema, therefore, can be understood as a constructive mapping from socio-cognitive phenomena to mathematical structures. The universality of the modeling schema is a constructive proof of the possibility of mathematical social and cognitive sciences.

Examples

These considerations show that and how mathematical models of social-cognitive processes can be constructed. Yet the most convincing proof is of course the demonstration of concrete examples, i. e., the models of specific complex social-cognitive systems. More detailed descriptions of these and other examples can be found in [20] and [21].

The Differentiating of Social Groups into Subgroups

A certain social group that consist of more than, say, three or four members is usually not a homogenous entity but is differentiated into several subgroups. The criteria for the generation of such subgroups are, of course, not always the same. In working teams for example the different qualifications of the members might be a criterion for the forming of specialized sub-teams. Yet in many informal groups a principle is certainly valid that may be called Homans' Principle [17] and that can be stated as follows: "Members of a social group favor interactions with people they like and they avoid interactions, if possible, with people they do not like." Everyday experience shows that this principle is indeed in many social groups an important criterion.

Since the foundation of sociometry by Moreno [28] the usage of a so-called sociomatrix or Moreno matrix respectively for the representation of social relations has become customary. If one wants to represent the mutual feelings

of the members of a certain group via a sociomatrix the matrix may be:

	a	*b*	*c*
a	0	1	−1
b	1	0	1
c	−1	0	0

The matrix is coded with three values. Value 1 means a positive feeling or a liking respectively, value 0 means neutral or indifferent feelings, and −1 means negative feelings or disliking respectively. Hence the matrix can be understood as member *a* likes *b* and dislikes *c*, member *b* likes *a* and *c*, and member *c* dislikes *a* and is indifferent to *b*. The usage of such a sociomatrix as a mirror of the emotional structure of a certain group can be very important for, e. g. teachers, group leaders, group therapists or trainers of sport teams.

According to the Homans principle our research group COBASC constructed a cellular automaton with the goal to simulate and predict the emerging of different subgroups in specific groups, which we called in honor to Moreno the Moreno-CA (COBASC is the abbreviation for *Computer Based Analysis of Social Complexity*). It is a two-dimensional cellular automaton whose cells represent the members of the particular group that shall be modeled. The state of the cells represent the emotional states of the group member that range from "very dissatisfied" to "very satisfied." The cells state is dependent of the eight adjacent cells (a "Moore neighborhood") in the following way: the values of the emotional relations of the center cell to the eight cells of the Moore neighborhood are summed up; the emotional state of the center cell then is the arithmetical mean of these values. (The cells are usually modeled as squares; hence there are eight adjacent cells). The values of the emotional relations are coded as 1, 0, or −1 as in the example above. The task of each cell is to find a Moore neighborhood where its own emotional state is a maximum. Accordingly the rules of the Moreno-CA are:

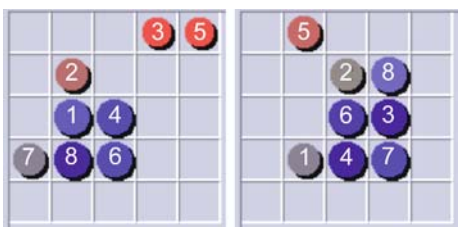
a) Place as many cells on the grid as there are group members; the rest of the grid's cells are free places.
b) Compute the emotional state for each "member" cell; vacant cells get the value 0.
c) Let each cell look on the grid for a Moore neighborhood where it can obtain a maximum state value.
d) Move each cell into that neighborhood.
e) If an attractor state, preferably a point attractor, has been reached, stop; if no attractor state could be reached after a certain number of runs, stop too. (A point attractor is a state the system will not leave although the rules are still operating; an attractor of longer period consists of several different states that the system will generate in cyclic order.)

The last rule is necessary because it can happen that the group reaches no attractor state at all or only those with rather long periods [21]. Factually the rules are a bit more complicated but this simplified version is the essence.

We validated this model with different social groups and obtained in all cases very satisfactory results. One empirical example shall serve to illustrate the methodical procedure:

With the permission of the teacher we asked a group of eight pupils of a primary school, with which pupils they would like to share a room in a youth hostel, with which pupils they would not share a room under any circumstances, and with which pupils they would share or not, according to the circumstances in the youth hostel. (It is often problematic to ask children (in this case 10 years old) directly if they like or dislike other children. That is why we chose this procedure of asking). These answers were transferred in the respective sociomatrix of the group. Afterwards the pupils were ordered to go into a strange classroom and choose places in order to wait for the teacher. The hypothesis was of course that the pupils would choose places according to their emotional relations to the other pupils. Then the values of the factual sociomatrix were inserted into the Moreno-CA; it reached a point attractor after 8 runs. The prediction of the program was then compared with the factual sitting order of the pupils (Fig. 6).

The program of course did not predict the absolute spatial distribution of the pupils in the classroom, for example if one pupil prefers to sit at a window. The Moreno-CA just predicts the relative positions of the pupils, i. e., which pupils are sitting together or which pupils are outsiders and are sitting alone. A comparison of the factual distribution and the prediction of the cellular automaton shows that the program indeed correctly prognosticated the sitting order for nearly all pupils with the exception of



Social Cognitive Complexity, Figure 6
The factual sitting order of the pupils (*left side*) and the prediction of the Moreno-CA

pupil 3. He is factually sitting besides the outsider pupil 5. However, the teacher of the children could give an explanation for this error of the program: Pupil 3 is known to be a practical jester who wanted to win the applause of the other children by teasing the outsider. Such melancholy character traits of course the program could not know.

We inserted the same sociomatrix into another program, namely a so-called self-organizing map or Kohonen feature map respectively (see next example). This is a special type of artificial neural nets. This net obtained even better results than the Moreno-CA because it placed pupil 3 beside pupil 5. An explanation for this result is given in [21].

To be sure, we did not want to predict just sitting orders of children in a classroom. The social experiment and the simulation by the two programs should show which subgroups existed in the whole group of children; this differentiation would become visible in the little experiment. The teacher, by the way, confirmed the results of the experiment and the prediction because according to her thorough knowledge the group of pupils factually consisted of an outsider and a homogenous group of all other pupils.

With respect to the modeling schema of the previous section this model uses only level 1, namely the social one. To be sure, in an implicit way, the emotional state of the children also plays a decisive role: there are no explicit social rules in this situation and the desire of the children to be in the neighborhood of their friends and to avoid other children is, according to Homans' principle, the main determination of the factual behavior. But in the model it was sufficient to represent the children as finite state automata, i. e., units whose state values are a direct consequence of their respective social milieu – in the model the Moore neighborhood. The Moreno-CA and the social experiment with the children demonstrate that indeed in many cases it is sufficient to construct models with only one level, in particular the social one.

Socialization by Learning from a Model

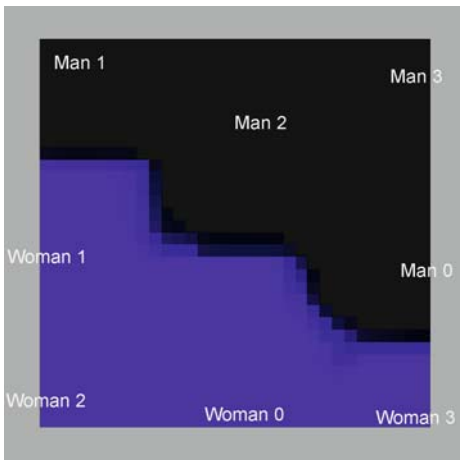
In Sect. "Introduction" it was demonstrated why theories of socialization always must take account of both social level(s) and the cognitive one. The social level represents the social milieu that is decisive for the specific socialization process; if one wants to explain the particular characteristics of this milieu one must consider additional social levels. The cognitive level of course represents the personality or identity of the individual that is formed or influenced by the respective milieu. The term formed though does not necessarily mean that socialization is a passive process. The socialized individual has the task to actively construct a worldview consisting of cultural values, social rules, and a culture specific knowledge.

One of the most known theories of socialization is the theory of learning from a model by Albert Bandura [4]. This theory states that socialization is to a high degree dependent on other people in the social milieu, who can act as models, i. e. as paradigms. The socialized individual perceives these models as positive or negative examples for its form of thinking and behavior. In other words, young people and children do not learn cultural values and specific forms of social behavior by internalizing general rules but by concrete examples they can observe and imitate. (There is a striking similarity between the socialization theory of Bandura and the cognitive theory of prototypes by Eleanor Rosch [23], which can only be mentioned here.)

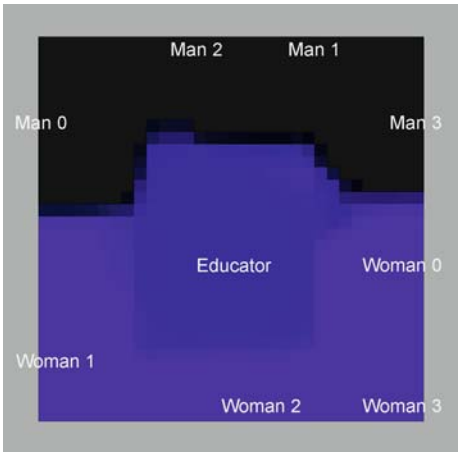
We applied this theory to a real case, namely a male youth from the *Ruhrgebiet*, an industrial area in the West of Germany. The youth, who shall be named Tom, of course a fictitious name, grew up in a workers family; he told us his story in an interview that was performed by one of our students. The father was an aggressive and violent man who often hit Tom's mother and frequently got involved in brawls and fights in the pubs. The mother was a submissive person who tried to keep up the family by working as a cleaner. Tom, therefore, had to care for his two younger sisters, which he very much disliked although he liked his sisters. When Tom was 10 years old his father left the family and disappeared. At the time of the interview Tom was 17 years old and had never heard any more from his father. Because Tom joined a gang of hooligans and got into trouble with the authorities and because the mother was unable to care for the two girls and Tom alike, Tom was admitted to a hostel of delinquent youths.

Despite the behavior of the father Tom adored him because *the father was a real man*. "Real" men are not afraid of other men, they become violent and hit if they feel insulted and in particular they dominate women, if necessary also by violence. "Real" women, on the other hand, are submissive and obedient to their men; they cook, wash and care for the children while real men work outside the house. As a consequence Tom developed a dichotomous worldview: men are strong and women are weak; women should obey and men command. The worlds of men and women must not be mixed in the sense that women take over tasks of men and men in turn care for the household. Consequently Tom despises the educators in the hostel because the female educators try to give orders to young men like Tom and the male educators perform household tasks like cleaning and cooking.

Tom apparently obtained his worldview by learning from the positive model of his father – "I want to be a man



Social Cognitive Complexity, Figure 7
The reconstruction of a dichotomous worldview: female versus male clusters



Social Cognitive Complexity, Figure 8
Increasing the female cluster by adding other "real women" and "false" men

like my father" – and the negative model of his mother. He also remarked that most men he met in his childhood were similar to his father and the women were similar to his mother. Hence he was early influenced by a homogenous milieu, where men and women all occupied the same gender specific roles.

For a simulation of Tom's biography with the result of a dichotomous worldview we used a Self-Organizing Map (SOM) or Kohonen Feature Map respectively that belongs to the class of neural networks, which learn in a non-supervised manner. In contrast to the forms of supervised learning non-supervised learning is not determined by externally given targets but the learning process is characterized by an internal logic. To put it into a nutshell, non-supervised learning is the ordering of a set of data according to certain internal criteria. There are different types of SOMs; we used a so-called Ritter–Kohonen model [38]. The technical details do not matter here; they can be looked up in any textbook on neural nets. The main intention of the simulation was to reconstruct the genesis of Tom's worldview and to demonstrate the according behavior in the present. Hence we inserted formal representations of the father and the mother as positive and negative models and inserted additional persons into the SOM. The SOM then constructed clusters of the different persons who are distinguished according to the criterion "real men" versus "real women". A first result is shown in Fig. 7.

The spatial nearness on the screen represents similarities or differences between the clustered persons respectively.

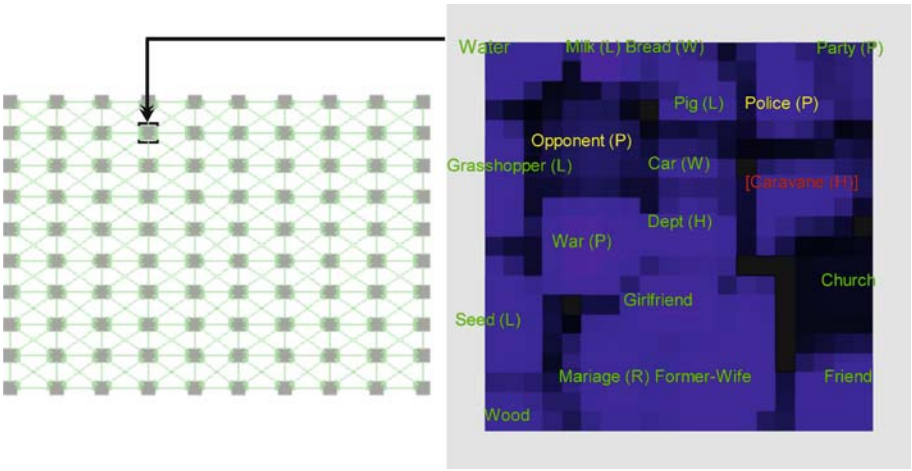
Afterwards we inserted other persons who differed from the first in certain personal respects. In particular we inserted formal representations of male educators who had female characteristics like cooking or cleaning. The SOM placed these men near the female cluster according to Tom's worldview: Tom does not accept these educators as real men but sees them more as women in male disguise. This is shown in Fig. 8.

The SOM apparently is able to simulate the important parts of Tom's socialization processes. As a SOM is a purely deterministic algorithm the successful simulation of Tom's cognitive genesis must be seen as a deterministic process too. Tom, so to speak, had no chance to become another person because the social circumstances determined his biography in a complete way.

The model uses only level 0, i. e., the cognitive one. Yet the social level is implicitly present too because of the inputs we inserted to the program; these inputs represent the social milieus in which Tom grew up and presently lives. In addition, the worldview of Tom has certainly social consequences not only for his biography. His membership in a hooligan gang, where he feels at home, indicates that his biography is representative for many youth members of such gangs. The according consequences for societal problems are obvious. In this respect the example of Tom also contains level 2, e. g. the level of collective actors like the different hooligan gangs and the social institutions that have to deal with them.

Modeling Social-Cultural Evolution

The last example is a rather complex model whose theoretical foundations and historical arguments can be read up in [20]. In contrast to the two preceding examples that be-



Social Cognitive Complexity, Figure 9
Cognitive state of an individual actor and his place in the artificial society

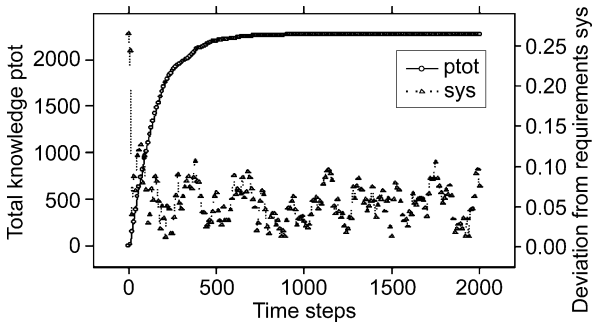
long to the field of micro-sociology this model deals with the evolution of whole societies. The basic assumption is that socio-cultural evolution is generated by creative inventions of individual actors and by the acceptance of these new ideas by the respective society. The social structure and the cultural norms of a society may promote or hinder the generation and acceptance of new ideas. In particular, the decisive factor is the degree of role autonomy for the occupants of creative roles like engineers, artisans, artists or scientists. If this degree is sufficiently high then the respective society evolves, i. e. increases its cultural capacity and generates social structures that are more efficient to deal with the problems of the society than old ones. If the degree is low, then the society stagnates, i. e. it becomes caught in a cultural and socio-structural attractor.

Comparative studies of the history of many societies always showed the same results, namely an early blossoming and cultural growth, then a period of consolidation, and finally stagnation or even regression, i. e. the return to some previous states [40,42]. According to our theoretical assumptions in these cases the degree of role autonomy was not high enough to allow an unhindered evolution. Only one case is known where apparently cultural evolution did not get caught in an attractor, namely the culture of European Modernity that was the historical basis for the contemporary Western culture. Historical studies confirm our hypothesis: In all cases of socio-cultural evolution the degree of role autonomy was much lower than in the European Modernity and the Western culture in general.

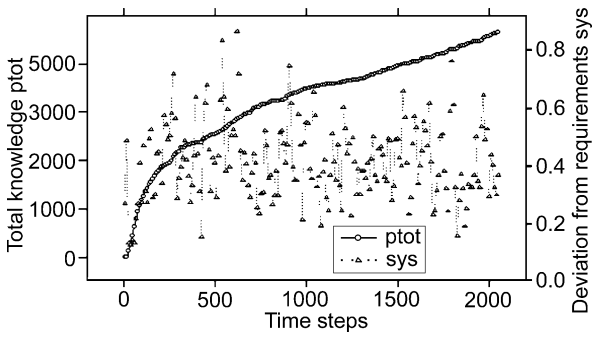
To validate these considerations by the use of a mathematical model we constructed the so-called socio-cultural-cognitive algorithm (SCCA). The model consists of a) a social level that is constructed as a cellular automaton. As usual the cells represent individual actors, i. e. members of a certain society. The rules of the cellular automaton determine the degree of role autonomy, i. e. the proportional ability of the actors to learn and create new ideas on the second level. b) The second level is a cognitive one: Each cell consists of a combination of different neural networks, namely several so-called bi-directional associative nets (BAM) and a SOM from the type used in example "Socialization by Learning from a Model". The task of the BAM is to associate observed characteristics from certain objects with the according semantic concepts – small, furry, and meowing is associated with cat. The task of the SOM is to order the concepts into a semantical network. These cognitive processes can be performed either in the way of learning, i. e. taking over new concepts from other actors, or by creating new concepts. The sum of the results of the individual cognitive processes is defined as the state or level that the respective culture has reached. Figure 9 shows an individual actor and its place on the grid of the cellular automaton.

The degree of role autonomy is inserted into the rules of the cellular automaton. We did many experiments with different values of this degree. The most important result is that in most cases the according culture indeed showed an evolutionary pattern known from human history, in particular described by the great British historian Arnold Toynbee (Fig. 10).

Only in the few cases with a high degree of role autonomy our model showed an evolutionary path known from the history of European Modernity. This is shown in Fig. 11.



Social Cognitive Complexity, Figure 10
A normal path (Toynbee path) of cultural evolution



Social Cognitive Complexity, Figure 11
A "Western development"

Societies with a sufficient high degree of role autonomy are not hindered in their cultural development although of course such societies may be stopped in their evolution by external forces. That could have been the reason for the final stagnation of the ancient Greek culture that also was characterized by a high degree of role autonomy for the occupants of the creative roles. Our model confirms that such societies are very seldom, which may be indeed the main reason for the fact, as Toynbee [42] observed, that the overwhelming majority of historically known societies stagnation was their inevitable fate.

The model is characterized by interplay of the two levels with consequences for a third one because the social structure determines the individual creative processes that in turn determine the fate of the whole society. In addition, sufficient developments on the cognitive level may also change the social structure: the better the cognitive improvements are, the higher becomes the degree of role autonomy on the social level. Such a rule is confirmed too by historical observations of the European history: In the Middle Ages the degree of role autonomy was not much higher than in contemporary cultures like feudal China or the Islamic societies. By the cultural progress during European Modernity the degree of role autonomy permanently increased. The model, hence, shows the consequences of such interdependent dynamics of the different levels that may explain important courses of history with respect to the fate of whole cultures.

Conclusions and Future Directions

The general possibility of a mathematically formulated science of social and cognitive processes was demonstrated in the Universal Modeling Schema and the fact that suited formal models like cellular automata or neural networks are equivalent to Universal Turing Machines. How many levels one needs for a concrete model, which ones, and how the relations between the levels are to be formulated is of course a question of the specific research problem. Constructing models of socio-cognitive processes will never be a trivial task. The modeling schema gives general orientations, but the detailed knowledge of the problem and in particular of the empirical data is always a necessary condition for the construction of social-cognitive models. That is of course a truism if it is stated in such a general manner but one that must always be taken into account. In particular, the knowledge about model construction and of suited algorithms never substitutes the profound knowledge about the specific fields of research.

The general relation between a social level and a cognitive one can also be stated in terms of individual actors. If one takes for example a most basic social situation, i. e. the communicative interaction between an actor A and another actor B, and if one assumes that during the interaction neither the social rules of this situation nor the cognitive processes of the actors change, then this interaction can be written the following way:

If f designates the set of social rules decisive for this situation and if g_A and g_B designate the cognitive processes of A and B respectively, then

$$f(A) = g_B(B)\,.$$

In other words, the application of the social rules f on the personality of A causes A to send an appropriate message to B, who in turn is moved to perform the cognitive processes g_B. In this case the social level determines the cognitive one, as is certainly the case in most forms of everyday communication. Accordingly the equations can be written if the cognitive processes determine the social rules, for example if a student becomes more intelligent than his professor by the teaching of the professor. (For general formulations of these equations see also [21]).

The most important problem for a socio-cognitive science that operates with mathematical models and accord-

ing computer simulations is without doubt the proper relation between theory, formal model and simulation program. A complete research program that has to take care of all three components follows the order: theory construction or theory selection respectively – transformation of the theory into a suited model – construction of a simulation program – comparing the results of the simulation runs with known empirical data. If everything went right, the last step will be a confirmation of the whole process. Yet since Murphy's law we know that each step may contain errors and if the data are not compatible with the simulation results then the error(s) may be in the theory, the model, the simulation program, or in several of the steps.

In particular the fundamental role of theory has frequently been neglected in the many attempts to construct a mathematical or computational respectively social science by using computer models of socio-cognitive phenomena. The mentioned dominance of RC-models on the foundations of Game Theory is just one indicator for this problem. Yet the social and cognitive sciences are much more than theories about rational and egoistical players in strategic games. Therefore, the most important problem for the future is the careful analysis of the research program formulated in the three or four steps mentioned above. Only then the use of mathematical models and suited simulation programs will be the methodical basis for a real socio-cognitive science.

Yet we have no alternative to the development of such sciences. If this research program is not successful we will always be the unconscious slaves of the social conditions we produced ourselves, in particular of a future we did not wish for. A science of the complex socio-cognitive reality is certainly not a sufficient condition for the solving of all our social problems. But it is a necessary condition for the task that we must understand our own species, that is to understand our own social and cognitive processes.

Bibliography

Primary Literature

1. Alexander JC, Giesen B, Münch R, Smelser NJ (eds) (1987) The Micro-Macro-Link. University of California Press, Berkeley
2. Axelrod R (1984) The Evolution of Cooperation. Basic Books, New York
3. Axelrod R (1987) The Evolution of Strategies in the Iterated Prisoner's Dilemma. In: Davis L (ed) Genetic Algorithms and Simulated Annealing. Morgan Kauffman, Los Altos
4. Bandura A (1986) Social Foundations of Thought and Action. A Social Cognitive Theory. Prentice Hall, Englewoods Cliff
5. Beltrami E (1993) Mathematical models in the social and biological sciences. Jones and Bartlett, London
6. Berger P, Luckmann T (1966) The Social Construction of Reality. Doubleday, New York
7. Dahrendorf Lord R (1958) Homo Sociologicus. Ein Versuch zur Geschichte. Bedeutung und Kritik der Kategorie der sozialen Rolle. Leske Buderich, Opladen
8. Fararo TJ (2001) Social Action Systems. Foundation and Synthesis in Sociological Theory. Praeger, London
9. Fogel DB (1993) Evolving Behaviors in the Iterated Prizoner's Dilemma. Evol Comput 1:77–97
10. Gehlen A (1956) Urmensch und Spätkultur. Frobenius, Bonn
11. Geulen D, Hurrelmann K (1982) Zur Programmatik einer umfassenden Sozialisationstheorie. In: Hurrelmann K, Ulich K (eds) Handbuch der Sozialisationsforschung. Beltz, Weinheim, pp 5–67
12. Giddens A (1984) The Constitution of Society. Outline of the Theory of Structuration. Polity Press, Cambridge
13. Gilbert N, Troitzsch KG (1999) Simulation for the Social Scientist. Open University Press, Buckingham
14. Goonatilake S, Khebbal S (eds) (1995) Intelligent Hybrid Systems. Wiley, London
15. Habermas J (1981) Theorie des kommunikativen Handelns, vol 2 (Theory of communicative action). Suhrkamp, Frankfurt
16. Hegselmann R, Mueller U, Troitzsch KG (eds) (1996) Modeling and Simulation in the Social Sciences from the Philosophy of Science Point of View. Kluwer, Dordrecht
17. Homans GC (1950) The Human Group. Harcourt, Brace, Jovanovich, New York
18. Kauffman S (1993) The Origins of Order. Oxford University Press, London
19. Klüver J (2000) The Dynamics and Evolution of Social Systems. Kluwer, Dordrecht
20. Klüver J (2002) An Essay Concerning Socio-Cultural Evolution. Theoretical Principles and Mathematical Models. Kluwer, Dordrecht
21. Klüver J, Klüver C (2007) On communication. An Interdisciplinary and Mathematical Approach. Springer, Dordrecht
22. Knorr-Cetina K, Cicourel AV (eds) (1981) Advances in Social Theory and Methodology. Toward an Integration of Micro- and Macro-Sociologies. Routledge and Kegan Paul, Boston
23. Lakoff G (1987) Women, Fire, and Dangerous Things. What Categories reveal about the Mind. The University of Chicago Press, Chicago
24. Langton C (1988) Preface. In: Langton C (ed) Artificial Life. Addison-Wesley, Redwood
25. Levy S (1993) Artificial Life. Penguin Books, London
26. Luhmann N (1984) Soziale Systeme. Suhrkamp, Frankfurt
27. Marx K, Engels F (1969) Deutsche Ideologie (German Ideology). MEW 3
28. Moreno JL (1953) Who shall survive? Foundations of sociometry, group psychotherapy and sociodrama. Beacon House Inc., Beacon
29. Münch R (1988) Theorie des Handelns. Suhrkamp, Frankfurt
30. Nowak A, Levenstein M (1996) Modeling Social Change with Cellular Automata. In: Hegselmann R, Trotzsch K, Muller U (eds) Computer simulations from the philosophy of science point of view. Kluwer, Dordrecht
31. Nowak MA, May RM (1993) The Spatial Dilemma of Evolution. Int J Bifurc Chaos 3:35–78
32. Parsons T (1968) The Structure of Social Action. Academic Press, New York
33. Parsons T, Platt GM (1973) The American University. Harvard University Press, Cambridge

34. Piaget J (1972) The Principles of Genetic Epistemology. Routledge, London
35. Piaget J (1972) The psychology of intelligence. Littlefield Adams, Totowa
36. Rasmussen S, Knudsen C, Feldberg R (1992) Dynamics of Programmable Matter. In: Langton C, Taylor C, Farmer JD, Rasmussen S (eds) Artificial Life II. Addison-Wesley, Redwood
37. Read DW (2005) Change in the Form of Evolution: Transition from Primate to Hominid Form of Social Organization. J Math Sociol, Special Issue Emergence Soc Order 25:3
38. Ritter H, Kohonen T (1989) Self-organizing Semantic Maps. Biol Cybern 61:241–254
39. Schelling TC (1971) Dynamical Models of Segregations. J Math Sociol 1:143–186
40. Spengler O (1926) The Decline of the West. Deutsches Original: (1922) Der Untergang des Abendlandes. Alfred Knopf, New York
41. Suleiman R, Trotzsch KG, Gilbert N (eds) (2000) Tools and Techniques for Social Science Simulation. Physica, Heidelberg
42. Toynbee A (1934–1961) A Study of History (12 vols). Oxford University Press, Oxford
43. Troitzsch KG, Mueller U, Gilbert N, Doran J (eds) (1996) Social Science Microsimulation. Springer, Heidelberg
44. Vanberg VJ (1994) Rules and choice in economics. Routledge, New York
45. von Neumann J, Morgenstern O (1944) Theory of Games and Economic Behavior. Princeton University Press, Princeton
46. Weber M (1981) Die protestantische Ethik. Mohr, Tübingen
47. Weingart P, Mitchell SD, Richerson PJ, Maasen S (eds) (1997) Human by Nature. Between Biology and the Social Sciences. Lawrence Erlbaum, Mahwah
48. Zadeh LA (1994) Fuzzy Logic, Neural Networks and Soft Computing. Comun of the ACM 37:77–84

Books and Reviews

Alexander JC (1983) Theoretical Logic in Sociology, vol 3. The Classical Attempt at Theoretical Synthesis: Max Weber. Melbourne-Henly, London
Brennan G, Buchanan M (1985) The Reason of Rules. Constitutional Political Economy. Cambridge University Press, Cambridge
Coleman JS (1974) Power and the Structure of Society. Norton, New York
Coleman JS (1990) Foundations of Social Theory. Cambridge, London
Diekmann A, Schmidt P (1998) Environmental Sociology, Special Issue of Rationality and Society. Sage, London
Ginsberg A (1990) Connecting Diversification to Performance. A Sociocogn Approach. The Acad Manag Rev 15:3:514–535
Hamilton D, Devine P, Ostrom T (eds) (1994) Social Cognition: Impact on Social Psychology. Academic Press, London
Howard J (1994) A social cognitive conception of social structure. Soc Psychol Q 57:210–27
Howard RW (1995) Learning and Memory. Major Ideas, Principles, Issues and Applications. Praeger, Westport
Parsons T (1951) The Social System. The Free Press of Glencoe, Glencoe
Resnick LB, Levine JM, Reasley SD (ed) (1991) Perspectives on socially shared cognition. American Psychological Association, Washington
Rogoff B (1990) Apprenticeship in thinking: Cognitive development in social context. Oxford University Press, New York
Schech S, Haggis J (2000) Culture and Development. A Critical Introduction. Blackwell Publishers, Malden
Singley MK, Anderson JR (eds) (1989) The transfer of cognitive skill. Harvard University Press, Cambridge
Snijders T, Snijders TAB, Bosker R (2000) Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. Sage Publications, London
Walsh DM (ed) (2001) Naturalism, Evolution and Mind. University Press, Cambridge
Wertsch JV (ed) (1985) Culture, communication and cognition: Vygotskian perspectives. Cambridge University Press, Cambridge
Winch C (1998) The Philosophy of Human Learning. Routledge, London

Social Coordination, from the Perspective of Coordination Dynamics

Olivier Oullier[1,2], James A. S. Kelso[2]
[1] Laboratoire de Neurobiologie Humaine (UMR 6149), Aix-Marseille Université, Marseille, France
[2] Human Brain and Behavior Laboratory, Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, USA

Article Outline

Glossary
Definition of the Subject
Introduction
Intentional Interpersonal Coordination
Issues in Quantifying Spontaneous Interpersonal Coordination
Human Spontaneous Synchronization
Shared Behavioral and Neural Social Coordination Dynamics
Social Memory and the Dependence on Initial Conditions
Future Directions
Acknowledgments
Bibliography

Glossary

Self-organization Self-organization lies behind all structure and pattern formation in nature's complex systems, including the human brain. Self-organization is a principle governing a system where no agent-like entity is ordering the elements, telling them where and what to do. In self-organizing systems, low-di-

mensional dynamics are revealed by changing one (or more) control parameter(s) whose role is simply to move the system through a series of state changes without prescribing its behavioral patterns.

Coordination dynamics Coordination dynamics seeks the laws, principles and mechanisms underlying the coordinated behavior of different kinds of components at multiple levels of description (molecules, cells, circuits, etc). It is an overarching conceptual framework that describes, explains and predicts how patterns of coordination form and change at multiple levels of brain and behavior. The brain, mind and behavior are linked by virtue of sharing a common underlying coordination dynamics.

Information exchange A remarkable fact is that in contrast to classical dynamics that deal with fundamental quantities such as mass, length and time and their relations, coordination dynamics is informational in nature, dealing with informational quantities of a relational kind that couple different parts of a system or different systems.

Phase transitions Phase transitions are the true illustration that a system is self-organizing. They are spontaneous qualitative pattern changes occurring as parameters are changed quantitatively. When they occur, abrupt switches from one coordinated pattern to another are observed and the dynamics of the entire self-organizing system is dominated by one or a few collective variables: the order parameters.

Stability Stability is a key concept in coordination dynamics. Here the stability is of coordination or collective variables. The (loss of) stability of a self-organizing system indicates whether a phase transition is to occur. In order to evaluate the stability of a system, one can perturb it and measure the time it takes for the system to return to its initial state, i. e. its *relaxation time*. A number of other converging measures have been used to measure stability in coordination dynamics such as *switching time* (the time it takes for the system to switch from one pattern to another when phase transitions occur) and *critical fluctuations* (the increase of variability of the collective variable in the vicinity of the phase transition).

Definition of the Subject

Social Coordination Dynamics (SCD) explores, at both behavioral and neural levels, the mechanisms mediating the formation and dissolution of bonds between individuals. SCD applies the concepts, methods and tools of informationally coupled self-organizing systems (coordination dynamics) to quantify real time social processes. Just as coordination dynamics deals with how the parts of complex systems work together in a meaningful way to achieve goals, so SCD aims to understand the interplay of forces operating at both individual and collective levels to produce effective social behavior. SCD offers a novel perspective and new metrics to explore systematically a fundamental form of human bonding (or lack thereof), and the self-organizing processes that underlie its persistence and change over space and time. SCD therefore complements recent developments in several fields such as sociology, social cognitive neuroscience, behavioral economics, game theory and neuroeconomics.

Introduction

Coordination can be broadly defined as a functional ordering among interacting components in space and time. Coming in many guises, coordination represents one of the most striking features of living organisms. The science of coordination, *Coordination Dynamics* (abbreviated CD) [36,37,38,41] stems from a complex systems framework based on the theory and methods of informationally coupled self-organizing dynamical systems (see ► Coordination Dynamics). CD explores a number of basic coordination phenomena that cut across a wide range of levels, creatures and functions. Of particular relevance to social coordination are: (i) patterned states of coordination remain stable in time despite perturbations; (ii) component parts and processes (dis)engage in a flexible fashion depending on functional demands and/or changes in environmental conditions; (iii) multiple coordination states exist rendering living things *multi-functional*, effectively satisfying the same (or different) set of circumstances; (iv) switching from partially to fully coordinated states and vice versa is commonplace; (v) selection of coordination patterns is tailored to suit the current needs of the organism; (vi) coordination patterns adapt to changing internal and external contingencies; (vii) depending on a balance between competitive and cooperative processes, learning may take the form of abrupt transitions from one coordinated pattern to another; and (viii) the system may remain in the current pattern of coordination even when conditions change thus exhibiting memory.

The foregoing list contains some of the core aspects of CD reflecting its inherently nonlinear and emergent character. Such phenomena appear so spontaneously and so consistently as to suggest the existence of an underlying lawfulness or regularity that transcends the multitude of differences between different systems and the settings in which they can be observed [41,45].

Coordination achieves its pinnacle in the vast array of cells and connections called the human brain, and in the collection of human beings we call society [41,45]. How social interactions form and change in complex systems and contexts is of great interest to many disciplines, particularly psychology, biology, physics, economics and the social sciences. The primary focus of the present article is to review recent work investigating the coordination dynamics of individuals interacting with each other in real time. At the core of all personal relationships is how the other becomes intertwined with the self. *Social coordination* is the tendency of two or more individuals to coordinate their ongoing actions with each other based on mutual information exchange. *Social Coordination Dynamics* (abbreviated SCD) is a theoretical-empirical framework that investigates the behavioral and neural dynamics of bond formation between individuals, operationalized in terms of how they spontaneously synchronize their behavioral and neural patterns [82,99].

Synchronization is a form of spontaneous pattern formation that operates according to general principles of self-organization described by nonlinear dynamics [25,26,72]. Following on Huygens's analysis of two clocks synchronizing on a wall, many studies have framed the problem of mutual synchronization in terms of a network of oscillators each of whose individual behavior is altered by nearest neighbor interaction [5,7,30,57,106,107]. Under that framework synchronization has been observed among very different entities in a broad range of physical, biological and social systems. Human brains (and behavior) have proven no exception to these principles [19,27,41,50,51,89]. Experiments have revealed that humans exchange information – whether uni- or multimodal in nature – to spontaneously adopt and switch coordination patterns (e. g. [37,53,58]).

The validity of the measures and constructs from coordination dynamics are worth mentioning because they speak to the appropriateness of a dynamical framework for investigating social situations. Whereas it is easy to justify the physical existence of linkages between components in the coordinated behavior of a single entity, no such linkage typically exists between people. *Social coordination occurs via information exchange*, typically through vision, touch and sound. Emotional interactions may also be involved. A natural measure that describes this informational exchange is the relative phase between coordinating behaviors. The relative phase is an informational variable whose dynamics and has been shown to capture quantitatively coordinated patterns of brain and behavior among different kinds of components, events and processes (see ► Coordination Dynamics [26,40,41]). For coupled rhythmic behaviors, the relative phase dynamics is often adequate not only for uncovering basic mechanisms underlying synergy formation and behavioral change but also the strength and directionality of influences during social interaction ([82,99]; see also [62,87]) for recent reviews.

Intentional Interpersonal Coordination

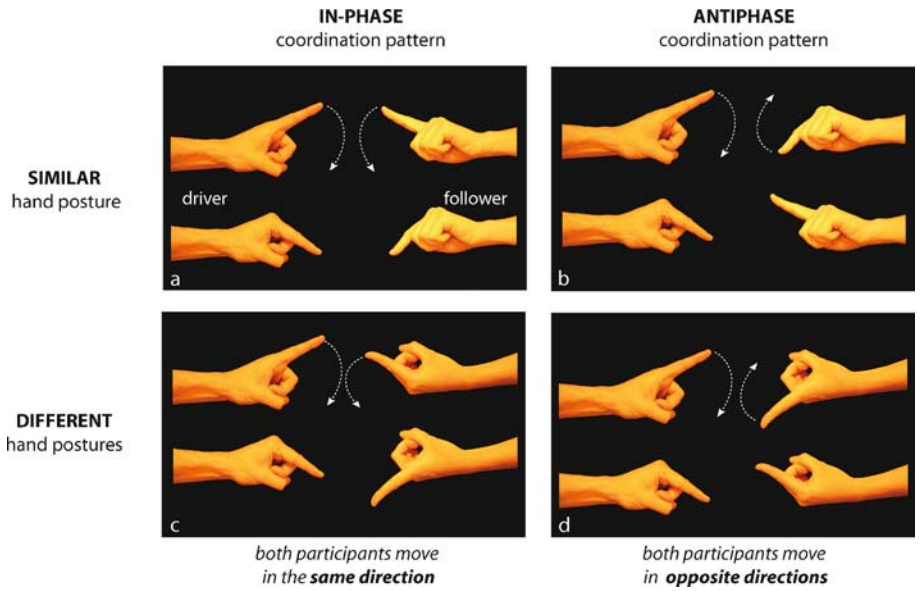
Among the many phenomena of human social coordination, one that most of us have experienced is the synchronized clapping of an audience. Néda and colleagues [70,71] have investigated why applause occurs in unison, with individual "clappers" sometimes acting as a single synchronized ensemble. Although synchronized clapping may vary little from one situation to another, the mechanisms governing the phenomenon are nuanced and context-dependent, even within the same audience. An illustration of this context-dependence comes from the world-famous *New Year's Concert* given every year by the Vienna Philarmonic Orchestra in Austria. Traditionally the concert ends with the Radetzky March by Johann Strauss Jr. This piece of classical music is performed in quite an unusual way. For instance, the conductor leads not only the orchestra but also the audience. Upon a visual cue from the maestro, the audience claps in synchrony with the music. The collective clapping is synchronized both with the music and the visual signals given by the conductor. The reader who is not really into classical music might prefer the example of the song *'We will rock you'* by Queen. Except for a final and unique guitar solo, this song is constituted by a powerful rhythm and a poignant vocal performance by lead singer Freddie Mercury. When this very rhythmic song was performed live, the audience intentionally coordinated its movements with the sound of the drums and the pattern of movements visually provided by the singer. People were therefore intentionally clapping their hands on the first two beats and extending their arms on the third. In the coordination dynamics literature, this is referred to as *intentional sensorimotor coordination* of individuals with external events.

Several studies have employed the sensorimotor coordination tasks to investigate interpersonal coordination dynamics for the case when a person intentionally synchronizes her movements with another by means of visual information exchange see [62] for a review. Following the bimanual and sensorimotor paradigms introduced by Kelso and colleagues [36,37,46,48], Schmidt, Carello and Turvey [88] asked two individuals sitting next to each other to swing their legs in an in-phase or antiphase fashion with respect to the leg movements of the other member of the dyad. As movement frequency was increased

(or decreased) by means of an auditory metronome, they found many of the predicted features of nonequilibrium phase transitions [27,47,89]: (i) *differential stability* between the two coordination patterns; (ii) *phase transitions* from the less stable coordination pattern (antiphase) to the more stable one (in-phase); (iii) *critical fluctuations* (i. e. increase in coordination variability) in the vicinity of the transition region and (iv) *hysteresis* (i. e. a sensitivity to the history of the system). All such hallmarks of coordination dynamics (and others such as *critical slowing down* in the vicinity of transitions) have been repeatedly found in a huge number of studies covering various experimental settings. Those settings include, but are not restricted to inter- (e. g. [36,37]) and intra-limb coordination (e. g. [9]) and coordination beween a limb and its uni- (e. g. [3,9,48]) or multi-modal environment [58] to name only a few. The main contribution of Schmidt and colleagues' research was to demonstrate that coordination phenomena found within a person's brain or body, extend to the interactions between people. It is noteworthy that the observed effects extend outside the typical laboratory setting to include coordination phenomena between an individual and an animal as in Lagarde and colleagues' investigation of the coordination dynamics of the horse–rider system [59]. In this unique experiment horses were ridden while walking, trotting and running on a treadmill. The movement dynamics of the horse, the rider and the horse–rider pair were recorded and analyzed revealing that the human–animal dyad exhibits similar coordination dynamics to human interpersonal coordination [59]. In this respect, it cannot be overemphasized that coordination dynamics deals with emergent cooperative effects across very different coordinating elements from neurons to muscles to limbs to people and across the animal–environment divide ([41,42]; for an excellent discussion, see Turvey [100] and commentaries in Vallacher and Nowak [101]). Both the 'intrinsic dynamics' of the individual elements and the nature of the coupling between different elements must be identified for a full account of the phenomena observed.

Several experiments by Schmidt and co-workers, as well as by other groups, have explored the effects on interpersonal coordination of variables such as the manipulation of objects (e. g. hand-held pendulums) or visual surroundings ([74,86]; see [87] for a recent review). Incorporating both aspects de Rugy and colleagues developed a neuro-mechanical model of visually mediated intentional interpersonal coordination [16]. Their model consists of two cross-coupled neuro-mechanical units, each composed of a neural oscillator driving a wrist-pendulum system moved by a different person. Taken individually, each unit reproduces the natural tendency of the participants to freely oscillate close to resonance frequency. When cross-coupled through the vision of movements of the other individual, each person entrains the other as they adopt a common frequency influenced by their own mechanical properties. Although important, neuromechanical properties are not the only factors that determine the stability of coordination patterns between individuals: attentional load and egocentric constraints also influence interpersonal coordination dynamics [96,97].

A series of experiments has investigated whether the motoric and perceptual constraints that shape the dynamics of inter- and intra-limb coordination play a similar role in the coordination between people (e. g. [12,13,53,58,66,78]). In intrapersonal bimanual coordination the preference for co-activation of homologous muscles appears to be mediated by general principles of symmetry in neural organization such as reciprocal connectivity between homologous brain areas. In a study by Oullier and colleagues [76] investigating the relative role of visual/directional and motor (a) symmetries in interpersonal coordination, two participants made index finger flexions while seated facing each other. One acted as a driver (D) by synchronizing to a metronome that systematically increased in rate. The second participant, or follower (F), was required to coordinate finger movements with D via visual coupling only. F participated in four conditions (Fig. 1) determined by a combination of coordination pattern (in-phase or antiphase) and hand posture (supination or pronation). The relative phase requirement was defined by the spatial configuration (i. e. the position of the endpoint of the finger). In this way, co-activation of homologous muscles (finger flexion by F and D) produced both an in-phase and antiphase relationship between the effector endpoints depending on the experimental condition. If purely directional constraints [92] determine the stability of interpersonal coordination, and D functions only as a generic rhythmic stimulus, perceptual antiphase coordination should display decreased stability regardless of the relative hand position of the participants. Contrary to this hypothesis, a strong role was found for interpersonal homologous muscle co-activation. Coordination between individuals was most stable when they were activating similar muscle groups such that co-flexion was always more stable regardless of the resulting spatial pattern. Directional constraints played only a modulatory role. These initial results are at odds with the concept of social coordination as a form of simple perceptual-motor coupling. Rather, it appears that perception of homologous muscular activation acts as a constraint on coordinative stability, creating a "functional homology" to bimanual coordination. Thus, social coordination may be differentiated from



Social Coordination, from the Perspective of Coordination Dynamics, Figure 1

Participants show a preference for homologous muscular activation, irrespective of the visuospatial congruency of their movement. Each column describes a possible configuration of interpersonal coordination. The left participant is paced with a metronome whose frequency increases (driver). Oullier and colleagues [76] studied the frequency at which the right subject (follower) loses stability in each condition. In columns a and b, both subjects are in the same hand position. The pattern a (both flex then both extend: in-phase coordination) is more stable than the pattern b (when one extends the other flexes: anti-phase coordination). In columns c and d, the participants adopt a different hand posture. The pattern d (both flex then both extend: anti-phase coordination) is more stable than the pattern c (where both subjects move in the same direction). These results suggest that coordinative stability is not purely governed by visuospatial congruency (cf. [67]). Rather, the embodiment of the other's movement leads the follower to adopt an anatomically homologous movement. This pattern of behavior is unique to the fact that the follower and the entity with which he/she coordinates are both humans [76]

simple perceptual-motor coupling by virtue of the biological and functional relevance provided when viewing another person.

Although all these studies have employed comparable experimental settings and the common theoretical framework of coordination dynamics with the aim of better understanding intentional interpersonal coordination, it is not yet clear whether spontaneous mutual entrainment actually occurs in a true two-way interaction, or whether one individual simply acts as a pacing stimulus or 'driver' for the other (e. g. [48]). A similar concern can be raised regarding the behavior of the audience during the Radetzky March at the New Year's Concert in Vienna. It seems unlikely that audience members spontaneously synchronized with each other while music was played, since their primary intent was to respond by clapping in rhythm with the music and with the visual cues coming from the stage. This process has been well described in human movement (neuro)science and coordination dynamics and occurs when an individual intentionally coordinates his movements with external physical stimuli [3,41,48]. A sensorimotor interpretation of audience participation is strengthened by results of an experiment in which the auditory metronome used to pace the interpersonal coordination was silenced at times [76]. The study revealed that the presence of an external pacing stimulus (an auditory metronome in that case) actually reduced interpersonal coordinative stability regardless of the adopted directional or muscular pattern adopted. Oullier and colleagues [76] provided evidence for stronger mutual entrainment when no external information could perturb the dyadic interactions, analogous to what the music and the conductor would do during a concert. Hence, from an experimental perspective, this phenomenon is not social interaction per se but rather sensorimotor coordination to an external event. In the case of the "observed" audience clapping in unison *during* the performance, one could argue that the audience is constituted by a collection of individuals coordinating mainly with the music and the conductor with little contribution from neighbor-to-neighbor interactions, (A similar concern can be raised in the study by Schmidt and co-workers [88] as partic-

ipants were instructed to intentionally coordinate with each other *and* with an auditory metronome. Hence, one participant could serve as a visual metronome to the other (and reciprocally) and/or the phase transitions observed could either be interpersonal in nature (from interpersonal antiphase to in-phase) or from syncopation to synchronization as in a single individual coordinating with an auditory metronome (cf. [48]).

A different scenario, however, is characteristic of the end of the performance, when the audience expresses its approval of the orchestra and conductor through applause. At this moment each person applauds according to her preferred/intrinsic pace with no driving stimuli – whether visual or auditory – coming from the stage. In spite of the absence of pacing information, the audience quickly and spontaneously entrains to a common rhythm such that everyone is clapping in unison. Note that at this moment, the only information that can alter an individual's behavior is the sound (and possibly the vision) of the movements made by their neighbors [70,71]. Thus, we have units involved in individual rhythmic behaviors communicating via, at least, one means of information exchange. According to Winfree [108], this is a minimum requirement for self-organized spontaneous synchronization to emerge (see also [41]). In that case, any collective pattern that emerges is more likely to be unintentional compared to situations where the audience follows the conductor and the music.

Issues in Quantifying Spontaneous Interpersonal Coordination

An abundant literature exists addressing unintentional interpersonal coordination in experimental paradigms ranging from people swinging pendulums [86], dancing [33], walking [102] or rocking chairs [84] to performing joint Fitts' tasks [69], talking to each other [83,91] or even boxing [60]. However, many questions remain regarding the nature of the behavioral and neural processes mediating the formation and dissolution of unintended synchronous behavior between individuals and how such processes may be quantified [2,55].

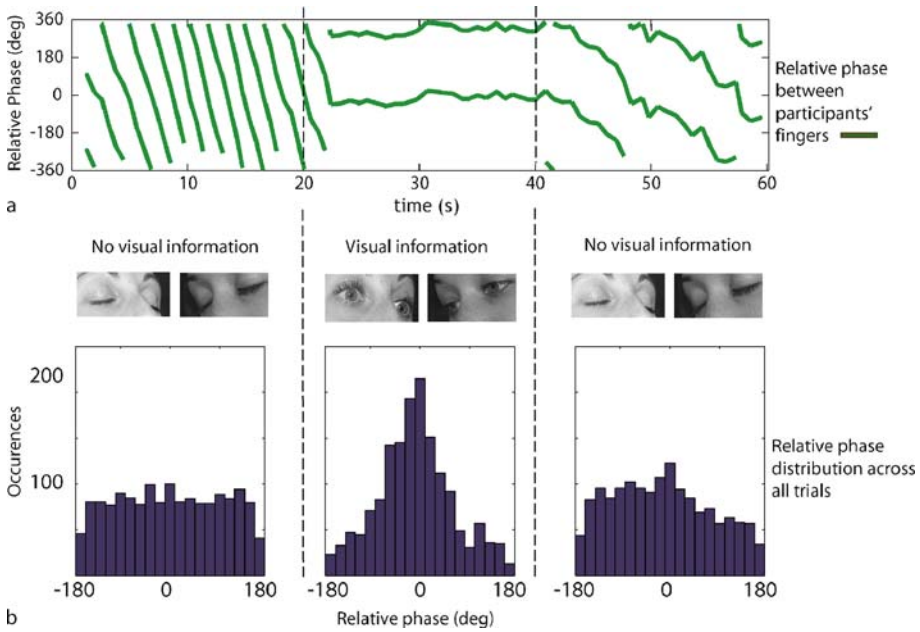
Oullier and colleagues [82] have identified three major problems in investigating spontaneous synchronization in social settings. First, even when the source and nature of the coupling has been identified, it is difficult to manipulate experimentally relevant variables such as the coupling strength (e. g. [71]). Almost by definition, spontaneous behavior is not externally goal directed or explicitly controlled. Most of the results reporting unintentional synchronization in humans are based on observation and categorization methods that rely primarily on the experimenter's appreciation of a given exemplar behavior rather than a quantitative measure of coupling and individual behavior (e. g. [4,14]).

A second problem is the challenge of complexity, both in terms of the large number of units to analyze (e. g. thousands of pairs of clapping hands [71]) and the complexity of the behavior itself (e. g. mother-infant synchronization [14]). Such compositional and behavioral complexity has hindered experimental attempts to record and quantify both the individual and social dynamics. Even the reduction in dimensional complexity afforded in coordinated behavior can only go so far in elucidating the relationship between group behavior and the individual units of which it is composed.

A third problem comes from the possibility that any change in a person's behavior induced by interacting with another may persist even after the encounter is over. So far, there has been very little precise quantification of the mutual influence people have on each other's behavior a posteriori, i. e. how individual behavior is affected after the social encounter when people no longer exchange information (but see Sect. "Social Memory and the Dependence on Initial Conditions").

Human Spontaneous Synchronization

In behavioral experiments that revealed spontaneous interpersonal synchronization, Oullier and colleagues [77,80,82] explored coordinative patterns that emerge only as a function of visual information exchange. The main hypothesis was that even without instructions to do so, spontaneous synchronization between partners would occur as soon as they coupled visually while moving in front of each other. On the other hand, spontaneous interpersonal coordination should disappear whenever exchange of information is no longer possible. In Oullier et al.'s behavioral experiments, pairs of participants executed movements while in full view (or not) of each other's ongoing actions as well as their own [77,80,82]. Each member of the dyad executed movements at their own preferred frequency and amplitude without any external pacing from a metronome or any other sort. Movements were required to be as smooth and continuous as possible throughout an experimental trial. What is important here is that participants were not given any instructions regarding the way to move with respect to each other. The experimental protocol consisted of participants moving with no vision of the other's movements before being allowed to see their own actions at the same time as they saw the other person's. Finally, visual information was removed again. Experi-



Social Coordination, from the Perspective of Coordination Dynamics, Figure 2

Relative phase between the movements of two individuals. a This panel illustrates the evolution of the relative phase as a function of time in a representative trial of the SCD paradigm. *Left column* No visual information exchange, each individual moves independently, movements are uncoordinated. *Middle column* as soon as people exchange visual information, they spontaneously couple. Their relative phase is therefore close to 0° . *Right column* When visual information is removed, they are no longer synchronized. b This panel represents distributions of relative phase for all the subjects and all the trials (adapted from [82])

mental trials were therefore equally partitioned into three contiguous segments each of equal duration within which both subjects either were allowed to exchange information with each other or not. When visual information was available, participants looked at each other's finger motion and were also able to see their own finger [77,80,82].

In SCD, following theories of cooperative phenomena in open systems [25,26] a central idea is that the behavior of a complex dyadic system may be captured by the value of a low-dimensional collective variable known as the *order parameter*. In the vicinity of critical points, emergent behavior is governed by the dynamics of this collective variable e. g. [25,41]. In experimental cases the order parameters are not known in advance but have to be discovered. For the situation of social coordination as in many other cases treated by CD, an appropriate order parameter describing the system dynamics is the relative phase ϕ between the movements of each member of the pair [80,82]. The relative phase measure allows for a reduction of a potentially very high dimensional system (e. g. where one has to consider, among other components, the neurons, joints and muscles of both individuals) as it captures the macroscopic spatio-temporal behavioral pattern (see Fig. 2). Even at an overt behavioral level, four degrees of freedom (position and velocity of each component) may be compressed onto a single relative phase value that summarizes the organization of the dyadic system. Quantitative evaluation of spontaneous synchrony is also provided by the FFT power spectrum overlap between the movements of each person. The spectrum overlap measures the percentage of movement frequencies common to both partners in a pair [82]. Defined as the area of intersection between each participant's normalized spectral plots, it serves an indicator of the strength of the frequency entrainment between the two participants (see Fig. 4).

When no visual exchange was allowed, each subject produced movements independently at their own frequency. As a result, the relative phase ϕ between the subjects' finger motions exhibited phase wrapping (Fig. 2, left column). However, following a simple auditory cue to open their eyes, subjects spontaneously adopted in-phase motion, ϕ stabilizing around 0° (Fig. 2, middle column). On a signal to close the eyes again, the individual movement frequencies diverged and ϕ fell back into phase wrapping (Fig. 2, right column). These initial results were corroborated by a subsequent more extensive study, in which the order of the vision and no-vision segments was changed. Once again, spontaneous synchro-

nization emerged as soon as vision of the other's movements was allowed [82]. Overall, results reveal that with visual information exchange, participants tend to mutually couple at a common phase and frequency, whereas in the absence of vision, participants' movement trajectories diverge and behave independently. Such emergent mutual coupling is truly a result of spontaneous social interaction and may be distinguished from previous dyadic studies in which one person may simply be intentionally tracking (or driving) the other [16,77,88,97] or maintaining their own rhythm [86].

Why does spontaneous interpersonal coordination occur at all? Compelling examples stretching from human evolution through religious ritual and sports to political, war and economical strategy suggest that *keeping together in time* is one of the most powerful ways to create and sustain communities and communication [65]. Moreover, *not* moving in synchrony may be too costly for the dyad see, (e. g., [56]).

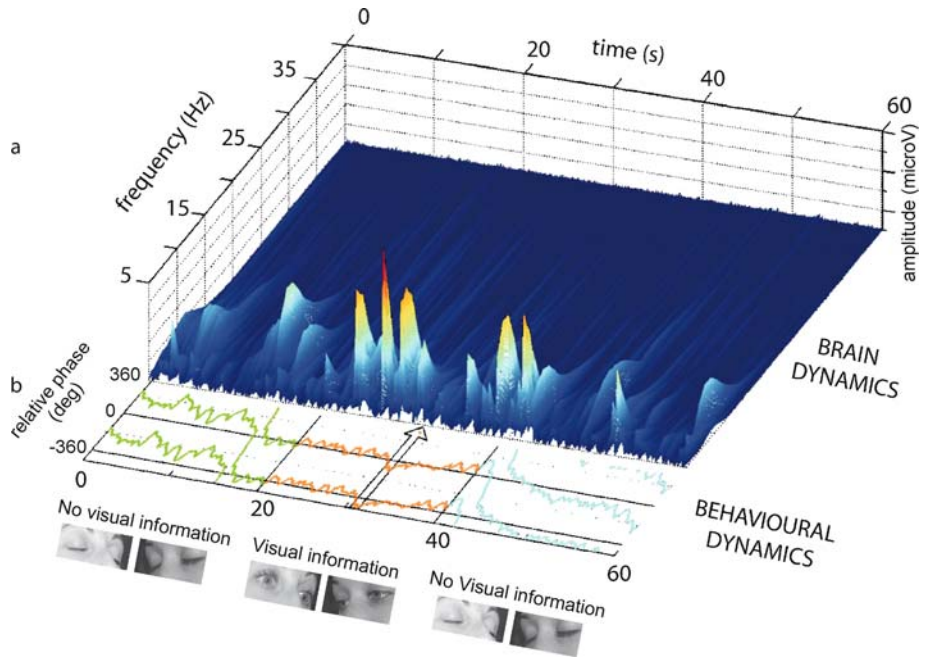
In order to better understand which features of visual information exchange may facilitate spontaneous social coordination one has to bear in mind that human movements can be unintentionally affected by the vision of an object oscillating in the environment. This is illustrated by experiments using the moving-room paradigm in which the walls of the room move but not the floor (e. g. [61,75,78]). Body sway of the observer's couples in time spontaneously with small oscillatory motions of the room. In addition, experimental data show that the mere observation of the movements of another person interferes with one's execution of a similar action [54]. Interestingly, such interference is less noticeable when the movements observed are not generated by humans [15]. In the latter work, one of the members of the dyad was replaced by a computer-generated moving hand, the trajectory of which was driven either by a sinusoidal function or a pre-recorded real finger trajectory. The stimulus movement frequency in the study by de Guzman and coworkers [15] was fixed at either 10% below or 10% above the subject's self-paced rate as determined at the start of the experiment. Results revealed that the human–avatar coordination was strongest when the latter was an image of a hand driven by real movement data. The weakest coupling occurred when the visual stimulus followed a sinusoidal trajectory. Unlike the interpersonal situation [82], spontaneous synchronization was not found for all trials and, when it happened, was supported by a significantly lower frequency overlap [15]. One may invoke a one-way coupling to explain these findings, since the motion of the computer generated hand could not be influenced by the movement of the participant. Taken together, the foregoing results support the hypothesis that biological relevance in general, and biological motion in particular – including its natural variation– play a key role in social coordination.

Shared Behavioral and Neural Social Coordination Dynamics

One explanation for the emergence of spontaneous social coordination may be found at the neurophysiological level. For instance, some areas of the brain are known to be associated with the perception (but not the execution) of biological motion including the posterior superior temporal sulcus (abbreviated STS) [1,23,24,31]. STS is known to be a major source of visual information for the so-called *human mirror system* (abbreviated HMS) [85]. Originally identified in monkeys, *mirror neurons* are (sensori)motor neurons discharging both when one performs a given action and sees the same action performed by someone else. They have been identified primarily in the ventral premotor cortex and the rostral region of the inferior parietal lobule [20]. The HMS constitutes a neural mechanism that is automatically activated by the sight of somebody else's actions, even when the observer does not make overt movements. The main idea is that during observation the HMS provides a simulation of the actions of other people potentially providing a basis for understanding the intentions of others [31].

Since the foregoing behavioral experiments allow participants to both produce and observe movement at the same time it seems possible that the HMS is at least partially involved in the spontaneous coordination observed. In order to investigate this question, Tognoli and colleagues [99] recorded brain activity of each member of the dyad using a specially designed dual-electroencephalography (EEG) system. Each participant wore a 60-electrode EEG-cap that enabled simultaneous recording of their brains to accompany kinematic measurements of their behavior.

To grasp the significance of the work by Tognoli and colleagues, we need to revert to earlier studies conducted within the framework of Coordination Dynamics have employed instabilities in coordination as a means to uncover the link between the dynamics of behavior and the dynamics of the brain [39,42], with the goal of relating levels by virtue of their *shared dynamical properties* (e. g. [19,39,49,50,52]). In such research, the high temporal resolution of electroencephalography (EEG) and magnetoencephalography (MEG) was exploited to quantify the relationship between the large scale neural dynamics emerging from billions of interconnected neurons and the behavioral dynamics revealed in experiments on coordi-



Social Coordination, from the Perspective of Coordination Dynamics, Figure 3

Relation between Phi_2 and social coordination. a Time-frequency spectrum from electrode CP4 (located over parietal brain regions) from a single trial. Phi_2 is low before and after vision but increases during vision. b Corresponding relative phase between finger movements. Synchronized in-phase behavior is observed during visual contact. Notice the temporary disengagement of the rhythm when coordination is lost briefly (adapted from [99])

nation [41]. Observed features of the dynamics expressed at both levels of description such as multistability and phase transitions (i. e. the spontaneous switch from one pattern to another due to loss of stability), were taken as evidence that principles of self-organization govern pattern formation in both brain and behavior [26,41]. Of particular initial interest was the identification of qualitative changes in the pattern of neural activity that occurred simultaneously with transitions between behavioral patterns [19,49,50,63,104]. On the basis of this work, an exciting hypothesis is that the transitions from uncoordinated to spontaneous coordination observed in the SCD paradigm may be accompanied by similar events at the brain level.

In an effort to shed new light on how social processes are integrated in the brain, Tognoli and colleagues [99] identified several neural mechanisms or *neuromarkers* that appear and disappear with the emergence and dissolution of coordinated behavior between two people. Interestingly, these social neuromarkers consist of brain rhythms in the 10 Hz frequency range located over right centro-parietal areas of the cerebral cortex. In particular, a social brain rhythm termed the *Phi Complex* consists of two components: the first, Phi_1, increases during independent behavior i. e. before information exchange between members of the dyad. When subjects saw each other's finger movements and coordinated together, Phi_1 disappeared and Phi_2, a different rhythm within the same frequency band appeared (Fig. 3) [99].

In a subsequent study, Tognoli and colleagues [98] explored the dynamics of the *Phi Complex* by instructing participants to *intentionally* synchronize when visual information exchange was allowed. In this case, participants interact to accomplish a shared goal. Again, Phi_1 appeared during uncoordinated behavior and Phi_2 when social coordination occurred. Analysis of dyads who participated in both experiments [98,99] revealed that the amplitude of Phi_2 was higher during intentional than spontaneous coordination. Thus, Phi_2 appears to be a neural signature of social coordination whether it emerges spontaneously or not.

The cortical location of the *Phi Complex* appears to be consistent with neuro-anatomical sources within the human mirror system. One of the conclusions drawn by Tognoli and colleagues is that Phi_1 might have an inhibiting role on the mirror system. Previous claims by Brass and Heyes [8] have argued that the mirror system is always active by default and thus must be inhibited in non-social contexts. Hence Phi_1 could inhibit Phi_2, the latter being seen as a facilitator of social coordination that participates

in information exchange between the motor cortex and the mirror system [98,99].

In summary, experiments using dual high density electrode arrays to record and measure brain activity from two persons in conjunction with motion capture technology, have allowed an exploration of shared behavioral and neural social coordination dynamics [98,99]. Transitions from uncoordinated dyadic behavior to interpersonal synchronization have been demonstrated to accompany the emergence of a new brain rhythm – the Phi complex – located in the human mirror system. Such work suggests that SCD may serve as a novel framework for identifying behavioral and neural signatures in reciprocal interactions and allows for a more dynamical approach to the study of the mirror neuron system.

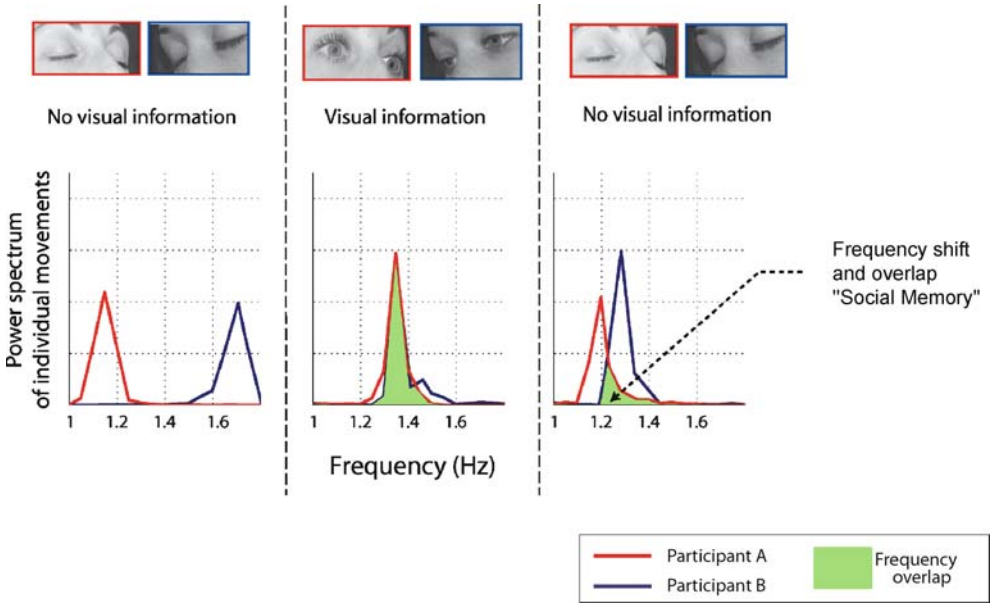
Social Memory and the Dependence on Initial Conditions

At first blush, the emergence of spontaneous coordination between individuals [77,80,82,99] might be seen as an instantiation of mutual entrainment that entails nothing more than a couple of oscillators and a medium of information exchange [41,108]. In such generic cases, once the coupling is removed, each oscillator should return to its own intrinsic frequency, that is, any influence of the interaction should disappear. However, the situation between two people is different (see Fig. 4). Theoretically, in a typical coupled clocks scenario, there should be no difference between the movement periods of the 'clocks' before or after coupling-induced synchronization. However, a serendipitous experimental finding [82] was the consistent and persistent influence of the social interaction on subsequent rhythmic behavior despite the absence of information exchange between the pair (Fig. 4). This remnant of a prior social interaction may qualify as a kind of *social memory* [82].

Social memory is thought to play an important role in human actions, and, to a larger extent, on the way we live [32]. In the context of SCD, social memory implies that the intrinsic parameters of the individual components have been altered by virtue of the social interaction. Mathematically [42], one may represent the situation before the interaction as follows:

$$\begin{aligned} \ddot{x}_1 + i_1(x_1, \dot{x}_1, a_1) + \omega_1^2 x_1 &= 0 \\ \ddot{x}_2 + i_2(x_2, \dot{x}_2, a_2) + \omega_2^2 x_2 &= 0\,. \end{aligned} \tag{1}$$

Where x_1 and x_2 represent the coordinating components,



Social Coordination, from the Perspective of Coordination Dynamics, Figure 4

Evidence for social memory. Illustrated is an example of frequency overlap between the movements of both subjects in a representative trial of the SCD paradigm. *Left column* no vision of each other's movements; each individual moves at their own intrinsic frequency so there is no frequency overlap. *Middle column* visual information is exchanged between participants, causing spontaneous synchronization to occur at a common frequency (and phase, see Fig. 2). *Right column* visual information is no longer exchanged but individuals do not revert back to their initial intrinsic frequency. This remnant of frequency overlap as a result of prior social interaction suggests a kind of 'social memory' (adapted from [82])

$i_{1,2}$, $a_{1,2}$ and $\omega_{1,2}$ refer to individually chosen intrinsic parameters such as the chosen frequency and amplitude.

During the interaction, the system is visually coupled, F_1 and F_2 representing a coupling function such as the well-known HKB-coupling [27,44]:

$$\begin{aligned} \ddot{x}_1 + I_1(x_1, \dot{x}_1, A_1) + \omega_1^2 x_1 &= F_1(x_1, \dot{x}_1, x_2, \dot{x}_2) \\ \ddot{x}_2 + I_2(x_2, \dot{x}_2, A_2) + \omega_2^2 x_2 &= F_2(x_2, \dot{x}_2, x_1, \dot{x}_1) \,. \end{aligned} \tag{2}$$

Now notice that the interactive context has formed a coupling (the right hand side of Eq. (2)) but also led of a modification of the individual component parameters, I and A (on the left hand side of Eq. (2)). One may say that the boundary conditions of Eq. (1) have been altered by the social interaction.

After the interaction, the coupling function disappears (F_1 and F_2 terms on the right hand side are zero) and the system is "uncoupled" (cf. Fig. 1):

$$\begin{aligned} \ddot{x}_1 + I_1(x_1, \dot{x}_1, A_1) + \omega_1^2 x_1 &= 0 \\ \ddot{x}_2 + I_2(x_2, \dot{x}_2, A_2) + \omega_2^2 x_2 &= 0 \,. \end{aligned} \tag{3}$$

However, notice in Eq. (3) the individual intrinsic parameters of the system which were modified by the interaction are still in place. Though uncoupled, the individual components are still affected by the interaction. How this internalization process occurs remains open to empirical investigation.

A benefit of the SCD paradigm is that one is able to quantify the strength and persistence of prior social influences on an individual's behavior. The finding that the modification of the neural network depends on which modality is engaged during the mutual encounter suggests that additional cortical areas may have been recruited and included into the initial global neural assembly due to social context [32]. However, beyond the Phi complex, and perhaps due to inherent limitations in spatial resolution, examination of the dual EEG data showed no evidence of further cortical engagement. Another possibility (in line with the foregoing mathematical analysis) is that the connectivity and dynamics of the initial network is modified by social interaction, and the new organization retained after the interaction is over. Recent evidence in support of this hypothesis suggests that two people engaging in a common task share a representation of each other's movement dynamics, including trajectory amplitude and frequency [6,17]. Such a (shared) representation may persist when vision is removed, i. e. when information exchange is no longer possible [21]. Moreover, representations at the neural level have been shown to be highly flexible and context-dependent [34,35], influenced both by environmental [105] and task demands [79].

The extent and duration of the carryover or remnant effects observed in behavioral experiments may reflect many factors, including the strength of the bond that is formed between people, place in the social hierarchy, the willingness of each participant to cooperate, gender differences, personality characteristics and the significance each participant attaches to the social encounter [32]. An additional finding from our work favors a motor contribution to social memory as well: the persistence effect was found to be independent of the duration of movement that followed the social encounter [82]. This hypothesis is strengthened by results showing that observation of another person performing movements generates a kinematically specific memory of the observed motions in primary motor cortex [95].

The systematic directionality effect observed in the SCD paradigm is revealing also [82]: the extent to which one member of the dyad is influenced by the other was shown to depend on initial conditions. Obviously for synchronization to occur, the person moving with the lowest/highest intrinsic movement frequency must speed up/slow down during information exchange. A surprising result is that the difference between the initial and the final intrinsic movement frequencies (vision absent) was always greater for the person starting with the higher compared to the lower movement frequency [82]. The extent to which initial, so-called 'intrinsic dynamics' determine behavior after the social encounter is over may be of great interest to understanding social interactions in more complex settings where hierarchical relations are involved.

An important problem in human social behavior concerns understanding the degree to which an individual influences the actions of a group (e. g. peer group, family, class) he/she is in. Due to several factors (personality, situational), a person (the leader) may affect the behavior of others more than the others affect her or him. The concept of leadership is commonly associated with interactions taking place in hierarchical settings such as typical organizations, but is actually broader than that. Strength of behavioral influence is overlooked because behavioral interactions have not been systematically studied. Contemporary complex systems approaches (e. g. [28]) view the formation of leader-follower roles as interactive and emergent but in so doing may have undermined the significance of individual dimensions. The approach of SCD is rather to ask: *What makes two people behave independently and what makes them behave as a unit?* The paradigm of social coordination dynamics exploits inherent asymmetry between two people during behavioral interactions and gauges, e. g. using directional coupling measures, which of the two has a stronger influence.

Future Directions

Human beings are social by nature, and interactions with others represent a substantial portion of their many daily activities. A common and well described consequence of interpersonal activity is that an individual's behavior, whether intentional or not, is modified by interactions with others [32]. Alterations of individual and collective behaviors range from imitation and mimicry to spontaneous synchronization, and have been observed in groups varying in size from dyads to thousands of individuals e. g. [4,68].

Social Coordination Dynamics investigates how the natural (uninstructed) social influence of one person on another evolves in real time and has led to a number of new findings. The first is that humans immediately and spontaneously coordinate their actions with each other when provided vision of the movements of the other together with their own. The second is that a specific brain rhythm underlies social coordination. Transitions from individual to coordinated social behavior are observed at both behavioral and brain levels.

The third finding is that an individual's intrinsic behavior is altered by social interaction: the effect of the previous social encounter persists when vision of the other's movements is no longer available. A fourth and final finding is that social coordination is affected by initial conditions, enabling one to predict which individual is most affected by the social encounter.

Insights into elementary forms of social interaction have been obtained by applying the concepts, methods and tools of coordination dynamics. A notable feature of coordination dynamics is its ability to uncover mechanisms and principles common to different kinds of complex systems at different levels of observation and to relate them by virtue of shared behavioral and neural dynamics [41]. SCD and its dynamical measures have proven to provide adequate quantification of the spontaneous coupling between individuals, the transition to loss of entrainment and the effect of the social encounter at both behavioral and brain levels. The same basic patterns of coordinated behavior and pattern dynamics (multistability, critical fluctuations accompanied by a temporary loss of stability, phase transitions, hysteresis and critical slowing down) have been observed within an individual, between an individual and the environment, and between individuals. In this respect SCD complements recent developments in social cognitive neuroscience, behavioral economics, game theory, socio-economics and neuroeconomics (e. g. [10,11,18,73,81,94,103]).

The field of *social neuroeconomics* serves to illustrate the benefits of considering SCD in contexts other than interpersonal sensorimotor interaction. Social neuroeconomics investigates the neural correlates of economical decision making [18]. One particular feature of this nascent field is that decision making processes are always studied in a body- and movement-independent fashion. Why is that? After all, from the very first months of life, individuals live vicariously through one another adopting, if only temporarily, a similar posture or tempo during interactions with a peer, or yawning [4,64,90]. As Henry Greely [22] recently reminded the readership of Science Magazine *"Human society is the society of human brains. Of course those brains are encased in, affected by, and dependent on the rest of the body, but our most important interactions are with other people's brains, as manifested through their bodies."* Although this statement sounds like common sense, thus far the coordination dynamics between bodies has remained unexplored in the field of social neuroeconomics. Yet how many times have we experienced the feeling that trusting someone will be difficult even before talking to them? Whether it was the way she moved or some other factor, body-related cues play a key role in modulating economic decisions (e. g. [93]). A scientific approach to "body language" might aim to understand how perceived actions of others affect the cognitive and emotional processes involved in economical decision-making. For instance, a finding such as the Phi Complex – especially the modulation of Phi_2 when individuals intentionally coordinate [98,99] – could turn to be crucial to better competition–cooperation mechanisms underlying decision in economic contexts such as public coordination games [29]. In sum, as a conceptual framework that encompasses the dynamics of both neural and behavioral levels, SCD promises to bridge the gaps between levels of analysis [41,81] and clear a path for new multi-level, interdisciplinary investigations of social interactions. Like synchronization itself, the function of SCD is to facilitate communication across heretofore unrelated fields.

Acknowledgments

The authors wish to thank their collaborators at the *Human Brain and Behavior Laboratory* (Center for Complex Systems and Brain Sciences) who co-authored the articles on social coordination dynamics reported in this article: Drs. Gonzalo C. de Guzman (*Florida Atlantic University*), Kelly J. Jantzen (*Washington Western University*), Julien Lagarde (*Université de Montpellier I*), Cyrille Magne (*Middle Tenesse State University*) and Emmanuelle Tognoli (*Florida Atlantic University*). Amélie L. and Au-

drey O. should also be thanked for letting the authors use their beautiful hands and eyes in all figures. We wish to thank the following institutions for supporting our research and the writing of this article. The *US National Institutes of Health* (NIMH Innovation Grant MH 42900 and MH080838), the Pierre de Fermat Chair of the Région Midi-Pyrénées (France) to J.A.S.K, and the *Centre National de la Recherche Scientifique* (Programme CNRS Neuroinformatique to O.O.). Correspondence should be addressed to: Olivier Oullier (olivier@oullier.fr) or J. A. Scott Kelso (kelso@ccs.fau.edu). For additional information please visit the following websites: www.oullier.fr and www.ccs.fau.edu/section_links/HBBLv2/index.html

Bibliography

Primary Literature

1. Allison T, Puce A, McCarthy G (2000) Social perception from visual cues: Role of the STS region. Trends Cogn Sci 4:267–278
2. Balaban E (2004) Neurobiology - Why voles stick together. Nature 429:711–712
3. Bardy BG, Oullier O, Bootsma RJ, Stoffregen TA (2002) Dynamics of human postural transitions. J Exp Psychol Hum Percept Perform 28:499–514
4. Barsalou LW, Niedenthal PM, Barbey AK, Ruppert JA (2003) Social embodiment. Psychol Learn Motiv 43:43–92
5. Bennett M, Schatz MF, Rockwood H, Wiesenfeld K (2002) Huygens's clocks. Proc Roy Soc Lond A Math Physic Engineer Sci 458:563–579
6. Bosbach S, Cole J, Prinz W, Knoblich G (2005) Inferring another's expectation from action: The role of peripheral sensation. Nat Neurosci 8:1295–1297
7. Bottani S (1996) Synchronization of integrate and fire oscillators with global coupling. Phys Rev e 54:2334–2350
8. Brass M, Heyes C (2005) Imitation: is cognitive neuroscience solving the correspondence problem? Trends Cogn Sci 9:489–495
9. Buchanan JJ, Kelso JA (1993) Posturally induced transitions in rhythmic multijoint limb movements. Exp Brain Res 94:131–142
10. Camerer CF (2003) Behavioral game theory: Experiments in strategic interaction. Princeton University Press, Princeton
11. Camerer CF, Loewenstein G, Prelec D (2005) Neuroeconomics: How neuroscience can inform economics. J Econ Lit XLIII:9–64
12. Carson RG (2004) Governing coordination. Why do muscles matter? In: Jirsa VK, Kelso JAS (eds) Coordination dynamics: Issues and trends. Springer, Berlin, pp 141–154
13. Carson RG, Kelso JAS (2004) Governing coordination: behavioural principles and neural correlates. Exp Brain Res 154:267–274
14. Condon WS, Sander LW (1974) Neonate movement is synchronized with adult speech - Interactional participation and language acquisition. Science 183:99–101
15. de Guzman GC, Tognoli E, Lagarde J, Jantzen KJ, Kelso JAS (2005) Effects of biological relevance of the stimulus in mediating spontaneous visual social coordination. Society Neurosci Program 867(21)
16. de Rugy A, Salesse R, Oullier O, Temprado JJ (2006) A neuromechanical model for interpersonal coordination. Biol Cybern 94:427–443
17. Decety J, Sommerville JA (2003) Shared representations between self and other: A social cognitive neuroscience view. Trends Cogn Sci 7:527–533
18. Fehr E, Camerer CF (2007) Social neuroeconomics: the neural circuitry of social preferences. Trends Cogn Sci 11:419–427
19. Fuchs A, Kelso JAS, Haken H (1992) Phase transitions in the human brain: Spatial mode dynamics. Int J Bifurc Chaos 2:917–939
20. Gallese V, Fadiga L, Fogassi L, Rizzolatti G (1996) Action recognition in the premotor cortex. Brain 119:593–609
21. Goldman MS, Levine JH, Major G, Tank DW, Seung HS (2003) Robust persistent neural activity in a model integrator with multiple hysteretic dendrites per neuron. Cereb Cortex 13:1185–1195
22. Greely H (2007) On neuroethics. Science 318:533
23. Grèzes J, Fonlupt P, Bertenthal B, Delon-Martin C, Segebarth C, Decety J (2001) Does perception of biological motion rely on specific brain regions? Neuroimage 13:775–785
24. Grèzes J, Armony JL, Rowe J, Passingham RE (2003) Activations related to mirror and canonical neurons in the human brain: An fMRI study. Neuroimage 18:928–937
25. Haken H (1983) Synergetics: An introduction. Springer, Berlin
26. Haken H (1996) Principles of brain functioning: A synergetic approach to brain activity, behavior and cognition. Springer, Berlin
27. Haken H, Kelso JAS, Bunz H (1985) A theoretical-model of phase-transitions in human hand movements. Biol Cybern 51:347–356
28. Hazy JK, Goldstein JA, Lichtenstein BB (2007) Complex systems leadership theory: New perspectives from complexity science on social and organizational effectiveness. ISCE Publishing Company, Mansfield
29. Hichri W, Kirman AP (2007) The emergence of coordination in public good games. Europ J Phys B 55:149–159
30. Hugenii C (1673) Horologium oscillatorium. Apud F. Muguet, Paris
31. Iacoboni M, Molnar-Szakacs I, Gallese V, Buccino G, Mazziotta JC, Rizzolatti G (2005) Grasping the intentions of others with one's own mirror neuron system. PLoS Biol 3:529–535
32. Insel TR, Fernald RD (2004) How the brain processes social information: Searching for the social brain. Annu Rev Neurosci 27:697–722
33. Issartel J, Marin L, Cadopi M (2007) Unintended interpersonal co-ordination: can we march to the beat of our own drum? Neurosci Lett 411:174–179
34. Jantzen KJ, Steinberg FL, Kelso JAS (2004) Brain networks underlying human timing behavior are influenced by prior context. Proc Nat Acad Sci USA 101:6815–6820
35. Jantzen KJ, Steinberg FL, Kelso JAS (2005) Functional MRI reveals the existence of modality and coordination-dependent timing networks. Neuroimage 25:1031–1042
36. Kelso JAS (1981) Contrasting perspectives on order and regulation in movement. In: Long J, Baddeley A (eds) Attention and performance IX. Erlbaum, Hillsdale, pp 437–457
37. Kelso JAS (1984) Phase-transitions and critical behavior in human bimanual coordination. Am J Physiol 246:1000–1004

38. Kelso JAS (1991) Behavioral and neural pattern generation: The concept of NBDS. In: Koepchen HP, Huopaniemi T (eds) Cardiorespiratory and motor coordination. Springer, Munich, pp 224–238
39. Kelso JAS (1992) Coordination dynamics of human brain and behavior. Springer Proc Phys 69:223–234
40. Kelso JAS (1994) The informational character of self-organized coordination dynamics. Hum Mov Sci 13:393–413
41. Kelso JAS (1995) Dynamic Patterns: The self-organization of brain and behavior. MIT Press, Cambridge
42. Kelso JAS (2000) Principles of dynamic pattern formation and change for a science of human behavior. In: Bergman R, Cairns RB, Nilsson LG, Nystedt L (eds) Developmental science and the holistic approach . Lawrence Erlbaum Associates, Mahaw, pp 63–83
43. Kelso JAS (2005) Context, components and complexity. Invited paper, Plexus Institute Summit, Delray Beach, September 11
44. Kelso JAS (2007) The Haken–Kelso–Bunz Model. Scholarpedia (Computational Neuroscience/Dynamical Systems). http://www.scholarpedia.org/article/Haken-Kelso-Bunz_model
45. Kelso JAS, Engstrøm DA (2006) The complementary nature. MIT Press, Cambridge
46. Kelso JAS, Scholz JP, Schöner G (1986) Dynamics govern switching among patterns of coordination. Phys Lett A 134:8–12
47. Kelso JAS, Schöner G, Scholz JP, Haken H (1987) Phaselocked modes, phase transitions and component oscillators in coordinated biological motion. Physica Scripta 35:79–87
48. Kelso JAS, DelColle J, Schöner G (1990) Action-perception as a pattern formation process. In: Jeannerod M (ed) Attention and Performance XIII. Erlbaum, Hillsdale, pp 139–169
49. Kelso JAS, Bressler SL, de Guzman GC, Ding M, Fuchs A, Holroyd T (1991) Cooperative and critical phenomena in the human brain revealed by multiple SQUIDS. In: Duke D, Pritchards W (eds) Measuring chaos in the human brain. World Scientific, New Jersey, pp 97–112
50. Kelso JAS, Bressler SL, de Guzman GC, Ding M, Fuchs A, Holroyd T (1992) A phase-transition in human brain and behavior. Phys Lett A 169:134–144
51. Kelso JAS, Fuchs A, Lancaster R, Holroyd T, Cheyne D, Weinberg H (1998) Dynamic cortical activity in the human brain reveals motor equivalence. Nature 392:814–818
52. Kelso JAS, Fuchs A, Jirsa VK (1999) Traversing scales of brain and behavioral organization. I.-III. In: Uhl C (ed) Analysis of neurophysiological brain functioning. Springer, Heidelberg, pp 73–125
53. Kelso JAS, Fink PW, DeLaplain CR, Carson RG (2001) Haptic information stabilizes and destabilizes coordination dynamics. Proc R Soc Lond B Biol Sci 268:1207–1213
54. Kilner JM, Paulignan Y, Blakemore SJ (2003) An interference effect of observed biological movement on action. Curr Biol 13:522–525
55. Konner M (2004) The ties that bind - Attachment: the nature of the bonds between humans are becoming accessible to scientific investigation. Nature 429:705
56. Körding KP, Fukunaga I, Howard IS, Ingram JN, Wolpert DM (2004) A neuroeconomics approach to inferring utility functions in sensorimotor control. PLoS Biol 2:1652–1656
57. Kuramoto Y (1984) Chemical oscillations, waves, and turbulences. Springer, Berlin
58. Lagarde J, Kelso JAS (2006) Binding of movement, sound and touch: multimodal coordination dynamics. Exp Brain Res 173:673–688
59. Lagarde J, Kelso JAS, Peham C, Licka T (2005) Coordination dynamics of the horse-rider system. J Mot Behav 37:418–424
60. Lagarde J, de Guzman GC, Oullier O, Kelso JAS (2006) Interpersonal interactions during boxing: Data and model. J Sport Exerc Psychol 28: S108
61. Lee DN, Lishman JR (1975) Visual proprioceptive control of stance. J Hum Mov Stud 1:87–95
62. Marsh KL, Richardson MJ, Baron RM, Schmidt RC (2006) Contrasting approaches to perceiving and acting with others. Ecol Psychol 18:1–37
63. Mayville JM, Bressler SL, Fuchs A, Kelso JAS (1999) Spatiotemporal reorganization of electrical activity in the human brain associated with a timing transition in rhythmic auditory-motor coordination. Exp Brain Res 127:371–381
64. McGarva AR, Warner RM (2003) Attraction and social coordination: Mutual entrainment of vocal activity rhythms. J Psycholing Res 32:335–354
65. McNeill WH (1995) Keeping together in time. Harvard University Press, Cambridge
66. Mechsner F (2004) A perceptual-cognitive approach to bimanual coordination. In: Jirsa VK, Kelso JAS (eds) Coordination dynamics: Issues and Trends. Springer, Berlin, pp 177–195
67. Mechsner F, Kerzel D, Knoblich G, Prinz W (2001) Perceptual basis of bimanual coordination. Nature 414:69–73
68. Motter AE, Nishikawa T, Lai YC (2003) Large-scale structural organization of social networks. Phys Rev E 68:e036105-e036110
69. Mottet D, Guiard Y, Ferrand T, Bootsma RJ (2001) Two-handed performance of a rhythmical fitts task by individuals and dyads. J Exp Psychol Hum Percept Perform 27:1275–1286
70. Néda Z, Ravasz E, Vicsek T, Brechet Y, Barabasi AL (2000) Physics of the rhythmic applause. Phys Rev e 61:6987–6992
71. Néda Z, Ravasz E, Brechet Y, Vicsek T, Barabasi AL (2000) The sound of many hands clapping - Tumultuous applause can transform itself into waves of synchronized clapping. Nature 403:849–850
72. Nicolis G, Prigogine I (1977) Self-organization in non-equilibrium systems. Wiley, New York
73. Oullier O, Kelso JAS (2006) Neuroeconomics and the metastable brain. Trends Cogn Sci 10:353–354
74. Oullier O, Temprado JJ (2005) The structure of visual backgrounds modulates interpersonal coordination dynamics in a virtual environment. J Sport Exerc Psychol 27:S118-S119
75. Oullier O, Bardy BG, Stoffregen TA, Bootsma RJ (2002) Postural coordination in looking and tracking tasks. Hum Mov Sci 21:147–167
76. Oullier O, de Guzman GC, Jantzen KJ, Kelso JAS (2003) On context dependence of behavioral variability in inter-personal coordination. Int J Comput Sci Sport 2:126–128
77. Oullier O, de Guzman GC, Jantzen KJ, Lagarde JF, Kelso JAS (2004) Spontaneous interpersonal synchronization is modulated by the degree of visual coupling. J Sport Exerc Psychol 26:S11-S11
78. Oullier O, Bardy BG, Stoffregen TA, Bootsma RJ (2004) Task-specific stabilization of postural coordination during stance on a beam. Mot Control 8:174–187

79. Oullier O, Jantzen KJ, Steinberg FL, Kelso JAS (2005) Neural substrates of real and imagined sensorimotor coordination. Cereb Cortex 15:975–985
80. Oullier O, de Guzman GC, Jantzen KJ, Lagarde J, Kelso JAS (2005) Spontaneous interpersonal synchronization. In: Peham C, Schöllom WI, Verwey W (eds) European workshop on movement sciences: Mechanics-Physiology-Psychology. Sportverlag, Köln, pp 34–35
81. Oullier O, Kelso JAS, Kirman AP (2008) Neuroeconomics: A dynamical systems perspective. Rev Econ Pol 118:51–62
82. Oullier O, de Guzman GC, Jantzen KJ, Lagarde J, Kelso JAS (2008) Social coordination dynamics: Measuring human bonding. Soc Neurosci 3:178–192
83. Richardson MJ, Marsh KL, Schmidt RC (2005) Effects of visual and verbal interaction on unintentional interpersonal coordination. J Exp Psychol Hum Percept Perform 31:62–79
84. Richardson MJ, Marsh KL, Isenhower RW, Goodman JR, Schmidt RC (2007) Rocking together: dynamics of intentional and unintentional interpersonal coordination. Hum Mov Sci 26:867–891
85. Rizzolatti G, Craighero L (2004) The mirror-neuron system. Annu Rev Neurosci 27:169–192
86. Schmidt RC, O'Brien B (1997) Evaluating the dynamics of unintended interpersonal coordination. Ecol Psychol 9:189–206
87. Schmidt RC, Richardson MJ (2008) Dynamics of interpersonal coordination. In: Fuchs A, Jirsa VK (eds) Springer, Heidelberg, pp 281–308
88. Schmidt RC, Carello C, Turvey MT (1990) Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. J Exp Psychol Hum Percept Perform 16:227–47
89. Schöner G, Kelso JAS (1988) Dynamic pattern generation in behavioral and neural systems. Science 239:1513–1520
90. Schurmann M, Hesse MD, Stephan KE, Saarela M, Zilles K, Hari R, Fink GR (2005) Yearning to yawn: the neural basis of contagious yawning. Neuroimage 24:1260–1264
91. Shockley K, Baker AA, Richardson MJ, Fowler CA (2007) Articulatory constraints on interpersonal postural coordination. J Exp Psychol Hum Percept Perform 33:201–208
92. Simon JR (1969) Reactions toward the source of stimulation. J Exp Psychol 81:174–176
93. Solnick SJ, Schweitzer ME (1999) The Influence of Physical Attractiveness and Gender on Ultimatum Game Decisions. Organ Behav Hum Decis Process 79:199–215
94. Sommerville JA, Decety J (2006) Weaving the fabric of social interaction: Articulating developmental psychology and cognitive neuroscience in the domain of motor cognition. Psychon Bull Rev 13:179–200
95. Stefan K, Cohen LG, Duque J, Mazzocchio R, Celnik P, Sawaki L, Ungerleider L, Classen J (2005) Formation of a motor memory by action observation. J Neurosci 25:9339–9346
96. Temprado JJ, Laurent M (2004) Attentional load associated with performing and stabilizing a between-persons coordination of rhythmic limb movements. Acta Psychol 115:1–16
97. Temprado JJ, Swinnen SP, Carson RG, Tourment A, Laurent M (2003) Interaction of directional, neuromuscular and egocentric constraints on the stability of preferred bimanual coordination patterns. Hum Mov Sci 22:339–363
98. Tognoli E, Magne C, de Guzman GC, Kelso JAS (2007) Brain rhythms underlying intentional social coordination. Society Neurosci Program 304
99. Tognoli E, Lagarde J, de Guzman GC, Kelso JA (2007) The phi complex as a neuromarker of human social coordination. Proc Nat Acad Sci USA 104:8190–8195
100. Turvey MT (2004) Impredicativity, dynamics and the perception-action divide. In: Jirsa VK, Kelso JAS (eds) Coordination Dynamics: Issues and Trends. Springer, Berlin
101. Vallacher RR, Nowak A (1997) The emergence of dynamical social psychology. Psychol Inquiry 8:73–99 (and commentaries therein)
102. van Ulzen NR, Lamoth CJ, Daffertshofer A, Semin GR, Beek PJ (2008) Characteristics of instructed and uninstructed interpersonal coordination while walking side-by-side. Neurosci Lett 432:88–93
103. Vinkovic D, Kirman A (2006) A physical analogue of the Schelling model. Proc Nat Acad Sci USA 103:19261–19265
104. Wallenstein GV, Kelso JAs, Bressler SL (1995) Phase transitions in spatiotemporal patterns of brain activity and behavior. Physica D 84:626–634
105. Wheeler ME, Peterson SE, Buckner RL (2004) Memory's echo: Vivid remembering reactivates sensory-specific cortex (vol 97, pg 11125, 2000). Proc Nat Acad Sci USA 101:5181
106. Winfree AT (1967) Biological rhythms and behavior of populations of coupled oscillators. J Theor Biol 16:15–42
107. Winfree AT (1980) The geometry of biological time. Springer, New York
108. Winfree AT (2002) On emerging coherence. Science 298: 2336–2337

Books and Reviews

Frith C, Wolpert DM (2004) The neuroscience of social interaction: Decoding, imitating and influencing the actions of others. Oxford University Press, Oxford

Fuchs A, Jirsa VK (eds) (2008) Coordination: Neural, behavioral and social dynamics. Springer, Berlin

Jantzen KJ, Kelso JAS (2007) Neural coordination dynamics of human sensorimotor behavior: A Review. In: Jirsa VK, McIntosh AR (eds) Handbook on brain connectivity. Springer, Berlin, pp 421–461

Jeannerod M (2006) Motor cognition: What actions tell to the self. Oxford University Press, Oxford

Jirsa VK, Kelso JAS (2004) Coordination dynamics: Issues and trends. Springer, Berlin

Kelso JAS (1995) Dynamic Patterns: The self-organization of brain and behavior. MIT Press, Cambridge

Kelso JAS, Engstrøm DA (2006) The complementary nature. MIT Press, Cambridge

Oullier O, Jantzen KJ (2008) Neural indices of behavioral instability in coordination dynamics. In: Fuchs A, Jirsa VK (eds) Springer, Berlin, pp 205–227

Pikovsky A, Rosenblum M, Kurths J (2001) Synchronization: A universal concept in nonlinear science. Cambridge University Press, Cambridge

Rizzolatti G, Craighero L (2004) The mirror-neuron system. Ann Rev Neurosci 27:169–192

Sommerville JA, Decety J (2006) Weaving the fabric of social interaction: Articulating developmental psychology and cognitive neuroscience in the domain of motor cognition. Psychon Bull Rev 13:179–200

Strogatz SH (2003) Sync: The emerging science of spontaneous order. Hyperion Press, New York

Social Network Analysis, Estimation and Sampling in

Ove Frank
Statistics Department, Stockholm University, Stockholm, Sweden

Article Outline

Glossary
Definition of the Subject
Introduction
Background
Basic Terminology and Notation
Sampling Designs
Snowball Sampling Designs
Estimation Based on Snowball Samples
Network Walk Sampling Designs
Estimation Based on Network Walk Samples
Methods for Estimating Hidden Populations
Probabilistic Network Models
Bipartite Network Sampling Designs
A Bipartite Network Model for Crime Participation
Future Directions
Bibliography

Glossary

Network model A network model describes how an observed network structure with its data could be generated from probabilistic assumptions.

Network population A network population is a finite or infinite collection of existing or conceivable networks. For instance, the population could consist of network instances of a process changing with time, or the population could consist of network realizations in different sets of units.

Observation scheme An observation scheme describes what variables are known or can be observed for units and relations in the population and in the sample.

Population network A population network is a population of units equipped with relational structure between them. Usually several attributes are attached both to units and to pairs of units, and the network can be described as a valued graph on the population of units with vertex and edge or arc variables defined in the graph.

Sampling design A sampling design describes how units are sampled from a population of units and what the selection probabilities are for all the possible sample sequences or sample subsets. The sampling procedure is usually controlled by the investigator. If the sampling is not controlled by the investigator but voluntary or "controlled by nature", the sample network can be considered as an outcome of a probabilistic network model, and there is no clear separation between design-based and model-based inference.

Sampling frame A sampling frame is a list of identifying labels of all possible sample units.

Social network A social network is a network consisting of units and relations of interest in the social sciences. The term is used both for population networks and for networks belonging to a network population.

Survey sampling Survey sampling of networks refer to the investigation of networks sampled from a network population while survey sampling in networks refer to the investigation of sub-networks sampled from a population network. Survey sampling of networks is usually model-based with an infinite network population. Survey sampling in networks is usually design-based with a finite population network.

Definition of the Subject

Sampling in network analysis can refer either to the sampling of networks from a population of networks or to the sampling of one or more sub-networks from a population network. Sampling designs and model approaches based on probabilistic methods allow statistical tools to be developed for the analysis of data obtained in surveys carried out in a network setting.

From about the 1960s the rapid development of computers and computer power changed the scene for statistics and applied mathematics in general. New possibilities of processing large amounts of data by using fast computer algorithms for calculations, sorting, and searching contributed to an accelerated interest in computer science and discrete mathematics. Simulation techniques and other computer intensive numerical methods revolutionized the arsenal of tools available in statistics and other parts of applied mathematics. Graph theory was established not only as an expanding part of discrete mathematics but also became applied as a tool for investigating relational structures in the social sciences. Friendship, co-operation, dominance, support and other relations between individuals were studied as sociograms or as more general networks in which both the individuals and the relations were attributed with qualitative and quantitative variables like gender and age of the individuals and type and strength of the relations. Other network applications include transmission of infectious diseases between people, work scheduling and other planning, routing systems

for transportation of goods, traffic flows in urban areas, information flows in communication channels, administrative and organizational structure, capital transfer and surveillance systems. Characteristic of such network applications is the inherent complexity of co-variation or correlation that could exist within and between variables defined for different elements of the network.

Introduction

Special sampling designs using network information for successive link tracing or sample adjustment have been described in the literature under various names such as multiplicity sampling, snowball sampling, and Respondent driven sampling. Such designs have been used for investigations of hard-to-reach populations of various kinds, like drug addicts, homeless people and individuals at risk for hiv-infection, tuberculosis, or hepatitis. Such designs have also been proved useful in investigations of rare properties of individuals in populations that could be sampled from available frames but would require very large samples with conventional sampling designs.

This presentation focuses on sampling designs in finite population networks and network models developed for survey sampling of social networks. Estimation methods are described that can be applied under different assumptions about what population data are known and what sample data can be observed.

The next three sections give background material of various kinds. A brief historical background is given that describes the development of a survey sampling theory for networks. Basic terminology and some general notation needed is then specified and serves the purpose of unifying and simplifying some of the concepts and arguments used later. A common and characteristic feature of network surveys is that they allow non-orthodox sampling methods. Hence it is important to specify when such methods are justified. Some fundamentals about sampling designs are given together with an overview of some standard designs.

After this background follows five sections on network sampling designs. Sections "Snowball Sampling Designs"–"Estimation Based on Network Walk Samples" treat snowball sampling designs and the important sub-class of network walk designs. For various special cases and modifications of basic snowball and walk samples, standard estimation methods are discussed. Section "Methods for Estimating Hidden Populations" is devoted entirely to the estimation of sizes and properties of hidden populations and rare parts of large populations.

The discussion of network designs is followed by a presentation of various probabilistic network models used especially in the social sciences. Some of these models are appropriate in combination with designs or as tools for deriving so-called model-assisted estimators. The main interest in these network models stems from their use when an observed network is considered as an instant of a network process or as a sample network from a super-population of networks.

Next there are two sections on sampling designs and models for bipartite networks. Bipartite networks have units of two kinds which generally require very different treatment so that it is not appropriate to consider them as special cases of networks. Section "Bipartite Network Sampling Designs" describes sampling designs and estimators in bipartite networks. A combination of sampling designs and random modeling in bipartite networks is illustrated in Sect. "A Bipartite Network Model for Crime Participation" which treats a specific criminological application.

Finally, Sect. "Future Directions" comments briefly on possible future directions for the development of network surveys in the social sciences.

Background

Some early attempts to incorporate graphs in a statistical investigation are the studies on sociograms [49], on snowball sampling [33], and on sampling in graphs [3]. Of particular interest is a paper by Stephan [60] pointing to the need for network methods in sample survey theory. He suggests the term nexus sampling for surveys that consider data from units as well as from relations between units. Sampling in populations with graph structure is also discussed in the articles [6,14], and a comprehensive monograph on statistical inference in graphs is the thesis [16].

An independent discussion of some first steps in network sampling is given in [34,42]. Various problems in survey sampling and estimation in networks were discussed in a series of papers during the 1970s and 1980s: [7, 17,18,19,20,21,22,23,24,25,26,27,48]. These papers use essentially various network sampling designs and no elaborate probabilistic model assumptions.

The theory of random graphs developed mainly from two models introduced by Erdös and Renyi in 1959 and 1960 [12,13], the uniform graph with n vertices and N randomly selected edges among the unordered pairs of vertices, and the Bernoulli graph with n vertices and independent insertions of edges with a common probability p for all the unordered pairs of vertices. These random graphs have challenged many mathematicians, and there is now an extensive literature on combinatorial and asymptotic results for random graphs. A brilliant comprehensive ref-

erence on general graph theory is [11]. Monographs concentrating on random graphs are [4,37,46].

Statistical inference based on probabilistic network models without explicitly addressing the sampling design problem has been treated for some variants of Bernoulli graphs, but generally such models have too little structure to be of interest in a statistical context with network data. A network model with a more elaborate and flexible parametrization that make it applicable in the social sciences is the random digraph model of Holland and Leinhardt introduced in [36]. A comprehensive monograph with many references to the social network literature is [67]. The introduction of Markov graphs in [30] and the log-linear modeling of random networks in [47,57,68] have had a major impact on the development of statistical methods for social network analysis. The edited volumes [5,8] are good sources for the methodological development. Various aspects of network modeling and statistical inference from network samples are the topics of several theses in recent years. Well known to me are the following important contributions to the methodological development: [10,35,38,40,42,55,59,62].

Basic Terminology and Notation

Survey sampling concepts of basic importance are populations of units, sample selection, sampling designs, observation schemes, and variables of interest defined on the units. Two standard reference works for survey sampling in finite populations are [54,63]. In order to make network sampling possible, the population has to be equipped with a relational structure between its units. Usually it is given as a binary variable defined on the ordered pairs of units. A population with a binary relation can be represented as a graph in which the vertices represent the units and the edges (arcs) indicate the unordered (ordered) pairs of units that are related. In a network setting vertices and edges (arcs) are sometimes called nodes and links.

The population from which samples are drawn can be different from the target population of the survey. It can also be changing with time so that particular precautions need to be taken. Disregarding such complications we assume that there is a finite population of N units labeled by integers $1, \ldots, N$. The population is denoted $U = \{1, \ldots, N\}$. Samples are drawn from U either as ordered sequences of units with or without repetitions allowed or as subsets of U. In the first case, a sample sequence of m units drawn is denoted $u = (u_1, \ldots, u_m)$, and in the second case a sample subset is denoted S. Both u and S are random entities when the sample is drawn according to a probability design. The probability design of u is given by a probability distribution over the set U^m of all ordered m-sequences of units from U, and the probability design of S is given by a probability distribution over the class of all subsets of U. If the size n of S is fixed, all subsets of other sizes have zero probability.

The observation scheme specifies which variables are known or observable. A variable defined on the units (a vertex variable) is given by a function x from U to some range space R. Formally,

$$x = \{(i, x_i) \colon i \in U\} ,$$

is a set of pairs (i, x_i) assigning a value x_i in R to each unit i in U. If there is a specified ordering of the units, x could also be represented as a sequence of values ordered in the same way, e. g. an ordered sequence $(x_1, \ldots, x_N)$ corresponding to the order $1, \ldots, N$ of the units. The first representation is more convenient when restrictions to different subsets of U are considered. For any subset A of U we denote by

$$x(A) = \{(i, x_i) \colon i \in A\} ,$$

the restriction of the variable x to units in A. In particular, $x(U) = x$ and $x(\emptyset) = \emptyset$. A binary variable x with $R = \{0, 1\}$ and, more generally, a K-category variable x with $R = \{0, \ldots, K-1\}$ partitions the population U into K disjoint subsets

$$U_j = \{i \in U \colon x_i = j\} \quad \text{for} \quad j \in R .$$

A variable defined on the ordered pairs of units (an arc variable) is given by a function y from U^2 to some range space R. Formally,

$$y = \{(i, j, y_{ij}) \colon i \in U, j \in U\} ,$$

is a set of triplets (i, j, y_{ij}) assigning a value y_{ij} in R to each ordered pair (i, j) in U^2. If there is a specified ordering of the units in U, y could also be represented as a matrix array (y_{ij}) with N rows and N columns ordered according to the unit-ordering. For any two subsets A and B of U we denote by

$$y(A, B) = \{(i, j, y_{ij}) \colon i \in A, j \in B\} ,$$

the restriction of the variable y to ordered pairs of units (i, j) with $i \in A$ and $j \in B$. The set $y(A, B)$ is simply referred to as the set of values of y from A to B. In particular, $y(U, U) = y$ and $y(A, \emptyset) = y(\emptyset, B) = \emptyset$ for any subsets A and B of U. A binary variable y with $R = \{0, 1\}$ and, more generally, a K-category variable y with $R = \{0, \ldots, K-1\}$ partitions U^2 into K disjoint subsets

$$\{(i, j) \in U^2 \colon y_{ij} = r\} \quad \text{for} \quad r \in R .$$

Sampling Designs

In order to gain general information about population units with access only to the information from a sample of units, it is evident that the sample should represent the population in some specified way. Even if the population is partitioned into subpopulations according to several attributes of the units, and the sample is selected so that some units from each subpopulation are included in the sample, one cannot be sure that the sample allows inference to be drawn with confidence about collective properties of all the population units.

Survey sampling theory is based on information from samples that are selected according to probabilistic designs. A sampling design that gives all population units a positive probability of being included in the sample is required to achieve a sample that is representative of the population. The essential property of a sampling design that makes it appropriate for population inference is that its inclusion probabilities should be known or estimable for all units. This allows sample data to be weighted so that they accurately represent population data. A probabilistic sampling design not only makes inference possible; it also makes it possible to specify how inference uncertainty can be judged and quantified. During the early days of survey sampling, the benefits of probabilistic sampling designs were not always properly understood and had to be emphasized. Today this is so well known that it may seem dubious if someone tries to use non-probability samples.

Therefore, it is important to clarify the legitimacy of network sampling designs that involve arbitrarily or conveniently selected samples. The essential principle is still that inclusion probabilities should be known or estimable. Some network sampling designs are based on an initial probability sample followed by sequential samples that may be systematically or probabilistically selected. As a consequence, it might be hard to estimate inclusion probabilities for the final network sample, but there are no doubts about its legitimacy.

When the network sample is generated from an initial sample that is not a probability sample, it is important that the sequential samples generated after the initial sample are selected according to specified probabilistic rules. Under some fairly general assumptions to be specified, the inclusion probabilities for the final network sample can be proved to converge towards limits that are independent of the way the initial sample was selected. Thus, the lack of knowledge about the initial conditions is compensated for by letting the network sample consist of several waves of units sequentially selected according to a known or estimable probability design. This wave design determines the limiting inclusion probabilities of the network sampling design, and the accuracy of these probabilities increases with the number of waves. As we will see later, it is possible to modify the probabilistic wave design so that the limiting inclusion probabilities of the network sample become uniform or equal to any other desired probability distribution over the population of units.

It is convenient to specify a few standard sampling designs and their inclusion probabilities which are fundamental for some of the estimators discussed later. Consider a population U of N units and a sample subset S selected according to a uniform probability design over the n-subsets of U. This design is specified as

$$S \sim \mathrm{Unif}(n, U) \,.$$

It is often called simple random sampling without replacement. Thus, the probability that $S = s$ is equal to

$$P(s) = n!/N(N-1)\ldots(N-n+1) \,,$$

for any n-subset s of U. The probability of inclusion in S of a fixed unit i from U is equal to the sum of the selection probabilities $P(s)$ over all s containing unit i. It is equal to n/N and denoted by

$$\pi_i = n/N \,,$$

for any i in U. The probability of inclusion in S of two distinct units i and j from U equals

$$\pi_{ij} = n(n-1)/N(N-1) \,,$$

for any distinct i and j in U.

Another subset design called Bernoulli sampling with parameter $p (0 < p < 1)$ is specified as

$$S \sim \mathrm{Bern}(p, U) \,,$$

and is defined by selection probabilities

$$P(s) = p^n (1-p)^{N-n} \,,$$

for any subset s of U of size n for $n = 0, 1, \ldots, N$. Its inclusion probabilities are

$$\pi_i = p \,,$$

for any i in U, and

$$\pi_{ij} = p^2 \,,$$

for any distinct i and j in U. An important property of Bernoulli sampling is that its inclusion indicators

$I_i = I(i \in S)$ are independent. The inclusion indicators are 1 or 0 according to whether unit i is included in S or in its complement $U - S$.

A generalization of Bernoulli sampling is obtained if the inclusion indicators are independent but not necessarily identically distributed Bernoulli variables. The sampling design with independent Bernoulli(p_i) distributed inclusion indicators for i in U is sometimes called Poisson sampling (perhaps referring to the fact that the random sample size is approximately Poisson distributed with parameter $\lambda = p_1 + \cdots + p_N$). We specify this design as a Bernoulli sampling with general parameter $p = \{(i, p_i)\colon i \in U\}$ satisfying $0 < p_i < 1$ for all i in U; in short

$$S \sim \text{Bern}(p_i\colon i \in U)\,.$$

Its selection probabilities are

$$P(s) = \prod_{i \in s} p_i \prod_{j \in U - s} (1 - p_j)\,,$$

for any subset s of U, and its inclusion probabilities are

$$\pi_i = p_i\,,$$

for any i in U, and

$$\pi_{ij} = p_i p_j\,,$$

for any distinct i and j in U.

A probability design for a sample sequence $u = (u_1, \ldots, u_m)$ of m units independently drawn according to a uniform distribution over U has selection probabilities

$$P(i_1, \ldots, i_m) = 1/N^m\,,$$

for any sequence $(i_1, \ldots, i_m)$ in U^m. This design specified by

$$u_k \sim \text{Unif}(U) \quad \text{independent for } k = 1, \ldots, m\,,$$

is often called simple random sampling with replacement. Its inclusion probabilities are given by

$$\pi_i = 1 - (1 - 1/N)^m\,,$$

for any i in U, and

$$\pi_{ij} = 1 - 2(1 - 1/N)^m + (1 - 2/N)^m\,,$$

for any distinct i and j in U.

A generalization of this design to m independent identically distributed (IID) draws according to a probability distribution with N arbitrary positive probabilities $p(i)$ summing to 1 for i in U has selection probabilities

$$P(i_1, \ldots, i_m) = p(i_1) \ldots p(i_m)\,,$$

for any sequence $(i_1, \ldots, i_m)$ in U^m. We specify this design by

$$u_k \sim \text{IID}(p(i)\colon i \in U) \quad \text{for } k = 1, \ldots, m\,.$$

Its inclusion probabilities are

$$\pi_i = 1 - [1 - p(i)]^m\,,$$

for any i in U, and

$$\pi_{ij} = 1 - [1 - p(i)]^m - [1 - p(j)]^m + [1 - p(i) - p(j)]^m\,,$$

for any distinct i and j in U. A further generalization would allow the sequence of independent draws to be made according to different distributions.

For any probability design of a sample sequence $u = (u_1, \ldots, u_m)$ we define the inclusion count or multiplicity m_i of unit i as the number of draws equal to that unit. The multiplicities sum to m, and a positive multiplicity of a unit indicates its inclusion in the sample sequence. The set of distinct units in a sample sequence u is denoted by $s(u)$. This is a random subset of U of size $n(u)$ and expected size

$$\text{E}\ n(u) = \sum_{i \in U} \pi_i\,,$$

and variance

$$\text{Var}\ n(u) = \sum_{i \in U} \sum_{j \in U} (\pi_{ij} - \pi_i \pi_j)\,.$$

Any sample sequence $u = (u_1, \ldots, u_m)$ with $u_k \sim \text{IID}(p(i)\colon i \in U)$ for $k = 1, \ldots, m$ has multiplicities $(m_1, \ldots, m_N)$ that are multinomially distributed with parameters m and $(p(1), \ldots, p(N))$. The number of distinct units $n(u)$ in u is a sum of independent Bernoulli(π_i)-variables, which is asymptotically Poisson($\sum_i \pi_i$). With dependent draws it is generally difficult to derive the simultaneous probability distribution of the multiplicities or the distribution of the number of distinct units in u.

Snowball Sampling Designs

Let the population $U = \{1, \ldots, N\}$ have a graph structure given by a binary variable y defined on the pairs of population units. Let the sets of units adjacent after and before unit i be

$$A_i = \{j \in U\colon y_{ij} = 1\} \quad \text{and} \quad B_i = \{j \in U\colon y_{ji} = 1\}\,,$$

for $i \in U$. For any subset s of U, we denote by $A(s)$ the union of the A_i for $i \in s$. A snowball sampling design starts with an initial subset S_0 of U. This subset can either be

a convenience sample or the outcome of a probability sampling design with a known or unknown probability distribution P_0 over the class of all subsets of U. In the latter case there are selection probabilities $P_0(s) \geq 0$ summing to 1 for all subsets s of U. In the former case S_0 is equal to a fixed set s_0, and it is convenient to put $P_0(s_0) = 1$. Note that in either case we don't even know that all units in the population have positive probabilities of being included in the initial sample.

The snowball sample is obtained by successively extending the initial sample S_0 with units selected from the sets of units that are adjacent after previously joined units. More precisely, a first wave W_1 is selected according to a probability distribution over the class of subsets of

$$(U - S_0) \cap A(S_0) .$$

The union of the initial sample S_0 and the disjoint first wave W_1 is equal to the one-wave snowball S_1. Generally, the kth wave W_k is selected according to a probability distribution over the class of subsets of

$$(U - S_{k-1}) \cap A(W_{k-1}) ,$$

for $k = 1, 2, \ldots$ with $W_0 = S_0$. The k-wave snowball S_k is the union of the $(k-1)$-wave snowball S_{k-1} and the kth wave W_k. The waves are disjoint and

$$W_k = S_k - S_{k-1} \quad \text{for} \quad k = 1, 2, \ldots$$

The k-wave snowball S_k is called saturated if the kth wave W_k is the last wave that is non-empty, so that the sequence of snowballs has reached an absorbing state

$$S_k = S_{k+1} = \cdots$$

According to the selections of the waves, it follows that the k-wave snowball S_k evolves as a stochastic process depending on the $(k-1)$-wave snowball S_{k-1} and the kth wave which depends on S_{k-1} and $W_{k-1} = S_{k-1} - S_{k-2}$. Thus, the snowball process S_k has a two-step memory: The probability distribution of S_k conditional on $S_0, \ldots, S_{k-1}$ depends on S_{k-1} and S_{k-2} only, for $k = 2, 3, \ldots$ Another equivalent way to express this is to say that the stochastic process with states consisting of the current snowball and its last wave (S_k, W_k) is a Markov chain.

The transition probabilities of this Markov chain are

$$\begin{aligned} P(S_{k+1} = s_{k+1}, W_{k+1} = w_{k+1} | S_k = s_k, W_k = w_k) \\ = P(W_{k+1} = w_{k+1} | S_k = s_k, W_k = w_k) \\ = P(w_{k+1} | s_k, w_k) , \end{aligned}$$

for s_{k+1} equal to the union of s_k and w_{k+1}. The Markov chain is assumed to be time-homogeneous so that the transition probabilities do not depend on the stage parameter k. The transition probabilities do depend on graph parameters

$$y(w_k, U - s_k) = \{(i, j, y_{ij}) \colon i \in w_k, j \in U - s_k\} ,$$

governing possible wave selections and possibly also on other parameters governing properties of the wave selections. Assume that saturation occurs at stage k or later. Then the sequence of snowball samples $(S_0, S_1, \ldots, S_k)$ has a probability distribution given by

$$\begin{aligned} P(S_0 = s_0, S_1 = s_1, \ldots, S_k = s_k) = P(s_0, s_1, \ldots, s_k) \\ = P_0(s_0) P(w_1 | s_0, w_0) P(w_2 | s_1, w_1) \\ \cdots P(w_k | s_{k-1}, w_{k-1}) , \end{aligned}$$

for any strictly increasing subset sequence $(s_0, s_1, \ldots, s_k)$ with $s_j - s_{j-1} = w_j$ for $j = 1, 2, \ldots, k$ and $s_0 = w_0$.

The probability distribution of the k-wave snowball S_k is obtained by summing the probabilities $P(s_0, s_1, \ldots, s_k)$ over all strictly increasing subset sequences $(s_0, s_1, \ldots, s_{k-1})$ with s_{k-1} strictly included in s_k. The initial sample and each new wave need to contain at least one unit, so that s_k contains at least $k + 1$ units. If s_k contains more units, the sum is extended over all possible distributions of the non-initial units among the k waves.

In order to be able to perform such calculations, we need to specify the common sampling design of the waves. At stage k, the next wave W_{k+1} has a probability distribution over the class of subsets of that part of U that is not included in the current snowball S_k but is adjacent after its last wave W_k. Now the selection of a subset for W_{k+1} can be considered as obtained by the recruitment of new units made by units in W_k. Each unit i in W_k can recruit only from A_i, the set of units adjacent after unit i. Assume that the recruited set R_i is a Bernoulli sample with inclusion probabilities p_{ij} for j in A_i. The recruited sets R_i are assumed to be selected independently for all i in U. The union of the sets R_i recruited by the units i in W_k is the set $R(W_k)$. This set is a subset of $A(W_k)$, and the part of it that is not included in S_k is the new wave W_{k+1}. Thus

$$W_{k+1} = (U - S_k) \cap R(W_k) ,$$

where $R(W_k)$ as a union of independent Bernoulli samples is itself a Bernoulli sample conditional on W_k. The probability that unit j is included in $R(W_k)$ equals the probability that j is included in at least one of the R_i for i in W_k. This probability is denoted

$$p(W_k, j) = 1 - \prod (1 - p_{ij}) ,$$

where the product is over i in W_k and j is a unit in $A(W_k)$. It follows that conditional on the current state of the Markov chain (S_k, W_k), the next wave W_{k+1} is a Bernoulli sample with inclusion probabilities

$$p(W_k, j) \quad \text{for } j \text{ in } (U - S_k) \cap A(W_k)\,.$$

The transition probabilities of the Markov chain (S_k, W_k) can now be specified as

$$P(w_{k+1}|s_k, w_k) = \left[\prod p(w_k, j)\right]\left[\prod q(w_k, j)\right]\,,$$

where the first product is over j in w_{k+1} and the second product is over j in $U - s_{k+1}$ with probabilities denoted $q(w_k, j) = 1 - p(w_k, j)$. The basic parameters here are the probabilities p_{ij} that unit i would recruit unit j. Only units j in A_i can be recruited, and only units i included in the initial sample or any of the waves will have an opportunity to recruit. We can interpret the parameters p_{ij} as probabilities of independent recruitment arcs in the population graph. When y is considered as specifying a fixed population graph, p_{ij} is zero whenever y_{ij} is zero. If all units in A_i should have the same probability of being recruited, we could put $p_{ij} = \alpha_i y_{ij}$ where α_i is a probability governing the recruitment or co-operation activity of unit i. If recruitment also depends on availability, status or attraction of the recruited unit, we could put $p_{ij} = \alpha_i \beta_j y_{ij}$. The parameters α_i and β_i for units i in U could be related to a categorical variable x defined on U. For a K-category variable x the parameters $\alpha_i = \alpha(x_i)$ and $\beta_i = \beta(x_i)$ have K possible values each, and the parametrization involves only $2K$ parameters.

Recruitment or selection of adjacent members of the population can be considered as a sampling performed by the investigator or by the respondent. The units of the population can be people that could be interviewed by the investigator. The units can also be geographical regions or other entities for which the adjacent units can either be observed directly or determined by interviews with some unit representative. In such cases when the investigator can get access to information about all units that are adjacent after a sampled unit, then the recruitment of further units can be made by the investigator. There are situations when information about adjacent units might be sensitive or embarrassing to the respondent. If the respondent is unwilling to release information about all adjacent units, a possibility might be to ask the respondent just to select some of them at random. Such respondent driven sampling might in some situations give more reliable data and could provide a reasonable way to avoid some problems of integrity. For a more thorough discussion of this, see, for instance, [53]. There are techniques developed with distribution of recruitment cards and various benefits to participants in order to help the investigator to get new waves of the sample. Even if respondent i provides only R_i and not A_i, information could be collected about out-degrees or other statistics needed for the respondent driven probability design of the wave sampling.

The combination of design based and model based probabilities are standard in survey sampling and should be so also when network sampling is applied. Respondent driven sampling could be such a case when a random graph approach would be appropriate for the wave sampling. An illustration of combination of design-based and model-based approaches to network surveys is given in [66].

Estimation Based on Snowball Samples

Estimation of population data from snowball sample data requires inclusion probabilities of all population units. Any fixed unit j is included in a k-wave snowball sample if it is included in the initial sample or in any of its k following waves. Since the initial sample and the waves are disjoint, their inclusion events are mutually exclusive, and the probability that the snowball includes j equals the sum of the probabilities that the initial sample or any of the waves include j:

$$P(j \in S_k) = P(j \in S_0) + P(j \in W_1) + \cdots + P(j \in W_k)\,.$$

Obviously this inclusion probability increases strictly with the number of waves until the snowball is saturated. If the initial sample is a convenience sample, its inclusion probability is 1 or 0 depending on whether j belongs to it or not. According to the results derived above for the conditional distribution of the next wave W_{k+1} given the current state (S_k, W_k), the conditional probability that a unit j not in S_k should be included in W_{k+1} is given by

$$p(W_k, j) = 1 - \prod(1 - p_{ij})\,,$$

where the product extends over units i in W_k. Therefore

$$P(j \in W_{k+1}) = \mathrm{E}\ [1 - I(j \in S_k)]p(W_k, j)\,,$$

where the expectation is over outcomes of (S_k, W_k). The probability of inclusion of unit j in S_{k+1} is obtainable as the sum of the probabilities of inclusion in the waves, but usually it is easier to first consider the conditional probability of inclusion of j in S_{k+1} given the current state (S_k, W_k). This probability is given by $p(S_k, j)$ if we define $p(S_k, j) = 1$ for $j \in S_k$. This is achieved by defining $p_{jj} = 1$ for $j \in U$. Hence, the unconditional inclusion

probability is given by

$$P(j \in S_{k+1}) = \mathrm{E}\ p(S_k, j) = 1 - \mathrm{E} \prod_i [1 - I(j \in S_k) p_{ij}],$$

where the expectation is over outcomes of S_k. Even if the initial sample and all the recruitment samples are independent Bernoulli samples, it doesn't follow that S_k is a Bernoulli sample, and we cannot carry out the expectation factor by factor to obtain a recurrence relation between single unit inclusion probabilities of the current and the next snowball. In fact, it can be shown that for distinct i and j in U

$$\pi_{ij}(k) \geq \pi_i(k)\pi_j(k)\,,$$

where $\pi_j(k) = P(j \in S_k)$ and $\pi_{ij}(k) = P(i \in S_k \ \& \ j \in S_k)$ for $k = 0, 1, \ldots$. Equality holds true if and only if there are no units of positive probability of inclusion in S_0, which are k steps adjacent before both i and j. An approximation that deserves to be further investigated for snowballs with few waves is

$$\pi_j(k+1) = 1 - \prod_i [1 - \pi_i(k) p_{ij}]\,,$$

or, equivalently, using logarithms

$$\log[1 - \pi_j(k+1)] = \sum_i \log[1 - \pi_i(k) p_{ij}]\,.$$

When the inclusion probabilities are not too close to 1, we can obtain a linear approximation to the recurrence relation by series expansions of the logarithmic functions:

$$\pi_j(k+1) = \sum_i \pi_i(k) p_{ij}\,.$$

Since the inclusion probabilities are fundamental for many estimators, it would be of interest to know if there are sampling designs for which this recurrence relation is an acceptable approximation.

Consider a one-wave snowball with Bernoulli samples as described above. There are several possibilities for data collection that need to be distinguished. Assume that the value of a vertex variable x and the degree are observed for sample units. If the units are identifiable and individual recruitments are observed, data consist of (i, x_i, y_i) for $i \in S_0$ and (j, x_j, y_j) for $j \in R_i$ and $i \in S_0$. Units have multiplicities 0 or 1 in S_0, but their multiplicities in S_1 might be larger and that applies also to units in S_0 if they are recruited in the first wave. If only collective recruitments are observed, data consist of (i, x_i, y_i) for $i \in S_0$ and (j, x_j, y_j) for $j \in R(S_0)$. In this case, multiplicities in S_1 can be 0, 1, or 2, and 2 applies only to units in S_0. If only collective recruitments not included in S_0 are observed, data consist of (i, x_i, y_i) for $i \in S_0$ and (j, x_j, y_j) for $j \in W_1$. Now multiplicities are never larger than 1. If units are not separated between S_0 and W_1, data consist of (i, x_i, y_i) for $i \in S_1$. Also in this case multiplicities are never larger than 1. If the units are not identifiable and individual recruitments are observed, data consist of (x_i, y_i) for $i \in S_0$ and (x_j, y_j) for $j \in R_i$ and $i \in S_0$ with no possibility to observe multiplicities and no possibility to reduce data to $R(S_0)$, W_1, or S_1. If only collective recruitments are observed, we cannot separate S_0 and W_1 and we don't know S_1 or any multiplicities. If only collective recruitments of units not included in S_0 are observed, we can separate S_0 and W_1 and we know S_1 and we know that no multiplicities are larger than 1. Without identities and without separation of S_0 and W_1, data consist of (x_j, y_j) for $j \in S_1$. Depending on the observation scheme available, different estimators can be used.

To illustrate various estimators, it is convenient to use snowball sample indicators $S_{kj} = I(j \in S_k)$ for $j \in U$ and $k = 0, 1, \ldots$ and recruitment indicators $z_{ij} = I(j \in R_i)$ for $i \in U$ and $j \in U$. According to our assumptions S_{0j} are Bernoulli(p_j) for $j \in U$, and z_{ij} are Bernoulli(p_{ij}) for $i \in U$ and $j \in U$, and all these variables are independent. Here $0 \leq p_j \leq 1$ for $j \in U$, and by allowing the values 0 and 1 initial convenience samples are possible. Recruitments are restricted by the presence of edges in the acquaintance graph, which implies that $0 \leq p_{ij} \leq y_{ij}$ for distinct i and j in U. For convenience we define $p_{jj} = 1$ for $j \in U$. Note that even if the acquaintance graph is undirected with $y_{ij} = y_{ji}$, the recruitment graph given by arc indicators z_{ij} is directed.

Consider the total $T = \sum x_i$ of a vertex variable x and the total $D = \sum y_i$ of degrees in the acquaintance graph. Let us specify $p_j = y_j/D$ and $p_{ij} = y_{ij}/y_i$ for $i \neq j$. Assuming D known and using only data from S_0, we have an unbiased estimator of T given by

$$T' = \sum_i (x_i S_{0i}/p_i) = \sum_i (x_i S_{0i} D/y_i)\,.$$

Similarly the total population size N has an unbiased estimator

$$N' = \sum_i (S_{0i} D/y_i)\,.$$

Without assuming D known, we have an asymptotically unbiased estimator of the population mean of x according to

$$T'/N' = \sum_i (x_i S_{0i}/y_i) \Big/ \sum_i (S_{0i}/y_i)\,.$$

Thus, the average of x-values in S_0 is weighted by reciprocal degrees.

Using individual recruitment data, an unbiased estimator of T is given by

$$T'' = \sum_j \left(x_j \sum_i S_{0i} z_{ij} \Big/ \sum_i p_i p_{ij} \right) = \sum_i \sum_j (x_j S_{0i} z_{ij} D / y_j) \, .$$

Also here the degree total D cancels from the asymptotically unbiased estimator of the population mean of x. We get

$$T''/N'' = \sum_i \sum_j (x_j S_{0i} z_{ij} / y_j) \Big/ \sum_i \sum_j (S_{0i} z_{ij} / y_j) \, .$$

Thus, the average of x-values in S_1 is counted with multiplicities and weighted by reciprocal degrees.

Using collective recruitments and only data from S_1, an unbiased estimator of T is given by

$$T''' = \sum_j [x_j S_{1j} / \pi_j(1)] \, ,$$

where

$$\pi_j(1) = 1 - \prod_i (1 - p_i p_{ij}) = 1 - \prod_i (1 - y_{ij}/D) = 1 - (1 - 1/D)^{y_i} \, .$$

This time D does not appear as a factor in T''' or in the corresponding N''', so it does not cancel in the population mean estimator T'''/N'''. Replacing x_j by y_j in the estimators T', T'', and T''', we obtain D', D'', and D''' that all depend on D. The first two wouldn't provide any possibility to iteratively determine a value of a D-estimator. For D''' an iterative method would start with an initial value of D in $\pi_j(1)$, find D''', use this D''' for D in $\pi_j(1)$, find a new D''', etc. Convergence properties of such algorithms and properties of estimators that can be obtained from them haven't been explored. We will comment on similar iterative methods in Sect. "Estimation Based on Network Walk Samples".

When data from snowballs with more than one wave are available, it might be comparatively easy to set up estimators based on a first few waves, but more complicated to find the required inclusion probabilities for snowballs of many waves. An important technique described in [64] is to use Rao-Blackwellization to improve estimators based on data from only a few waves to obtain estimators based on all data. Rao-Blackwellization of an estimator requires extensive numerical calculations of its possible values for all arrangements of available data that are consistent with the reduced data remaining when multiplicities of sample units and order between the sample units are ignored. The expected value of an estimator conditional on a sufficient statistic in the form of the reduced data is the improved estimator. Sufficiency in survey sampling from a finite population was treated in [1]. More recent references are [9,44].

Network Walk Sampling Designs

Assume that the population $U = \{1, \ldots, N\}$ has a graph structure given by a binary variable y defined on the set of ordered pairs of population units. Let A_i and B_i be the sets of units adjacent after and before unit i. The size of A_i is the out-degree of unit i, and the size of B_i is the in-degree of unit i. When the graph is undirected, $A_i = B_i$ and out- and in-degrees are equal and called degrees.

A sampling design based on a random walk in a graph selects the sample as a sequence of units $u_1, u_2, \ldots$ according to a time-homogeneous Markov chain with transition probabilities

$$P(u_{k+1} = j | u_k = i) = P_{ij} \, ,$$

that are independent of the stage parameter k for j in A_i and $k = 1, 2, \ldots$ If the graph is directed and strongly connected so that every unit has a directed path to every other unit, and if not all cycle lengths are multiples of a common factor larger than one, then the Markov chain will be irreducible and aperiodic. If the graph is undirected, connected, and contains at least one cycle of length one or three, then the Markov chain will be irreducible and aperiodic. A time-homogeneous, irreducible, and aperiodic Markov chain has a unique limiting distribution

$$\lim P(u_k = j | u_1 = i) = p_j > 0 \, ,$$

independent of the initial unit i as k tends to infinity. The limiting distribution is equal to the stationary distribution that can be obtained by solving the equation system

$$p_j = \sum p_i P_{ij} \quad \text{for} \quad j = 1, \ldots, N \, ,$$

subject to $\sum p_j = 1$. For instance, let the transition probability be $P_{ij} = y_{ij}/y_i$ for an undirected graph with degree y_i of unit i. This means that the transition from unit i is made with equal probabilities to each one of the y_i units that are adjacent to it. It follows that $p_i = y_i/(y_1 + \cdots + y_N)$. Hence, transition from any unit to one of its adjacent units according to a uniform probability distribution implies that the limiting probabilities of the units are proportional to their degrees.

If the limiting probabilities p_i for the units i in U cannot be determined from known or observable data, ergodic theory proves that they can be estimated by the relative frequencies in a long sample sequence obtained according to any time-homogeneous, irreducible, and aperiodic Markov chain. To be more specific, consider a sequence $u = (u_1, \ldots, u_m)$ of length m. The multiplicities of the units are m_i for i in U. The relative multiplicity sequence $(m_1/m, \ldots, m_N/m)$ converges towards the limiting distribution $(p_1, \ldots, p_N)$ when m tends to infinity.

Let $n(u)$ be the size of the set $s(u)$ of distinct units in u. The probability π_i that unit i is included in $s(u)$ equals the probability that $m_i > 0$. The multiplicity sequence is a reduction of the information in the sample sequence u that ignores the order of selection. The set $s(u)$ is a further reduction that ignores both the order of selection and the multiplicities of the units selected.

The graph employed for random walk sampling determines the zero entries among the transition probabilities $P = (P_{ij})$. These network-induced transition probabilities determine the limiting stationary distribution $p = (p_1, \ldots, p_N)$. If the initial unit u_1 is selected according to the stationary distribution p, this distribution applies to the marginal distribution of any position u_k of the walk. Another initial distribution implies that the stationary distribution is approached asymptotically on the walk for positions u_k as k increases. The limiting stationary distribution is important for appropriate weighting of data collected by network walk samples. Therefore, it is of special interest that it is possible to achieve any pre-assigned limiting distribution by adjusting the transition probabilities P induced by the network.

Assume that it is desirable to obtain a limiting distribution $q = (q_1, \ldots, q_N)$ that is different from the stationary distribution p derived from the transition probabilities P. Modified new transition probabilities $Q = (Q_{ij})$ satisfying the reversibility condition $q_i Q_{ij} = q_j Q_{ji}$ would lead to the stationary distribution q, which is also the limiting distribution when transitions are made according to the new Q. If we define Q by

$$Q_{ij} = P_{ij} \min(1, q_j P_{ji} / q_i P_{ij}) ,$$

for $i \neq j$ and

$$Q_{ii} = 1 - \sum_{i \neq j} Q_{ij} ,$$

it follows that

$$q_i Q_{ij} = \min(q_i P_{ij}, q_j P_{ji}) = q_j Q_{ji} ,$$

i. e. the reversibility condition is satisfied. To implement transitions according to Q, we can tentatively apply P and then decide to accept or reject a suggested transition from i to j with probability Q_{ij}/P_{ij} and $1 - Q_{ij}/P_{ij}$, respectively. More specifically, if the walk at stage k is in position $u_k = i$, and the next transition according to P would lead to unit j, the investigator might have to collect information from j to be able to calculate P_{ji} and compare it with P_{ij}. If $q_j P_{ji} \geq q_i P_{ij}$, then the transition to j has a probability valid also according to Q, and the walk is allowed to jump to $u_{k+1} = j$. Otherwise, the transition to j has not a valid probability according to Q, and, unless a transition is accepted, the walk has to stay at $u_k = i$ for another trial. Equivalently we could put $u_{k+1} = i$ and let the stage parameter count the number of trials instead of the number of jumps to other positions. Note that the probability that the modified walk stays for another trial equals the sum of the probabilities that the walk according to P jumps to a unit to be rejected, which is

$$\sum_{i \neq j} P_{ij}(1 - Q_{ij}/P_{ij}) = 1 - \sum_{i \neq j} Q_{ij} .$$

Consider the following illustrations for a directed network with transition probabilities $P_{ij} = y_{ij}/a_i$ for $j \in A_i$. Here the stationary probabilities can be difficult to find. Assume that we want a stationary distribution with probabilities q proportional to the out-degrees: $q_j = a_j/(a_1 + \cdots + a_N)$. The modified transition probabilities are

$$Q_{ij} = P_{ij} \min(1, q_j P_{ji} / q_i P_{ij}) = \min(y_{ij}, y_{ji})/a_i ,$$

and only jumps to mutual contacts are accepted. We need to collect information about out-degrees and be able to verify mutuality. Assume instead that we want a stationary distribution proportional to in-degrees: $q_j = b_j/(b_1 + \cdots + b_N)$. Then

$$Q_{ij} = (y_{ij} y_{ji}/b_i) \min(b_i/a_i, b_j/a_j) .$$

Only jumps to mutual contacts are possible, and only those with $b_i/a_i \leq b_j/a_j$ have transition probabilities that are valid also according to Q. Thus, unit j can be rejected unless it is a mutual contact with at least the same ratio of out-degree to in-degree as unit i. In this case we need to collect information about out- and in-degrees and be able to verify mutuality. Finally, assume that a uniform limiting distribution $q_j = 1/N$ is wanted. Then

$$Q_{ij} = \min(P_{ij}, P_{ji}) = y_{ij} y_{ji} / \max(a_i, a_j) ,$$

and j can be rejected unless it is a mutual contact that has at most the same out-degree as unit i. This case requires information about out-degrees and mutuality. Further illustrations are provided in [65]. A general presentation

of methods for constructing transition kernels (probabilities or probability densities) with desired properties, including the Metropolis–Hastings algorithm and the Gibbs sampler, can be found in Sect. 3.6 of [39].

The random walk sampling discussed so far has been in undirected graphs that are connected or in directed graphs that are strongly connected. If the connectedness assumption is uncertain for the graph being employed, or if it is known that it is not valid, then the Markov chain is reducible, and the transitions cannot reach beyond the connected component induced by the initial sample. In order not to confine the chain to a part of the population only, transitions to adjacent units could sometimes be replaced by transitions to other units.

A way to control the balance between staying on the current walk and starting on a new one, is to use a mixture distribution for the transitions. Let $P' = (P'_{ij})$ be a matrix of probabilities for transitions to adjacent units, and let $P'' = (P''_{ij})$ be a matrix of probabilities for transitions to arbitrary units in the population. A mixture distribution with transition probabilities given by

$$P = \alpha P' + (1 - \alpha)P'' ,$$

where $0 < \alpha < 1$, is said to have a damping factor α measuring the proportion of transitions that are selected according to P'. For instance, $P'_{ij} = y_{ij}/a_i$ for j in A_i and $P''_{ij} = 1/N$ for j in U represent a mixture between uniform transitions from i to A_i and uniform transitions from i to U. The walk stops when it reaches a unit with $a_i = 0$. A slight modification is to define P'_{ij} also when $a_i = 0$, so that the walk is never forced to stop. For α close to 1, the damping effect will be small, and a population with a disconnected graph will be poorly represented by walk sample data. For decreasing α, the damping effect increases, the influence of the graph decreases, and the importance of having an appropriate matrix P'' increases. It might be appropriate to assume, for instance, that global transitions are made with probabilities proportional to out-degrees or in-degrees of the units, which would reflect that activity or attraction of the units is likely to influence the selections. In this way, the damping effect needed to handle a disconnected graph does not necessarily diminish the influence of the graph.

When network walk sampling is applied to webpage transitions in [70], a model is used that in our notation can be given as $P'_{ij} = y_{ij}/a_i$ if $a_i > 0$ and $P'_{ij} = 1/N$ if $a_i = 0$, and $P''_{ij} = p''_j$. Here p''_j reflects the web surfer's preferences for various pages, and the damping factor α reflects tendencies to use page links for surfing rather than own preferences.

Estimation Based on Network Walk Samples

Assume that sample data consist of the values of a variable x defined on the population of units. When units in the sample sequence can be identified, sample data consists of m values of x that can be referred to the units. The values (i, x_i, m_i) are known for all units i in $s(u)$. When units in the sample sequence cannot be identified, sample data consist of m values of x that cannot be referred to the units. Neither the multiplicities nor the set of distinct units can be determined. Sample data are restricted to the frequency distribution of the x-values.

There is an intermediate situation that could appear when sample data consist of the values of a variable x defined on U. If units in the sample sequence cannot be identified, it could still be possible to successively mark them in some way and observe when the same unit is sampled again. Data then consist of a sequence of m values on x that cannot be referred to the units but that can be partitioned into n sub-sequences of constant values belonging to the same unknown unit. Notice that the same value might appear in different sub-sequences. The multiplicities are known but their unit affiliations are not, and the number of distinct units is known but not their identities. The values (x_i, m_i) are known for n unknown distinct units i in U. If x is a K-category variable, the frequency distribution of these values (x_i, m_i) can be displayed in a K by m array representing available sample data when identities of units can not be observed but similarity between units is observed. Let $(n_1, \ldots, n_m)$ be the marginal frequency distribution of the multiplicities. Then

$$n_1 + n_2 + \cdots + n_m = n ,$$

and

$$n_1 + 2n_2 + \cdots + mn_m = m .$$

Knowledge of the frequencies $(n_1, \ldots, n_m)$ represents a reduction of the information in the multiplicity sequence that ignores the identities of the units. The marginal frequency distribution of the K-category variable x is the only information left when neither similarities nor identities between units can be observed.

Consider a network walk sample sequence $u = (u_1, \ldots, u_m)$ obtained by a time-homogeneous, irreducible, and aperiodic Markov chain with transition probabilities $P_{ij} = y_{ij}/a_i$ for $j \in A_i$ and initial probabilities equal to the stationary probabilities $p_j = y_j/(y_1 + \cdots + y_N)$ for $j \in U$. The following reasoning applies asymptotically for large sample sequences even if the initial unit is not selected according to the stationary distribution. Data consist of the values of a vertex variable x and the degrees for

each unit in the sample sequence u, but the units need not be identified. The population mean of x is $\mu = T/N$ where $T = x_1 + \cdots + x_N$. The mean of x in the sample sequence is

$$\sum_{i\in u}(x_i/m) = \sum_{i\in U}(m_i x_i/m)\,,$$

where m_i is the unknown number of units equal to i in u. Since

$$\mathrm{E}\ m_i = mp_i = my_i/(y_1 + \cdots + y_N)\,,$$

the sample mean is biased as an estimator of the population mean μ unless all degrees are equal. Now

$$T' = \sum_{i\in u}(x_i/mp_i) = \sum_{i\in U}(m_i x_i/mp_i)\,,$$

has expected value T, and

$$N' = \sum_{i\in u}(1/mp_i) = \sum_{i\in U}(m_i/mp_i)\,,$$

has expected value N. Moreover,

$$\mu' = T'/N' = \sum_{i\in u}(x_i/y_i)\Big/\sum_{i\in u}(1/y_i)\,,$$

has asymptotically expected value μ for large values on m. The quantities T' and N' are not available from data since they contain the unknown total $D = y_1 + \cdots + y_N$ of all degrees. However, this total cancels in the ratio $\mu' = T'/N'$ so μ' is an asymptotically unbiased estimator of μ. We note that the estimator μ' is a weighted mean of the values of x in the sample sequence u with weights that are inversely proportional to the degrees.

The results obtained can be applied with $x_i = y_i$ to get an estimator of the average degree D/N in the population. Its estimator is

$$D'/N' = m\Big/\sum_{i\in u}(1/y_i)\,,$$

which is the harmonic mean of the degrees in the sample sequence. The results can also be applied with $x_i = I(i \in U_1)$, i. e. with x as an indicator variable for a specific subset U_1 of U of unknown size N_1. Then $T = N_1$ and the relative size of U_1 has an asymptotically unbiased estimator

$$N_1'/N' = \sum_{i\in u}(x_i/y_i)\Big/\sum_{i\in u}(1/y_i)\,,$$

which is the ratio between the sum of inverted degrees of units in the sample sequence that belong to U_1 and the sum of inverted degrees of all units in the sample sequence. The average degree in U_1 is given by

$$D_1/N_1 = \left[\sum_{i\in U} x_i y_i\right]\Big/\left[\sum_{i\in U} x_i\right]\,,$$

where $x_i = I(i \in U_1)$, and it can be estimated by

$$D_1'/N_1' = \left[\sum_{i\in u} x_i\right]\Big/\left[\sum_{i\in u}(x_i/y_i)\right]\,,$$

which is the harmonic mean of the degrees of the units in the sample sequence that belong to U_1. We note that the corresponding arithmetic mean is larger than or equal to the harmonic mean and therefore over-estimates the (arithmetic) population mean. This is because we over-sample units with large degrees. The bias is reduced by taking the harmonic sample mean.

If the units in the sample sequence are identifiable, it is possible to observe the multiplicities and the set of distinct sample units $s(u)$. Should the inclusion probabilities $\pi_i = P(i \in s(u))$ be known, this would allow the population total T to be estimated by the unbiased estimator

$$T'' = \sum_{i\in s(u)}(x_i/\pi_i)\,.$$

Similarly, N would have an unbiased estimator

$$N'' = \sum_{i\in s(u)}(1/\pi_i)\,,$$

and μ an asymptotically unbiased estimator

$$\mu'' = \left[\sum_{i\in s(u)}(x_i/\pi_i)\right]\Big/\left[\sum_{i\in s(u)}(1/\pi_i)\right]\,.$$

The inclusion probabilities $\pi_i = P(i \in s(u))$ are approximately $1 - \exp(-mp_i)$ and according to our assumptions about the Markov chain, the stationary probabilities are $p_j = y_j/D$ where $D = \sum y_j$ is the degree total. Now the average degree D/N was shown above to be estimated by the harmonic mean of the degrees in the sample sequence. Altogether, this allows us to set up the estimator N'' as a function of N, and by replacing N by N'' we might have a possibility to solve iteratively for N'' in the equations

$$N'' = \sum_{i\in s(u)}[1 - \exp 1/(-my_i/D')]\,,$$
$$D' = mN''\Big/\sum_{i\in u}(1/y_i)\,.$$

Possibilities of finding estimators iteratively, when the inclusion probabilities are functions of some of the parameters we want to estimate, deserve further study.

Methods for Estimating Hidden Populations

Hidden populations usually refer to populations of human individuals sharing a stigmatizing, illegal, risky, or embarrassing behavior that make them hard to find. There are no lists or frames to sample them from. Investigation of such populations might be of interest to society and necessary for determining appropriate resources to be allocated for social support and other types of social action. Standard survey sampling methods that require population frames cannot be used to study homeless people, drug users, sexual workers, individuals exposed to criminal or medical risks, etc. Various alternative approaches are discussed in [61].

Convenience samples of such hidden populations might be obtained by social field workers or by the police. The information gained from such sources could indicate the need for further study and data collection to obtain reliable inference about the hidden population. Members of the hidden population that are approached by social field workers might not be unwilling to co-operate and give away information about other members of the hidden population. Several practical methods described in the literature use various kinds of payments and other benefits to individuals in the hidden population that help the investigator to reach other members of the population for interview. When respondents are willing to reveal which other members they know of, the investigator can select the waves of a snowball sample. Respondents might be more willing to co-operate if they are asked to select at random one or two of the other members they know of. Thus, respondent driven random walk or small wave snowball sampling can be applied to the population network.

In cases when members of hidden populations are not willing to reveal themselves, it might still be possible to get information about them by applying network sampling. Assume for instance that the hidden population of interest is part of a larger population that can be sampled conventionally. A respondent sampled from the larger population might without revealing names or identities be able and willing to give some information about how many members of the hidden subpopulation she knows about. Such local network information about the hidden subpopulation can sometimes be used to infer information about the entire hidden subpopulation. As an illustrative example, consider the following setup considered in [41].

The hidden population U_1 consists of an unknown number N_1 of individuals. In order to estimate N_1, a larger population U of N individuals is defined so that it comprises U_1 as a subset; U_1 is embedded in U in such a way that another subset U_2 of U can be assumed to be equally common in U_1 and in U. The subset U_2 should consist of individuals having some easily identified characteristic. Neither the size N of U nor the size N_2 of U_2 need to be known. Select a sample S from U with positive inclusion probabilities π_i for all i in U. Each respondent is asked whether she belongs to U_2 and whether she has some acquaintances in U_2. She is not asked whether she belongs to U_1 but only whether she has some acquaintances in U_1. Let x_1 and x_2 be indicator variables of the sets U_1 and U_2, i. e. $x_{jk} = I(j \in U_k)$ for $j = 1, \ldots, N$ and $k = 1, 2$. Let further $y = (y_{ij})$ be the symmetric adjacency matrix of the acquaintance graph with degree y_i equal to the unknown number of acquaintances of individual i in U. This number is usually quite large and the respondent need not be asked to estimate it but rather to concentrate on giving reliable estimates of her numbers of acquaintances that belong to U_1 and U_2, say

$$a_{ik} = \sum_{j \in U} x_{jk} y_{ij} ,$$

for $i = 1, \ldots, N$ and $k = 1, 2$. If the independence assumptions are met for x_1, x_2, and y, then it could be expected under a randomization model that proportionality is valid according to

$$a_{ik}/y_i = N_k/N \quad \text{for} \quad k = 1, 2 \quad \text{so that} \quad a_{i1}/a_{i2} = N_1/N_2.$$

The sum of the numbers a_{ik} for $i \in U$ equals the degree sum in U_k, i. e.

$$\sum_{i \in U} a_{ik} = \sum_{j \in U} x_{jk} y_j = D_k .$$

The standard Horvitz–Thompson estimator of D_k is

$$D'_k = \sum_{i \in S} (a_{ik}/\pi_i) ,$$

for $k = 1, 2$. Similarly Horvitz–Thompson estimators of N and N_2 are given by

$$N' = \sum_{i \in S} (1/\pi_i) \quad \text{and} \quad N'_2 = \sum_{i \in S} (x_{i2}/\pi_i) .$$

From the proportionality assumption we get that the average degrees are equal in U, U_1, and U_2:

$$D/N = \sum_{i \in U} y_i/N = D_k/N_k ,$$

for $k = 1, 2$. Hence, N_1 can be estimated by

$$N'_1 = D'_1 N'_2 / D'_2 .$$

By summing $a_{i1}/a_{i2} = N_1/N_2$ over i in U_2 and over i in U we get

$$\sum_{i\in U}(a_{i1}x_{i2}/a_{i2}) = N_1 \quad \text{and} \quad \sum_{i\in U}(a_{i1}/a_{i2}) = NN_1/N_2\,,$$

which leads to two alternative estimators of N_1, namely

$$N_1'' = \sum_{i\in S}(a_{i1}x_{i2}/a_{i2}\pi_i)$$

$$\text{and} \quad N_1''' = (N_2'/N')\sum_{i\in S}(a_{i1}/a_{i2}\pi_i)\,.$$

Here the sums are understood to be over sample units $i \in S$ with acquaintances in U_2. Should $a_{i2} = 0$ we expect $y_i = 0$ and $a_{i1} = 0$. If the three estimators N_1', N_1'', and N_1''' turn out to give very different estimates, the assumptions underlying the calculations are doubtful and another approach is needed.

This example illustrates that properties of hidden groups that are hard to access can still be estimated if the hidden group is embedded in a convenient network. While snowball sampling and random walk sampling use the network to successively extend an initial sample, the hidden group example uses randomization assumptions about stable expected average network degrees as a possible bridge between different subgroup sizes. If the network assumptions needed for the randomization procedure are not valid, then a natural alternative approach would be to combine the probability design of the sampling with a probability model of the network.

An entirely different kind of hidden population are animal populations. In wildlife surveys the technique of catch-recapture sampling is an attempt to handle the lack of population frames. Bird counting, for instance, relies on marking the birds caught in a first sample by identifying rings before they are released again. After some time a new sample is taken, and the number of marked birds in the first sample is divided by the proportion of marked birds in the second sample to get an estimator of the size of the bird population. This idea can be elaborated on by specifying more carefully various assumptions about how the population changes with time and how the risk of being caught depends on different environmental factors that can be controlled for. Catch-recapture techniques have also been discussed as a tool in criminology surveys to estimate the number of offenders of specific types of crimes in different regions and in different periods of time. Such surveys cannot be carried out without a thorough model specification as a complement to the sampling design. In Sect. "Bipartite Network Sampling Designs" an illustration is given of utilizing network models and network sampling methods in a criminological setting.

Area sampling is another technique that can be used for counting animal populations and other populations distributed over geographical areas. A population region is partitioned into a number of area sites. Neighboring sites can be considered as adjacent vertices in a graph. The size of the area population is a vertex variable, and the total population size is a vertex variable total that can be estimated by a snowball sample of area sites with counts of their area populations.

Probabilistic Network Models

Survey sampling in finite populations distinguishes between design-based and model-based inference. Design-based inference uses probabilistic sampling designs but considers population data as fixed, while model-based inference also makes probabilistic assumptions about population data. Essential properties of population data might be summarized in a few parameters in an appropriate model and thereby allow a convenient and simplified description of the population. The graph employed in a network sampling design can be exposed to various unknown influences and uncontrolled effects that make it natural to try to capture its essential features in a random graph model. Random graph models can also be used to describe the outcome when a graph is sampled from a population of graphs, or when an observed graph is generated by some complex process, for instance, being an instant realization of a network changing with time. Random graphs useful for statistical data analysis need to be sufficiently rich in parameters so that meaningful fits can be made. However, too many degrees of freedom in the model make it vulnerable to accidental fits with no explanatory power.

A simple Bernoulli graph on $U = \{1, \ldots, N\}$ with a symmetric adjacency matrix $y = (y_{ij})$ has $N(N-1)/2$ independent edge indicators y_{ij} that are Bernoulli(p)-distributed for $1 \le i < j \le N$. The only parameter p represents edge density. A block model variant of the Bernoulli graph has a K-category vertex variable x and assumes y_{ij} to be independent and Bernoulli distributed with a parameter $p(x_i, x_j) = p(x_j, x_i)$ for $1 \le i < j \le N$. Now there are $(K+1)K/2$ parameters representing edge densities within and between different blocks of vertices. Depending on whether block affiliation is known, can be observed, or is not observable, the possibilities to estimate the parameters are very different. Statistical inference for stochastic block models is considered in many papers. An extensive account with many references can be found in [67].

For a directed graph on U with adjacency matrix y there are $N(N-1)/2$ dyads (y_{ij}, y_{ji}) for $1 \le i < j \le N$. The simple model with IID dyads has four parameters

$p_{00}, p_{01}, p_{10}, p_{11}$ that are probabilities $p_{kl} = P(y_{ij} = k, y_{ji} = l)$, and the probabilities sum to 1 leaving 3 degrees of freedom. More generally, with independent dyads and probabilities $P_{ijkl} = P(y_{ij} = k, y_{ji} = l)$, there are $2N(N-1)$ parameters subject to $N(N-1)/2$ restrictions $\sum_k \sum_l P_{ijkl} = 1$, which leaves $3N(N-1)/2$ degrees of freedom. Assuming a block structure with a K-category vertex variable x and $P_{ijkl} = p_{kl}(x_i, x_j) = p_{lk}(x_j, x_i)$ there are 2 degrees of freedom within each category of size larger than one and 3 degrees of freedom between any pair of distinct categories, in total $2K + 3K(K-1)/2 = K(3K+1)/2$ degrees of freedom. In particular, $x_i = i$ for $i = 1, \ldots, N$ implies that each category consists of only one vertex, and there are 3 degrees of freedom for each pair of vertices, i. e. $3N(N-1)/2$ degrees of freedom in total in accordance with the previous result. Thus, a categorization of 20 vertices based on two binary vertex variables reduces the degrees of freedom from 570 to 26, and a simple dichotomy reduces the degrees of freedom to 7, so block models represent a substantial simplification of dyad independence models. A practical approach might be to start with a rather large number of tentative vertex variables and apply cluster analysis to their dyad distributions in order to find how many of these distributions need to be distinguished. This technique is applied in [31,32].

The block model approach can be considered even if the category affiliation is uncertain and described by independent vertex variables x_i with a probability distribution $P(x_i = k) = p_k$ for $k = 1, \ldots, K$. Let N_k be the number of vertices in category k and N_{kl} the number of pairs of vertices in categories k and l, i. e. $N_{kl} = N_k N_l$ for $k \neq l$ and $N_{kk} = N_k(N_k - 1)/2$. The log-likelihood function when vertex variables and graph adjacencies are observed is given by

$$\log L = \sum_{k=1,\ldots,K} N_k \log p_k + \sum_{1 \le k \le l \le K} \sum_{m=0,1} \sum_{n=0,1} N_{klmn} \log p_{mn}(k,l) ,$$

where

$$N_{klmn} = \sum_{1 \le i < j \le N} \sum I(x_i = k, x_j = l, y_{ij} = m, y_{ji} = n) .$$

This model has $K-1$ degrees of freedom in addition to the previous $K(3K+1)/2$, so in total $3(K+1)K/2 - 1$ degrees of freedom. This model is treated for instance in [45,58,69].

Consider now the model with independent dyads and probabilities $P_{ijkl} = P(y_{ij} = k, y_{ji} = l)$. A convenient way to re-parametrize this model is to introduce odds and an odds ratio. Let a_{ij} be the odds of y_{ij} when $y_{ji} = 0$, and let b_{ij} be the odds of y_{ij} when $y_{ji} = 1$:

$$a_{ij} = P_{ij10}/P_{ij00}, b_{ij} = P_{ij11}/P_{ij01} .$$

The ratio between these odds is

$$c_{ij} = b_{ij}/a_{ij} = P_{ij11}P_{ij00}/P_{ij01}P_{ij10} .$$

The corresponding odds of y_{ji} are

$$a_{ji} = P_{ji10}/P_{ji00} = P_{ij01}/P_{ij00} ,$$
$$b_{ji} = P_{ji11}/P_{ji01} = P_{ij11}/P_{ij10} ,$$

and the odds ratio is the same

$$c_{ji} = b_{ji}/a_{ji} = c_{ij} .$$

The three parameters a_{ij}, a_{ji}, c_{ij} are arbitrary non-negative numbers, and together with the normalizing constant

$$P_{ij00} = 1/(1 + a_{ij} + a_{ji} + a_{ij}a_{ji}c_{ij}) ,$$

we retain the other probabilities as

$$P_{ij10} = a_{ij}P_{ij00} , \quad P_{ij01} = a_{ji}P_{ij00} ,$$
$$P_{ij11} = a_{ij}a_{ji}c_{ij}P_{ij00} .$$

Finally we put

$$\lambda_{ij} = -\log P_{ij00} , \quad \alpha_{ij} = -\log a_{ij} , \quad \gamma_{ij} = -\log c_{ij} ,$$

where $\lambda_{ij} = \lambda_{ji} > 0$ and $\gamma_{ij} = \gamma_{ji}$. Using the three parameters $\alpha_{ij}, \alpha_{ji}, \gamma_{ij}$ and the normalizing constant λ_{ij} determined by them, it is possible to express the likelihood function L according to

$$-\log L = \sum_{1 \le i < j \le N} \sum (\lambda_{ij} + \alpha_{ij}y_{ij} + \alpha_{ji}y_{ji} + \gamma_{ij}y_{ij}y_{ji}) .$$

There are $3N(N-1)/2$ degrees of freedom represented by $\alpha_{ij}, \alpha_{ji}, \gamma_{ij}$ for $1 \le i < j \le N$. By introducing an out-degree effect α_i, an in-degree effect β_i, and a mutuality effect γ_i for each vertex, we assume that

$$\alpha_{ij} = \alpha_i + \beta_j , \gamma_{ij} = \gamma_i + \gamma_j .$$

Since $\alpha_{ij} = \alpha + \beta + \alpha_i - \alpha + \beta_j - \beta$ for arbitrary α and β, α_i and β_j are identifiable only up to a translation

$\alpha = -\beta$. If we introduce an arc density effect δ and put $\alpha_{ij} = \delta + \alpha_i + \beta_j$ we can impose the restrictions $\sum \alpha_i = 0$ and $\sum \beta_j = 0$. In this way we have reduced the degrees of freedom from $3N(N-1)/2$ to $3N-1$, and the likelihood L is a log-linear function of the parameters given by

$$-\log L = \sum_{1 \le i < j \le N} \sum \lambda_{ij} + \delta \sum_{i \ne j} \sum y_{ij} + \sum_i \alpha_i \sum_j y_{ij} + \sum_j \beta_j \sum_i y_{ij} + \sum_i \gamma_i \sum_j y_{ij} y_{ji} ,$$

with out-degree, in-degree, and number of mutual arcs at each vertex as the sufficient statistics. A further reduction is obtained by assuming all mutuality effects equal, $\gamma_i = \gamma$, which leads to a model with $2N$ degrees of freedom. This is the original Holland–Leinhardt model for directed graphs introduced and investigated in [36]. Various extensions to valued graphs and block models are possible, and algorithms for model fitting and parameter estimation are described and illustrated in the reference book [67] and in several references given there.

Random graph models with probability distributions of exponential type have been suggested with more elaborate network statistics. Some examples of investigations of such models are [47,52,57,68].

Instead of specifying which statistics are likely to be essential for the application considered and use an exponential type model with these statistics as sufficient statistics, Frank and Strauss in [30] derived the sufficient statistics from basic assumptions about the dependence structure in the adjacency matrix. They specified a Markov dependence structure for the adjacency matrix y and proved that it implied a log-linear likelihood with certain sufficient statistics. The Markov assumption is that y_{ij} and y_{kl} are independent whenever i, j, k, l are all distinct, but dependence is possible between any arc indicators with some common vertex. The sufficient statistics are then given by star and triangle statistics. Adding a homogeneity assumption saying that isomorphic structures have the same probability, the sufficient statistics are given by star counts and triangle counts. Here stars and triangles should be understood in a broad sense so that they can consist of dyads of any type. The fundamental tool for this result in [30] is a theorem of Besag given in [2] together with a dependence graph specifying the dependencies among arc indicators in y. This dependence graph is in fact the so-called line-graph of the graph given by y. The maximal cliques of the dependence graph are the minimal sufficient statistics of y. Further exploration of these ideas can be found in the thesis [50] and the article [51].

Bipartite Network Sampling Designs

A bipartite network consists of vertices of two kinds and vertex and edge variables defined on them. If U is a set of N vertices of the first kind, and V is a set of M vertices of the second kind, different vertex variables could be defined on U and V. Edge variables are defined between U and V. As a simple example, take U as a set of households, V as a set of individuals, and define a binary edge variable y_{ij} that indicates whether individual j belongs to household i for $i \in U$ and $j \in V$. Each individual belongs to exactly one household. Assume that U is a target population, and that we are interested in some variable x defined on U. Only a sample S from V can be obtained. For each individual sampled, the value of x is reported that belongs to the individual's household. Thus

$$\sum_{i \in U} x_i y_{ij} ,$$

is reported by $j \in S$. If the sampling design has inclusion probabilities

$$\pi_j = P(j \in S) ,$$

the household total $T = \sum_{i \in U} x_i$ has a design unbiased estimator

$$T' = \sum_{j \in S} \sum_{i \in U} x_i y_{ij} / \pi_j = \sum_{i \in U} x_i \sum_{j \in S} y_{ij} / \pi_j .$$

If more than one individual from the same household are sampled, the estimator contains this household value x with multiplicity. If household identities are checked and no repeated household values should be used, another design unbiased estimator is available as

$$T'' = \sum_{i \in B(S)} x_i / P(i \in B(S)) ,$$

where $B(S)$ is the set of distinct households reported by individuals in S. The probability $P(i \in B(S))$ is equal to $P(S \cap A_i \ne \emptyset)$, where A_i is the set of individuals belonging to household i. This is the probability that none of those individuals are included in the sample S. If we introduce multiplicities m_i for the number of individuals in S belonging to household i, there is an alternative design unbiased estimator

$$T''' = \sum_{i \in U} x_i m_i / \mathrm{E} m_i ,$$

that requires the expected value of the multiplicity. Comparisons between these estimators and further results on

bipartite network sampling can be found in Sirken's pioneering paper on multiplicity sampling [56].

The possibility to use bipartite networks in survey sampling is not confined to cases with sampling units that are different from the units in the target population. As an illustration of a more complex combination of sampling from both U and V, we consider in the next section a model approach to a simultaneous estimation of crime rates and offender activity that relies on a bipartite network model.

A Bipartite Network Model for Crime Participation

Consider crimes of a particular kind in a specified geographical region during a specified period of time, say assaults in Canada during year 2000. Let M denote the unknown total number of incidents and m the number known to the police. Thus, $M - m$ is an unknown so-called "dark figure" of incidents, and m is the number of identified incidents. The identified incidents involve r incidents reported to the police and $m - r$ incidents identified by other means. Let m_i be the number of identified incidents having i identified offenders, and let r_i be the number of reported incidents having i identified offenders for $i = 0, 1, \ldots$ Incidents with at least one identified offender are said to be cleared by the police. There are unidentified offenders for all the unidentified incidents, and there are unidentified co-offenders for the identified incidents. The unknown numbers of incidents M_{ij} with i actual offenders and j identified offenders for $j = 0, 1, \ldots, i$ and $i = 1, 2, \ldots$ involve many dark numbers in addition to $M - m$. The unknown numbers of incidents

$$M_i = \sum_{j=0,\ldots,i} M_{ij} ,$$

with i actual offenders for $i = 1, 2, \ldots$ provide a partition of M that specifies the distribution of incidents according to size, i. e. their number of co-offenders.

The same offender can be involved in several offences. Let N denote the unknown total number of distinct individuals participating in the M offences. The known number n of distinct offenders identified in the cleared incidents imply that there is a dark number of $N - n$ non-identified offenders. The number of incidents in which an offender participates is called the activity of the offender. We need to distinguish between identified activity and actual activity. Let N_{ij} be the number of offenders with activity i and identified activity j for $j = 0, 1, \ldots, i$ and $i = 1, 2, \ldots$. The partition of N into the numbers

$$N_i = \sum_{j=0,\ldots,i} N_{ij} ,$$

of offenders of activity i for $i = 1, 2, \ldots$ specifies the distribution of offenders according to activity. This distribution involves further dark numbers that cause its discrepancy from the known numbers n_j of identified offenders participating in j identified offences for $j = 1, 2, \ldots$

In order to estimate the distribution of incidents according to size and the distribution of offenders according to activity, a bipartite network model can be defined with the set U of M incidents as vertices of the first kind and the set V of N offenders as vertices of the second kind. Criminal participation can be specified by an edge from $i \in U$ to $j \in V$ if incident i involves offender j. Let A_i be the subset of V specifying the co-offenders of incident i, and let B_j be the subset of U specifying the re-offences of offender j.

Now we can consider the set of identified incidents as a sample S from U, and the set of identified co-offenders of identified incident $i \in S$ as a sample S_i from A_i. These samples have sampling designs depending on factors governing how incidents are reported to the police or identified by other means and how the co-offenders are identified. In a situation like this when sampling is not determined by the investigator it is natural to incorporate the sampling design in a probabilistic network model.

An approach taken in [28] specifies the incidents in U as generated by a homogeneous Poisson process with parameter λ representing the expected number of incidents per time unit. The sample S of identified incidents is a Bernoulli sample with inclusion probability α, which means that identified incidents occur according to a homogeneous Poisson process with parameter $\alpha\lambda$. The parameter α measures the proportion of identified incidents. At least for crimes with a victim, α can be estimated by comparing incidence reporting frequencies and incidence identifications by other means with incidence frequencies estimated from victimization studies.

The set V of offenders is assumed to consist of a fixed number N of individuals that have a positive probability of offending. Independently for each incident $i \in U$, a set A_i of co-offenders in that incident is selected from V according to some common probability sampling design with inclusion probabilities $\pi_j > 0$ for $j \in V$. The sampling design might involve unknown parameters reflecting what is known about structure and behavior in the population V. For each identified incident $i \in S$, a set S_i of identified co-offenders is modeled as an independent Bernoulli sample of A_i with parameter β governing the chance that the police can identify an offender. An optimistic and a pessimistic view of this chance distinguishes between whether or not the identification of one co-offender in an incident should imply the identification of all co-offenders in that incident.

With this model, it turns out to be possible to estimate simultaneously both the distribution of incidents according to size and the distribution of offenders according to activity. The basic results are the following. From the distribution of identified size among identified incidents, it is possible to estimate the parameter β and the distribution of true size among all incidents. Furthermore, if θ is the average identified activity among identified offenders and $\gamma = n/N$ is the unknown proportion of identified offenders, then identified activity has a Poisson$(\gamma\theta)$ distribution that is truncated at zero, and true activity has a Poisson$(\gamma\theta/\alpha\beta)$ distribution. This implies that

$$\gamma = 1 - \exp(-\theta\gamma)\,,$$

from which γ can be solved, and N estimated by n/γ.

An interesting feature of this bipartite network model is that it makes explicit assumptions about both incident size and offender activity, and these assumptions lead to estimators of all the dark numbers related to unidentified incidents, unidentified offenders, and unidentified co-offenders in identified incidents. In particular, the estimator of the hidden population size N of offenders provides a special insight into how this estimator can be modified if not all parts of the distribution of identified activity are reliable. This also makes interesting connections with the literature that derives or suggests such modifications by using other approaches.

Future Directions

It is likely that the development of more sophisticated network models for specific applications will appear as improvements of Markov models and other network models belonging to the exponential family. Such development is also expected to contribute to more efficient algorithms for computer intensive estimation and goodness-of-fit testing of network models.

The developments of design-based inference methods that are appropriate for sampling of sub-networks from a finite population network are especially challenging for few-waves snowball sampling and for network walk sampling. The iterative estimation methods discussed in Sects. "Estimation Based on Snowball Samples" and "Estimation Based on Network Walk Samples" are examples of problems that deserve to be investigated further. More important is perhaps the need to develop efficient algorithms for Rao-Blackwellization of various initially obtained estimators and to gain more information about minimal sufficient statistics in various snowball sampling designs.

Bibliography

1. Basu D, Ghosh JK (1967) Sufficient statistics in sampling from a finite universe. Proc 36th Session Int Stat Inst, pp 850–859
2. Besag JE (1974) Spatial interaction and the statistical analysis of lattice systems. J Royal Stat Soc B 36:192–236
3. Bloemena AR (1964) Sampling from a graph. Math Centre Tracts, Amsterdam
4. Bollobas B (2001) Random graphs. Cambridge University Press, Cambridge
5. Brandes U, Erlebach T (eds) (2005) Network analysis. Springer, Berlin
6. Capobianco M (1970) Statistical inference in finite populations having structure. Trans New York Acad Sci 32:401–413
7. Capobianco M, Frank O (1982) Comparison of statistical graph-size estimators. J Stat Plan Inference 6:87–97
8. Carrington PJ, Scott J, Wasserman S (eds) (2005) Models and methods in social network analysis. Cambridge University Press, Cambridge
9. Cassel CM, Särndal CE, Wretman J (1993) Foundations of inference in survey sampling. Krieger Publ Comp, Malabar
10. Corander J (2000) On Bayesian graphical model determination. Dissertation, Stockholm University
11. Diestel R (2005) Graph theory. Springer, Berlin
12. Erdös P, Renyi A (1959) On random graphs I. Publ Math Debrecen 6:290–297
13. Erdös P, Renyi A (1960) On the evolution of random graphs. Publ Math Inst Hungar Acad Sci 5:17–61
14. Frank O (1969) Structure inference and stochastic graphs. Swedish Research Institute of National Defence, Stockholm, FOA-Reports 3(6):1–10
15. Frank O (1970) Sampling from overlapping subpopulations. Metrika 16:32–42
16. Frank O (1971) Statistical inference in graphs. Dissertation, Stockholm University
17. Frank O (1977) Estimation of graph totals. Scand J Stat 4:81–89
18. Frank O (1977) A note on Bernoulli sampling in graphs and Horvitz–Thompson estimation. Scand J Stat 4:178–180
19. Frank O (1977) Survey sampling in graphs. J Stat Plan Inference 1:235–264
20. Frank O (1978) Sampling and estimation in large social networks. Soc Netw 1:91–101
21. Frank O (1978) Estimation of the number of connected components in a graph by using a sampled subgraph. Scand J Stat 5:177–188
22. Frank O (1979) Estimating a graph from triad counts. J Stat Comput Simul 9:31–46
23. Frank O (1979) Estimation of population totals by use of snowball samples. In: Holland P, Leinhardt S (eds) Perspectives on Social Network Research. Academic Press, New York, pp 319–347
24. Frank O (1980) Sampling and inference in a population graph. Int Stat Rev 48:33–41
25. Frank O (1980) Estimation of the number of vertices of different degrees in a graph. J Stat Plan Inference 4:45–50
26. Frank O (1981) A survey of statistical methods for graph analysis. In: Leinhardt S (ed) Sociological methodology. Jossey-Bass, San Francisco, pp 110–155
27. Frank O (1987) Random sampling and social networks – a survey of various approaches. Math Inform Sci Hum 26(104):19–33

28. Frank O, Carrington PC (2007) Estimation of offending and co-offending using available data with model support. J Math Sociol 31:1–46
29. Frank O, Snijders T (1994) Estimating the size of hidden populations using snowball sampling. J Off Stat 10:53–67
30. Frank O, Strauss D (1986) Markov graphs. J Am Stat Assoc 81:832–842
31. Frank O, Komanska H, Widaman K (1985) Cluster analysis of dyad distributions in networks. J Classif 2:219–238
32. Frank O, Hallinan M, Nowicki K (1985) Clustering of dyad distributions as a tool in network modelling. J Math Sociol 11:47–64
33. Goodman LA (1961) Snowball sampling. Ann Math Stat 32:148–170
34. Granovetter M (1976) Network sampling: Some first steps. Am J Sociol 81:1287–1303
35. Hagberg J (2003) On degree variance in random graphs. Dissertation, Stockholm University
36. Holland PW, Leinhardt S (1981) An exponential family of probability distributions for directed graphs. J Am Stat Assoc 76:33–65
37. Janson S, Luczak T, Rucinski A (2000) Random graphs. Wiley, New York
38. Jansson I (1997) On statistical modeling of social networks. Dissertation, Stockholm University
39. Kaipio J, Somersalo E (2005) Statistical and computational inverse problems. Springer, New York
40. Karlberg M (1997) Triad count estimation and transitivity testing in graphs and digraphs. Dissertation, Stockholm University
41. Killworth PD, McCarty C, Bernard HR, Johnsen EC, Domini J, Shelley GA (2002) Two interpretations of reports of knowledge of subpopulation sizes. Soc Netw 25:141–160
42. Koskinen J (2004) Essays on Bayesian inference for social networks. Dissertation, Stockholm University
43. Morgan DL, Rytina S (1977) Comment on "Network sampling: Some first steps, by M Granovetter". Am J Sociol 83:722–727
44. Mukhopadhyay P (2001) Topics in survey sampling. Springer, New York
45. Nowicki K, Snijders TAB (2001) Estimation and prediction for stochastic blockstructures. J Am Stat Assoc 96:1077–1087
46. Palmer EM (1985) Graphical evolution. Wiley, New York
47. Pattison P, Wasserman S (1998) Logit models and logistic regressions for social networks. II: Multivariate relations. Br J Math Stat Psychol 52:169–193
48. Proctor CH (1979) Graph sampling compared to conventional sampling. In: Holland PW, Leinhardt S (eds) Perspective on social network research. Academic Press, New York, pp 301–318
49. Proctor CH, Loomis CP (1951) Analysis of sociometric data. In: Jahoda M, Deutsch M, Cook SW (eds) Research methods in social relations. Dreiden Press, New York, pp 561–586
50. Robins GL (1998) Personal attitudes in inter-personal contexts: Statistical models for individual characteristics and social relationships. Dissertation, University of Melbourne
51. Robins GL, Pattison P (2005) Interdependencies and social processes: Dependence graphs and generalized dependence structures. In: Carrington PJ, Scott J, Wasserman S (eds) Models and methods in social network analysis. Cambridge University Press, Cambridge, pp 192–214
52. Robins GL, Pattison P, Wasserman S (1999) Logit models and logistic regressions for social networks. III: Valued relations. Psychometrika 64:371–394
53. Salganik MJ, Heckathorn DD (2004) Sampling and estimation in hidden populations using respondent-driven sampling. Sociol Methodol 34:193–239
54. Särndal CE, Swensson B, Wretman J (1992) Model assisted survey sampling. Springer, New York
55. Schweinberger M (2007) Statistical methods for studying the evolution of networks and behaviour. Dissertation, University of Groningen
56. Sirken MG (1970) Household surveys with multiplicity. J Am Stat Assoc 63:257–266
57. Snijders TAB (2002) Markov chain Monte Carlo estimation of exponential random graph models. J Soc Struct 3(2)
58. Snijders TAB, Nowicki K (1997) Estimation and prediction of stochastic blockmodels for graphs with latent block structure. J Classif 14:75–100
59. Spreen M (1999) Sampling personal network structures: Statistical inference in Ego-graphs. Dissertation, University of Groningen
60. Stephan FF (1969) Three extensions of sample survey technique. In: Johnson NL, Smith H Jr (eds) New developments in survey sampling. Wiley, New York
61. Sudman S, Kalton G (1986) New developments in the sampling of special populations. Ann Rev Sociol 12:401–429
62. Tallberg C (2003) Bayesian and other statistical approaches for analyzing network block-structures. Dissertation, Stockholm University
63. Thompson ME (1997) Theory of sample surveys. Chapman & Hall, London
64. Thompson SK (2006) Adaptive web sampling. Biometrics 62(4):1224–1234
65. Thompson SK (2006) Targeted random walk designs. Surv Methodol 32:11–24
66. Thompson S, Frank O (2000) Model-Based Estimation with Link-Tracing Sampling Designs. Surv Methodol 26:87–98
67. Wasserman S, Faust K (1994) Social network analysis. Cambridge University Press, Cambridge
68. Wasserman S, Pattison P (1996) Logit models and logistic regressions for social networks. I: An introduction to Markov random graphs and p*. Psychometrika 60:401–426
69. Wellman B, Frank O, Espinoza V, Lundquist S, Wilson C (1991) Integrating individual, relational and structural analysis. Soc Netw 13:223–249
70. Wills RS (2006) Google's page rank: The math behind the search engine. Math Intell 28(4):6–11

Social Network Analysis, Graph Theoretical Approaches to

WOUTER DE NOOY
University of Amsterdam, Amsterdam, The Netherlands

Article Outline

Glossary
Definition of the Subject
Introduction

Cohesion
Centrality and Brokerage
Prestige and Ranking
Analyzing Complexity in Social Networks
Future Directions
Bibliography

Glossary

Adjacent Two vertices are adjacent if they are connected by a line.

Arc An arc is a directed line. Formally, an arc is an ordered pair of vertices.

Attribute An attribute is a characteristic of a vertex measured independently of the network.

Bipartite network See: Two-mode network.

Clique A clique is a maximal complete subnetwork containing three vertices or more.

Complete A (sub)network is complete if it has maximum density: All possible lines occur.

Component A (weak) component is a maximal (weakly) connected subnetwork.

Degree The degree of a vertex is the number of lines incident with it.

Density Density of a simple network is the number of lines, expressed as a proportion of the maximum possible number of lines.

Digraph A digraph or directed graph is a graph containing one or more arcs.

Distance The distance from vertex u to vertex v is the length of the geodesic from u to v.

Edge An edge is an undirected line. Formally, an edge is an unordered pair of vertices.

Ego-network The ego-network of a vertex contains this vertex, its neighbors and all lines among the selected vertices.

Geodesic A geodesic is the shortest path between two vertices.

Graph A graph is a set of vertices and a set of lines between pairs of vertices.

Incident A line is defined by its two endpoints, which are the two vertices that are incident with the line.

Indegree The indegree of a vertex is the number of arcs it receives.

Line A line is a tie between two vertices in a network: Either an arc or an edge.

Loop A loop is a line that connects a vertex to itself.

Neighbor A vertex that is adjacent to another vertex is its neighbor.

Network A network consists of a graph and additional information on the vertices or the lines of the graph.

One-mode network In a one-mode network, each vertex can be related to each other vertex.

Outdegree The outdegree of a vertex is the number of arcs it sends.

Path A path is a semipath with the additional condition that none of its lines is an arc of which the end vertex is the arc's tail.

Reachable We say that a vertex is reachable from another vertex if there is a path from the latter to the former.

Semicycle A semicycle is a closed semipath ending at the vertex at which it starts.

Semipath A path is a closed sequence of lines such that the end vertex of one line is the starting vertex of the next line and no vertex in between the first and last vertex of the sequence occurs more than once.

Signed graph A signed graph is a graph in which each line carries either a positive or a negative sign.

Simple graph A simple undirected graph contains neither multiple edges nor loops. A simple directed graph does not contain multiple arcs.

Star-network A star-network is a network in which one vertex is connected to all other vertices but these vertices are not connected among themselves.

Strong component A strong component is a maximal subnetwork in which each pair of vertices is connected by a path.

Strongly connected A (sub)network is strongly connected if each pair of vertices is connected by a path.

Structural property A structural property is a characteristic (value) of a vertex that is a result of network analysis.

Triad A triad is a (sub)network consisting of three vertices.

Two-mode network In a two-mode network, vertices are divided into two sets and vertices can only be related to vertices in the other set.

Undirected graph An undirected graph does not contain arcs: All of its lines are edges.

Vertex (vertices) A vertex (singular of vertices) is the smallest unit in a network.

Weakly connected A (sub)network is weakly connected if each pair of vertices is connected by a semipath.

Definition of the Subject

Social network analysis (SNA) focuses on the structure of ties within a set of social entities or actors, e. g., persons, groups, organizations, and nations, or the products of human activity or cognition such as semantic concepts, web sites, and so on. In a graph theoretical approach, a social network is conceptualized as a graph, that is, a set of ver-

tices (or nodes, units, points) representing social actors and a set of lines representing one or more social relations among them.

A network, however, is more than a graph because it contains additional information on the vertices and lines. Characteristics of the social actors, for instance a person's sex, age, or income, are represented by discrete or continuous attributes of the vertices in the network, and the intensity, frequency, valence, or type of social relation are represented by line weight or value, line sign, or line type. Formally (see pp. 94–95, 127–128 in [1]), a network $\mathbf{N}$ can be defined as $\mathbf{N} = (U, L, F_U, F_L)$ containing a graph $\mathbf{G} = (U, L)$, which is an ordered pair of a unit or vertex set U and a line set L, extended with a function F_U specifying a property of the units ($f: U \rightarrow X$) and a function F_L specifying a property of the lines ($f: L \rightarrow Y$). The set of lines L may be regarded as the union of a set of undirected edges E and a set of directed arcs $A (L = E \cup A)$. Each element e of E is an unordered pair of units u and v (vertices) from U, that is, $e(u: v)$, and each element a of A is an ordered pair of units u and v (vertices) from U, that is, $a(u: v)$.

The application of graph theory to social relations can be traced back to at least the 1940s (see pp. 69–72 in [2]) when the mathematician R. Duncan Luce and the engineer Albert Perry teamed up with the social psychologist Leon Festinger [3] and when the mathematician Frank Harary started his collaboration with Leon Festinger and afterwards with Dorwin Cartwright [4]. They extended pioneering work in SNA that had been done notably in sociometry [5,6] and anthropology [7,8,9]. In the 1960s, advances in graph theoretical approaches to SNA such as the contributions by Øystein Ore [10], Claude Flament [11], Frank Harary [12], and innovative applications such as Everett M. Rogers' work on the diffusion of innovations [13], prepared the ground for the rise of SNA in both the USA [14] and Europe [15] as a new set of methods or a new methodology [16] in the 1970s.

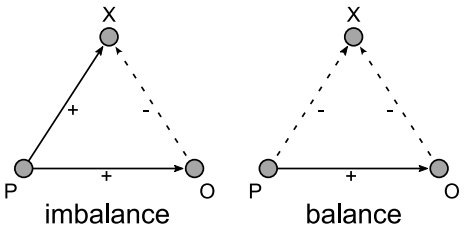
Introduction

The conceptualization of social systems as graphs and networks offered the opportunity for systematic investigation and theorizing of the structure of ties among social actors beyond the pair. Whereas classical sociology tended to make a quantum leap from the individual and the pair to the triple, group, or society [17], graph theory offered the tools to formally describe and visualize social structure consisting of three and more actors. This led to a new awareness of social structure as a system of ties that is both the product of human action and the context and condition for human action. Because this point of view is very relevant to the issue of complexity in social networks, it is briefly presented in the next paragraph.

The prevalent action theory in SNA conceptualizes collective behavior as the socially 'orchestrated' behavior of individuals or other actors. Actors adjust their behavior and attitudes, opinions, and beliefs to the behavior (etc.) of other members of the social system in which they participate. The system itself is not supposed to behave but it constrains actor behavior: It is the social context within which actors operate. As a network of ties, the system defines to whom an actor is exposed. The immediate contacts – the neighbors in graph theory – of an actor are usually most important to its behavior, but indirect contacts such as the neighbors' neighbors may be taken into account as well. In other words, an actor's local context or ego-network is likely to affect its behavior. At the same time, however, by ending ties or creating new ones, the individual changes both local network structure and overall network structure, that is, the system. Thus, individual action changes the local context for its neighbors' (neighbors' etc.) action. Complexity arises in the interplay between individual behavior and the system both as the overall structure of the network of social ties and as the local context for each actor within the system. To the actors, the change of network structure is not necessarily predictable, so the interplay between individual action and network structure may offer surprising results.

Let us turn to an example now, which is one of the earliest applications of graph theory to social networks. This example nicely illustrates both the transition from a focus on the tie within a single pair to the study of group structure in the social sciences and the interplay between local action and overall network structure. In 1946, the psychologist Fritz Heider formulated the theory of psychological balance [18], which stated among other things that a person (P) feels uncomfortable when he or she disagrees with his or her friend (O) on a particular topic (X). Person P is hypothesized to be stressed and to try to change this situation either by adapting its opinion on topic X, so it matches O's opinion, or by changing its opinion on O, regarding him or her no longer as a friend. Figure 1 represents both a situation of imbalance and a situation of balance as a signed digraph. The circles and arrows represent the vertices and arcs of the graph and the valence of the opinions or affections are shown both by the labels and the style of the arrows: Solid arcs show positive opinions or sentiments, dotted arcs show negative opinions or sentiments.

In 1956, the mathematician Frank Harary and the psychologist Dorwin Cartwright realized that psychological



Social Network Analysis, Graph Theoretical Approaches to, Figure 1
The principle of imbalance and balance conceptualized as signed digraphs

balance in this triad (three vertices and the lines among them) may be conceptualized as a specific pattern of arcs in a signed graph, viz., in a balanced triad, the $P - O - X$ semicycle (a closed semipath) always contains zero or an even number of negative arcs, whereas an imbalanced triad is characterized by a semicycle with an uneven number of negative arcs [4]. As a next step, replacing the topic by a third person and perceptions of liking or disliking (by the focal person *P*) by expressed liking or disliking as ties, they easily generalized this idea to a network of arbitrary size. They proved that a signed network is balanced if and only if all semicycles contain no or an even number of negative arcs.

In addition, they proved that a balanced network either contains one set of vertices with just positive arcs among them, or two sets of vertices with all positive arcs within the sets and all negative arcs between the sets, which is a polarized network. In 1967, this result was generalized to polarization among three or more groups by James A. Davis, who showed that a network can be partitioned into an arbitrary number of subsets of vertices such that all positive ties are within the subsets and all negative arcs are among the subsets if the network does not contain semicycles with exactly one negative line [19]. Figure 2 shows an example from Samuel F. Sampson's [20] study of a network of sentiments among novices in a monastery. It depicts the situation at the fourth measurement wave, which was highly polarized at that time. Vertex color indicates whether the novice had previously attended one particular seminary (black: Yes, white: No).

In this case, a mathematical relation was established between a behavioral hypothesis at the level of the individual, viz., adjusting ones affect relations such that ones situation is balanced, and overall network structure, viz., polarization. Individual behavior proved to have an unexpected outcome at the level of overall network structure and graph theoretical concepts, in this case semicycles, provided the link. In many cases, however, the link between individual behavior and overall network structure has not been established formally and is sometimes even hard to predict intuitively.

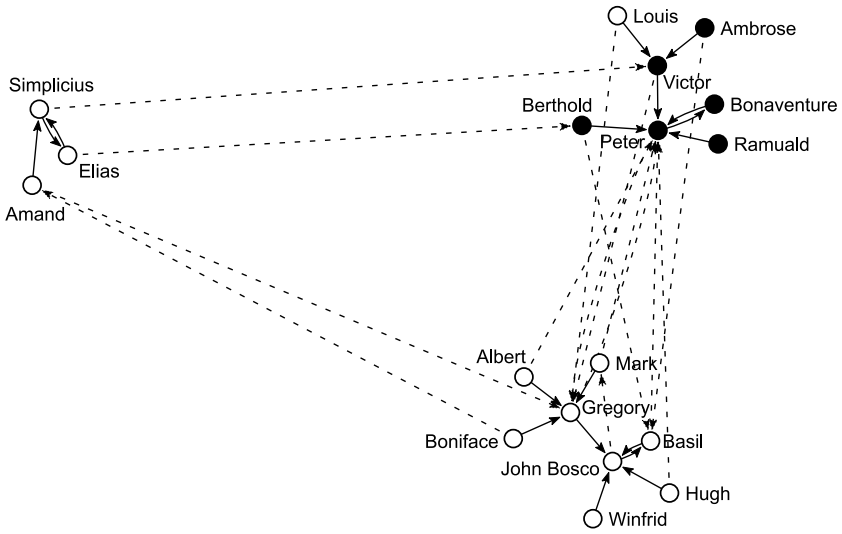
This entry of the encyclopedia aims to present an overview of graph theoretical approaches to SNA, highlighting the complex relations between individual action and overall network structure. It intends to explain why current developments focus on local structure rather than overall network structure to unravel the complexity of social networks. For each of the main topics in SNA (cohesion, brokerage, and prestige), behavioral hypotheses are presented stating why actors create, maintain or dissolve ties. The typical local structure of ties produced by this behavior is sketched in combination with graph theoretical indices for measuring it, and finally the expected consequences to the overall structure of the social network are discussed in combination with the graph theoretical measures developed for measuring them. Note that this approach more or less reverses the historical development of SNA, which focused on overall network structure first and gradually became more interested in the behavior of actors that created, maintained, or changed network structure.

Cohesion

One of the first intuitions in SNA concerns the tendency of human beings to form cohesive subgroups. This is a classical topic in the social sciences, see, for instance, George C. Homans' book *The Human Group* [21], and it was central to the sociometry tradition [22]. But where do cohesive subgroups come from and what do they do?

The first and most general behavioral hypothesis merely states that similar people tend to interact more easily and people who interact tend to become or perceive themselves as more similar provided that the interaction is characterized as positive, friendly, and so on. In SNA, this tendency is mainly known as homophily, a concept coined by Paul F. Lazarsfeld and Robert K. Merton [23,24], but it is known under other names in several scientific disciplines, e. g., the phenomenon of attribution [25] and affect control [26,27] in social-psychology, assortative or selective mixing in epidemiology and ecology ([28], p. 2 in [29]), and assortative mating in genetics with efforts at statistical modeling at least as early as 1985 [30].

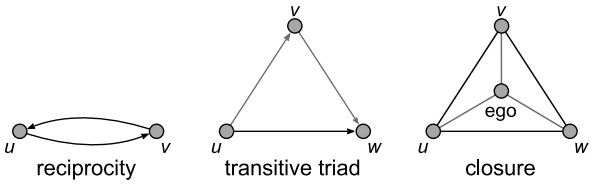
It is important to note that there are two sides to this behavioral hypothesis, a selection effect, that is, the impact of similarities on the ties that are created, sustained, or broken [31], and an influence effect [32], which hypothesizes that perceived or actual similarities such as the socially constructed identities or opinions [33] result from



Social Network Analysis, Graph Theoretical Approaches to, Figure 2
Almost perfect polarization in the network of sentiments among novices

the ties among actors. According to these hypotheses, people who are directly linked are or become more similar because of their interaction, so they become more likely to engage in ties and maintain ties among them. Thus, social groups form and persist as tightly linked sets of people that tend to share social and psychological characteristics, producing solidarity.

If we concentrate on the graph theoretical aspects of this behavioral hypothesis, that is, the structure of ties, and ignore the (dis)similarities among the actors for the moment, we find several characteristics of local structures that measure cohesive subgroup formation. At the level of a pair of actors, reciprocity of ties in directed networks signals subgroup formation: Both actors are hypothesized to choose each other when they are similar. At the level of the triple, transitivity results from tendencies toward cohesion. If actor u establishes a tie with actor v because they are similar, and actor v establishes a tie to actor w for the same reason, actors u and w are also similar, so actor u is expected to establish a tie with w as well, creating a so-called transitive triad (Fig. 3). Stated differently, the path from u via v to w is closed by an arc from u to w. In general, the closure of paths or semipaths both in directed and undirected networks signals cohesive subgroup formation at the local level. Closure within an ego-network may be calculated as the percentage of all possible ties among a vertex' neighbors that are present, which is one of the definitions of the clustering coefficient (see pp. 32–33 in [34]) but the concept of closure can be extended beyond a vertex' immediate neighbors, e. g., the number of semi-



Social Network Analysis, Graph Theoretical Approaches to, Figure 3
Reciprocity, transitivity, and closure in a directed ego-network

cycles of length 4 or larger in which a vertex is involved, e. g., balanced semicycles in signed networks.

If we include vertex attributes in our measures of cohesive subgroup formation, we can calculate homophily quite simply as the probability or ratio of ties between vertices that share a particular characteristic to ties between vertices that do not. Extending this idea to the ego-network, the homogeneity of actors involved in an ego-network may be taken as a measure of tendencies toward homophily. For qualitative attributes of the actors, Blau's index of variability or heterogeneity [35] can be used $(1 - \sum p_i^2)$ where p is the proportion of group members in a particular category and i is the number of different categories), which is conceptually related to the Herfindahl–Hirschman Index in economics measuring the extent of monopoly within an industry. It is interesting to note that Blau's theory hypothesizes that heterogeneity rather than homogeneity of actors within a group enhances the operation and efficiency of the group. If improving group

efficiency is the aim of actors, we would have to use a behavioral hypothesis that is quite the opposite of the homophily hypothesis.

Tendencies toward cohesive subgroup formation at the level of actor behavior are most likely to produce sets of densely connected vertices in the overall network and, if we add attributes of the actors, the vertices within sets tend to have similar characteristics. In the most extreme case, the sets are disconnected, so the network consists of several weak components, that is, maximal weakly connected subnetworks, or they are connected only by ties with a negative social meaning, as in the example of polarization presented in the Introduction.

In the history of SNA, the concept of relatively densely connected subnetworks has yielded a large number of graph theoretical ways for identifying cohesive subgroups at the level of overall network structure. Limiting the discussion to one-mode networks, that is networks in which there can be a tie between any pair of vertices (for two-mode or bipartite networks, ► Social Network Analysis, Two-Mode Concepts in), Wasserman and Faust (see pp. 251–252 in [36]) distinguish 4 approaches to defining cohesive subgroups.

In the first and strictest approach, a cohesive subgroup is defined as a set of vertices in which all vertices are adjacent, that is, directly linked, to one another. In other words, cohesive subgroups are maximal complete subgraphs, which are called cliques [3].

The second approach is based on the notion of reachability and closeness of members within a subgroup. Members of a subgroup must be reachable in the sense that there are paths between them, i. e., a sequence of lines such that the end vertex of one line is the starting vertex of the next line (following the direction of the lines if they are arcs) such that no vertex in between the first and last vertex occurs more than once. In addition, the shorter the geodesics (shortest paths) between them, the closer the vertices are in a graph theoretical sense, so the more cohesive the subgroup is supposed to be.

The criterion of reachability does not necessarily yield very dense subgroups. In sparse networks, any maximal connected subgraph (strong component) may represent a cohesive subgroup: There is a path between each pair of vertices within a component. Increasing the required number of independent paths between any pair of vertices within a cohesive subgroup yields slightly denser subgroups, e. g., requiring at least two independent paths produces bi-components, which may be generalized to k-components for higher minimum numbers of independent (vertex-disjoint) paths between all vertices within a subgroup [37].

Focusing on graph-theoretical distance between vertices usually yields denser subgroups. One may, for instance, set a maximum n to the distance between any two vertices within a subgroup, which is the concept of an n-clique [38,39]. Adding the restriction that the diameter of an n-clique is n or less, one obtains n-clans [38,40]. Alternatively, one may define a cohesive subgroup as a maximal subgraph of diameter n, which is called an n-club [40].

The third approach focuses on the minimum number, strength, or multiplicity of ties among subgroup members. Subgraphs that are maximal with respect to the minimum number of neighbors within the subgraph are called k-cores [41] and a maximal subgraph with respect to the maximum number of vertices in the subgraph that are not adjacent are known as k-plex [42]. In a similar vein, restrictions can be imposed on the minimum strength or multiplicity of ties among members of a cohesive subgroup, generalizing Seidman's concept of a k-core to a valued core, which is called an m-core (see pp. 115–116 in [43]) or m-slice (see pp. 109–110 in[44]): Maximal connected subgraphs considering only lines with minimum value (or multiplicity) m.

In the fourth approach, cohesive subgroups are based on the relative frequency of ties among subgroup members in comparison to non-members: Cohesive subgroups are relatively dense sections within the network, that is, relative to the sections outside (and between) subgroups. An *LS* set [45] is a maximal subgraph such that any of its subsets has more ties to its complement within the *LS* set than to vertices outside the *LS* set. Borgatti, Everett and Shirey [46] generalized this idea to the concept of the lambda set, which requires the number of edge-disjoint paths between any pair of vertices within the lambda set to be larger than between any vertex within and any vertex outside the lambda set. A probabilistic version of plus-clusters in signed networks, discussed in the Introduction, can also be subsumed under this approach as it requires relatively many positive lines within cohesive subgroups and relatively many negative lines among cohesive subgroups [47]. Finally, clustering techniques and some types of blockmodels (see pp. 741–742 in [14]) also detect clusters of vertices that have relatively many ties within clusters and few among clusters. These models offer an alternative way for finding cohesive subgroups (see pp. 133–246 in [1]), ► Positional Analysis and Blockmodeling.

The large number of alternative measures for cohesive subgroups attests to the fact that behavioral tendencies at the actor level do not play out into nicely structured overall networks in a standard way. Especially the density of the social relation under investigation has an impact on

the extent and ways in which cohesive subgroups can be found in the overall structure of the network. As a consequence, measures of network cohesion such as the clustering coefficient (averaged over all vertices in the network) can be quite uninformative about overall network structure: Loosely knit cohesive subgroups which are clearly identified by some of the techniques presented above, yield low clustering coefficients in a sparse network. Cohesion in a network is better summarized by calculating the percentage of vertices that are part of identified subgroups, the number and sizes of cohesive subgroups, and so on [48].

In addition to homophily, there is a second behavioral hypothesis related to cohesion in SNA. This hypothesis is based on the idea that social action is embedded in networks [49,50,51]. Named after the sociologist Georg Simmel, Simmelian ties are ties that are embedded in other ties, e. g., business ties are embedded in family ties, or in complete triads and cliques. They are hypothesized to enforce group norms and enhance trust, hence pressure people into the same behavior because the two actors involved in a tie share common neighbors who supervise their behavior.

Just as with the homophily hypothesis, the embeddedness hypothesis predicts that tightly connected actors will be more similar in their behavior and attitudes. In addition, it predicts that embedded ties are more stable and new ties are more likely to be established when they are embedded in existing cliques or existing ties. Closure again is an important indicator of tendencies to establish and maintain embedded ties but so is the multiplicity of relations: The extent to which a tie on one social relation duplicates a tie on another social relation. At the level of overall network structure, we should expect relatively dense sections, especially cliques, and in a multirelational network, that is, a network containing ties on two or more social relations, we should find that the same subsets of vertices are clustered on each relation. Graph-theoretical measures of the latter are rare. The stochastic blockmodeling technique developed by Krysztof Nowicki and Tom A.B. Snijders [52] is an example. See ► Social Networks, Algebraic Models for.

If data on vertex attributes are available, especially if they concern public behavior, that is, behavior that is easily noticed by third parties such as publicly expressed opinions and statements, Simmelian ties are hypothesized to produce a special effect. Involvement in different groups (cliques) then exposes actors to possibly conflicting sets of norms and loyalties, which may urge them to cut their ties with some or all of these groups [53]. In this case, actors are hypothesized to withdraw from stressful relations, so they discontinue ties that incorporate them into cliques (with a preference for cutting a minimum number of ties) or they discontinue ties such that they are no longer connected to actors voicing different opinions or norms. At the macro level, this would produce disconnected sets of cliques instead of overlapping cliques.

Centrality and Brokerage

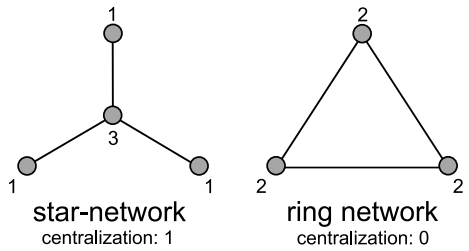
The notion of centrality in social networks has a long history in SNA. It is attributed to Alex Bavelas [54]. In discussions of centrality, network ties are usually regarded as channels for the exchange of information, goods, services, and so on. Being central in this exchange system has always been hypothesized to be related to influence and satisfaction. Centralization, as a characteristic of a network, has been linked to the efficiency of a network as an exchange system. More centralized groups, for example, have often been shown to be more efficient.

Linton C. Freeman [55] argued that the approaches to centrality are based on three ideas about what being central means: (1) being active within the network, that is, maintaining many ties, (2) being efficient or independent of go-betweens by having short distances to other vertices in the network, and (3) being an important go-between, that is, being part of many paths between other vertices in the network. Although alternative classifications and approaches exist, for instance, Noah E. Friedkin's alternative classification [56] and the formal graph theoretical approach to centrality by Stephen P. Borgatti and Martin G. Everett [57], Freeman's classification is used here, adding concepts of brokerage that have been developed elsewhere in SNA.

Activity

Being active or prominent in the network means that an actor has many ties, hence access to many sources of information (etc.). As a consequence, this actor is more attractive as a neighbor for other actors, which translates to the behavioral hypothesis that actors have a preference for ties with vertices that already have many ties. The degree of a vertex (the number of lines incident with a vertex), then, is the relevant graph theoretical measure of local structure, which is also known as degree centrality. Note that the concept of centrality expresses a structural property of a vertex.

Centralization is the corresponding structural property of a network and it is defined as the variation in the centrality scores of the vertices in the network because this variation shows the extent to which there is a center (very central vertices) and a periphery (vertices with very low centrality scores). The star and ring networks are defined



Social Network Analysis, Graph Theoretical Approaches to, Figure 4
Centrality and centralization in a star network and a ring network

as respectively the most and least centralized networks and they are known to exhibit the highest and lowest variation in centrality scores in simple networks, that is, networks without multiple lines (and loops in the case of a directed network). See Fig. 4 for an illustration, showing a star and a ring network labeling the vertices with their degree centrality scores.

From research on the power law in networks [58,59], discussed in another entry of this encyclopedia, it is known that preferential attachment to degree creates networks with a peculiar degree distribution, including many vertices with low or modest degree and few vertices with high degree. According to the definition of centralization in network analysis, this implies large variation in degree centrality scores, hence high degree centralization. In other words, the behavioral hypothesis of preference for high degree actors produces centralized overall network structure.

Note, however, that the way in which the 'power law networks' are assembled – growing from an initial seed without context – is hardly ever applicable to social networks, which usually have no discernable starting point (networks originate from networks) and are always constrained by the historical context. It remains to be seen whether power law distributions in empirical social networks are created by preferential attachment. Some results indicate that even though the degree distributions of cross-sectional snapshots of a large social network follow the power law, there is hardly any continuity in degree centrality of vertices over time, which does not suggest that the actors are driven by preferential attachment [60].

Efficiency and Weak Ties

The second approach to centrality focuses on graph theoretical distances between vertices. The central idea here is that actors try to improve access to information and efficient spreading of information by minimizing the number of go-betweens needed to reach or be reached by all other actors in the network. The behavioral hypothesis states that a vertex prefers to link to actors that give access to parts of the network that are presently remote to this vertex and that can only be reached through some or many go-betweens that may withhold or distort information. A minimum number of go-betweens yields maximum independence and maximum efficiency in the exchange network.

Graph theoretical measures of local structure focus on graph theoretical distance, that is, the minimum length of paths between vertices because path length equals the number of go-betweens in the network plus one. Linton Freeman's closeness centrality [55] is a straightforward implementation of this idea because it merely normalizes the average graph theoretical distance between a vertex and all other reachable vertices in the network. In addition, paths can be weighted by the centrality of the vertices on them, which is done by Phillip Bonacich's eigenvector centrality [61,62]. Finally, the difference between incoming paths and outgoing paths may be added, which is implemented in the version of closeness centrality developed by Thomas W. Valente and Robert K. Foreman [63].

Again, the normalized variation of closeness centrality scores of the vertices in the network yields the appropriate measure of centralization of overall network structure. It is not known yet, however, whether and how tendencies to reduce distances to other vertices at the level of individual actors in the network play out into the closeness centralization of the overall network. On the one hand, if the network contains some vertices with high closeness centrality, they offer rather short paths toward many vertices so they should attract a lot of new ties. This would enhance their centrality and possibly the variation in centrality scores although a general rise in closeness centrality scores of all vertices may also decrease the variation. On the other hand, if the network has low centralization, it is more likely that vertices connect directly to remote parts, reducing path lengths among remote parts, which would not raise the variation in closeness centrality scores and yield or sustain low closeness centralization of overall network structure.

The strength of weak ties argument proposed by Mark Granovetter [64,65] may be regarded as a special application of the notion of efficiency. In his research on finding a job, Granovetter noticed that relatively superficial ties, ties with infrequent contact, give access to new information because they are more likely to link you to someone with whom you are not linked directly or at a short distance. Strong or intense ties tend to be situated within cohesive subgroups, so they are more likely to offer redun-

dant information already received through other ties. Granovetter is only interested in the effects of having weak ties, but if we turn his idea into a behavioral hypothesis, it suggests a preference to relate to distant vertices that are neither connected to yourself nor to your neighbors (or your neighbors' neighbors, and so on).

A tendency to connect to the most remote parts of the network means that actors tend to establish links to vertices at a large or maximum distance in the network. As in the case of maximizing closeness centrality, it is not clear whether this leads to identifiable patterns in overall network structure. It is quite obvious, however, that this tendency acts as a counterforce against tendencies toward cohesion as densely connected subnetworks: Actors are hypothesized to span gaps rather than to close local configurations. Different hypotheses must be developed for strong ties, which are hypothesized to contribute to subgroup formation, and weak ties linking remote parts. The weak ties will probably increase the number of links between dense parts of the network, increasing the k-connectivity (minimum number of node-independent paths between any two vertices) of the network. If so, we should expect high k-connectivity of the network if the weak ties hypothesis is true [37].

Control and Structural Holes

The third approach to centrality focuses on control over flows within the network: The more you are in between other vertices in the network, the more they depend on you to pass on information, the more you are able to control exchange within the network and profit from your control. Using this type of control is called brokerage.

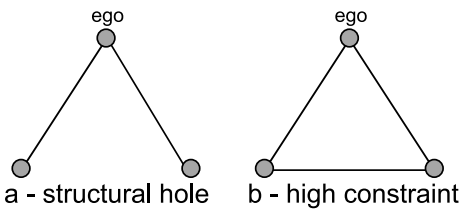
The notion of being in between other vertices has a straightforward translation to graph theory as being part of a path between two other vertices. Limiting paths to the shortest paths between vertices both Linton C. Freeman's [55] betweenness centrality and Jac M. Anthonisse's [66] rush of a vertex are based on the proportion of all geodesics between other vertices that include this vertex. This measure has been extended to handle directed ties [67,68]. Information centrality [69,70] takes into account all paths between vertices, not just the geodesics and flow betweenness [71] or entropy [72] also consider the values of lines. See Stephen P. Borgatti's [73] classification for more details.

Betweenness centralization as the normalized variation of betweenness centrality of the vertices in the network offers a measure of centralization at the network level and so do center-periphery blockmodels (see pp. 741–742 in [14]), e. g., in the world trade system, central countries are able to profit from the lack of trade among countries in the periphery [74,75,76,77]. The link between strategies for maximizing betweenness centrality by actors and the overall structure of the resulting network is unclear. It is not to be expected that control behavior will produce a highly centralized structure because no actor will be satisfied with low betweenness centrality and as a consequence high betweenness centrality for some vertices is unlikely. Therefore, the variation in betweenness centrality scores will not be high.

While betweenness centrality situates an actor with respect to all other actors in the network, Ronald S. Burt [78,79] proposed a local variant, focusing on control within the ego-network of an actor. His behavioral hypothesis rests on the *tertius gaudens* principle: The benefits that accrue to an actor that is in between two actors that are not directly linked because of the opportunity to broker information between them or, in a more malicious variant, to divide and conquer.

This hypothesis translates quite easily into graph theoretical structure. The absence of a tie between two neighbors of an actor is called a structural hole (Fig. 5a). The behavioral hypothesis states that actors try to increase the structural holes that they can exploit. At the same time, however, they try to minimize the structural holes through which they can be exploited. This means, among other things, that an actor will not end a tie to one of its neighbors if the two neighbors are directly linked (see Fig. 5b): That would create an opportunity to broker at the expense of the actor. In this situation, the actor is constrained in its opportunities to change ties.

Structural holes and constraint are the flip sides of the same coin. A tie with low constraint indicates that the tie is involved in (many) incomplete triads (such as Fig. 5a), so there are (many) structural holes offering the actor options for brokering. High constraint on a tie means that it is part of (many) complete triads (as in Fig. 5b), so there are few or no possibilities for brokerage. Because the presence or absence of ties among an actor's neighbors is key



Social Network Analysis, Graph Theoretical Approaches to, Figure 5
A structural hole (a) and a triad with high constraint (b)

to the argument, network analysts have also used the density of the ego-network without the ego as a proxy of constraint: The higher the density, the higher the constraint on the ego. Alternatively, betweenness centrality for ego-networks [80,81] can be used.

If the structural holes hypothesis governs actors' behavior, what overall network structure should we expect to find? A strong tendency toward brokerage at the micro level is not likely to produce a centralized network because each actor would try to maximize the number of structural holes around itself, which would yield a bipartite graph in the extreme case (no ties among any vertex' neighbors), or it would minimize its constraint by ending all ties to neighbors that have contacts outside ego's immediate neighborhood, which would produce a highly clustered network consisting of isolated cliques or isolated vertices in the extreme case.

In summary, the relation between local action and overall network structure is simple and clear only in the case of preferential attachment to high-degree vertices. Tendencies to maximize centrality that looks beyond the immediate neighbors such as closeness and betweenness centrality, do not necessarily yield centralized networks. Even for local structures, alternative hypotheses for actor's behavior are available that are unlikely to produce centralized networks, e. g., a preference to avoid constraint. The interplay between local action and overall structure is quite complex.

Prestige and Ranking

The preceding sections have not distinguished between directed and undirected networks. For prestige and ranking, however, the direction of ties is crucial because asymmetry in networks is assumed to be linked to social prestige [82]. The general idea here is that social inequalities are reflected and possibly created by asymmetric ties, e. g., everyone invites the most popular boy or girl in class but s/he doesn't return each invitation. Of course, the nature of the social relation determines the direction of choices; ties like "reports to" or "pays respect to" point toward higher levels in a hierarchy while "beating up" points in the opposite direction.

A central behavioral hypothesis concerns the popularity or attractiveness of actors. Actors tend to (want to) relate to actors with attractive structural properties or attributes, so attributes related to power or social status increase the probability that an actor will be chosen. From a constructivist point of view, however, being chosen often is also interpreted as a sign of importance and prestige, so receiving many choices (ties) increases the probability of receiving even more. In this way, networks may produce informal status hierarchies. The Matthew Effect, proposed by Robert K. Merton [83], comes to mind here: "For unto every one that hath shall be given, and he shall have abundance: But for him that hath not shall be taken away even that which he hath" (gospel of Matthew XXV, 29).

In graph theoretical terms, the structural attractiveness of an actor refers to the number of incoming arcs on vertices, which is simply the indegree of a vertex. This is called the popularity of a vertex and, of course, we must replace it by the vertex' outdegree if the relation is negative, e. g., submission, beating up, criticizing. If indirect choices must be taken into account as well, attractiveness is measured by proximity prestige [84], which is based on the average distance from all other vertices in the network – a directed variant of closeness centrality. Proximity prestige captures the idea that nominations or choices by actors who are themselves popular, contribute more to ones structural prestige. Bonacich's measure of power [62,85] adds the idea that power may also be derived from being connected to powerless actors rather than to other powerful people.

The popularity hypothesis is another example of preferential attachment to degree with the restriction that we focus either on indegree or outdegree. Therefore, the indegree (or outdegree) distribution of the overall network is expected to follow the power law and network structure will be characterized by high degree centralization.

Adding data on social attributes that make some actors more prestigious such as wealth, social class, beauty, and so on, we should expect preferential attachment to vertices that score high on these attributes. Note that vertex attributes play a slightly different role here than in the case of cohesion. Now we are concerned with attributes of (at least) ordinal level, expressing prestige that an actor possesses to a higher or lower degree. In the case of cohesion, we deal with nominal attributes, which merely express an identity. In contrast to homophily, the attributes of the actor who initiates the directed tie (the tail of the arc) does not matter here because the effect is solely related to structural characteristics or attributes of the actor at the receiving end of the tie (the arc's head).

There is a second, slightly different behavioral hypothesis relating to deference or submission rather than attractiveness. The idea is that actors mainly tend to create positive ties to other actors in their own status group or to actors in a higher status group – the people they are looking up to – to consolidate and improve their social position. Similarly, they tend to direct negative ties to actors in lower status groups. Note the difference with attractiveness: It is hypothesized that actors choose upwardly but

they need not prefer the most attractive (top) actors in the network as they are supposed to do according to the attractiveness hypothesis.

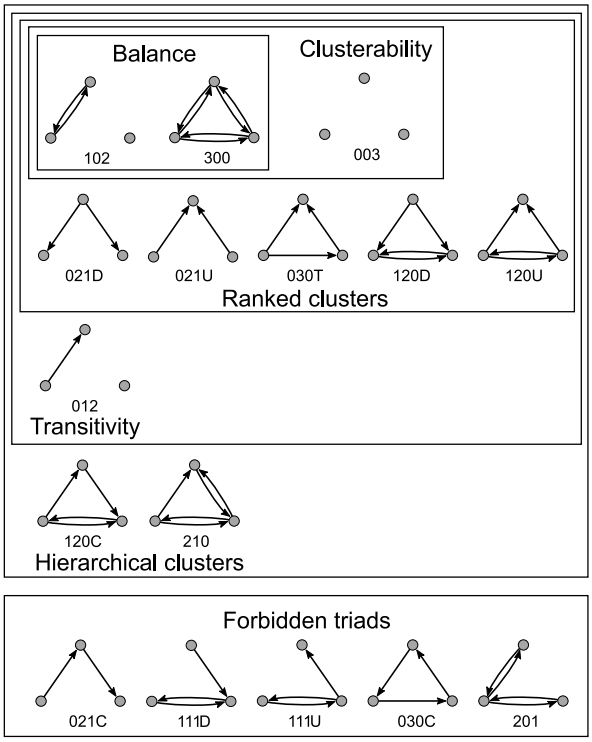
The main difference between attractiveness and deference is that the former only takes into account structural properties or attributes of the tie's receiver, whereas the characteristics of both the sender and receiver of the tie matter to the latter. This distinction has important consequences to the structure of the overall network. Whereas the attractiveness hypothesis yields centralized networks, the deference hypothesis yields layered networks. The layers consist of sets of vertices that are symmetrically linked, e. g., by reciprocal ties, while the ties between layers are asymmetric, all pointing in the same direction. This behavior may be both a consequence of a formal hierarchy, e. g., positions within an organization with formalized relations such as reports to, or actually show an informal social hierarchy, e. g., status differences between men and women in a particular social setting.

In simple directed networks, triads, that is, three vertices and the lines among them, are the key to measuring tendencies toward ranking at the local level. Translating the concepts of balance and clusterability (see Sect. "Introduction") from signed digraphs to unsigned digraphs, James A. Davis and Samuel Leinhardt [86,87] replaced positive ties by symmetric ties within cohesive subgroups and negative ties by no ties among subgroups. Thus, ties within a layer, either within clusters or among clusters, are symmetric. They assumed that asymmetric ties represent the ranking of clusters into a hierarchy, introducing the model of ranked clusters. Moreover, they showed that a network with a perfectly fitting ranked clusters model contains only certain types of triads, whereas other types do not occur (Fig. 6).

The ranked clusters model requires arcs from each vertex to all vertices on higher ranks. This requirement is usually too strict for empirical social networks and it is relaxed in the transitivity model proposed by Paul W. Holland and Samuel Leinhardt [88], which requires that clusters of vertices on different ranks are either completely linked or not linked at all, yielding a partial order, by simply adding one type of triad to the set of allowed triads (Fig. 6). Later, Eugene Johnsen [89] proposed the model of hierarchical clusters to account for asymmetries within clusters. Note the nesting of the models for overall network structure, which is why the sets of permitted triads is extended for more general models. Finally, it was shown that the models developed for unsigned digraphs could also be detected in incomplete signed digraphs using types of semicycles [90].

The triads characterizing ranked structures serve as models for tie creation, maintaining, and breaking behavior of actors under the deference hypothesis. For instance, triad 120U (Fig. 6) predicts that a member of a cluster is likely to establish or maintain a tie to an actor at a higher rank if its neighbors within the cluster have such a tie. In the perfect case, there is a one-to-one relation between sets of occurring types of triads and the overall structure of the network. Therefore, triad census [91,92], which is the frequency distribution of the sixteen types of triads in a directed network, offers an indication of overall network structure. In the imperfect case, the triad census of a network may be compared to the frequency distributions of triad types in randomly generated networks to test the tendency toward ranking.

The triad census does not show the composition of the clusters and ranks; it does not identify the vertices belonging to particular clusters and ranks. This can be done in several ways. Realizing that ranks should be connected asymmetrically in directed networks, strong components cannot include more than one rank because vertices within strong components are by definition mutually reachable. Ties between strong components, then, are asymmetric, so it is easy to establish the ranking among strong components. Strong components, however, do not



Social Network Analysis, Graph Theoretical Approaches to, Figure 6
Triad types and balance-theoretic models

require a lot of symmetry in the ties; actually, no tie needs to be reciprocated. The symmetric-acyclic decomposition proposed by Patrick Doreian, Vladimir Batagelj and Anuška Ferligoj [93] does require at least some symmetric ties (mutual choices) within clusters because they define a symmetric cluster as a maximum subset of vertices that are directly or indirectly linked by symmetric ties. Generalized blockmodels [1] that are asymmetric with respect to off-diagonal blocks offer another way to identify hierarchical relations (see ► Positional Analysis and Blockmodeling).

Analyzing Complexity in Social Networks

The preceding sections presented behavioral hypotheses that have similar, different, or even opposite consequences for overall network structure. It is not plausible that one particular type of behavior dominates network formation. Therefore, it is not likely that overall structure of empirical social networks will display one particular form that can be hypothesized in advance or that behavioral tendencies can be adequately tested on particular characteristics of overall network structure. It has been shown, for example, that the degree distribution of a network does not reveal tendencies toward cohesive subgroup formation that are operative during network evolution [94].

Even if overall network structure displays certain characteristics, they may be produced by different types of behavior. Centralization in a social network, for instance, may arise from a tendency of actors to minimize paths to all other actors or from a preference for prestigious actors. Alternatively, it may be a by-product of cohesive subgroup formation: Actors that are marginal to the cohesive subgroups may directly or indirectly connect different subgroups, which gives them a central position with respect to betweenness. Furthermore, pronounced overall network patterns may occur only temporarily in empirical social networks when they create socially unstable situations. The polarization predicted by balance theory in Sampson's network of novices, used as an example in the Introduction, was only temporary. After this polarization and most likely due to it, many novices left the monastery. The network, so to speak, fell apart.

For these reasons, SNA increasingly focuses on local structure using overall network structure merely as a collection of (overlapping) local structures. Behavioral hypotheses translate much more directly to local structure, that is, to the ties of the actor and those of its neighbors (and possibly their neighbors), as we have seen in the preceding sections. Local structure is the part of the network that an actor can easily survey and actually change.

The latest developments in techniques for modeling network structure and evolution apply this actor-oriented approach either in statistical models, see ► Network Analysis, Longitudinal Methods of and ► Social Networks, Exponential Random Graph (p^*) Models for. The techniques test behavioral tendencies by relating the creation, maintenance, and ending of ties by individual actors to the local configuration of ties, to previous ties between the actor and the alter or to present ties on another social relation, and to characteristics of both the actor and the alter.

In principle, the actor-oriented approach is able to test all behavioral hypotheses presented in the preceding sections on characteristics of local structure in which the actors are embedded and properties of the actors themselves. If hypothesized local configurations appear more often than expected by chance, the underlying behavioral hypothesis is assumed to guide individual behavior at least to some extent. If the behavior in a set of actors or individual actor's behavior are in line with several behavioral hypotheses at the same time, the effect of each behavioral tendency can be separated. Thus, it is possible to link complex overall network structure to compound behavior of the actors in the network.

Future Directions

The techniques for analyzing local structure are in development. Models for the co-evolution of relations and quantitative attributes of vertices over time have just been introduced [95]. Not all behavioral hypotheses have been included yet, for instance, because they involve non-standard types of networks such as signed relations, and new ones are bound to be proposed. Incomplete data and external constraints on data collection or conceptual constraints on network structure such as two-mode networks (see ► Social Network Analysis, Two-Mode Concepts in) may limit the applicability of current models and spur the development of new ones.

If the actor-oriented approach is successful, will overall network structure become completely redundant? Will it only serve network exploration – looking for behavioral hypotheses rather than testing them – and for analyzing the consequences of network position on behavior, attitudes, or esteem, for instance, does the subgroup to which an actor belongs or its centrality correlate with its subsequent behavior or attitudes?

Let us return to the balance example, presented in the Introduction. The high degree of polarization among Sampson's novices was followed by the voluntary or forced exit of several novices. This suggests that people are able to survey overall network structure and draw conclusions

from it, rather than just react to their local network environment. A highly polarized social group does not seem to be the kind of situation that we like to be living in. An actor's perception of overall network structure may also affect its behavior.

In this respect, the balance example is interesting because the original theory by Heider referred to perceptions of affect relations and unit relations (having characteristics or possessing items) rather than objectively measured relations. The importance of perception to network analysis is stressed in a more general way by Harrison White, who argues that social ties are stories [96]. According to him, people are linked into social networks by the stories that they tell about their ties. Thus, we should expect people to react to the ties as they remember and tell them rather than to the ties as they are observed by the researcher or registered in, for instance, membership lists.

A similar argument was made by David Krackhardt when he reconsidered balance theory [97]. According to him, we should measure each actor's perception of network structure, for which he proposed the concept of Cognitive Social Structures, compare these perceptions, and use them to explain why actors behave as they do. In his approach, actors are assumed to be able and active in forming impressions of overall network structure.

If humans are capable of imagining overall network structure, then communication of these perceptions may also play a role in network formation, including the accidental and deliberate distortions or simplifications that are likely to happen. This brings us to the link between network structure and mental categories such as social classifications or culture in a more general sense as argued by Ronald L. Breiger [98,99]. If, for example, members of the network perceive and discuss cohesive subgroups, they assign names and meanings to social configurations. Thus, social meanings and identities are created in the process of establishing social ties and interpreting them. The duality of social structure on the one hand and the structure of symbolic categories on the other hand as proposed by John W. Mohr [100], may be the essential condition for cultural meanings that are social in the sense that they affect actors' sense of identity and behavior.

These discrete and qualitative rather than continuous and quantitative classifications are very likely to affect the network behavior of actors: They define the categories that are experienced as being similar or dissimilar in the case of homophily and group formation, or superior versus inferior in the case of prestige and ranking. As a consequence, it is to be expected that the present focus on local structure will be complemented by a focus on overall network structure, especially perceived and communicated network structure.

Bibliography

Primary Literature

1. Doreian P, Batagelj V, Ferligoj A (2005) Generalized blockmodeling. Cambridge University Press, Cambridge
2. Freeman LC (2004) The development of social network analysis. A study in the sociology of science. Empirical, Vancouver
3. Luce RD, Perry A (1949) A method of matrix analysis of group structure. Psychometrika 14:95–116
4. Cartwright D, Harary F (1956) Structural balance: A generalisation of Heider's theory. Psychol Rev 63:277–293
5. Murphy G (1937) Editorial foreword. Sociometry 1:5–7
6. Moreno JL (1937) Sociometry in relation to other social sciences. Sociometry 1:206–219
7. Warner WL, Lunt PS (1941) The social life of a modern community. Yale University Press, New Haven
8. Davis A, Gardner BB, Gardner MR (1946) Deep south: A social anthropological study of caste and class. University of Chicago, Chicago
9. Mitchell JC (1969) The concept and use of social networks. In: Mitchell JC (ed) Social networks in urban settings. Manchester University Press, Manchester
10. Ore Ø (1963) Graphs and their uses. Mathematical Association of America, Washington
11. Flament C (1963) Applications of graph theory to group structure. Prentice-Hall, Englewood Cliffs
12. Harary F, Norman RZ, Cartwright D (1965) Structural models: An introduction to the theory of directed graphs. Wiley, New York
13. Rogers EM (1962) Diffusion of innovations. Free Press, New York
14. White HC, Boorman SA, Breiger RL (1976) Social structure from multiple networks. I. Blockmodels of roles and positions. Am J Sociol 81:730–780
15. Helmers HM, Mokken RJ, Plijter RC, Stokman FN, Anthonisse JM (1975) Graven naar macht: Op zoek naar de kern van de Nederlandse economie. Van Gennep, Amsterdam
16. Wellman B (1988) Structural analysis: From method and metaphor to theory and substance. In: Wellman B, Berkowitz SD (eds) Social structures: A network approach. Cambridge University Press, Cambridge, pp 19–61
17. Simmel G (1902) The number of members as determining the sociological form of the group. I. Am J Sociol 8:1–46
18. Heider F (1946) Attitudes and cognitive organization. J Psychol 21:107–112
19. Davis JA (1967) Clustering and structural balance in graphs. Hum Relat 20:181–187
20. Sampson SF (1968) A novitiate in a period of change. An experimental and case study of social relationships. Cornell University, Ithaca, p 575
21. Homans GC (1950) The human group. Harcourt Brace, New York
22. Moreno JL (1953) Who shall survive?: Foundations of sociometry, group psychotherapy and sociodrama, 1st edn 1934. Beacon House, Beacon
23. Lazarsfeld PF, Merton RK (1954) Friendship as a social process: A substantive and methodological analysis. In: Berger M, Abel

T, Page CH (eds) Freedom and control in modern society. Van Nostrand, Princeton, pp 18–66
24. McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: Homophily in social networks. Ann Rev Sociol 27:415–444
25. Festinger L (1962) A theory of cognitive dissonance, 1st edn 1957. Tavistock, London
26. Heise DR (1979) Understanding events: Affect and the construction of social action. Cambridge University Press, New York
27. Smith-Lovin L, Heise DR (1988) Analyzing social interaction: Advances in affect control theory. Gordon Breach Science, New York, pp 192
28. Morris M (1991) A log-linear modeling framework for selective mixing. Math Biosci 107:349–377
29. Newman MEJ (2003) Mixing patterns in networks. Phys Rev E 67
30. Tallis GM (1985) Transfer systems and covariance under assortative mating. Theor Appl Genet 70:497–504
31. Robins G, Elliott P, Pattison P (2001) Network models for social selection processes. Soc Netw 23:1–30
32. Robins G, Pattison P, Elliott P (2001) Network models for social influence processes. Psychometrika 66:161–189
33. Friedkin NE, Johnsen EC (1990) Social influence and opinions. J Math Sociol 15:193–205
34. Watts DJ (1999) Small worlds. The dynamics of networks between order and randomness. Princeton University Press, Princeton
35. Blau PM (1977) Inequality and heterogeneity: A primitive theory of social structure. Free Press, New York
36. Wasserman S, Faust K (1994) Social network analysis: Methods and applications. Cambridge University Press, Cambridge
37. White DR, Harary F (2001) The cohesiveness of blocks in social networks: Node connectivity and conditional density. Sociol Methodol 31:305–359
38. Alba RD (1973) A graph-theoretical definition of a sociometric clique. J Math Sociol 3:113–126
39. Luce RD (1950) Connectivity and generalized cliques in sociometric group structure. Psychometrika 15:169–190
40. Mokken RJ (1979) Cliques, clubs and clans. Qual Quant 13:161–173
41. Seidman SB (1983) Network structure and minimum degree. Soc Netw 5:269–287
42. Seidman SB, Foster BL (1978) A graph-theoretic generalization of the clique concept. J Math Sociol 6:139–154
43. Scott J (1991) Social network analysis: A handbook. Sage, London
44. de Nooy W, Mrvar A, Batagelj V (2005) Exploratory social network analysis with Pajek. Cambridge University Press, Cambridge
45. Seidman SB (1983) Internal cohesion of ls sets in graphs. Soc Netw 5:97–107
46. Borgatti SP, Everett MG, Shirey PR (1990) LS sets, lambda sets and other cohesive subsets. Soc Netw 12:337–357
47. Doreian P, Mrvar A (1996) A partitioning approach to structural balance. Soc Netw 18:149–168
48. White DR, Owen-Smith J, Moody J, Powell WW (2004) Networks, fields and organizations: Micro-dynamics, scale and cohesive embeddings. Comput Math Organ Theory 10:95–117
49. Friedkin NE (1984) Structural cohesion and equivalance explanations of social homogeneity. Sociol Methods Res 12:235–261
50. Granovetter M (1985) Economic action and social structure: The problem of embeddedness. Am J Sociol 91:481–510
51. Granovetter M (1992) Problems of explanation in economic sociology. In: Nohria N, Eccles RG (eds) Networks and organizations: Structure, form, and action. Harvard Business School Press, Boston, pp 25–56
52. Nowicki K, Snijders TAB (2001) Estimation and prediction for stochastic blockstructures. J Am Statist Assoc 96:1077–1087
53. Krackhardt D (1999) The ties that torture: Simmelian tie analysis in organizations. Res Sociol Organ 16:183–210
54. Bavelas A (1948) A mathematical model for group structures. Hum Organ 7:16–30
55. Freeman LC (1979) Centrality in social networks conceptual clarification. Soc Netw 1:215–239
56. Friedkin NE (1991) Theoretical foundations for centrality measures. Am J Sociol 96:1478–1504
57. Borgatti SP, Everett MG (2006) A graph-theoretic perspective on centrality. Soc Netw 28:466–484
58. Albert R, Barabasi A-L (2000) Topology of evolving networks: Local events and universality. Phys Rev Lett 85:5234–5237
59. Barabási A-L (2002) Linked: The new science of network. Perseus, Cambridge
60. Braha D, Bar-Yam Y (2006) From centrality to temporary fame: Dynamic centrality in complex networks. Complexity 12:59–63
61. Bonacich P (1972) Factoring and weighting approaches to clique identification. J Math Sociol 2:113–120
62. Bonacich P (1987) Power and centrality: A family of measures. Am J Sociol 92:1170–1182
63. Valente TW, Foreman RK (1998) Integration and radiality: Measuring the extent of an individual's connectedness and reachability in a network. Soc Netw 20:89–105
64. Granovetter M (1973) The strength of weak ties. Am J Sociol 78:1360–1380
65. Granovetter M (1995) Getting a job: A study of contacts and careers, 1st edn 1974. The University of Chicago Press, Chicago
66. Anthonisse JM (1971) The rush in a directed graph. Stichting Mathematisch Centrum, Amsterdam, p 10
67. Gould RV (1987) Measures of betweenness in non-symmetric networks. Soc Netw 9:277–282
68. White DR, Borgatti SP (1994) Betweenness centrality measures for directed graphs. Soc Netw 16:335–346
69. Stephenson K, Zelen M (1989) Rethinking centrality: Methods and examples. Soc Netw 11:1–37
70. Newman MEJ (2005) A measure of betweenness centrality based on random walks. Soc Netw 27:39–54
71. Freeman LC, Borgatti SP, White DR (1991) Centrality in valued graphs: A measure of betweenness based on network flow. Soc Netw 13:141–154
72. Tutzauer F (2007) Entropy as a measure of centrality in networks characterized by path-transfer flow. Soc Netw 29:249–265
73. Borgatti SP (2005) Centrality and network flow. Soc Netw 27:55–71
74. Snyder D, Kick E (1979) Structural position in the world system and economic growth 1955–70. A multiple network analysis of transnational interactions. Am J Sociol 84:1096–1126

75. Breiger RL (1981) Structures of economic interdependence among nations. In: Blau PM, Merton RK (eds) Continuities in structural inquiry. Sage, Newbury Park, pp 353–380
76. Bornschier V, Trezzini B (1997) Social stratification and mobility in the world system – Different approaches and recent research. Int Sociol 12:429–455
77. Borgatti SP, Everett MG (2000) Models of core/periphery structures. Soc Netw 21:375–395
78. Burt RS (1992) Structural holes: The social structure of competition. Harvard University Press, Cambridge/London
79. Burt RS (1992) The social structure of competition. In: Nohria N, Eccles RG (eds) Networks and organizations: Structure, form, and action. Harvard Business School Press, Boston, pp 57–91
80. Everett M, Borgatti SP (2005) Ego network betweenness. Soc Netw 27:31–38
81. Marsden PV (2002) Egocentric and sociocentric measures of network centrality. Soc Netw 24:407–422
82. Knoke D, Burt RS (1983) Prominence. In: Burt RS, Minor MJ (eds) Applied network analysis. A methodological introduction. Sage, Beverly Hills, pp 195–222
83. Merton RK (1968) The Matthew effect in science. Science 159:56–63
84. Lin N (1976) Foundations of social research. McGraw-Hill, New York
85. Bonacich P, Lloyd P (2001) Eigenvector-like measures of centrality for asymmetric relations. Soc Netw 23:191–201
86. Davis JA, Leinhardt S (1968) The structure of positive interpersonal relations in small groups. Annual Meeting of the American Sociological Association, Boston, August 1968
87. Davis JA, Leinhardt S (1972) The structure of positive interpersonal relations in small groups. In: Berger J (ed) Sociological theories in progress, vol 2. Houghton Mifflin, Boston, pp 218–251
88. Holland PW, Leinhardt S (1971) Transitivity in structural models of small groups. Comp Group Stud 2:107–124
89. Johnsen EC (1985) Network macrostructure models for the Davis–Leinhardt set of empirical sociomatrices. Soc Netw 7:203–224
90. de Nooy W (1999) The sign of affection: Balance-theoretic models and incomplete signed digraphs. Soc Netw 21:269–286
91. Davis JA (1979) The Davis/Holland/Leinhardt Studies: An overview. In: Holland PW, Leinhardt S (eds) Perspectives on social network research. Academic, New York, pp 51–62
92. Holland PW, Leinhardt S (1978) An omnibus test for social structure using triads. Sociol Methods Res 7:227–256
93. Doreian P, Batagelj V, Ferligoj A (2000) Symmetric-acyclic decompositions of networks. J Classif 17:3–28
94. Snijders TAB (2003) Accounting for degree distributions in empirical analysis of network dynamics. In: Breiger R, Carley K, Pattison P (eds) Dynamic social network modeling and analysis: Workshop summary and papers. National Academic, Washington DC, pp 146–161
95. Snijders TAB, Steglich CEG, Schweinberger M (2007) Modeling the co-evolution of networks and behavior. In: Montfort KV, Oud H, Satorra A (eds) Longitudinal models in the behavioral and related sciences. Erlbaum, Mahwah, pp 41–71
96. White HC (1992) Identity and control; A structural theory of social action. Princeton University Press, Princeton
97. Krackhardt D (1987) Cognitive social structures. Soc Netw 9:109–134
98. Breiger RL (2004) The analysis of social networks. In: Hardy MA, Bryman A (eds) Handbook of data analysis. Sage, London, pp 505–526
99. Bian Yj, Breiger RL, Davis D, Galaskiewicz J (2005) Occupation, class, and social networks in urban China. Social Forces 83:1443–1468
100. Mohr JW (1998) Measuring meaning structures. Ann Rev Sociol 24:345–370

Books and Reviews

Berkowitz SD (1982) An introduction to structural analysis: The network approach to social research. Butterworths, Toronto
Brandes U, Erlebach T (2005) Network analysis: Methodological foundations. Springer, Berlin
Breiger RL (2004) The analysis of social networks. In: Hardy MA, Bryman A (eds) Handbook of data analysis. Sage, London, pp 505–526
Carrington PJ, Scott J, Wasserman S (2005) Models and methods in social network analysis. Cambridge University Press, Cambridge
Degenne A, Forsé M (1999) Introducing social networks. Sage, London
de Nooy W, Mrvar A, Batagelj V (2005) Exploratory social network analysis with Pajek. Cambridge University Press, Cambridge
Freeman LC, White DR, Romney AK (1992) Research methods in social network analysis, 1st edn 1989. Transaction, New Brunswick, pp 530
Knoke D, Kuklinski JH (1982) Network analysis. Sage, Beverly Hills
Marsden PV, Lin N (1982) Social structure and network analysis. Sage, Beverly Hills
Scott J (1991) Social network analysis: A handbook. Sage, London
Wasserman S, Faust K (1994) Social network analysis: Methods and applications. Cambridge University Press, Cambridge
Wellman B, Berkowitz SD (1988) Social structures: A network approach. In: Granovetter M (ed) Structural analysis in the social sciences. Cambridge University Press, Cambridge, p 497

Social Network Analysis, Large-Scale

VLADIMIR BATAGELJ
University of Ljubljana, Ljubljana, Slovenia

Article Outline

Glossary
Definition of the Subject
Introduction
Large Networks and Complexity of Algorithms
Decompositions
Connectivity
Cuts
Dense Groups – Cores and Short Rings
Islands

Pattern Searching
Two Mode Networks
Multiplication of Networks
Statistical Approach
Future Directions
Bibliography

Glossary

For the basic notions on graphs and networks see the Articlea Wouter de Nooy: ▶ Social Network Analysis, Graph Theoretical Approaches to.

Network consists of vertices linked by lines and additional data about vertices and/or lines.

Network decomposition identification of parts of network and their interconnections. Usually it is described by a partition of set of vertices or set of lines.

Time complexity of algorithm describes how the time needed to run the algorithm depends on the size of the input data.

Reduction of network a network obtained by shrinking each cluster from a given partition into a vertex.

Condensation a reduction for strong connectivity partition.

Cut a subnetwork of vertices/lines with values of selected property above given threshold.

Island a connected subnetwork of selected size of (locally) important, with respect to selected property, vertices/lines.

Pattern searching identification of all appearances of selected small subnetwork (pattern or fragment) in a given network.

Topological sort procedure to determine a compatible ordering in acyclic network.

Definition of the Subject

A *network* is based on two sets: a set of *vertices* (nodes), that represent the selected *units*, and a set of *lines* (links), that represent *ties* between units. Each line has two vertices as its *end-points*; if they are equal it is called a *loop*. Vertices and lines form a *graph*. A line can be *directed* – an *arc*, or *undirected* – an *edge*.

Additional data about vertices or lines are usually known – their *properties* (attributes). For example: name/label, type, value, position, … In general

$$\text{Network} = \text{Graph} + \text{Data}\,.$$

The data can be measured or computed.

Formally, a *network* $\mathcal{N} = (\mathcal{V}, \mathcal{L}, \mathcal{P}, \mathcal{W})$ consists of the following:

- A *graph* $\mathcal{G} = (\mathcal{V}, \mathcal{L})$, where $\mathcal{V}$ is the set of vertices and $\mathcal{L} = \mathcal{E} \cup \mathcal{A}$. $\mathcal{E} \cap \mathcal{A} = \emptyset$ is the set of lines. $\mathcal{A}$ is the set of *arcs* and $\mathcal{E}$ is the set of *edges*.
- $\mathcal{P}$ – set of *vertex value functions* or properties: $p\colon \mathcal{V} \to A$
- $\mathcal{W}$ – set of *line value functions* or weights: $w\colon \mathcal{L} \to B$

The size of a network/graph is expressed by two numbers: number of vertices $n = |\mathcal{V}|$ and number of lines $m = |\mathcal{L}|$. In a *simple undirected* graph (no parallel edges, no loops) $m \le \frac{1}{2}n(n-1)$; and in a *simple directed* graph (no parallel arcs) $m \le n^2$.

For a family of graphs $\mathbb{G}$, we define a *density* of graph $\mathcal{G}$ as $\gamma(\mathcal{G}) = \frac{m(\mathcal{G})}{m_{\max}(\mathbb{G})}$.

Introduction

Small networks (some tens of vertices) can be represented by a picture and analyzed by many algorithms (***UCINET***, ***NetMiner***). Also *middle size* networks (some hundreds of vertices), if they are not dense, can still be represented by a picture, but some analytical procedures can't be used.

Till 1990 most networks were small – they were collected by researchers using surveys, observations, archival records, etc. The advances in IT allowed one to create networks from the data already available in the computer(s) or by browsing on the Internet. *Large* networks became reality. Large networks are too big to be displayed in detail; special algorithms are needed for their analysis (Pajek). The availability of large data sets also provided incentives to the boost of theoretical research in (large) network analysis (not only in the social sciences).

A recent overview of social network analysis software is given in Huisman and Van Duijn [28].

Large Networks and Complexity of Algorithms

Large networks have several thousands or millions of vertices. The upper limit to their size is technologically dependent – they can be stored in computer's memory; otherwise we deal with a *huge* network (see Abello et al. [1]).

Large networks are usually sparse $m \ll n^2$; typically $m = O(n)$ or $m = O(n \log n)$, see Table 1.

A collection of large networks is available from Pajek's datasets.

The *time complexity* of an algorithm describes how the time needed to run the algorithm depends on the size of the input data. In computer science the problems for which only algorithms of exponential (or higher) complexity are known are considered hard or intractable, since the speed-up of a computer only additively increases the size of problems that can be solved in a given period of

Social Network Analysis, Large-Scale, Table 1
Examples of large networks

Network	$n = \|\mathcal{V}\|$	$m = \|\mathcal{L}\|$	Source
ODLIS dictionary	2909	18419	ODLIS online
Citations SOM	4470	12731	Garfield's collection
Molecula 1ATN	5020	5128	Brookhaven PDB
Comput geometry	7343	11898	BiBTEX bibliographies
English words 2-8	52652	89038	Knuth's English words
Internet traceroutes	124651	207214	Internet Mapping Project
Franklin genealogy	203909	195650	Roperld.com gedcoms
World-Wide-Web	325729	1497135	Notre Dame Networks
Internet Movie DB	1324748	3792390	IMDB
Wikipedia	659388	16582425	Wikimedia
US patents	3774768	16522438	Nber
SI internet	5547916	62259968	Najdi Si

time, but the problems for which an algorithm of polynomial complexity exists are considered 'nice'. When dealing with large instances of problems this isn't always true anymore. Let us look to time complexities of some typical algorithms in Table 2.

For the interactive use on large networks already quadratic algorithms, $O(n^2)$, are too slow – we have to restrict our 'toolbox' to a selection of efficient, *subquadratic* algorithms.

How can we deal with large structures? Already Romans knew – ***divide et impera*** (divide and conquer). In case of networks ***divide*** means the use of (recursive) *decomposition* of a large network into several smaller networks (see Fig. 1) that can be visualized and treated further using more sophisticated methods; ***impera*** means that we have to take care about the interlinks among so obtained parts.

Another approach is the use of different *statistical* quantities to describe the properties of a network and using probabilistic models to derive the answers to some questions.

Decompositions

Decompositions of a network are usually described by clusterings of vertices or lines. In the following we shall use mainly the clusterings of vertices.

A nonempty subset $C \subseteq \mathcal{V}$ is called a *cluster* (group). A nonempty set of clusters $\mathbf{C} = \{C_i\}$ forms a *clustering*.

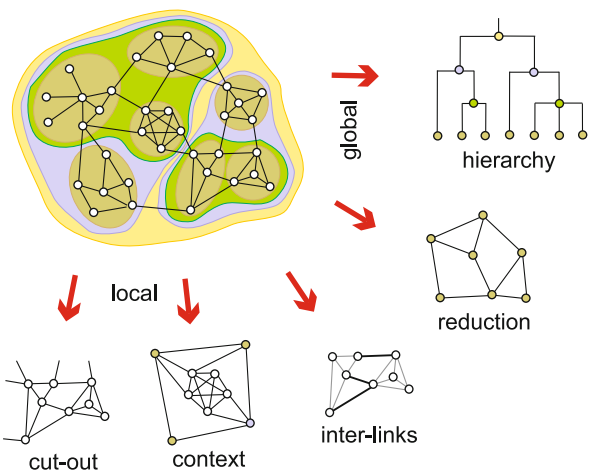
Clustering $\mathbf{C} = \{C_i\}$ is a *partition* iff

$$\cup\mathbf{C} = \bigcup_i C_i = \mathcal{V} \quad \text{and} \quad i \neq j \Rightarrow C_i \cap C_j = \emptyset\,.$$

Clustering $\mathbf{C} = \{C_i\}$ is a *hierarchy* iff $C_i \cap C_j \in \{\emptyset, C_i, C_j\}$. In other words, in a hierarchy two clusters are either disjoint or is one contained in the other.

Hierarchy $\mathbf{C} = \{C_i\}$ is *complete*, iff $\cup\mathbf{C} = \mathcal{V}$; and is *basic* if for all $v \in \cup\mathbf{C}$ also $\{v\} \in \mathbf{C}$.

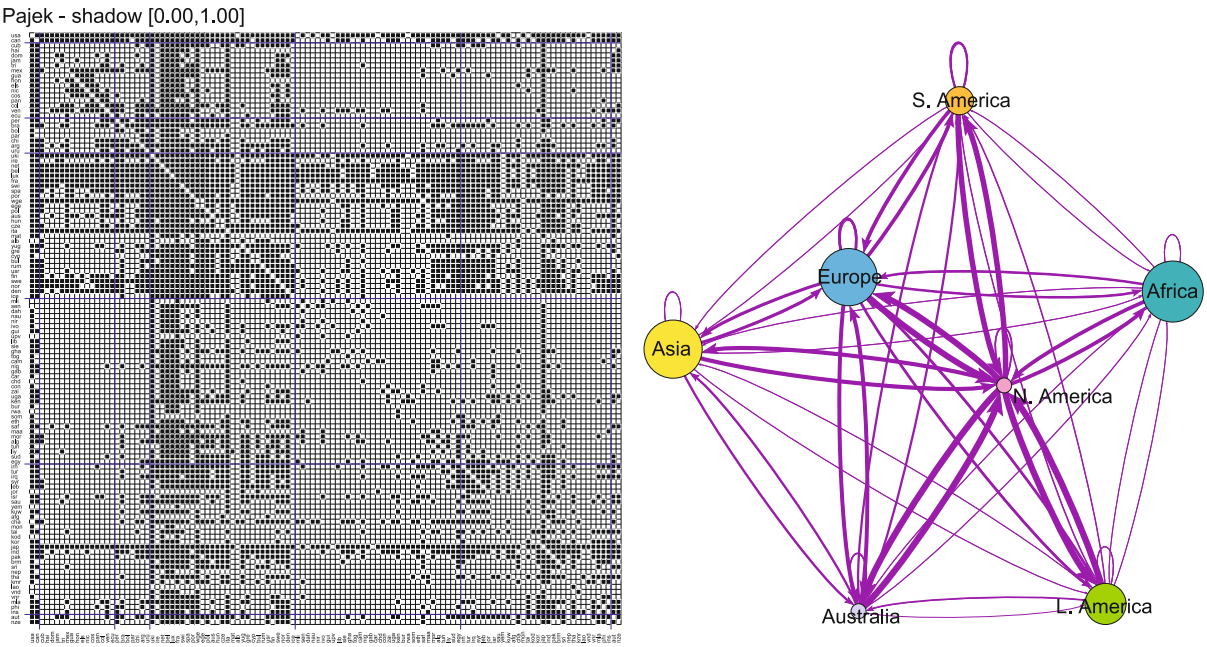
Contraction of cluster C in a graph $\mathcal{G}$ is called a graph $\mathcal{G}/C$, in which all vertices of the cluster C



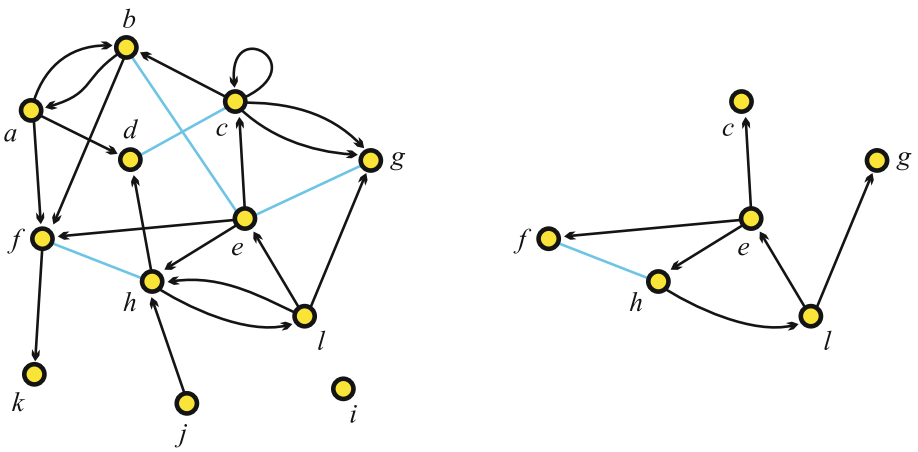
Social Network Analysis, Large-Scale, Figure 1
Decompositions

Social Network Analysis, Large-Scale, Table 2
Complexities of some typical algorithms

Algorithm	T(*n*)	1000	10 000	100 000	1 000 000	10 000 000
Alg-A	$O(n)$	0.00 s	0.015 s	0.17 s	2.22 s	22.2 s
Alg-B	$O(n \log n)$	0.00 s	0.06 s	0.98 s	14.4 s	**2.8 m**
Alg-C	$O(n\sqrt{n})$	0.01 s	0.32 s	10.0 s	**5.27 m**	**2.78 h**
Alg-D	$O(n^2)$	0.07 s	7.50 s	**12.5 m**	**20.8 h**	**86.8 d**
Alg-E	$O(n^3)$	0.10 s	**1.67 m**	**1.16 d**	**3.17 y**	**3.17 ky**



Social Network Analysis, Large-Scale, Figure 2
Snyder and Kick's international trade; matrix display and reduction



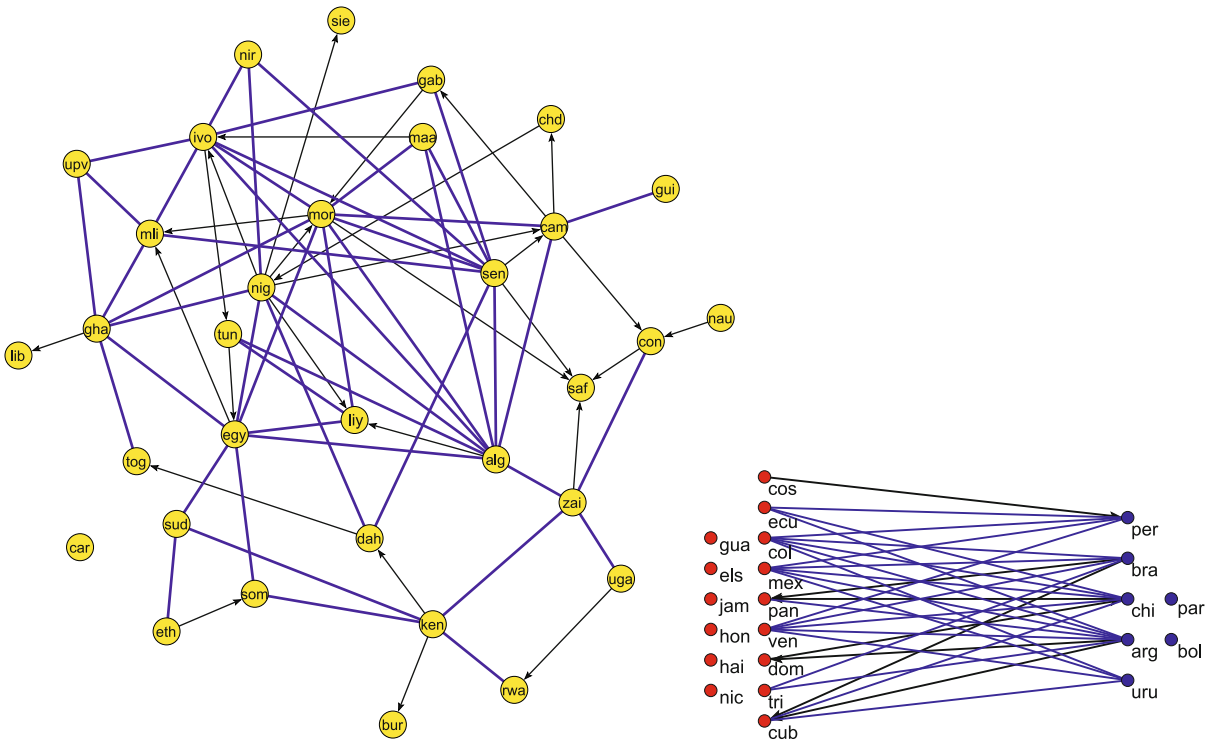
Social Network Analysis, Large-Scale, Figure 3
Graph and its subgraph

are replaced by a single new vertex, say c. More precisely: $\mathcal{G}/C = (\mathcal{V}', \mathcal{L}')$, where $\mathcal{V}' = (\mathcal{V} \setminus C) \cup \{c\}$ and $\mathcal{L}'$ consists of lines from $\mathcal{L}$ that have both end-points in $\mathcal{V} \setminus C$. Beside these it contains also a 'star' with the center c and: arc(v, c), if $\exists p \in \mathcal{L}, u \in C$: $p(v, u)$; or arc(c, v), if $\exists p \in \mathcal{L}, u \in C$: $p(u, v)$. There is a loop (c, c) in c if $\exists p \in \mathcal{L}, u, v \in C$: $p(u, v)$.

In a network over graph $\mathcal{G}$, we have also to specify how the new values/weights are determined in the shrunk part of the network. Usually as the sum or maximum/minimum of the original values.

For a given partition if we contract all clusters except few selected we obtain their *context*; and if we contract all clusters we obtain the *reduction* of a given network.

On the left side of the Fig. 2 the *matrix display* of Snyder and Kick's [40] international trade network is presented. Vertices in the display are reordered according to the partition by (sub)continents. On the right side the



Social Network Analysis, Large-Scale, Figure 4
Africa cut-out and inter-links between South and Latin America

corresponding reduction of the network is presented. The lines in the reduction have the thickness proportional to the weights

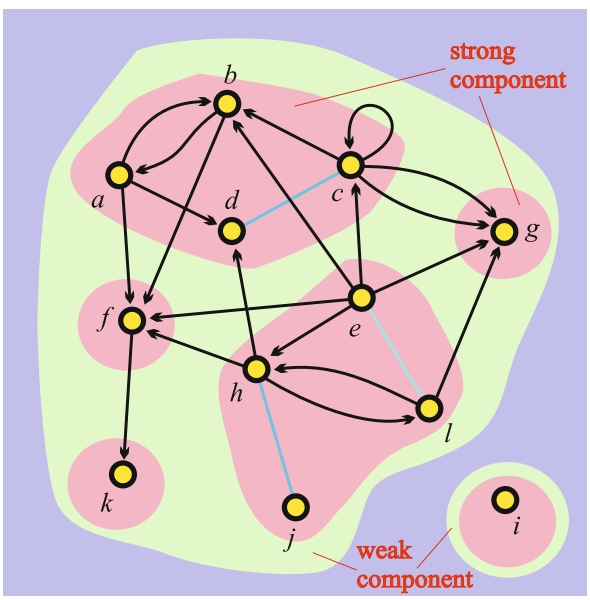
$$w(C_i, C_j) = \frac{n(C_i, C_j)}{n(C_i) \cdot n(C_j)}$$

where $n(C_i, C_j)$ is the number of lines from cluster C_i to cluster C_j; and $n(C_i)$ is the number of lines inside the cluster C_i.

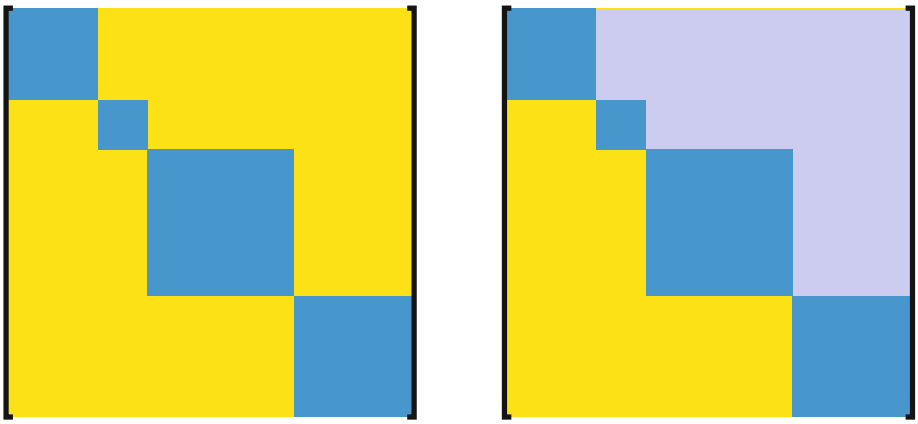
A *subgraph* $\mathcal{H} = (\mathcal{V}', \mathcal{L}')$ of a given graph $\mathcal{G} = (\mathcal{V}, \mathcal{L})$ is a graph which set of lines is a subset of set of lines of $\mathcal{G}$, $\mathcal{L}' \subseteq \mathcal{L}$, its vertex set is a subset of set of vertices of $\mathcal{G}$, $\mathcal{V}' \subseteq \mathcal{V}$, and it contains all end-vertices of $\mathcal{L}'$. The graph on the right side of Fig. 3 is a subgraph of the graph on the left side.

A subgraph can be *induced* by a given subset of vertices $\mathcal{V}'$, then $\mathcal{L}' = \mathcal{L}|\mathcal{V}'$ consists of all lines from $\mathcal{L}$ which have both end-points in $\mathcal{V}'$; or lines $\mathcal{L}'$, then $\mathcal{V}' = \mathcal{V}|\mathcal{L}'$ consists of all end-points of lines from $\mathcal{L}'$. It is a *spanning* subgraph iff $\mathcal{V}' = \mathcal{V}$.

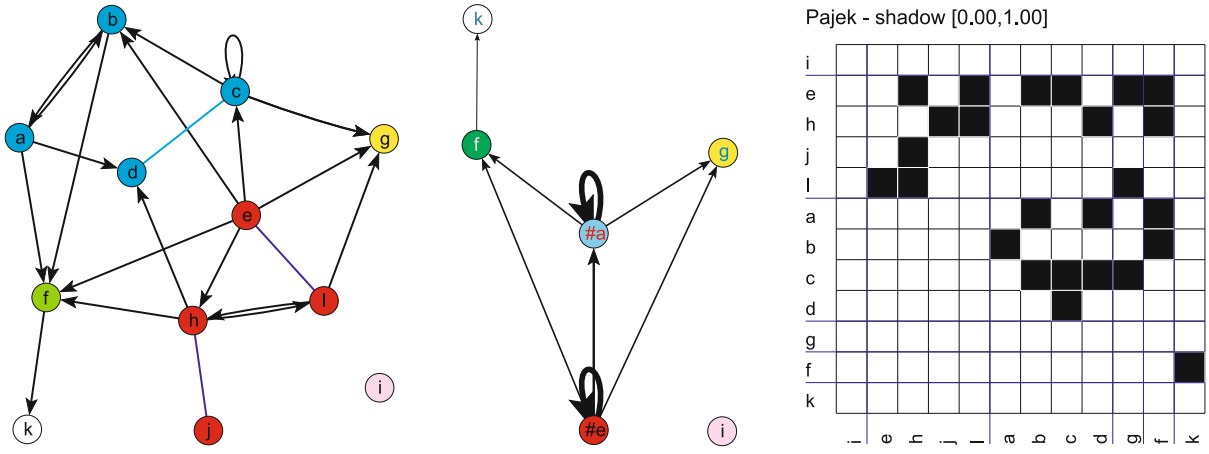
On the left side of Fig. 4 the *cut-out* of African countries from the Snyder and Kick's network is presented – the induced subgraph by Africa cluster; and on the right side the *inter-links* between Latin America and South America – the induced subgraph by Latin America and South America clusters with inside cluster lines removed.



Social Network Analysis, Large-Scale, Figure 5
Weak and strong components



Social Network Analysis, Large-Scale, Figure 6
Weak and strong components in matrix display



Social Network Analysis, Large-Scale, Figure 7
Condensation

Connectivity

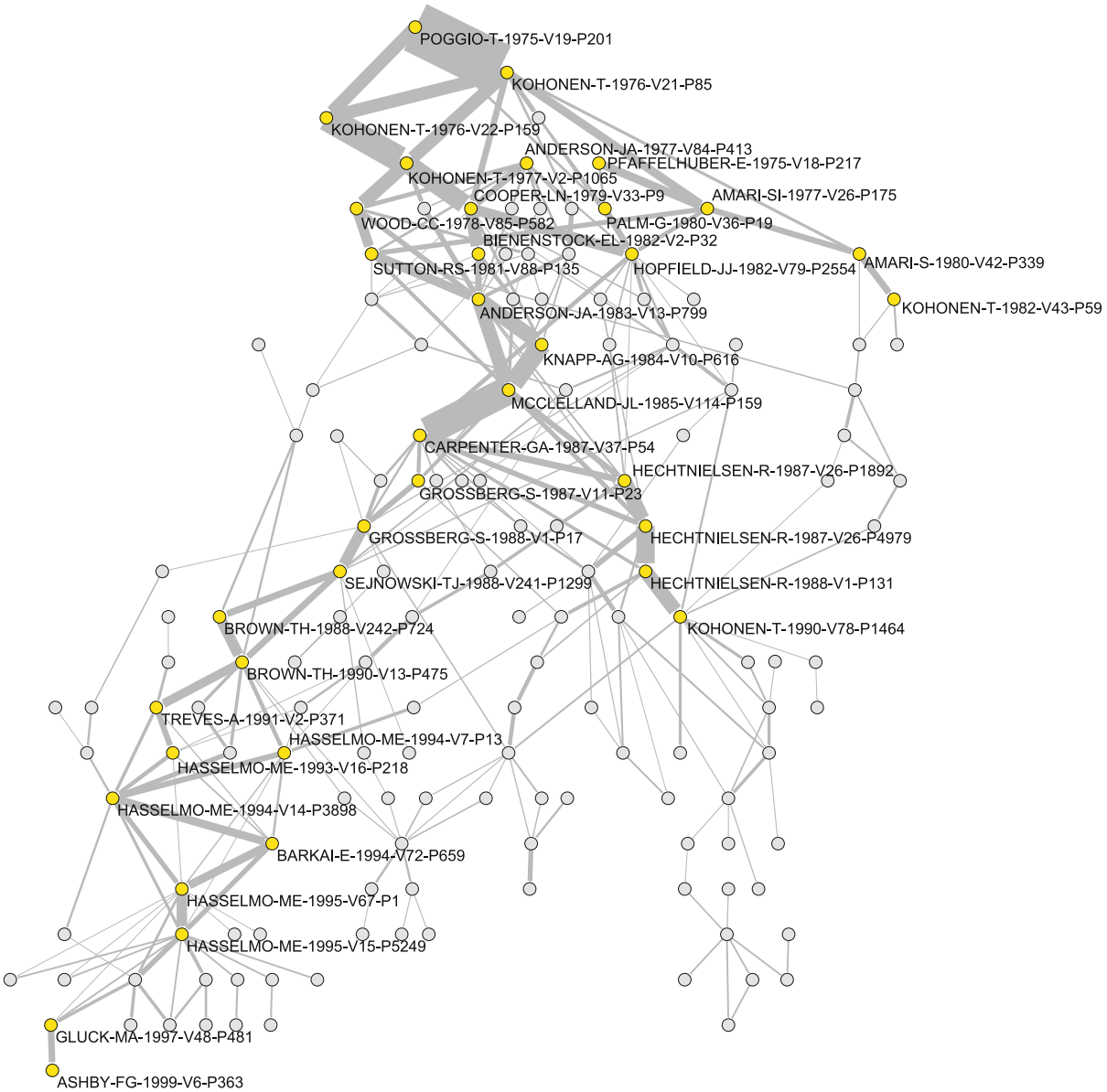
A *walk* from vertex u to vertex v is a sequence of lines $l(v_{i-1}, v_i)$, $i = 1, \ldots, k$ such that $v_0 = u$ and $v_k = v$. k is called the *length* of the walk. If in the definition of a walk we don't care about the direction of its lines we get a *semi-walk*. A walk is *closed* iff $u = v$. A graph is *acyclic* iff it doesn't contain any closed walk. A walk in which all vertices are different is a *path*.

Vertex u is *reachable* from vertex v iff there exists a walk with initial vertex v and terminal vertex u. Vertex v is *weakly connected* with vertex u iff there exists a semi-walk with v and u as its end-vertices. Vertex v is *strongly connected* with vertex u iff they are mutually reachable.

Weak and strong connectivity are equivalence relations. Equivalence classes induce weak/strong *components* (See Fig. 5).

Reordering the vertices of network such that the vertices from the same class of weak partition are put together we get a matrix representation (left side of Fig. 6) consisting of diagonal blocks – weak components. The out-diagonal blocks are zero-blocks. Most problems can be solved separately on each component and afterward these solutions combined into final solution.

If we shrink every strong component of a given graph into a vertex, delete all loops and identify parallel arcs the obtained reduced graph, called also the *condensation* of a given graph, is acyclic [27]. For every acyclic



Social Network Analysis, Large-Scale, Figure 8
Main component of arc cut at level 0.007 of the SOM citation network

graph an *ordering/level* function $i\colon \mathcal{V} \to \mathbb{N}$ exists such that $(u, v) \in \mathcal{A} \Rightarrow i(u) < i(v)$. The procedure to determine such ordering is called *topological sort* [19]. Reordering in matrix display the vertices of a network by this ordering we obtain a representation as at the right side of Fig. 6 – the blocks below the diagonal are zero-blocks.

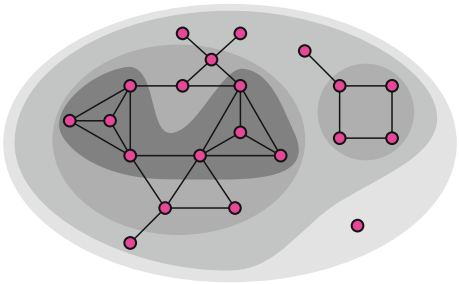
A directed graph, its condensation and its topologically ordered matrix display are presented in Fig. 7.

For several network analysis problems more efficient algorithms exist for acyclic networks.

Cuts

The basic approach to find interesting groups inside a network is to express our intentions (question) with an appropriate property/weight (measured or computed from network structure) and then identify the substructures of elements with the highest (lowest) values of the selected property. This approach is known as a method of *cuts*.

There exist several measures of importance of vertices in a network such as: degree, betweeness, closeness [16,24], hubs and authorities [30], clustering coefficient, …



Social Network Analysis, Large-Scale, Figure 9
Cores

The *degree* $\deg(v)$ of vertex v equals to the number of lines having vertex v as their end-point. The *maximum degree* of a graph is denoted by Δ. Similarly the in-degree $\text{indeg}(v)$ of vertex v equals to the number of lines having vertex v as their terminal point, and the out-degree $\text{outdeg}(v)$ ...

The *vertex-cut* of a network $\mathcal{N} = (\mathcal{V}, \mathcal{L}, p)$, for a property $p\colon \mathcal{V} \to \mathbb{R}$, at selected level t is a subnetwork $\mathcal{N}(t) = (\mathcal{V}', \mathcal{L}(\mathcal{V}'), p)$, determined by the set

$$\mathcal{V}' = \{v \in \mathcal{V}\colon p(v) \geq t\}$$

and $\mathcal{L}(\mathcal{V}')$ is the set of lines from $\mathcal{L}$ that have both end-points in $\mathcal{V}'$.

The *line-cut* of a network $\mathcal{N} = (\mathcal{V}, \mathcal{L}, w)$, for a weight $w\colon \mathcal{L} \to \mathbb{R}$, at selected level t is a subnetwork $\mathcal{N}(t) = (\mathcal{V}(\mathcal{L}'), \mathcal{L}', w)$, determined by the set

$$\mathcal{L}' = \{e \in \mathcal{L}\colon w(e) \geq t\}$$

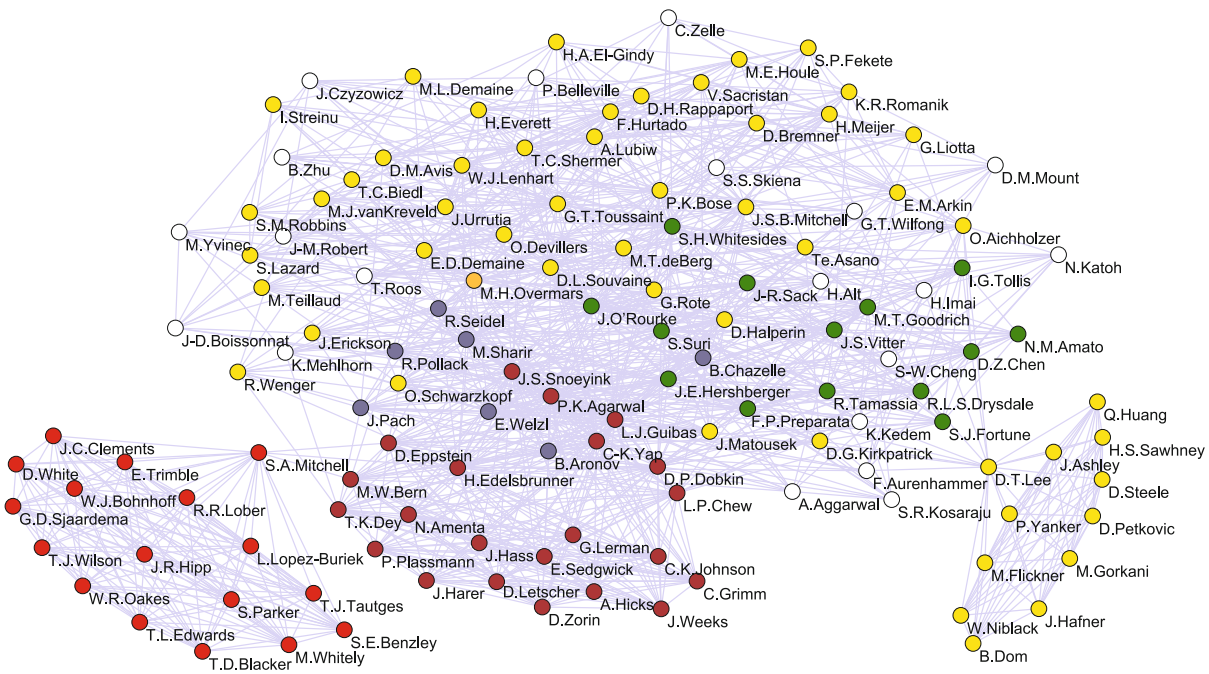
and $\mathcal{V}(\mathcal{L}')$ is the set of all end-points of the lines from $\mathcal{L}'$.

In the analysis of a cut $\mathcal{N}(t)$, we look at its components. Their number and sizes depend on t. Usually there are many small components. Often we consider only components of size at least k and not exceeding K. The components of size smaller than k are discarded as 'less interesting'; and the components of size larger than K are cut again at some higher level.

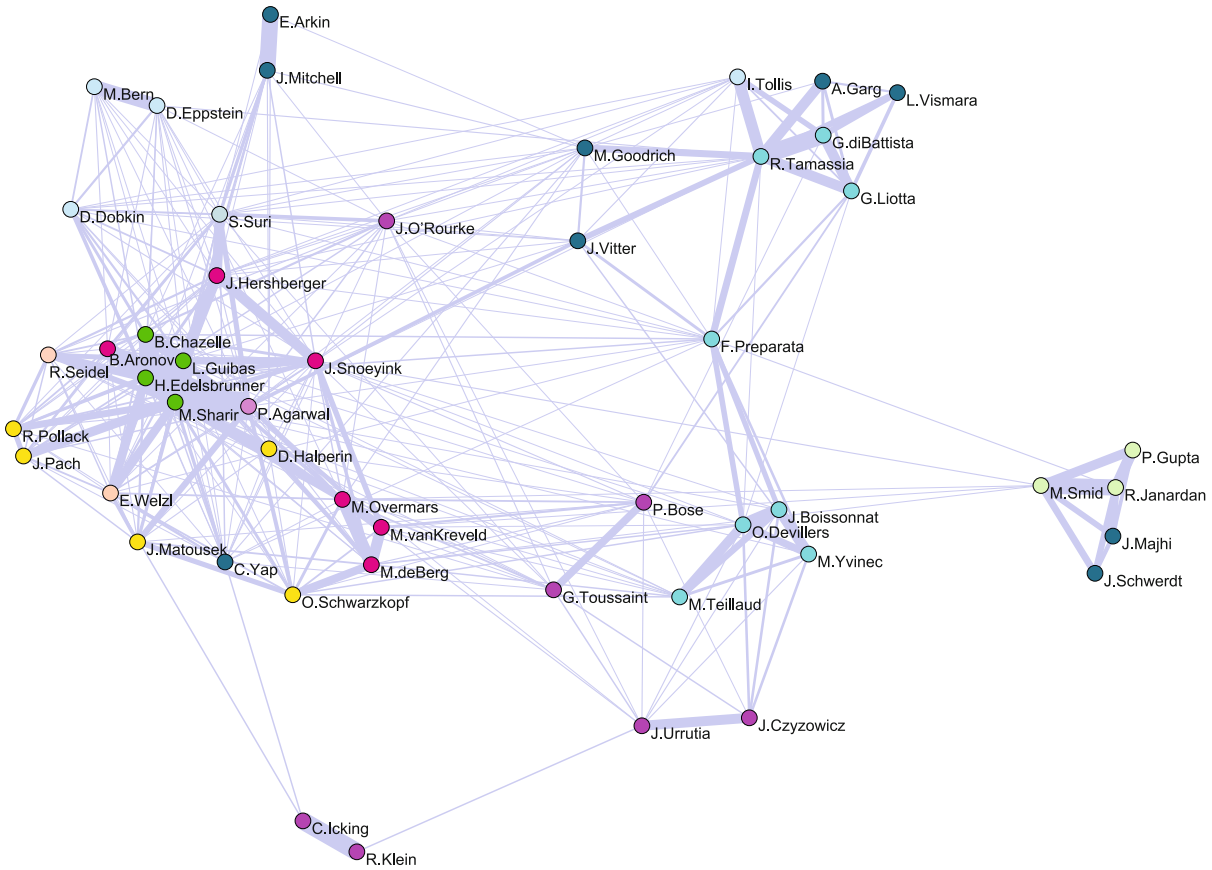
The values of threshold t and size bounds k and K are determined by inspecting the distribution of vertex/line-values and the distribution of component sizes and considering additional knowledge on the nature of network or goals of analysis.

The p_S-core at level 46 (see Fig. 11) of the collaboration network in the field of computational geometry is an example of vertex cut.

The citation network analysis started in 1964 with the paper of Garfield et al. [25]. In 1989 Hummon and Doreian [29] proposed three indices – weights of arcs that are proportional to the number of different source-sink paths passing through the arc. In Fig. 8 the main component



Social Network Analysis, Large-Scale, Figure 10
Cores of orders 10–21 in Computational Geometry collaboration network



Social Network Analysis, Large-Scale, Figure 11
***p*$_5$-core at level 46 in Computational Geometry collaboration network**

of the arc cut at level 0.007 for SPC (search path count) weights of the SOM (selforganizing maps) citation network (4470 vertices, 12,731 arcs) is presented.

Dense Groups – Cores and Short Rings

Several notions were proposed in attempts to formally describe dense groups in graphs.

Clique of order k, $k \geq 3$, is a maximal complete subgraph (isomorphic to complete graph K_k – graph with k vertices and all possible edges among them).

Other notions are: s-plexes, s-clans, LS sets, lambda sets, cores, … (Wasserman and Faust [43]). For all of them, except for cores, it turned out that they are difficult (no fast algorithm exists) to determine.

The notion of core was introduced by Seidman in 1983 [38]. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph. A subgraph $\mathcal{H}_k = (\mathcal{W}, \mathcal{E}|\mathcal{W})$ induced by the set $\mathcal{W}$ is a *k-core* or a *core of order k* iff for all $v \in \mathcal{W}$: $\deg_{\mathcal{H}_k}(v) \geq k$, and $\mathcal{H}_k$ is a maximal subgraph with this property. The core of maximum order is also called the *main* core. The *core number* of vertex v is the highest order of a core that contains this vertex. In general graphs instead of the degree $\deg(v)$ we can also use: in-degree, out-degree, in-degree + out-degree, etc., determining different types of cores.

From Fig. 9, representing 0, 1, 2 and 3 core, we can see the following properties of cores:

- The cores are nested: $i < j \Longrightarrow \mathcal{H}_j \subseteq \mathcal{H}_i$. They form a hierarchy.
- Cores are not necessarily connected subgraphs.

An efficient algorithm for determining the cores hierarchy is based on the following property: If from a given graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ we recursively delete all vertices, and edges incident with them, of degree less than k, the remaining graph is the k-core.

The Fig. 10 presents the cores of orders 10 to 21 in the collaboration network ($n = 7343$, $m = 11{,}898$) for the field of computational geometry – two authors are linked

iff they wrote a paper together. The weight of the edge equals to the number of joint papers.

The notion of core can be generalized to networks. Let $\mathcal{N} = (\mathcal{V}, \mathcal{E}, w)$ be a network, where $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a graph and weight $w\colon \mathcal{E} \to \mathbb{R}$ is a function assigning values to edges. A *vertex property function* on $\mathcal{N}$, or a *p-function* for short, is a function $p(v, U), v \in \mathcal{V}, U \subseteq \mathcal{V}$ with real values. Let $N_U(v) = N(v) \cap U$, where $N(v)$ is the set of neighbors of v. Besides degrees, here are some other examples of p-functions [13]:

$$p_S(v, U) = \sum_{u \in N_U(v)} w(v, u), \text{ where } w\colon \mathcal{E} \to \mathbb{R}_0^+$$

$$p_M(v, U) = \max_{u \in N_U(v)} w(v, u), \text{ where } w\colon \mathcal{E} \to \mathbb{R}$$

$$p_k(v, U) = \text{number of cycles of length } k \text{ through vertex } v \text{ in } (U, \mathcal{E}|U)$$

$$p_\gamma(v, U) = \begin{cases} \dfrac{\deg(v, U)}{\max_{u \in N(v)} \deg(u)}, & \text{if } \deg(v) > 0; \\ 0, & \text{otherwise} \end{cases}$$

$$p_\delta(v, U) = \max_{u \in N_U^+(v)} \deg(u) - \min_{u \in N_U^+(v)} \deg(u)$$

$$p_a(v, U) = \begin{cases} \dfrac{1}{|N_U(v)|} \displaystyle\sum_{u \in N_U(v)} w(v, u), & \text{if } N_U(v) \neq \emptyset; \\ 0, & \text{otherwise} \end{cases}$$

The subgraph $\mathcal{H} = (C, \mathcal{E}|C)$ induced by the set $C \subseteq \mathcal{V}$ is a *p-core at level p*-core $t \in \mathbb{R}$ iff for all $v \in C\colon t \leq p(v, C)$ and C is a maximal such set.

The function p is *monotone*, iff it has the property

$$C_1 \subset C_2 \Rightarrow \forall v \in \mathcal{V}\colon (p(v, C_1) \leq p(v, C_2))$$

The degrees and the functions p_S, p_M, p_k, p_γ and p_δ are monotone; and p_a is not. For a monotone function the p-core at level t can be determined, as in the ordinary case, by successively deleting vertices with value of p lower than t; and the cores on different levels are nested

$$t_1 < t_2 \Rightarrow \mathcal{H}_{t_2} \subseteq \mathcal{H}_{t_1}$$

The p-function is *local*, iff $p(v, U) = p(v, N_U(v))$. The degrees, $p_S, p_M, p_\gamma, p_\delta$ and p_a are local; but p_k is **not** local for $k \geq 4$. For a local monotone p-function an $O(m \max(\Delta, \log n))$ algorithm for determining the p-core levels exists, assuming that $p(v, N_C(v))$ can be computed in $O(\deg_C(v))$.

Figure 11 presents the p_S-core at level 46 of the collaboration network in the field of computational geometry. Note, for example, that R. Klein (lower left) has in-core degree only 2, but its in-core sum of weights is at least 46 – he wrote most of his papers with C. Icking.

A k-ring is a simple closed chain of length k. Using k-rings we can define a weight of an edge e as $w_k(e) = \#$ of different k-rings containing the edge $e \in \mathcal{E}$.

Since for each edge e of complete graph K_r, $r \geq k \geq 3$ we have $w_k(e) = (r-2)!/(r-k)!$, the edges belonging to cliques have large weights. Therefore these weights can be used to identify the dense parts of a network. For example: all r-cliques of a network belong to $(r-2)$-edge cut for the weight w_3.

Related to triangular (3-rings) network is the notion of *triangular connectivity* that can be used to operationalize the notion of Granovetter's strong and weak ties [26]. This notion can be generalized to short cycle connectivity. For details see [14]. For efficient algorithms for computing triangles in networks see [10,39], and Latapy.

In Fig. 12 the edge-cut at level 16 of triangular network of Erdős collaboration graph (without Erdős, $n = 6926$, $m = 11{,}343$) is presented [9].

In directed networks there are two types of triangles or 3-rings (cyclic and transitive, see Fig. 14).

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a simple undirected graph. *Clustering* in vertex v is usually measured as a quotient between the number of lines in subgraph $\mathcal{G}^1(v) = \mathcal{G}(N(v))$ induced by the neighbors of vertex v and the number of lines in the complete graph on these vertices:

$$C(v) = \begin{cases} \dfrac{2|\mathcal{L}(\mathcal{G}^1(v))|}{\deg(v)(\deg(v)-1)}, & \deg(v) > 1 \\ 0, & \text{otherwise} \end{cases}.$$

For simple directed graphs, we have to omit the number 2.

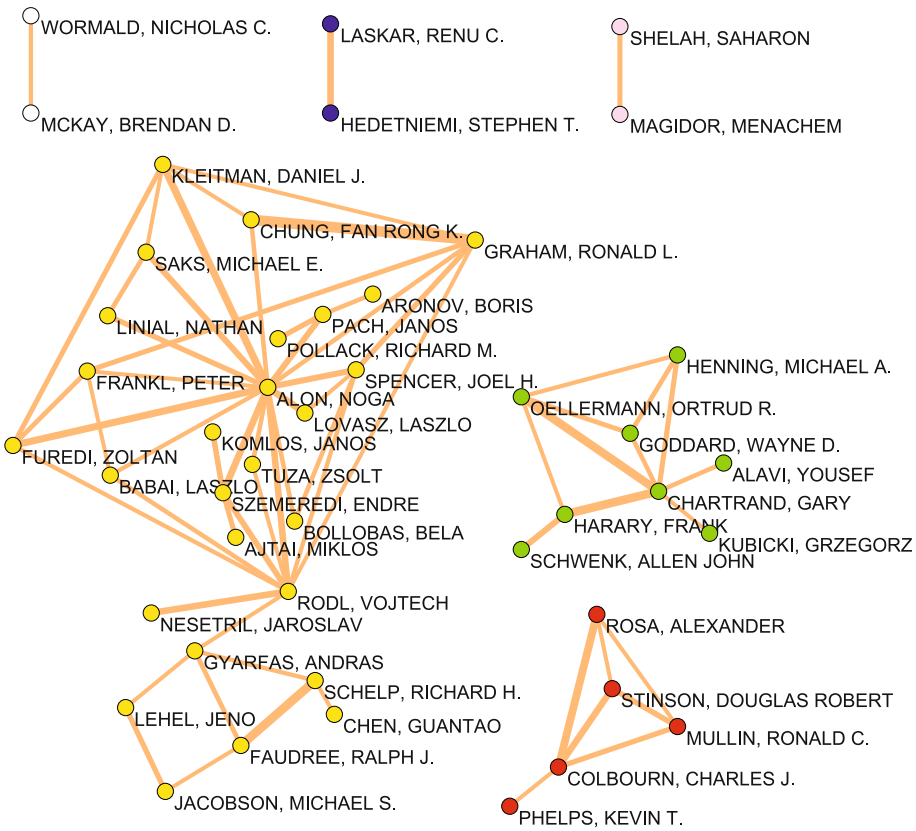
So defined clustering coefficient attains largest values mostly on vertices of low degree – it is not useful for data analysis task. A better coefficient is obtained by the following correction

$$C_1(v) = \frac{\deg(v)}{\Delta} C(v)$$

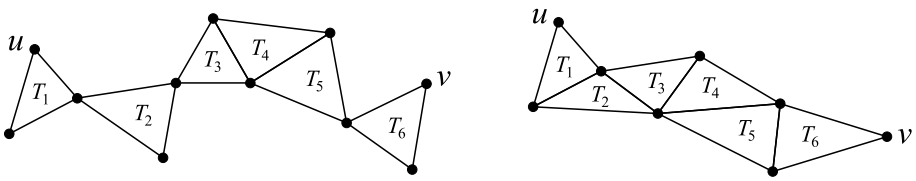
where Δ is the maximum degree in graph $\mathcal{G}$. This measure attains its largest value in vertices that belong to an isolated clique of size Δ.

Islands

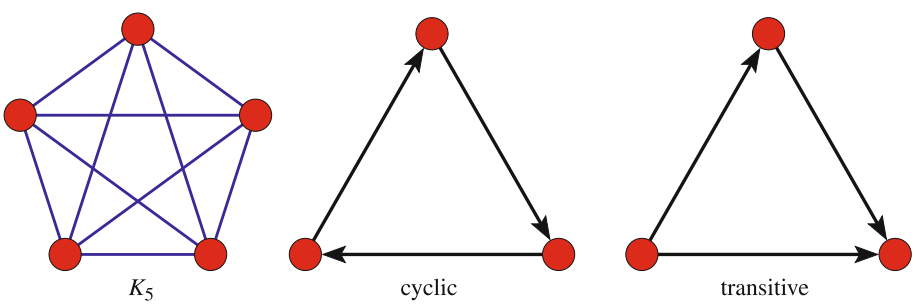
Islands are very general and efficient approach to determine the 'important' subnetworks in a given network with respect to a given property of vertices or lines. It is an improvement of the cuts approach. If we represent a given or computed value of vertices/lines as a height of vertices/



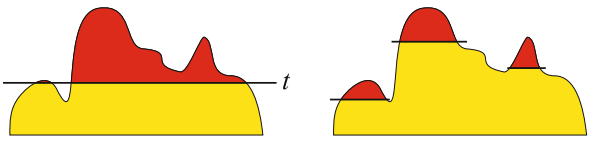
Social Network Analysis, Large-Scale, Figure 12
Edge-cut at level 16 of triangular network of Erdős collaboration graph



Social Network Analysis, Large-Scale, Figure 13
Vertex and edge triangular connectivity



Social Network Analysis, Large-Scale, Figure 14
***K_5* and cyclic and transitive 3-ring**

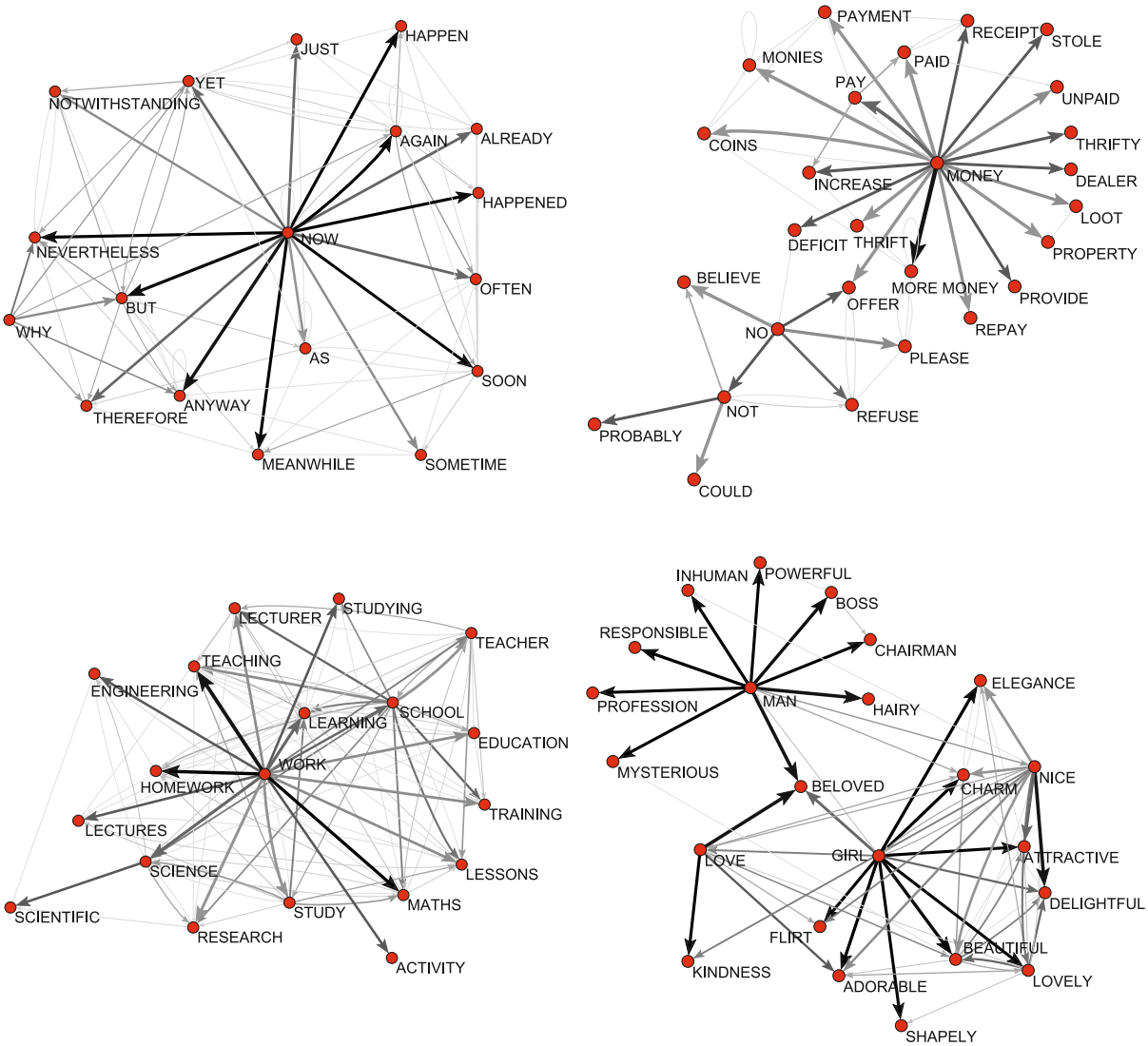


Social Network Analysis, Large-Scale, Figure 15
Cuts and islands

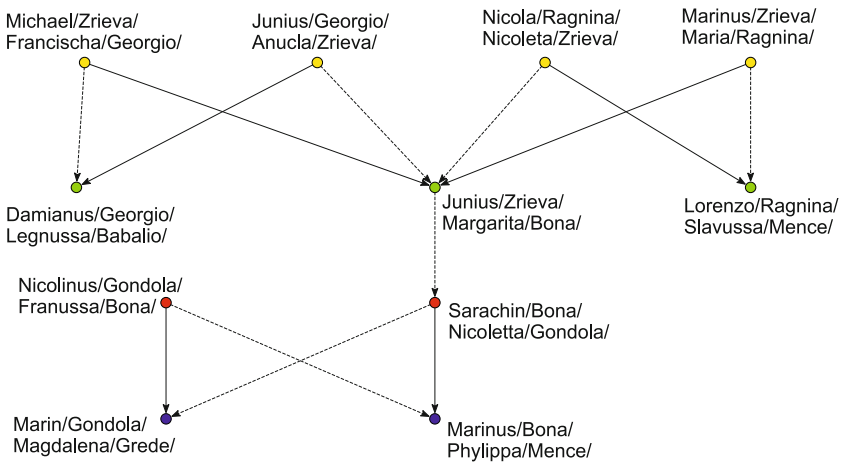
lines and we immerse the network into a water up to selected level we get *islands*. Varying the level we get different islands [45].

In the islands approach we select only maximal islands of sizes inside the given size bounds k to K, but on different levels. In this way we bypass the problems of the cuts approach: determining the 'right' threshold value and too small/large sizes of obtained components. Besides this we can also identify locally important islands with small heights – emerging groups. Very efficient algorithms exist to determine the islands hierarchy and to list all the islands of selected sizes. An island is *simple* iff it has only one peak.

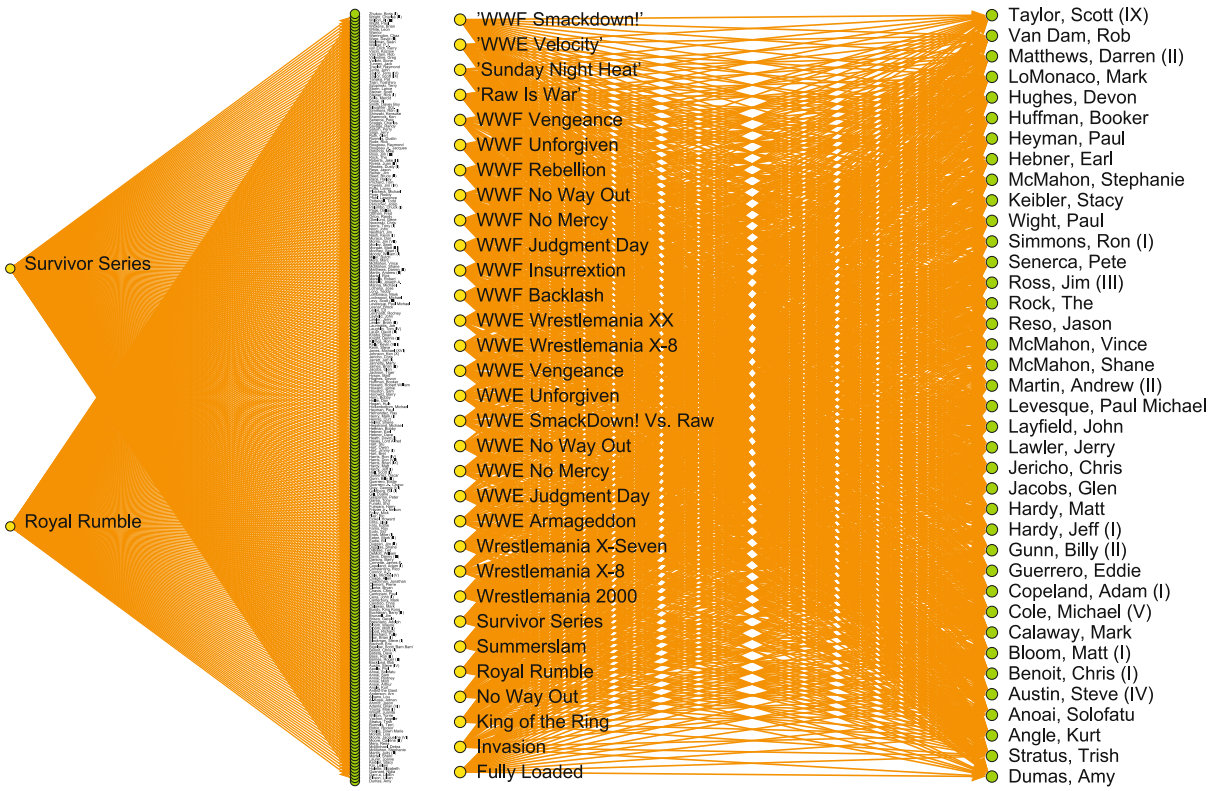
As an example, let us take the Nber network of US Patents. It has 3,774,768 vertices and 16,522,438 arcs. We computed SPC weights in it and determined all (2,90)-islands. The reduced network has 470,137 vertices, 307,472 arcs and for different k: $C_2 = 187{,}610$, $C_5 = 8859$, $C_{30} = 101$, $C_{50} = 30$ islands. The main island turns out



Social Network Analysis, Large-Scale, Figure 16
Selected islands from The Edinburgh Associative Thesaurus



Social Network Analysis, Large-Scale, Figure 17
Marriages among relatives in Ragusa

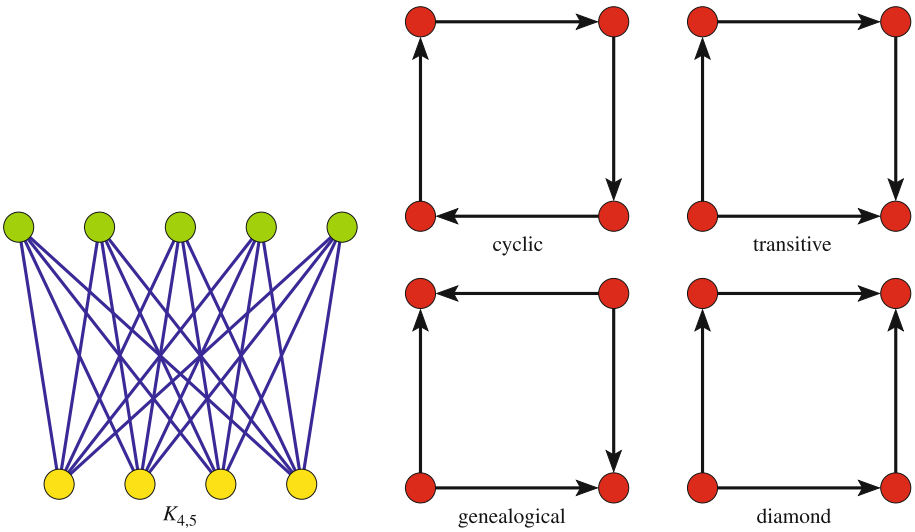


Social Network Analysis, Large-Scale, Figure 18
(247,2)-core and (27,22)-core of IMDB – wrestling

to be the island on the theme *LCD – Liquid crystal display*. In Fig. 16 four islands for transitivity triangular weight from *The Edinburgh Associative Thesaurus* network ($n = 23{,}219$, $m = 325{,}624$) are presented. From the left bottom island of words around the leader 'WORK' we see that the data were collected asking students.

Pattern Searching

If a selected *pattern* determined by a given graph does not occur frequently in a sparse network the straightforward backtracking algorithm applied for pattern searching finds all appearances of the pattern very fast even in



Social Network Analysis, Large-Scale, Figure 19
***$K_{4,5}$* and directed 4-rings**

the case of very large networks. Pattern searching was successfully applied to searching for patterns of atoms in large organic molecula (carbon rings) and searching for relinking marriages in genealogies (Batagelj and Mrvar [12]; Batagelj [6]).

The Fig. 17 presents three connected relinking marriages in the genealogy (represented as a p-graph) of Ragusan noble families. In a p-graph the vertices represent married couples or nonmarried individuals. A solid arc indicates the __ *is a son of* __ relation, and a dotted arc indicates the __ *is a daughter of* __ relation. In all three patterns a brother and a sister from one family found their partners in the same other family.

Two Mode Networks

A network $\mathcal{N} = (\mathcal{V}, \mathcal{L}, w)$ in which the set of vertices $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ is composed of two disjoint sets $\mathcal{V}_1$ and $\mathcal{V}_2$, and $\mathcal{L}$ is a set of *lines* linking $\mathcal{V}_1$ and $\mathcal{V}_2$ is called a *two-mode* or bipartite network.

The two-mode networks often appear in applications, but till recently no directed methods for analysis of larger two-mode networks were available. To identify dense parts of two-mode network, we can use the adapted cores and short rings approaches (Ahmed et al. [3]).

The subset of vertices $C \subseteq \mathcal{V}$ is a (p, q)-core in a two-mode network $\mathcal{N} = (\mathcal{V}_1, \mathcal{V}_2; \mathcal{L})$, $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ iff

a. in the induced subnetwork $\mathcal{H} = (C_1, C_2; \mathcal{L}(C))$, $C_1 = C \cap \mathcal{V}_1$, $C_2 = C \cap \mathcal{V}_2$ it holds for all $v \in C_1$: $\deg_{\mathcal{H}}(v) \geq p$ and for all $v \in C_2$: $\deg_{\mathcal{H}}(v) \geq q$;

b. C is the maximal subset of $\mathcal{V}$ satisfying condition **a**.

The two-mode cores have the following properties:

- $C(0, 0) = \mathcal{V}$
- $C(p, q)$ is not always connected
- $(p_1 \leq p_2) \wedge (q_1 \leq q_2) \Rightarrow C(p_2, q_2) \subseteq C(p_1, q_1)$.

To determine a (p, q)-core an algorithm similar to the ordinary core algorithm can be used: recursively remove from the first set all vertices of degree less than p, and from the second set all vertices of degree less than q. It can be implemented to run in $O(m)$ time.

The main question when applying the bipartite cores is what are the right values of p and q? The most interesting are the values on the 'border' that don't produce too large cores.

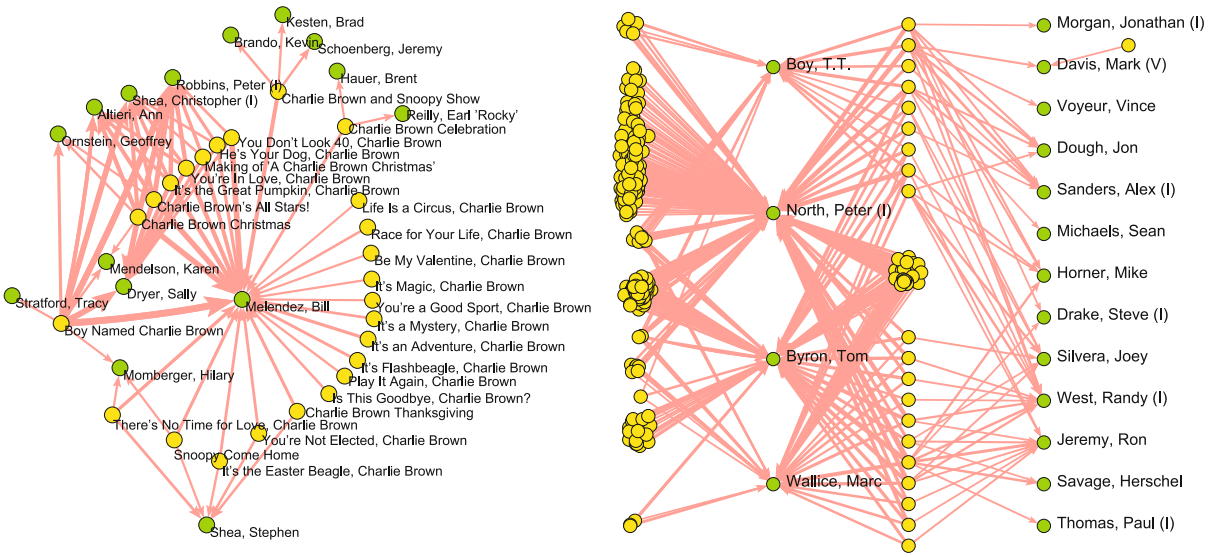
In Fig. 18 the (247,2)-core and (27,22)-core from the Internet Movie Database (two-mode network actors × movies, $n = 1{,}324{,}748 = 428{,}440 + 896{,}308$ vertices and $m = 3{,}792{,}390$ arcs) are presented. Both deal with wrestling.

In 2-mode network there are no 3-rings. The densest substructures are complete bipartite subgraphs $K_{p,q}$ – see $K_{4,5}$ on the left side of Fig. 19. They contain many 4-rings

$$w_4(e) = (p-1)(q-1)\,, \quad \text{for } e \in K_{p,q}\,.$$

There are four types of directed 4-rings – see the right side of Fig. 19. In the case of transitive rings we can count also on how many transitive rings the arc is a *shortcut*.

In the Internet Movie Database, we obtained for w_4 12,465 simple line islands on 56,086 vertices; 30 among



Social Network Analysis, Large-Scale, Figure 20
Islands for w_4 / Charlie Brown and adult

them have size at least 50. Two of them are presented on Fig. 20.

Multiplication of Networks

To a simple two-mode *network* $\mathcal{N} = (\mathcal{I}, \mathcal{J}, \mathcal{E}, w)$; where $\mathcal{I}$ and $\mathcal{J}$ are sets of *vertices*, $\mathcal{E}$ is a set of *edges* linking $\mathcal{I}$ and $\mathcal{J}$, and $w\colon \mathcal{E} \to \mathbb{R}$ is a *weight*; we can assign a *network matrix* $\mathbf{W} = [w_{i,j}]$ with elements: $w_{i,j} = w(i, j)$ for $(i, j) \in \mathcal{E}$ and $w_{i,j} = 0$ otherwise.

Given a pair of compatible networks $\mathcal{N}_A = (\mathcal{I}, \mathcal{K}, \mathcal{E}_A, w_A)$ and $\mathcal{N}_B = (\mathcal{K}, \mathcal{J}, \mathcal{E}_B, w_B)$ with corresponding matrices $\mathbf{A}_{\mathcal{I}\times\mathcal{K}}$ and $\mathbf{B}_{\mathcal{K}\times\mathcal{J}}$ we call a *product of networks* $\mathcal{N}_A$ and $\mathcal{N}_B$ a network $\mathcal{N}_A \star \mathcal{N}_B = \mathcal{N}_C = (\mathcal{I}, \mathcal{J}, \mathcal{E}_C, w_C)$, where $\mathcal{E}_C = \{(i, j)\colon i \in \mathcal{I}, j \in \mathcal{J}, c_{i,j} \neq 0\}$ and $w_C(i, j) = c_{i,j}$ for $(i, j) \in \mathcal{E}_C$. The product matrix $\mathbf{C} = [c_{i,j}]_{\mathcal{I}\times\mathcal{J}} = \mathbf{AB}$ is defined in the standard way

$$c_{i,j} = \sum_{k\in\mathcal{K}} a_{i,k} \cdot b_{k,j} \,.$$

In the case when $\mathcal{I} = \mathcal{K} = \mathcal{J}$ we are dealing with ordinary one-mode networks (with square matrices).

The standard matrix multiplication is too slow to be used for large networks. For sparse large networks, we can multiply much faster considering only nonzero elements. In general the multiplication of large sparse networks is a 'dangerous' operation since the result can 'explode' – it is not sparse. But in many interesting cases, we can assure that also the product is sparse. For example, we can prove:

If at least one of the sparse networks $\mathcal{N}_A$ and $\mathcal{N}_B$ has small maximal degree on K then also the resulting product network $\mathcal{N}_C$ is sparse.

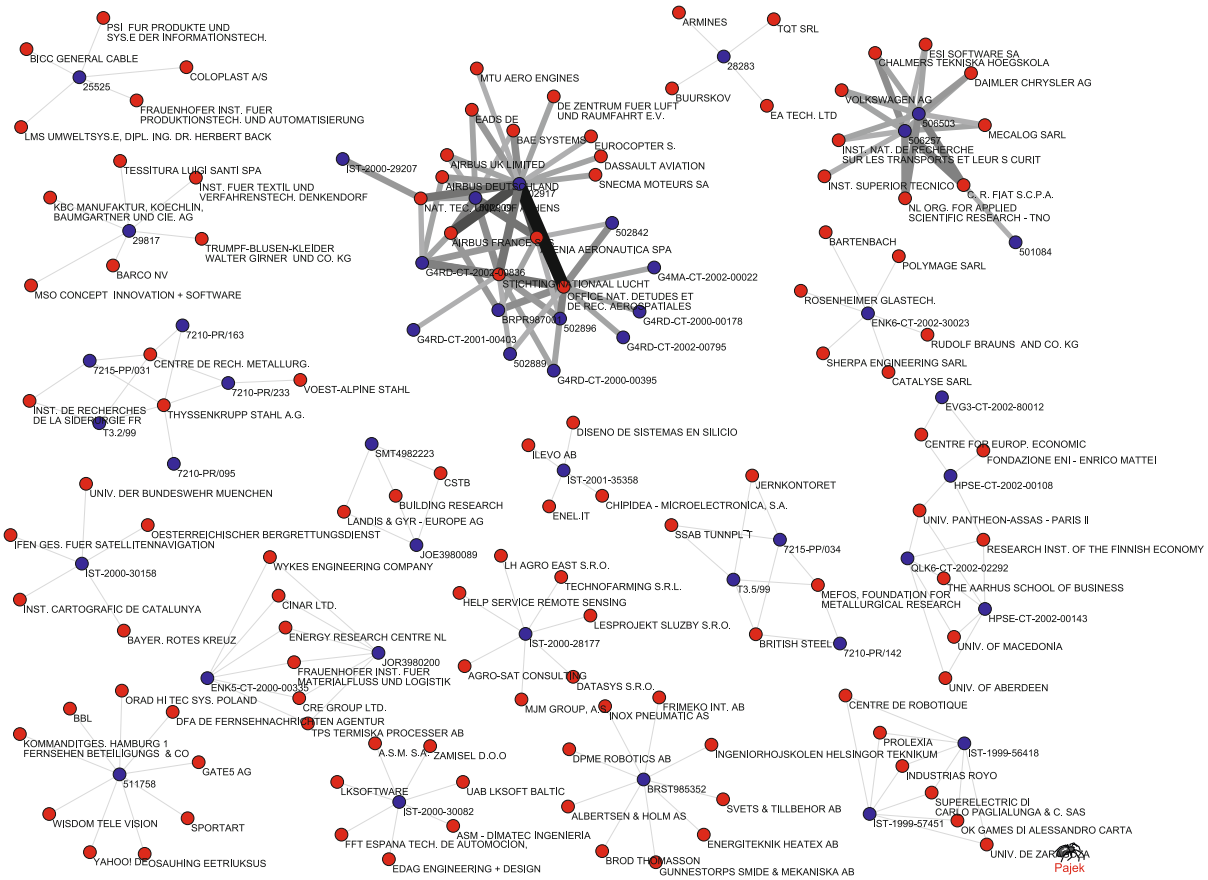
A more detailed analysis gives: Let for $k \in \mathcal{K}$ be $d_{\min}(k) = \min(\deg_A(k), \deg_B(k))$, $\Delta_{\min} = \max_{k\in\mathcal{K}} d_{\min}(k)$, $d_{\max}(k) = \max(\deg_A(k), \deg_B(k))$, $\mathcal{K}(d) = \{k \in \mathcal{K}\colon d_{\max}(k) \geq d\}$, and $d^* = \operatorname{argmin}_d (|\mathcal{K}(d)| \leq d)$. If for the sparse networks $\mathcal{N}_A$ and $\mathcal{N}_B$ the quantities $\Delta_{\min}$ and d^* are small then also the resulting product network $\mathcal{N}_C$ is sparse.

For example, using network multiplication we can in a given genealogy from the basic relations (P – parent-of, L – is a man, J – is a woman) compute all other kinship relations. For details see [12].

An important application of network multiplication is conversion of two-mode network to the corresponding one-mode networks. Often we transform a two-mode network $\mathcal{N}$ into an ordinary (one-mode) network $\mathcal{N}_1 = (\mathcal{I}, \mathcal{E}_1, w_1)$ or/and $\mathcal{N}_2 = (\mathcal{J}, \mathcal{E}_2, w_2)$, where $\mathcal{E}_1$ and w_1 are determined by the matrix $\mathbf{W}^{(1)} = \mathbf{W}\mathbf{W}^{\mathrm{T}}$, $w_{ij}^{(1)} = \sum_{k\in\mathcal{J}} w_{ik} \cdot w_{kj}^{\mathrm{T}}$ and $\mathbf{W}^{\mathrm{T}}$ is the transpose of matrix $\mathbf{W}$. Evidently the matrix $\mathbf{W}^{(1)}$ is symmetric $w_{ij}^{(1)} = w_{ji}^{(1)}$. There is an edge $(i, j) \in \mathcal{E}_1$ in $\mathcal{N}_1$, iff $N(i) \cap N(j) \neq \emptyset$. Its weight is $w_1(i, j) = w_{ij}^{(1)}$. The network $\mathcal{N}_2$ is determined in a similar way by the matrix $\mathbf{W}^{(2)} = \mathbf{W}^{\mathrm{T}}\mathbf{W}$.

The networks $\mathcal{N}_1$ and $\mathcal{N}_2$ are analyzed using standard methods for one-mode networks.

Another very important application of network multiplication is producing different networks from data tables. A *data table* $\mathcal{T}$ is a set of *records* $\mathcal{T} = \{T_k\colon k \in \mathcal{K}\}$,



Social Network Analysis, Large-Scale, Figure 21
The main two islands in ProjInst

where $\mathcal{K}$ is the set of *keys*. A record has the form $T_k = (k, q_1(k), q_2(k), \ldots, q_r(k))$ where $q_i(k)$ is the value of the *property* (attribute) $\mathbf{q}_i$ for the key k.

Suppose that the property $\mathbf{q}$ has the range $2^{\mathcal{Q}}$. For example: Authors[WasFau] = {S. Wasserman, K. Faust}, PubYear [WasFau] = {1994}, … If $\mathcal{Q}$ is finite (it can always be transformed in such set by partitioning the set $\mathcal{Q}$ and recoding the values) we can assign to the property $\mathbf{q}$ a two-mode network $\mathcal{K} \times \mathbf{q} = (\mathcal{K}, \mathcal{Q}, \mathcal{E}, w)$ where $(k, v) \in \mathcal{E}$ iff $v \in q(k)$, and $w(k, v) = 1$.

Also, for properties $\mathbf{q}_i$ and $\mathbf{q}_j$ we can define a two-mode network $\mathbf{q}_i \times \mathbf{q}_j = (\mathcal{Q}_i, \mathcal{Q}_j, \mathcal{E}, w)$ where $(u, v) \in \mathcal{E}$ iff $\exists k \in \mathcal{K}\colon (u \in q_i(k) \wedge v \in q_j(k))$, and $w(u, v) = \mathrm{card}\big(\{k \in \mathcal{K}\colon (u \in q_i(k) \wedge v \in q_j(k))\}\big)$.

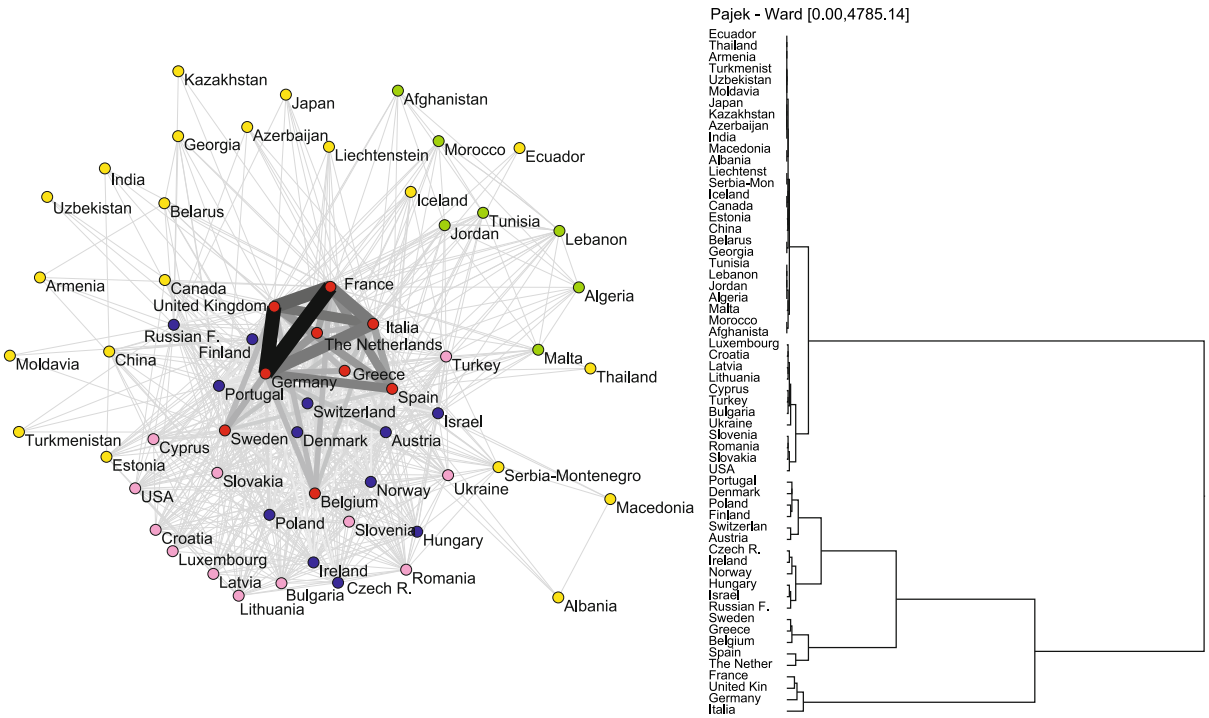
It holds $[\mathbf{q}_i \times \mathbf{q}_j]^{\mathrm{T}} = \mathbf{q}_j \times \mathbf{q}_i$ and $\mathbf{q}_i \times \mathbf{q}_j = [\mathcal{K} \times \mathbf{q}_i]^{\mathrm{T}} \star [\mathcal{K} \times \mathbf{q}_j] = [\mathbf{q}_i \times \mathcal{K}] \star [\mathcal{K} \times \mathbf{q}_j]$.

We can join a pair of properties $\mathbf{q}_i$ and $\mathbf{q}_j$ also with respect to the third property $\mathbf{q}_s$: we get a two-mode network $[\mathbf{q}_i \times \mathbf{q}_j]/\mathbf{q}_s = [\mathbf{q}_i \times \mathbf{q}_s] \star [\mathbf{q}_s \times \mathbf{q}_j]$.

For the meeting ***The Age of Simulation*** at Ars Electronica in Linz, January 2006, a dataset of EU projects on simulation was collected by FAS research, Vienna and stored in the form of Excel table. The rows are the descriptions of projects participants (idents) and columns correspond to different their properties. From this table three two-mode networks were produced: Project – $\mathbf{P}$ = [idents × projects]; Country – $\mathbf{C}$ = [idents × countries]; and Institution – $\mathbf{U}$ = [idents × institutions]; where $|idents| = 8869$, $|projects| = 933$, $|institutions| = 3438$, and $|countries| = 60$.

Since all three networks have the common set (idents) we can derive from them using network multiplication several interesting networks, such as: ProjInst – $\mathbf{W}$ = [projects × institutions] = $\mathbf{P}^{\mathrm{T}} \star \mathbf{U}$; Countries – $\mathbf{S}$ = [countries × countries] = $\mathbf{C}^{\mathrm{T}} \star \mathbf{C}$; and Institutions – $\mathbf{Q}$ = [institutions × institutions]/projects = $\mathbf{W}^{\mathrm{T}} \star \mathbf{W}$.

For identifying important parts of ProjInst network the 4-rings weights were computed and in the obtained



Social Network Analysis, Large-Scale, Figure 22
Collaboration among countries – graph and hierarchical clustering

network the line islands were determined. 101 islands were obtained, 18 of the size at least 5 (see Fig. 21). The two most important islands are: aviation companies and car companies.

In Fig. 22 the collaboration among countries is presented. For dense (sub)-networks we get better visualization by using matrix display. To determine the ordering of vertices we used Ward's clustering procedure with corrected Euclidean distance as dissimilarity measure [21]. The permutation determined by hierarchy can often be improved by changing the positions of clusters in the clustering tree. We get a typical center-periphery structure (see Fig. 23).

Note that in matrix display some details become apparent, such as the collaboration inside the peripherical group Afghanistan, Morocco, Malta, Tunisia, Lebanon, Jordan and Algeria; or a collaboration of Russian Federation with ex-Soviet republics Turkmenistan, Uzbekistan, Moldavia, Kazakhstan, Azerbaijan and Japan.

Statistical Approach

There are many properties *computed* from the network data that give us different information about it. For example:

Global properties Number of vertices, lines (edges/arcs), components; diameter; centralization; maximum core number, ...

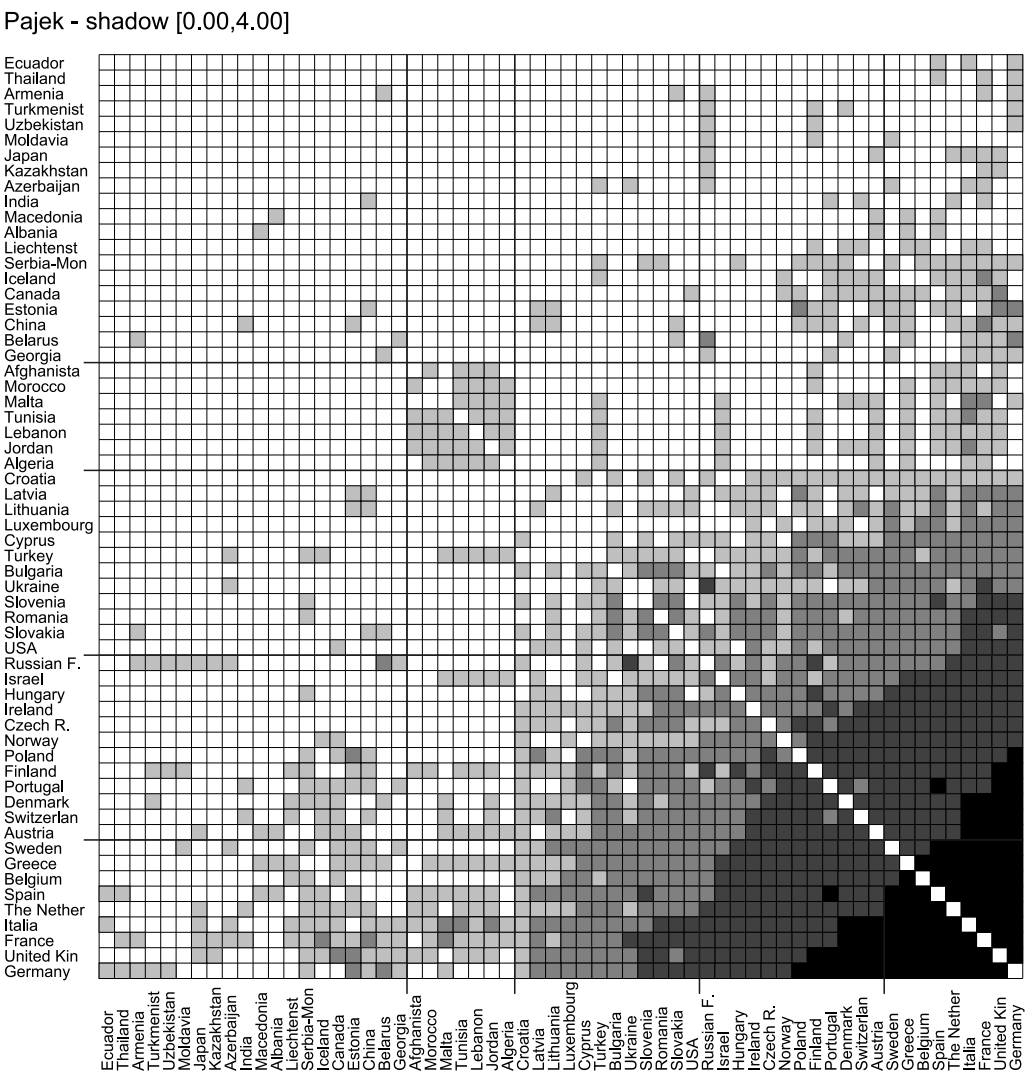
Local properties Degrees, core numbers, indices (betweeness, hubs, authorities, ...). Usually we look at their *distributions* or *inspect* the values of interesting elements.

Another interesting task is searching for associations between computed (structural) data and input (measured) data.

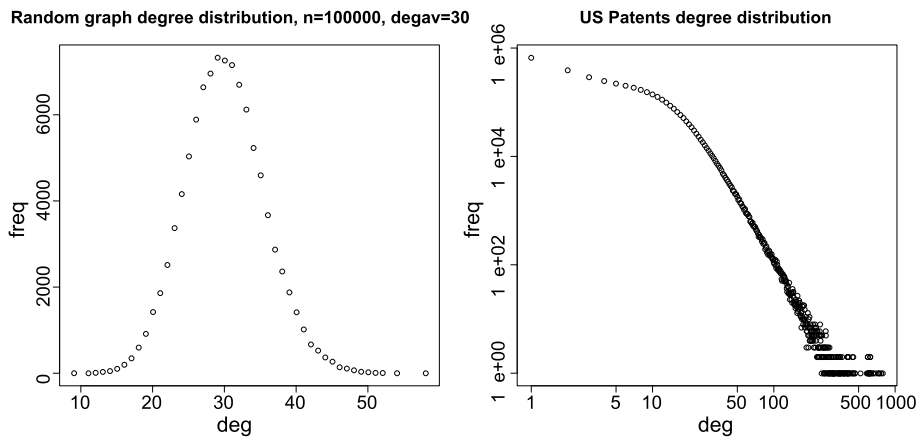
Paul Erdős and Alfréd Rényi introduced in 1959 the notion of random graph in which each pair of vertices is linked with a given probability p. The theory of ER random graphs is well developed (see [15]). Some characteristic results:

- The degree distribution is binomial (in the limit Poisson's) and most of the vertices have degree (very) close to the average degree;
- For $p \geq \frac{1}{n}$ cycles appear in the graph, and soon also the *giant component*;
- For $p \geq \frac{\log_2 n}{n}$ almost all graphs are connected.

Real-life networks are usually not random in the Erdős–Rényi sense. The analysis of their distributions gave new views about their structure.



Social Network Analysis, Large-Scale, Figure 23
Matrix display of collaboration among countries



Social Network Analysis, Large-Scale, Figure 24
Distributions: ER-random and US patents

On the left side of Fig. 24 a degree distribution in ER graph on $n = 100{,}000$ vertices with average degree $\overline{\deg} = 30$ is presented. On the right side a degree distribution for US patents citation network is presented (in log–log scale). Evidently this distribution is very far away from Poisson distribution.

In 1967 a psychologist Stanley Milgram made his experiment with letters. The letter should reach a target person. The persons involved in experiment were asked to send the letter with these instructions to his or her acquaintance that is supposed to be closer (in the acquaintances network) to the target person. The letter was sent from Boston to Omaha. The average length of the successful paths was six – *six degrees of separation*. The average path length on the internet is 19 clicks.

The networks in which the average shortest path length is small are called *small worlds*. Duncan Watts and Steven Strogatz developed in late 90-ties a procedure for construction of (random) small worlds by *rewiring* – an edge is randomly selected and one of its endpoints is attached to same other vertex. After each rewiring step the average length of geodesics is usually decreased because the rewiring creates shortcuts.

Albert-László Barabási from University of Notre Dame in 1998 analyzed several networks and noticed:

- The degree distribution follows the *power law* – the probability p_d that a vertex has a degree d equals to $p_d = cd^{-\gamma}$. In a log–log scale diagram it is represented by a line.
- In a network there exist some vertices with *large degree* (very improbable in ER graphs). These vertices link the network into a single component.

It turned out that most of real life networks (persons – e-mail, phone calls, sexual contacts (drug users, AIDS), collaboration; movie actors – playing in the same movie; proteins – interactions; words – semantic relations; ...) have such characteristics. Because for these networks their degree distribution has no natural scale they were named *scale-free* networks. For a discussion about the notion of scale-free network see [33].

The following was the first explanation (Barabási) of scale-free nature of many real-life networks:

- These networks are growing.
- In this process new vertices are added and linked with new edges to already existing vertices. The random selection of vertex to which a new vertex is attached is not uniform but follows the *preferential attachment* rule – the selection probability is proportional to the degree of a vertex.

Based on this model it can be shown the following:

- The degree distribution is the power law.
- The average length of geodesics is $O(\log n)$.
- These networks are resilient against random vertex or edge removals (random attacks), but quickly become disconnected when large degree nodes (Achilles' heel) are removed (targeted attacks).

Mark Granovetter noticed in 1973 that in social networks groups appear linked with *strong ties* [26]. They link in larger networks with *weak ties*. Also in other real-life networks vertices often form groups – the clustering coefficient is larger than in ER networks.

Several improvements and alternative models were proposed that also produce scale-free networks with some additional properties characteristic for real-life networks: copying [31], combining random and preferential attachment [37], R-mat [18], forest fire [32], aging, fitness, nonlinear preferences, ...

There are several applications of the scale-free networks theory. For example searching (Adamic et al. [2]) and spreading of epidemies (Barthélemy, Barrat, Pastor-Sattoras, Vespignani, Complex Networks Collaboratory).

For general overviews see [4,22,35,36].

Future Directions

In 2005 the support for *multi-relational* networks was introduced in Pajek. Combined with *temporal* networks it enables analysis of new kinds of networks – such as KEDS networks (*Kansas Event Data System* or *Tabari*). These networks are usually small in terms of vertices but can be (very) large in terms of lines – different interaction events among actors.

The last developed approach for analysis of large networks is adaptation of hierarchical clustering with relational constraints based on Ferligoj and Batagelj [23] to large networks. The basic idea to get a fast algorithm is to compute the dissimilarities between units (vertices) only for the linked pairs of units [11]. This approach is one of the possible approaches to analysis of spatial networks.

There are still several fields of social network analysis for which efficient approaches to deal with large networks have to be developed such as blockmodeling, probabilistic models, ...

In the near future new versions of network analysis software will appear using very large computer memories enabled by the new 64-bit computer architecture. A special challenge is development of methods and software for analysis of huge networks.

Bibliography

Primary Literature

1. Abello J, Pardalos PM, Resende MG (2002) Handbook of Massive Data Sets. Springer, Heidelberg
2. Adamic LA, Lukose RM, Huberman BA (2002) Local Search in Unstructured Networks. In: Bornholdt S, Schuster HG (eds) Handbook of Graphs and Networks: From the Genome to the Internet. Wiley-VCH, Berlin
3. Ahmed A, Batagelj V, Fu X, Hong SH, Merrick D, Mrvar A (2007) Visualisation and Analysis of the Internet Movie Database. Proceedings of the Asia-Pacific Symposium on Visualisation (APVIS2007), Sydney, Australia, 5–7 Feb. IEEE, New York, pp 17–24
4. Albert R, Barabási AL (2002) Statistical mechanics of complex networks. Rev Mod Phys 74:47–97
5. Alvarez-Hamelin JI, Dall'Asta L, Barrat A, Vespignani A (2005) *k*-core decomposition: a tool for the visualization of large scale networks. cs.NI/0504107 published in Advances in Neural Information Processing Systems 18, Canada, 2006
6. Batagelj V (1989) Similarity measures between structured objects. In: Graovac A (ed) Proceedings of International Course and Conference on the Interfaces between Mathematics, Chemistry and Computer Science, Dubrovnik, 20–25 June 1988. Studies in Physical and Theoretical Chemistry, vol 63. Elsevier/North-Holland, Amsterdam, pp 25–40
7. Batagelj V, Brandes U (2005) Efficient Generation of Large Random Networks. Phys Rev E 71:036113
8. Batagelj V, Ferligoj A (2000) Clustering relational data. In: Gaul W, Opitz O, Schader M (eds) Data Analysis. Springer, Berlin, pp 3–15
9. Batagelj V, Mrvar A (2000) Some Analyses of Erdős Collaboration Graph. Soc Netw 22:173–186
10. Batagelj V, Mrvar A (2001) A Subquadratic Triad Census Algorithm for Large Sparse Networks with Small Maximum Degree. Soc Netw 23:237–43
11. Batagelj V, Mrvar A (2007) Hierarchical clustering with relational constraints of large data sets. 6th Slovenian International Conference on Graph Theory, Bled, 24–30 June
12. Batagelj V, Mrvar A (2008) Analysis of kinship relations with Pajek. Soc Sci Comput Rev 26(2):224–246
13. Batagelj V, Zaveršnik M (2002) Generalized Cores. arxiv cs.DS/0202039
14. Batagelj V, Zaveršnik M (2007) Short cycle connectivity. Discret Math 307(3–5):310–318
15. Bollobás B (2001) Random Graphs. Cambridge University Press, Cambridge
16. Brandes U (2001) A Faster Algorithm for Betweenness Centrality. J Math Soc 25(2):163–177
17. Breiger RL (2004) The analysis of social networks. In: Hardy M, Bryman A (eds) Handbook of data analysis. Sage, London, pp 505–526
18. Chakrabarti D, Zhan Y, Faloutsos C (2004) R-MAT: A Recursive Model for Graph Mining. In: SIAM Data Mining 2004, Orlando, Florida, SIAM
19. Cormen TH, Leiserson CE, Rivest RL, Stein C (2001) Introduction to algorithms. MIT Press, Cambridge
20. Doreian P, Batagelj V, Ferligoj A (2000) Symmetric-Acyclic Decompositions of Networks. J Classif 17(1):3–28
21. Doreian P, Batagelj V, Ferligoj A (2005) Generalized Blockmodeling. Cambridge University Press, Cambridge
22. Dorogovtsev SN, Mendes JFF (2003) Evolution of networks: from biological nets to the internet and www. Oxford University Press, Oxford
23. Ferligoj A, Batagelj V (1983) Some types of clustering with relational constraints. Psychometrika 48(4):541–552
24. Freeman LC (1979) Centrality in Social Networks: A Conceptual Clarification. Soc Netw 1:211–213
25. Garfield E, Sher IH, Torpie RJ (1964) The Use of Citation Data in Writing the History of Science. The Institute for Scientific Information, Philadelphia
26. Granovetter M (1973) The Strength of Weak Ties. Am J Sociol 78:1360–80
27. Harary F, Norman RZ, Cartwright D (1965) Structural Models: An Introduction to the Theory of Directed Graphs. Wiley, New York
28. Huisman M, van Duijn MAJ (2005) Software for social network analysis. In: Carrington PJ, Scott J, Wasserman S (eds) Models and methods in social network analysis. Cambridge University Press, Cambridge, pp 270–316
29. Hummon NP, Doreian P (1990) Computational Methods for Social Network Analysis. Soc Netw 12:273–288
30. Kleinberg J (1998) Authoritative sources in a hyperlinked environment. Proc 9th ACM-SIAM Symposium on Discrete Algorithms
31. Kleinberg J, Kumar R, Raghavan P, Rajagopalan S, Tomkins A (1999) The Web as a graph: measurements, models and methods. Proc of the 5th International Computing and combinatorics Conference
32. Leskovec J, Kleinberg J, Faloutsos C (2006) Laws of Graph Evolution: Densification and Shrinking Diameters. ACM Transactions on Knowledge Discovery from Data (TKDD) vol 1, issue 1, article 2
33. Li L, Alderson D, Tanaka R, Doyle JC, Willinger W (2007) Towards a Theory of Scale-Free Graphs: Definition, Properties, and Implications. cond-mat/0501169, Internet Math 2(4):431–523
34. Mane KK, Börner K (2004) Mapping topics and topic bursts in PNAS. Proc Natl Acad Sci USA 101:5287–5290
35. Newman MEJ (2003) The structure and function of complex networks. SIAM Rev 45:167–256
36. Newman MEJ, Barabási AL, Watts D (2006) The Structure and Dynamics of Networks. Princeton Studies in Complexity. Princeton University Press, Princeton
37. Pennock DM, Flake GW, Lawrence S, Glover EJ, Giles CL (2002) Winners don't take all: Characterizing the competition for links on the web. Proc Natl Acad Sci USA 99(8):5207–5211
38. Seidman SB (1983) Network Structure And Minimum Degree. Soc Netw 5:269–287
39. Schank T, Wagner D (2005) Finding, counting and listing all triangles in large graphs, an experimental study. In: Workshop on Experimental and Efficient Algorithms (WEA). Lecture Notes in Computer Science, vol 3503, Springer, pp 606–609
40. Snyder D, Kick E (1979) The World System and World Trade: An Empirical Exploration of Conceptual Conflicts. Sociol Q 20(1):23–36
41. Snijders TAB (2005) Models for Longitudinal Network Data. In: Carrington P, Scott J, Wasserman S (eds) Models and methods in social network analysis. Cambridge University Press, New York

42. Stuckenschmidt H, Klein M (2004) Structure-Based Partitioning of Large Concept Hierarchies. Proc of the 3rd International Semantic Web Conference ISWC 2004, Hiroshima, Japan
43. Wasserman S, Faust K (1994) Social Network Analysis: Methods and Applications. Cambridge University Press, Cambridge
44. White DR, Batagelj V, Mrvar A (1999) Analyzing Large Kinship and Marriage Networks with Pgraph and Pajek. Soc Sci Comput Rev 17:245–274
45. Zaveršnik M, Batagelj V (2004) Islands. Slides from Sunbelt XXIV, Portorož, Slovenia, 12–16 May

Books and Reviews

Ahuja RK, Magnanti TL, Orlin JB (1993) Network Flows: Theory, Algorithms, and Applications. Prentice Hall, Englewood Cliffs
Batagelj V, Mrvar A (2003) Pajek – Analysis and Visualization of Large Networks. In: Jünger M, Mutzel P (eds) Graph Drawing Software. Springer, Berlin, pp 77–103
Brandes U, Erlebach T (2005) Network Analysis: Methodological Foundations. Lecture Notes in Computer Science. Springer, Berlin
Carrington PJ, Scott J, Wasserman S (2005) Models and Methods in Social Network Analysis. Cambridge University Press, Cambridge
Degenne A, Forsé M (1999) Introducing Social Networks. SAGE Publications, London
de Nooy W, Mrvar A, Batagelj V (2005) Exploratory Social Network Analysis with Pajek. Cambridge University Press, Cambridge
Knuth DE (1993) The Stanford GraphBase: A Platform for Combinatorial Computing. Addison-Wesley, Reading
Scott JP (2000) Social Network Analysis: A Handbook. SAGE Publications, London

Web Resources

Center for Complex Network Research, Notre Dame: http://www.nd.edu/networks/
Center for Spatially Integrated Social Science: http://www.csiss.org/
Complex Networks Collaboratory: http://cxnets.googlepages.com/
Internet Movie Database http://www.imdb.com/
Matthieu Latapy. Triangle computation web page. http://www-rp.lip6.fr/latapy/Triangles/
Nber: http://www.nber.org/patents/
Netminer: http://www.netminer.com/
Pajek: http://pajek.imfm.si data sets: http://vlado.fmf.uni-lj.si/pub/networks/data/
The Edinburgh Associative Thesaurus: http://www.eat.rl.ac.uk/
The Kansas Event Data System: http://web.ku.edu/keds/
UCINET: http://www.analytictech.com/

Social Network Analysis, Overview of

JOHN SCOTT
School of Law and Social Science,
University of Plymouth, Plymouth, UK

Article Outline

Glossary
Definition of the Subject
The Development of Social Network Analysis
Graph Theory and Ideas of Balance
Diffusion Processes
Algebraic Models and Blockmodeling
Scaling models and Visualization
Statistical Models for Hypothesis Testing
Agent-Based Computational Models and Temporal Processes
Small World Models and Network Dynamics
Bibliography

Glossary

Algebraic models Approaches to network analysis that use algebraic methods for studying sets of points to produce positional analyses such as blockmodels.

Agent-based computational models Models of dynamic systems that focus on agent-level properties, seeing system level changes as consequences of the interaction of rule-following agents.

Balance theories A number of related theories concerning both the psychological state of consonance or dissonance found in a person's ideas and affects and the equilibrium or disequilibrium of one person's relations with another. The exploration of the mathematical principles of these states.

Blockmodeling An approach to positional analysis that uses algebraic methods to construct image graphs in which blocks represent connections or the absence of a connection between sets of points.

Graph theory A basic method for the analysis of networks in which the relational properties of the members of a set are seen in terms of points and lines. Pairwise connections among points are used to generate and explore system-level phenomena such as density and centralization.

Diffusion processes The processes through which innovations and other changes spread through a network, the flow and pace of change being determined by the structure of the network.

Scaling methods Geometrical techniques for displaying and analyzing a network as a mapping of points located in a multidimensional space.

Small-world models Network models based on graphs of relatively low density but high reachability. Originated in psychological experiments on the communicative effectiveness of interpersonal acquaintance.

Definition of the Subject

The formal idea of using a network as a tool of analysis originated in electrical engineering as a way of envisaging and modeling the flows of electricity through national power grids and in domestic and industrial settings. It was merely a short step to the application of the idea in civil engineering, where water supply, drainage, roads and railways could all be regarded as networks. Much later this idea was central to electronic and computer systems engineering. Parallel methodological innovations had occurred in areas of physical geography, where river drainage systems had been envisaged as networks of flows that could be modeled in mathematical terms, and this led to interest among human geographers about the potential for applying the idea. The idea that social relationships could be seen as forming a network or reticule of relations that are capable of analysis in formal, mathematical terms suggested novel ways of considering social phenomena and promised insights that were unavailable through alternative approaches.

Social network analysis first emerged as an approach distinctive from statistical analysis within anthropology, social psychology, and sociology. It went on to influence theoretical and methodological developments in many other areas of social sciences. In all these areas, researchers took up mathematical models from graph theory, which they used to investigate structures of relations. Researchers who were working in the area of social network analysis soon began to broaden their framework of analysis by exploring the implications of other mathematical models, such as algebraic set theory, scaling methods, and statistical and stochastic methods. In this way, social network analysis became linked to the larger intellectual debates going on in other areas of the natural and social sciences where these same methods were being drawn upon, and a number of social network analysts have contributed to the development of these and other mathematical models. Social network analysts have promoted and undertaken the application of these ideas in many areas of sociology, most notably in economic sociology, political sociology, the study of family and kinship, community studies, the investigation of world systems, and numerous other areas.

Investigation into social networks has now become far more fashionable than ever before. However, it is in the nature of intellectual fashions that much of the new interest comes from those who are least familiar with the area. This is certainly true of social network analysis, where many newcomers were unaware of its long history. They were influenced, in particular, by the exaggerated claims to intellectual novelty that were being made by some physicists who have advocated random-network and small-world models as the basis of what they call a "new science of networks". Indeed, many people have assumed that social network analysis began de novo with these models and they know little about the applications already made using other models. Despite their exaggerated claims, the work of the physicists has provided us with some powerful new models, and these have had a major impact beyond the specifically social sphere – this fact will be apparent from considering chapters in a number of the sections in this *Encyclopedia*. However, the physicists have not provided a complete replacement for social network analysis, as they sometimes imply, but have offered a broadening of its range and concerns.

This overview article will discuss the development of social network analysis and will sketch the various mathematical models that have been used to explore the nature of complex social systems. It begins with an overview of the history of social network analysis and its main applications. This is followed by a consideration of graph theory, the earliest and principal approach to the study of social networks. The following sections look at later and more advanced methods: algebraic methods and blockmodeling, scaling methods and visualization techniques, statistical methods, agent-based computational methods, and finally the small-world methods of the physicists. The latter discussion will place the new social physics in the context of earlier discussions and it will show the variety of mathematical models that have a part to play in contemporary social network analysis. The various methods and models are considered in outline and without technical detail. The various specialist chapters included in this and other sections of the Encyclopedia cover these methods in greater detail.

The Development of Social Network Analysis

Social network analysis has a long history, though it appeared under its specific name only far more recently (see Chap. 2, p. 23 in [51]). Although structural thinking was central to classical sociology, the intellectual roots of a specifically network perspective can be traced back to those sociologists of the late nineteenth and early twentieth centuries who used the idea of a network of relations as a metaphor for understanding social relations as *structures* of relations. Georg Simmel and other German sociologists, in particular, wrote extensively on the "interweaving" of actions to form complex configurations or "webs" of connection. Their metaphorical usage of the idea drew explicitly from the textile industry and its ideas of the weft and warp of woven fabrics. The first proper usage of the

actual word "network" in a sociological context, however, was in the work of the anthropologist Alfred Radcliffe-Brown [45], who used it to suggest a more formal focus for structural analysis. Simultaneously with this, a number of social psychologists working in the area of child study and field theory began to construct diagrams of social relations that they gradually formalized as "sociograms" to show the chains of positive and negative connections that they were finding among the individuals in their experimental studies.

The crucial social psychological work on social networks was that of the psychotherapist Jacob Moreno [41]. Following his arrival in the United States he began to combine psychodrama with studies of the influence of friendship choices on personality development among schoolchildren. His key innovation was to make use of sociograms to depict the friendship patterns that he was able to discover within small groups such as a school class or a neighborhood peer group. Moreno made many intellectual contacts, most importantly with the *Gestalt* psychologist Kurt Lewin, who had himself recently migrated to the United States. Moreno and Lewin exchanged many ideas – and the nature of the exchange became a matter of contention between them in later years. Lewin soon began to move towards a more social form of psychology that stressed the importance of looking at organized patterns of group relations in particular spheres of activity as forming systemic wholes. Such wholes could be seen, by analogy with electromagnetic and gravitational fields, as social fields through which influences are able to flow from one area or part to another. The aim of Lewin's field theory was to map and to explore the "dynamics" of these fields of interpersonal influence.

Both Moreno and Lewin used diagrammatic representations of social relations, but it was Moreno who invented the classic form of the sociogram. A sociogram consists of points, representing individuals, and lines, representing the social relationships among them. Moreno argued that these patterns of points and lines – later described more formally as "vertices" and "edges" – could be described as "social configurations" in which could be charted such things as the direction of friendship choices and their intensity. This work came to be referred to as "sociometry", and a journal with that name was founded in 1937. Through the work published in this journal, the ideas of Moreno and Lewin on configurations of social relations in small groups gradually gave rise to a more mathematically oriented approach that went under the name of "group dynamics". It was in this school of thought that ideas from graph theory – first formulated by René König [33] – began to be employed from the late 1940s, these ideas becoming the basis of a more formal and rigorous study of the properties of the social networks of small groups. By treating the sociogram as a graph, it proved possible to chart dyadic, triadic, and more complex structures and to analyze the states of "balance" and imbalance that existed among positive and negative relations within each network [3,13,28].

It was during the 1930s and 1940s that anthropological and sociological work also began to move in a similar direction. In 1929, the anthropologist W. Lloyd Warner, a former colleague of Radcliffe-Brown, had moved to Harvard University in order to work with the psychologist Elton Mayo on a series of experimental studies of work behavior. Their aim was to use the techniques already employed by anthropologists in studies of small tribes and villages, but to carry out new studies of small groups in the factory and office settings of urban societies. Influenced also by the early system theory constructed by the physiologists Lawrence Henderson and Walter Cannon, they began a series of empirical studies at the Hawthorne electrical works in Chicago. The so-called Hawthorne studies became a milestone in social research – not least because of their significance in developing the idea of the social network.

The Hawthorne studies [47] had begun as management-run investigations into worker efficiency and its relationship to the physical conditions of work. The early experiments examined such things as the consequences of changes in levels of heating and lighting and the provision of work breaks. The studies turned from physical to social factors under Warner and Mayo, and Mayo famously reported that the crucial factor in increasing worker productivity had been their involvement in the experiment itself. The paradoxical results showing increased productivity even when physical conditions worsened were interpreted as showing the high levels of satisfaction that workers felt simply from being part of the management experiment: it was the first time that anyone from management had shown any direct interest in them. Adopting more anthropological methods of observation in the bank wiring room, the researchers documented the gradual building of informal work group relations and forms of social solidarity among those in the experimental group. These informal relations ran in parallel with the formal relations of the managerial organization chart, and the researchers used sociograms to chart the informal social networks among workers.

The sociograms drawn by the Hawthorne researchers allowed them to identify "cliques" and other sub-groups within the larger group, and Warner went on to investigate this kind of group structure further in a series of

investigations into community relations in the New England city of Newburyport (referred to as "Yankee City" in [59,60,61,62]), the Mississippi city of Natchez ("Old City" in [15]), and the Indiana city of Morris ("Jonesville" in [58]). In this work, Warner and his colleagues used network ideas to explore the formation of cliques, "crowds", and "circles" among those they studied, and they examined these in relation to the more formal structures of economic and political life. Their investigations identified these social circles through a rough and ready use of the Venn diagrams of set theory and the methods of matrix display. Using these simple methods, they were able to identify such things as class divisions in clique membership and the effects of ethnicity on the structure of network relations.

It was during the late 1940s and early 1950s that a significant consolidation of the social network approaches that had been developing began to be forged. It was to be another decade, however, before a recognizable tradition of social network analysis was properly established. George Homans, a colleague of Warner and Mayo at Harvard, attempted a synthesis of the anthropological and sociological ideas as the basis of an alternative sociology to the system theory of Talcott Parsons [30]. Homans re-analyzed some of the earlier data using rudimentary techniques of matrix rearrangement, his aim being to bring out the clique structure more clearly. Homans' method of manually "reshuffling" the 0s and 1s in a matrix to bring out blocks of high and low density was a pioneering move towards the more formal methods of blockmodeling that developed during the 1970s. Homans himself interpreted his data on social relations in terms of the frequency and duration of the relationships involved and went on to develop rational choice models of action to explain the findings [31].

The early 1950s also saw the development of more systematic anthropological work on social networks in Britain. The Manchester University anthropologist Max Gluckman established the Rhodes-Livingstone Institute in Northern Rhodesia (now Zambia) to study African social structures in village and town settings. An early associate of the Institute was John Barnes, who undertook independent fieldwork in Bremnes, Norway. It was here that he recognized the importance of the network metaphor – famously while observing the repaired fishing nets of the local fishermen – and he advocated the study of whole communities as networks of points and lines [2]. Elizabeth Bott, a student of Lloyd Warner and then working at the Tavistock Institute (associated with the Michigan Research Center for Group Dynamics), took up this metaphor in her own investigations into the kinship networks of London families [6,7]. Because of her position at the Tavistock, she was encouraged to link this to the work of Moreno and to the early field theory. This work was systematized by Siegfried Nadel and Clyde Mitchell. Nadel [42] aimed to establish a mathematical basis for social network research, but his early death prevented him from carrying this further. Mitchell [39], however, made the network idea the basis for a series of studies at the Rhodes-Livingstone Institute and set out a number of conceptual innovations – density, multiplexity, durability, direction, reachability and so forth – that echoed and took further the ideas of Homans.

By 1969, then, a clear body of network ideas had been established. in anthropology and social psychology, and this was coming to be organized around the mathematical theory of graphs. It was at this point, however, that an alternative framework began to develop in the United States, and this eventually subsumed the earlier approach in a comprehensive framework of social network analysis. John Boyd, François Lorrain, and Harrison White had begun to explore some aspects of Lévi-Strauss's analysis of kinship structures using algebraic set theory [72], while Edward Laumann [34] and Joel Levine [35] began to use multidimensional scaling to investigate social structure. White moved to Harvard University and brought together a strikingly original group of young researchers who were interested in using mathematical methods to explore aspects of social networks and social relations.

One of the earliest products of this group was Mark Granovetter's [26,27] influential argument about "the strength of weak ties" – the first, and most important, of the counter-intuitive results to come out of this more sophisticated form of network analysis. Granovetter's argument held that the chances that a person will receive useful information from those in his or her social network is greater when a person has acquaintances rather than close and intimate friends. One's intimates tend to have access to the same pool of information through their dense pattern of mutual connections. Acquaintances, on the other hand are connected to quite diverse and distant social networks and so can communicate a much greater range of information. Paradoxically, then, it is the weak acquaintanceship ties that are the "strongest" or most important carriers of information. Having a small number of acquaintances can be more useful in a job search than having a large number of intimates.

The ability of White's group to pursue this program of social network analysis was made possible by the advances that were being made in computing technology and that were then becoming available in software programs. Earlier work had used manual techniques and so

had been limited to relatively small groups. Mainframe computers and Fortran programming made possible the analysis of much larger data sets and the use of more powerful mathematics. As the members of the Harvard group spread to Graduate centers across North America, their work became a major influence within sociology and encouraged the global spread of formal social network analysis. Barry Wellman and Steven Berkowitz were particularly active in setting up the *International Network for Social Network Analysis* (http://www.insna.org/) and this became the basis for the founding of the journal *Social Networks* and for their own powerful syntheses in social network analysis [4,71] (see also [32]). Powerful methodological ideas and sociological applications were developed in the United States by Ronald Breiger, Ronald Burt [10,11], Patrick Doreian, and others. Social network analysis was pursued in Canada by Barry Wellman and Peter Carrington, by John Scott in Britain [51], by Rob Mokken and Frans Stokman in the Netherlands (the developers of the *Gradap* program used in [29]), and by Charles Crothers and Malcolm Alexander in Australia. Stanley Wasserman led the way in producing a standard textbook for social network methods [64] and other consolidating texts appeared [24,37]. By the 1990s, social network analysis had become one of the most strongly established research specialisms in Sociology, with a strong intellectual base, advanced methods and texts [12,18], and a comprehensive range of exemplary studies.

During the 1990s, however, a rather paradoxical development occurred. The exhaustion of many of the conventional research programmes in physics had led a number of theoretical physicists to search out new areas in which to apply their theories. In 1998, Duncan Watts and Steven Strogatz published a paper [68] that revisited some of the ideas on random networks that had grown out of Stanley Milgram's work on "small worlds" [38,56]. This caused a great stir among physicists and led Albértó-Laszlo Barabási to propose the building of a "new science of networks" with the potential for applications across the social sciences [1]. In apparent ignorance of the prior existence of social network analysis, he claimed that it was only during the 1990s – and thanks to physicists' models – that investigators had become aware of the fact that social networks have structures and show orderly patterns of development.

This claim startled and dumbfounded most sociologists working on social networks, for whom the existence of social structure was a fundamental axiom. They resented the ignoring of what they had already achieved and the misrepresentation of sociological understandings of the social world that it involved. Despite its shaky historiography, the arguments of Bárabasi and other social physicists had a major influence outside the discipline and among those in sociological specialisms that were, so far, untouched by social network analysis. Works by Duncan Watts were particularly important [66,67]. The new social physics of networks was especially well-received in the popular and scientific press and resulted in a number of popularizing works (see, for example [8]).

It is, nevertheless, true that the work of the physicists has certainly been responsible for a substantial growth of interest in social network analysis during the last decade. The reason for the appeal of this new work lies in its promise to provide a dynamic model of social networks, rather than the more static ones of the past. While the claim that social network analysis lacks a concern for dynamic processes is, perhaps, overstated, it is certainly the case that the new methods do seem to offer the possibility of a nuanced account of structural change in social networks. While much of the work is still presented without any recognition of prior sociological contributions, sociologists, for their part, are beginning to discuss how these ideas might enrich social network analysis. The arguments about network dynamics are, in fact, converging with some of the work on agent-based computational models already being undertaken within Sociology.

A very important issue that has arisen in the use of the mathematics of graphs is that of whether it provides a substantive *theory* or simply a *method* of analysis. Some researchers hold that the mathematical theorems that can be derived from graph theory can be directly translated into substantive theorems concerning particular domains of application. According to this point of view, sociological laws can be derived directly from mathematical laws. Those who see social network analysis as a method – perhaps a majority of those working in the area – hold that mathematical theorems provide simply the formal constraints to which social processes must conform: a knowledge of mathematics does not, in itself, obviate the need to study social phenomena empirically in order to understand the implications of these theorems for particular types of social relations.

The remainder of this overview article will set out the main intellectual strands within contemporary social network analysis. These areas are discussed in greater detail in the specialist articles in this and other sections of the *Encyclopedia*, and their mathematical bases will not be presented in any detail. The overview will concentrate on the general principles involved in each area and on the ways in which they relate to each other.

Graph Theory and Ideas of Balance

Graph theory, as already noted, is that branch of mathematics that had the earliest influence on the development of social network analysis (see ► Social Network Analysis, Graph Theoretical Approaches to). In graph theory, a graph or network is simply a collection of points (vertices) connected by lines (edges). Graph theory comprises a set of mathematical axioms and derivations that describe the actual and potential patterns formed by these points and lines. This basic model of a graph can be applied widely in many substantive domains: electrical wiring networks, river drainage networks, road and transportation networks, and computer networks have all been investigated using graph theory.

A basic graph consists of simple points and lines, but the lines in a graph can be assigned positive or negative attributes, directions, or numerical values, and the attributes of points can be recorded. Complex mathematical theorems can be constructed for each of the different types of graph. Some of the earliest work to use these ideas focused on the question of "balance" in graphs. Balance is a term that describes the pattern of signs and numerical values in a graph in terms of the reciprocity and transitivity of the relations. A graph of friendship relations, for example, might be unbalanced if two people who each like one another have opposite relationships to a third person. If, on the other hand, all three people liked one another with a similar intensity, the graph would be balanced. Balance theorists have analyzed complex graphs in terms of such "triads" of relations. By producing a census of triads and computing the balance or imbalance in each triad, it is possible to arrive at a calculation of the overall balance in the graph as a whole. Such ideas of balance have generally been assimilated to the idea of equilibrium analyzed in classical mechanical models. A graph that is in a state of disequilibrium, therefore, will be marked by tensions and stresses that push and pull it towards a state of equilibrium. Where a social network is mapped as an unbalanced graph, it can be hypothesized that the participants in the network will experience these stresses as push and pull factors on the social relations that they maintain with each other. The overall network can, therefore, be seen as a dynamic field of causal influences.

Much of the work that has been carried out using the idea of balance has considered focal actors to be the points of reference, and much of the early work that applied graph theory pursued such "ego-centric" ideas. From this point of view, the chosen actor is seen as a point that is "adjacent" to certain other points that comprise its immediate "neighborhood". Each point can, therefore, be given an adjacency score. Actors vary in terms of the number of others with whom they are connected – the number of friends, number of enemies, number of political associates, and so on. People or groups may, therefore, be more or less connected into their neighborhood of others and so they may be compared in terms of their adjacency scores. The early work of Moreno had identified school children as sociometric "stars" or "isolates" according to the number of friends they had. The stars were the especially popular class members, while the isolates had few or no friends in the class. The adjacent others may have more or fewer connections among themselves: my friends may or may not be friendly with each other, my business associates may or may not do business with each other, and so on. Thus, a basic ego-centric measure is the "density" of an actor's neighborhood. This measure is the number of links that actually exist among immediate contacts expressed as a proportion of the number of possible contacts that they could, in principle, sustain among themselves. Where a large proportion of connections intersect, the actor's neighborhood has a high density.

Researchers recognized fairly early on the importance of going beyond such ego-centric measures to more global or socio-centric measures of the network as a whole. In considering a village, for example, a researcher may want to know not only the distribution of individuals according to the densities of their friendship relations, but also the density of friendship connections in the village considered as a whole. Such measures of density would indicate the state of social cohesion or social solidarity in the network. The basic ego-centric measure of density can easily be extended to the network as a whole – if appropriate data are available – by computing the proportion of possible links within the village that are actually established. Calculations of density can also be made for large social networks that must be studied through sample investigations. If it can be assumed that the sample is representative, then the distribution of adjacency scores for sample members can be used to compute the overall density of the network.

Density has become one of the most widely used measures in studies that employ graph theory and it has been used to indicate the changing character of social solidarity and the state of "community" that exists in a group. Wellman's work [69,70] has demonstrated that growing amounts of geographical mobility have not led to a commensurate decline in community but have resulted, instead, in the transformation of patterns of solidarity. People in large cities, he shows, are likely to interact less frequently and less intensely with numerous others on a day to day basis as they move to live and work in other parts of the city. Thanks to cars, the telephone, and the internet,

however, they are able to maintain more connections, but on a less frequent and less intense basis.

Particularly important measures of global connectedness within a graph are based on the "distance" from one point to another and the corresponding "centrality" that points have in the overall pattern of connection. The distance between two points in graph theoretical terms is simply the number of lines that must be traversed in moving from one point to another. Thus, two adjacent points stand at distance 1 from each other, while two points that are connected through a common neighbor are at distance 2 from each other. It is generally assumed that influence and communication within a social network attenuates with increasing distance. Nevertheless, the rate of attenuation and the particular level of distance that it is reasonable to regard as constituting a socially significant connection are quite variable from one type of network to another. Distant friends of friends, for example, may be insignificant, while distant financial donors to a political party or candidate may be more significant. The criteria of significance, it will be apparent, rests always on a sociological judgment and never on a mathematical measure alone.

The principal use of distance measures has been to assess the relative centrality or peripherality of points in a graph. A point whose average distance to all other points in the graph is low can be regarded as occupying a central position in the network as a whole, while a point with a very high average distance from all others is peripheral to it. Applications of these ideas in studies of interlocking directorships – the lines created when company directors sit on two or more company boards – have made great use of measures of centrality. They have to documented the existence of "hubs" within large spheres of adjacent companies. Such hubs – typically banks or other financial institutions – are the focal centers for business decision-making and for the allocation of credit at the level of the intercorporate network as a whole [50].

Users of graph theory have often distinguished between measures of centrality based on overall "closeness" and an idea of centrality based on "betweenness" [22]. A betweenness measure of centrality is one that is based on a calculation of the probability that any particular point lies on a path between any two other points. This latter measure can be used to operationalize the idea of the broker or intermediary, a person who may not be central in the network as a whole, but is an important point of contact between parts of the network.

A further extension of the ideas of distance and point centrality has been the attempt to measure the overall centralization of the graph as a whole. A measure of centrality grasps the extent to which a whole graph is organized around a small number of central points that constitute its structural center.

The early work of Warner and his associates was concerned with the formation of cliques within networks. Their use of this idea of the clique was, however, relatively loose and essentially commonsensical. Developments in graph theory have since made possible the construction of a number of structural concepts for sub-graphs that are both precise and distinct from each other. The most straightforward of these sub-graph concepts is that of the "component". A component is a maximally connected sub-graph, the set of all those points that are connected, directly or indirectly, by a continuous sequence of lines. A graph will typically consist of a number of such components, together with a variable number of completely isolated points. Component boundaries may, therefore, indicate the boundaries to the flow of communication, influence, or resources within the corresponding social network. A mapping of the size distribution of components will show the extent to which these flows are fragmented. It is also possible to map the internal structures of the various components by taking account of the sign, value, and direction of the constituent lines in order to uncover any blockages or distortions in the flow of communication, influence, or resources. Using a measure of value, for example, the nesting of intensely connected components within more weakly connected components can be disclosed. The "slicing" techniques available in many software packages allow contour maps of component structure to be built, the resulting map showing the peaks of intensity and the troughs or valleys that separate them.

Component structure can also be studied through an investigation of the "cycles" that it contains. A cycle is a path that returns to its own starting point, and a cyclic component consists of a set of intersecting cycles. A component may consist of a number of such cyclic components that are connected only through non-cyclic "bridges".

Cliques, as defined in graph theory, are sub-sets of points in which every possible pair of points is connected at a specified distance. Thus, a 1-clique is a sub-set of point in which all points are directly connected to each other. A 2-clique, on the other hand, is a sub-set of points that are connected through common neighbors. There is, therefore, a whole family of clique concepts based on varying distance measures. These are generically referred to as n-cliques and any large component in a graph will typically consist of many overlapping *n*-cliques, as well-connected points will tend to be members of large numbers of cliques. An analysis of cliques must always specify the value of n that will be used. Once more, the choice of a value for

identifying cliques is not a matter that can be determined on mathematical grounds alone. The investigator must always decide on the sociological meaning – if any – that can be given to particular values of n.

This basic idea of the clique has been extended into a variety of related ideas, most notably called clans, clubs, plexes, and circles, each of which differs according to the precise criterion that defines sub-set membership. These various sub-graph and sub-set concepts are particularly important in studies of diffusion processes in social networks as they define the "obstacles" and barriers that prevent the smooth diffusion of ideas or resources. This was the insight that lay behind Granovetter's [26] recognition of the strength of weak ties: it is the possession of "weak" connections that bridge a person into sub-graphs that would otherwise be unable to transfer information to the person.

Most uses of graph theory in social network analysis have employed one-mode data. That is to say, they have transformed the initial bipartite or two-mode data into a simpler form. An initial incidence matrix showing, for example, the memberships of people (as columns) in particular organizations (as rows) is transformed into a column-by-column adjacency matrix of people and a row-by-row adjacency matrix of organizations. Thereafter, the analyses of the people and the organizations proceeds separately (see ► Social Network Analysis, Two–Mode Concepts in). Recently, however, attempts have been made to develop graph theory approaches to the analysis of two-mode data, treating the bipartite data as if it were one-mode data. The adjacency matrices are, in effect, seen as sub-matrices of a larger adjacency matrix in which people and organizations, for example, comprise *both* the columns and the rows. Once the larger matrix is generated, many of the conventional graph theory measures and visualization techniques can be applied to it. Not all such measures, however, are sensibly or usefully seen in two-mode terms.

Many of the more complex graph theoretical measures for whole graphs can be difficult, if not impossible, to use in large-scale networks where there is incomplete data. This is typically the case where a sampling methodology has been followed (see ► Social Network Analysis, Estimation and Sampling in).

Some network parameters can be estimated with sample data. If information on the neighborhood or neighborhood density of individuals are collected from a sample, for example, it is straightforward to compare estimates of these measures for the whole population. It would also be possible to estimate the density of the network as a whole from such sample data, as density is an adjacency-based measure. So long as the normal questions of sample size and representativeness are considered, such measures pose few problems. Matters are more complex, however, with many global network measures, where the structure of relations is more likely to be lost in the process of sampling. It is, for example, impossible to measure the number of separate components in a network using sample data. Some ingenious methods of estimation have, however, been suggested as being useful for such measures, so long as appropriate sampling designs are used.

An alternative to sampling is the adoption of explicit methods of data reduction that allow large-scale networks to be analyzed more easily. As Vladimir Batagelj shows in ► Social Network Analysis, Large-Scale, it is especially useful to have methods of visualization for such networks. The *Pajek* program [16] was designed specifically for this purpose. Large networks are those with thousands, or perhaps millions, of points and the computing time required for such networks has been a barrier to studying them. Batagelj uses recursive decomposition methods to divide a large network into several smaller ones. This can be achieved by the clustering of points or lines, compressing them into compound points or lines. Varying forms of reduction can be achieved by adopting different criteria for reduction: for example, collapsing all members of a component onto a point, or collapsing all members of a clique into a single point. Such reduced graphs can then be analyzed with the familiar techniques of graph theory. He also introduces a number of concepts specific to reduced networks. The concept of an island, for example, refers to peaks at specified cut-off levels on a measure of the properties of the compared points or lines.

Diffusion Processes

A key issue in sociological studies of influence has been the ways in which innovations and other social changes are diffused through the social relations that people have established. Similar processes of diffusion were investigated in geography and anthropology, and a number of investigators began to explore mathematical models of diffusion, that had already been begun to be applied in epidemiological studies of the spread of disease through a population [48,49] (see also [14]). Everett Rogers documented the existence of an S-curve describing the pace of innovation: a slow initial uptake was followed by more rapid adoption, and then a tailing-off as diffusion produced a relative saturation of the population and correspondingly fewer people available to take up the changes. The key parameters were the speed of the initial take-off and the speed of subsequent adoption.

Contemporary approaches to diffusion [57] begin from measures of spatial autocorrelation to investigate the speed of the spread of innovations (see ► Social Networks, Diffusion Processes in). Spatial autocorrelation is the correlation of a variable with itself measured spatially. That is, its dependence on neighboring points. Such models measure the extent to which changes spread through contiguous regions of space. Within networks, this idea is generalized to analyze contiguity in social network terms. It is in this way that a "contagion" model of social influence can be operationalized. Network parameters can be weighted in various ways in a model, taking, for example, direct links as more significant than indirect links, and predictive conclusions about the temporal and spatial spread of changes can be drawn. As Tom Valente shows, event history analysis has been used to allow more sophisticated temporal measures of diffusion, and this has become a key area of innovation within the area of diffusion studies. Such work is also now developing closely with the agent-based computational models discussed below.

Algebraic Models and Blockmodeling

The early work on network structures that was carried out by Harrison White and his associates [36,72] made use of algebraic ideas to model graphs of social relations. Their argument was that algebra provided a means for analyzing the properties of sets of points. This work laid the basis for what came to be known as blockmodeling. This is a way of exploring the relations among structurally defined positions in social networks, the positions being understood as "blocks" or sets of points with certain common relational attributes that make them, in crucial respects, structurally equivalent to each other [5,73]. This form of positional analysis has subsequently been elaborated and enlarged in a variety of ways, resulting in the production of a whole family of models for network structure.

An algebraic representation of a social network involves identifying a set of points that is defined by the specific relation that unites its members [43] (see ► Social Networks, Algebraic Models for). Thus, a network is seen as consisting of a set of ordered pairs. This approach can easily be generalized to explore situations in which more than one relation is involved and units can be understood as connected into multiple networks, each of which is defined by a particular type of relationship. From this point of view, a multiple network is a set of sets, and each set is defined by a "primitive relation" or generator that contributes to determining the overall network structure. Therefore, the individuals or groups that comprise the points in a social network are treated as standing in compound relations to each other. Paths in the network may be composed from lines that represent a variety of different relations. The basic step in identifying the structure of such a complex network is the multiplication of the initial binary relations to convert them into binary compound relations. The structure of relationships among all possible paths in a multiple network is termed a partially ordered semigroup and this comprises the underlying relational structure of the network. As with graph theory, this algebraic method allows such global properties of social networks to be linked to the local, egocentric properties of specific units within the network. Thus, the immediate environment of any particular member of the network can be seen as a "local role algebra" that represents the role set of the focal individual or group. Such an algebra defines the vectors that slice through the whole network and so summarize an ordering of the constituent relation vectors in which the individual point is involved.

The basic elements of blockmodeling are based on the ideas first set out by the anthropologist Nadel, who held that social network analysis must take account of social positions and their associated roles rather than simply with individuals [42] (see ► Positional Analysis and Blockmodeling). Graph theory is especially well-suited to the identification of individual-level phenomena and the characteristics of the social groups formed by individuals. It is much less well-suited to the identification of structural positions. For example, a large number of individuals may each occupy the socially determined position of "father", but a graph theoretical analysis of each individual male adult's connections to children will not easily identify this shared social role. This is because graph theory searches for clusters of connections based on distance measures. Blockmodeling, on the other hand, starts out from a recognition that individuals that occupy the same social position are "equivalent" or "substitutable" in some way. Thus, each father relates in similar ways to particular children, but the fathers do not all relate to the *same* children. Positions and their associated roles are enduring structural regularities among sets of points and differ markedly from cliques, components and other such subgraphs.

The typical strategy for constructing a blockmodel is to apply an algorithm to a bipartite, two-mode, matrix of individual-level data in order to reduce it to an "image graph" that comprises blocks of equivalent points. Though typically employing two-mode data, blockmodels can be produced for one-mode data and, with greater difficulty, for three-mode data. Simple binary data have been most widely used in blockmodeling, with valued data being han-

dled by slicing procedures that partition the matrix into 1s and 0s on the basis of a chosen threshold of significance. Attempts are being made, however, to extend blockmodeling directly to valued data. Approaches to blockmodeling vary from each other in terms of the particular criteria that are used in the algorithm. The earliest approach – the CONCOR method of White and his associates – used iterated correlation measures to produce the image graph. In this method, repeated correlations computed on the rows and columns of the initial matrix converge towards a rearranged matrix that shows blocks of 0s and blocks of 1s. The blocks of 0s – "zero-blocks" – are the structural holes in the network that divide the various social positions (the blocks of 1s) from each other. Later work associated with Ronald Burt [9,10], has used simpler clustering approaches that are based on Euclidean distance measures calculated from path distances among pairs of points. Burt's method also allows the construction of an image graph for the identification of blocks of equivalent positions, and Burt terms this a social topology.

Applications of these ideas to intercorporate business connections has shown that blocks in networks of shareholdings and interlocking directorships can be seen as comprising structurally similar agents within the economy. The blocks uncovered in such an analysis are the dominant investors and subordinate enterprises that stand in similar relations to each other but may have few mutual or reciprocal connections among themselves. They are social positions but do not necessarily constitute social groups. The key text on blockmodeling has been produced by Doreian and colleagues [19].

The construction of an algebraic representation of a social network comprises a depiction of the "deep structure" of the network in the sense that was intended by Lévi-Strauss and many structuralist writers. The method allows a formal construction and representation of the deep structures that have, more typically, been identified through more purely interpretative means. As in these structuralist approaches, a key aim of algebraic network analysis is the identification of isomorphic subgroups and local role algebras across different empirical networks that can be compared for commonalities of structure. This comparative approach allows the formulation of statements concerning generic processes in social systems, modeled as lattices. Hierarchical structures, for example, may be common to a number of different empirical domains and may exhibit similar formal properties, regardless of their specific empirical content. Similarly, empirically variant role sets may exhibit a similarity of underlying structure, such as patterns of conflict or cooperation with different categories of others.

This approach does not, however, reduce the variation found in actual social networks to invariant mathematical properties. There can be no derivation of a network *theory* directly from a mathematical theory. As Pattison (see p. 135 in [43]) argues, algebraic approaches have sacrificed mathematical power in order to reflect more adequately the social relationships in the particular field under investigation. Algebraic methods have tended to be used – as shown in the discussion of blockmodeling – as methods of data reduction, as means for decomposing or fragmenting networks into simpler structural components. The construction of empirically relevant theories from these components is far more difficult and has, so far, rarely been pursued.

Scaling models and Visualization

Both graph theory and algebraic approaches offer rudimentary techniques for the visualization of social networks (see ► Social Network Visualization, Methods of). Graph theory, for example, is based around the idea of the sociogram of connected points. For any but the smallest networks, however, sociograms become a confused jumble of cross-cutting lines, and a rearrangement of points aimed at reducing the amount of overlap can actually destroy any visual representation of the structure of the network by distorting the distances between and the relative locations of the points. A blockmodeling approach to large networks reduces the amount of data to be handled and allows simple image graphs of the relations among blocks to be drawn, but this is limited to positional analysis and loses sight of the details that are present in the relations among the individual points. In both forms of analysis, therefore, social network analysts have tended to rely on the purely mathematical manipulation of large networks and many have almost abandoned the attempt to visually display whole complex social networks. The desire to recapture the simple visual impact of the sociogram, however, has motivated a number of researchers to investigate ways of drawing network diagrams that retain the spatial sense inherent in relational data. Unsurprisingly, perhaps, the techniques discovered have followed the methods used in cartography to represent physical landscapes on the page of a map or atlas.

The earliest and most influential visualization techniques drew on multidimensional scaling (MDS), as this was seen as a method for mapping data that retained its fundamental metrics of space and distance. Essentially similar techniques are those of principal component analysis or factor analysis. Instead of calculating distance simply by the number of lines connecting two points – "path

distance" – MDS approximates to the measures of physical distance used in conventional spatial thinking. MDS techniques convert raw data on social networks into Euclidean distances. In this conversion, points are regarded as "close" to each other by virtue of their position in the overall configuration of relations. This is typically measured on the basis of a proximity measure such as the frequency of contact between two points or the intensity of a relationship between them, and these are usually given a strictly metric form by computing correlation coefficients. For example, two points with identical patterns of connection are perfectly correlated with each other, and decreasing similarity is reflected in declining closeness. Using such similarity (or dissimilarity) data, MDS procedures search for the best fit within a Euclidean space of specified dimensions. Once a solution is obtained it can be displayed on the page and rotated into an informative orientation that discloses the meaning of the dimensions that have been identified in the analysis. Non-metric MDS, using the rank order of distances rather than imputed actual distances, is particularly appropriate for much social network data and operates in very much the same way. Useful implementations of this approach are now available in the major software packages, and novel techniques such as multiple correspondence analysis are also beginning to become available.

The use of MDS and similar scaling or dimensional models provides a visual image of a network, but it also provides measures of distance, direction, and location that are not available within graph theory or algebraic approaches. This can allow a more rigorous embedding of social networks in social space and can permit intriguing hypotheses to be formed. For example, cliques and components, centrality and density, can all be mapped onto an MDS solution, and the relations among cliques or central points appear strikingly and can be measured determinately. The procedures made available within the *Pajek* program [18] have been designed specifically to allow such possibilities for large-scale networks. Graph theory and algebraic methods can be used together with the spring embedding of data through a variant of MDS to produce powerful network representations.

Particular problems of visualization occur when research is concerned with two-mode data rather than simple one-mode data. The most easily interpretable results are those that display a bipartite graph by using different symbols or colors for the two sets of data. Such graphs are a logical extension of graph theory approaches to two-mode data. Some interesting work is now being undertaken, however, using Galois lattices. Abandoning graph theoretical and spatial representations, the lattices aim to disclose the pattern of connections and their hierarchical order of "containment" from core to periphery.

One particularly striking advance in visualization techniques has been the use of models that allow the tracking of change over time. These have typically involved the adaptation of animation methods. Further advances in this area might be expected as some of the techniques discussed below begin to offer improved ways of handling temporal data on social networks.

Statistical Models for Hypothesis Testing

Much work on social networks has remained at the level of description and has failed to move towards explanatory concerns. Only rarely have researchers gone beyond a description to formulate hypotheses and test theories aimed at explaining the observed structures. One reason for this absence of theoretical work and hypothesis testing has been the weakness of the available statistical techniques for undertaking this work. While attempts have sometimes been made to use basic statistical measures of probability and significance to test hypotheses about network structure, these involve a number of difficulties. Standard statistical procedures such as significance tests, regression, and the analysis of variance all assume the independence of observations. The relational data studied by social network analysts, however, cannot be seen in this way. Relational data are, by definition, non-independent observations. Thus, novel statistical techniques have been required if methods of statistical inference are to be used. The most important work in recent years has been the generalization of Markov graphs to a larger family of models that has been undertaken by Stanley Wasserman and his colleagues [44,46,65], building on earlier work by Frank and Strauss [21]. Their $\boldsymbol{p}^*$ models – now termed exponential random graph models (ERG models or ERGM) – are designed specifically to allow the easier formulation and testing of theories. An ERGM defines a probability distribution on the set of all networks that can be constructed on a given set of points in terms of a particular parameter vector. ERGMs comprise a family of statistical models that promise a great advance in the understanding of social networks and have the potential to connect with the "small world" methods discussed below. The models are discussed in ► Social Networks, Exponential Random Graph ($\boldsymbol{p}^*$) Models for.

For any set of points it is possible to generate through simulation a set of all the possible graphs connecting them. These randomly generated graphs vary along the full range from completely unconnected to completely connected. A very large number of such graphs can exist along that

range, as the number of possible graphs varies with the size of the network at an exponential rate. Thus, the number of possible configurations for a three-point graph is 64, that of a four-point graph is 1024, and the number for a five point graph is in excess of a million. An actually observed network can be treated as one realization from this set of logically possible graphs. If its probability of occurring by chance is low, then its actual occurrence may be regarded as statistically significant.

To assess this statistical significance, it is necessary to construct a probability distribution of the random graphs. Each random graph has a specific probability of occurring, depending on its particular structural properties. The aim in constructing these distributions is to see whether a particular structural property, such as a particular level of density, centralization, or clustering, is more or less likely to occur by chance than any other. These probabilities are estimated through log linear analysis from the statistical dependencies among the members of the network. The log odds ratios of the conditional probabilities of the relational ties associated with each element constructed from this dependence graph are used to produce an approximate likelihood function that can be used with Monte Carlo estimation techniques to produce maximum likelihood estimates. In this process the successive approximations constitute chains of graphs – a Markov chain – that converge towards a stationary distribution and a stable estimate of network parameters. Thus, ERGM provides a means for testing the probability that the observed network will occur with precisely the structural characteristics that it has on the basis of chance alone. This enables the researcher to highlight the statistically significant – non-chance – results apparent in the data.

Agent-Based Computational Models and Temporal Processes

As well as its descriptive focus, much social network analysis has also concentrated on the static features of social networks. This has also begun to change in recent years as more attention has been given to the dynamic processes involved in changes over time. A key advance in this direction has come from the use of models that depict the ways in which the behavior of individual agents results in global transformations of network structure. In so-called agent-based computational models, agents (whether individuals or groups) are seen as rule-following entities whose decisions to act in one way or another are consequential for the overall network by virtue of their concatenation with the action consequences of others. Therefore, a knowledge of the rules under which agents act can be used to predict broad patterns of change in network structure.

Agent-based models originated in simulation studies, where a simple and determinate set of rules can be applied to generate specific network structures (see also pp. 99ff in [25]; [40]). Applications of this idea to human social networks have tended to rely on assumptions drawn from rational choice theory, according to which individual agents are assumed to act on certain simple principles of utility maximization. However, the more complex models allow for the handling of different categories of agent, the members of each category following a different set of rules, and for a relaxation of the assumption of pure rationality by modeling rules of non-rational or ritualistic behavior. Such models are stochastic – specifying the probability of certain courses of action but not implying a complete determinism.

A particularly powerful example of such an approach is that of Tom Snijders [52,53,54] (see ► Network Analysis, Longitudinal Methods of). Change in network structure is regarded as a probabilistic process, describable by an objective function that measures the attractiveness to the agents of changes in their connectivity. The model sees an incremental adjustment of individual action to the changing network structure, resulting in a continuous – but often non-linear – process of network evolution. It is assumed that agents control the outgoing relations to others within the constraints on their opportunities that are set by the current network structure. However, they act "myopically" in relation to the immediate environment and the immediate consequences of their actions, having no conception of the wider consequences of their choices. At each point in time, the outcome for the overall network is a result of the intersection of the courses of action pursued by these myopic agents and so it will rarely correspond to the intentions of any one actor. At the same time, each actor has only a partial awareness of the changes that have resulted from their actions and so their next acts must be regarded as equally myopic and as likely, once again, to have unintended consequences at the network level. As a result of the continual iteration of actions, there is a co-evolution of network and behavior. Complex patterns of change are likely to occur. Snijders shows that small, incremental changes can accumulate to a point of "crisis" [55], a tipping point at which a radical, non-linear transformation of network structure occurs.

Snijders' longitudinal approach is especially powerful as it explicitly takes account of global network parameters such as density, reciprocity, distance, and balance. It can also be reprogrammed to recognize a wide range of different rule systems and decision procedures. Snijders

also suggests that the approach can handle a greater degree of far-sightedness in agents, though with a corresponding decline in the determinateness of the analyses that can be drawn. Current work is making important connections with the early work of Wasserman [63] and the subsequent development of ERGM procedures. The approach is implemented in the *SIENA* program and has been used in a number of simulation studies.

Small World Models and Network Dynamics

The most recent area of development in social network analysis is the investigation of small world networks inspired by a number of physicists interested in network dynamics (explored in various entries in other sections of this *Encyclopedia*). Although this has, as yet, led to relatively few empirical investigations, there is growing interest in these applications. Many advocates of the approach have failed to discover or to appreciate the amount of work already undertaken in social network analysis and have assumed a very truncated history for the approach. There are, however, some signs of a closer appreciation of social science work, and the approaches are beginning to have an impact in the mainstream of social network analysis.

The "small world" hypothesis suggests that any two people chosen at random are likely to be connected to each other by a path of relatively short length. The hypothesis was formulated by Ithiel de Sola Pool and Manfred Kochen in a working paper of 1958 that was not published for another 20 years [17]. Their ideas achieved a wide currency and in 1967 Stanley Milgram at Harvard undertook some experimental studies to test the hypothesis [38]. He gave experimental subjects the names of people initially unknown to them and asked them to get a message to them using only known persons as intermediaries. That is, the messages were to pass through friend of friends of friends until they reached their destinations. Milgram discovered that subjects and targets were typically separated by paths of distance 6, no matter what the physical distances separating them. This finding embodied the common exclamation "it's a small world" uttered when people discover a mutual connection. Milgram hypothesized that the density of links of acquaintanceship is such that a considerable amount of redundancy is built into social networks: it is level of density in a network that explains its small world characteristics.

In the late 1990s, Watts and Strogatz [68] undertook some empirical studies in the mathematics of random networks. The structural properties of random networks, produced through the random linking of one point to another, had first been studied by Paul Erdös and Alfréd Rényi [20] and had led to a number of interesting conclusions. Watts and Strogatz discovered that a particular subset of random graphs corresponded closely to those actually found in natural and social phenomena, and they demonstrated that these had the small world characteristics identified by Milgram; though they were skeptical about the precise finding of "six degrees of separation". They showed, however, that the high levels of clique and sub-graph formation that were basic to the redundancy found in real networks gave them a number of characteristic features. Crucial to the structure of these networks was the existence of well-connected points – which they termed "hubs" – that make the density and, therefore, the small world properties, possible. It is the existence of hubs that explains why the density does not need to be extremely high for small world conditions to apply. On the contrary, there is a middle range of density in which efficient communication is possible.

The physicist Albértó-Laszlo Barabási, searching for a new topic to investigate, began to look at the findings of Watts and Strogatz. Barabási's central idea was that of the "power law" or "scale free" distribution, according to which small world networks have a frequency distribution of point degrees showing a small number of well-connected points and a large number of less well-connected points. These well-connected points are the hubs that generate the small world properties. He concluded that these findings provided the basis for a completely new approach to networks [1]. As has already been noted, his claims to novelty and to the revolutionary character of the work seriously underestimated the amount of prior work undertaken on social networks and their properties. Nevertheless, he did formulate some intriguing ideas and the work subsequently taken up by Watts does offer the prospect of genuine advances in social network analysis.

Watts holds that a small world graph is one in which there are a large number of "short cuts". These are lines that connect points that would otherwise be quite distant from each other. Neither a completely connected graph nor a sparse graph have short cuts, and so small world properties occur in an intermediate range of graphs. They are locally clustered but globally sparse. It is for these graphs, he holds, that precise mathematical conclusions about structure and structural development can be drawn. He holds, for example, that relatively small changes in network connectivity, such as the addition of a small number of shortcuts, can dramatically change its properties through "phase transitions" at critical threshold points. There are considerable possibilities inherent in using these ideas alongside those from agent-based computational studies already discussed. The common thread

is a concern for the investigation of longitudinal processes in which micro-level incremental changes result in non-linear macro-level transformations of structure.

Bibliography

1. Barabási A-L (2002) Linked: The new science of networks. Perseus, Cambridge
2. Barnes JA (1954) Class and committee in a Norwegian island parish. Hum Relat 7:39–58
3. Bavelas A (1948) A mathematical model for group structures. Appl Anthr 7:16–30
4. Berkowitz SD (1982) An introduction to structural analysis. Butterworths, Toronto
5. Boorman SA, White HC (1976) Social structure from multiple networks, vol II. Am J Sociol 81:1384–1446
6. Bott E (1955) Urban families: Conjugal roles and social networks. Hum Relat 8:253–292
7. Bott E (1956) Urban families: The norms of conjugal roles. Hum Relat 9:325–342
8. Buchanan M (2002) Small world: Uncovering nature's hidden networks. Weidenfeld and Nicolson, London
9. Burt RS (1980) Models of network structure. Ann Rev Sociol 6:79–141
10. Burt RS (1982) Towards a structural theory of action. Academic Press, New York
11. Burt RS (1992) Structural holes. Cambridge University Press, New York
12. Carrington PJ, Scott J, Wasserman S (2005) Models and methods in social network analysis. Cambridge University Press, Cambridge
13. Cartwright D, Zander A (1953) Group dynamics. Tavistock, London
14. Coleman JS, Katz E, Menzel H (1966) Medical innovation: A diffusion study. Bobbs-Merrill, New York
15. Davis AB, Gardner BB, Gardner MR (1941) Deep south. University of Chicago Press, Chicago
16. De Nooy W, Mrvar A, Batagelj V (2005) Exploratory social network analysis with Pajek. Cambridge University Press, New York
17. De Sola Pool I, Kochen M (1978) Contacts and influence. Soc Netw 1(1):5–51
18. DeNooy W, Mrvar A, Batagelj V (2005) Exploratory social network analysis with Pajek. Cambridge University Press, New York
19. Doreian P, Batagelj V, Ferligoj A (2005) Generalized blockmodelling. Cambridge University Press, New York
20. Erdös P, Rényi A (1959) On the evolution of random graphs. Publ Math Inst Hung Acad Sci 5:17–61
21. Frank O, Strauss D (1986) Markov Graphs. J Am Stat Assoc 81:832–842
22. Freeman LC (1979) Centrality in social networks: Conceptual clarification. Soc Netw 1:215–239
23. Freeman LC (2004) The development of social network analysis: A study in the sociology of science. Empirical Press, Vancouver
24. Freeman LC, White DR, Romney AK (1989) Research methods in social network analysis. Transaction Books, New Brunswick
25. Gilbert N, Troizch KG (1999) Simulation for the social scientist. Open University Press, Buckingham
26. Granovetter M (1973) The strength of weak ties. Am J Sociol 78:1360–1380
27. Granovetter M (1974) Getting a job. Harvard University Press, Cambridge
28. Harary F, Norman RZ (1953) Graph theory as a mathematical model in social science. Institute for Social Research, Ann Arbor
29. Helmers HM (1975) Graven naar macht. Van Gennep, Amsterdam
30. Homans G (1950) The human group. Routledge and Kegan Paul, London
31. Homans G (1961) Social behaviour: Its elementary forms. Routledge and Kegan Paul, London
32. Knoke D, Kuklinski JH (1982) Network analysis. Sage Publications, Beverley Hills
33. König D (1936) Theorie der endlichen und unendlichen Graphen. Chelsea, New York
34. Laumann EO (1966) Prestige and association in an urban community. Bobbs-Merrill, Indianapolis
35. Levine JH (1972) The sphere of influence. Am Sociol Rev 37:14–27
36. Lorrain F, White HC (1971) Structural equivalence of individuals in social networks. J Math Sociol 1:49–80
37. Marsden PV, Lin N (1982) Social structure and network analysis. Sage, Beverley Hills
38. Milgram S (1967) The small world problem. Psychol Today 2:60–67
39. Mitchell JC (1969) Social networks in urban situations. Manchester University Press, Manchester
40. Monge PR, Contractor NS (2003) Theories of communication networks. Oxford University Press, Oxford
41. Moreno JL (1934) Who shall survive? Beacon Press, New York
42. Nadel SF (1957) The theory of social structure. Free Press, Glencoe
43. Pattison P (1993) Algebraic models for social networks. Cambridge University Press, Cambridge
44. Pattison P, Wasserman S (1999) Logit models and logistic regressions for social networks: II Multivariate relations. Br J Math Stat Psychol 52:169–193
45. Radcliffe-Brown AR (1940) On social structure. In: Radcliffe-Brown AR (ed) Structure and function in primitive society. Cohen and West, London
46. Robins GL, Pattison P, Wasserman S (1999) Logit models and logistic regressions for social networks: III Valued relations. Psychometrika 64:371–394
47. Roethlisberger FJ, Dickson WJ (1939) Management and the worker. Harvard University Press, Cambridge
48. Rogers E (1962) Diffusion of innovations, 5th edn. Free Press, New York
49. Ryan R, Gross N (1943) The diffusion of hybrid seed corn in two Iowa communities. Rural Sociol 8:15–24
50. Scott J (1991) Networks of corporate power: A comparative assessment. Ann Rev Sociol 17:181–203
51. Scott J (2000) Social network analysis, 2nd edn. Sage, London
52. Snijders TAB (2001) The statistical evaluation of social network dynamics. In: Sobel ME, Becker MP (eds) Sociological methodology. Basil Blackwell, Oxford
53. Snijders TAB (2005) Models for longitudinal network data. In: Carrington PJ, Scott J, Wasserman S (eds) Models and methods in social network analysis. Cambridge University Press, Cambridge

54. Snijders TAB, Van Duijn MAJ (1997) Simulation for statistical inference in dynamic network models. In: Conte R, Hegelmann R, Terna P (eds) Simulating social phenomena. Springer, Berlin
55. Thom R (1972) Structural stability and morphogenesis. Benjamin, London
56. Travers J, Milgram S (1969) An experimental study of the small world problem. Sociometry 32:425–443
57. Valente TW (1995) Network models of the diffusion of innovations. Hampton Press, Cresskill
58. Warner WL (1949) Democracy in Jonesville. Harper and Brothers, New York
59. Warner WL, Low JO (1947) The social system of a modern factory. Yale University Press, New Haven
60. Warner WL, Lunt PS (1941) The social life of a modern community. Yale University Press, New Haven
61. Warner WL, Lunt PS (1942) The status system of a modern community. Yale University Press, New Haven
62. Warner WL, Srole L (1945) The social systems of american ethnic groups. Yale University Press, New Haven
63. Wasserman S (1980) Analyzing social networks as stochastic processes. J Am Stat Assoc 75:280–294
64. Wasserman S, Faust K (1994) Social network analysis: Methods and applications. Cambridge University Press, New York
65. Wasserman S, Pattison P (1996) Logit models and logistic regressions for social networks, vol I. An introduction to Markov random graphs and $\boldsymbol{p}^*$. Psychometrika 60:401–426
66. Watts D (1999) Small worlds: The dynamics of networks between order and randomness. Princeton University Press, Princeton
67. Watts D (2003) Six degrees. The science of a connected age. Norton WW, New York
68. Watts DJ, Strogatz SH (1998) Collective dynamics of small-world networks. Nature 393:440–442
69. Wellman B (1979) The community question: The intimate networks of East Yorkers. Am J Sociol 84:1201–1231
70. Wellman B (1988) Networks as personal communities. In: Wellman B, Berkowitz SD (eds) Social structures. Cambridge University Press, Cambridge
71. Wellman B, Berkowitz S (1988) Social structures. Cambridge University Press, New York
72. White HC (1963) An anatomy of kinship. Prentice-Hall, Englewood Cliffs
73. White HC, Boorman SA, Breiger RL (1976) Social structure from multiple networks, I. Am J Sociol 81:1384–1446

Social Network Analysis, Two-Mode Concepts in

Stephen P. Borgatti
Gatton College of Business and Economics, University of Kentucky, Lexington, USA

Article Outline

Glossary
Definition of the Subject
Introduction
Basic Concepts
Two-Mode Data in Social Network Analysis
Unimodal Approaches to Two-Mode Data
Bimodal Approaches to Two-Mode Data
Conclusion
Future Directions
Bibliography

Glossary

2-Mode matrix A (2-dimensional) matrix is said to be 2-mode if the rows and columns index different sets of entities (e. g., the rows might correspond to persons while the columns correspond to organizations). In contrast, a matrix is 1-mode if the rows and columns refer to the same set of entities, such as a city-by-city matrix of distances.

Blockmodel A blockmodel is a partitioning of the cells of a matrix into blocks that is induced by the partitioning the rows and columns into classes and sorting the matrix such that rows (and columns) that belong to the same class are next to each other. More specifically, two matrix cells x_{ij} and x_{mp} are in the same block if $\text{class}(i) = \text{class}(m)$ and $\text{class}(j) = \text{class}(p)$.

Centrality A family of concepts for characterizing the structural importance of a node's position in a network.

Graph cohesion A family of concepts characterizing the extent of connectedness of a graph, such as density (the proportion of pairs of nodes that have ties with each other), or average path distance.

Multidimensional scaling (MDS) A method of locating points in space such that Euclidean distances between the points correspond to a matrix of input similarities/distances among objects. Used to provide visual representations of 1-mode matrices such as correlation matrices or perceptual distances among objects.

Regular equivalence The definition of regular equivalence is recursive. If two nodes are regularly equivalent, then they are connected to regularly equivalent nodes. Regular equivalence is used to identify nodes that are playing the same structural role, even if they are not connected to each other.

Social network (or, in graph theory, a graph) A collection of nodes (also referred to as vertices or actors) together with a set of ties (also known as edges or links) that connect pairs of nodes. Typically used to represent social relations such as who is friends with whom, or who is the supervisor of whom.

Structural equivalence At an intuitive level, a pair of nodes is said to be structurally equivalent to the ex-

Names of Participants of Group I	Code Numbers and Dates of Social Events Reported in *Old City Herald*													
	(1) 6/27	(2) 3/2	(3) 4/12	(4) 9/26	(5) 2/25	(6) 5/19	(7) 3/15	(8) 9/16	(9) 4/8	(10) 6/10	(11) 2/23	(12) 4/7	(13) 11/21	(14) 8/3
1. Mrs. Evelyn Jefferson	×	×	×	×	×	×		×	×					
2. Miss Laura Mandeville	×	×	×		×	×	×	×						
3. Miss Theresa Anderson		×	×	×	×	×	×	×	×					
4. Miss Brenda Rogers	×		×	×	×	×	×	×						
5. Miss Charlotte McDowd			×	×	×		×							
6. Miss Frances Anderson			×		×	×		×						
7. Miss Eleanor Nye					×	×	×	×						
8. Miss Pearl Oglethorpe						×		×	×					
9. Miss Ruth DeSand					×		×	×	×					
10. Miss Verne Sanderson							×	×	×			×		
11. Miss Myra Liddell								×	×	×		×		
12. Miss Katherine Rogers								×	×	×		×	×	×
13. Mrs. Sylvia Avondale							×	×	×	×		×	×	×
14. Mrs. Nora Fayette						×	×		×	×	×	×	×	×
15. Mrs. Helen Lloyd							×	×		×	×	×		
16. Mrs. Dorothy Murchison								×	×					
17. Mrs. Olivia Carleton									×		×			
18. Mrs. Flora Price									×		×			

Social Network Analysis, Two-Mode Concepts in, Figure 1
DGG women-by-events matrix

tent that they occupy identical locations in a network, meaning that they are connected to exactly the same others. Structurally equivalent nodes are identical with respect to all structural properties, such as centrality or subgroup membership.

Definition of the Subject

In social network analysis, 2-mode data refers to data recording ties between two sets of entities. In this context, the term "mode" refers to a class of entities – typically called actors, nodes or vertices – whose members have social ties with other members (in the 1-mode case) or with members of another class (in the 2-mode case). Most social network analysis is concerned with the 1-mode case, as in the analysis of friendship ties among a set of school children or advice-giving relations within an organization. The 2-mode case arises when researchers collect relations between classes of actors, such as persons and organizations, or persons and events. For example, a researcher might collect data on which students in a university belong to which campus organizations, or which employees in an organization participate in which electronic discussion forums. These kinds of data are often referred to as affiliations. Co-memberships in organizations or participation in events are typically thought of as providing opportunities for social relationships among individuals (and also as the consequences of pre-existing relationships). At the same time, ties between organizations through their members are thought to be conduits through which organizations influence each other.

Introduction

Perhaps the best known example of 2-mode network analysis is contained in the study of class and race by Davis, Gardner and Gardner (henceforth DGG) published in the 1941 book *Deep South* [6]. They followed 18 women over a nine-month period, and reported their participation in 14 events, such as a meeting of a social club, a church event, a party, and so on. Their original figure is shown in Fig. 1.

DGG used the data to investigate the extent to which social relations tended to occur within social classes.

Basic Concepts

A typical data matrix has two dimensions or *ways*, corresponding to the rows and columns of the matrix. The number of ways in a matrix X can be thought of as the number of subscripts needed to represent a particular datum, as in x_{ij}. If we stack together a number of similarly sized 2-dimensional matrices, we can think of the result as a 3-dimensional or 3-way matrix.

The *modes* of a matrix correspond to the distinct sets of entities indexed by the ways. In the DGG dataset described above, the rows correspond to women and the columns to a different class of entities, namely events.

Hence, the matrix has two modes in addition to two ways; it is 2-way, 2-mode. In contrast, a persons-by-persons matrix A, in which $a_{ij} = 1$ if person i is friends with person j, is a 2-way, 1-mode matrix, because both ways point to the same set of entities.

In a sense, what constitutes different modes is up to the researcher. If we collect romantic ties among a group of heterosexual people of both genders, we could construct a 2-mode men-by-women matrix X in which $x_{ij} = 1$ if a romantic tie was observed between man i and woman j, and $x_{ij} = 0$ otherwise. Or, one could construct a larger 1-mode person-by-person matrix B also consisting of 1s and 0s in which it just happens that 1s only occur in cells where the row and column correspond to persons of different gender. Use of the men-by-women matrix would imply that same-gender relations were impossible, whereas use of the person-by-person matrix would suggest that same-gender relations were logically possible, even if actually not observed.

Matrices recording relational information such as romantic ties can be represented as mathematical graphs as well. A graph $G(V, E)$ consists of a set of nodes or vertices V together with a set of lines or edges E that connect them. An edge is simply an unordered pair of nodes (u, v). (In directed graphs or *digraphs* we use *ordered* pairs to indicate direction of the tie.) To indicate a tie between two nodes u and v, we simply include the pair (u, v) in the set E. The number of nodes in a graph is denoted by $|V|$ or n.

A bipartite graph is a graph in which we can partition all nodes into two sets, V_1 and V_2, such that all edges include a member of V_1 and a member of V_2. The number of nodes in each vertex set is denoted n_1 and n_2, respectively.

Two-Mode Data in Social Network Analysis

Most social networks are conceived of as relations among a set of nodes, and therefore represented as a 1-mode matrix (typically of 1s and 0s) or a simple graph or digraph. For example, we might collect data on who is friendly with whom within an organization, or who injects drugs with whom in a neighborhood.

However, 2-mode data are common in social network contexts as well. Typical examples include, actor-by-event attendance (as in the DGG data), actor by group membership (such as managers sitting on corporate boards), and actor by trait possession (such as adjective checklist data), and actor by object possession (such as material style of life scales in which inventories are made of household possessions).

In many cases when 2-mode data are collected, the analytical interest is focused on one mode or the other. For example, in the DGG dataset, person-by-event attendances were collected in order to understand social relations among the women, specifically, whether women tended to have social relations primarily within their own social classes. In the interlocking directorate literature, membership of executives on corporate boards is collected mainly in order to understand how corporations are intertwined, and how the structure of this connectivity affects corporate control of society.

However, it can also occur that neither mode dominates our analytical focus and the primary interest is in the correspondence of one mode to the other. For example, a university might ask its faculty which courses they prefer to teach. Here, the objective is typically not to understand how faculty are related to each other through courses, nor how courses are related via faculty, but in the optimal assignments of persons to courses so that courses are staffed and faculty are not complaining.

Unimodal Approaches to Two-Mode Data

One approach to handling 2-mode data in social network analysis is to convert the data to 1-mode data. This is especially appropriate when the analytical interest focuses primarily on just one of the modes. Consider, for example, the case of a person by group membership matrix X in which $x_{ij} = 1$ if person i belongs to group j. Let us assume that the groups are small and everyone in a group knows everyone else. In that case, we could try to infer an acquaintance network by constructing a 1-mode matrix A such that $a_{ij} = 1$ if person i is in at least one group with person j. Better yet, we can construct a valued matrix A such that a_{ij} gives the number of groups that i and j are both members of. In other words,

$$a_{ij} = \sum_k x_{ik} x_{jk} \text{ or } A = \mathrm{XX}' . \tag{1}$$

We might regard a_{ij} as a proxy for the social proximity of i and j, or perhaps as a rough indicator of the potential for information flow between them. In this approach, we analyze each mode of the data separately. Figure 2 shows the values of A for the 2-mode data shown in Fig. 1.

It should be noted that A can be seen as a matrix of profile similarities or correlations among pairs of rows in X. For example, the matrix of Pearson correlations among rows of X is defined as follows:

$$r_{ij} = \frac{\frac{1}{m}\sum_k x_{ik} x_{jk} - u_i u_j}{s_i s_j} , \tag{2}$$

where u_i is the mean of row i and s_i is the standard deviation of row i. It is evident that the correlation r_{ij} is es-

	EVE	LAU	THE	BRE	CHA	FRA	ELE	PEA	RUT	VER	MYR	KAT	SYL	NOR	HEL	DOR	OLI	FLO
EVELYN	8	6	7	6	3	4	3	3	3	2	2	2	2	2	1	2	1	1
LAURA	6	7	6	6	3	4	4	2	3	2	1	1	2	2	2	1	0	0
THERESA	7	6	8	6	4	4	4	3	4	3	2	2	3	3	2	2	1	1
BRENDA	6	6	6	7	4	4	4	2	3	2	1	1	2	2	2	1	0	0
CHARLOTTE	3	3	4	4	4	2	2	0	2	1	0	0	1	1	1	0	0	0
FRANCES	4	4	4	4	2	4	3	2	2	1	1	1	1	1	1	1	0	0
ELEANOR	3	4	4	4	2	3	4	2	3	2	1	1	2	2	2	1	0	0
PEARL	3	2	3	2	0	2	2	3	2	2	2	2	2	2	1	2	1	1
RUTH	3	3	4	3	2	2	3	2	4	3	2	2	3	2	2	2	1	1
VERNE	2	2	3	2	1	1	2	2	3	4	3	3	4	3	3	2	1	1
MYRNA	2	1	2	1	0	1	1	2	2	3	4	4	4	3	3	2	1	1
KATHERINE	2	1	2	1	0	1	1	2	2	3	4	6	6	5	3	2	1	1
SYLVIA	2	2	3	2	1	1	2	2	3	4	4	6	7	6	4	2	1	1
NORA	2	2	3	2	1	1	2	2	2	3	3	5	6	8	4	1	2	2
HELEN	1	2	2	2	1	1	2	1	2	3	3	3	4	4	5	1	1	1
DOROTHY	2	1	2	1	0	1	1	2	2	2	2	2	2	1	1	2	1	1
OLIVIA	1	0	1	0	0	0	0	1	1	1	1	1	1	2	1	1	2	2
FLORA	1	0	1	0	0	0	0	1	1	1	1	1	1	2	1	1	2	2

Social Network Analysis, Two-Mode Concepts in, Figure 2
Women-by-women matrix of overlaps across events

		Row j	
		1	0
Row i	1	a	b
	0	c	d

Social Network Analysis, Two-Mode Concepts in, Figure 3
Contingency table *T*

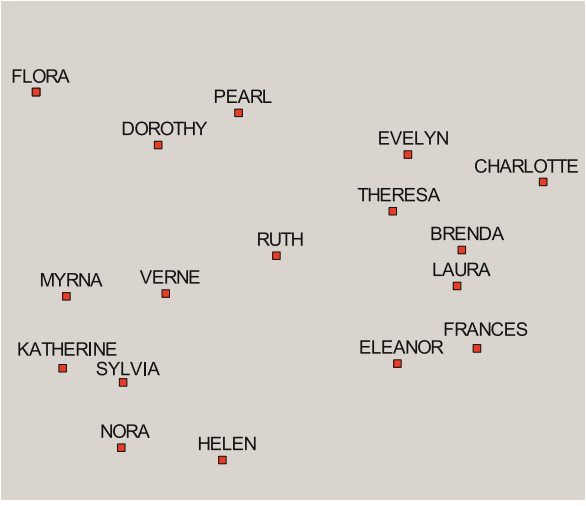
sentially a_{ij} corrected for the number of groups that each belongs to. This kind of correction seems eminently desirable, but of course there are many ways of doing this. For example, consider a cross-tabulation T of row i and row j, such that t_{uv} gives the number of columns k of X for which $x_{ik} = u$ and $x_{jk} = v$, as shown in Fig 3.

In the table, the first row marginal $(a + b)$ gives the number of groups that person i belongs to while the first column marginal gives the number of groups that person j belongs to. Note that a_{ij} of Eq. (1) corresponds to a in Fig. 3. An obvious approach is the Jaccard coefficient, which may be defined as

$$c_{ij} = \frac{a}{a + b + c} . \tag{3}$$

Thus, c_{ij} is essentially the cardinality of the intersection of the groups belonged to by both persons, divided by the cardinality of their union. The Jaccard coefficient is often recommended when the number of columns in X is large and the number of 1s in each row is highly limited.

An alternative specifically designed for 2-mode affiliation data is given by Bonacich [1]. It normalizes a_{ij} as



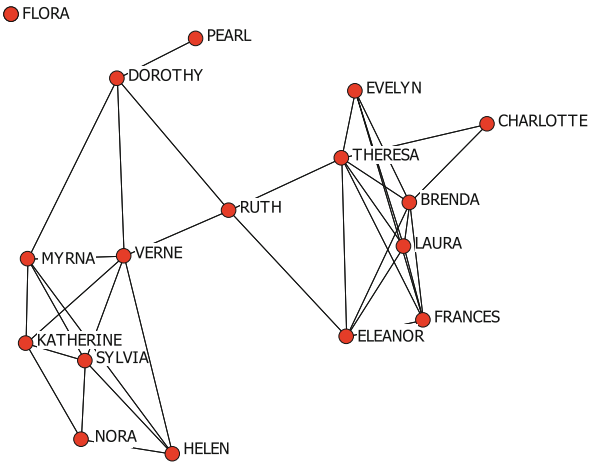
Social Network Analysis, Two-Mode Concepts in, Figure 4
Multidimensional scaling of Jaccard coefficients among rows of DGG matrix

follows:

$$a^*_{ij} = \frac{a_{ij} - \sqrt{adbc}}{ad - bc} , \quad \text{for } ad \neq bc . \tag{4}$$

Unimodal Visualization of Two-Mode Data

Given that a 2-mode matrix has been transformed into a 1-mode matrix by taking similarities among the rows (or columns), one can visualize the network using all the usual techniques for visualization of valued networks. For exam-



Social Network Analysis, Two-Mode Concepts in, Figure 5
Multidimensional scaling of Jaccard similarity coefficients, with lines indicating similarity scores greater than 0.4. Actor Olivia is obscured by Flora

ple, a standard approach is to use metric multidimensional scaling (MDS) on the matrix A, generating a map in which points corresponding to nodes (e. g., persons) appear close to each other to the extent that they share many groups. Figure 4 shows such an MDS map based on Jaccard similarities.

To highlight structure, it is common to overlay lines between pairs of nodes with a similarity score greater than a certain level. Figure 5 shows the result for similarities greater than 0.4. The diagram effectively shows the bridging role of Ruth, who, based on ethnographic evidence, was seen by DGG as a member of both groups of women.

Alternatively, one can use a standard graph layout algorithm (GLA) to draw the graph induced by dichotomizing the Jaccard similarity matrix. For example, define $(u, v) \in E$ if and only if (iff) $c_{ij} > 0.4$. Compared to multidimensional scaling representations, GLAs have the disadvantage that distances between points cannot strictly be interpreted, but this property also means that nodes need not obscure each other. Figure 6 shows the results of applying a spring-embedding [10] GLA to the dichotomized data.

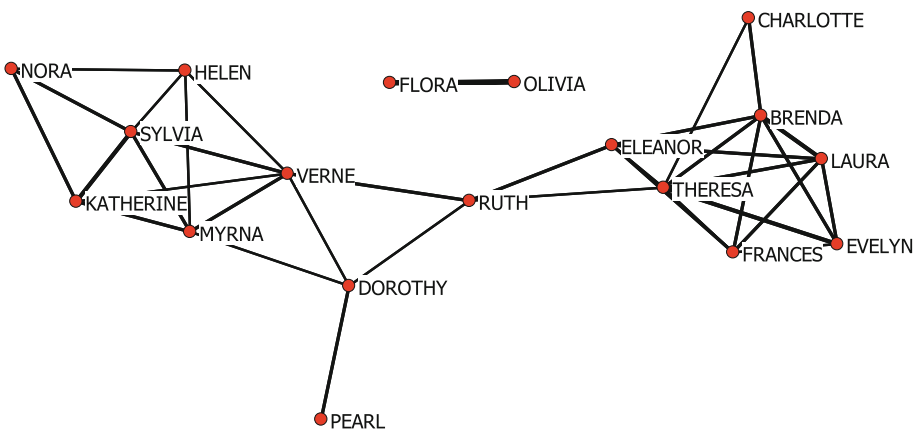
A similar analysis can be carried out on the events rather than the women. Applying Eq. (3) to the columns of the 2-mode matrix in Fig. 1 yields a matrix of Jaccard coefficients which can be visualized using the same methods used for the women. Figure 7 shows events with Jaccard overlaps greater than 0.35.

Unimodal Analysis of Two-Mode Data

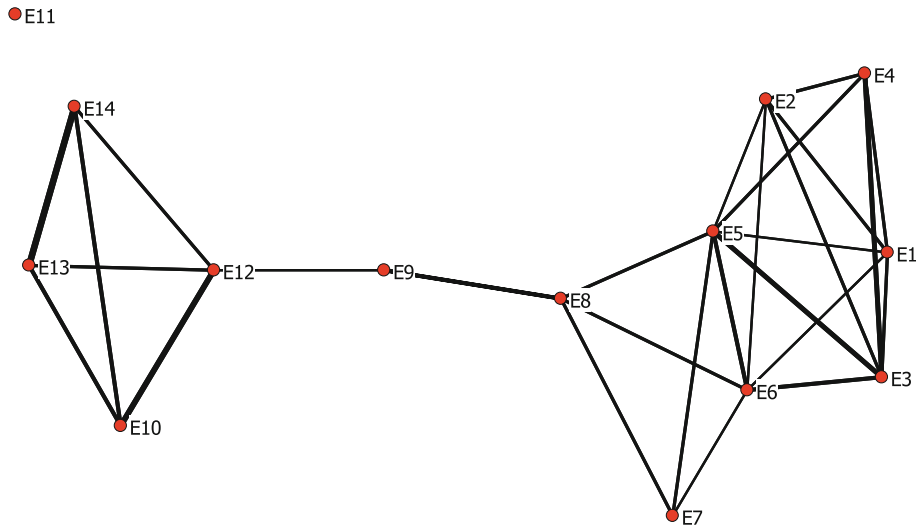
In general, analysis of 2-mode data transformed into valued 1-mode networks proceeds like any other valued network. As with visualization, this often means generating a graph from the valued data via some rule such as $(u, v) \in E$ iff $a_{ij} > q$, where q is chosen by the researcher. Typically, there is no theoretical reason for choosing any particular value of q; hence a series of different values is generally chosen and the analysis repeated for each.

There are, however, a few consequences that stem from the 2-mode origin of the data. By their very nature, many commonly used measures of similarity and dissimilarity satisfy triangle inequality laws. For example, for Euclidean distance, every triple of nodes i, j, k satisfy the following rule:

$$d_{ik} \leq d_{ij} + d_{jk} \,. \tag{5}$$



Social Network Analysis, Two-Mode Concepts in, Figure 6
Spring-embedding representation of Jaccard similarities dichotomized at > 0.4



Social Network Analysis, Two-Mode Concepts in, Figure 7
Spring-embedding representation of "ties" among events ($c_{ij} > 0.35$)

As a result, the 1-mode data (especially if not dichotomized) artifactually exhibit a certain level of transitivity that may be higher than baseline models built on simple sociometric choice data would expect. Statistics based on transitivity, such as structural holes and clustering coefficients, must similarly be interpreted with some caution in such data.

Bimodal Approaches to Two-Mode Data

Another approach to working with 2-mode data seeks to analyze both modes simultaneously. The data are seen to represent relations between two sets of nodes, forming a bipartite graph $G_B(V_1 + V_2, E)$ in which, or all u and v, $(u, v) \in E$ if and only if u and v belong to different vertex sets. In other words, all ties are between vertex sets and none are within-group. The matrix representation of such a graph can be a rectangular incidence matrix X (as in Fig. 1) or a square bipartite adjacency matrix B with $n = n_1 + n_2$ rows representing both modes, and an equal number of columns, also representing both modes. In the latter case, the original matrix X forms a submatrix of the larger adjacency matrix B in which both rows and columns index the $V_1 + V_2$ entities. The matrix B is composed of four blocks, two of which are empty, as shown in Fig. 8 Note that the original matrix X forms the top right quadrant of B, and its transpose forms the bottom left quadrant.

Bimodal Visualization of Two-Mode Data

All of the standard ways to visualize networks, such as MDS and GLAs, apply to bipartite graphs. For example, Fig. 9 shows a spring-embedding layout of the bipartite graph represented by the matrix in Fig. 8.

In the representation, two nodes are near each other roughly to the extent that the geodesic distance between them is short. Thus, events are near each other if they are attended by the same women (distance 2), and women are near each other if they attend the same events. In this example, the representation makes clear that there is a set of women on the left (Mryna, Helen, Katherine, Nora, Silvia, etc.) that attend a set of events exclusive to them (events 10 through 13), and another set of women (Evelyn, Theresa, Laura, Brenda, etc.) that have their own events (E1 through E4), and finally a set of events that both "circles" of women attend (events E6 through E9).

For small datasets, this bimodal visualization is often extremely effective for transmitting a holistic understanding of the whole dataset.

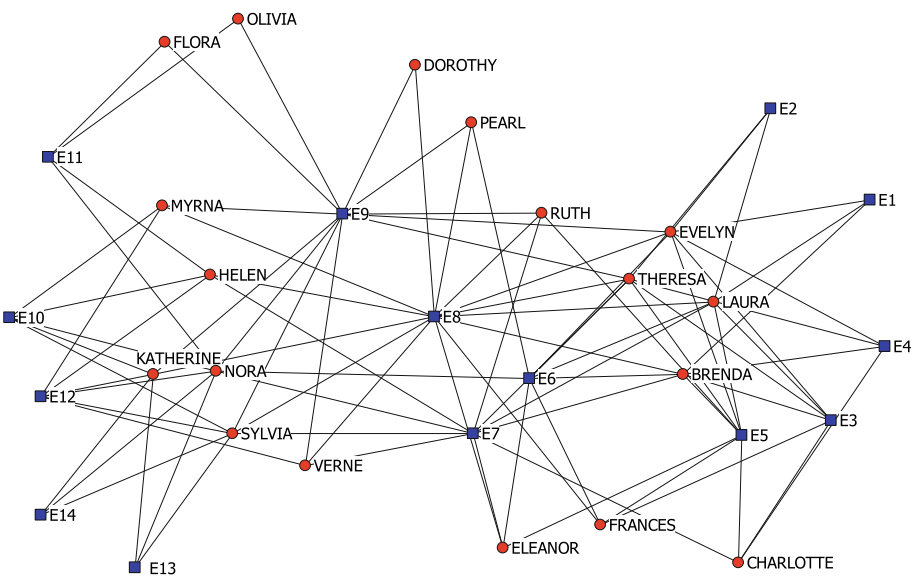
It is worth noting that there is a simple mathematical relationship between pairwise overlaps a_{ij} as defined in Eq. (1) and paths in the bipartite graph. Specifically, the number of 2-step paths between any pair of women i and j in the bipartite graph is equal to a_{ij}, the number of events they attended in common. Of course, the number of 2-step paths is simply the matrix product BB, the bipartite adjacency matrix multiplied by itself. As shown in Fig. 10, the top left block and bottom right blocks of BB are the matrices A as calculated by Eq. (1) for rows and columns respectively of X.

Bimodal Analysis of Two-Mode Data

Given a bimodal perspective, one approach to analyzing 2-mode data is to develop entirely new metrics and algo-

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	1	1	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	0	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	1	1	1	0	1	0	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	1	1	1	1	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Social Network Analysis, Two-Mode Concepts in, Figure 8
Bipartite adjacency matrix *B* created from the original DGG 2-mode matrix *X*



Social Network Analysis, Two-Mode Concepts in, Figure 9
Spring-embedding representation of bipartite graph

rithms designed specifically for 2-mode data. Such techniques take cognizance of the fact that the observed network is not just bipartite by happenstance, but could not have been any other way. In other words, taking account of the fact that the observed lack ties between certain nodes (namely, those belonging to different modes) was by design – similar to the concept of structural zeros in log-linear modeling. To date, few techniques of this kind have

	EVE	LAU	THE	BRE	CHA	FRA	ELE	PEA	RUT	VER	MYR	KAT	SYL	NOR	HEL	DOR	OLI	FLO	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14
EVELYN	8	6	7	6	3	4	3	3	3	2	2	2	2	2	1	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LAURA	6	7	6	6	3	4	4	2	3	2	1	1	2	2	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
THERESA	7	6	8	6	4	4	4	3	4	3	2	2	3	3	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BRENDA	6	6	6	7	4	4	4	2	3	2	1	1	2	2	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CHARLOTTE	3	3	4	4	4	2	2	0	2	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FRANCES	4	4	4	4	2	4	3	2	2	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ELEANOR	3	4	4	4	2	3	4	2	3	2	1	1	2	2	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PEARL	3	2	3	2	0	2	2	3	2	2	2	2	2	2	1	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RUTH	3	3	4	3	2	2	3	2	4	3	2	2	3	2	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
VERNE	2	2	3	2	1	1	2	2	3	4	3	3	4	3	3	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MYRNA	2	1	2	1	0	1	1	2	2	3	4	4	4	3	3	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
KATHERINE	2	1	2	1	0	1	1	2	2	3	4	6	6	5	3	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SYLVIA	2	2	3	2	1	1	2	2	3	4	4	6	7	6	4	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NORA	2	2	3	2	1	1	2	2	2	3	3	5	6	8	4	1	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HELEN	1	2	2	2	1	1	2	1	2	3	3	3	4	4	5	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DOROTHY	2	1	2	1	0	1	1	2	2	2	2	2	2	1	1	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OLIVIA	1	0	1	0	0	0	0	1	1	1	1	1	1	2	1	1	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FLORA	1	0	1	0	0	0	0	1	1	1	1	1	1	2	1	1	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	2	3	2	3	3	2	3	1	0	0	0	0	0
E2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	3	2	3	3	2	3	2	0	0	0	0	0
E3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3	6	4	6	5	4	5	2	0	0	0	0	0
E4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	4	4	4	3	3	3	2	0	0	0	0	0
E5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3	6	4	8	6	6	7	3	0	0	0	0	0
E6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3	5	3	6	8	5	7	4	1	1	1	1	1
E7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	4	3	6	5	10	8	5	3	2	4	2	2
E8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3	5	3	7	7	8	14	9	4	1	5	2	2
E9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	2	2	3	4	5	9	12	4	3	5	3	3
E10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	4	4	5	2	5	3	3
E11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	1	3	2	4	2	1	1
E12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	4	5	5	5	2	6	3	3
E13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	2	3	3	1	3	3	3
E14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	2	3	3	1	3	3	3

Social Network Analysis, Two-Mode Concepts in, Figure 10
Matrix *BB* giving the number of paths of length 2 between all pairs of nodes

been developed, the exception being the area of 2-mode centrality measures, which has received significant attention (e. g., [2]).

Another approach is to treat the bipartite graphs as ordinary graphs and apply all the standard algorithms and techniques of social network analysis. Effectively this assumes either that the special nature of the graphs will not affect the techniques, or that we can pretend that ties within modes could have occurred and just didn't. This approach works for a small class of methods, but by no means all. For example, calculating transitivity fails because transitive triples are impossible in bipartite graphs (all ties are between modes, which means that if $a \to b$ and $b \to c$ then a and c must be members of the same class, and therefore cannot be tied, making transitivity impossible). In contrast, if we were to adapt the definition of transitivity to be based on quadruples such that a quad is transitive if $a \to b$, $b \to c$, $c \to d$ and $a \to d$, this would be an example of the first approach.

Finally, a compromise approach is to use the standard metrics and algorithms that apply to general graphs, and then either adjust the outputs via normalization or adjust the baseline expectations for the results [5,9]. In the former case, we develop normalizations that adjust the metrics (typically by dividing by theoretical bipartite maxima), and in the latter case we derive different theoretical distributions for the statistics in question. In general, we choose the former approach when statistics are bounded between 0 and 1 and can be interpreted as proportions of maximum possible values. We choose the latter approach when the statistics are unbounded and have direct interpretations (see example of cohesion below).

Bimodal Approaches to Graph Cohesion The simplest and most common measure of network cohesion is *density* – the number of edges divided by the number of pairs of nodes (using ordered pairs in the case of directed graphs and unordered pairs in the case of undirected graphs). In bipartite graphs, of course, only edges between vertex sets are possible. As a result, the maximum possible undirected ordinary density is

$$\frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 1)} \,. \tag{6}$$

Thus, if density were calculated on a 2-mode network as if it were an ordinary graph, we would probably want to normalize the result by dividing by the sum above, otherwise the calculated density would appear misleadingly low. This is a case where the "compromise approach" discussed above is effective. (Of course, in terms of computation, it would be easier to simply calculate the average of the 2-mode matrix X rather than convert to bipartite form and then perform this adjustment.)

For measures of cohesion that are not expressed as fractions of maximum possible values, such as average geodesic path length, the need to renormalize is not as

great, since the raw values are directly interpretable. In doing so, however, we must be careful not to compare the results with standard rules of thumb or theoretical models based on simple random graphs, as these have not assumed the bipartite restrictions.

Bimodal Approaches to Centrality Four measures of centrality are used with high frequency in social network analysis today: degree, closeness, betweenness and eigenvector. We discuss each in turn.

Degree centrality, d_i, is defined as the number of ties incident upon node i. Degree centrality is typically normalized by dividing by the maximum number of ties possible, which in a graph of n nodes is $n-1$. Hence

$$d_i^* = \frac{d_i}{n-1} \,. \tag{7}$$

However, in any (non-trivial) bipartite graph, no node can be connected to all others, since this would mean within-mode ties. Instead, for a node in V_1, the maximum number of ties it could have is n_2, whereas for a node in V_2, the maximum number of ties is n_1. Hence a natural adjustment is to provide two separate normalization formulas as follows:

$$\begin{aligned} d_i^* &= \frac{d_i}{n_2} \,, \quad \text{for } i \in V_1 \,, \\ d_j^* &= \frac{d_j}{n_1} \,, \quad \text{for } j \in V_2 \,. \end{aligned} \tag{8}$$

Closeness centrality is ordinarily defined as the sum of geodesic distances from a node to all others, as shown in Eq. (8), where d_{ij} is the length of the shortest path from i to j and $n = n_1 + n_2$ is the total number of nodes. The best (smallest) score possible occurs when the node has a tie to every other node, in which case the total distance to all others is $n-1$. Thus, closeness centrality is usually normalized by dividing c_i into $n-1$.

$$\begin{aligned} c_i &= \sum_j^n d_{ij} \\ c_i^* &= \frac{n-1}{c_i} \,. \end{aligned} \tag{9}$$

In the bipartite case, the maximum number of nodes that a node can be distance 1 from is the number of nodes in the other class. For other nodes in its own class, the closest it can be is two links away. Thus, the theoretical bipartite minimum value of c_i (where $i \in V_1$) is $n_2 + 2(n_1 - 1)$ and the minimum for c_j (where $j \in V_2$) is $n_1 + 2(n_2 - 1)$. Therefore, closeness centrality can be normalized in bipartite graphs as follows:

$$\begin{aligned} c_i^* &= \frac{n_2 + 2(n_1 - 1)}{c_i} \,, \quad \text{for } i \in V_1 \,, \\ c_j^* &= \frac{n_1 + 2(n_2 - 1)}{c_j} \,, \quad \text{for } j \in V_2 \,. \end{aligned} \tag{10}$$

Betweenness centrality refers to the sum of shares of shortest paths that pass through a given node. The betweenness of node k in an ordinary graph is defined by Eq. (11), where g_{ij} is the number of geodesic paths from node i to node j, and g_{ikj} is the number of geodesic paths from i to j that pass through k.

$$b_k = \frac{1}{2} \sum_{i \neq k}^{n} \sum_{j \neq k,i}^{n} \frac{g_{ikj}}{g_{ij}} \,. \tag{11}$$

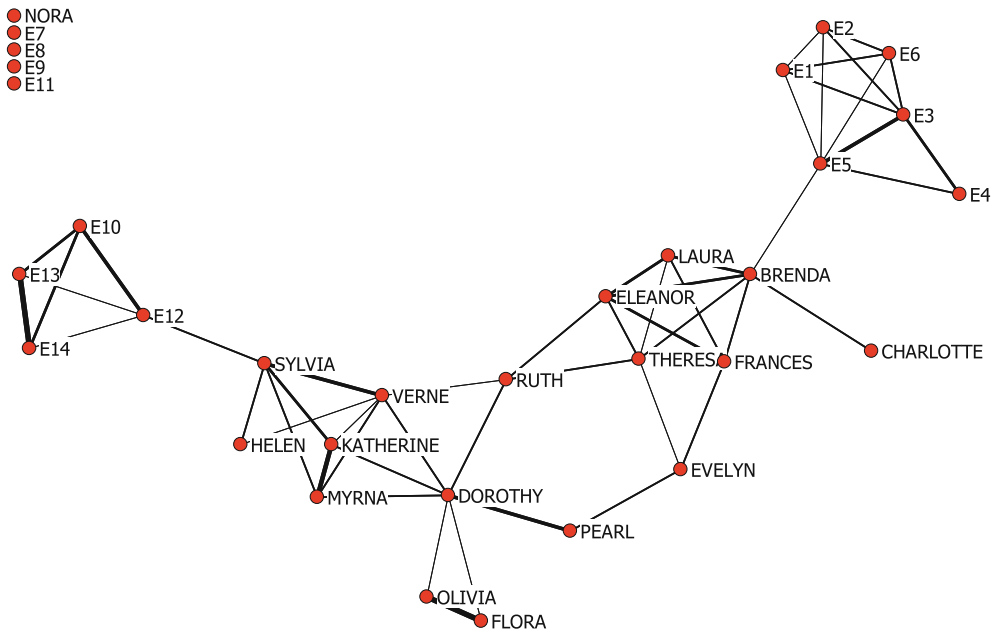
Betweenness is ordinarily normalized by dividing by $(n-1)(n-2) = n^2 - 3n + 2$, which is the maximum betweenness that any node can achieve in a graph with n nodes, which occurs for the node at the center of a star-shaped graph. This maximum is appropriate for bipartite graphs only when one mode has just one node; otherwise we must take account of the sizes of each vertex set. Eq. (12) gives the maximums for nodes in each vertex set as a function of the vertex set sizes. In the equation, x div y refers to integer division of x by y and x mod y refers to the remainder of an integer division of x by y.

$$\begin{aligned} b_{V_1 \max} &= \tfrac{1}{2}\big[n_2^2(s+1)^2 + n_2(s+1)(2t-s-1) \\ &\qquad - t(2s-t+3)\big] \\ s &= (n_1 - 1) \text{ div } n_2 \,, \quad t = (n_1 - 1) \mod n_2 \\ b_{V_2 \max} &= \tfrac{1}{2}\big[n_1^2(p+1)^2 + n_1(p+1)(2r-p-1) \\ &\qquad - r(2p-r+3)\big] \\ p &= (n_2 - 1) \text{ div } n_1, \quad r = (n_2 - 1) \mod n_1 \end{aligned} \tag{12}$$

Given these maxima, we can normalize standard betweenness centrality for bipartite graphs by dividing by these maxima, as shown in Eq. (13).

$$\begin{aligned} b_i^* &= \frac{b_i}{b_{V_1 \max}} \,, \quad \text{for } i \in V_1 \,, \\ b_j^* &= \frac{b_j}{b_{V_2 \max}} \,, \quad \text{for } j \in V_2 \,. \end{aligned} \tag{13}$$

Cohesive Subgroups Detecting cohesive subgroups is somewhat more difficult in bipartite graphs than in ordinary graphs. For example, one of the earliest formal definitions of a subgroup is the clique [11] which is defined



Social Network Analysis, Two-Mode Concepts in, Figure 11
Graph induced by Pearson correlations across biclique membership profiles. Edges indicate a correlation greater than 0.4

as a maximal complete subgraph – i. e., a set of nodes that is as large as possible subject to the condition that every pair is adjacent. However, in bipartite graphs, cliques are impossible, since in any connected triple, two nodes must be non-adjacent. Thus, cliques are not useful in this context. In addition, relaxations of the clique concept based on density or frequency of paths (such as k-plexes, lambda sets and ls-sets) do not work well, due to the sparse nature of bipartite graphs. For example, if we calculate k-plexes for the DGG dataset, we find huge numbers of very small groups. For $k = 2$, there are 394 2-plexes of size 3 or greater. For $k = 3$, there are 5,553 3-plexes, and for $k = 4$, there are 37,633 4-plexes of size 3 or greater.

However, other classical notions of cohesive subgroups make more sense for bipartite graphs. In particular, relaxations of the clique concept based on distance, such as n-cliques, n-clans and n-clubs have good interpretations in bipartite graphs and work well in practice. In fact, 2-cliques defined on the bipartite graph have the property of bipartite completeness, meaning that, within the bipartite subgraph induced by the 2-clique, every node of one mode has a tie with every node of the other mode, which is to say that all possible ties are present. In this sense, for bipartite graphs, 2-cliques capture the underlying idea of a clique better than cliques do. This notion has been formalized in the definition of a *biclique*, which is defined as a maximal complete bipartite subgraph. Mathematically, it is identical to a 2-clique computed on the bipartite representation.

As a practical example, a biclique analysis of the DGG dataset finds 68 bicliques, and a secondary analysis of similarities of nodes across biclique membership profiles does a good job of revealing structure, as shown in Fig. 11. In the figure, an edge is shown between two nodes if the Pearson correlation between their biclique membership profiles is greater than 0.4 (i. e., $(u, v) \in E$ iff $r_{ij} > 0.4$). The two groups are women are clearly shown, as well as the two groups of events that are associated with each of the groups of women. In addition, the separation of Flora and Olivia is clearly shown, as well as the bridging position of Ruth and Pearl.

The success of bicliques as 2-mode analogues of cliques suggests the possibility of creating 2-mode analogues for other cohesive subgroup concepts that we previously dismissed as unusable in the 2-mode context. For example, a *k-plex* is defined in ordinary network analysis as a maximal subgraph $S(V, E)$ of size n such that each member of the k-plex is adjacent to at least $n - k$ others. By analogy, for 2-mode networks we can define a (k_1, k_2)-*biplex* as a maximal bipartite subgraph $G(V_1, V_2, E)$ such that each node in V_1 is adjacent to $|V_2| - k_2$ others and each node in V_2 is adjacent is adjacent to $|V_1| - k_1$ others. Obviously, a (0,0)-biplex is a biclique. A useful feature of this relaxation of a biclique is that we can set different standards of

cohesiveness for each mode of the data network, perhaps reflecting the relative sizes of the modes, or the affiliative capabilities of the nodes themselves.

Bimodal Approaches to Positions and Roles Two concepts of position and role are particularly well known in social network analysis. These are structural equivalence and regular equivalence. Generalizations to 2-mode (and higher) data were developed by Borgatti [3] and Borgatti and Everett [4] and Everett and Borgatti [7]. We consider each of these in turn.

Structural Equivalence In the best definition of structural equivalence [8], two nodes u and v are said to be structurally equivalent if there exists a graph automorphism that would be an identity except that u and v are mapped to each other. In other words, given a diagram of the graph, you could swap nodes u and v (and no other nodes) without changing the structure of the network one iota. A key implication of this definition is that structurally equivalent nodes have identical relational environments – aside from each other, they are connected and not connected to exactly the same third parties. As a result, one approach to identifying structurally equivalent nodes is to compute a similarity measure among rows and columns of the adjacency matrix defining the graph.

This approach requires no modification for the bipartite case. In fact, the unimodal analysis described in Sect. "Unimodal Approaches to Two-Mode Data" (applied to each mode in turn) is precisely the same as computing structural equivalence on the 2-mode incidence matrix (Fig. 1). In addition, for certain measures such as the Jaccard coefficient, computing structural equivalence on the bipartite adjacency matrix B of Fig. 8 gives exactly the same results as computing it on the 2-mode incidence matrix X.

Another approach to structural equivalence is known as blockmodeling. Instead of (or as a result of) measuring the extent of structural equivalence between nodes, we partition the nodes into classes such that nodes in the same class are structurally equivalent (or nearly so). Given such a partition of nodes, we can reorder and partition the corresponding rows and columns of the adjacency matrix. This in turn induces to partition of matrix cells into *matrix blocks*. A characteristic property of perfect structural equivalence partitions is that matrix blocks are necessarily homogeneous with respect to cell values, and each block will consist of all 1s or all 0s (known as 1-blocks and 0-blocks respectively). Figure 12 gives an example with three equivalence classes.

In the 2-mode case, one approach is to blockmodel the bipartite adjacency matrix B. This can be done, but the bipartite structure imposes certain constraints. For example, blocks involving within-mode ties must be 0-blocks. In addition, the best 2-block partition will almost certainly be the mode partition (except in trivial cases), and in general, all other partitions will be refinements of the mode partition (i. e., they will be nested hierarchically within the mode partition).

	A1	A2	A3	B1	B2	B3	B4	C1	C2	C3
A1	0	0	0	1	1	1	1	0	0	0
A2	0	0	0	1	1	1	1	0	0	0
A3	0	0	0	1	1	1	1	0	0	0
B1	1	1	1	0	0	0	0	1	1	1
B2	1	1	1	0	0	0	0	1	1	1
B3	1	1	1	0	0	0	0	1	1	1
B4	1	1	1	0	0	0	0	1	1	1
C1	1	1	1	1	1	1	1	0	0	0
C2	1	1	1	1	1	1	1	0	0	0
C3	1	1	1	1	1	1	1	0	0	0

Social Network Analysis, Two-Mode Concepts in, Figure 12
Perfect structural equivalence blockmodel

A more elegant (and computationally efficient) approach is to work from the 2-mode incidence matrix X. In this case, we redefine a blockmodel to refer to a coordinated pair of partitions, one for the rows and one for the columns. We then redefine structural equivalence as follows. Given a 2-mode matrix X with modes V_1 and V_2, a 2-mode blockmodeling consists of a pair of partitions P_1 and P_2 such that (a) for all u, v in V_1, $p_1(u) = p_1(v)$ iff $x_{uw} = x_{vw}$ for all w in V_2 and (b) for all a, b in V_2, $p_2(a) = p_2(b)$ iff $x_{za} = x_{zb}$ for all z in V_1. (The notation $p(u)$ indicates the equivalence class that node u belongs to in partition P.) Restating this in words, a 2-mode structural equivalence blockmodeling is one in which row nodes in the same class if and only if they have identical rows, and column nodes are in the same class if and only if they have identical columns. An example involving 4 classes of rows and 3 classes of columns is shown in Fig. 13.

In empirical work, of course, we do not expect to obtain perfect 1-blocks and 0-blocks. Instead we seek partitions that minimize the number of errors (where we define an error as either a 1 inside a 0-block or a 0 inside a 1-block).

Regular Equivalence Two nodes are said to be *regularly equivalent* (i. e., play the same structural role) to each other to the extent that they have ties (lack of ties) to corresponding others who are themselves regularly equivalent to each other. That is, u and v are regularly equivalent

	E1	E2	E3	F1	F2	F3	F4	G1	G2	G3
A1	1	1	1	1	1	1	1	0	0	0
A2	1	1	1	1	1	1	1	0	0	0
A3	1	1	1	1	1	1	1	0	0	0
B1	1	1	1	0	0	0	0	0	0	0
B2	1	1	1	0	0	0	0	0	0	0
B3	1	1	1	0	0	0	0	0	0	0
B4	1	1	1	0	0	0	0	0	0	0
C1	0	0	0	1	1	1	1	0	0	0
C2	0	0	0	1	1	1	1	0	0	0
C3	0	0	0	1	1	1	1	0	0	0
D1	0	0	0	1	1	1	1	1	1	1
D2	0	0	0	1	1	1	1	1	1	1

Social Network Analysis, Two-Mode Concepts in, Figure 13
A 2-mode structural equivalence blockmodel in which the row partition has 4 classes and the column partition has 3 classes

if, (a) for all x such that u has a tie to x, there exists a y that is regularly equivalent to x such that v has a tie to y, and (b) for all x such that u has an incoming tie from x, there exists a y that is regularly equivalent to x such that v has an incoming tie from y. Whereas in structural equivalence equivalent nodes are connected to the same others, in regular equivalence equivalent nodes are connected to the same types of others.

Alternatively, we can define regular equivalence in terms of a partition C of nodes into labeled classes (known as colors) such that regularly equivalent nodes are required to have the same colors in their neighborhoods. That is, for all nodes u and v,

$$C(u) = C(v) \text{ implies } C(N(u)) = C(N(v)) , \qquad (14)$$

where $C(N(u))$ is the set of distinct colors found among the nodes constituting u's immediate neighborhood.[1] Taking a blockmodeling perspective, it can readily be seen that Eq. (14) implies a blockmodel in which each matrix block is either entirely zero (a "zeroblock") or contains at least one 1 in every column and in every row (a "regular oneblock"), as illustrated in Fig. 14.

Equation (14) lends itself easily to the bipartite case. Instead of seeking a single partition of nodes, we seek a different partition for each mode. Labeling these partitions R and C, we modify our definition such that for all $u, v \in V_1$ and $x, y \in V_2$,

$$\begin{aligned} &C(u) = C(v) \text{ implies } R(N(u)) = R(N(v)) , \quad \text{and} \\ &R(x) = R(y) \text{ implies } C(N(x)) = C(N(y)) . \end{aligned} \qquad (15)$$

[1]For directed graphs, we require $C(N^o(u)) = C(N^o(v))$ and $C(N^i(u)) = C(N^i(v))$, where $N^o(v)$ refers to the set of nodes that v sends a tie to, and $N^i(u)$ refers to the set of nodes that v receives a tie from.

	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10
N1	0	0	0	1	0	0	1	0	0	0
N2	0	0	0	0	1	0	0	0	0	0
N3	0	0	0	1	1	1	1	0	0	0
N4	1	0	0	0	0	0	0	0	1	1
N5	1	0	1	0	0	0	0	1	1	0
N6	0	1	0	0	0	0	0	0	1	1
N7	0	1	1	0	0	0	0	1	0	1
N8	1	0	1	0	1	1	0	0	0	0
N9	0	0	1	0	0	1	0	0	0	0
N10	0	1	0	1	1	0	1	0	0	0

Social Network Analysis, Two-Mode Concepts in, Figure 14
Perfect regular equivalence blockmodel on an ordinary graph

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
C1	1	0	1	0	1	1	0	0	0	0
C2	0	0	1	0	0	1	0	0	0	0
C3	0	1	0	1	1	0	1	0	0	0
C4	1	0	0	0	0	0	0	0	1	1
C5	1	0	1	0	0	0	0	1	1	0
C6	0	1	0	0	0	0	0	0	1	1
C7	0	1	1	0	0	0	0	1	0	1
C8	0	0	0	0	1	1	0	0	0	0
C9	0	0	0	0	0	1	0	0	0	0
C10	0	0	0	1	1	0	1	0	0	0
C11	1	0	1	0	0	0	0	0	1	1
C12	0	1	0	0	0	0	0	1	0	1

Social Network Analysis, Two-Mode Concepts in, Figure 15
A 2-mode regular equivalence blockmodel

Of course, the neighbors N() of any node consist entirely of nodes belonging to the other mode. Thinking in terms of the 2-mode incidence matrix shown in Fig. 1, Eq. (15) implies that we can section the matrix into rectangular blocks such that each block is zeroblock or a regular 1-block. An example of a 2-mode regular blockmodel is shown in Fig. 15.

As with structural equivalence, in empirical work we do not expect to obtain perfect 1-blocks and 0-blocks. Instead we seek partitions induce matrix blocks that are as nearly regular as possible.

Conclusion

In this chapter we have considered methods of visualizing and analyzing 2-mode network data. An examination of the chapter suggests that four strategies are employed for handling 2-mode data. First, there is converting the data into two separate 1-mode datasets, and analyzing each separately. Second, there is using existing 1-mode methods on a bipartite representation of the 2-mode data, and essentially ignoring its special features. Third, there is using 1-mode methods but either adjusting the interpretation or

applying some kind of normalization to adjust the results. Finally, there is developing new methods designed specifically for the 2-mode case. The chapter has provided examples of all four strategies.

It is worth reminding the reader that only 2-mode data viewed in a network context has been considered here. In principle, most data dealt with by social scientists is structured as a 2-mode matrix of cases (rows) and variables (columns) and most statistical techniques assume such data. This chapter obviously does not discuss this vast pantheon of techniques, although it is clear that many would be appropriate, particularly structure-finding techniques such as factor analysis and correspondence analysis.

Future Directions

Although it would be a mistake to think of 2-mode data as an advance over 1-mode data, it is important to note that there are many cases were extending network analysis methodology to more than 2 modes is desirable. For example, we might analyze membership in organizations over time, yielding a 3-way, 3-mode data matrix of relations that is person by organization by period. Similarly, we might be interested in modeling the intellectual landscape of an academic field, representing publications as k-mode bundles of authors, journals, years, institutional affiliations and topics.

Bibliography

1. Bonacich P (1972) Techniques for analyzing overlapping memberships. In: Costner HL (ed) Sociological methodology. Jossey-Bass, San Francisco, pp 176–185
2. Bonacich P (1991) Simultaneous group and individual centralities. Soc Netw 13:155–168
3. Borgatti SP (1989) Regular equivalence in graphs, hypergraphs, and matrices. Doctoral Dissertation, Univ of California, Irvine, Ann Arbor
4. Borgatti SP, Everett M G (1992) Regular blockmodels of multiway, multimode matrices. Soc Netw 14:91–120
5. Borgatti SP, Everett MG (1997) Network analysis of 2-mode data. Soc Netw 19(3):243–269
6. Davis A, Gardner B, Gardner M (1941) Deep South: A Social Anthropological Study of Caste and Class. University of Chicago Press, Chicago
7. Everett MG, Borgatti SP (1993) An extension of regular colouring of graphs to digraphs, networks and hypergraphs. Soc Netw 15:237–254
8. Everett MG, Borgatti SP (1994) Regular equivalence: General theory. J Math Sociol 19(1):29–52
9. Faust K (1997) Centrality in affiliation networks. Soc Netw 19:157–191
10. Kamada T, Kawai S (1989) An algorithm for drawing general undirected graphs. Inf Process Lett 31:7–15
11. Luce RD, Perry AD (1949) A method of matrix analysis of group structure. Psychometrika 20:319–327

Social Networks, Algebraic Models for

PHILIPPA PATTISON
Department of Psychology, University of Melbourne, Parkville, Australia

Article Outline

Glossary
Definition of the Subject
Introduction
Algebras from Networks
Algebraic Structure
Algebraic Analysis
Algebraic and Network Mappings
Examples
Future Directions
Bibliography

Glossary

Social network A *social network* comprises a set of relationships among members of a set $N = \{1, 2, \ldots, n\}$ of actors. It can be represented by an $n \times n$ binary array X recording the presence or absence of a social relationship, or *tie*, between each pair (or ordered pair) of members of $N = \{1, 2, \ldots, n\}$. If there is a relationship from actor k to actor l, we write $X(k, l) = 1$; otherwise, $X(k, l) = 0$. If the relationship is a property of a pair of actors, the network is *nondirected*; if it is a property of an ordered pair, the network is termed *directed*. The directed network X may also be regarded as a *binary relation* R_X on the set N with $(k, l) \in R_X$ if and only if $X(k, l) = 1$; equivalently, it may be construed as a *directed graph* with *node set* N and *arc set* R_X, with an *arc* from node k to node l if and only if $(k, l) \in R_X$.

Affiliation network An *affiliation network* is an $n \times g$ binary array X recording the membership of each of a set N of actors in a prescribed set G of *groups*, with $X(k, l) = 1$ if actor k is a member of group l, and $X(k, l) = 0$ otherwise.

Multiple network A *multiple network* is a collection of networks for each of a set of r relations. We let $X_m(k, l) = 1$ if the tie from k to l corresponding to the relation of *type* m is present; and $X_m(k , l) = 0$ if the tie is absent. Nodes k and l are joined by a *labeled walk* with *label* $Y_1\ Y_2\ \ldots\ Y_j$ if there is a sequence of nodes $k = k_0, k_1, \ldots, k_j = l$, for which $Y_h(k_{h-1}, k_h) = 1$ for $h = 1, 2, \ldots, j$.

Local network The *(1-)neighborhood* of a subset P of actors in a network is defined to be the set $P \cup \{l \in N\colon X_m(k,l) = 1$ for some $k \in P$ and some relation $m\}$. The *q-neighborhood* of P is then defined recursively as the 1-neighborhood of the $(q-1)$-neighborhood of P. The *q-local network* of the subset P of N is the network restricted to its q-neighborhood.

Algebra A (*partially ordered*) *algebra* is a triple $[S, F, \leq]$, where S is a nonempty set of elements (usually assumed to be finite), F is a specified set of operations, f_α, each mapping a power $S^{n(\alpha)}$ of S into S, for some non-negative finite integer $n(\alpha)$, and $\leq$ is a partial order on S. Each operation f_α is assumed to be *isotone* in each of its variables: that is, if $x_i \leq y_i\,(x_i, y_i \in S;\ i = 1, 2, \ldots, n(\alpha))$, then $f_\alpha(x_1, x_2, \ldots, x_{n(\alpha)}) \leq f_\alpha(y_1, y_2, \ldots, y_{n(\alpha)})$. A *family of algebras* is a collection of algebras each having the same set F of operations and satisfying a specified set of postulates. Two algebras belonging to the same family are termed *similar*.

Partial algebra A *partial algebra* is a triple $[S, F, \leq]$, where S is a nonempty set of elements (usually assumed to be finite), F is a specified set of *partial* operations, f_α, each mapping some subset $T^{(\alpha)}$ of $S^{n(\alpha)}$ into S, for some non-negative finite integer $n(\alpha)$, and $\leq$ is a partial order on S. Each partial operation f_α is assumed to be *isotone* in each of its variables: that is, if $x_i \leq y_i\,(x_i, y_i \in S;\ i = 1, 2, \ldots, n(\alpha))$, then $f_\alpha(x_1, x_2, \ldots, x_{n(\alpha)}) \leq f_\alpha(y_1, y_2, \ldots, y_{n(\alpha)})$ provided that both $(x_1, x_2, \ldots, x_{n(\alpha)})\,, (y_1, y_2, \ldots, y_{n(\alpha)}) \in T^{(\alpha)}$. A *family of partial algebras* is a collection of algebras each having the same set F of partial operations defined on the same subsets $T^{(\alpha)}$ of the power sets $S^{n(\alpha)}$.

Semigroup A *(partially ordered) semigroup* is an algebra $[S, F, \leq]$ in which F comprises a single binary operation f satisfying the *associativity* condition:

$$f(f(x, y), z) = f(x, f(y, z))\,.$$

Lattice A *lattice* L is an algebra $[S, F, \leq]$ in which F comprises two associative and commutative binary operations, $\wedge$ and $\vee$ (termed *meet* and *join*, respectively) satisfying the identities:

$$x \wedge x = x\,, \quad x \vee x = x$$

and

$$x \wedge (x \vee y) = x \vee (x \wedge y) = x\,.$$

The operations are isotone, so that $x \leq y$ is equivalent to the pair of conditions:

$$x \wedge y = x \quad \text{and} \quad x \vee y = y\,,$$

and the operations of meet and join may be interpreted as the greatest lower bound and least upper bound, respectively. A lattice L is *distributive* if the identity

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$$

holds. A lattice L is *modular* if, whenever $x \leq z$, then

$$x \vee (y \wedge z) = (x \vee y) \wedge z\,.$$

Role algebra A *role algebra* is an algebra $[S, F, \leq]$ in which F comprises a single binary composition operation satisfying the condition: $s \leq t$ in S implies $su \leq tu$ in S, for any $u \in W$.

Homomorphism An (*isotone*) *homomorphism* from an algebra $A = [S, F, \leq]$ onto a similar algebra $B = [T, F, \leq]$ is a mapping $\phi\colon S \to T$ such that,

(i) for all $f_\alpha \in F$ and $x_i \in S$,

$$\phi(f_\alpha(x_1, x_2, \ldots, x_{n(\alpha)})) = f_\alpha(\phi(x_1)\,, \phi(x_2), \ldots, \phi(x_{n(\alpha)}))\,; \quad \text{and}$$

(ii) $x \leq y$ in S implies $\phi(x) \leq \phi(y)$ in T.

The algebra B is termed a (*homomorphic*) *image* of A, and we write $B = \phi(A)$. Each homomorphism ϕ from $A = [S, F, \leq]$ onto $B = [T, F, \leq]$ has a corresponding binary relation π on S (termed here a *π-relation*) in which $(x, y) \in \pi$ if and only if $\phi(y) \leq \phi(x)$. The equivalence relation σ_π defined by $(x, y) \in \sigma_\pi$ if and only if $(x, y) \in \pi$ and $(y, x) \in \pi$ is termed a *congruence relation*.

Homomorphism lattice The *homomorphism lattice* $L(A)$ of the algebra $A = [S, F, \leq]$ is the collection of all homomorphisms of A partially ordered by the relation: $\phi_1 \leq \phi_2$ if, for all $x, y \in S$, $\phi_2(x) \leq \phi_2(y)$ implies $\phi_1(x) \leq \phi_1(y)$. The *lattice $L_\pi(A)$ of π-relations* on A, dual to $L(A)$, has the partial ordering: $\pi_1 \leq \pi_2$ if $(x, y) \in \pi_1$ implies $(x, y) \in \pi_2$, for any $x, y \in S$.

Definition of the Subject

Algebraic models have been proposed to represent structure in social networks. They are usually constructed from a set of operations and relations defined on network constituents such as paths or walks in the network, or vectors of ties directed to or from individual network members. The algebra represents the relationships among these

constituents, for example, relations of ordering among all possible walks in a multiple network, or relations of overlap and ordering among profiles of group membership. To date, they have been used to represent kinship structures [6,9,48], role structures and stability in multiple networks [1,5,8,36], states of diffusion processes in networks [30], informal hierarchy [17], connectivity structures [13] and the structure of membership in affiliation networks [16,38]. In several cases, partial algebraic representations have also been proposed [35,39,40].

Algebraic models have usually been proposed as exact structural representations and, in many cases, therefore, their application to network data has been coupled with some form of prior data aggregation such as blockmodeling [5,50]; see also the entry on ► Positional Analysis and Blockmodeling. More rarely, algebraic construction and data aggregation have been combined into a single step [8,33].

Introduction

Algebraic models for relational structures such as networks have their foundations in anthropology and were developed for general network structures by Lorrain, White and colleagues [5,26,27]. These general network representations based on path structures in networks emerged from more specific models for path structures in kinship systems, including permutation group models for Australian aboriginal kinship systems [48]. Boyd extended this approach in an important way when he proposed that the structural evolution of kinship systems could be analyzed in terms of homomorphic mappings [6]; applications to network semigroups were described later [5,37].

Lattice representations for relational structures have a long history in mathematics [2] and applications to relational data in many domains have been developed extensively by Wille and collaborators in formal concept analysis [19,20,51]. They have also been applied to affiliation network structures [16,45,47].

The form of algebraic structure regarded as appropriate in a particular context is a matter for substantive consideration. The large majority of forms that have been proposed to date have depended upon the importance of paths or walks in the network, the symmetric or asymmetric nature of network ties, and the patterns of overlap among tie partners. A focus on one or more of these features has resulted in algebraic groups, semigroups, partially ordered semigroups, local role algebras and lattices as representations of different aspects of structure in multiple social networks and affiliation networks [5,8,14,16,29,36,48] and it is on these forms that we focus here.

Algebras from Networks

A Language for Discriminating Among Relations

We begin with a finite set $\Sigma = \{x_1, x_2, \ldots, x_r\}$ of r relational terms. We term Σ a relational *alphabet* and call each term x_i in Σ a *letter* of the alphabet. The letters $x_1, x_2, \ldots$ denote relational terms such as "friendship", "co-worker", "confidant", and so on. If $r = 1$, the set Σ simply identifies a single network relation.

In order to construct algebraic representations that are sensitive to social structural form, claims in the social network literature can be used to introduce a *free word algebra* on the alphabet Σ. The algebra is intended to provide a language for describing more complex relational terms that can be constructed from the primitive set Σ and so to provide the means for discriminating among social structural forms. The claims about appropriate ways of constructing new terms from the primitive set may be drawn from a number of sources, but one set of such claims is helpfully laid out by White [49]. White argues that there are three universal ways in which types of tie (i. e., relational terms) are discriminated. The first is in terms of patterns of overlap; the second in terms of asymmetry of relation; and the third in terms of the institutionalization of indirect ties. The overlap of two relational ties refers to their joint occurrence and can arguably be described using an *intersection* operation, $\cap$. The asymmetry of a tie can be described by the relationship between the tie and its *converse*, that is the coding of whether, for two individuals k and l, there is a tie from k to l and from l to k. The tie is *symmetric* if the tie from k to l occurs only in conjunction with the tie from l to k, and the tie is *strictly asymmetric* if a tie from k to l occurs only in the absence of a tie from l to k. We use the expression x' to denote the converse of the relational term x. Indirect ties may be described by a *composition*, or *concatenation* operation, denoted $^\circ$. If x_1 refers to the term "friendship" and x_2 to "coworker", then the concatenation $x_1{}^\circ x_2$ may be used to refer to a relation that holds between a person and his or her friend's co-workers.

We denote by F the set of operations from which our descriptive language is to be built. Unless otherwise specified, we assume here that F comprises the three operations just described; that is $F = \{\cap, ', ^\circ\}$. Each operation f_α in F possesses an *arity*, $n(\alpha)$, that determines the form of new terms created from Σ by f_α: in particular, f_α leads to expressions of the form $f_\alpha(y_1, y_2, \ldots, y_{n(\alpha)})$; $y_m \in \Sigma$, $m = 1, 2, \ldots, n(\alpha)$. The operations $\cap$ and $^\circ$ have arity 2, that is, they are *binary* and lead to expressions that can be written as $y_1 \cap y_2$, and $y_1{}^\circ y_2$, respectively, whereas

the converse operation $'$ is *unary* (that is, it has an arity of 1) and leads to an expression that can be denoted y_1' (where $y_1, y_2 \in \Sigma$).

We describe a recursive construction of a free word algebra $W = [\Sigma, F]$ from the alphabet Σ and the set F of operations [2]. We call each letter in Σ an F-polynomial of *rank* 0. For any positive integer ρ, we define an F-polynomial of rank ρ recursively as an expression, termed a *word*, of the form $f_\alpha(u_1, \ldots, u_{n(\alpha)})$, where each u_j is an F-polynomial of rank $\leq \rho - 1$ (and at least one u_j is an F-polynomial of rank $\rho - 1$). The *(free) word algebra* $W = [\Sigma, F]$ comprises all distinct F-polynomials of finite rank (where $f_\alpha(u_1, \ldots, u_{n(\alpha)})$ and $f_\beta(v_1, \ldots, v_{n(\beta)})$ are distinct unless $\alpha = \beta$ and $u_k = v_k$ for all $k = 1, \ldots, n(\alpha)$).

It is sometimes convenient to restrict attention to the collection of words of rank no greater than some fixed integer ρ. So we also define $W_\rho = [\Sigma, F]_\rho$ as the subset of the free word algebra $W = [\Sigma, F]$ comprising words of rank $\leq \rho$.

For example, suppose that F comprises the operations $\{\cap, ', \circ\}$ and that $\Sigma = \{x, y\}$. Then the elements of rank 0 in the word algebra $W = [\Sigma, F]$ are $\{x, y\}$; the elements of rank 1 are $\{x \cap x, x \cap y, y \cap x, y \cap y, x', y', x^\circ x, x^\circ y, y^\circ x, y^\circ y\}$; elements of rank 3 include $x \cap x \cap x, x \cap (x^\circ x), (x^\circ x) \cap x, x \cap x', x' \cap x$, etc; and so on. In this case, the free word algebra $W = [\Sigma, F]$ comprises all composite relational terms that can be constructed from the primitive terms x and y using the intersection, converse and concatenation operations.

The free word algebra simply provides a language that permits certain forms of discrimination among types of tie to be encoded. It makes no structural assumptions whatever, although we might anticipate from the intended interpretation of the terms in F that certain terms in W will not need to be discriminated. For instance, we might suppose that: the intersection operation can be assumed to be *idempotent* ($x \cap x = x$, for all $x \in W$), *commutative* ($x \cap y = y \cap x$, for all $x, y \in W$) and *associative* ($(x \cap y) \cap z = x \cap (y \cap z)$, for all $x, y, z \in W$); that the converse operation satisfies $x'' = x$, for all $x \in W$; and that the concatenation operation is also associative. Rather than impose such postulates at this abstract level, though, we move directly to the evaluation of the terms in W for an empirical realization of a (multiple) network.

Algebras for Multiple Networks

The language provided by the free algebra $W = [\Sigma, F]$ can be used to determine all possible discriminations among ties that our structural claims suggest as appropriate. To describe the structural forms to which observed network ties give rise, each of the derived relations for an observed multiple network can be computed. There is a straightforward correspondence between the words u in $W = [\Sigma, F]$ expressed in terms of letters in Σ and operations in F, on the one hand, and Boolean matrix manipulations of the binary constituents in the multiple network, on the other. The correspondence is achieved by replacing:

- the relational term x_m by the array X_m;
- the intersection operation $\cap$ by Boolean intersection of arrays (i. e., element-wise Boolean multiplication, denoted $\bullet$); and
- the composition operation $\circ$ by Boolean matrix multiplication.

Boolean intersection, transposition and composition are defined for $n \times n$ binary arrays Y_1 and Y_2 by:

- $Y_1 \bullet Y_2(k, l) = Y_1(k, l) Y_2(k, l)$;
- $Y_1'(k, l) = Y_1(l, k)$; and
- $Y_1 Y_2(k, l) = Y_1(k, 1) Y_2(1, l) + Y_1(k, 2) Y_2(2, l) + \cdots + Y_1(k, n) Y_2(n, l)$.

The term $u \in W$ then corresponds to an $n \times n$ Boolean array, denoted X_u, with $X_u(k, l) = 1$ if and only if there is a relational tie described by the word u from k to l. We also define a partial order on the set of relations $\{X_u : u \in W\}$ by:

$$X_u \leq X_v \quad \text{if and only if } X_u(k, l) \leq X_v(k, l) \text{ for all } k,\ l \in N\ .$$

Suppose that we evaluate X_u for all $u \in W$. We define a generalized version of the *Axiom of Quality* by setting $u \leq v$ in W whenever $X_u \leq X_v$ [5].

Defining an induced partition θ on W by $(u, v) \in \theta$ if and only if $u \leq v$ and $v \leq u$ and letting $S = W/\theta$ be the set of associated classes with the induced partial order:

$$s \leq t \text{ in } S \quad \text{if } u \in s\ ,\ v \in t \text{ and } u \leq v \text{ in } W\ ,$$

we can then define a general (partially ordered) algebraic structure on S.

A (*partially ordered*) *algebra* is a triple $[S, F, \leq]$, where S is a nonempty set of elements (usually assumed to be finite), F is a specified set of operations, f_α, each mapping a power $S^{n(\alpha)}$ of S into S, for some non-negative finite integer $n(\alpha)$, and $\leq$ is a partial order on S. Each operation f_α is assumed to be *isotone* in each of its variables: that is, if $x_i \leq y_i$ ($x_i, y_i \in S$; $i = 1, 2, \ldots, n(\alpha)$), then $f_\alpha(x_1, x_2, \ldots, x_{n(\alpha)}) \leq f_\alpha(y_1, y_2, \ldots, y_{n(\alpha)})$. A *family of algebras* is a collection of algebras each having the

same set F of operations and satisfying a specified set of postulates. Two algebras belonging to the same family are termed *similar*.

For $F = \{^\circ\}$, the algebra $[S, F, \leq]$ is a *partially ordered semigroup* [36].

The Semigroup of a Multiple Network

For multiple networks, a composite relation constructed as the concatenation of observed relations records the presence of a particular form of *labelled walk*. Specifically, actor k is connected to actor l by a walk with label u if and only if $X_u(k, l) = 1$. If $X_u \leq X_v$, two actors who are joined by a walk with label u are also necessarily joined by a walk with label v.

An algorithm for constructing the semigroup of a multiple network $[S, \{^\circ\}, \leq]$ is as follows:

(S1) Let $W_1 = \{X_1, X_2, \ldots, X_r\}$ and set $i = 1$.
(S2) Construct the networks with edge sets WZ for each $W \in W_i$ and each $Z \in W_1$. Place WZ in W_{i+1} if $WZ \neq Y$, for any $Y \in W_j$, $j = 1, 2, \ldots, i+1$; otherwise, $WZ = Y$ for some $Y \in W_j$ $(j = 1, 2, \ldots, i+1)$ is an *equation* in the semigroup S.
(S3) If $W_{i+1} = \emptyset$, stop and compute the partial order among elements in $W_1 \cup W_2 \cup \ldots$; otherwise, set i to i+1 and return to step S2.

The equations $WZ = Y$ may be recorded in a *right multiplication table* in which Y appears as the table entry corresponding to the row labelled W and the column labelled Z. The partial ordering on S is constructed from the collection of networks in $W_1, W_2, \ldots$.

Lattices from Networks

Lattice representations for networks arise in two important ways. In the first, elements of the lattices are elements of $S = W/\theta$ for $F \supseteq \{\cap, ^\circ\}$. The meet of two elements is their intersection, and their join is the least relation in S which is greater than or equal to both of them. Indeed, the algebra $[S, F, \leq]$ is both a lattice and a partially ordered semigroup in this case.

In the second type of lattice construction, the elements correspond to rows or columns of an affiliation network. Since the affiliation array provides an explicit representation of the co-constitution of groups by actors and actors by groups, the lattice structures to which it gives rise can be used to analyze the duality of actors and groups [10,15,16,31,33,38,47].

A lattice denoted $L(X)$ is generated by the rows of a matrix X under the intersection operation; it is termed the *Galois* or *concept lattice* of X [20]. In $L(X)$, the meet of two elements is equal to their intersection, while their join is the minimal element in the lattice which is greater than or equal to both of them. A lattice can also be constructed from the rows of the matrix X under the union operation. In this lattice, termed the *Zareckii lattice* by Boyd, the join of two vectors is equal to their union, and their meet is the maximal vector which is less than or equal to both of them [8].

To construct the Galois lattice, one simply needs to compute all distinct intersections among rows of X, add the vector $[1\ 1 \ldots 1]$ and compute the partial order among the resulting set of row vectors.

Role Algebras from Networks

Role algebras have been proposed as local network analogues of the semigroup representation of relational structure in an entire social network. The representation has been presented in several slightly different forms but each is derived from an original formulation proposed by Mandel [11,29,53,54].

Denote by $k * X_u$ the vector $[X_u(k, 1), X_u(k, 2), \ldots, X_u(k, n)]$ for $u \in W$. The vector indicates whether there is a relationship of type u from actor k to each other actor in the network. As before, we can envisage evaluating $k * X_u$ for all $u \in W$. We then define a k-centered version of the *Axiom of Quality* by setting

$$u \leq v \text{ in } W \quad \text{whenever} \quad k * X_u \leq k * X_v$$

and an induced partition θ on W by $(u, v) \in \pi$ if and only if $u \leq v$ and $v \leq u$. Also as before, we let $S = W/\theta$ denote the set of associated classes with the induced partial order:

$$s \leq t \text{ in } S \quad \text{if } u \in s\,,\ v \in t \ \text{ and } \ u \leq v \text{ in } W\,.$$

The algebra $[S, F, \leq]$ is termed a *role algebra*.

The local role algebra for actor k may be constructed as follows.

(RA1) Let $W_1 = \{X_1, X_2, \ldots, X_r\}$ and set $i = 1$.
(RA2) Construct the vectors $k * WZ$ for each $W \in W_i$ and each $Z \in W_1$. Place WZ in W_{i+1} if $k * WZ \neq k * Y$, for any $Y \in W_j$, $j = 1, 2, \ldots, i+1$; otherwise, $k * WZ = k * Y$ for some $Y \in W_j$ $(j = 1, 2, \ldots, i+1)$ and $WZ = Y$ is an *equation* in the role algebra.
(RA3) If $W_{i+1} = \emptyset$, stop and compute the partial order among elements in $W_1 \cup W_2 \cup \ldots$; otherwise, set i to i+1 and return to step RA2.

Partial Algebras

To achieve algebraic closure, it may be necessary to compute relations corresponding to words of high rank in W and, in the case of some representations at least, it can be argued that some restriction on these repeated operations should be imposed. Mandel, for instance, made such an argument for the composition operation in social networks; he claimed that short network paths are likely to be more salient to the members of the network than longer ones [29]. These considerations can be formalized with the introduction of partial algebraic structures, whose elements are subject to rank restrictions on some or all of the operations used to construct the algebra.

Let $W_\rho = [\Sigma, F]_\rho$ be the subset of the free word algebra $W = [\Sigma, F]$ comprising words of rank $\leq \rho$. Define a partial order on the elements of W_ρ by the generalized Axiom of Quality:

$$u \leq v \text{ in } W_\rho \quad \text{whenever} \quad X_u \leq X_v \,.$$

The relation $\leq$ is clearly both *reflexive* ($u \leq u$ for all u) and *transitive* ($u \leq v$ and $v \leq w$ implies $u \leq w$ for any u, v, w), that is, a *quasi-order*. Further, each operation f_α is *isotone* in each of its variables whenever the result of the operation is contained in W_ρ; that is, if $x_i \leq y_i$ ($x_i, y_i \in W_\rho$; $i = 1, 2, \ldots, n(\alpha)$), then $f_\alpha(x_1, x_2, \ldots, x_{n(\alpha)}) \leq f_\alpha(y_1, y_2, \ldots, y_{n(\alpha)})$, provided that both $f_\alpha(x_1, x_2, \ldots, x_{n(\alpha)}), f_\alpha(y_1, y_2, \ldots, y_{n(\alpha)}) \in W_\rho$. Thus it follows that the partial ordering $\leq$ on W_ρ gives rise to a *partial* algebra in the following sense.

A *partial algebra* is a triple $[S, F, \leq]$, where S is a nonempty set of elements (usually assumed to be finite), F is a specified set of *partial* operations, f_α, each mapping some subset $T^{(\alpha)}$ of $S^{n(\alpha)}$ into S, for some non-negative finite integer $n(\alpha)$, and $\leq$ is a partial order on S. Each partial operation f_α is assumed to be *isotone* in each of its variables: that is, if $x_i \leq y_i$ ($x_i, y_i \in S$; $i = 1, 2, \ldots, n(\alpha)$), then $f_\alpha(x_1, x_2, \ldots, x_{n(\alpha)}) \leq f_\alpha(y_1, y_2, \ldots, y_{n(\alpha)})$ provided that both $(x_1, x_2, \ldots, x_{n(\alpha)}), (y_1, y_2, \ldots, y_{n(\alpha)}) \in T^{(\alpha)}$. A *family of partial algebras* is a collection of algebras each having the same set F of partial operations defined on the same subsets $T^{(\alpha)}$ of the power sets $S^{n(\alpha)}$.

Partial semigroup algebras may be constructed from labelled walks in a multiple network of some maximum length ρ [40]. For role algebras, the partial role algebra for a node k derived from W_ρ is the role algebra associated with the ρ-local neighborhood of k [39].

Algebraic Structure

In some cases, it is possible to analyze the structural properties of algebraic representations of networks.

The Semigroup of a (Single) Network Under Composition

We first consider the semigroup for a single network array X. We say that there is a *walk* from actor k to actor l if $X(k, l) = 1$. There is a walk of *length* p from k to l if there is a sequence of actors $k = k_0, k_1, \ldots, k_p = l$ such that $X(k_{i-1}, k_i) = 1$, for $i = 1, 2, \ldots, p$. It is readily seen that $X^p(k, l) = 1$ if and only if there is a walk of length p from k to l. The sequence $X, X^2, \ldots$ generated by the composition operation is therefore a sequence recording the presence of walks of different lengths between all pairs of network actors. The semigroup structure of this sequence therefore describes relations of equality and ordering among walks in the network.

The sequence can be described with the help of the following definition. Let p be the least integer such that

$$X^p = X^q \quad \text{for some} \quad q > p \,.$$

Also let $q = p + d$ ($d \geq 1$) be the least integer satisfying this relation. The integers p and d are termed the *index* and *period* of the network, respectively. It can readily be seen that the sequence has the form

$$X, \ldots, X^{p-1}, X^p, \ldots, X^{p+d-1}, X^p, \ldots, X^{p+d-1}, \ldots \,.$$

The set $\{X, X^2, \ldots, X^{p+d-1}\}$ defines a semigroup S of which $G = \{X^p, \ldots, X^{p+d-1}\}$ is a group (defined below).

It is natural to ask what is the relationship between properties of a network and its index and period.

For a strongly connected network X (that is, a network in which there is a walk of some length from each actor to each other actor), the period d is the greatest common denominator of all integers m such that m is the length of a cycle in X (that is, a walk with the same initial and final node). If X has more than one strong component, its period is the least common multiple of the periods of those strong components [24].

A network X with index 1 and period 1 is termed *idempotent*; for a network of index p and period 1, the network X^p is idempotent. It may be shown that every idempotent network is a pseudo-order, defined as follows [43]. Let Z be a quasi-order on N and let $e_Z = Z \cap Z'$. An actor $k \in N$ is termed *Z-strict* if there is no distinct actor $l \neq k$ such that $Z(k, l) = 1$. The subset H of N is *Z-permissible* if each of its actors is Z-strict and there are no pairs $k, l \in N$ such that k covers l or l covers k (recall that k covers l if $Z(k, j) = 1$ and $Z(j, l) = 1$ implies $j = k$ or $j = l$). Define the relation Z_H by

$$Z_H(k, l) = 1 \quad \text{iff} \quad k = l \text{ and } k \in H \,.$$

Then a network X is a *pseudo-order* if

$$X = Z \backslash Z_H$$

for some quasi-order Z and Z-permissible subset H of N.

Structure of Multiple Network Semigroups

In some cases, networks define semigroups with certain structural features.

An element s in a semigroup S possesses a (generalized) *inverse* if there exists some $t \in S$ such that $sts = s$ and $tst = t$. The element t is termed the inverse of s. If every $s \in S$ possesses an inverse, then S is termed an *inverse semigroup*. If, in addition, S possesses an *identity* element, that is, an element e such that for any $s \in S$, there exists some $t \in S$ such that $ts = e$, then S is a *group*. The element t is termed a *(group-)inverse* of s. An element s in a semigroup S is *regular* if there exists some $t \in S$ such that $sts = s$. The semigroup S is *regular* if each of its elements is regular.

It is clear from these definitions that any group is also an inverse semigroup and that any inverse semigroup is also regular. The results below describe some of the conditions under which mutiple networks give rise to regular semigroups, inverse semigroups and groups.

We first describe the conditions under which S is a group [28,42]. Let X be a network relation having a group-inverse Y, that is, a relation Y such that $XY' = I$ where I is the identity relation (for which $I(k, l) = 1$ if $k = l$; $I(k, l) = 0$ otherwise). Then it follows that (i) $XX' = I = X'X$ and $Y = X$ (that is, X' is a two-sided group-inverse for X and is unique); and (ii) X is a permutation relation (that is a relation in which, for each node k, there is exactly one node l for which $X(k, l) = 1$ and exactly one node j for which $X(j, k) = 1$). Hence if each relation X_i of a multiple network is a permutation relation and X_i' is included in S for each i, then the semigroup S is a group.

An important class of social relations which are constructed as permutation relations are "marriage class systems". These systems were proposed by White as models for kinship relations in certain Australian aboriginal tribes [48]. In such systems, persons are uniquely and permanently assigned to a *clan* and clans are related to one another by marriage and descent rules which can be represented as permutation relations. The marriage relation specifies the clan containing the wives of the male members of each clan, while the descent relation identifies the clan containing the children of male members of each clan. It follows that marriage class systems give rise to groups that represent the kinship structures of the societies concerned.

Another structure arising in the description of kinship systems is the inverse semigroup [9]. One can verify whether a network semigroup is an inverse semigroup by determining whether each of its elements has an inverse in S. The following is a characterization of conditions under which a relation possesses an inverse [24].

A network relation X has a *Thierrin–Vagner inverse* if there exists a relation Y satisfying $XYX = X$ and $YXY = Y$. Clearly, X has a Thierrin–Vagner inverse if and only if X possesses an inverse in S. If X is a regular element of S, so that $XYX = X$ for some Y in S, then YXY is a Thierrin–Vagner inverse for X. Thus determining whether X is regular will reveal whether X has an inverse. The regularity of X can be ascertained as follows.

The network relation Y is a *subinverse* of X if $XYX \leq X$. The set of subinverses of X is closed under union (that is, if Y and Z are subinverses of X then so is the relation $Y \cup Z$); hence there is a largest subinverse X^* for any relation X. The largest subinverse X^* of a relation X can be calculated from X according to $X^* = (X^T X^c X^T)^c$ where X^c denotes the complement of X. It therefore follows that a network X is regular if and only if $X \leq X(X^T X^c X^T)^c X$, and that the largest Thierrin–Vagner inverse of a regular relation X is $X^* X X^*$ [24,44].

The Algebraic Structure of Local Role Algebras

Less is known of the general algebraic properties of local role algebras. It is easily seen, however, that any orderings among relations present in the semigroup of a multiple network are also present in the local role algebra of each actor in the network. As a result, a number of the algebraic properties of local role algebras are inherited from the "parent" network semigroup. For example, if $s \leq t$ in the partially ordered semigroup S constructed from the composition operation, then $s \leq t$ in the local role algebra of every node in the network. Indeed, the converse is readily seen to also hold, so that $s \leq t$ in S if and only if $s \leq t$ in the local role algebra of every node in the network.

Algebraic Analysis

In order to describe an algebraic representation, general procedures for analyzing finite algebraic representations into simpler, independent components have been developed [20,36,37,39]. These procedures have been applied to the decomposition of lattice representations [20], partially ordered semigroups and local role algebras [34,35,36]. The procedures extract maximally independent components

of a given algebraic representation and lead to a detailed structural analysis.

The account of these procedures for decomposing a finite algebra into simpler components is based on definitions adapted from [2,18]. Clearly, any decomposition procedure depends on the synthesis rules by which the components are assumed to be combined to produce the algebra. Two important synthesis rules are the direct and subdirect product.

Let $A_1 = [S_1, F, \leq]$, $A_2 = [S_2, F, \leq]$, $\ldots, A_q = [S_q, F, \leq]$ be a collection of similar algebras. The *direct product* $A_1 \times A_2 \times \cdots \times A_q$ of $A_1, A_2, \ldots, A_q$ is the algebra comprising the set $S_1 \times S_2 \times \cdots \times S_q$ and the operations $f_\alpha \in F$ given by

$$f_\alpha([x_1, x_2, \ldots, x_q], \ldots, [z_1, z_2, \ldots, z_q]) = [f_\alpha(x_1, \ldots, z_1), \ldots, f_\alpha(x_q, \ldots, z_q)] .$$

The partial order for $A_1 \times A_2 \times \cdots \times A_q$ is given by

$$[x_1, x_2, \ldots, x_q] \leq [z_1, z_2, \ldots, z_q]$$

if and only if

$$x_1 \leq z_1, x_2 \leq z_2, \ldots, \text{ and } x_q \leq z_q .$$

An algebra $B = [T, F, \leq]$ is a *subalgebra* of $A = [S, F, \leq]$ if T is a subset of S (possibly empty) which is closed under the operations of F (that is, $f_\alpha(x_1, x_2, \ldots, x_{n(\alpha)}) \in T$, for any $x_1, x_2, \ldots, x_{n(\alpha)} \in T$, and any $f_\alpha \in F$); in addition, $x \leq y$ in B if and only if $x \leq y$ in A, for any $x, y \in T$. A subalgebra $C = [S, F, \leq]$ of a direct product $A_1 \times A_2 \times \cdots \times A_q$ of similar algebras $A_i = [S_i, F, \leq]$ $(i = 1, 2, \ldots, q)$ is a *subdirect product* of $A_1, A_2, \ldots, A_q$ if for each $x_i \in S_i$, there exists an element $c \in S$ having x_i as its component in A_i.

In the case of both the direct and subdirect product of algebras, the operations in the composite algebra are performed as the conjunction of their independent operation in the component algebras. The difference between the two constructions lies in the set on which these operations are defined. For the direct product, the appropriate set is the full Cartesian product of elements from each component algebra; in the subdirect product case, the appropriate set is simply a subset of the full Cartesian product with the additional condition that each element in each component algebra appears in *some* element of the subset.

Some well-known theorems in universal algebra establish that the existence and nature of direct and subdirect representations of an algebra are determined by the lattice of homomorphisms of the algebra, or, equivalently, by a lattice of relations on the algebra associated with its homomorphisms.

An (*isotone*) *homomorphism* from an algebra $A = [S, F, \leq]$ onto a similar algebra $B = [T, F, \leq]$ is a mapping $\phi : S \to T$ such that,

(i) for all $f_\alpha \in F$ and $x_i \in S$,

$$\phi(f_\alpha(x_1, x_2, \ldots, x_{n(\alpha)})) = f_\alpha(\phi(x_1), \phi(x_2), \ldots, \phi(x_{n(\alpha)})) ; \quad \text{and}$$

(ii) $x \leq y$ in S implies $\phi(x) \leq \phi(y)$ in T.

The algebra B is termed a (*homomorphic*) *image* of A, and we write $B = \phi(A)$.

Further, each homomorphism ϕ from $A = [S, F, \leq]$ onto $B = [T, F, \leq]$ has a corresponding binary relation π on S (termed here a *π-relation*) in which $(x, y) \in \pi$ if and only if $\phi(y) \leq \phi(x)$; it may readily be established that π is transitive and reflexive (and hence a quasi-order) and that the equivalence relation σ_π defined by $(x, y) \in \sigma_\pi$ if and only if $(x, y) \in \pi$ and $(y, x) \in \pi$ has the substitution property, namely:

$$(x_1, y_1), (x_2, y_2), \ldots, (x_{n(\alpha)}, y_{n(\alpha)}) \in \sigma_\pi$$

implies

$$(f_\alpha(x_1, x_2, \ldots, x_{n(\alpha)}), f_\alpha(y_1, y_2, \ldots, y_{n(\alpha)})) \in \sigma_\pi .$$

The relation σ_π is termed a *congruence relation*. The collection of π-relations may be partially ordered by: $\pi_1 \leq \pi_2$ if $(x, y) \in \pi_1$ implies $(x, y) \in \pi_2$, for any $x, y \in S$. Under this partial ordering, the relations π form a lattice $L_\pi(A)$, the *lattice of π-relations* on A.

The conditions under which an algebra A may be represented as a direct or subdirect product of similar algebras $A_1, A_2, \ldots, A_r$ (that is, possesses *direct* or *subdirect representations*) can be described as follows.

Let $X = \{\pi_1, \pi_2, \ldots, \pi_q\}$ be a set of π relations in $L_\pi(A)$ for which

$$\pi_1 \wedge \pi_2 \wedge \cdots \wedge \pi_q = \pi_{\min},$$

where $\pi_{\min}$ is the minimal element of $L_\pi(A)$. Then A may be represented as a subdirect product of the algebras $\phi_1(A), \phi_2(A), \ldots, \phi_q(A)$, where $\phi_i(A)$ is the image of A under the homomorphism ϕ_i corresponding to the relation π_i $(i = 1, 2, \ldots, q)$. If, in addition, for all i,

(a) $\pi_1 \wedge \pi_2 \wedge \cdots \wedge \pi_{i-1}$ and π_i are *permutable* (i. e., $(\pi_1 \wedge \pi_2 \wedge \cdots \wedge \pi_{i-1})\pi_i = \pi_i(\pi_1 \wedge \pi_2 \wedge \cdots \wedge \pi_{i-1})$); and

(b) $(\pi_1 \wedge \pi_2 \wedge \cdots \wedge \pi_{i-1}) \vee \pi_i = \pi_{\max}$, where $\pi_{\max}$ is the maximal element of $L_\pi(A)$, then A may be represented as the direct product of the algebras $\phi_1(A), \phi_2(A), \ldots, \phi_q(A)$ [2].

An algebra which possesses no nontrivial representation as the direct product of similar algebras is termed *directly irreducible*, while an algebra which cannot be expressed as a nontrivial subdirect product of similar algebras is *subdirectly irreducible*. Pattison and Bartlett advocated the use of subdirect representations, but sought to restrict attention to a small set of such representations rather than to the (possibly large) class of all subdirect representations of the algebra [37]; see also [7]. Their restriction entailed (a) identifying all *irredundant* subdirect representations of an algebra, and then (b) defining a partial ordering on the irredundant representations and selecting only minimal members of the resulting partially ordered set. The resulting set of minimal, irredundant subdirect representations were termed the factorizations of the algebra.

More formally, an element x of a lattice L is *meet-irreducible* if $x = x_1 \wedge x_2$ implies $x = x_1$ or $x = x_2$. A subset of lattice elements $X = \{x_1, x_2, \ldots, x_r\}$ is *irredundant* if

a) each x_i is meet-irreducible; and
b) $x_1 \wedge \cdots \wedge x_{i-1} \wedge x_{i+1} \wedge \cdots \wedge x_q, \neq x_{\min}$, for each $i = 1, 2, \ldots, q$; where $x_{\min}$ is the minimal element of the lattice L.

The collection of all irredundant subsets of a lattice L may be partially ordered by the relation:

$X \leq Y$ iff, for each $j = 1, 2, \ldots, p$, there exists some i $(i = 1, 2, \ldots, q)$ such that $y_j \leq x_i$ in L, where $X = \{x_1, x_2, \ldots, x_q\}$ and $Y = \{y_1, y_2, \ldots, y_p\}$.

A *factorization* of an algebra A is then the subdirect representation corresponding to any minimal, irredundant subset of elements of $L_\pi(A)$ whose meet is the minimal relation $\pi_{\min}$.

An Algorithm for Factorization

A general algorithm for constructing the set of all factorizations of an algebra has also been developed [37]. The algorithm operates on a subset of relations in the lattice $L_\pi(A)$ and conducts a reasonably efficient search for subsets of π relations which satisfy the conditions of the factorization definition. In particular, the algorithm constructs the set of factorizations of an algebra A from the atoms of its lattice $L_\pi(A)$ of π-relations and their maximal meet-complements.

An *atom* of a lattice L is an element that covers $x_{\min}$. A *meet-complement* of an element $x \in L$ is an element x^* such that $x^* > x_{\min}$ and $x \wedge x^* = x_{\min}$. A meet-complement x^* of x is *maximal* if x has no other meet-complement z such that $z > x^*$.

Let $Z = \{z_1, z_2, \ldots, z_a\}$ be the set of atoms of $L_\pi(A)$. Define the collection of sets of the form $\{z_1^*, z_2^*, \ldots, z_a^*\}$ where z_i^* is a maximal meet-complement of z_i. Then it can be shown that any factorization of A is a subset of such a set [36]. This result is the basis for the algorithm for finding factorizations. Under certain circumstances, we can write down an explicit expression for π-relations in the factorization of A.

Let A be an algebra with π-relation lattice $L_\pi(A)$. Let π_{ab} be the least π-relation in which $a \geq b$ for elements $a, b \in A$; we term π_{ab} the π-relation generated by the ordering $a \geq b$.

Let z be an atom of $L_\pi(A)$; define the relation $\pi(z) = \{(a, b) : \pi_{ab} \wedge z = \pi_{\min}\}$. If z has a unique maximal meet-complement, then $\pi(z)$ is a π-relation and is a unique maximal meet-complement of z. Conversely, if $\pi(z)$ is a π-relation, then it is the unique maximal meet-complement of z. Further, if $Z = \{z_1, z_2, \ldots, z_a\}$ is the set of atoms of $L_\pi(A)$ and if, for each $z_i \in Z$, z_i has a unique maximal meet-complement $\pi(z_i)$, then the factorization of A is unique and corresponds to the π-relations $\{\pi(z_1), \pi(z_2), \ldots, \pi(z_a)\}$ [36]. In other words, if the factorization is unique, it can be identified immediately from the maximal meet-complements of the atoms of $L_\pi(A)$.

Under some circumstances, we can make more general claims about the uniqueness of factorizations. Not surprisingly, the structure of the lattice $L_\pi(A)$ plays a major part. It is known, for instance, that if the lattice $L_\pi(A)$ is distributive, then A possesses a unique irredundant subdirect representation, and hence a unique factorization. In particular, since the lattice of π-relations for a lattice is necessarily distributive, any finite lattice has a unique factorization. If the lattice $L_\pi(A)$ is modular, then any irredundant subdirect representation of A has the same number of components [2].

Factorization for Partial Algebras

Analogous decompositions can be described for partial algebras. An (*isotone*) *homomorphism* from a partial algebra $A = [S, F, \leq]$ onto a similar partial algebra $B = [T, F, \leq]$ is a mapping $\phi: S \to T$ such that,

(i) for all $f_\alpha \in F$ and $x_i \in S$,

$$\phi(f_\alpha(x_1, x_2, \ldots, x_{n(\alpha)})) = f_\alpha(\phi(x_1), \phi(x_2), \ldots, \phi(x_{n(\alpha)}))$$

whenever the operations on both sides of the expression are defined; and

(ii) $x \leq y$ in S implies $\phi(x) \leq \phi(y)$ in T.

The algebra B is termed a (*homomorphic*) *image* of A, and we write $B = \phi(A)$.

Further, each homomorphism ϕ from $A = [S, F, \leq]$ onto $B = [T, F, \leq]$ has a corresponding binary relation π on S (termed here a π*-relation*) in which $(x, y) \in \pi$ if and only if $\phi(y) \leq \phi(x)$; it may readily be established that π is transitive and reflexive (so that π is a quasi-order). Further, the equivalence relation σ_π defined by $(x, y) \in \sigma_\pi$ if and only if $(x, y) \in \pi$ and $(y, x) \in \pi$ has the substitution property, namely:

$$(x_1, y_1), (x_2, y_2), \ldots, (x_{n(\alpha)}, y_{n(\alpha)}) \in \sigma_\pi$$

implies

$$(f_\alpha(x_1, x_2, \ldots, x_{n(\alpha)}),\ f_\alpha(y_1, y_2, \ldots, y_{n(\alpha)})) \in \sigma_\pi$$

whenever the latter term is defined. The relation σ_π is often termed a *congruence relation*. The collection of all π-relations form a lattice $L_\pi(A)$, the *lattice of* π*-relations* on A, under the partial ordering: $\pi_1 \leq \pi_2$ if $(x, y) \in \pi_1$ implies $(x, y) \in \pi_2$, for any $x, y \in S$.

In the case of partial algebras, it is convenient to express the algebra as an intersection of the π-relations corresponding to maximally independent homomorphic images of the algebra. Recall that the minimal element $\pi_{\min}$ of the lattice $L_\pi(A)$ corresponds to the algebra A itself. Thus, any expression of $\pi_{\min}$ in the form

$$\pi_1 \wedge \pi_2 \wedge \cdots \wedge \pi_q = \pi_{\min}$$

expresses A as the intersection of π-relations corresponding to homomorphic images of A. Any such expression that involves π-relations $\{\pi_i;\ i = 1, 2, \ldots, q\}$ satisfying the conditions of the factorization definition yields an expression in terms of a set of simple and maximally independent images of the algebra. Further, the same algorithm for constructing factorizations of an algebra that was described earlier may be used to find expressions of this type.

Algebraic and Network Mappings

The semigroup of a multiple network records the orderings and equations among different types of labelled paths in the network. A natural question to arise is how properties of a multiple network constrain the relational structure of its semigroup. We therefore review some general conditions under which the semigroup of one network is a homomorphic image of the semigroup of another.

We define two multiple networks to possess the same structure if their semigroups are isomorphic, that is, if the network relations are in one-to-one correspondence and they have identical right multiplication tables and partial orders.

One strategy for investigating whether networks have similar structures is to determine the conditions under which network homomorphisms induce homomorphisms of the network semigroup [4]. A *network homomorphism* from a multiple network $\{X_1, X_2, \ldots, X_r\}$ on actor set N to a multiple network $\{Y_1, Y_2, \ldots, Y_r\}$ on actor set M is a mapping ψ from N onto M such that: (a) $X_m(k, l) = 1$ implies $Y_m(\psi(k), \psi(l)) = 1$, for any k, l, m; and (b) $Y_m(i, j) = 1$ for some i, j implies that $X_m(k, l) = 1$, for some k, l such that $\psi(k) = i$ and $\psi(l) = j$. The network on M is termed the *image* of the network on N under the mapping ψ.

The mapping ψ satisfies the *structural equivalence* condition if for any m, and for any $k, l \in N$, $\psi(k) = \psi(l)$ if and only if:

- $X_m(k, j) = 1$ iff $X_m(l, j) = 1$ for any $j \in N$; and
- $X_m(j, k) = 1$ iff $X_m(j, l) = 1$ for any $j \in N$.

Lorrain and White observed that if two multiple networks are related by a network homomorphism satisfying the structural equivalence condition, then their semigroups are isomorphic and we may argue that they possess the *same* relational structure [27]. More generally, we can ask whether a homomorphism between two networks induces a homomorphism between their semigroups. It is readily seen that a homomorphism is not always guaranteed.

Network homomorphisms induced by certain blockmodels lead to semigroup homomorphisms (see the entry on ► Positional Analysis and Blockmodeling). More general conditions under which a semigroup homomorhism is guaranteed were established by Kim and Roush [25].

A network homomorphism ψ satisfies Kim and Roush's *condition* G_i if the following holds for any pair of equivalence classes ρ_1 and ρ_2 on N induced by the mapping ψ (so that $k, l \in \rho_h$ for some h iff $\psi(k) = \psi(l)$). Let the number of elements in ρ_1 and ρ_2 be n_1 and n_2, respectively. Let D be any subset of ρ_1 of i elements or, if $i > n_1$, let $D = \rho_1$. Then, for any $m = 1, 2, \ldots, r$, the set $\{l : l \in \rho_2$ and $X_m(k, l) = 1$ for some $k \in D\}$ has at least $\min(i, n_2)$ elements. The condition G_1 is also known as the *outdegree condition* and G_n is also termed the *indegree condition*. A network homomorphism that satisfies both G_1 and G_n is termed *regular*.

Kim and Roush demonstrated that if ψ is a network homomorphism from one multiple network onto another that satisfies the condition G_i, then there is a homomor-

phism mapping the semigroup of the first onto the semigroup of the second.

A more general condition combines the condition G_i with what Pattison termed the *central representatives condition* [34]. Let ψ be a network homomorphism. Then ψ satisfies the condition G_{im} if, for each class ρ of elements of N induced by ψ,

- there exists a *central subset* C of ρ such that for any X_m
  - $X_m(k, l) = 1$ for some $k \in \rho$ implies $X_m(k^*, l) = 1$ for some $k^* \in C$, and
  - $X_m(l, k) = 1$ for some $k \in \rho$ implies $X_m(l, k^*) = 1$ for some $k^* \in C$; and
- if C^* denotes the union of central subsets C, then the central subsets C satisfy Kim and Roush's condition G_i on the network defined on C^*.

If each central subset C comprises a single actor, then the condition is equivalent to Pattison's central representatives condition, while if each central subset C comprises the whole of the equivalence class on N induced by ψ, it is equivalent to the condition G_i. Kim and Roush showed that if one network can be mapped onto another by a network homomorphism satisfying the condition G_{im}, then there is a homomorphism from the semigroup of the first to the semigroup of the second [25]. The condition G_{im} is the most general condition known that guarantees the existence of such a homomorphism.

Comparing Network Semigroups

Two network semigroups S_1 and S_2 are strictly comparable only if $S_1 \leq S_2$ or $S_2 \leq S_1$. Loosely speaking, though, they are "similar" if they share many homomorphic images. Boorman and White proposed that the largest shared homomorphic image (the so-called *joint homomorphism*) is a useful construction for comparing network semigroups [5]. Bonacich and McConaghy, on the other hand, argued that the smallest semigroup containing each of S_1 and S_2 (the *common structure semigroup*) was a more appropriate representative of common semigroup structure [3,32]. The resolution of the problem of finding a representative of common structure depends on how a semigroup is viewed [36]. If a semigroup is seen as a list of the homomorphic images that it admits, then the joint homomorphism corresponds to the list of shared features, while if it seen as a collection of semigroup equations and orderings, then the common structure semigroup serves as a representation of shared structure. The common structure semigroup of two network semigroups is the semigroup of the disjoint union of the underlying networks.

Both the joint homomorphic image and common structure semigroup constructions have been shown to be useful in particular applications, the former in identifying common reductions of the networks giving rise to identical semigroups, and the latter in identifying shared cultural forms (such as "the friend of a friend is always an associate").

Linking Algebraic and Network Homomorphisms

The factorization of an algebra identifies relatively independent components of the algebra generated by constituents of the data array under the selected operation. It is natural to ask whether this analysis induces a corresponding decomposition of the network itself. That is, if an algebra A has a homomorphic image B (such as one of the factors of A), is it possible to find one or more homomorphisms of the network that is consistent with the congruence relation σ for the homomorphism and that generates the algebra B? If so, we can argue that there is an association between such a reduction of the network and the homomorphic image.

In the case of the lattice of an affiliation network, such a reduction can always be uniquely found [20]. For semigroups and role algebras, there is no guarantee that such a reduction can be found, but it is possible to identify the smallest homomorphic image of the network whose algebra contained B as a homomorphic image [34].

Examples

We illustrate a number of the algebraic constructions just described for a multiple network blockmodel and an approximation to an affiliation network.

Network Semigroup

Table 1 corresponds to a multiple network on 7 nodes with two kinds of edges, F and N. The network is a blockmodel reported by Vickers for relations of friendship and negative ties among members of a high school class in an Australian country town [46]. The questions from which the blockmodel was constructed were (a) "Who are your best friends?" and (b) "Who would you rather not have as a friend?".

The semigroup generated by the multiple network in Table 1 has the right multiplication table and partial order shown in Table 2. The process of constructing the semigroup generates the sets $W_1 = \{F, N\}$, $W_2 = \{F^2, FN, NF, N^2\}$, $W_3 = \{F^3, FNF, FN^2, NF^2, N^2F, N^3\}$, $W_4 =$

Social Networks, Algebraic Models for, Table 1
A multiple network on 7 nodes

		Block									Block						
Relation	Block	1	2	3	4	5	6	7	Relation	Block	1	2	3	4	5	6	7
F	1	1	1	1	0	0	0	0	*N*	1	0	0	0	1	1	1	1
	2	0	1	0	0	0	0	1		2	0	0	0	1	1	1	0
	3	1	1	1	0	0	0	0		3	0	0	0	1	1	1	0
	4	0	0	0	1	0	0	0		4	0	1	1	0	0	1	1
	5	0	0	0	0	1	0	0		5	0	0	0	0	0	0	1
	6	1	0	0	0	0	1	0		6	0	0	0	1	0	0	1
	7	1	1	1	0	0	0	1		7	0	0	0	1	1	1	1

Social Networks, Algebraic Models for, Table 2
Right multiplication table and partial order for the semigroup of the multiple network of Table 1

Right multiplication table				Partial order														
		Generators			Elements													
Element	Word	*F*	*N*		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	*F*	3	4	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	*N*	5	6	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0
3	*FF*	7	4	3	1	0	1	0	0	0	0	0	0	0	0	0	0	0
4	*FN*	8	9	4	0	1	0	1	0	0	0	0	0	0	0	0	0	0
5	*NF*	10	9	5	0	1	0	0	1	0	0	0	0	0	0	0	0	0
6	*NN*	11	12	6	0	0	0	0	0	1	0	0	0	0	0	0	0	0
7	*FFF*	7	4	7	1	0	1	0	0	0	1	0	0	0	0	0	0	0
8	*FNF*	8	9	8	0	1	0	1	1	0	0	1	0	1	0	0	0	1
9	*FNN*	13	12	9	0	0	0	0	0	1	0	0	1	0	0	0	0	0
10	*NFF*	14	9	10	0	1	0	0	1	0	0	0	0	1	0	0	0	0
11	*NNF*	11	12	11	1	0	1	0	0	1	1	0	0	0	1	0	0	0
12	*NNN*	13	12	12	0	1	0	1	0	1	0	0	1	0	0	1	0	0
13	*FNNF*	13	12	13	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	*NFFF*	14	9	14	0	1	0	0	1	0	0	0	0	1	0	0	0	1

$\{FN^2F, NF^3\}$ and $W_5 = \emptyset$; equations include $F^2N = FN, NFN = FN^2$, and $F^4 = F^3$.

It is evident from Table 2 that the relations F and N both have index 3 and period 1. The relation F is regular, whereas the relation N is not.

Each of the equations of Table 2 specifies the empirical coincidence of potentially distinct labelled walks. The equation $F^2N = FN$, for example, indicates the those individuals nominated by one's friends as preferred non-friends coincide with individuals so nominated by the friends of one's friends.

Role Algebra

The role algebra for block 4 in the network of Table 1 is presented in Table 3. Equations in the right multiplication table indicate equalities among labelled paths having block 4 as the source. For example, the equation $FF = F$ indicates that friendship paths of length 2 emanating from block 4 reach exactly the same set of blocks as direct friendship ties. Likewise, the equation $FN = N$ allows us to infer that, namely block 4 would prefer not to have as friends precisely those blocks so nominated by their friends. The partial order diagram indicates that block 4's friendship ties are a subset of those to whom block 4 has negative tie paths (block 4) of length 2.

The role algebra of block 4 has a unique factorization. The lattice of π-relations has two atoms and each atom possesses a unique maximal meet-complement; the atoms and their maximal meet-complements are shown in Table 4. These unique maximal meet-complements are associated with the role algebras shown in Table 5. In the first factor, both *FN* and *NF* ties from block 4 coincide with *N* ties, and *NNF* ties include all others. In the second factor, any path from block 4 whose last tie is labelled *N* coincides with an *N* tie.

Social Networks, Algebraic Models for, Table 3
The role algebra for block 4 in the network of Table 1

Right multiplication table				Partial order						
		Generators			Elements					
Element	Word	*F*	*N*		1	2	3	4	5	6
1	*F*	1	2	1	1	0	0	0	0	0
2	*N*	3	4	2	0	1	0	0	0	0
3	*NF*	3	4	3	0	1	1	0	0	0
4	*NN*	5	6	4	1	0	0	1	0	0
5	*NNF*	5	6	5	1	1	1	1	1	1
6	*NNN*	5	6	6	1	1	0	1	0	1

Social Networks, Algebraic Models for, Table 4
Atoms and unique maximal-meet complements for the block 4 role algebra

	Partial order							Partial order						
		Elements							Elements					
	Elements	1	2	3	4	5	6	Elements	1	2	3	4	5	6
Atom	1	1	0	0	0	0	0	1	1	0	0	0	0	0
	2	0	1	0	0	0	0	2	0	1	0	0	0	0
	3	0	1	1	0	0	0	3	0	1	1	0	0	0
	4	1	1	0	1	0	0	4	1	0	0	1	0	0
	5	1	1	1	1	1	1	5	1	1	1	1	1	1
	6	1	1	0	1	0	1	6	1	1	1	1	0	1
Maximal meet-complement	1	1	0	0	0	0	0	1	1	0	0	0	0	0
	2	0	1	1	0	0	0	2	1	1	0	1	0	1
	3	0	1	1	0	0	0	3	1	1	1	1	1	1
	4	1	0	0	1	0	0	4	1	1	0	1	0	1
	5	1	1	1	1	1	1	5	1	1	1	1	1	1
	6	1	1	1	1	1	1	6	1	1	0	1	0	1

Lattice of an Affiliation Network

The affiliation network in Table 6 is an approximation to the so-called "Southern Women" data [12,23]. The approximation was obtained by a method described in [52], and is akin to a form of dual clustering of the rows and columns of the original data array. The Galois lattice of the (approximate) affiliation array is presented in Fig. 1. In Fig. 1, known as the *line diagram* of the lattice, one element s is drawn above and connected to another element t if s covers t, (that is if $s \geq t$ and there is no element u distinct from s and t for which $s \geq u \geq t$) [51]. The meet of two elements is the greatest element lying below both elements and connected to them by a descending path. The join of two elements is the least element to which they are both connected by an ascending path.

Social Networks, Algebraic Models for, Table 5
Factors of the role algebra for block 4

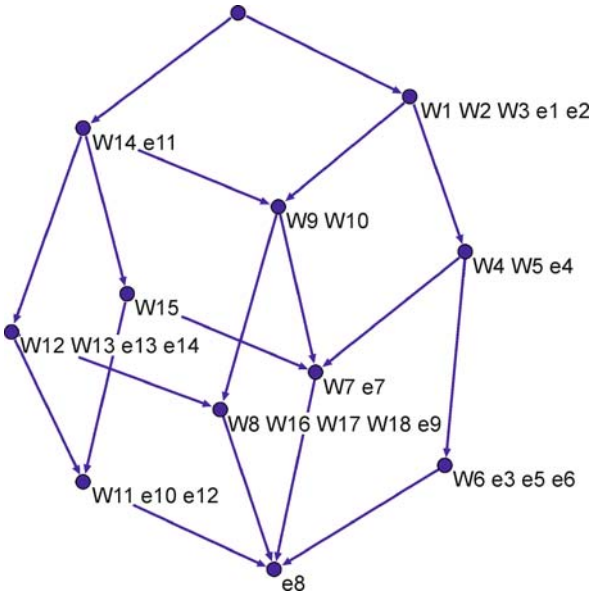
Factor 1	Right multiplication table			Partial order				
		Generators			Elements			
Element	Word	*F*	*N*		1	2	3	4
1	*F*	1	2	1	1	0	0	0
2	*N*	2	3	2	0	1	0	0
3	*NN*	4	4	3	1	0	1	0
4	*NNF*	4	4	4	1	1	1	1

Factor 2	Right multiplication table			Partial order			
		Generators			Elements		
Element	Word	*F*	*N*		1	2	3
1	*F*	1	2	1	1	0	0
2	*N*	3	2	2	0	1	0
3	*NF*	3	2	3	0	1	1

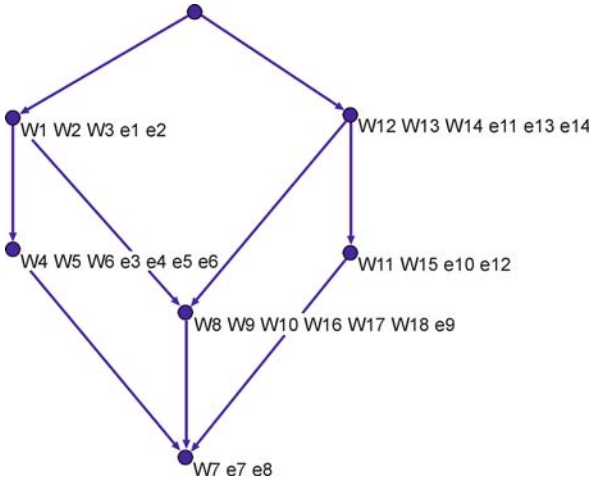
Social Networks, Algebraic Models for, Table 6
An affiliation network: an approximation to the Southern Women data; the approximation differs from the original data in the 20 underlined values

Event														
Woman	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1	1	1	1	1	1	1	1	1	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	0	0	0	0	0
3	1	1	1	1	1	1	1	1	1	0	0	0	0	0
4	0	0	1	1	1	1	1	1	0	0	0	0	0	0
5	0	0	1	1	1	1	1	1	0	0	0	0	0	0
6	0	0	1	0	1	1	0	1	0	0	0	0	0	0
7	0	0	0	0	0	0	1	1	0	0	0	0	0	0
8	0	0	0	0	0	0	0	1	1	0	0	0	0	0
9	0	0	0	0	0	0	1	1	1	0	0	0	0	0
10	0	0	0	0	0	0	1	1	1	0	0	0	0	0
11	0	0	0	0	0	0	0	1	0	1	0	1	0	0
12	0	0	0	0	0	0	0	1	1	1	0	1	1	1
13	0	0	0	0	0	0	0	1	1	1	0	1	1	1
14	0	0	0	0	0	0	1	1	1	1	1	1	1	1
15	0	0	0	0	0	0	1	1	0	1	1	1	0	0
16	0	0	0	0	0	0	0	1	1	0	0	0	0	0
17	0	0	0	0	0	0	0	1	1	0	0	0	0	0
18	0	0	0	0	0	0	0	1	1	0	0	0	0	0

The rows and columns of the affiliation array may be mapped to lattice elements in such a way that any row element (in this case, a woman) is at or above any column elements (in this case, events) with which it is affiliated; likewise, any column element (event) is at or below any affiliated row elements (women). It can be seen from Fig. 1, for example, that all women attend event e8, whereas only women W1, W2 and W3 attend events e1 and e2.



Social Networks, Algebraic Models for, Figure 1
The line diagram of the Galois lattice of Table 4

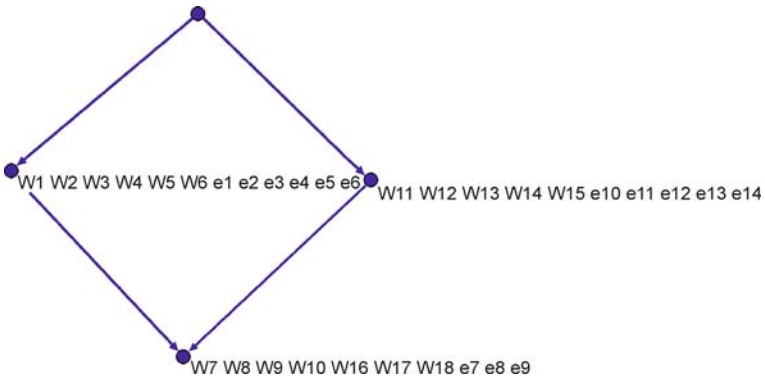


Social Networks, Algebraic Models for, Figure 2
Largest homomorphic image of the lattice of Fig. 1

The lattice displayed in Fig. 2 is a maximal homomorphic image of the lattice of Fig. 1 and provides a simplification of the lattice. In this diagram, a division of the women in the network into three ordered clusters is apparent. One cluster comprises the women W1 to W6, with women W4, W5 and W6 attending only a subset of the events attended by women W1, W2 and W3. A second cluster comprises women W11 to W15, with women W11 and W15 attending only a subset of the events attended by women W12, W13 and W14. The third cluster of women comprises W7, W8, W16, W17 and W18 and is again divided into two sub-clusters, one of whom (W7) attends a subset of the events attended by the other (W8, W16, W17, W18). It is also clear from the lattice of Fig. 2 that the first two clusters of women attend a mixture of distinct events (e1 to e6 in the case of the first cluster, and e10 to e14 in the case of the second) as well as common events (e7 to e9), whereas the third cluster of women attend only those events also attended by women in the first two clusters (e7 to e9). This structural feature of the affiliation network is drawn out very clearly in the unique maximal homomorphic image of the lattice of Fig. 2, shown in Fig. 3.

Future Directions

The algebraic constructions described above provide exact and detailed representations of structural relationships



Social Networks, Algebraic Models for, Figure 3
Largest homomorphic image of the lattice of Fig. 2

among network constituents. Of course, this level of detail assumes that the relational observations comprising the network are accurate, as no allowance is made for variability or error. This is often a tenuous assumption, and two important directions for further development utilize the structural sensitivities of algebraic representations in the presence of network tie variability.

The first direction is to use statistical criteria to generate (partial) algebraic structures that summarize key structural features of a network [1,8,39,40]. Instead of the generalized Axiom of Quality introduced earlier, an *Approximate Axiom of Quality* is proposed in which $u \leq v$ in W whenever there is "sufficient evidence" that the relation $X_u \leq X_v$ holds. Such an approach can lead to a theoretically-guided and structurally-focused form of exploratory data analysis for multiple networks. Theoretical guidance comes from choice of operations in the set F, while structural focus resides in the (partial) algebra to which the approach gives rise.

The second direction is to understand how exact algebraic representations can emerge as special, so-called *degenerate*, cases of statistical models such as exponential random graph models for networks (see the entry on ► Social Networks, Exponential Random Graph (p^*) Models for *for Networks*). An exponential random graph model defines a probability distribution on the set of all networks on a node set N as a function of some parameter vector. For any parameter vector, a subset of networks have what can be defined as minimum "energy". As the parameter vector is scaled by an increasingly large multiplier, the probability associated with any network whose energy exceeds the minimum tends to zero. The minimum energy networks can therefore be seen as highly constrained or "frozen" structural forms associated with the stochastic model [21,22,41]. The structure of these forms will often warrant algebraic analysis of the type described earlier; in addition, we can understand the non-frozen models as stochastic generalizations of these structural forms.

Bibliography

1. Barnes G, Cerrito P, Levi I (1998) A mathematical model for interpersonal relationships in social networks. Soc Netw 20:179–196
2. Birkhoff G (1963) Lattice theory, 3rd edn. American Mathematical Society, Providence
3. Bonacich P (1980) The common structure semigroup, a replacement for the Boorman and White joint reduction. Am J Soc 86:159–166
4. Bonacich P (1983) Representations for homomorphisms. Soc Netw 5:173–192
5. Boorman SA, White HC (1976) Social structures from multiple networks II. Role structures. Am J Soc 81:1384–1446
6. Boyd JP (1969) The algebra of group kinship. J Math Psychol 6:139–167
7. Boyd JP (1989) Social semigroups and Green relations. In: Freeman LC, White DR, Romney AK (eds) Research methods in social network analysis. George Mason University Press, Fairfax, pp 215–254
8. Boyd JP (1990) Social semigroups: A unified theory of scaling and blockmodelling as applied to social networks. George Mason University Press, Fairfax
9. Boyd JP, Haehl JH, Sailer LD (1989) Kinship systems and inverse semigroups. J Math Soc 2:37–61
10. Breiger RL (1974) The duality of persons and groups. Soc Forces 53:181–190
11. Breiger RL, Pattison PE (1986) Cumulated social roles: the duality of persons and their algebras. Soc Netw 8:215–256
12. Davis A, Gardner B, Gardner M (1941) Deep south. Chicago University Press, Chicago
13. Doreian P (1981) Polyhedral dynamics and conflict mobilization in social networks. Soc Netw 3:107–116
14. Duquenne V (1996) On lattice approximations: syntactic aspects. Soc Netw 18:189–200
15. Freeman LC (1996) Cliques, Galois lattices, and the structure of human social groups. Soc Netw 18:173–87
16. Freeman LC, White DR (1993) Using Galois lattices to represent network data. In: Marsden P (ed) Sociological methodology. American Sociological Association, Washington, pp 127–146
17. Friedell M (1967) Organizations as semilattices. Am Soc Rev 32:46–54
18. Fuchs L (1963) Partially ordered algebraic systems. Pergamon, Oxford
19. Ganter B, Wille R (1989) Conceptual scaling. In: Roberts F (ed) Applications of combinatorics and graph theory in the biological and social sciences. Springer, New York, pp 139–167
20. Ganter B, Wille R (1999) Formal concept analysis: mathematical foundations. Springer, New York
21. Grenander U (1993) General pattern theory. Oxford University Press, Oxford
22. Handcock MS (2003) Statistical models for social networks. In: Breiger RL, Carley KM, Pattison PE (eds) Dynamic social network modeling and analysis. National Academies Press, Washington
23. Homans G (1951) The human group. Routledge & Kegan Paul, London
24. Kim KH (1982) Boolean matrix theory and applications. Dekker, New York
25. Kim KH, Roush FW (1984) Group relationships and homomorphisms of Boolean matrix semigroups. J Math Psychol 28:448–452
26. Lorrain F (1975) Reseaux sociaux et classifications sociales. Hermann, Paris
27. Lorrain F, White HC (1971) Structural equivalence of individuals in social networks. J Math Soc 1:49–80
28. Luce RD (1956) A note on Boolean matrix theory. Proc Am Math Soc 3:382–388
29. Mandel M (1983) Local roles and social networks. Am Soc Rev 48:376–386
30. Martin JL (2002) Some algebraic structures for diffusion in social networks. J Math Soc 26:123–146

31. Martin JL (2006) Jointness and duality in algebraic approaches to dichotomous data. Soc Method Res 35:159–192
32. McConaghy M (1981) The common role structure: improved blockmodelling methods applied to two communities' elites. Soc Method Res 9:267–285
33. Mische A, Pattison PE (2000) Composing a civic arena: Publics, projects and social settings. Poetics 27:163–194
34. Pattison PE (1982) The analysis of semigroups of multirelational systems. J Math Psychol 25:87–118
35. Pattison PE (1989) Mathematical models for local social networks. In: Keats J, Taft R, Heath R, Lovibond S (eds) Mathematical and theoretical systems. North Holland, Amsterdam
36. Pattison PE (1993) Algebraic models for social networks. Cambridge University Press, New York
37. Pattison PE, Bartlett WK (1982) A factorization procedure for finite algebras. J Math Psychol 25:51–81
38. Pattison PE, Breiger RL (2002) Lattices and dimensional representations: matrix decompositions and ordering structures. Soc Netw 24:423–444
39. Pattison PE, Wasserman S (1995) Constructing algebraic models for local social networks using statistical methods. J Math Psychol 39:57–72
40. Pattison PE, Wasserman S, Robins G, Kanfer A (2000) Statistical evaluation of algebraic constraints for social networks. J Math Psychol 44:536–568
41. Robins G, Woolcock J, Pattison P (2005) Small and other worlds: Global network structures from local processes. Am J Sociol 110:894–936
42. Rutherford D (1963) Inverses of Boolean matrices. Proc Glasgow Math Assoc 6:49–53
43. Schein B (1970) A construction for idempotent binary relations. Proc Jpn Acad 46:246–247
44. Schein B (1976) Regular elements of the semigroup of all binary relations. Semigr Forum 13:95–102
45. Schweizer T (1993) The dual ordering of actors and possessions. Curr Anthropol 34:469–483
46. Vickers M (1981) Relational analysis: an applied evaluation. MSc Thesis, University of Melbourne, Melbourne
47. White DR, Duquenne V (1996) Special issue on social networks and discrete structure analysis. Soc Netw 18:169–318
48. White HC (1963) An anatomy of kinship. Prentice-Hall, Englewood Cliffs
49. White HC (1992) Identity and control. University of Chicago Press, Chicago
50. White HC, Boorman SA, Breiger RL (1976) Social structure from multiple networks: I. Blockmodels of roles and positions. Am J Sociol 81:730–780
51. Wille R (1984) Line diagrams of hierarchical concept systems. Intern Classif 11:77–86
52. Wilson S, Bladin P, Saling M, Pattison P (2005) Characterizing psychosocial outcome trajectories following seizure surgery. Epilepsy Behav 6:570–580
53. Winship C, Mandel M (1983) Roles and positions: a critique and extension of the blockmodelling approach. In: Leinhardt S (ed) Sociological methodology. Jossey-Bass, San Francisco, pp 314–344
54. Wu L (1983) Local blockmodel algebras for analyzing social networks. In: Leinhardt S (ed) Sociological methodology. Jossey-Bass, San Francisco, pp 272–313

Social Networks, Diffusion Processes in

THOMAS W. VALENTE
Department of Preventive Medicine, School of Medicine, University of Southern California, Alhambra, USA

Article Outline

Glossary
Definition of the Subject
Introduction
Network Models
Empirical Studies
Agent Based Models
Conclusions
Future Directions
Acknowledgments
Bibliography

Glossary

Adoption A person's change in behavior.

Agent based models Creation of hypothetical (sometimes prototypical) network structures and the simulation of diffusion within those structures.

Homophily Tendency for people to be connected to others like themselves.

Incidence The percent of new adopters at each time period.

Internal versus external influence Internal influence posits that adoption is driven by person-to-person persuasion whereas external influence posits it is driven by sources outside the network such as mass media.

Event history analysis Transformation of data to represent person-time observations.

Network exposure The degree of behavioral adoption in each person's network neighborhood.

Network threshold The number or percent of adopters in a person's neighborhood necessary for a person to adopt the innovation.

Rate The speed of diffusion.

Prevalence The cumulative percent of adopters in the population.

Weight matrix Any N×N matrix representing potential distances or similarities that models potential pathways for adoption influence (e. g., a structural equivalence matrix derived from an adjacency matrix to model influence of structural equivalence relations on adoption).

Definition of the Subject

Diffusion of innovations is the study of how new ideas and behaviors spread within a population, both within and between communities. Diffusion research spans many disciplines since understanding how new ideas and behaviors spread is an important aspect in many different fields of study. For example, marketers are often interested in how and why people change their purchasing habits and sociologists concerned with how people come to accept new norms. The first empirical investigations date back 100 years, while contemporary studies mark the beginning to the 1943 study of diffusion of hybrid seed corn in Iowa. A subset of the hundreds of empirical diffusion studies have focused specifically on the collection and analysis of social network data combined with data on when individuals adopted selected behaviors.

Introduction

Diffusion of innovations theory attempts to explain how new ideas and practices spread within and between communities. The theory has its roots in anthropology, economics, geography, sociology, marketing, mathematics, among other disciplines [1,2,3,4], and has in some ways been adapted from epidemiology [5,6]. The premise, confirmed by empirical research, is that new ideas and practices spread through interpersonal contacts largely consisting of interpersonal communication [1,7,8,9,10,11].

In their 1943 pioneering study, Ryan and Gross [10] laid the groundwork for the diffusion paradigm by showing that, among other things, social factors rather than economic ones were important influences on adoption [11]. Hundreds of diffusion studies were conducted in the 1950s and early 1960s to examine the diffusion process in more detail across a variety of settings [1]. Many studies sought to understand how information created in government or otherwise sponsored programs could be disseminated more effectively. Diffusion research peaked in the early 1960s, but has been reinvigorated recently with the advent of more sophisticated network models and technology making it possible to study the diffusion process more explicitly.

Most diffusion studies focus on trying to understand the factors that lead some members of a population to adopt a new idea, while others do not. Further, studies try to understand why some people adopt the behavior early while others wait a substantial amount of time before accepting the new practice. For example, Ryan and Gross [10] wanted to know why some farmers purchased hybrid seed corn almost immediately upon its availability while others waited until almost all the farmers in the area purchased it before they were willing to do so. Similarly, Coleman and others [12] wanted to know why some physicians began prescribing tetracycline as soon as it was available, while others waited until most physicians prescribed it before they were willing to do so.

This chapter describes a variety of mathematical and network models used to study the diffusion of these and other innovations. The Coleman and others [12] study provided a conceptual leap from other diffusion studies by explicitly measuring who talked to whom within the community about the innovation. Rogers [13] also collected network data to study the diffusion of weed spray in Iowa in his dissertation. Burt [14] unearthed the data and made it available to the network community so that scholars could debate various models used to describe the network diffusion process. Although having data has been useful for clarifying diffusion models, the limitations of these data and this study, make it a poor choice for studying adoption behavior. Rather, scholars should have focused on collecting better data, or reanalyzing diffusion network data in which contagion was more likely.

This chapter chronicles the development of network diffusion models and indicates where such progress is being made. I first present macro models used to estimate the speed of diffusion and, with the Bass [15] model, to estimate rates of innovation and imitation. Next, spatial autocorrelation is presented which is used to estimate the degree contiguous nodes adopt innovations. Spatial autocorrelation led to the network autocorrelation model which is presented statically (cross-sectional data only) and then with one time lag. I then discuss event history analysis applications of network autocorrelation and its extension by including time-based network interaction terms. Throughout the chapter, I attempt to provide a review of recent research conducted in a variety of domains, but mostly drawn from the public health field.

Macro Models

One consistent finding of diffusion research has been that the cumulative pattern of diffusion follows a growth pattern approximated by a simple one-parameter logistic function such as:

$$y_t = b_0 + \frac{1}{1 + e^{-b_1 t}}, \tag{1}$$

where y is the proportion of adopters, b_0 the y intercept, t is time, and b_1 the rate parameter to be estimated. This simple model can be used to compare growth rates for various innovations, but is extremely limited in its applicability. A considerable improvement was advanced by Bass [15]

Social Networks, Diffusion Processes in, Table 1
Diffusion Rate Parameter Estimates and Moran's *I* Estimates for Two Data Sets

	Medical Innovation	Cameroon Tontine 1 Simulation
One Parameter Model		
Coefficient (95% Confidence Intervals)	0.23 (–.053–0.51)	0.06 (.01–0.12)
N	17	50
R^2	0.76	0.71
Two Parameter (Bass) Model		
Innovation Coefficient (95% C. I.)	–0.43(–0.83–0.03)	–0.20 (–0.30–0.09)
Imitation Coefficient (95% C. I.)	4.09 (3.05–5.12)	2.96 (2.58–3.34)
N	16	49
R^2	0.89	0.89
Moran's *I*	–.13	–.08
z-score	–6.73	–7.80

and many others [16,17,18] by creating a two parameter model:

$$y_t = b_0 + (b_1 - b_0)Y_{t-1} - b_1(Y_{t-1})^2 \,, \tag{2}$$

where y is the proportion of adopters, b_0 a rate parameter for innovation, and b_1 a rate parameter for imitation (the degree of adoption due to prior adopters). The Bass model incorporates the percentage adopters at each time point and thus makes a better estimate of the growth attributable to personal network persuasion. The mathematical model in Eq. 2 can be used to: (1) forecast expected levels of diffusion [17], (2) estimate the rate of diffusion attributed to different theoretical aspects of the diffusion processes, b_0, external influence or innovativeness, and b_1, internal influence or interpersonal persuasion [15,16,18]. This model can be used to estimate rate of disease spread from a central source such as contaminated food or from infectious spread through interpersonal contact. In the social realm, one can use the model to estimate rate of adoption from a mass media advertisement or from interpersonal influence. Rate parameter estimates from both models for two diffusion datasets are provided in Table 1. Interpretation of these estimates is highly dependent on the time scale used to measure diffusion.

These rate parameter estimates can be used as outcomes to study factors associated with diffusion at the macro-level by comparing rates between groups and/or populations. For example, parameter estimates for different countries can be compared to study factors associated with the spread of behaviors in different countries. Modeling at this macro-level, however, is imprecise at best because it assumes perfect social mixing, everyone interacting with everyone else [19,20]. These macro models do not measure whether people who are connected to one another engage in the same behaviors. Geographers have devoted considerable attention to trying to determine whether innovations spread between contiguous areas.

Spatial Autocorrelation

Rather than just estimate rate of diffusion, spatial models measure whether artifacts, diseases, farming practices, and other behaviors spread between contiguous areas [2,21,22]. Proximity data are easy to obtain and are relatively unambiguous, thus providing a network of connections based on distance. Moran's I was an early model developed to test for spatial association, geographic clustering of adoption:

$$I = \frac{N\sum_{i}^{N}\sum_{j}^{N} D_{ij}(y_i - \bar{y})(y_j - \bar{y})}{S\sum_{i}^{N}(y_i - \bar{y})^2} \,, \tag{3}$$

where N is the sample size, D a distance matrix (as proximities), y indicates adoption, and S the sum of the distances in the distance matrix. Moran's I measures the degree nodes that are connected to one another deviate from the average behavior in the network similarly or differently. Moran's I is high when connected nodes (positive elements of D) are either positively or negatively different from the average score. The statistical significance of Moran's I can be calculated in two ways: via permutation methods or analytically.

To use a permutation method to calculate the significance of Moran's I assume adoption (y_i) is randomly distributed and calculate I repeatedly to get a sample of estimates based on D and the number of adopters. If Moran's I calculated is significantly different than the random sample generated, Moran's I is considered significant (z-scores can be obtained). The logic then is to calculate the degree neighbors (however defined) have similar adoption behavior compared to that expected if adoption were distributed randomly. Variance estimators for Moran's I can be found in spatial statistics textbooks [21,23] and used to calculate exact significance tests. Moran's I is useful, and has been extended considerably, yet this approach often assumes that geographic proximity equates with communication and influence, which may not be true.

The spatial autocorrelation methodology was seen as a useful approach to measuring network autocorrelation,

the bias inherent in a regression model when "y" appears as both the dependent and independent variable. Erbing and Young [24] wrote an influential paper on measuring network effects and using network autocorrelation methods. Dow [25,26] demonstrated the effects of network autocorrelation on estimate errors, and Doriean and others [27] found considerable bias in the point estimates and their standard errors. Exactly how network autocorrelation applied to diffusion of innovations was not clear since spatial autocorrelation measured diffusion at the macro level, but did not show whether specific individuals were more or less likely to adopt based on their network position. Further, spatial autocorrelation did not show how network structure influenced diffusion. To do so, we turn to network models.

Network Models

Figure 1 displays two networks from a study conducted in Cameroon among women in voluntary organizations [28]. Women were asked to name their friends in the organization in an attempt to determine if friendship ties were associated with contraceptive choices (they were). The diffusion network model posits that initial contraceptive choices would be made by some women based on their innovativeness and exposure to outside sources of influence such as their cosmopoliteness, media use, or greater need for the innovation. The new idea, and its practice, then spreads through the network as users persuade nonusers to adopt either by exhortation, entreaty, enticement, or example.

These influences are captured by an exposure or contagion model (Fig. 2), each individuals likelihood of adoption increases as the proportion (or number) of users in his/her personal network increases. Personal network exposure is the proportion or number of adopters in each person's network that provide information and influence with regard to some behavior. The equation for non-random mixing, or personal network exposure is:

$$E_i = \frac{\sum w_{ij} \boldsymbol{y}_j}{\sum w_i}, \tag{4}$$

where w is the social network weight matrix, and y is vector of adoptions. For an individual who reported five contacts, network exposure (E_i) is the proportion of those contacts that have adopted (Fig. 2). When network exposure is measured on direct contacts it captures social influence conveyed through overt transmission of information, persuasion or direct pressure. Alternatively, exposure can be calculated by transforming the social network, W, to reflect other social influence processes. For example, W can be transformed to represent the degree of structural equivalence (similarity in network position) among people in the network. Exposure calculated on this network captures social influence conveyed via comparison to equivalent others such as by social comparison or competition [14]. Exposure can also be weighted by network properties such as centrality to reflect social influence by opinion leaders.

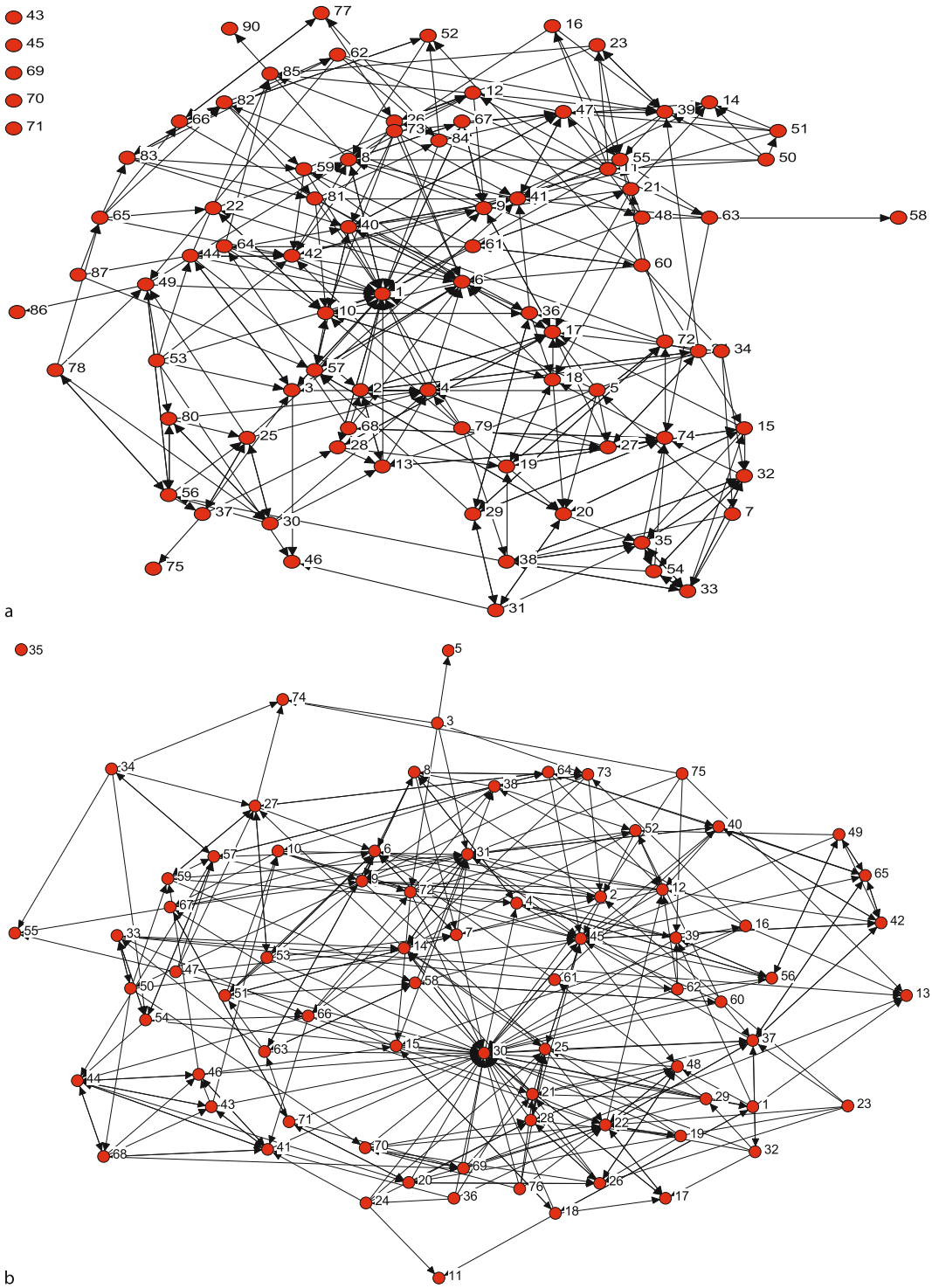
Social Networks, Diffusion Processes in, Table 2
Social network influence weightings

Relational	Positional	Central
1. Direct Ties	1. Percent Positive Matches (Tie Overlap)	1. Degree
2. Indirect Ties	2. Euclidean Distance	2. Closeness
3. Joint Participation in Groups or Events	3. Regular Equivalence	3. Betweenness
		4. Flow
		5. Integration/radiality
		6. Information
		7. Power

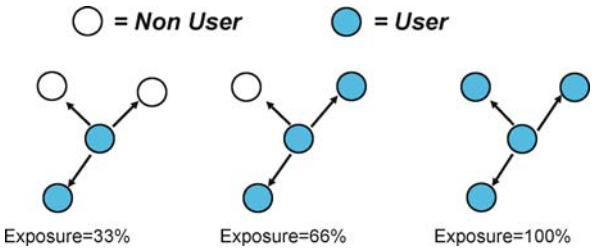
These three social influence processes are modeled with three different classes of network weight matrices (relational, positional, and central), constructed from the same social network data (Table 2). All three can be justified theoretically as sources of influence on adoption behavior and all three can be calculated various ways (there are at least 10 centrality measures). It is possible that all three operate for different people or at different times during the diffusion process.

In addition to the social influence process, a second dimension to these influence mechanisms is the weights attached to each based on social distance. For example, in relational influence models, different weights can be assigned to direct ties, ties-of-ties, and even the ties-of-ties-of-ties; in positional equivalence models, different weights can be assigned to those that are more equivalent than others [7]. A potential line of diffusion network research then is to compare different network weighting mechanisms to model and compare different social influence processes.

Diffusion was simulated through the two Cameroon networks in Fig. 1 to illustrate how network exposure and network structure influence diffusion. At each time period, adoption occurred for the non-adopter with the most nominations received, then network exposure was calculated, all nodes with exposure of 50% or higher were categorized as adopters and the process repeated. We compared diffusion in this network to that simulated in a net-



Social Networks, Diffusion Processes in, Figure 1
Networks 1 and 2 from the Cameroon Voluntary Association Study



Social Networks, Diffusion Processes in, Figure 2
Personal network exposure from direct contacts

work of the same size and density, but with links allocated randomly. Both conditions were averaged across 1000 runs. Figure 3 shows that, in Network 1, initially the diffusion trajectories are similar, but at about time 10, diffusion in the actual network accelerated.

The network accelerated diffusion because it is somewhat centralized (in-degree 21.7%) and once diffusion reaches the center of the network it can propagate rapidly. Notice that at about time 20 diffusion slowed, accelerated again at about 25, and then slowed from about time 30 to 40. These "fits and starts" are a product of the network structure: diffusion reaches pockets of inter-connectivity and spreads rapidly within these dense pockets, but slows between groups. Network 2 (Fig. 3) had more rapid and sustained diffusion because it was even more centralized (in degree, 47.2%). Note that in the spatial autocorrelation model adoptions were randomized to measure statistical significance and in this simulation, the network structure was randomized to illustrate its influence on the rate of diffusion.

Simulation assumptions regarding influences on adoption could easily be changed to achieve different outcomes. For example, when adoptions were assigned randomly, diffusion was constant in network one (and saturation lower) and similar to the random network in network two. The validity of these diffusion models rests partly on determining whether network exposure influences adoption. To that end, a number of empirical studies have been conducted to measure the degree social network exposure is associated with adoption.

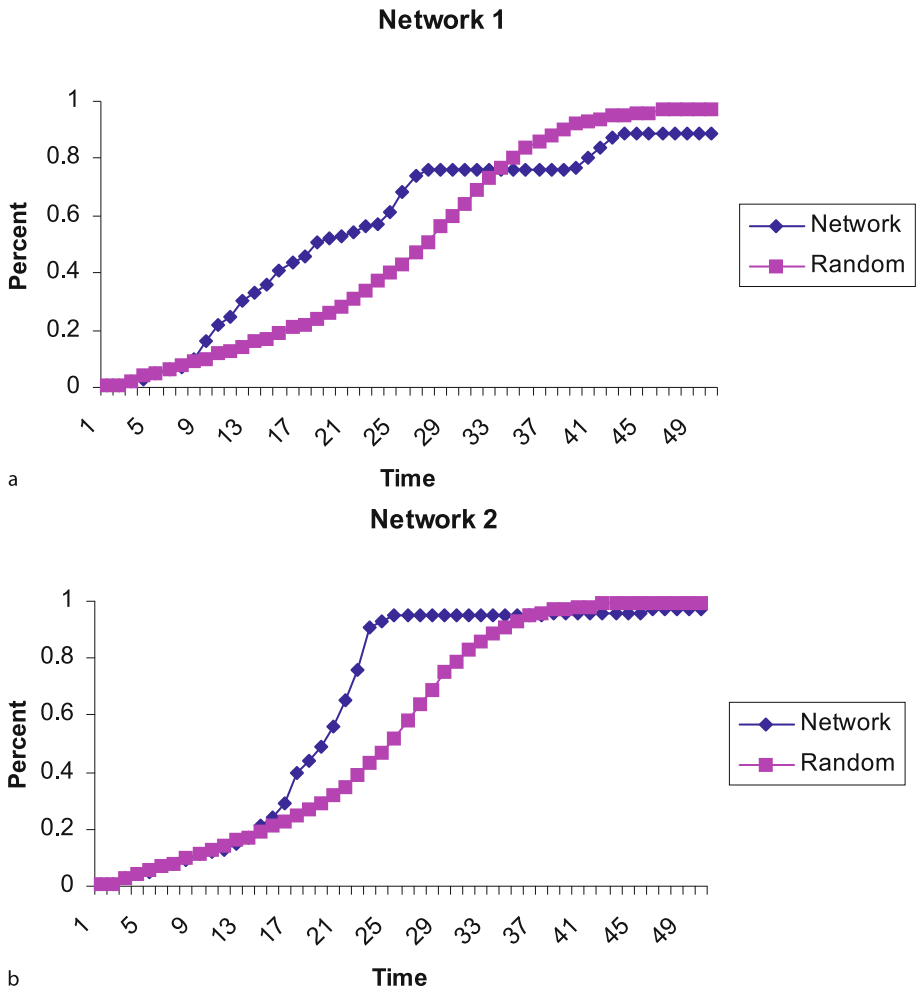
Empirical Studies

Empirical support for an association between individual own behavior and that of his/her peers can be found throughout the behavioral sciences literature. While many scholars assume adoption is associated with network exposure, few studies have traced an innovation through a network of social contacts to empirically validate this proposition. The lack of data on diffusion within an entire network stems largely from the difficulty of trying to collect data over a time period long enough for diffusion to occur. Consequently, most studies have relied on retrospective data which introduces some but not much bias [29,30]. It has also meant that several scholars have re-analyzed two studies that collected network and adoption data, (1) medical innovation study [12], reanalyzed by Burt [14], Marsden and Podolny [31], Strang and Tuma [32], Valente [7,33], and Van den Bulte and Lillien [34]; (2) Korean family planning study [35], reanalyzed by Dozier [36], Valente [7,33], Montgomery and Chung [37], and Kohler [38]. Recent studies in the fields of reproductive health [39] and substance abuse [40] have provided new data, but these classics remain classic.

Because collecting complete network data can be difficult, most empirical research has been ego-centric [41,42], based on respondent reports of their behavior and that of their network peers who are not necessarily connected to one another and not interviewed. Social influence is often based on respondent reports of perceptions of peer behavior or perceptions of peer influence [43,44]. Comparison of exposure scores based on respondent perceptions with alters' reports in one study found that perceptions were more strongly related to behavior than exposure based on alter reports [28,37,45].

Sociometric studies interview members of a bounded community and attempt to gather information from everyone in the community (typically conducted in schools, organizations, and small communities) and record their time of adoption [1,12,46,47,48]. Sociometric studies are useful for understanding how an innovation flows within the community and how certain network structural variables influence the diffusion process. Sociometric data capture network influences by the alters' reports since they were also interviewed. For example, sociometric studies can determine whether structural positions such as centrality are associated with adoption and/or whether centralization is associated with more rapid diffusion [7].

A number of recent diffusion network studies have been cross-sectional and in many cases retrospective involving only one time point. For example, a study in Thailand by Entwisle and others [49] found that contraceptive choices made by early adopters contributed significantly to the contraceptive choices made by later adopters (also see [35]). Valente and others [28] collected sociometric data on contraceptive use among women in voluntary associations in Cameroon and showed that perceptions of these friends' behavior, and in particular perceptions that these friends encouraged contraceptive use, were significantly associated with behavior. In general, statistical ana-



Social Networks, Diffusion Processes in, Figure 3
Simulated diffusion in two networks, each compared to random networks of the same size and density. Network 2 is more centralized than Network 1

lyzes use the following model:

$$\log \frac{\Pr(\boldsymbol{y}_t = 1)}{(1 - \Pr(\boldsymbol{y}_t = 1))} = \alpha + \sum B_k X_k + B_{(k+1)} \omega \boldsymbol{y}_t, \quad (5)$$

where $\boldsymbol{y}$ is a binary vector of adoption behavior, α is the intercept, β_k are parameter estimates for vectors of K socio-demographic characteristics (Xs), and ω represents the social network matrix. The $\omega \boldsymbol{y}_t$ term represents the calculation of contemporaneous network exposure and this vector is usually divided by a count of the number of nominations sent (alternatively the number of nominations can be entered into the regression separately).

Significant estimates for β_{k+1} indicate contagion effects by showing that network exposure is associated with adoption. The variances for these estimates, however, are usually biased since the observations are not independent and hence the errors in prediction are not independent. One partial solution is to obtain robust estimates by controlling for clustering. Clustering is the degree that elements from the same cluster are similar compared to those of different clusters. For example, two individuals chosen at random from the same organization are more likely to be similar than two chosen at random from different organizations. Table 3 reports regression results of the Cameroon data with and without correction for clustering. Without correction, network exposure is strongly and significantly associated with adoption, but with the correction it is only marginally statistically significant (p=.04). Controlling for clustering is particularly important in network exposure models because network choices are often restricted to the cluster.

Social Networks, Diffusion Processes in, Table 3
Logistic regression on the likelihood of contraceptive behavior on controls and network exposure with and without correction for clustering (N=555; Groups=9)

	Contraceptive Method Use			
	Without Correction		With Correction	
	Adjusted		Adjusted	
	Odds Raios	P-value	Odds Raios	P-value
Age	0.97	0.001	0.97	0.010
Education	0.91	0.247	0.91	0.184
Possessions	1.39	0.000	1.39	0.000
Network Exposure	1.14	0.005	1.14	0.047

Even with clustering controlled, social influence as measured through social networks seems to be strongly associated with behavior. For example, a school-based sociometric study was conducted by Alexander and colleagues [50] using the Adolescent health data [51] to show that students with a majority of network ties who were smokers were almost two times as likely to smoke themselves with an additional two times greater likelihood of smoking for those with best friends who smoke. Clustering was controlled and the multi-level model accurately captured micro-level effects within the context of macro-level influences. The study measured the influence of peers on smoking while conditioning on the smoking rate within the school [50].

Estimating the network exposure (autocorrelation) term with a multi-level model can provide contagion estimates across settings and estimate the degree it varies between settings (i. e., communities, schools, organizations, and so on). The models are incomplete however, because there may be factors that influence both adoption and choice of social network contacts. For example, the decision to smoke and to nominate friends who smoke may both be a function of delinquency or rebellion. Hence an association between behavior and peer behavior can be spurious. Testing social influence with network methods then requires longitudinal data involving at least two time points. Boulay and Valente [52] collected data among women in three villages of Nepal and found that having discussion partners who used contraception influenced information seeking behavior and contraceptive choice. Having data from two time points allows testing of a simple dynamic model on adoption:

$$\log \frac{\Pr(\boldsymbol{y}_t = 1)}{(1 - \Pr(\boldsymbol{y}_t = 1))} = \alpha + \sum B_k X_k + B_{(k+1)}\omega_t \boldsymbol{y}_t + B_{(k+2)}\omega_{(t-1)}\boldsymbol{y}_{(t-1)}, \quad (6)$$

where y is a binary indicator of behavior, α is the intercept, β_k are parameter estimates for vectors of K socio-demographic characteristics (Xs), and ω represents the social network matrix. A positive and significant β_{k+2} indicates that respondents with high network exposure at baseline were more likely to adopt at time two. A positive and significant β_{k+1} indicates that change in network exposure is associated with change in behavior. This may indicate contagion but still may be a product of some omitted factor. Panel data collected at two time periods is adequate for most research needs, and can provide evidence of network influences on behavior. However, since there is often a considerable time between the two measures, many factors may account for simultaneous change in behavior and network exposure. To cope with this threat, data can be collected on time of adoption, expanding the micro level dynamic analysis by using event history analysis [53].

Event History Analysis

Event history analysis techniques have been developed to analyze data with a substantive number of time points, estimating coefficients with maximum likelihood estimators [32,53,54,55,56]. There are two types of event history analysis, discrete time, in which the outcome is binary, and continuous time, in which the outcome is time-to-an-event. Since diffusion occurs over time, there is an explicit time dimension in diffusion studies captured by both discrete and continuous time models. The time of adoption variable is the dependent variable and may be influenced by both time-varying and time-constant factors. Some individuals may not have adopted by the time of data collection giving rise to time-censored observations. Right censoring occurs when data are collected before the innovation has finished diffusing or does not diffuse to all members of the community or study. Left censoring occurs when the data are incomplete at the beginning of the process. For example, adoption data for the period 1993 to 2000 may have some people who adopted in 1989–1992 classified as 1992 adopters.

There are a variety of event history techniques including hazard models developed in epidemiology used to understand the hazard or risk to disease or injury over time. Hazard and/or event history analysis generally requires that the data be reshaped from simple observations to a case-time format such that there is a case in the data for each individual at each time period of study up to and including that person's time of adoption. The time-varying and time-constant independent variables are included in each case as well as a binary indicator for whether the individual adopted or not (got sick or not).

Maximum-likelihood estimation can determine whether the independent variables are associated with the dependent variable (adopt/not adopt) [57]. A study of 100 people with an average adoption time of seven translates into 700 person-time cases. Each person-time case has a variable for the network exposure at that time period plus an indicator for whether the person adopted or not (plus additional time-constant and time-varying covariates as desired). The event history model is:

$$\log \frac{\Pr(\boldsymbol{y}_t = 1)}{(1 - \Pr(\boldsymbol{y}_t = 1))} = \alpha + \sum B_j X_j + \sum B_{kt} X_{kt} + \sum B_{(k+1)} \omega \boldsymbol{y}_t \,, \tag{7}$$

where y is a binary indicator of behavior, α is the intercept, β_j are parameter estimates for vectors of J socio-demographic characteristics (X_j), and β_{kt} are parameter estimates for the matrix of time-varying socio-demographic characteristics (X_{kt}) and ω represents the social network weight matrix, and t a time indicator. Note here we have assumed a static (constant) network. Standard statistical packages allow testing of event history or survival data in a relatively straightforward manner, once the data are reformatted. Event history analysis requires the construction of exposure matrices for each time period which can be a formidable task particularly if one uses more than one network weight matrix.

Marsden and Podolny [31] used event history analysis and tested network exposure's association with adoption in the medical innovation data. Results showed that exposure was not associated with adoption in that study. Strang and Tuma [32] revisited the issue with the same data by postulating time variance in network influence, (i. e., how much lag time, if any, is there in the influence). They found evidence of contagion. Van den Bulte and Lillien [34] supplemented the medical innovation data with archival data on media promotion by pharmaceutical firms at the time of the original study and showed that network contagion effects disappear once these data are added. Their analysis demonstrates the importance of omitted variables when studying diffusion through networks. The rapid diffusion measured in the medical innovation study indicates that contagion was probably not the primary factor driving diffusion.

Event history analysis of the three classic diffusion network datasets has been conducted [58]. The analysis controlled for within village and within person covariation, and terms for time and a logistic transformation of time were included to control for macro-level effects. Terms for infection and susceptibility [32,59] were included to measure whether adoption by central individuals (high in-degree) influenced subsequent adoption, infection, and whether centrality (out-degree) influenced a person's likelihood to adopt as diffusion occurred, susceptibility. In-degree and out-degree were also included in the model. Two network exposure terms were computed, direct ties and structural equivalence. Structural equivalence was computed as in Burt's [14] measure, Euclidian distance raised to the 16th power. For this analysis, network exposure was calculated using contemporaneous measures since two of the datasets recorded adoption in one-year intervals. Two control variables representing individual characteristics were included. Analysis was conducted only on those who adopted. The following model was estimated:

$$\log(\Pr(\boldsymbol{y}_t = 1)) = \alpha + \sum \beta_{lm} X_{lm} + \sum \beta_{lmt} V_{lmt} + \sum \beta_{lmt} \omega_s \boldsymbol{y}_t + \lambda_{lm1} C_D(\boldsymbol{y}_+) + \lambda_{lm2} C_D(\boldsymbol{y}_+) \,, \tag{8}$$

where y is a binary indicator of behavior, α is the intercept, Xs are vectors or time-constant socio-demographic and network characteristics, V represents vectors of time-varying terms, in this case time and its transformation, ω_s represents the social network matrices, and λ estimates the effects of centrality degree variables multiplied by the time varying proportion of adopters in the network (infection and susceptibility). Results were mixed, but seem to indicate that both infection and susceptibility effects are present. In all three datasets, infection is positively associated with adoption indicating that as those with high in-degree adopt, the likelihood others in the network will adopt increases. In two studies, Brazilian farmers and Korean women, susceptibility is associated with adoption indicating that those with a high number of nominations sent are more likely to adopt as the innovation diffuses. Ties sent and received are marginally associated with adoption, and only for the Brazilian data is exposure, through structural equivalence, associated with adoption. These results, however, change dramatically when non-adopters are included or when a term for the average exposure at each time period is included such that infection and susceptibility effects disappear.

The event history analysis approached has also been used by Montgomery and others (2001) using ego-centric data to study network exposure's influence on contraceptive use in Ghana. Current analysis of four rounds of data over two years has shown that contraceptive use is strongly associated with use by social network peers. The Ghana field study provides some of the most conclusive evidence of the magnitude of social influence on behavior change

by showing that as the number of social network contacts who use contraceptives increases, the likelihood of contraceptives use by ego also increases. Of all the variables, the network exposure variables were the most significant influences on contraceptive adoption. Another longitudinal field study in Kenya found similar results, again based on ego-centric network data [60].

Montgomery and others [61] also report preliminary analysis in which network influences are weighted by tie characteristics such as the frequency of communication. They found that adding these weights did not change the strength of peer influence. Similar results have been reported in Valente [7] and Valente and Saba, p. 109 in [43]. Consequently, it seems that the influence of social networks on behavior (contraceptive use in these cases) seems broad in nature, and are not conditioned on specific factors such as the frequency of communication between dyads or the their socio-demographic equality. These factors may play a strong, and even pervasive role in determining who is connected to whom [62], but they do not seem to determine the degree of influence social contacts provide.

Network exposure and adoption may not always be strongly correlated for a number of reasons. First, exposure may not be associated with adoption for everyone, but may be most influential during the middle stages of diffusion, when awareness is high, but uncertainty about its relative advantages is also high [63]. Exposure may have less of an effect early in the process when there are few adopters and obvious advantages to waiting; and late in the process when most people have a majority of adopters in their personal network anyway. Second, individuals may have varying thresholds to adoption such that some are innovative and others are not [19]. Valente [7,33] posited a social network threshold model in which contagion (majority rule) is a special case. Most simulation models assume majority influence on adoption decisions as was done in the beginning of this chapter. It is reasonable, however, to expect that individuals vary in the amount of network exposure needed to adopt an innovation. Disproving thresholds may not be possible, but construct validity for the concept has been demonstrated [33]. Valente and Saba [43] replicated the threshold model using egocentric data and showed that people with a minority of network members using contraception had higher campaign recall indicating that the media campaign could substitute for interpersonal sources of influences. If thresholds vary, network exposure is needed for people to reach those thresholds, if they do not and the special case of contagion exists, network exposure will determine when individuals adopt.

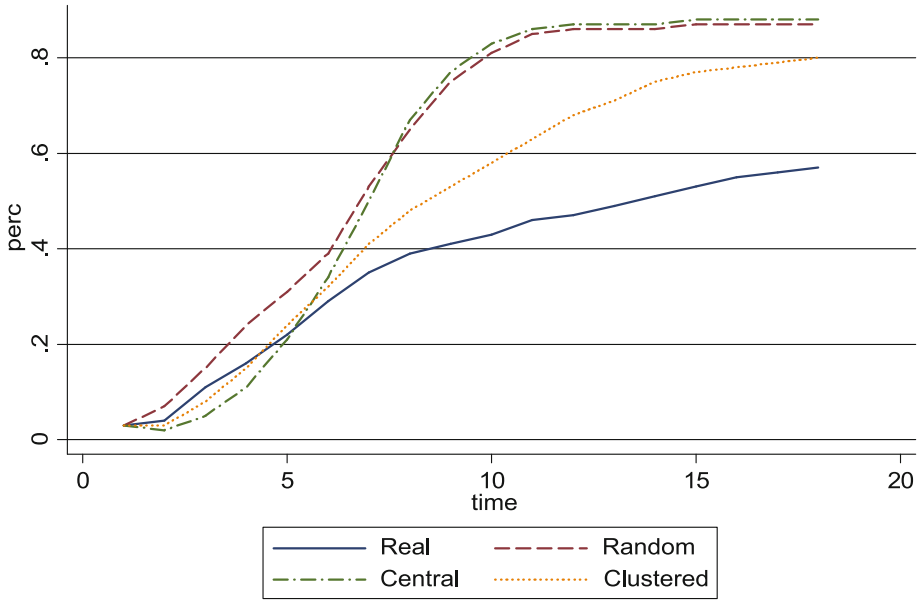
In spite of the impressive list of studies showing some support for an association between individual behavior and network exposure, and the theoretical simulations of network structure and thresholds, significant work remains to be done. Most scholars and lay people would agree that social networks influence behavior. The barriers to demonstrating this effect, however, have been challenges of data collection and agreement on appropriate statistical methodology. The most commonly analyzed dataset, Medical Innovation, is 45 years old, consists only of 125 respondents and arguably is not a diffusion study at all. The limitation of available data has forced researchers to rely on simulations and agent based modeling approaches to understand theoretical mechanisms driving diffusion.

Agent Based Models

Agent based models (ABM) are computer simulations created to understand how actors (usually though not necessarily people) behave given a set of preexisting conditions and rules for their behavior [64]. A researcher can hypothetically pose a population of people who have some degree of connectedness and some structure to those connections. Model parameters may be set that vary those structural conditions or they may be taken as fixed and other properties varied. To model the diffusion of innovations one might be interested in how network structure affects the rate of diffusion. So a simulation can be created which varies network structure and then simulates the spread of an innovation within those structures. One could also vary other properties of the diffusion by varying the number of initial adopters or the transmissibility of the innovation.

To illustrate, an empirical dataset was used to simulate theoretical network structures and potential diffusion within those structures. Data were collected among adolescents in ninth grade and they were asked to write the names of their five closest friends. A computer program was written to match the names to the roster in the school so a network could be constructed from the data. The network had 150 nodes with 327 friendship links. The agent based model had 2 parameters: (1) the initial 7% seed adopters would vary under 3 conditions: (a) randomly selected seeds, (b) the most central nodes, and (c) peripheral nodes; (2) the type of network structure would vary by 4 conditions (a) the real network, (b) a random one with the same links, (c) a centralized one, and (d) a clustered one.

A simple diffusion model was then run setting a low threshold for adoption such that at each time interval after the initial seeding those nodes connected to 15% other



Social Networks, Diffusion Processes in, Figure 4
Comparison of simulated diffusion trajectories for a real network, and 3 simulated ones based on the real network's density. The three simulated networks are random links, centralized links (centralization = 3%), and clustered links (clustering coefficient=12%). Diffusion occurs most rapidly in random and centralized networks

nodes would become adopters. The model was run for 20 time periods which was generally long enough for diffusion to occur in most situations. After a run of 20 time periods was completed, results were tabulated and then the simulation for each condition run 25 times and the results aggregated. The model was run 25 times for each condition to ensure that the results were not the product of some anomaly. It is customary in agent based models to run the simulations for hundreds or thousands to times to guarantee the results are robust.

Figure 4 shows illustrative findings by plotting the diffusion trajectories in the different conditions. The data show that simulated diffusion occurs most rapidly in the random and centralized networks. It is slowest in the real network. The network structure of real networks retards diffusion because of homophily – the tendency for people to be friends with others like themselves. This homophily tendency creates clustered networks so that diffusion gets trapped in these pockets of interconnectivity. Real networks also retard diffusion because they are symmetric (If $A \rightarrow B$ then $B \rightarrow A$) and so adoption by a pair leads to reinforcement, not more diffusion. Bridges are needed to carry the innovation across the gap between clusters and these bridges act as bottlenecks that slow diffusion.

Statistical analysis can be conducted on the data to determine the effect on diffusion. Outcomes can be the final cumulative percent of adopters (the prevalence), or the number of time periods till diffusion has reached 50%, or the rate of adoption estimated using one of the mathematical models described above. In these data we find that all 3 simulated networks have more rapid diffusion than the real network, and we find that seeding the network with central adopters accelerates diffusion while seeding it with peripheral actors slows diffusion.

Conclusions

Much progress has been made since 1943 when Ryan and Gross first laid the foundation for diffusion of innovations theory. Rogers [1] chronicled the many studies conducted since then and helped shaped a general diffusion model with wide applicability now being renewed and reinvigorated with fresh theory and analytic models. Overall, results indicate that social network influences on behavior are important and have consequences for the health and well being of populations and individuals. These new insights have shed light on important aspects of how new ideas and practices spread within and between communities.

Along with new insights have come new questions and new perspectives to be addressed. It is clear that a lack of data on time of adoption coupled with information on net-

work relations has hampered developments. Few diffusion or behavioral studies collect information on networks, conversely few network studies record time of adoption. There are advantages to marrying these two ideas, however, and future research will hopefully try to collect both types of data.

It is also clear that our understanding of how diffusion occurs is still somewhat limited. The Medical Innovation data have often been used to demonstrate the importance of networks in adoption yet analyses by Valente [7] and by Van den Bulte and Lillien [34] have shown that contagion via social influence in this setting was unlikely. Given the number of confounding factors, and some of the data requirements, it may be prohibitively difficult to substantiate the role of social networks in innovation adoption via survey methods alone. Purposively intervening on social networks, however, may prove to be a fruitful avenue of research. If network-based interventions can be used to accelerate innovation diffusion, then a stronger case can be made for the importance of social contagion in the diffusion process.

Nonetheless, it is clear that networks are important influences on behavior since most people acknowledge that they receive information and influence via their social networks and that they model the behavior of others. What is less clear is how to capture that influence in quantitative terms that mimic the theoretical progress made in the network field. Further, verbal accounts on how people make decisions and adopt behaviors usually reveal non-linearities, chance circumstances, and whims that are not independent of networks, but not easily captured in social influence models.

The link between micro and macro levels of analysis represents an opportunity for study of diffusion processes. The opportunity lies in the fact that multi-level modeling techniques enable the separation of micro-level network exposure influences from macro-level contextual factors. Yet both are social network influences and both represent elements of the diffusion paradigm. It is hoped that by controlling for contextual effects we don't "throw the baby out with the bathwater," by eliminating the micro-level influences that provide expressions for those contextual effects.

In spite of controls for macro-level contextual effects, micro-level associations between peer network behavior, and those of respondents are still sometimes strong. Debate remains about the meaning of these associations, is it peer influence, peer selection, or further contextual effects? More rigorous studies may eventually tease this out, in the interim, better study designs and interventions will need to be created. This review has attempted to point out some of the challenges diffusion scholars face and some of the promising new directions it may take.

Scholars have turned to agent based modeling as a technique to investigate diffusion properties. Agent based models (ABM) allow the theoretical exploration of how network structure can be varied and shown to influence diffusion. ABM also enables us to study diffusion parameters such as thresholds, transmissibility, or the importance of initial adopters on diffusion speed and prevalence. Until empirical data can be gathered and analyzed ABM represent a promising approach to understanding how diffusion processes occur on networks.

Future Directions

Future developments in diffusion processes are likely to occur along two interrelated fronts. First, increasingly scholars are using network data to propose and implement social change practices [65]. These activities may involve organizational change in the business context or health promotion and treatment in the medical context. These network based interventions provide the opportunity to change networks and measure the resulting changes in performance or behavior. Thus, scholars can more definitively link changes in networks to changes in outcomes.

Second, considerable progress has been made in advancing statistical models that measure network evolution and behavioral changes that account for multiple structural influences [66]. For example, scholars can test the degree of similarity in behaviors while simultaneously estimating the influence of social network characteristics such as whether two people nominate one another or are part of a linked triad. These models (known as P* or exponential random graph models) provide a statistical estimation of concepts such as network influence, and network selection.

Acknowledgments

Support for this research was provided by NIDA grant #DA10172. Some portions of this chapter appeared in [58].

Bibliography

1. Rogers EM (2003) Diffusion of innovation, 5th edn. The Free Press, New York
2. Hägerstrand T (1967) Innovation diffusion as a spatial process. University of Chicago Press, Chicago
3. Brown L (1981) Innovation diffusion: a new perspective. Methuen, New York
4. Robertson TS (1971) Innovative behavior and communication. Holt, New York

5. Bailey NTJ (1975) The mathematical theory of infectious diseases and its applications. Charles Griffen, London
6. Morris M (1993) Epidemiology and social networks: modeling structured diffusion. Sociol Methods Res 22:99–126
7. Valente T (1995) Network models of the diffusion of innovations. Hampton Press, Cresskill
8. Bael GM, Bohlen JM (1955) How farm people accept new ideas. Cooperative Extension Service Report 15, Ames
9. Katz E, Levine ML, Hamilton H (1963) Traditions of research on the diffusion of innovation. Am Sociol Rev 28:237–253
10. Ryan R, Gross N (1943) The diffusion of hybrid seed corn in two Iowa communities. Rural Sociol 8(1):15–24
11. Valente TW, Rogers EM (1995) The origins and development of the diffusion of innovations paradigm as an example of scientific growth. Sci Commun 16(3):242–73
12. Coleman JS, Katz E, Menzel H (1966) Medical innovation: a diffusion study. Bobbs Merrill, New York
13. Rogers EM (1957) Personality correlates of the adoption of technological practices. Rural Sociol 22:267–268
14. Burt R (1987) Social contagion and innovation: cohesion versus structural equivalence. Am J Sociol 92:1287–1335
15. Bass FM (1969) New product growth for model consumer durables. Manag Sci Ser a-Theory 15(5):215–227
16. Hamblin RL, Jacobsen RB, Miller JLL (1973) A mathematical theory of social change. Wiley, New York
17. Mahajan V, Peterson RA (1985) Models of innovation diffusion. Sage, Newburg Park
18. Valente T (1993) Diffusion of innovations and policy decision-making. J Commun 43(1):30–41
19. Granovetter M (1978) Threshold models of collective behavior. Am J Sociol 83:1420–1443
20. Van den Bulte C, Lillien GL (1997) Bias and systematic change in the parameter estimates of macro-level diffusion models. Marketing Sci 16:338–353
21. Cliff A, Ord JK (1981) Spatial processes: models and applications. Pion, London
22. Griffith DA et al (1999) A casebook for spatial statistical data analysis: a compilation of analyses of different thematic data sets. Oxford University Press, New York
23. Bailey TC, Gatrell AC (1995) Interactive spatial data analysis. Longman, Essex
24. Erbing L, Young A (1979) Individuals and social structure: contextual effects as endogenous feedback. Sociol Methods Res 7:396–430
25. Dow M, Burton ML, White DR (1982) Network autocorrelation: a simulation study of a foundational problem in regression and survey research. Soc Netw 4:169–200
26. Dow M (1986) Model selection procedures for network autocorrelated disturbance models. Sociol Methods Res 14:403–422
27. Doreian P, Teuter K, Wang C (1984) Network autocorrelation models: some Monte Carlo results. Sociol Methods Res 13:155–200
28. Valente TW et al (1997) Social network associations with contraceptive use among Cameroonian women in voluntary associations. Soc Sci Med 45:677–687
29. Coughenour CM (1965) The problem of reliability of adoption data in survey research. Rural Sociol 30:184–203
30. Nischan P et al (1993) Comparison of recalled and validated oral contraceptive histories. 138(9):697–703. Am J Epidemiol 189(9):697–703
31. Marsden PV, Podolny J (1990) Dynamic analysis of network diffusion processes. In: Weesie J, Flap H (eds) Social networks through time. ISOR, Utrecht
32. Strang D, Tuma NB (1993) Spatial and temporal heterogeneity in diffusion. Am J Sociol 99:614–639
33. Valente T (1996) Social network thresholds in the diffusion of innovations. Soc Netw 18:69–89
34. Van den Bulte C, Lillien GL (2001) Medical innovation revisited: social contagion versus marketing effort. Am J Sociol 106:1409–1435
35. Rogers EM, Kincaid DL (1981) Communication Networks: A new paradigm for research. Free Press, New York
36. Dozier DM (1977) Communication networks and the role of thresholds in the adoption of innovations. Ph D Thesis, Stanford University
37. Montgomery MR, Chung W (1999) Social networks and the diffusion of fertility control in the Republic of Korea. In: Leete R (ed) Dynamics of values in fertility change. Oxford University Press, Oxford
38. Kohler HP (1997) Learning in social networks and contraceptive choice. Demography 34:369–383
39. Casterline J (2001) Diffusion processes and fertility transition: selected perspectives. National Academy Press, Washington
40. Neiagus A et al (2001) HIV risk networks and HIV transmission among injecting drug users. Eval Program Plan 24:221–226
41. Marsden PV (1987) Core discussion networks of Americans. Annu Rev Sociol 52:122–131
42. Marsden PV (1990) Network data and measurement. Annu Rev Sociol 16:435–463
43. Valente T, Saba W (1998) Mass media and interpersonal influence in a reproductive health communication campaign in Bolivia. Commun Res 25:96–124
44. Valente T, Vlahov D (2001) Selective risk taking among needle exchange participants in Baltimore: implications for supplemental interventions. Am J Public Health 91:406–411
45. Urberg KA, Degirmencioglu SM, Pilgrim C (1997) Close friend and group influence on adolescent cigarette smoking and alcohol use. Dev Psychol 33:834–844
46. Becker MH (1970) Sociometric location and innovativeness: reformulation and extension of the diffusion model. Am Sociol Rev 35:267–282
47. Scott J (2000) Network analysis: a handbook, 2nd edn. Sage, Newbury Park
48. Wasserman S, Faust K (1994) Social networks analysis: methods and applications. Cambridge University Press, Cambridge
49. Entwisle B et al (1996) Community and contraceptive choice in rural Thailand: a case study of Nang Rong. Demography 33:1–11
50. Alexander C et al (2001) Peers, schools, and adolescent cigarette smoking. J Adolesc Health 29(1):22–30
51. Bearman PS, Jones J, Udry JR (2000) National longitudinal study of adolescent health: research design. http://www.cpc.unc.edu/projects/addhealth/data
52. Boulay M, Valente TW (2005) Dynamic sources of information and dissonance in the discussion networks of women in rural Nepal. J Health Commun 10:519–536
53. Tuma NB, Hannan MT (1984) Social dynamics: models and methods. Academic Press, New York
54. Allison PD (1984) Event history analysis. Sage, Newberry Park
55. Bartholomew DJ (1982) Stochastic models for social processes. Wiley, New York

56. Teachman JD, Hayward MD (1993) Interpreting hazard rate models. Sociol Methods Res 21(3):340–371
57. Eliason SR (1979) Maximum likelihood estimation: logic and practice. Sage, Newburg Park
58. Valente T (2005) Models and methods for innovation diffusion. In: Carrington PJ, Scott J, Wasserman S (eds) Models and methods in social network analysis. Cambridge University Press, Cambridge
59. Myers DJ (2000) The diffusion of collective violence: infectiousness, susceptibility, and mass media networks. Am J Sociol 106:173–208
60. Kohler H-P, Behrman JR, Watkins SC (2001) The density of social networks and fertility decisions: evidence from South Nyanza District, Kenya. Demography 38:43–58
61. Montgomery MR et al (1999) Social networks and contraceptive dynamics in Southern Ghana. Paper presented at the annual meeting of the Population Association of America. Washington DC
62. White K, Watkins SC (2000) Accuracy, stability and reciprocity in informal conversational networks in rural Kenya. Soc Netw 22:337–356
63. Carley KM (2001) Learning and using new ideas: a socio-cognitive approach. In: Casterline J (ed) Diffusion processes and fertility transition: selected perspectives. National Academy Press, Washington
64. Epstein JM (2006) Generative social science. Princeton University Press, Princeton
65. Valente TW, Fosados R (2006) Diffusion of innovations and network segmentation: the part played by people in the promotion of health. J Sex Transm Dis 33:S23-S31
66. Robins G, Pattison P, Kalish Y, Lusher D (2007) An introduction to exponential random graph (p*) models for social networks. Soc Netw 29:173–191

Social Networks, Exponential Random Graph (p^*) Models for

Garry Robins
School of Behavioural Science, University of Melbourne, Melbourne, Australia

Article Outline

Glossary
Definition of the Subject
Introduction
Notation and Terminology
Dependence Hypotheses
Bernoulli Random Graph (Erdös–Rényi) Models
Dyadic Independence Models
Markov Random Graphs
Simulation and Model Degeneracy
Social Circuit Dependence:
  Partial Conditional Dependence Hypotheses
Social Circuit Specifications
Estimation
Goodness of Fit and Comparisons
  with Markov Models
Further Extensions and Future Directions
Bibliography

Exponential random graph models, also known as p^* models, constitute a family of statistical models for social networks. The importance of this modeling framework lies in its capacity to represent social structural effects commonly observed in many human social networks, including general degree-based effects as well as reciprocity and transitivity, and at the node-level, homophily and attribute-based activity and popularity effects. The models can be derived from explicit hypotheses about dependencies among network ties. They are parametrized in terms of the prevalence of small subgraphs (configurations) in the network and can be interpreted as describing the combinations of local social processes from which a given network emerges. The models are estimable from data and readily simulated. Versions of the models have been proposed for univariate and multivariate networks, valued networks, bipartite graphs and for longitudinal network data. Nodal attribute data can be incorporated in social selection models, and through an analogous framework for social influence models.

The modeling approach was first proposed in the statistical literature in the mid-1980s, building on previous work in the spatial statistics and statistical mechanics literature. In the 1990s, the models were picked up and extended by the social networks research community. In this century, with the development of effective estimation and simulation procedures, there has been a growing understanding of certain inadequacies in the original form of the models. Recently developed specifications for these models have shown a substantial improvement in fitting real social network data, to the point where for many network data sets a large number of graph features can be successfully reproduced by the fitted models.

Glossary

Alternating independent-2-paths A parameter (and statistic) in new specification models; a particular combination of *k-independent-2-path* counts into the one statistic.

Alternating k-stars A Markov parameter (and statistic) in the new specification models; a particular combination of Markov *k*-star counts into the one statistic; equivalent to *geometrically weighted degree counts*; useful for modeling the degree distribution.

Alternating k-triangles A parameter (and statistic) in the new specification models; a particular combination of *k-triangle* counts into the one statistic; equivalent to *weighted shared partners*.

Cyclic triad A Markov graph configuration: in a directed network, ties *ij, jk* and *ki* are observed among actors *i*, *j*, and *k*.

Degeneracy (or near-degeneracy) When a model implies that very few distinct graphs are probable, often only empty or complete graphs; degenerate models cannot be good models for social network data.

Dependence assumption Theoretical assumption about dependencies among possible network ties; determines the type of parameters in the model.

Dyad independence Assumes that dyads are independent of one another; the model includes edge and reciprocity parameters, and possibly also node or dyad attributes.

Dyad-wise shared partners A parameter (and statistic) in the higher order models; equivalent to *alternating independent 2-paths*.

Edge-wise shared partner distribution Distribution of the number of dyads who are themselves related and who have a fixed number of shared partners.

Edge-wise shared partners A parameter (and statistic) in the higher order models; equivalent to *alternating k-triangles*.

Geometrically weighted degree counts A statistic (and parameter) in the new specification models: a sum of degree counts with geometrically decreasing weights; equivalent to *alternating k-stars*.

Homogeneity assumption Assumption about which parameters to equate, to make a model identifiable.

***k*-independent-2-paths** Configurations in the higher order models; equivalent to *k*-triangles but without the base.

***k*-in-star** A Markov graph configuration: in a directed graph, *k* arcs are directed to the one actor.

***k*-out-star** A Markov graph configuration: in a directed graph, *k* arcs are expressed by the one actor.

***k*-triangle** A configuration in higher order models; in a non-directed graph, the combination of *k* triangles, each sharing the one edge (the base of the *k*-triangle).

***k*-star** A Markov graph configuration: in a non-directed graph, *k* edges are expressed by the one actor.

Markov dependence assumption Introduced by Frank and Strauss [9], proposes that, conditional on the rest of the graph, two possible ties are independent of each other unless they share an actor.

Mixed-star A Markov graph configuration: a two path in a directed graph.

Monte Carlo Markov chain maximum likelihood estimation (MCMCMLE) Method of estimation based on computer simulation; more principled than pseudolikelihood.

Network configuration A small subgraph that may be observed in the data and that is represented by parameters in the model: e. g. reciprocated ties, triangles.

Parameters Relate to specific network configurations that may be observed in the graph; a large positive parameter is interpreted as the presence of more of the configurations than might be expected from chance (given the other effects in the model); a large negative parameter signifies the relative absence of the configuration.

Partial dependence assumption Assumption for dependencies among possible ties created by the presence of other ties; permits models with higher order configurations than Markov configurations.

Pseudo-likelihood estimation An approximate method of estimation using logistic regression; does not produce reliable standard errors.

p_1 Model An early dyad independence model, including popularity and expansiveness effects.

p_2 Model Elaboration of p_1 model, where popularity and expansiveness effects are random, and independent variables may be used to predict ties.

Simple random graphs, Bernoulli graphs, Erdös–Rényi graphs Assume that edges are independent of one another and are observed with a given probability.

Social circuit dependence Two possible ties are conditionally dependent when, if observed, they would create a 4-cycle.

Transitive triad A Markov graph configuration: in a directed network, ties *ij, jk* and *ik* are observed among actors *i*, *j*, and *k*.

Triangle A Markov graph configuration: in a non-directed network, a clique of three actors, ties *ij, jk* and *ik* are observed among actors *i*, *j*, and *k*.

Definition of the Subject

Exponential random graph models, also known as p^* models, constitute a family of statistical models for social networks. These models take the form of a probability distribution of graphs:

$$\Pr(\mathbf{X} = \mathbf{x}) = (1/\kappa)\exp\{\boldsymbol{\eta}'\mathbf{g}\},$$

for a set of tie indicator variables $\mathbf{X}$ on a network of fixed node size n, where $\mathbf{x}$ is a realization, with a parameter vector $\boldsymbol{\eta}$ and a vector of network statistics $\mathbf{g}$. Each value of the parameter vector corresponds to a probability distribution on the set of all graphs with n nodes.

Introduction

As noted in [34], statistical approaches to social networks have quite a long history, stretching at least back to the work of Moreno in the 1930s [24]. Yet, that history is rather sparsely scattered across various literatures and different eras and, as a result, even today we see that older techniques and approaches are reinvented and repackaged as new. The *statistical modeling* of a social network – as distinct from the use of statistically-based approaches to understanding particular properties of a social network – has a shorter and somewhat more coherent pedigree.

Criteria for a successful statistical model of a social network have been proposed by [34]: it should be possible to estimate the parameters of the model from data in a principled way, and the fitted model should be a good representation of that data; the model should provide theoretically plausible interpretations about the type of effects that might have produced the network; and using the estimated model parameters it should be possible to draw inferences about competing explanations for the data. Most, perhaps all, models for social networks fall short of completely meeting these requirements. Some models are "thought experiments", intended to illuminate but with an uncertain link to data. Some models are estimable from data but still cannot adequately represent important features of the network. Models may imply interpretations that are just not theoretically plausible in terms of social science; and some models cannot include different effects simultaneously in order to test one against the other. Of course, by definition all models are imperfect (else, they are not "models"). So these criteria direct the aim in model development, even under the knowledge that we cannot always hit the target exactly.

A major insight in the statistical modeling of social network structure has been that structural effects can be detected as some form of deviation from what would be expected as "randomness". This was in fact the original proposal of Moreno and colleagues [24] who were the first to compare observed network data to what would be expected from null distributions. More explicitly, Rapoport's biased net theory [32,33] proposed certain structural "biases" away from random tie formation, including the important notions of transitivity (later to be described as clustering), the propensity for human social networks to exhibit a high proportion of triangulated ties.

Of course, in the graph theory literature the best-known model for randomness is the *simple random graph model*, often known as the Erdös–Rényi graph [7] or uniform Bernoulli graph distribution. This model proposes that for a given number of nodes network ties are observed between pairs of nodes independently and with a fixed probability p. Properties of this model have been examined extensively, utilizing its analytic tractability. It has also been used as a null model for various sampling distributions (e. g. see the summary in [8].) For an observed network, p can be readily estimated (the maximum likelihood estimate is simply the density of the network), but unfortunately this model is not a good representation of almost any human social network. There are no structural effects here, only randomness.

For directed networks, [15] extended the Bernoulli model in the early 1980s by parametrizing structural effects for reciprocity and for differential node-level activity (out-degree) and popularity (in-degree). Possible network ties within dyads were dependent on one another, but were independent between dyads. This dyad-independence assumption permitted the model to be estimated as a straight-forward loglinear model. Holland and Leinhardt called this model p_1, the subscripted '1' implying a program of further research, with progressively enlarged dependence assumptions (within arcs, within dyads, within triads, etc.). However, Holland and Leinhardt were uncertain how to progress beyond dyads. The problem was that standard statistical estimation required some level of independence and triads, unlike dyads, overlapped. (A subsequent, more sophisticated extension, the p_2 model, also has dyadic independence at its heart, but conditional on random node-level effects, so that the random effects indirectly introduce dependencies that may extend beyond dyads [53].)

Ove Frank and his colleague, David Strauss, provided the crucial insight: in the world of complex networks, it was dependence that mattered, and assumptions involving traditional statistical independence, although helpful for fitting models, were likely to be inadequate empirically [9]. In contrast, their approach centered on *conditional* independence, whereby contingencies among network ties may transmit across the network in the sense that the presence of one network tie may affect the presence of any other, but most pairs of ties were in some sense "remote", so that the contingencies affecting a given network tie had to be transmitted through "neighboring" ties. So once one had a concept of "neighboring ties", that is, some notion of which possible ties were conditionally independent after taking into account other observed network ties, the form of a model could be specified. Approaches from spatial statistics and statistical mechanics, with notions of dependence within neighborhoods (broadly defined) could be translated in this way into social network models. Frank and Strauss proposed Markov random graph mod-

els, based on an explicit argument about conditional independence among networks ties, that could be regarded as theoretically plausible for the first time. These models incorporated reciprocity, degree-based effects and various forms of triangulation or clustering.

In the 1990s, Markov random graph models became the accepted form of exponential random graph models and they have only recently been superseded. They could be estimated through the rough and ready procedure of pseudo-likelihood [52], a not particularly sophisticated approach but enough to get the field started empirically. Wasserman and Pattison popularized these models among social networks researchers as p^* models [55] and added further generalizations, with extensions to multivariate [29] and valued networks [36]. Nodal attributes were introduced in social selection [37] and social influence models [38].

In the last decade, especially with more principled methods of estimation and simulation, it has been shown that homogeneous Markov random graph models face considerable, often insurmountable, difficulties in dealing with real network data. Data with inhomogeneities – for instance, very high degree nodes, or dense regions of multiple triangulation, both of which are not uncommon in real social networks – typically result in *degeneracy* for Markov random graph models, where no set of parameter estimates can adequately represent the data. Newer specifications have been proposed by Snijders and colleagues [48] that help considerably with these issues of degeneracy, and have the capacity to enhance our ability to model small-scale to medium-sized social networks, in some cases remarkably well. These advances are based on new theorizations of dependence among network ties, echoing the original project of Holland and Leinhardt with a notion of increasingly extended dependence assumptions.

This article begins with a presentation of notation and terminology and then discusses how dependence hypotheses lead to the general form of the model. Specific examples of dependence hypotheses, and the resulting models, are then presented: Bernoulli graph distributions; dyadic independence models; and Markov random graph models. Methods of simulation are briefly introduced before a discussion of degeneracy issues, particularly as they apply to Markov random graph models. Recently proposed dependence hypotheses are introduced, including the so-called "social circuit" dependence, leading to additional specifications for exponential random graph models that substantially improve model performance. Estimation and goodness of fit approaches are discussed before a short empirical example is presented. Extensions are briefly described followed by some concluding remarks about future directions.

Notation and Terminology

A network comprises a set of relational ties between pairs of individual *actors* (be they people or other social entities). A network can be represented as a graph G with node set $N = \{1, 2, \ldots, n\}$ representing the individuals and edge set E representing the relational ties. For the statistical models of this article, relational ties are construed as a set of binary random variables X_{ij} such that $X_{ij} = 1$ if a tie is observed from node i to node j, for $i \neq j$, and $X_{ij} = 0$, otherwise. For a nondirected network X_{ij} and X_{ji} are equivalent, whereas for a directed network the two variables are distinct. X_{ii} is undefined (or may be considered as a structural zero). A nondirected tie is an *edge* and a directed tie an *arc*. We specify x_{ij} as the observed value of the variable X_{ij} and we let $\mathbf{X}$ be the matrix of all variables with $\mathbf{x}$ being a realization (where the diagonals of the matrices are forced to be zero). $\mathbf{X}$ and $\mathbf{x}$ are necessarily symmetric for nondirected networks, but not so for directed networks. More generally, $\mathbf{x}$ may be valued but in this article we restrict attention to binary ties.

There are natural extensions to this basic notation, depending on the data structures. For instance, suppose the data is in the form of a bipartite network, with two distinct sets of nodes and ties between nodes of different types but not between nodes of the same type (e. g. such a data structure can represent people's membership of clubs). Then $\mathbf{X}$ may comprise a set of variables X_{pa} indicating the presence or absence of a tie between a node p of the first type and a node a of the second type. The data may involve *multiple* or *multivariate networks*, with r different types of relations among the one set of nodes. In that case $\mathbf{X}$ may be thought of as a three-way array of variables X_{ijr} such that $X_{ijr} = 1$ indicates that presence of a tie of type r from node i to node j. An analogous three-way array may represent network data collected for the same set of nodes at different time points such that $X_{ijt} = 1$ indicates that presence of a tie at time t from node i to node j.

Nodes themselves can have certain properties (*attributes*) that may be measured as binary, categorical or continuous variables. We denote an attribute variable with the vector $\mathbf{Y}$, where $Y_i = y_i$ indicates that for attribute Y node i has a value y_i.

Dependence Hypotheses

Following work in spatial statistics [1], Frank and Strauss introduced the notion of a *dependence graph* into social network modeling [9] in order to represent possible de-

pendencies among network variables X_{ij}. The nodes of the dependence graph are the network variables X_{ij} and an edge between two nodes indicates dependence between the respective variables, even when the observed values for all remaining variables are known. These edges specify a *neighborhood* relationship between pairs of variables and the cliques of the dependence graph can be thought of as *local social neighborhoods* for the set of tie variables [27]. Conversely, the absence of an edge in the dependence graph indicates that two network variables are conditionally independent, and are not neighbors, so that their interaction need not be taken into account in a model based on local social neighborhoods. (A more technical discussion on dependence graphs for social network models can be found in [35]).

The previous paragraph describes one important result from the Hammersley–Clifford theorem [1], that the form of a probability distribution for a set of interacting variables relates solely to the neighborhood structure (or clique structure) of the dependence graph (with a single network variable also taken as a clique). For a given node set N, once a dependence hypothesis is specified and the dependence graph is defined, it follows *necessarily* from the Hammersley–Clifford theorem that:

$$\Pr(\mathbf{X} = \mathbf{x}) = (1/\kappa(\boldsymbol{\eta})) \exp\{\Sigma_A \eta_A g_A(\mathbf{x})\}\,, \qquad (1)$$

where:

(i) the summation is over all neighborhoods A;
(ii) η_A is a parameter corresponding to the neighborhood A (and there is no non-zero parameter for any set of ties that do not constitute a neighborhood);
(iii) $g_A(\mathbf{x}) = \prod_{x_{ij} \in A} x_{ij}$ is the *network statistic* corresponding to neighborhood A and indicates whether all the ties in A are observed in the network $\mathbf{x}$;
(iv) κ is a normalizing quantity, that is a function of the parameter vector $\boldsymbol{\eta}$, to ensure that (1) is a proper probability distribution.

All exponential random graph models take this general form, describing a probability distribution of graphs on n nodes. Neighborhoods A are subsets of possible ties, so they represent possible subgraphs that may or may not be observed in the network $\mathbf{x}$. There is one, and no more than one, parameter for each distinct neighborhood (although one neighborhood may be a subgraph of another and so both may have separate parameters).

The form of the model in (1) is too general to be identifiable and some homogeneity constraint needs to be imposed on what is otherwise a large number of neighborhoods A (and hence a large number of parameters). The original proposal [9] was that parameters should be equated across subsets of ties that were isomorphic to each other (that is, subgraphs that are indistinguishable once node labels are removed). With some further constraints (noted below), this assumption can produce identifiable models that can be fitted to data. (Other possible ways to constrain the number of parameters are also described below). Following the terminology first used by Moreno [24], we term these isomorphic neighborhoods as *network configurations* (although some of the network literature of the last decade has adopted the term *motif*). Counts of configurations become the sufficient statistics of the model.

Examination of the parameters for the various configurations of a fitted model can provide insight into the social processes that may underpin the network. The strength and direction of any particular parameter value will affect how frequently the corresponding configuration is observed. If the parameter is large and positive, we expect to observe the corresponding configuration to occur more frequently than if the parameter value were zero [40].

Obviously, good dependence hypotheses are necessary for this approach to work, as they crucially shape the nature of the model by defining the possible configurations.

Bernoulli Random Graph (Erdös–Rényi) Models

Suppose we hypothesize that there are no dependencies within the network, so that the dependence graph has no edges and the only neighborhoods are single tie variables X_{ij}. Then (1) becomes:

$$\Pr(\mathbf{X} = \mathbf{x}) = (1/\kappa) \exp(\Sigma_{i,j} \eta_{ij} x_{ij})\,.$$

Constraining parameters to be equal for isomorphic neighborhoods here implies equating all the parameters $\eta_{ij} = \theta$, as there is only one type of neighborhood. Then we have:

$$\Pr(\mathbf{X} = \mathbf{x}) = (1/\kappa) \exp(\Sigma_{i,j} \theta x_{ij}) = (1/\kappa) \exp(\theta L(\mathbf{x}))\,,$$

where $L(\mathbf{x})$ is the number of edges (arcs) in the network $\mathbf{x}$ and θ is a *density* or *edge* parameter (sometimes called a *choice* parameter in older literature). This graph distribution is equivalent to that of a distribution of simple random graphs where the probability of a tie between a pair of edges is $\exp\theta/(1 + \exp\theta)$.

The dependence assumption here is unrealistic and this model will fit few, if any, real networks, but it is often a useful null or baseline model.

Dyadic Independence Models

For directed networks, a dyadic independence hypothesis implies edges in the dependence graph between variables

X_{ij} and X_{ji}. Neighborhoods then take the form of single edges and dyadic pairs $\{X_{ij}, X_{ji}\}$. Applying constraints in the form of the one density parameter θ to all single edge neighborhoods and the one *reciprocity parameter* ρ to all dyadic neighborhoods results in the following two parameter model:

$$\Pr(\mathbf{X} = \mathbf{x}) = (1/\kappa)\exp(\theta L(\mathbf{x}) + \rho M(\mathbf{x}))\,,$$

where $L(\mathbf{x})$ is the number of arcs and $M(\mathbf{x})$ the number of mutual dyads (i. e. where $X_{ij} = X_{ji} = 1$) in the network $\mathbf{x}$. This model has the capacity to model reciprocity effects in directed networks. Slightly less sweeping homogeneity constraints result in the p_1 model [15].

The interpretation of a single parameter in an exponential random graph model is relative to other effects in the model, so that in this case, the presence of a large and positive ρ parameter is relative to the density effect. This is an advantage: in this model, for instance, we can infer that there is substantial reciprocity *given the density of the graph*. It is of little use to talk about a large reciprocity effect in absolute terms, because the number of arcs in the graph (i. e. the density) establishes the preconditions for reciprocity. In other words, graphs of very low density have very little opportunity to demonstrate reciprocity anyway, and the absence of mutual dyads in a low density graph may simply be a feature of the low density. Similarly the presence of many mutual dyads in a high density graph may simply be explained by the high density, without the need to infer a reciprocity effect. A large positive ρ, on the other hand, shows that the number of arcs in a graph are arranged in sufficiently many mutual dyads to suggest that a separate and substantial reciprocity process is required to explain the structure of this network, over and above any density effect.

Once again, the dyadic independence assumption does not do a good job of reproducing real networks, because the model has no capacity for triangulation. The model may be useful as a baseline model.

Markov Random Graphs

Frank and Strauss [9] proposed Markov dependence, by postulating that a possible tie from i to j is assumed to be contingent on any other possible tie involving i or j, even if all other ties in the network are fixed. Markov dependence implies that two possible network ties are conditionally independent unless they share a common actor. They showed that this assumption resulted in models with configurations for nondirected graphs of single edges, star-like structures and triangles (or 3-cycles). For directed graph models, configurations include directed counterparts of these as well as reciprocity (see [55] for a fuller description of directed graph parameters). Frank and Strauss termed these *Markov random graph models*.

Once homogeneity is imposed across isomorphic neighborhoods, we obtain configurations (and related parameters) for nondirected networks as depicted in Fig. 1. The resulting model is:

$$\Pr(\mathbf{X} = \mathbf{x}) = (1/\kappa)\exp(\theta L(\mathbf{x}) + \Sigma_{r=2,n-1}\sigma_r S_r(\mathbf{x}) + \tau T(\mathbf{x}))\,, \quad (2)$$

where:

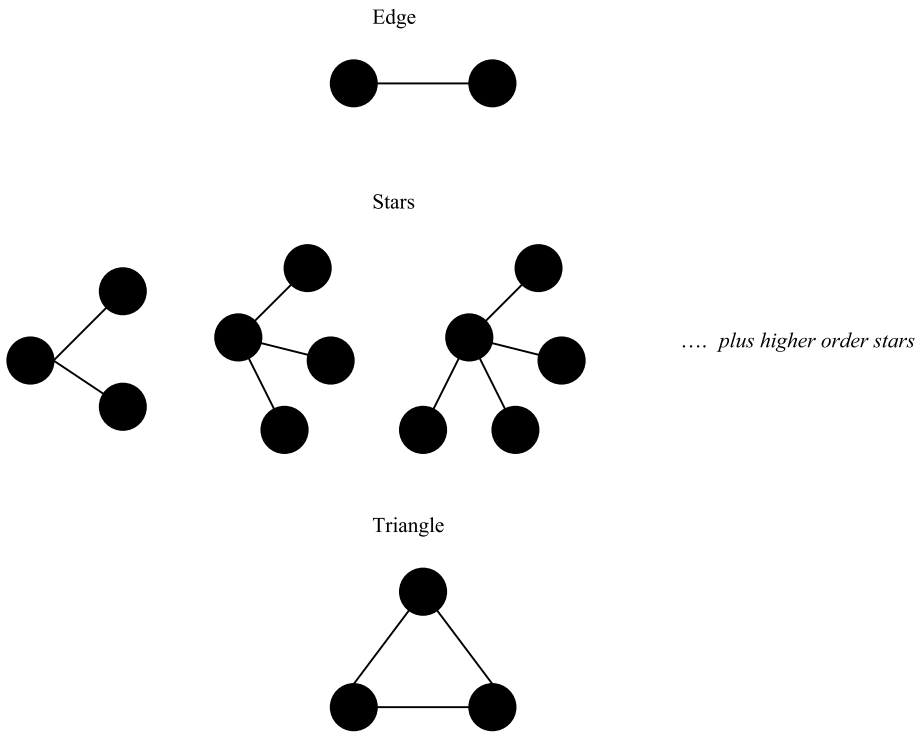
(i) θ is a density parameter, and $L(\mathbf{x})$ the number of edges in $\mathbf{x}$, as before;
(ii) σ_r is a parameter for a *star of size r*, and $S_r(\mathbf{x})$ is the number of stars of size r in $\mathbf{x}$;
(iii) a star of size r is a configuration centered on a single node i such that there are r edges emanating from i (note that stars of size larger than r contain many stars of size r, so that this is not simply a partition into nodes of different degrees, although the full set of counts $S_r(\mathbf{x})$ can be converted into the degree distribution, and vice versa);
(iv) τ is a triangle or clustering parameter and $T(\mathbf{x})$ is a count of the number of triangles in $\mathbf{x}$.

This model still has too many star parameters to be identifiable. Note that the expression $\theta L(\mathbf{x}) + \Sigma_{r=2,n-1}\sigma_r S_r(\mathbf{x})$ completely parametrizes the degree distribution (so the τ parameter can be interpreted as the strength of transitivity conditional on the degree distribution of the graph). An identifiable model can be obtained by restricting the number of star parameters to be substantially less than $n - 2$, for instance, to less than four (an alternative approach is described below). Then the model becomes:

$$\Pr(\mathbf{X} = \mathbf{x}) = (1/\kappa)\exp(\theta L(\mathbf{x}) + \sigma_2 S_2(\mathbf{x}) + \sigma_3 S_3(\mathbf{x}) + \tau T(\mathbf{x}))\,. \quad (3)$$

The parameters can be interpreted in ways that represent plausible social processes [39]:

(i) θ is a baseline propensity for social ties to form;
(ii) σ_2 represents a tendency for individuals to seek multiple partners (if it is positive);
(iii) whereas σ_3 (if negative as it often is when this model is fitted to real data) represents a ceiling effect against having too many partners; so, together the balance between σ_2 and σ_3 represents the benefits of multiple



Social Networks, Exponential Random Graph (p^*) Models for, Figure 1
Configurations for a nondirected Markov random graph model

contacts against the costs of maintaining too many contacts;

(iv) while τ represents tendencies towards clustering.

An alternative interpretation is to consider the θ, σ_2 and σ_3 parameters as controlling for the first three moments of the degree distribution, so that while model (2) in effect presents a full parametrization, model (3) provides a more parsimonious control over the degree distribution (saturated models of the degree distribution are discussed by [43] and [47], see also [11]). Of course, for any data set it is an empirical question whether the first three moments are sufficient to capture the degree distribution adequately. Below, we describe ways to investigate that question.

Extensions of this basic Markov random graph model include models for multivariate networks [29], for valued networks [36] and for affiliation networks [42] (see also [28]).

By tuning parameter values even the simple four parameter model (3) can represent many different types of networks, including small world networks, networks with long paths, and highly clustered ("caveman") networks [39]. For some parameter values, however, the models became "frozen" into certain unchanging highly structured patterns. This effect is an illustration of *degeneracy* for these models, which was a research issue given important attention in the early years of the 21st century. The upshot of this body of research was that homogeneous Markov random graph models were recognized as frequently inadequate to deal with real social network data.

Simulation and Model Degeneracy

The problems of degeneracy became apparent through a number of simulation studies. Strauss was the first to describe a rather straightforward application of standard statistical simulation techniques (e. g. Metropolis or Metropolis–Hastings algorithms) to simulate Markov random graph models for a fixed set of nodes and given set of parameter values [51]. More recent treatments and some important results are given in [13,14,17,39,45]. These methods provide a principled statistical means to produce a distribution of graphs with probabilities of each graph being observed in the distribution consistent with any particular model in the form of (1). Typically the simulation is run for a large number of iterations and a sample of graphs is extracted and examined to understand typical properties of the graphs in the distribution.

A graph distribution is termed as *near degenerate* [13,14] if it implies only a very few (possibly only one

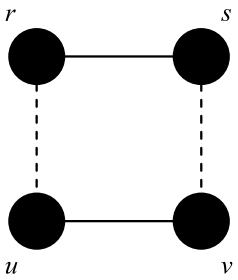
or two) distinct graphs with substantial non-zero probabilities. For instance, certain parameter values for Markov random graph models place almost all of the probability mass on either the empty or the full graph. These cannot be good models for human social networks. For certain parameter values, there may be two quite separate regions where graphs are most likely (possibly one region of low density and another of high density graphs), with the possibility of a dramatic phase transition from one region to the other (for instance, as parameter values change very slightly). Such phase transitions and other aspects of degeneracy have been studied for a variety of simple Markov random graph models by a number of authors [2,12,13,14,19,26,39,41,45,48].

Issues of degeneracy call into question whether Markov random graph models can adequately represent most network data. It would not matter if some regions of the parameter space produced degenerate models, as long as those regions applicable to empirical data were non-degenerate. Experience shows the opposite applies: for some data, Markov random graph models turn out to be well-behaved, but for many empirical networks they are degenerate. When applied to social networks with a few very high degree nodes, or with some regions of high triangulation, degeneracy frequently occurs for Markov models. Models such as (3) reflect homogeneity constraints, and they appear to have difficulty when the data presents inhomogeneity in the distribution of degrees or triangles. It is not entirely certain whether the problem with the models arises because of the Markov dependence assumption or the assumption of homogeneity. In any event, to deal more effectively with real networks, a different approach to these models seems necessary. To date, attention has been directed to revisiting the dependence assumption.

Snidjers and colleagues [48] drew two conclusions from the tendency for Markov random graph models to be degenerate when fitted to real networks with a high level of clustering: (1) the Markov dependence assumption may be too restrictive; (2) the representation of the social phenomenon of transitivity by the total number of triangles might be too simplistic. They proposed specifications that drew on both these possibilities. Before we introduce these new specifications, however, we turn to higher order network dependencies.

Social Circuit Dependence: Partial Conditional Dependence Hypotheses

There are several ways in which network dependence assumptions could be extended beyond Markov dependence [27]. One approach is of *partial conditional dependence*, whereby the presence of certain network ties *created* dependence among other possible network ties. This assumption permits the emergence of dependence structures as ties come into and go out of existence. This type of dependence can be incorporated into a version of the Hammersley–Clifford theorem and hence into exponential random graph models [27].



Social Networks, Exponential Random Graph (p^*) Models for, Figure 2
Social circuit dependence (*Broken lines* represent possible edges; *full lines* represent observed edges X_{ru} is conditionally dependent on X_{sv})

A particular hypothesis about this type of emergent dependence was used by [48] and described by [41] as *social circuit dependence*. Social circuit dependence is defined as two possible network ties being conditionally dependent if their observation would lead to a 4-cycle. This type of dependence is depicted in Fig. 2. Here $X_{rs} = X_{uv} = 1$, that is, in the network data there are observed edges between nodes r and s, and between nodes u and v. The presence of these edges leads to conditional dependence between variables X_{ru} and X_{sv} because if edges were also observed between r and u, and between s and v, a 4-cycle would result in the graph.

Interpretation of this hypothesis is discussed by [41]. Generally we would expect two distinct possible edges (r, u) and (s, v) to be conditionally independent. If, however, person r knows person s, and person u knows person v, then the presence of a tie between r and u can make the presence of a tie between s and v more likely, that is, they are conditionally dependent. It is simple to think of real social circumstances where this happens: for instance, in families the presence of a friendship between two children increases the chances of the two mothers coming to know one another; or in a business, cooperation between two bosses may lead to their employees working together; or in research, two postdoctoral researchers might discuss issues with each other because their academic mentors have been collaborators. In all of these cases, what makes the dependence come into effect is the presence of ties that

constitute part of the 4-cycle. Without those ties, the dependence does not arise. So, the mothers are the mothers of the two particular children, not the mothers of entirely different children; the employees are employed by those particular bosses; and the two post-docs are supported by those two mentors.

Social Circuit Specifications

Snijders and colleagues [48] proposed three new statistics for exponential random graph models for nondirected networks: *alternating k-stars*, *alternating k-triangles*, and *alternating independent two-paths*. To derive these statistics they use: (1) social circuit dependence *in addition to* Markov dependence; and (2) a non-linear functional form combining various configurations into the one statistic.

Alternating *k*-Stars

The alternating *k*-star parameter uses the second idea alone and does not extend dependence beyond the class of Markov random graph models. Rather than adopt the usual practice of limiting the number of higher order stars – as for example in model (3) where the star parameters are deliberately limited to no more than 3-stars – Snijders and colleagues returned to model (2), keeping all σ_r in the model but imposing constraints among these parameter values. Specifically, the *alternating k-star assumption* proposes that for all $k \geq 2, \sigma_{(k+1)} = -\sigma_k/\lambda$ for some λ greater than 1. Then in (1) there is one parameter σ for all star effects, with an associated statistic:

$$u = \sum_{k=2}^{n-1} (-1)^k \frac{S_k}{\lambda^{k-2}}, \tag{4}$$

where the parameter σ is referred to as the *alternating k-star parameter*. As λ is greater than 1, the direct impact of higher order stars (that is, very high degree nodes) is reduced for higher *k*. Of course, as noted above, high degree nodes still have substantial effect in the model as they produce many 2- and 3-stars, and so on. The alternating sign helps balance these overlapping star counts for high degree nodes. In [48] λ is set at 2 but [17] show how to estimate an optimal value of λ (see also [16]). The alternating *k*-star parameter is equivalent to a *geometrically weighted degree parameter* that explicitly models the degree distribution but puts more weight on the numbers of nodes with lower degrees, with weights decreasing geometrically as the degrees increase [16].

The interpretation of the parameter is as follows [41, 48]. If the alternating *k*-star parameter is positive, then highly probable networks are likely to contain some higher degree nodes. A positive alternating *k*-star parameter (together with a negative density parameter) implies graphs that exhibit preference for connections between a larger number of low degree nodes and a smaller number of higher degree nodes, akin to a core-periphery structure.

Alternating *k*-Triangles

The alternating k-triangle assumption incorporates both social circuit and Markov dependence. Snijders and colleagues show that with both these dependence assumptions operating simultaneously, configurations more complex than simple triangles are possible [48]. In particular, they introduced the notion of a *k-triangle*, a combination of *k* individual triangles that all share one edge (the common *base* of the *k* triangles), as represented in Fig. 3.

Let T_k be the count of *k*-triangles in a graph. Then the *alternating k-triangles assumption* combines these counts into the one statistic in an analogous way as for the alternating *k*-stars, that is:

$$t = 3T_1 + \sum_{k=1}^{n-3} (-1)^k \frac{T_{k+1}}{\lambda^k}, \tag{5}$$

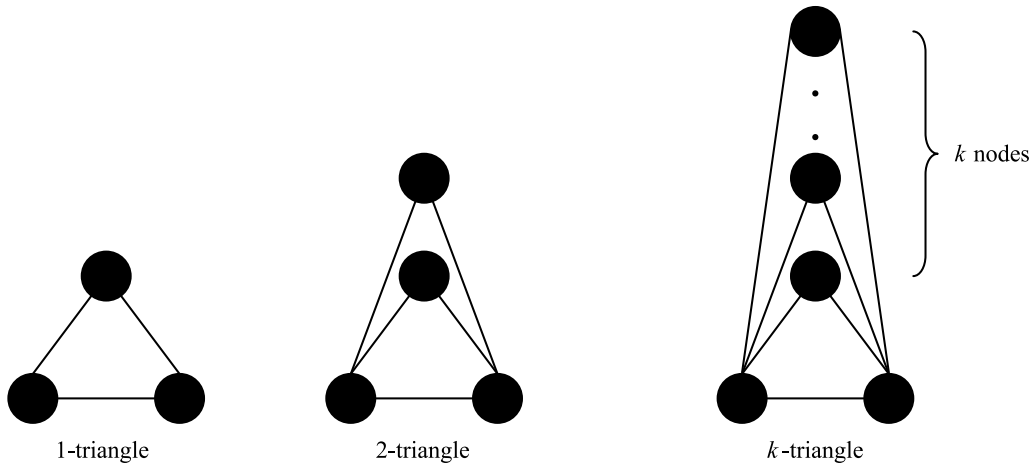
where the factor of 3 for T_1 is due to symmetry considerations in a non-directed triangle [48].

This is the *alternating k-triangle statistic* with an associated *alternating k-triangle parameter* τ. A positive *k*-triangle parameter indicates triangulation in the network but also tendencies for triangles themselves group together in larger higher order "clumps". A positive alternating *k*-triangle effect combined with a negative alternating *k*-star effect can produce a web of multiple smaller regions of triangulation [41,48].

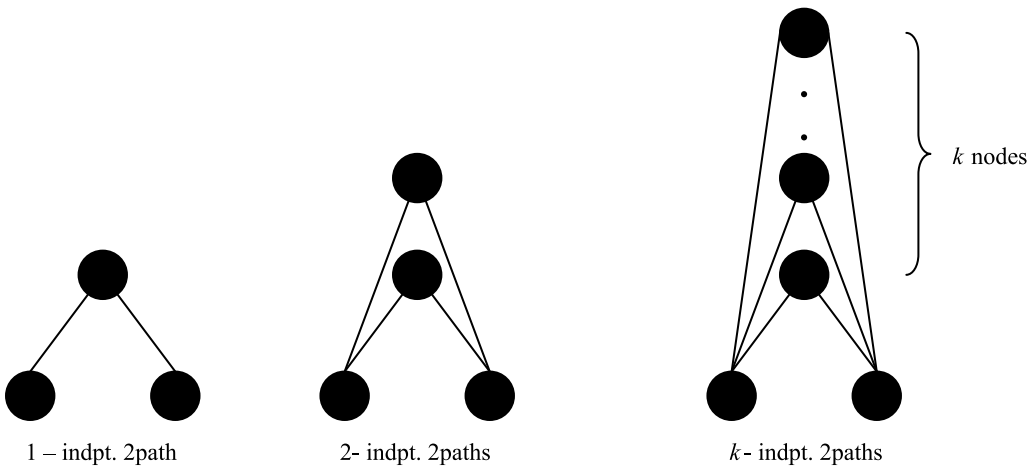
[16] shows that the alternating *k*-triangle parameter is equivalent to the *edge-wise shared partner* (or ESP) parameter that models the distribution of shared partners of tied actors, but with weights decreasing geometrically as the number of shared partners increase.

Alternating *k*-Two-Paths

Snijders and colleagues also proposed a parameter that represents a lower order configuration for a *k*-triangle, namely a *k-two-path* which is a *k*-triangle without the base [48]. These configurations, represented in Fig. 4, describe the number of distinct two-paths between a pair of nodes. The motivation was to provide a parameter that, in conjunction with *k*-triangles, would distinguish between tendencies to form edges at the base, or at the sides of a *k*-triangle.



Social Networks, Exponential Random Graph (p^*) Models for, Figure 3
Various *k*-triangles



Social Networks, Exponential Random Graph (p^*) Models for, Figure 4
Various alternating *k*-independent 2path configurations

Again the counts of these configurations are combined into the one alternating statistic in an analogous way as in the previous two cases, to produce an alternating 2-path statistic and parameter for the model. [16] shows that this new parameter is equivalent to the *dyad-wise shared partner* (or DSP) parameter that models the distribution of shared partners of actors who may or may not be tied, but with weights decreasing geometrically as the number of shared partners increase.

The parameter can be interpreted as representing localized multiple connectivity between nodes. When this parameter is negative, together with a positive alternating *k*-triangle parameter, there is a tendency against 4-cycles in the network, unless those cycles include triangles (alternatively, the presence of many 2-paths between nodes is related to the formation of triangles).

Specifications for Directed Graphs

Counterpart statistics and parameters for directed graphs have also been proposed. We do not discuss these here but refer readers to [48].

Estimation

Pseudo-likelihood estimation [52], based on logistic regression techniques, was commonly used until more principled methods became available in recent years. Preferred methods of estimation involve simulation proce-

dures (Monte Carlo Markov chain maximum likelihood estimation – MCMCMLE) which for these models has been discussed by a number of authors [4,5,6,17,45,47,56]. Software for Monte Carlo estimation techniques has recently become publicly available ([41] provides a review). The observed graph statistics are compared with those expected under a set of provisional parameter values, using a stochastic simulation based on those values. The goal is to find a set of parameter values whereby the observed statistics are equal to the expected statistics. Parameter values are refined until the observed and expected statistics can be equated, when the model is said to converge. If the parameter estimates never converge, the model is degenerate. Standard errors can also be estimated.

Goodness of Fit and Comparisons with Markov Models

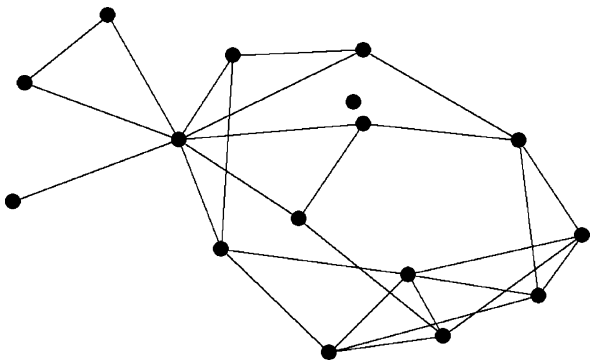
Simulation also provides an innovative means to assess how well the model can represent the data. By simulating the model from the parameter estimates, and extracting a sample of graphs it is possible to examine any graph statistic of interest (whether there is an associated parameter in the model or not). Any graph statistic for the observed graph can be compared to the distribution of such graph statistics from the simulated graphs. If the observed graph statistic is not extreme in the distribution of simulated statistics, then it is plausible that that particular feature of the observed graph could have come from the distribution of graphs implied by the model. In this way, it is possible to check which graph features the model successfully captures. This procedure is described in detail by [18] (see also [41]).

[11] provides a compelling example of how this procedure may be used to make decisions about model selection. On the basis of applying the new parametrization to a large school-based network of over 1000 nodes, [11] concludes that for this data Markov models were degenerate, but the newer parametrization had similar underlying interpretations, avoided degeneracy and was empirically validated. Other comparisons with Markov models demonstrate substantial improvements for the new parametrization in overcoming degeneracy [41].

A Simple Empirical Example

Figure 5 presents an empirical network of social relationships among 16 actors, in this case families from medieval Florence [25].

Markov models do not converge for this network. Table 1 provides parameter estimates and standard errors for models involving the social circuit parameters. The third column in the table provides a convergence statistic which indicates good convergence if it has an absolute value less than 0.1 [41]. In this case, the model exhibits excellent convergence.



Social Networks, Exponential Random Graph (p^*) Models for, Figure 5
Social relationships among 16 actors

Social Networks, Exponential Random Graph (p^*) Models for, Table 1
Parameter estimates: Florentine families network (NB: $\lambda = 2$)

Parameter	Estimate	Standard error	Convergence
Edge	−0.04	2.09	0.01
Alternating *k*-star	−1.01	0.78	0.02
Alternating *k*-triangle	0.68	0.33	0.02
Alternating *k*-2paths	0.18	0.17	0.03

A parameter may be inferred as important if its estimate in absolute terms is more than twice its standard error. So in this simple model, there is a substantial effect for alternating *k*-triangles. In Fig. 5, regions of triangulation can be seen, consistent with the alternating *k*-triangle effect, perhaps combined with the negative alternating *k*-star effect (although the reliability of that negative effect is not certain, given that the estimate is not more than twice the standard error). In sum, there is good evidence here that this network differs substantially from a simple random network, in particular, that there is substantially more triangulation in this network than would be expected in a simple random graph.

Table 2 shows the goodness of fit analysis for graph features other than those in the model. This analysis was based on extracting 1,000 sampled graphs from a simulation of 1,000,000 (after a burn-in of 100,000), using the parameter estimates in Table 1. The goodness of fit statistic takes the form of a *t*-ratio, calculated as the difference between the graph statistic and the mean from the

Social Networks, Exponential Random Graph (p^*) Models for, Table 2
Goodness of fit analysis: Florentine families network model

Graph feature	Goodness of fit
Number of 2-stars	−0.10
Number of 3-stars	−0.25
Number of triangles	−0.17
Standard deviation, degree distribution	−0.10
Skew, degree distribution	−0.12
Number of isolates	1.06
Clustering coefficient	−0.09

simulated graphs divided by the standard deviation as estimated from the sample. For graph features not in the model, the model is considered to reproduce the data well if the goodness of fit statistic is not extreme (for instance, with t-ratios of less than 2). As can be seen, on the features examined, including counts of various Markov configurations, aspects of the degree distribution, and clustering, the model reproduces the data well. There is one isolate in the data, whereas 71% of simulated graphs had no isolates, hence the rather larger goodness of fit statistic. Nevertheless, with 29% of graphs in the sample having at least one isolate, the data cannot be said to be extreme in this regard.

The clustering coefficient in the Table refers to the proportion of actual to possible triangles and is calculated as thrice the number of triangles as a ratio of the number of two-paths (the factor of three applies because there are three two-paths in any one triangle).

Table 2 does not show features of the geodesic distribution but a separate analysis reveals that the observed geodesic distribution is consistent with those from the simulated graphs: none of the geodesic quartiles of the observed graph is extreme compared to the distribution of quartiles from the simulated graphs. Overall, the model can be said to fit the data well.

Further Extensions and Future Directions

Multiple Networks

It is possible to use the same general framework as (1) when **X** is a three-way array representing multiple relations on the one set of nodes. The parametrization, however, becomes more complex. [29] showed how to develop Markov models along these lines, with these models subsequently used in some interesting empirical work, principally in organizations [22,23]. This empirical research, however, utilized pseudo-likelihood estimation and did not take into account possible model degeneracy. See also [20].

Early work on adapting the new specifications proposed by Snijders and colleagues to multiple networks has concentrated on the case of two networks, by using the alternating k-star, -triangle, and -2path parameters for effects within networks and then investigating dyadic associations between networks [57].

Bipartite Networks

Again (1) may be used as a basis for models for bipartite networks. [42] were the first to do this with Markov dependence. [28] proposed additional parameters that derived from partial conditional dependence assumptions. [54] incorporated these and other parameters, combined with an alternating combination of configuration counts analogous to those in the newer specifications, into a model for bipartite networks.

Longitudinal Networks

Perhaps the major approach to modeling network data collected at multiple time points is the actor-oriented models of Snijders and colleagues [44] which utilize a continuous time Markov chain technique. Certain specifications of actor-oriented dynamic models have exponential random graph models as their stationary distribution [44], so the links between the two approaches are quite strong. Accordingly, it is possible to develop longitudinal versions of exponential random graph models to investigate network dynamics. The conceptual difference with the actor-oriented approach is quite subtle: exponential random graph longitudinal models suppose that ties change in response to particular social neighborhoods of other ties, rather than in response to actors seeking to optimize particular structural positions. Early work along these lines, including the evolution of multiple networks, has been reported by [30]. For further discussion on the links between the actor-oriented and tie-based versions of the models, see [46].

Nodal Attributes

One of the important issues that researchers often wish to examine is whether and how nodal attributes relate to network structure. For instance, there are good empirical and theoretical grounds for expecting that people who have social relationships are more likely to be similar to each other (known as *homophily*). Conceptually, two possible processes apply: individuals may develop ties because they are similar (*social selection*) or individuals who are tied may influence each other to be similar (*social influence*).

Of course, both of these processes may occur simultaneously and it is a methodological challenge for a given set of data to disentangle the two to determine the respective strengths of selection and influence effects, respectively. Snijders and colleagues [31,49,50] have recently developed innovative methods for investigating this question for longitudinal data within actor-oriented longitudinal models.

[37] proposed social selection models within the framework of (1), having the form:

$$\Pr(\mathbf{X} = \mathbf{x}|\mathbf{Y} = \mathbf{y}) = (1/\kappa)\exp\{\Sigma_A \eta_A g_A(\mathbf{x}, \mathbf{y})\}\,, \qquad (6)$$

where $\mathbf{Y}$ is a vector of attribute variables (binary, categorical or continuous). Here the configurations A relate to combinations of nodal attributes and network subgraphs, and the network is modeled conditional on a fixed distribution of attributes. The intent of such models may be various: accounting for heterogeneity (that may otherwise cause difficulties in fitting models) as well as assessing selection, homophily, and covariate activity and popularity, while controlling for dependencies naturally occurring in the network setup.

An analogous approach to social influence models and network contagion was proposed by [38] where the patterns of attributes were modeled conditional on a fixed network:

$$\Pr(\mathbf{Y} = \mathbf{y}|\mathbf{X} = \mathbf{x}) = (1/\kappa)\exp\{\Sigma_A \eta_A g_A(\mathbf{x}, \mathbf{y})\}\,. \qquad (7)$$

Future Directions

Not only have the new specifications shown a remarkable improvement in the successful modeling of real networks, they have opened up a wide range of network modeling possibilities. The discussion of model extensions mentioned above gives some indication of current work in this area. Of course, the possibility of further improvements in model specification has not been closed. Future directions may include methods to examine very large, community-based networks, not only in terms of data collection (for instance, the development of model-based sampling methods) but also theoretical issues based on possible differences in dependence between very large scale and smaller scale structures. Model-based approaches to missing network data, perhaps within a Bayesian framework, are also in prospect [10,21].

The use of exponential random graph models in empirical research, especially in combination with multiple measures on individuals, opens a wide prospect for network statistical modeling to make a substantial contribution in many applied contexts.

Bibliography

Primary Literature

1. Besag J (1974) Spatial interaction and the statistical analysis of lattice systems. J R Stat Soc Ser B 36:96–127
2. Burda Z, Jurkiewicz J, Kryzwicki A (2004) Network transitivity and matrix models. Phys Rev E69:026106
3. Corander J, Dahmström K, Dahmström P (1998) Maximum likelihood estimation for Markov graphs, Research report. Department of Statistics, Stockholm University, Stockholm
4. Corander J, Dahmström K, Dahmström P (2002) Maximum likelihood estimation for exponential random graph models. In: Hagberg J (ed) Contributions to social network analysis, information theory and other topics in statistics: A festschrift in honour of Ove Frank. Department of Statistics, Stockholm University, Stockholm
5. Crouch B, Wasserman S (1998) Fitting p^*: Monte Carlo maximum likelihood estimation. International Conference on Social Networks, Sitges, Spain, May 28–31
6. Dahmström K, Dahmström P (1993) ML-estimation of the clustering parameter in a Markov graph model, Research report. Department of Statistics, Stockholm University, Stockholm
7. Erdös P, Rényi A (1959) On random graphs. I. Publicationes Mathematicae (Debrecen) 6:290–297
8. Frank O (1981) A Survey of Statistical Methods for Graph Analysis. Sociol Methodol 12:110–155
9. Frank O, Strauss D (1986) Markov graphs. J Am Stat Assoc 81:832–842
10. Gile K, Handcock M (2006) Model-based assessment of the impact of missing data on inference for networks. University of Washington, CSS working paper 66
11. Goodreau S (2007) Advanced in exponential random graph (p^*) models applied to a large social network. Soc Netw 29:231–248
12. Häggström O, Jonasson J (1999) Phase transition in the random triangle model. J Appl Probab 36:1101–1115
13. Handcock M (2002) Statistical models for social networks: Degeneracy and inference. In: Breiger R, Carley K, Pattison P (eds) Dynamic social network modeling and analysis. National Academies Press, Washington, pp 229–240
14. Handcock M (2003) Assessing degeneracy in statistical models of social networks. University of Washington CSS Working Paper no 39
15. Holland P, Leinhardt S (1981) An exponential family of probability distributions for directed graphs (with discussion). J Am Stat Assoc 76:33–65
16. Hunter D (2007) Curved exponential family models for social networks. Soc Netw 29:216–230
17. Hunter D, Handcock M (2006) Inference in curved exponential families for networks. J Comput Graph Stat 15:565–583
18. Hunter D, Goodreau S, Handcock M (2008) Goodness of fit of social network models. J Am Stat Assoc 103:248–258
19. Jonasson J (1999) The random triangle model. J Appl Probab 36:852–867
20. Koehly L, Pattison P (2005) Random graph models for social networks: multiple relations or multiple raters. In: Carrington P, Scott J, Wasserman S (eds) Models and methods in social network analysis. Cambridge University Press, New York, pp 162–191

21. Koskinen J (2004) Essays on Bayesian inference for social networks. Department of statistics, Stockholm University, Stockholm
22. Lazega E, Pattison P (1999) Multiplexity, generalized exchange and cooperation in organizations. Soc Netw 21:67–90
23. Lomi A, Pattison P (2006) Manufacturing relations: An empirical study of the organization of production across multiple networks. Organ Sci 17:313–332
24. Moreno J, Jennings H (1938) Statistics of social configurations. Sociometry 1:342–374
25. Padgett JF, Ansell CK (1993) Robust action and the rise of the Medici, 1400-1434. Am J Sociol 98:1259–1319
26. Park J, Newman M (2004) Solution of the 2-star model of a network. Phys Rev E70:066146
27. Pattison P, Robins G (2002) Neighborhood-based models for social networks. Sociol Methodol 32:301–337
28. Pattison P, Robins G (2004) Building models for social space: Neighbourhood based models for social networks and affiliation structures. Mathematiques des science humaines 168:11–29
29. Pattison P, Wasserman S (1999) Logit models and logistic regressions for social networks, II. Multivariate relations. Br J Math Stat Psychol 52:169–194
30. Pattison P, Robins G, Wang P, Snijders TAB, Koskinen J (2006) The co-evolution of multiple networks. Sunbelt XXVI International Social Networks Conference, Vancouver, April 2006
31. Pearson M, Steglich C, Snijders TAB (2006) Homophily and assimilation among sport-active adolescent substance users. Connections 27:51–67
32. Rapoport A (1953) Spread of information through a population with a socio-structural bias: I. assumption of transitivity. Bull Math Biophys 15:523–533
33. Rapoport A (1957) Contributions to the theory of random and biased nets. Bull Math Biophys 19:257–277
34. Robins G, Morris M (2007) Advances in exponential random graph (p^*) models. Soc Netw 29:169–172
35. Robins G, Pattison P (2005) Interdependencies and social processes: Generalized dependence structures. In: Carrington P, Scott J, Wasserman S (eds) Models and Methods in Social Network Analysis. Cambridge University Press, New York, pp 192–214
36. Robins G, Pattison P, Wasserman S (1999) Logit models and logistic regressions for social networks, III. Valued relations. Psychometrika 64:371–394
37. Robins G, Elliott P, Pattison P (2001) Network models for social selection processes. Soc Netw 23:1–30
38. Robins G, Pattison P, Elliott (2001) Network models for social influence processes. Psychometrika 66:161–190
39. Robins G, Pattison P, Woolcock J (2005) Social networks and small worlds. Am J Sociol 110:894–936
40. Robins G, Pattison, Kalish Y, Lusher D (2007) An introduction to exponential random graph (p^*) models for social networks. Soc Netw 29:173–191
41. Robins G, Snijders TAB, Wang P, Handcock M, Pattison P (2007) Recent developments in exponential random graph (p^*) models for social networks. Soc Netw 29:192–215
42. Skvoretz J, Faust K (1999) Logit models for affiliation networks. Sociol Methodol 29:253–280
43. Snijders TAB (1991) Enumeration and simulation methods for 0–1 matrices with given marginals. Psychometrika 56:397–417
44. Snijders TAB (2001) The statistical evaluation of social network dynamics. Sociol Methodol 31:361–395
45. Snijders TAB (2002) Markov chain Monte Carlo estimation of exponential random graph models. J Soc Struct 3:2
46. Snijders TAB (2006) Statistical methods for network dynamics. In: Luchini S et al (eds) Proceedings of the XLIII Scientific Meeting, Italian Statistical Society, pp 281–296
47. Snijders TAB, van Duijn MAJ (2002) Conditional maximum likelihood estimation under various specifications of exponential random graph models. In: Hagberg J (ed) Contributions to social network analysis, information theory, and other topics: A festschrift in honour of Ove Frank. Department of Statistics, University of Stockholm, Stockholm, pp 117–134
48. Snijders TAB, Pattison P, Robins G, Handcock M (2006) New specifications for exponential random graph models. Sociol Methodol 36:99–153
49. Snijders TAB, Steglich C, Schweinberger M (2007) Modeling the co-evolution of networks and behavior. In: van Montfort K, Oud H, Satorra A (eds) Longitudinal models in the behavioral and related sciences. Lawrence Erlbaum, Hillsdale, pp 41–71
50. Steglich C, Snijders TAB, West P (2006) Applying SIENA: An illustrative analysis of the coevolution of adolescents' friendship networks, taste in music, and alcolhol consumption. Methodology 2:48–56
51. Strauss D (1986) On a general class of models for interaction. SIAM Rev 28:513–527
52. Strauss D, Ikeda M (1990) Pseudo-likelihood estimation for social networks. J Am Stat Assoc 85:204–212
53. van Duijn M, Snijders TAB, Zijlstra B (2004) p2: a random effects model with covariates for directed graphs. Stat Neerlandica 58:234–254
54. Wang P (2006) Exponential random graph (p^*) models for affiliation networks. Postgraduate diploma thesis, University of Melbourne
55. Wasserman S, Pattison P (1996) Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and p^*. Psychometrika 61:401–425
56. Wasserman S, Robins G (2005) An Introduction to Random Graphs, Dependence Graphs, and p^*. In: Carrington P, Scott J, Wasserman S (eds) Models and Methods in Social Network Analysis. Cambridge University Press, New York, pp 148–161
57. Zhao Y, Robins G (2006) Multiple networks: Comparing QAP and exponential random graph (p^*) models. Sunbelt XXVI International Social Networks Conference, Vancouver, April 2006

Books and Reviews

Agneessens F, Roose H, Waege H (2004) Choices of theatre events: p^* models for affiliation networks with attributes. Metodoloski Zvezki 1:419–439

Anderson C, Wasserman S, Crouch B (1999) A p^* primer: Logit models for social networks. Soc Netw 21:37–66

Baddeley A, Möller J (1989) Nearest-neighbour Markov point processes and random sets. Int Stat Rev 57:89–121

Boer P, Huisman M, Snijders TAB, Zeggelink E (2003) StOCNET: an open software system for the advanced analysis of social networks. ProGAMMA/ICS, Groningen

Brieger R (1981) Comment on An exponential family of probability distributions for directed graphs. J Am Stat Assoc 76:51–53

Butts CT (2008) Social network analysis: A methodological introduction. Asian J Soc Psychol 11:13–41

Contractor N, Wasserman S, Faust K (2006) Testing multi-theoretical multilevel hypotheses about organizational networks: An analytic framework and empirical example. Acad Manag J 31:681–703
Frank O (1981) A survey of statistical methods for graph analysis. Sociol Methodol 11:110–155
Frank O (1991) Statistical analysis of change in networks. Stat Neerlandica 95:283–293
Frank O, Nowicki K (1993) Exploratory statistical analysis of networks. Quo Vadis, Graph Theory? Ann Discret Math 55:349–366
Handcock M, Hunter D, Butts C, Goodreau S, Morris M (2006) Statnet: An R package for the statistical analysis and simulation of social networks. University of Washington, Washington
Koehly L, Goodreau S, Morris M (2004) The link between exponential random graph models and loglinear models for networks. Sociol Methodol 34:241–270
Lazega E, van Duijn M (1997) Position in formal structure, personal characteristics and choices of advisors in a law firm: A logistic regression model for dyadic network data. Soc Netw 19:375–397
Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chkovskii D, Alon U (2002) Network motifs: simple building blocks of complex networks. Science 298:824–827
Monge P, Contractor N (2003) Theories of communication networks. Oxford University Press, New York
Newman M (2003) The structure and function of complex networks. SIAM Rev 45:167–256
Pattison P, Robins GL (2008) Probabilistic network theory. In: Rudas T (ed) Handbook of Probability: Theory with Applications. Sage, Thousand Oaks, pp 291–312
Pattison P, Wasserman S (2002) Multivariate random graph distributions: applications to social network analysis. In: Hagberg J (ed) Contributions to social network analysis, information theory and other topics in statistics: A festschrift in honour of Ove Frank. Department of Statistics, University of Stockholm, Stockholm, pp 74–100
Robins G, Pattison P (2001) Random graph models for temporal processes in social networks. J Math Sociol 25:5–41
Robins G, Pattison P, Woolcock J (2004) Models for social networks with missing data. Soc Netw 26:257–283
Schweinberger M, Snijders TAB (2003) Settings in social networks: A measurement model. Sociol Methodol 33:307–341
Wang P, Robins G, Pattison P (2006) PNet: Program for the estimation and simulation of *p*exponential random graph models, User Manual. School of Behavioural Science, University of Melbourne, Melbourne

Social Networks and Granular Computing

Churn-Jung Liau
Institute of Information Science, Academia Sinica, Taipei, Taiwan

Article Outline

Glossary
Definition of the Subject
Introduction
Mathematical Preliminaries
Rough Set Theory – Functional Granulation
Positional Analysis – Relational Granulation
Modal Logic – The Bridge
From Relational Granulation to Functional Granulation
Conclusion and Future Directions
Bibliography

Glossary

Granulation The process of drawing a set of objects (or points) together by indiscernibility, similarity, proximity, or functionality.

- **Functional granulation** The granulation process is functional if it is based on the attributes of the objects. It is called functional granulation because each attribute is a function from the set of objects to the set of values.
- **Relational granulation** The granulation process is relational if it is based on the relationships between objects.

Rough set A rough set is defined by the lower and upper approximations of a concept. The lower approximation contains all elements that necessarily belong to the concept, while the upper approximation contains those that possibly belong to the concept. In rough set theory, a concept is considered a classical set.

Social network A social network is comprised of a set of actors, called the domain, and a family of relations on the domain. It is usually represented as a graph, where each node represents an actor and an edge between two nodes represents a relational tie between these two actors. An edge can be labeled with the relation it represents.

Positional equivalence Two actors are in equivalent positions if their "pattern" of relationships with other actors is the same.

- **Structural equivalence** Two actors are structurally equivalent if they are related to the same actors.
- **Regular equivalence** Two actors are regularly equivalent if they are equally related to equivalent actors.
- **Exact equivalence** Two actors are exactly equivalent if they are related to the same number of equivalent actors.

Modal logic Modal logic was originally developed as a type of philosophical logic for reasoning about necessity and possibility. However, it has been extended to broadly cover a family of logics for reasoning about modalities including tense, obligation, belief, and knowledge. Semantically, it is also a powerful

mathematical discipline that deals with (restricted) description languages for discussing various kinds of relational structures, where a relational structure comprises a set of elements and a collection of relations on that set.

Hybrid logic Hybrid logic is a branch of modal logic that allows direct reference to the elements in a relational structure. Traditionally, only the properties of the elements could be represented by modal logic formulas.

Definition of the Subject

Granular computing (GrC) is a problem-solving concept deeply rooted in human thinking. Hence, it has played a major role in solving many important problems throughout the history of mathematics. GrC is concerned with the processing of information granules, which are groups of objects drawn together by indiscernibility, similarity, proximity, or functionality [37]. The process of forming information granules is called granulation. If the process is based totally on the attributes of the objects, it is called *functional granulation*, since attributes are mathematical functions from the set of objects to the set of values; if, in addition, the granulation process is also based on the relationship between objects, it is called *relational granulation* [11,20].

Interestingly, social scientists have applied the techniques of relational granulation (albeit by different names) to positional analysis in social networks [7,8,10,19,35]. Social network analysis (SNA) is a methodology used extensively in social and behavioral sciences, as well as in political science, economics, organization theory, and industrial engineering [15,31,34]. Positional analysis of a social network tries to find similarities between actors in the network. While many traditional clustering methods are based on the attributes of the individual actors, SNA is more concerned with the structural similarity between the actors. In SNA, a category, called a *social role* or *social position*, is defined in terms of the similarities of the patterns of relations among the actors, rather than the attributes of the actors.

Analyzing a social position from a GrC perspective has two major advantages. On the one hand, it extends the application scope of the GrC methodology. On the other hand, we can logically transform relational granulation of social networks into functional granulation; as a result, the well-developed techniques of data table analysis can be applied to social position analysis. Thus, this is an important cross-fertilization of GrC and SNA.

Introduction

GrC encompasses a wide range of techniques in computational intelligence, such as interval computing, classification, cluster analysis, and fuzzy and rough set theories. Among them, rough set theory [26,27] is one of the most well-known methods with successful applications in data analysis and uncertainty management. The theory was first proposed by Z. Pawlak as a mathematical tool to deal with vague concepts [26]. The basic assumption of the theory is that the effective use of knowledge is based on the capability to classify objects. Thus, knowledge consists of a family of classifications of a domain of interest. A classification of a domain is represented as a partition of (or an equivalence relation on) the domain. Each equivalence class of a classification is the elementary knowledge about the classification. For example, if the domain is classified according to colors, then the elementary knowledge may include "red", "green", "blue", etc. In this sense, a classification is derivable from a set of features or attributes of objects. An important task of data analysis is to define a concept via a set of features. If the concept can be precisely defined as the objects with some specific attribute values, then it is an exact concept with respect to the given attributes. However, in practice, most concepts are vague and can not be defined precisely by a given set of attributes. In such cases, rough set theory provides an effective way to approximate the vague concepts with some exact concepts in the domain. Because of its applications in data mining, rough set theory has flourished since 1990.

At about the time rough set theory was proposed, the pioneer of fuzzy set theory, L.A. Zadeh, emphasized the importance of information granularity in fuzzy reasoning [36]. Subsequently, T.Y. Lin proposed the term "GrC" to unify different techniques developed for processing information granules. From the GrC perspective, both classification in rough set theory and fuzzification in fuzzy set theory are types of granulation.

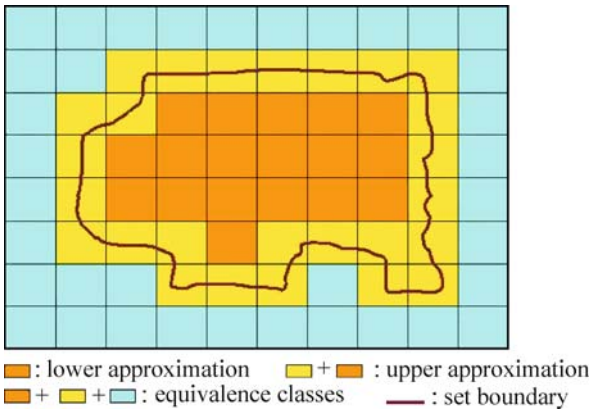
In rough set theory, objects are partitioned into equivalence classes based on their attribute values. While such values essentially represent functional information associated with the objects, more complicated information can be utilized in the classification of objects. In particular, relational information between objects can play an important role in the granulation process. The information is defined by general binary relations, which are extensions of the functional attributes of the objects. Geometrically, such granulation is derived from the neighborhood system of topological spaces [32], where each point/object is assigned at most one neighborhood/granule. This kind of granulation is called *relational granulation*, while granu-

lation based on attribute values only is called *functional granulation* [11,20].

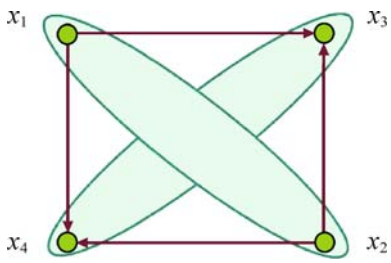
Relational granulation is especially useful for the analysis of network data because a network is a collection of relations between nodes. A concrete instance of the application of relational granulation is positional analysis of social networks. Positional analysis attempts to find actors occupying the same position in a social network based on the pattern of their relationships to other actors. Depending on the different patterns of relationships, different notions of positional equivalence have been proposed [19]. These notions can be easily regarded as instances of relational granulation. In this article, we consider three of the most important notions of positional equivalence – *structural equivalence*, *regular equivalence*, and *exact equivalence*.

Recently, it was shown that social positions based on regular equivalence can be syntactically expressed as well-formed formulas (wff) in a kind of modal logic [22]. Thus, actors occupying the same social position based on regular equivalence will satisfy the same set of modal formulas. Traditionally, modal logic has been considered the logic for reasoning about modalities, such as necessity, possibility, time, action, belief, knowledge, and obligation. However, semantically, it is essentially a language for describing relational structures [5]. A relational structure is simply a set combined with a collection of relations on that set. Thus, social networks are mathematically equivalent to relational structures. The logical characterization of social positions implies that relational granulation for positional analysis can be transformed into a functional granulation process. Indeed, since each actor in a social network may satisfy or falsify a modal formula, a modal formula can be seen as a binary attribute associated with each actor. Therefore, based on the results in [22], two actors are regularly equivalent if they have the same values with respect to all attributes defined by the language of the particular modal logic.

By extending previous results, we can find logical characterizations of different positional equivalences and transform them into a functional granulation process. Consequently, the techniques of rough set-based data analysis can also be applied to positional analysis. However, since the set of wffs of a modal logic is usually infinite, the functional granulations corresponding to positional equivalences need to consider infinite sets. To circumvent this problem, we also propose a new definition of positional equivalence, called *observational equivalence*. Let Σ be a set of wffs in a logical language that denotes the set of observable properties of the actors in a network. Two actors a and b are observationally equivalent with respect to Σ, if for each wff $\varphi \in \Sigma$, a satisfies φ iff b satisfies φ.

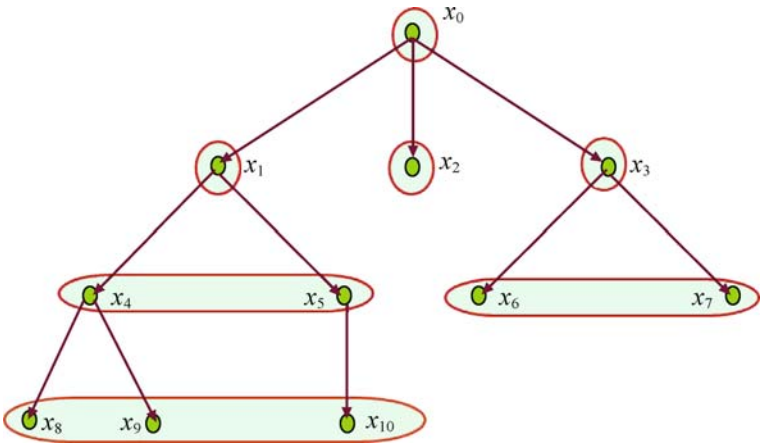


Social Networks and Granular Computing, Figure 1
A rough set: the set delimited by the *curve* (the set boundary) can not be represented exactly as a union of the basic building blocks. The *blocks* totally included in the set form the lower approximation of the set, while the *blocks* that intersect with the set (including the lower approximation) form the upper approximation of the set

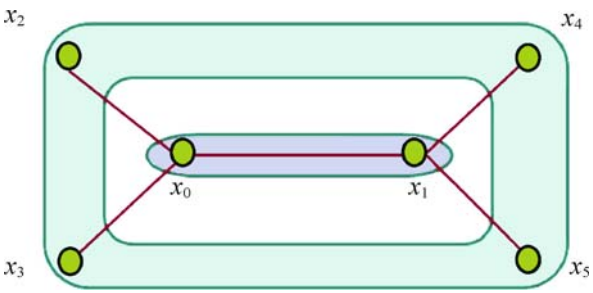


Social Networks and Granular Computing, Figure 2
Structural equivalence: the social network contains 4 actors denoted by the *nodes* and a single binary relation denoted by the *arcs*. Both the out-neighborhoods of x_1 and x_2 are $\{x_3, x_4\}$, and their common in-neighborhood is the empty set. On the other hand, both the in-neighborhoods of x_3 and x_4 are $\{x_1, x_2\}$, and their common out-neighborhood is the empty set. Thus, $x_1 \cong_s x_2$ and $x_3 \cong_s x_4$. The *green shadowed areas* indicate the equivalence classes

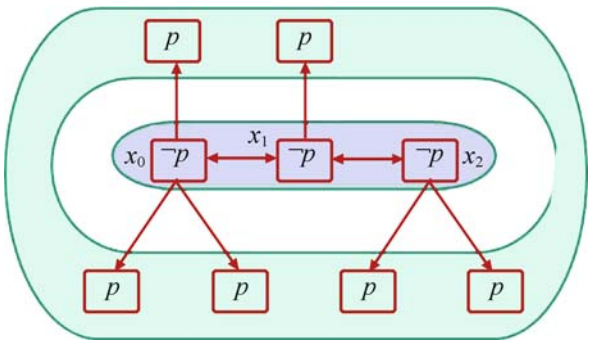
The remainder of this article is organized as follows. In Sect.“ Mathematical Preliminaries”, we introduce the background knowledge and mathematical notations of sets and relations. In Sects. “Rough Set Theory – Functional Granulation” to “Modal Logic – The Bridge”, we review rough set theory, social positional analysis, and modal logic respectively. In Sect. “From Relational Granulation to Functional Granulation”, we present the logical characterizations of three main positional equivalences and the definition of observational equivalence. Finally, in Sect. “Conclusion and Future Directions”, we present our conclusions and indicate some future research directions.



Social Networks and Granular Computing, Figure 3
Regular equivalence in a tree-structured social network: the root x_0 is discernible from other nodes because its in-neighborhood is empty, whereas all the other nodes have nonempty in-neighborhoods. Nodes x_1, x_2, x_3 are discernible from other nodes because they are the only nodes connected to x_0. Moreover, they are mutually discernible, since x_2 has no children, x_1 has children and grand-children, and x_3 has children, but not grand-children. Note that x_4 and x_5 are indiscernible, since they have the same in-neighborhood and equivalent out-neighborhoods; x_6 and x_7 are indiscernible, since they have the same in-neighborhood and both have empty out-neighborhoods; and x_8, x_9, and x_{10} are indiscernible, since they have equivalent in-neighborhoods and all have empty out-neighborhoods



Social Networks and Granular Computing, Figure 4
Exact equivalence: the social network is a undirected graph. In other words, the binary relation is symmetry. The exact equivalence partitions the nodes into two classes, $C_0 = \{x_0, x_1\}$ and $C_1 = \{x_2, x_3, x_4, x_5\}$. To verify that it is indeed an exact equivalence, we can check that both x_0 and x_1 connect to one C_0 node and two C_1 nodes, and all of x_2, x_3, x_4, and x_5 connect to one C_0 node



Social Networks and Granular Computing, Figure 5
A social network represented as a Kripke model: we assume the underlying modal language contains one propositional symbol p and one relational symbol r. Since $\{x_0, x_1, x_2\}$ is an equivalence class in the shown regular equivalence, the three nodes all satisfy the same set of PMML wffs. For example, they all satisfy $\neg p \wedge \langle r \rangle p$. However, they are not exactly equivalent, so they are distinguishable by GML wffs. For example, x_0 satisfies both $\langle r \rangle_1 p$ and $\langle r \rangle_2 p$, x_1 does not satisfy $\langle r \rangle_1 p$ or $\langle r \rangle_2 p$, and x_2 satisfies $\langle r \rangle_1 p$ but not $\langle r \rangle_2 p$

Mathematical Preliminaries

In this section, we introduce the basic knowledge about sets and relations used in this article.

Sets

A set is a collection of objects, called the elements of the set. We use capitals A, B, U, X, etc. to denote sets. The elements of a set are denoted by lower-case letters and we write $x \in U$ if x is an element of the set U. We usually use two notations to denote a set. The first one lists its elements. For example, we can write $\{x_1, x_2, \ldots, x_n\}$ for a finite set, where n is a positive integer, and $\{x_1, x_2, \ldots\}$ for an infinite set. The second notation specifies the defining property of the set. For example, we can write $\{x \mid x = 2n, n \in Z\}$ for the set of even numbers. The

number of elements in a set is called its cardinality. The cardinality of U is denoted by $|U|$. A set U is called an empty set if $|U| = 0$, and a singleton if $|U| = 1$. We denote the empty set by $\emptyset$.

The basic operations on sets are defined as follows:

1. Intersection of X and Y: $X \cap Y = \{u \mid u \in X \text{ and } u \in Y\}$,
2. Union of X and Y: $X \cup Y = \{u \mid u \in X \text{ or } u \in Y\}$,
3. Cartesian product of X and Y: $X \times Y = \{(x, y) \mid x \in X \text{ and } y \in Y\}$, and
4. Difference of X by Y: $X - Y = \{u \mid u \in X \text{ and } u \notin Y\}$.

Since intersection, union, and the Cartesian product are associative (i. e., $(X \cap Y) \cap Z = X \cap (Y \cap Z)$ etc.), we can eliminate the parentheses and apply the operations to more than two sets. For example, the Cartesian product of the sets $U_1, U_2, \ldots, U_k$ is denoted by

$$U_1 \times U_2 \times \cdots \times U_k = \{(x_1, x_2, \ldots, x_k) \mid x_i \in U_i, \forall 1 \leq i \leq k\},$$

where each element is called a k-tuple. A 2-tuple is also called a pair. If $U_1 = U_2 = \cdots = U_k = U$, the Cartesian product is written as U^k. A set X is said to be a subset of another set Y, denoted by $X \subseteq Y$, if for all $x \in X$, it is also the case that $x \in Y$. A function f from the set X to the set Y, denoted by $f: X \to Y$, is a mapping that associates each element in X with an element in Y. The set X is called the domain of the function, and Y is called the range of f. For each $x \in X$, we write $f(x)$ for the element in Y that is associated with x by f and call it the image of x under f. Also, the image of X under f is defined as $f(X) = \{f(x) \mid x \in X\}$.

Multisets

For sets, repeat occurrences of the same elements are only counted once. Thus, the set $\{x, x\}$ is the same as the set $\{x\}$. Sometimes, the number of occurrences of an element is important. In such cases, we consider multisets (or bags). A multiset is formally defined as a pair $M = (U, m)$, where U is some set and $m: U \to N$ is a function from U to the set $N = \{0, 1, 2, 3, \ldots\}$ of natural numbers. U is called the universe and for each element $x \in U$, $m(x)$ is the multiplicity (that is, the number of occurrences) of x in M. A set can be considered as a special multiset in which $m(x) = 0$ or 1 for each x in the universe. By generalizing the notation of set membership, we can write $x \in M$ if $m(x) > 0$. The cardinality of M is defined as $|M| = \sum_{x \in U} m(x)$. Let $M_1 = (U, m_1)$ and $M_2 = (U, m_2)$ be two multisets in the same universe; then, $M_1 \subseteq M_2$ iff $m_1(x) \leq m_2(x)$ for all $x \in U$. The intersection, union, and difference between $M_1 = (U, m_1)$ and $M_2 = (U, m_2)$ are defined as follows:

1. $M_1 \cap M_2 = (U, m)$, where $m(x) = \min(m_1(x), m_2(x))$ for all $x \in U$,
2. $M_1 \cup M_2 = (U, m)$, where $m(x) = \max(m_1(x), m_2(x))$ for all $x \in U$,
3. $M_1 - M_2 = (U, m)$, where $m(x) = (m_1(x) - m_2(x)) \cdot \chi(m_1(x) - m_2(x))$ for all $x \in U$, and

$$\chi(k) = \begin{cases} 1, & \text{if } k > 0, \\ 0, & \text{if } k \leq 0. \end{cases}$$

The Cartesian product of $M_1 = (U_1, m_1)$ and $M_2 = (U_2, m_2)$ is defined as $M_1 \times M_2 = (U_1 \times U_2, m)$, where $m(x, y) = m_1(x)m_2(y)$ for all $x \in U_1$ and $y \in U_2$.

Relations

In mathematics, a subset of $U_1 \times U_2 \times \cdots \times U_k$ is called a k-ary relation among the sets $U_1, U_2, \ldots, U_k$, and a subset of U^k is called a k-ary relation on U. We are particularly interested in binary relations. According to the definition, a binary relation on U is a set of pairs of elements in U. A pair (x, y) in a binary relation means that x is related to y by the relation. We use lower-case Greek letters, α, β, ρ, etc. to denote relations. The basic operations on binary relations include the three aforementioned set-theoretic operations – intersection, union, and difference, as well as converse and composition. The converse of a binary relation $\alpha \subseteq U_1 \times U_2$ is $\alpha^- \subseteq U_2 \times U_1$ such that

$$\alpha^- = \{(x, y) \mid (y, x) \in \alpha\}.$$

The composition of two binary relations $\alpha \subseteq U_1 \times U_2$ and $\beta \subseteq U_2 \times U_3$ is $\alpha\beta \subseteq U_1 \times U_3$ such that

$$\alpha\beta = \{(x, y) \mid \exists z \in U, (x, z) \in \alpha \wedge (z, y) \in \beta\}.$$

A binary relation $\rho \subseteq U^2$ is said to be reflexive if $(x, x) \in \rho$ for all $x \in U$; ρ is symmetric if for all $x, y \in U$, $(x, y) \in \rho$ implies $(y, x) \in \rho$; ρ is anti-symmetric if for all $x, y \in U$, $(x, y) \in \rho$ and $(y, x) \in \rho$ implies $x = y$; and ρ is transitive if for all $x, y, z \in U$, $(x, y) \in \rho$ and $(y, z) \in \rho$ implies $(x, z) \in \rho$. A binary relation that is reflexive, anti-symmetric, and transitive is called a partial order. Let $\leq$ be a partial order on the universe U and $X \subseteq U$, then, $u \in U$ is called an upper bound (resp. lower bound) of X if for all $x \in X$, $x \leq u$ (resp. $u \leq x$). The least upper bound (resp. the greatest lower bound) of X is the upper bound (resp. lower bound) u of X such that for any upper bounds (resp. lower bounds) u' of X, $u \leq u'$ (resp. $u' \leq u$). The least upper bound (resp. the greatest lower bound) of X is called

its supremum (resp. infimum), and is denoted by $\sup(X)$ (resp. $\inf(X)$). For any X (even though X is finite), $\sup(X)$ and $\inf(X)$ do not necessarily exist. If for all $x, y \in U$, $\sup(\{x, y\})$ and $\inf(\{x, y\})$ exist, then $\sup(\{x, y\})$ is called the join of x and y and denoted by $x \sqcup y$, while $\inf(\{x, y\})$ is called the meet of x and y and denoted by $x \sqcap y$. The structure $(U, \sqcap, \sqcup)$ is called a lattice. A subset $X \subseteq U$ is a sublattice of U if for any $x, y \in X$, $x \sqcap y$ and $x \sqcup y \in X$.

A binary relation that is reflexive, symmetric, and transitive is called an equivalence relation, which yields a partition of the universe. A partition of U is a class of subsets of U, $\{X_i \mid i \in I\}$ for some index set I, where $\bigcup_{i \in I} X_i = U$ and $X_i \cap X_j = \emptyset$ for any $i \neq j$. Let ρ be an equivalence relation on U, and let $x \in U$; then, the equivalence class containing x is defined as

$$[x]_\rho = \{y \in U \mid (x, y) \in \rho\} .$$

Obviously, $\{[x]_\rho \mid x \in U\}$ forms a partition of the universe. For any $X \subseteq U$, we denote $[X]_\rho$ by the set $\{[x]_\rho \mid x \in X\}$ and $[\![X]\!]_\rho$ by the multiset $([U]_\rho, m)$ such that $m([x]_\rho) = |[x]_\rho \cap X|$ for all $x \in U$.

Equivalence relations on a fixed domain are partially ordered by set-inclusion. Let ρ_1 and ρ_2 be two equivalence relations and define $\rho_1 \leq \rho_2$ iff $\rho_1 \subseteq \rho_2$; then, $\leq$ is a partial order on the set of all equivalence relations on a fixed domain. We say that ρ_1 is finer than ρ_2 or ρ_2 is coarser than ρ_1 if $\rho_1 \leq \rho_2$. Given a binary relation α, the composition of α with itself k times is denoted by α^k and the transitive closure of α is defined as $\alpha^\infty = \bigcup_{k \geq 1} \alpha^k$. Let ρ_1 and ρ_2 be two equivalence relations; then it can be shown that $\rho_1 \cap \rho_2$ and $(\rho_1 \cup \rho_2)^\infty$ are, respectively, the meet and the join of ρ_1 and ρ_2.

Rough Set Theory – Functional Granulation

The basic construct of rough set theory is the approximation space, which is a pair (U, ρ), where U is a finite set of objects (the universe) and ρ is an equivalence relation on U. From the GrC perspective, the equivalence classes of ρ are information granules of the approximation space.

In rough set theory, a concept is represented as a subset of the universe. For any concept $X \subseteq U$, we can associate the following two subsets with X,

$$\underline{\rho}X = \{x \in U \mid [x]_\rho \subseteq X\}$$
$$\overline{\rho}X = \{x \in U \mid [x]_\rho \cap X \neq \emptyset\} ,$$

where $\underline{\rho}X$ and $\overline{\rho}X$ are called the ρ-lower and ρ-upper approximation of X, respectively. From a practical viewpoint, ρ can be considered as an indiscernibility relation; thus, for a given concept X, we can only know that X contains all the elements in $\underline{\rho}X$ and does not contain any element outside $\overline{\rho}X$. In other words, the concept X cannot be defined exactly in the approximation space. The pair $(\underline{\rho}X, \overline{\rho}X)$ is considered a rough approximation of X, and any such pair is called a rough set (see Fig. 1). A concept X is called an exact concept in the approximation space if $\underline{\rho}X = X = \overline{\rho}X$; otherwise, it is called a rough concept.

Rough set theory is very useful in the analysis of data tables. A data table is a pair $S = (U, F)$, where U is a nonempty, finite set (the universe) and F is a nonempty, finite set of primitive attributes. Every $f \in F$ is a total function $f: U \to V_f$, where V_f denotes the possible values of f. If $E = \{f_1, \ldots, f_n\} \subseteq F$, then $V_E = V_{f_1} \times \cdots \times V_{f_n}$. Thus, E is also considered a function from U to V_E. An equivalence relation $\mathrm{IND}(E)$ is associated with every subset of attributes $E \subseteq F$, and defined by

$$(x, y) \in \mathrm{IND}(E) \Leftrightarrow f(x) = f(y) \quad \forall f \in E .$$

$\mathrm{IND}(E)$ is called an indiscernibility relation. We write $\mathrm{IND}(f)$ instead of $\mathrm{IND}(\{f\})$ for all $f \in F$. It is easy to show that $\mathrm{IND}(E) = \bigcap_{f \in E} \mathrm{IND}(f)$. Since $\mathrm{IND}(E)$ is an equivalence relation, $(U, \mathrm{IND}(E))$ is an approximation space and we can define the $\mathrm{IND}(E)$-lower and $\mathrm{IND}(E)$-upper approximation of X for any $X \subseteq U$. We write $\underline{E}X$ and $\overline{E}X$ instead of $\underline{\mathrm{IND}(E)}X$ or $\overline{\mathrm{IND}(E)}X$. Obviously, exact concepts in $(U, \mathrm{IND}(E))$ are those definable using only attributes in E. In effect, the indiscernibility relation $\mathrm{IND}(E)$ granulates (or partitions) the universe according to the functions in E; thus, it results in a functional granulation of the universe.

The definitions are used in the analysis of dependency between attributes in a data table. Let us say that attribute E_2 depends on E_1, denoted by $E_1 \Rightarrow E_2$, iff $\mathrm{IND}(E_1) \subseteq \mathrm{IND}(E_2)$, i. e., any two objects in U with the same attribute values in E_1 will also have the same attribute values in E_2. It is easy to show that $E_1 \Rightarrow E_2$ iff $\underline{E_1}X = X$ for all X that are equivalence classes of $\mathrm{IND}(E_2)$. Given an attribute $f \in F$, if $F - \{f\} \Rightarrow \{f\}$, then f can be eliminated from F without influencing the definability of F. In other words, by using the set of attributes, f is redundant for classification purposes. By repeated elimination of redundant attributes, we can eventually obtain an irreducible set of attributes with the same classification capability as the original set of attributes. Such an irreducible set of attributes is called a reduct in rough set theory. Note that the definition of dependency is relative to a data table, so it does not necessarily mean that an intrinsic connection exists between the inter-dependent attributes.

Positional Analysis – Relational Granulation

Social networks are defined by actors and relations (or nodes and edges in terms of graph theory) [15]. A social network is generally defined as a relational structure $\mathfrak{N} = (A, (\alpha_i)_{i\in I})$, where A is the set of actors in the network, I is an index set, and for each $i \in I$, $\alpha_i \subseteq A^{k_i}$ is a k_i-ary relation on the domain A. If $k_i = 1$, then α_i is also called an attribute. In practice, most SNA literature considers a simplified version of social networks with only binary relations. For ease of presentation, we focus on a social network with only unary and/or binary relations. Thus, the social network considered in this article is a structure $\mathfrak{N} = (A, (P_i)_{i\in I}, (\alpha_j)_{j\in J})$, where A is a *finite set* of actors, $P_i \subseteq A$ for all $i \in I$, and $\alpha_j \subseteq A \times A$ for all $j \in J$. In terms of graph theory, $\mathfrak{N}$ is a labeled graph, where A is a set of nodes labeled with subsets of I, and each α_j denotes a set of (labeled) edges. For each $a \in A$, the out-neighborhood and in-neighborhood of a with respect to a binary relation α, denoted respectively by $N^+_\alpha(a)$ and $N^-_\alpha(a)$, are defined as follows:

$$N^+_\alpha(a) = \{b \in A \mid (a, b) \in \alpha\} ,$$
$$N^-_\alpha(a) = \{b \in A \mid (b, a) \in \alpha\} .$$

If ρ is an equivalence relation on A and a is an actor, the ρ-equivalence class of a is equal to its neighborhood, i. e., $[a]_\rho = N^+_\rho(a) = N^-_\rho(a)$. Note that the latter equality holds because of the symmetry of ρ.

Several equivalence relations have been proposed for exploring the similarity between actors' roles. The simplest definition of positional equivalence is the concept of structural equivalence proposed in [21], which states that two actors are positionally equivalent if they are related to the same individuals (see Fig. 2).

Definition 1 Let $\mathfrak{N} = (A, (P_i)_{i\in I}, (\alpha_j)_{j\in J})$ be a social network and ρ be an equivalence relation on A; then ρ is a (strong) structural equivalence with respect to $\mathfrak{N}$ if $(a, b) \in \rho$ implies

1. $a \in P_i$ iff $b \in P_i$ for all $i \in I$, and
2. $N^+_{\alpha_j}(a) = N^+_{\alpha_j}(b)$ and $N^-_{\alpha_j}(a) = N^-_{\alpha_j}(b)$ for all $j \in J$.

By the definition, there may exist more than one structural equivalence for a given network. However, it has been shown that there always exists a maximum (i. e., coarsest) structural equivalence for a network [19]. In fact, the set of all structural equivalences form a sublattice of all equivalence relations on A. Indeed, if ρ_1 and ρ_2 are structural equivalences with respect to $\mathfrak{N}$, so too are $\rho_1 \sqcap \rho_2$ and $\rho_1 \sqcup \rho_2$. Since the set of all equivalence relations on a finite domain is finite, the join of all structural equivalences is the coarsest structural equivalence with respect to a network. Let $\mathfrak{N} = (A, (P_i)_{i\in I}, (\alpha_j)_{j\in J})$ be a social network and ρ be the coarsest structural equivalence with respect to $\mathfrak{N}$; then, two actors $x, y \in A$ are said to be structurally equivalent, denoted by $x \cong_s y$, if $(x, y) \in \rho$.

Although structural equivalence is conceptually simple, it is sometimes restrictive. Regular equivalence relaxes the requirement that equivalent actors must be connected with identical actors, and suggests that actors occupy the same position if they are connected to positionally equivalent actors. Regular equivalence has been studied extensively [7,8,10,19,35], and there are several alternative definitions of the concept. We present two here. The first is based on the characterization given by Boyd and Everett [8], which states that an equivalence relation ρ is a *regular equivalence* with respect to a binary relation α if it commutes with α, i. e.

$$\alpha\rho = \rho\alpha .$$

By this definition, if ρ is a regular equivalence with respect to α and $(a, b) \in \rho$, then for each $c \in N^+_\alpha(a)(\text{resp.}N^-_\alpha(a))$, there exists $c' \in N^+_\alpha(b)(\text{resp.}N^-_\alpha(b))$ such that $(c, c') \in \rho$. The property naturally leads to an alternative definition of regular equivalence [19], which states that an equivalence relation ρ is a *regular equivalence* with respect to a binary relation α if for $a, b \in A$,

$$(a, b) \in \rho \Rightarrow [N^+_\alpha(a)]_\rho = [N^+_\alpha(b)]_\rho \quad \text{and} \quad [N^-_\alpha(a)]_\rho = [N^-_\alpha(b)]_\rho .$$

According to this definition, if a and b are regularly equivalent, then they are connected to equivalent neighborhoods (see Fig. 3). Obviously, the above definitions are equivalent. Thus, we have the following definition.

Definition 2 Let $\mathfrak{N} = (A, (P_i)_{i\in I}, (\alpha_j)_{j\in J})$ be a social network and ρ be an equivalence relation on A; then ρ is a regular equivalence with respect to $\mathfrak{N}$ if

1. $(a, b) \in \rho$ implies $a \in P_i$ iff $b \in P_i$ for all $i \in I$; and
2. ρ is a regular equivalence with respect to α_j for all $j \in J$.

Analogous to the case of structural equivalences, the join of two regular equivalences is still a regular equivalence, even though their meet is not necessarily a regular equivalence. Consequently, we can find the coarsest regular equivalence of a network. Then, two actors, x and y, are regularly equivalent, denoted by $x \cong_r y$, if (x, y) is in the coarsest regular equivalence of the network.

For regular equivalence, only the occurrence or non-occurrence of a position in the neighborhood of an actor matters. However, the number of occurrences is sometimes an important factor in positional analysis. In such cases, a number restriction can be added to the definition of regular equivalences. An equivalence relation ρ is an *exact equivalence* with respect to a binary relation α if for $a, b \in A$,

$$(a,b) \in \rho \Rightarrow [\![N_\alpha^+(a)]\!]_\rho = [\![N_\alpha^+(b)]\!]_\rho \quad \text{and} \quad [\![N_\alpha^-(a)]\!]_\rho = [\![N_\alpha^-(b)]\!]_\rho .$$

By replacing the set equality in the definition with the multiset equality, the number of equivalent neighbors must be the same for two actors to be considered position-equivalent (see Fig. 4).

Definition 3 Let $\Re = (A, (P_i)_{i\in I}, (\alpha_j)_{j\in J})$ be a social network and ρ be an equivalence relation on A; then, ρ is an exact equivalence with respect to $\Re$ if

1. $(a,b) \in \rho$ implies $a \in P_i$ iff $b \in P_i$ for all $i \in I$; and
2. ρ is an exact equivalence with respect to α_j for all $j \in J$.

The set of all exact equivalences also forms a lattice, but it is not a sublattice of the set of all equivalence relations [10]. Consequently, the coarsest exact equivalence for a network exists and we can define two actors x and y as exactly equivalent, denoted by $x \cong_e y$, if they belong to the coarsest exact equivalence. The partition produced by an exact equivalence is called an equitable partition or a divisor of a graph [19].

From the above definition, it is clear that all the attributes in a social network are binary. In other words, for each attribute P_i, an actor either satisfies P_i (has the value 1) or falsifies P_i (has the value 0). When a social network $\Re = (A, (P_i)_{i\in I}, (\alpha_j)_{j\in J})$ has the property $J = \emptyset$, it is in fact a kind of data table. In this case, all three kinds of positional equivalence are reduced to the indiscernibility relation. This means that, if information about the actors' relationship is not available, social positions are fully determined by the attributes of the actors. Thus, without relational information, functional granulation is sufficient for positional analysis. However, if relational information is available, two actors can be further differentiated by their relationship with other actors, even though all their attributes have the same values. In this sense, the definition of positional equivalence is based on the process of relational granulation.

Modal Logic – The Bridge

Modal logics were originally developed as formalizations for reasoning about modalities. The formal study started with alethic modal logics (the logics of necessity and possibility). The modalities $\Box$ (for necessity) and $\Diamond$ (for possibility) have now become standard notations in modal logic literature, and many variants of modal logic have been proposed to deal with different classes of modalities; for example, temporal logic for tense operators, deontic logic for obligation and permission, dynamic logic for actions, and epistemic logic for beliefs and knowledge. The development of these logics was motivated by philosophical enquiry as well as by technical applications in computer science, artificial intelligence, and economic game theory.

While modal logics were initially presented in the form of reasoning systems, the invention of relational semantics has been influential in the continuing development of the field [17,18]. The semantics, proposed independently by J. Hintikka, S. Kanger, and S. Kripke, and now known as Kripke semantics, show that modal logics are in fact logics for reasoning about relational structures. From this semantic perspective, standard modal logics can be regarded as fragments of first-or second-order predicate logics, whereby the necessity and possibility modalities correspond to universal and existential quantifiers, respectively. In spite of this correspondence, quantification in modal logic tends to be bounded in some way to worlds that are "relevant to" or "accessible from" the current one. Consequently, although a number of properties of modal logics follow immediately from those of their classical quantificational counterparts, the modal operators typically have less expressive power than full quantification. This results in many interesting properties not available in classical predicate logic. One of the most striking results is that semantic invariances between models are actually various forms of *bisimulation*, which preserve the local properties of worlds and their transition patterns [5]. Interestingly, it has been shown that bisimulation in Kripke models corresponds exactly to regular equivalence in social networks [22]. The implication of the results is that position-equivalent actors can be characterized by an appropriate set of modal formulas. As each modal formula can be regarded as a binary attribute, from the GrC perspective, such characterization transforms relational granulation into functional granulation.

In this section, we present three modal logics – a multimodal logic, a graded modal logic, and a hybrid logic. The presentation of a modal logic, as with any other formal logic, requires the specification of its syntax and semantics. The syntactic aspect includes the language of the logic – its

alphabet and formation rules for wffs, as well as a deductive system for logical reasoning. For the purpose of this article, we need only present the language part. For the semantic aspect, we have to define the models within which the wffs can be interpreted, as well as the satisfaction of a wff in a model.

Propositional Multi-Modal Logic

We start with propositional modal logic (PML) [9]. The alphabet of PML consists of a set of propositional symbols, PV, and the logical symbols $\neg$(negation), $\wedge$(and), $\vee$(or), $\supset$(material implication), and $\Diamond$(possibility); $\neg$, $\wedge$, $\vee$, and $\supset$ are called Boolean connectives, whereas $\Diamond$ is the only primitive modality of PML. The set of wffs of PML is the smallest set containing PV that satisfies the following conditions:

- if φ is a wff, then $\neg\varphi$ and $\Diamond\varphi$ are wffs,
- if φ and ψ are wffs, then $\varphi \wedge \psi, \varphi \vee \psi$, and $\varphi \supset \psi$ are wffs.

As usual, we abbreviate $(\varphi \supset \psi) \wedge (\psi \supset \varphi)$ as $\varphi \equiv \psi$; $\neg \Diamond \neg\varphi$ as $\Box\varphi$; any tautology $p \vee \neg p$ as $\top$; and any contradiction $p \wedge \neg p$ as $\bot$. A Kripke model for PML is a triple $\mathfrak{M} = (W, \alpha, V)$, where W is a set of possible worlds, α is a binary relation on W, called an accessibility relation, and $V: W \times PV \to \{0, 1\}$ is a truth assignment for evaluating the truth value of each propositional symbol in each world. The satisfaction of a wff φ in a world w of the model $\mathfrak{M}$, denoted by $\mathfrak{M}, w \models \varphi$, is defined by the following clauses:

1. $\mathfrak{M}, w \models p$ if $V(w, p) = 1$ for each $p \in PV$;
2. $\mathfrak{M}, w \models \neg\varphi$ iff $\mathfrak{M}, w \not\models \varphi$;
3. $\mathfrak{M}, w \models \varphi \wedge \psi$ iff $\mathfrak{M}, w \models \varphi$ and $\mathfrak{M}, w \models \psi$;
4. $\mathfrak{M}, w \models \varphi \vee \psi$ iff $\mathfrak{M}, w \models \varphi$ or $\mathfrak{M}, w \models \psi$;
5. $\mathfrak{M}, w \models \varphi \supset \psi$ iff $\mathfrak{M}, w \not\models \varphi$ or $\mathfrak{M}, w \models \psi$;
6. $\mathfrak{M}, w \models \Diamond\varphi$ iff there exists $(w, u) \in \alpha$ such that $\mathfrak{M}, u \models \varphi$.

Propositional multi-modal logic (PMML) is an extension of PML that allows more than one modality [16]. PMML is particularly suitable for reasoning about relational structures that contain a number of binary relations, e. g., social networks. The alphabet of PMML consists of a set of propositional symbols PV, a set of relational symbols REL, the Boolean connectives, the relational converse symbol $^-$, and the modality-forming symbol $\langle\rangle$. The set of wffs of PMML is the smallest set containing PV that satisfies the following conditions:

- if φ is a wff, then $\neg\varphi$ is a wff;
- if φ is a wff and r is a relational symbol, then $\langle r\rangle\varphi$ and $\langle r^-\rangle\varphi$ are wffs;
- if φ and ψ are wffs, then $\varphi \wedge \psi, \varphi \vee \psi$, and $\varphi \supset \psi$ are wffs.

We abbreviate $\neg\langle r\rangle\neg\varphi$ and $\neg\langle r^-\rangle\neg\varphi$ as $[r]\varphi$ and $[r^-]\varphi$, respectively. A Kripke model for PMML is $\mathfrak{M} = (W, (\alpha_r)_{r\in REL}, V)$, where W and V are the same as above, and for each $r \in REL$, α_r is a binary relation on W. The satisfaction of PMML wffs is defined in the same way as that of PML wffs, except that clause 6 is replaced by

6. $\mathfrak{M}, w \models \langle r\rangle\varphi$ iff there exists $(w, u) \in \alpha_r$ such that $\mathfrak{M}, u \models \varphi$;
7. $\mathfrak{M}, w \models \langle r^-\rangle\varphi$ iff there exists $(u, w) \in \alpha_r$ such that $\mathfrak{M}, u \models \varphi$.

Note that we need the converse modalities $\langle r^-\rangle$ because all positional equivalences are defined with respect to their in-neighborhoods and out-neighborhoods. The modalities $\langle r\rangle$ allow us to access the out-neighborhood of a world, whereas $\langle r^-\rangle$ is needed to access the in-neighborhood.

Graded Modal Logic

Graded modal logic (GML) is a generalization of PMML that allows us to consider the number of worlds accessible from a given world [33]. The alphabet of GML is the same as that of PMML; however, the second formation rule of wffs is replaced by the following rule:

- if φ is a wff and r is a relational symbol, then $\langle r\rangle_n\varphi$ and $\langle r^-\rangle_n\varphi$ are wffs for each natural number n.

We also write $[r]_n\varphi$ and $[r^-]_\varphi$ as the abbreviations of $\neg\langle r\rangle_n\neg\varphi$ and $\neg\langle r^-\rangle_n\neg\varphi$, respectively.

The Kripke models for GML are the same as those for PMML, but clauses 6 and 7 for the conditions of satisfaction of PMML wffs are modified as follows:

6. $\mathfrak{M}, w \models \langle r\rangle_n\varphi$ iff $|\{u \in W \mid (w, u) \in \alpha_r \text{ and } \mathfrak{M}, u \models \varphi\}| > n$;
7. $\mathfrak{M}, w \models \langle r^-\rangle_n\varphi$ iff $|\{u \in W \mid (u, w) \in \alpha_r \text{ and } \mathfrak{M}, u \models \varphi\}| > n$.

According to the semantics, the modalities $\langle r\rangle_0$ and $\langle r^-\rangle_0$ in GML are respectively equivalent to $\langle r\rangle$ and $\langle r^-\rangle$ in PMML. Intuitively, $\mathfrak{M}, w \models \langle r\rangle_n\varphi$ means that φ is true in more than n worlds that are r-accessible from w, whereas $\mathfrak{M}, w \models [r]_n\varphi$ means that φ is false in no more than n worlds that are r-accessible from w. Thus, we abbreviate $\langle r\rangle_{n-1}\varphi \wedge [r]_n\neg\varphi$ (resp. $\langle r^-\rangle_{n-1}\varphi \wedge [r^-]_n\neg\varphi$) as $(r)_n\varphi$ (resp. $(r^-)_n\varphi$) to denote that there are exactly n r-accessible (resp. r^--accessible) worlds that satisfy φ.

Hybrid Logic

Hybrid logics (HL) are extensions of modal logics that include symbols to name possible worlds in models. The development of HL dates back to A. Prior's 1951 work, which had a philosophical foundation [29,30] (see [1] for a brief overview). The main idea is to use a special kind of atomic formula, called a nominal, to refer to possible worlds in models. The idea was adopted by the Sofia school in the late 80s and early 90s [14,24,25]. Subsequently, Blackburn and Seligman, who have been influential in the recent development of HL, adopted it in the late 90s [4,6].

Here, we only need the simplest hybrid logic (SHL). The alphabet of SHL is the same as that of PMML plus a set of nominals NOM that is disjoint with the set PV. The set of wffs of SHL is the smallest set containing $PV \cup NOM$ that satisfies the formation rules of PMML. The key semantic idea of HL is that each nominal must be true in exactly one possible world in any model. In other words, a nominal names a world by being true in that world and nowhere else. The idea can be easily formalized by extending PMML models with a component that assigns a nominal to a possible world. A hybrid model is then $\mathfrak{M} = (W, (\alpha_r)_{r\in REL}, V, f)$, where $(W, (\alpha_r)_{r\in REL}, V)$ is a PMML model and $f: NOM \to W$ is the mapping that assigns a nominal to a possible world. The satisfaction of SHL wffs is defined by clauses 1–7 for PMML, with the addition of the following clause for nominals:

8. $\mathfrak{M}, w \models i$ if $f(i) = w$ for each $i \in NOM$.

From Relational Granulation to Functional Granulation

In this section, we explain how to transform positional equivalences from relational granulation to functional granulation. Given a social network $\mathfrak{N} = (A, (P_i)_{i\in I}, (\alpha_j)_{j\in J})$, the steps are as follows:

1. For each kind of equivalence, find a corresponding modal language based on $\mathfrak{N}$. The three kinds of positional equivalence discussed earlier correspond to the following logical languages:
   (a) Structural equivalence: SHL
   (b) Regular equivalence: PMML
   (c) Exact equivalence: GML.
2. Consider $\mathfrak{N}$ as a model in the corresponding logic by a simple transformation. The set of actors is the set of possible worlds in the resultant model.
3. Find a sublanguage $\mathcal{L}$ of the corresponding logic. Prove that for any two actors x and y, $x \cong y$ iff x and y satisfy the same set of wffs in $\mathcal{L}$, where $\cong$ is $\cong_s$, $\cong_r$, or $\cong_e$ (see Fig. 5).

Structural Equivalences and Hybrid Logics

We use SHL to characterize structural equivalences. Given a social network $\mathfrak{N} = (A, (P_i)_{i\in I}, (\alpha_j)_{j\in J})$, we define an SHL language with the following alphabet:

1. $PV = \{p_i \mid i \in I\}$;
2. $REL = J$;
3. $NOM = A$.

The social network $\mathfrak{N}$ is transformed into a hybrid model $\mathfrak{M}_{\mathfrak{N}} = (A, (\alpha_j)_{j\in J}, V, id)$, where V is defined by $V(x, p_i) = 1$ iff $x \in P_i$ for all $x \in A$ and $i \in I$, and id is the identity function on A. Let us now define a sublanguage $\mathcal{H}_{\mathfrak{N}}$ of the SHL as the set of wffs $PV \cup \{\langle j\rangle a, \langle j^-\rangle a \mid j \in REL, a \in NOM\}$. We say that two actors, x and y, are $\mathcal{H}_{\mathfrak{N}}$-equivalent with respect to $\mathfrak{N}$ if for all $\varphi \in \mathcal{H}_{\mathfrak{N}}$, $(\mathfrak{M}_{\mathfrak{N}}, x \models \varphi$ iff $\mathfrak{M}_{\mathfrak{N}}, y \models \varphi)$. Then, we have the following characterization theorem.

Theorem 1 *Let $\mathfrak{N} = (A, (P_i)_{i\in I}, (\alpha_j)_{j\in J})$ be a social network; then, for all $x, y \in A$, $x \cong_s y$ in $\mathfrak{N}$ iff x and y are $\mathcal{H}_{\mathfrak{N}}$-equivalent with respect to $\mathfrak{N}$.*

Proof By the satisfaction condition, $(\mathfrak{M}_{\mathfrak{N}}, x \models \langle j\rangle a$ iff $(x, a) \in \alpha_j$. Thus, x and y are $\mathcal{H}_{\mathfrak{N}}$-equivalent iff the following two conditions are satisfied

1. $x \in P_i$ iff $y \in P_i$ for all $i \in I$ (as $PV \subseteq \mathcal{H}_{\mathfrak{N}}$),
2. for all $j \in J$ and $a \in A$, $(x, a) \in \alpha_j$ iff $(y, a) \in \alpha_j$ and $(a, x) \in \alpha_j$ iff $(a, y) \in \alpha_j$ (since both $\langle j\rangle a, \langle j^-\rangle a$ and $\langle j\rangle a, \langle j^-\rangle a$ are in $\mathcal{H}_{\mathfrak{N}}$).

These two conditions are exactly the same as those in Definition 1. Since $x \cong_s y$ iff (x, y) is in the coarsest structural equivalence, $x \cong_s y$ iff these two conditions are satisfied.□

Regular Equivalences and Multimodal Logics

The first logical characterization of positional equivalences is that of regular equivalences by dynamic logic – a particular type of PMML. This is shown in [22] by the connection between regular equivalence and the well-known notion of bisimulation in modal logics. We present the result here by following the general procedure of the transformation from relational granulation to functional granulation.

Given a social network $\mathfrak{N} = (A, (P_i)_{i\in I}, (\alpha_j)_{j\in J})$, we define a PMML language with the following alphabet:

1. $PV = \{p_i \mid i \in I\}$;
2. $REL = J$.

The social network $\mathfrak{N}$ is transformed into a Kripke model $\mathfrak{M}_{\mathfrak{N}} = (A, (\alpha_j)_{j\in J}, V)$, where V is defined by $V(x, p_i) = 1$ iff $x \in P_i$ for all $x \in A$ and $i \in I$. Let $\mathcal{L}_{\mathfrak{N}}$ be the set of wffs of the PMML and $\mathcal{L}_{\mathfrak{N}}$-equivalence be defined the same as $\mathcal{H}_{\mathfrak{N}}$-equivalence.

Theorem 2 ([22]) *Let $\mathfrak{N} = (A, (P_i)_{i\in I}, (\alpha_j)_{j\in J})$ be a social network; then, for all $x, y \in A$, $x \cong_r y$ in $\mathfrak{N}$ iff x and y are $\mathcal{L}_{\mathfrak{N}}$-equivalent with respect to $\mathfrak{N}$.*

Proof It is shown in [22] that regular equivalences are exactly bisimulations, i. e., $x \cong_r y$ iff x and y are bisimular. By Theorems 2.20 and 2.24 in [5], x and y are bisimular iff they are $\mathcal{L}_{\mathfrak{N}}$-equivalent with respect to $\mathfrak{N}$. □

Exact Equivalences and Graded Modal Logics

Following the procedure used in the previous subsection, we define $\mathcal{G}_{\mathfrak{N}}$ as the set of wffs of the GML based on the given network $\mathfrak{N}$. Then, we have the following theorem.

Theorem 3 *Let $\mathfrak{N} = (A, (P_i)_{i\in I}, (\alpha_j)_{j\in J})$ be a social network; then, for all $x, y \in A$, $x \cong_e y$ in $\mathfrak{N}$ iff x and y are $\mathcal{G}_{\mathfrak{N}}$-equivalent with respect to $\mathfrak{N}$.*

Proof

1. ($\Rightarrow$): Assume that $x \cong_e y$; then, there exists an exact equivalence ρ on $\mathfrak{N}$ such that $(x, y) \in \rho$. We show that $\mathfrak{M}_{\mathfrak{N}}, x \models \varphi$ implies $\mathfrak{M}_{\mathfrak{N}}, y \models \varphi$ by induction on the complexity of φ (the proof of the converse direction is exactly the same). The case where $\varphi \in PV$ follows from condition 1 of Definition 3. The cases for Boolean connectives are straightforward by the induction hypothesis. For formulas of the form $\langle j\rangle_n\psi$ or $\langle j^-\rangle_n\psi$, we prove the case for $\langle j\rangle_n\psi$ by using the condition $[\![N^+_{\alpha_j}(x)]\!]_\rho = [\![N^+_{\alpha_j}(y)]\!]_\rho$. The case for $\langle j^-\rangle_n\psi$ is proved similarly by applying $[\![N^-_{\alpha_j}(x)]\!]_\rho = [\![N^-_{\alpha_j}(y)]\!]_\rho$. Now, if $\mathfrak{M}_{\mathfrak{N}}, x \models \langle j\rangle_n\psi$, then there exist more than n worlds in $N^+_{\alpha_j}(x)$ that satisfy ψ. Assume $u_0, u_1, \ldots, u_n$ are among these worlds. Since $[\![N^+_{\alpha_j}(x)]\!]_\rho = [\![N^+_{\alpha_j}(y)]\!]_\rho$, there exist $u'_0, u'_1, \ldots, u'_n \in N^+_{\alpha_j}(y)$ such that $(u_i, u'_i) \in \rho$ for $0 \le i \le n$. By the induction hypothesis, $\mathfrak{M}_{\mathfrak{N}}, u'_i \models \psi$ for $0 \le i \le n$; thus, $\mathfrak{M}_{\mathfrak{N}}, y \models \langle j\rangle_n\psi$.
2. ($\Leftarrow$): Let us define an equivalence relation ρ on A by

   $(x, y) \in \rho$
   iff x and y are $\mathcal{G}_{\mathfrak{N}}$-equivalent with respect to $\mathfrak{N}$.

   We prove that ρ is an exact equivalence. The first condition of Definition 3 is straightforward. For the second condition, let us assume that $(x, y) \in \rho$, $[\![N^+_{\alpha_j}(x)]\!]_\rho = ([A]_\rho, m_1)$, and $[\![N^+_{\alpha_j}(y)]\!]_\rho = ([A]_\rho, m_2)$ for some $j \in J$. Assume also that $[A]_\rho = \{[z_i]_\rho \mid 1 \le i \le k\}$. First, we consider z_1 as an example. For each $2 \le i \le k$, since $[z_1]_\rho$ is an equivalence class distinct from $[z_i]_\rho$, there exists a $\mathcal{G}_{\mathfrak{N}}$-wff ψ_i such that $\mathfrak{M}_{\mathfrak{N}}, z_i \models \psi_i$, but $\mathfrak{M}_{\mathfrak{N}}, z_1 \models \neg\psi_i$. If $m_1([z_1]_\rho) = n_1 > 0$, then there exist exactly n_1 α_j-successors of x that are $\mathcal{G}_{\mathfrak{N}}$-equivalent to z_1. Thus, $\mathfrak{M}_{\mathfrak{N}}, x \models (j)_{n_1}(\psi \wedge \bigwedge_{i=2}^k \neg\psi_i)$ for all ψ satisfied in z_1. Since $(x, y) \in \rho$, it is that $\mathfrak{M}_{\mathfrak{N}}, y \models (j)_{n_1}(\psi \wedge \bigwedge_{i=2}^k \neg\psi_i)$ for all ψ satisfied in z_1. This means there are exactly $n_1\alpha_j$-successors of y that satisfy $\bigwedge_{i=2}^k \neg\psi_i$ (by letting $\psi = \top$). Thus, these n_1 α_j-successors of y must satisfy all ψ satisfied in z_1 and be $\mathcal{G}_{\mathfrak{N}}$-equivalent to z_1. Since no other α_j-successors of y can be $\mathcal{G}_{\mathfrak{N}}$-equivalent to z_1 (otherwise there would be more than $n_1\alpha_j$-successors of y satisfying $\bigwedge_{i=2}^k \neg\psi_i$), $m_2([z_1]_\rho) = n_1$. In the same way, we can prove that if $m_2([z_1]_\rho) > 0$, then $m_1([z_1]_\rho) > 0$. Thus, $m_1([z_1]_\rho) = m_2([z_1]_\rho)$. Similarly, we can prove $m_1([z_i]_\rho) = m_2([z_i]_\rho)$ for $2 \le i \le k$. Consequently, we have $m_1 = m_2$ and $[\![N^+_{\alpha_j}(x)]\!]_\rho = [\![N^+_{\alpha_j}(y)]\!]_\rho$. The proof of $[\![N^-_{\alpha_j}(x)]\!]_\rho = [|N^-_{\alpha_j}(y)|]_\rho$ is analogous, except that the modality $(j)_{n_1}$ is replaced by $(j^-)_{n_1}$. □

Observational Equivalences

We have presented the logical characterizations of three kinds of positional equivalence. The theoretical implication of these characterizations is that positional equivalences based on relational granulation can be transformed into indiscernibility relations based on a set of binary attributes. Let $\mathfrak{N} = (A, (P_i)_{i\in I}, (\alpha_j)_{j\in J})$ be a social network and let F denote $\mathcal{H}_{\mathfrak{N}}$, $\mathcal{L}_{\mathfrak{N}}$, or $\mathcal{G}_{\mathfrak{N}}$. Then, (A, F) is a (generalized) data table such that IND(F) is the corresponding positional equivalence. However, from a practical viewpoint, since both $\mathcal{L}_{\mathfrak{N}}$ and $\mathcal{G}_{\mathfrak{N}}$ are infinite sets, it is generally difficult (if not impossible) to decide IND(F) based on such a data table representation. This motivates us to define a new kind of positional equivalence based on observations.

Let $\mathfrak{N}$ be a social network and F be a *finite* set of wffs in an appropriate logic language. Then, two actors, x and y, are *F-observationally equivalent*, denoted by $x \cong_F y$, if for any wff $\varphi \in F$, x satisfies φ iff y satisfies φ. The wffs in F denote the observable properties of the actors in the network. Due to the limited capability of the observer, the set of observable properties is usually finite. Depending on the observable properties chosen, we can focus on different parts of the social network. For example, positional analysis for a genealogical study and a political investigation

may have to consider different relationships between actors. As shown by the above results, the choice of an appropriate logic language is also crucial, since it may result in different notions of positional equivalence.

A recent approach to structural equivalence that conforms with our definition of observational equivalence is reported in [28]. In that approach, a set of institutions E is given and each actor may belong to several institutions. Two actors are structurally equivalent if they belong to the same set of institutions. To avoid confusion with the structural equivalence defined in Sect. "Positional Analysis – Relational Granulation", we call the equivalence relation defined in [28] *institutional equivalence*. Note that institutional equivalence is in fact the same as structural equivalence defined in Sect. "Positional Analysis – Relational Granulation", if we consider the universe as the union of A and E and represent the network as a bipartite graph. A place (social position) is then defined as an equivalence class of actors (or equivalently, the set of institutions that those actors belong to). Here, our set of observation sentences corresponds to the set of institutions in [28]. While institutions are primitive entities, our observation sentences are formed from the social relations and attributes of actors. In fact, a wff in F can also be regarded as an intensional definition of an institution. In this sense, observational equivalence is a kind of intensionally institutional equivalence.

Conclusion and Future Directions

This article starts with an introduction to rough set theory – the basic theory of functional granulation. We then review several notions of positional equivalence in social network analysis and consider them as relational granulation from the GrC perspective. By using modal logics as a bridge, we transform the relational granulation-based positional equivalences into functional granulation. However, the transformation is mainly theoretically interesting, since the set of functional attributes used to granulate a social network may be infinite. This practical consideration motivates us to propose the notion of observational equivalence. It conforms with an alternative definition of structural equivalence, which we call institutional equivalence to avoid confusion.

More generally, both observational equivalence and institutional equivalence can be analyzed in the framework of formal concept analysis (FCA) [13]. FCA is a conceptual knowledge representation and data analysis method in which a *formal context* is defined as a triple (A, F, R), where A and F are sets of objects and attributes respectively; and $R \subseteq A \times F$ is a binary relation between A and F, called the incidence of the formal context. For $A' \subseteq A$ and $F' \subseteq F$, let us denote $R(A') = \{f \in F \mid \forall a \in A', (a, f) \in R\}$ and $R^-(F') = \{a \in A \mid \forall f \in F', (a, f) \in R\}$. Then, a *formal concept* is defined as a pair (A', F') such that $A' \subseteq A$, $F' \subseteq F$, $R(A') = F'$ and $R^-(F') = A'$. Here, we call A' the extent and F' the intent of the formal concept. The application of FCA to sociology has been proposed in [12]. In that application, the formal context (also called the Galois lattice) is used to model two-mode network data, where the set of actors is a mode of the data, the set of events is the other mode of data, and the relation between the two modes is the participation relationship between the actors and events. When we regard A as the set of actors, F as the set of observation sentences (resp. institution), and R as the satisfaction (resp. membership) relation, a formal context can be constructed out of a social network. However, a social position (or place) defined by observational or institutional equivalences may not be the extent of a formal concept. Thus, it is tempting to define a social position alternatively as a formal concept of such formal contexts. Unlike other positional analysis techniques that define a social position as an equivalence class of some positional equivalence, the extents of different formal concepts may overlap. This means that social positions may not be disjoint in such a definition. The theoretical and practical implications of such positional analysis requires further research.

Another future research direction concerns the practical computation of positional equivalences. As mentioned earlier, when a social network is transformed into a data table (A, F) for positional analysis and F is infinite, it is difficult to decide if two actors are IND(F)-equivalent by exhaustively comparing the satisfaction of all wffs in x and y. While observational equivalence provides a practical approximation of the computation, we would like to know if precise computation is possible. To answer this question, the inference of the *most specific concept* (msc) in description logics [2] may be helpful. Description logics are a family of knowledge representation formalisms that have strong connections with modal logics [3,23]. Indeed, the description languages $\mathcal{ALC}$ and $\mathcal{ALCN}$ correspond to the sublanguages of PMML and GML respectively [3,23]. Given a network represented in a description language, the msc of an actor x, denoted by $msc(x)$, corresponds to a wff that satisfies x and implies any wffs that satisfy x. Thus, if $msc(x)$ and $msc(y)$ are logically equivalent, then x and y will satisfy the same set of wffs, i. e., $(x, y) \in \text{IND}(F)$. The problem for researchers is to find logical languages that can characterize the aforementioned positional equivalences adequately and whose msc problems can be solved effectively.

In summary, we have proposed a logical approach to positional analysis in social networks from a GrC perspective. The approach facilitates a connection between more complicated relational granulation and simpler functional granulation. While there remain problems to be solved in the research programme, the approach seems promising for cross-fertilization between the fields of GrC and SNA.

Bibliography

1. Areces C, ten Cate B (2007) Hybrid logics. In: Blackburn P, Van Benthem J, Wolter F (eds) Handbook of Modal Logic. Elsevier, Amsterdam, pp 821–868
2. Baader F, Küsters R (2007) Nonstandard inferences in description logics: the story so far. In: Gabbay DM, Goncharov SS, Zakharyaschev M (eds) Mathematical Problems from Applied Logic I: Logics for XXIst Century. Springer, Berlin
3. Baader F, Nutt W (2002) Basic description logics. In: Baader F, Calvanese D, McGuinness DL, Nardi D, Patel-Schneider PF (eds) Description Logic Handbook. Cambridge University Press, Cambridge, pp 47–100
4. Blackburn P (2000) Representation, reasoning, and relational structures: a hybrid logic manifesto. Logic J IGPL 8(3):339–625
5. Blackburn P, de Rijke M, Venema Y (2001) Modal Logic. Cambridge University Press, Cambridge
6. Blackburn P, Seligman J (1995) Hybrid languages. J Log Lang Inf 4:251–272
7. Borgatti SP, Everett MG (1989) The class of all regular equivalences: algebraic structure and computation. Soc Netw 11(1):65–88
8. Boyd JP, Everett MG (1999) Relations, residuals, regular interiors, and relative regular equivalence. Soc Netw 21(2):147–165
9. Chellas BF (1980) Modal Logic: An Introduction. Cambridge University Press, Cambridge
10. Everett MG, Borgatti SP (1994) Regular equivalences: general theory. J Math Sociol 18(1):29–52
11. Fan TF, Liau CJ, Liu DR, Tzeng GH (2006) Granulation based on hybrid information systems. In: Proc of the 2006 IEEE International Conference on Systems Man and Cybernetics, Taipei, pp 4768–4772
12. Freeman LC, White DR (1993) Using Galois lattices to represent network data. Sociol Methodol 23:127–146
13. Ganter B, Wille R (1999) Formal Concept Analysis: Mathematical Foundations. Springer, Berlin
14. Gargov G, Goranko V (1993) Modal logic with names. J Philos Log 22(6):607–636
15. Hanneman RA, Riddle M (2005) Introduction to Social Network Methods. University of California, Riverside
16. Horrocks I, Hustadt U, Sattler U, Schmidt R (2007) Computational modal logic. In: Blackburn P, Van Benthem J, Wolter F (eds) Handbook of Modal Logic. Elsevier, Amsterdam, pp 181–245
17. Kripke S (1959) A completeness theorem in modal logic. J Symb Log 24(1):1–14
18. Kripke S (1963) Semantic analysis of modal logic I: normal propositional calculi. Z Math Log Grundl Math 9:67–96
19. Lerner J (2005) Role assignments. In: Brandes U, Erlebach T (eds) Network Analysis. LNCS, vol 3418. Springer, Berlin, pp 216–252
20. Liau CJ, Lin TY (2005) Reasoning about relational granulation in modal logics. In: Proc of the First IEEE International Conference on Granular Computing, Beijing, pp 534–558
21. Lorrain F, White HC (1971) Structural equivalence of individuals in social networks. J Math Sociol 1:49–80
22. Marx M, Masuch M (2003) Regular equivalence and dynamic logic. Soc Netw 25(1):51–65
23. Nardi D, Brachman RJ (2002) An introduction to description logics. In: Baader F, Calvanese D, McGuinness DL, Nardi D, Patel-Schneider PF (eds) Description Logic Handbook. Cambridge University Press, Cambridge, pp 5–44
24. Passy S, Tinchev T (1985) Quantifiers in combinatory PDL: completeness, definability, incompleteness. In: Fundamentals of Computation Theory FCT 85. LNCS, vol 199. Springer, Berlin, pp 512–519
25. Passy S, Tinchev T (1991) An essay in combinatory dynamic logic. Inf Comput 93:263–332
26. Pawlak Z (1982) Rough sets. Int J Inf Comput Sci 11(15):341–356
27. Pawlak Z (1991) Rough Sets – Theoretical Aspects of Reasoning about Data. Kluwer, Dordrecht
28. Pizarro N (2007) Structural identity and equivalence of individuals in social networks: Beyond duality. Int Sociol 22(6):767–792
29. Prior A (1967) Past, Present and Future. Oxford University Press, London
30. Prior A (1968) Now. Nous 2:101–119
31. Scott J (2000) Social Network Analysis: A Handbook, 2nd edn. SAGE Publications, London
32. Sierpinski W, Krieger C (1956) General Topology. University of Toronto Press, Toronto
33. van der Hoek W (1992) On the semantics of graded modalities. J Appl Non-Class Log 2(1):81–123
34. Wasserman S, Faust K (1994) Social Network Analysis: Methods and Applications. Cambridge University Press, Cambridge
35. White DR, Reitz KP (1983) Graph and semigroup homomorphisms on netwoks and relations. Soc Netw 5(1):143–234
36. Zadeh LA (1979) Fuzzy sets and information granularity. In: Gupta N, Ragade R, Yager R (eds) Advances in Fuzzy Set Theory and Applications. North-Holland, Amsterdam, pp 3–18
37. Zadeh LA (1997) Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets Syst 19:111–127

Social Network Visualization, Methods of

Linton C. Freeman
Department of Sociology and Institute for Mathematical Behavioral Science, School of Social Sciences, University of California, Irvine, USA

Article Outline

Glossary
Definition of the Subject
Introduction

Visualization in Social Network Analysis
Images Based on One Mode Undirected Relations
Images Based on One Mode Directed Relations
Images Based on Two Mode Relations
Images Based on One or Two Mode Data Matrices
Future Directions
Bibliography

Glossary

Adjacent A node is adjacent to another if there is an edge connecting them.

Arrow A line directed from one node to another.

Binary relation A two valued yes/no or on/off relation.

Bipartite graph A graph, $B = \langle N, E \rangle$ where N is a finite set of nodes and E is a collection of pairs of nodes in which N is partitioned into two disjoint subsets, N_1 and N_2, and no edge in E has both end points in the same subset.

Blockmodeling A procedure for clustering actors such that the actors in each cluster share similar patterns of ties both within and between clusters.

Connected Any two nodes in a graph are said to be connected if there is a path from one to the other; a graph is connected if there is a path connecting every pair of nodes.

Cycle Any path that begins and ends at the same node.

Digraph A directed graph.

Directed graph A graph $D = \langle N, A \rangle$ where N is a finite collection of nodes and A is a set of pairs linked by directed lines or arrows.

Directed line A line going from a node to another representing a non-reciprocated link.

Edge A line connecting two nodes representing a reciprocated link.

Edge labeled graph A graph in which at least two kinds of connections between nodes are identified.

Formal concept analysis A method of data analysis based on Galois lattice structure.

Galois lattice A dual structure that displays the dependencies of both objects and their properties.

Geodesic The shortest path between two nodes.

Graph A graph $G = \langle N, E \rangle$ where N is a finite set of nodes and E is a collection of pairs of nodes represented as edges.

Hyperedge An edge in a hypergraph that can enclose more than two nodes.

Hypergraph A hypergraph, $F = \langle N, H \rangle$, consists of a set of nodes N and a collection of hyperedges, H.

Indegree The indegree of a node is the number of directed lines it receives.

Irreflexive A relation in which no edge connects any node with itself.

Multidimensional scaling A search procedure designed to represent an observed set of proximities or distances in a small number of dimensions.

Node A point in a graph.

One mode matrix A data matrix in which the rows and columns both represent the same objects.

Outdegree The outdegree of a node is the number of directed lines it sends out.

Path A path is a sequence of nodes and edges beginning with a node that has an edge connecting it to the next node in the sequence and so on.

Path length The length of a path connecting two nodes is the number of edges it contains.

Permutation A reordering of the rows, columns, or rows and columns of a matrix.

Principle diagonal The set of cells in a square matrix that runs from the upper left to the lower right.

Relation A collection of ordered or unordered pairs of nodes.

Singular value decomposition an algebraic procedure that decomposes a data matrix into its "basic structure".

Sociometry An early version of social network analysis introduced by Jacob Moreno and Helen Jennings.

Spring embedder A kind of multidimensional scaling based on a model in which it is assumed that nodes are connected by springs that pull and push on them.

Symmetric A relation in which if a node a is adjacent to another, b, then b is adjacent to a.

Tree A graph is a tree if it is connected and contains no cycles.

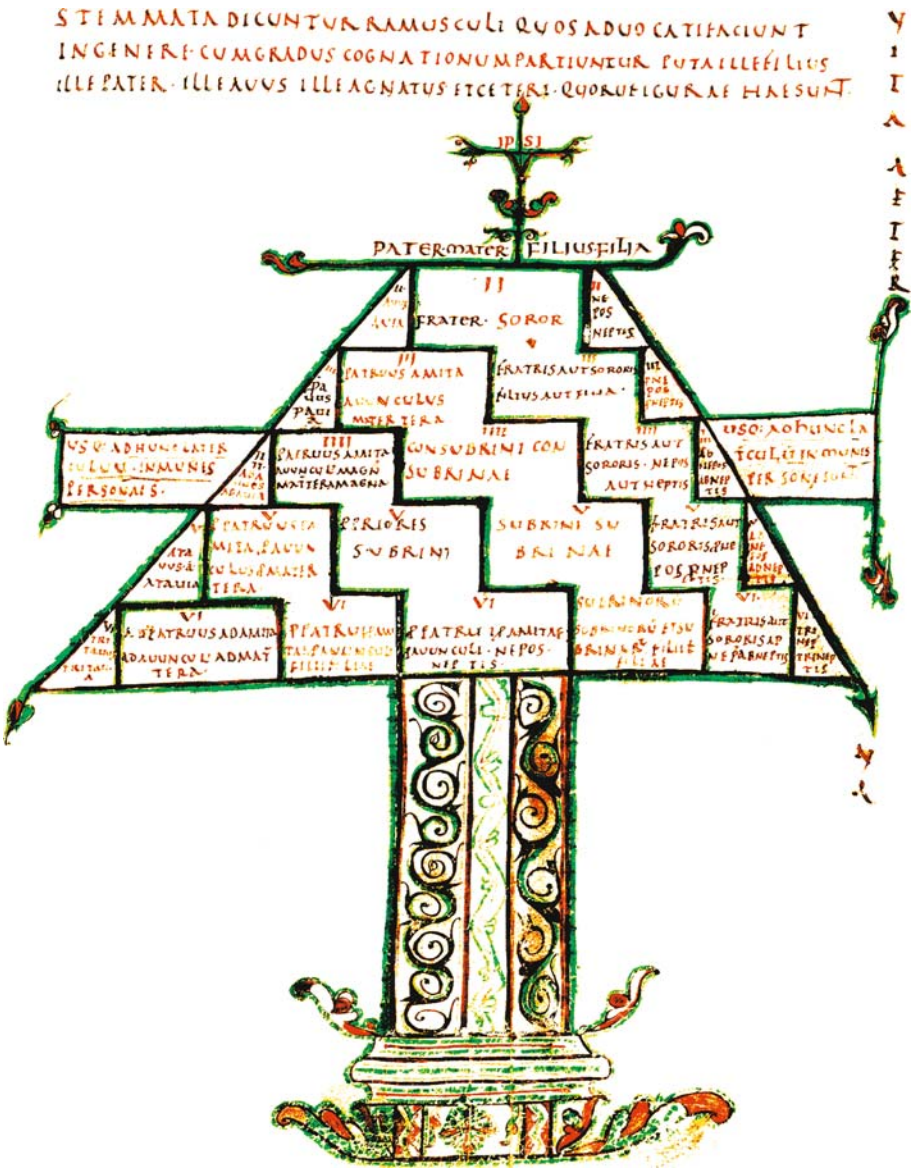
Two mode matrix A data matrix in which the rows and columns represent different objects.

Definition of the Subject

Social network visualization refers to the practice of constructing pictorial images of the connections linking social actors. The use of such images provides two benefits. It allows investigators to gain new insights into the patterning of social connections, and it helps investigators to communicate their results to others.

Introduction

Social network analysis did not emerge as a systematic field of research until early in the twentieth century [1]. But visual images of social networks were produced more than a millennium earlier. The earliest of these images that I have uncovered was produced in Spain in the middle of



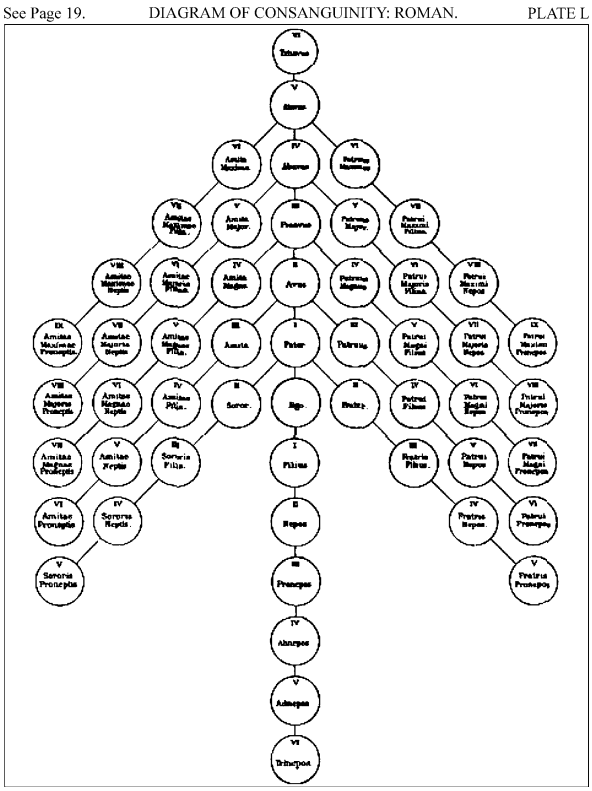
Social Network Visualization, Methods of, Figure 1
Tree of consanguinity with six degrees of relationship

the ninth century. That image is attributed to the prolific writer and Roman Catholic Saint, Isidore de Séville. It is reproduced here as Fig. 1.

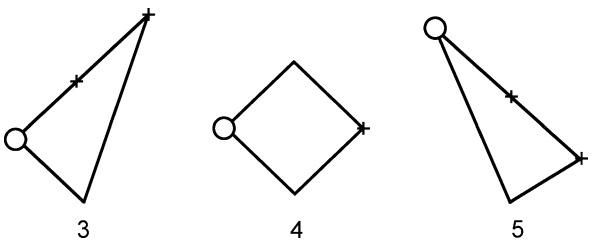
The image shown in Fig. 1 displays relationships based on genealogical descent. From the earliest times, people have been interested in kinship ties – in who is related to whom. This interest is evident in the descent lists found in the Christian bible and in the oral genealogies that were required to be memorized by Hawaiian nobles [2].

The fact that Fig. 1 takes the form of a tree shows that as early as the ninth century people saw the analogy between the branching structure of descent and that of trees. This notion was captured in a mathematical formalization in 1857 by Arthur Cayley [3]. Cayley defined a tree in mathematical graph theoretic terms. Biggs, Lloyd and Wilson (p. 38 in [4]) characterized Cayley's definition by saying that his "… use of the word 'tree' in this context is presumably derived from the diagrammatic form of these graphs, and is akin to the traditional use of the word in describing genealogical or 'family' trees."

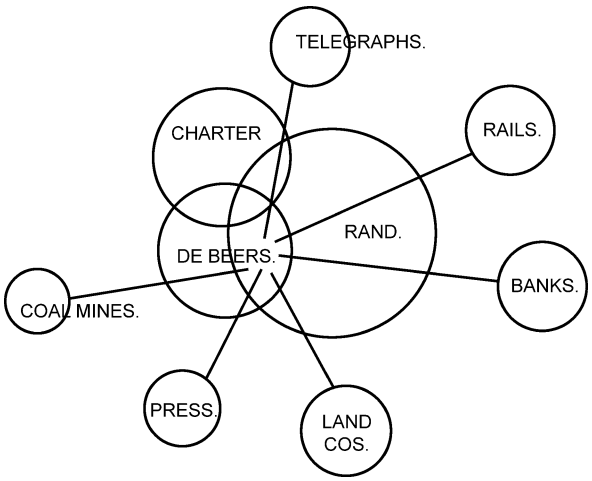
The use of trees to depict descent was, of course, continued. As time passed, however, their form became sim-



Social Network Visualization, Methods of, Figure 2
Descent in ancient Rome



Social Network Visualization, Methods of, Figure 3
Macfarlane's images of two-step marriage prohibitions



Social Network Visualization, Methods of, Figure 4
Hobson's image of corporate interlocks

plified. Lewis Henry Morgan [5] was an attorney and an anthropologist. He was interested in comparing how different peoples reckoned kinship and in 1871 he published a mammoth work containing a collection of kinship trees. Each tree depicted descent as conceived by a society somewhere in the world. Morgan's trees are quite simple. Figure 2 shows descent as it was reckoned in ancient Rome.

Twelve years later a mathematician-physicist, Alexander Macfarlane [6], produced a different kind of graphic image based still on kinship. Macfarlane set out to examine British marriage prohibitions and he represented them both algebraically and visually. His visual images depict males using plus signs (+) and females with circles (o). Earlier generations he placed higher on the page. Descent is shown by lines connecting points. A short line crossing a descent line indicates another person, of either sex, in an intermediate generation. And the lowest point is always the prohibited offspring.

The illustration shown in Fig. 3 displays all the two-step marriage relations that are prohibited by British law. The left image shows that a male may not marry his granddaughter. The middle image shows that he may not marry his sister. And the right image shows that he may not marry his grandmother. Or, put the other way, a woman may not marry her grandfather, her brother or her grandson.

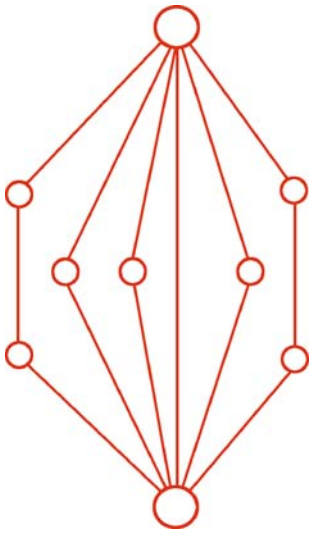
Macfarlane's paper also included algebraic expressions that captured all of the same marriage prohibitions. But Sir Francis Galton [7], who attended Macfarlane's presentation, declared that his "diagrammatic form" seemed "the most distinctive and self-explanatory" of the two treatments.

Finally, in 1894, John Hobson [8] produced a visual image of a social network that was not based on kinship. He had collected two mode (corporation by director) data on interlocking corporate directorates. He reasoned that, to the degree that corporations shared directors, they could be expected to cooperate and work together.

Hobson's illustration was designed to show the interlock among, as he put it, "the small inner ring of South African finance." Corporations are depicted as circles, and interlock is shown by overlapping or by a line connecting two circles. Hobson's image is reproduced here as Fig. 4.

The important feature of this image is that it displays a connection linking more than two corporations. Hobson's data showed that three corporations, Charter, Rand and De Beers, all shared directors in common. And, at the same time, Rand and De Beers also both shared directors with coalmines, telegraphs, rails, and others. The overlaps in his image allowed him to display which companies shared with which others.

It is clear, then, that a concern with connections among social actors and the use of visual images have a long history of intimate association. It should come as no surprise therefore that images played an important part in the development of social network analysis when it did emerge as an organized field of research.



Social Network Visualization, Methods of, Figure 5
Image of a Pattern of Linkages

Visualization in Social Network Analysis

In the book cited above (p. 3 in [1]), I described the modern science of social network analysis as possessing four defining properties. They were:

1. It embodies ideas about the importance of social ties linking social actors.
2. It collects data reflecting those ties.
3. It involves the use of graphic imagery.
4. It employs mathematical and/or computational models.

Pre-network research often included one or two of those properties, but in the late 1920s each of two independent research teams came up with efforts that included all four.

One took place in the early 1930s. It involved a psychiatrist, Jacob L. Moreno, and a psychologist, Helen H. Jennings. Together, they developed an approach they called "sociometry." They reported two huge studies, both focused on examining the structure of social ties. One was conducted among prisoners at Sing Sing Prison in Ossining, New York [9] and the other among young delinquents at the New York State Training School for Girls in Hudson, New York [10].

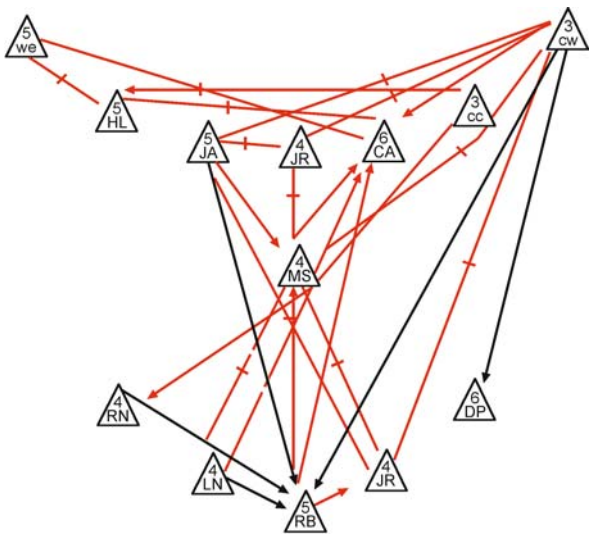
Both Moreno–Jennings studies involved the extensive use of graphic images. The image shown in Fig. 5 was included in their report on the research at Sing Sing prison (Moreno, 1932). In that figure, individuals or other kinds of social actors are represented as points or *nodes* and links between pairs of social actors are lines or *edges* connecting pairs of nodes. In Fig. 5 Moreno was concerned with the positions of individuals and the patterning of their ties. As he put it, the individuals at the top and the bottom were "dominant" and the image showed that those dominant individuals were linked both "directly" and "indirectly".

Most of the data collected by Moreno and Jennings involved asking individuals whom they liked or disliked. In data of that sort, choices are seldom reciprocated. So Moreno and Jennings drew lines with arrowheads to reveal who chose whom. Mutual choices were drawn without arrowheads and they also included a small line bisecting the main line connecting the two nodes.

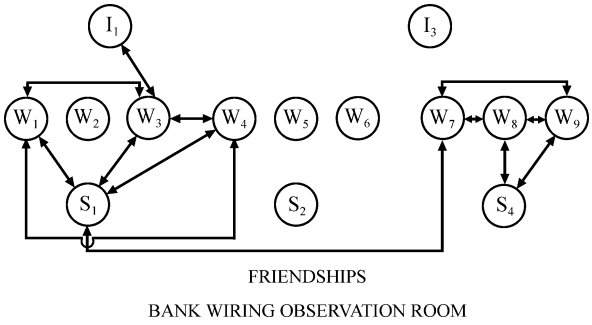
Moreno and Jennings often required subjects to report both their likes and their dislikes. By using different colors, red for likes and black for dislikes, a single image could display both. The image shown in Fig. 6 was published in the Moreno–Jennings report on the Hudson School (Moreno, 1934). It depicts positive and negative choices among 13 members of an American football team. Moreover, it contains another innovation. The various team members are placed in the drawing in approximately the same relative locations that they occupied on the football field. That arrangement shows the players' positions and permits the viewer to evaluate the impact of physical proximity on the patterning of social linkages.

Figures of the sort used by Moreno and Jennings had a major impact on the style of graphic imagery used subsequently in social network analysis. For the most part, social network analysts have represented social actors as nodes and links between actors as edges or as directed lines with arrowheads.

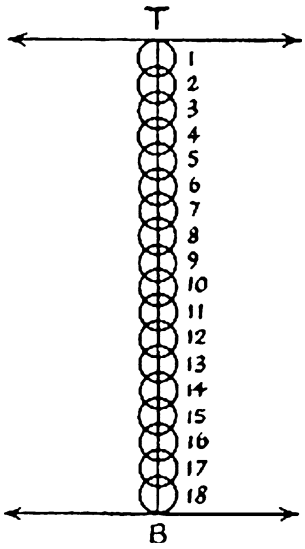
The second introduction of the social network approach also occurred in the early 1930s. An anthropologist, W. Lloyd Warner, and a collection of his colleagues and students at Harvard, conducted three elab-



Social Network Visualization, Methods of, Figure 6
Positive and Negative Choices in a Football Team



Social Network Visualization, Methods of, Figure 7
Friendships linking factory workers



VII. *The Clique*
Vertical Axis

Social Network Visualization, Methods of, Figure 8
An idealized pattern of overlapping 'cliques'

orate network analytic projects. One was a study of an industrial factory, the Western Electric plant in Cicero, Illinois [11]. The other two were studies of communities: one focused on a New England town, Newburyport, Massachusetts [12], and the other on a southern town, Natchez, Mississippi [13].

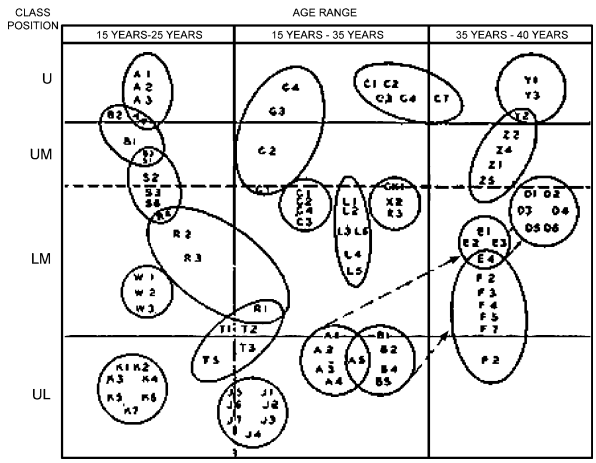
The image shown in Fig. 7 was produced as part of the factory study. It displays observed friendship ties among pairs of individuals who worked together in the same workroom. It was drawn using nodes and two-headed lines instead of edges, but it is very similar to the images produced by Moreno and Jennings. In addition, as in Fig. 6 above, the impact of physical space was displayed; workers were placed in the drawing in positions that reflected the locations of their workstations.

In reporting their study of Newburyport, Warner and Lunt used the kind of drawing of overlapping circles that Hobson had used to construct Fig. 4. But here that image was introduced, not to describe data, but to propose an idea they had about social structure. The diagram in Fig. 8 represents the investigators' idealized version of the expected structure of overlaps among subgroups in the presence of social class. The idea is that only subgroups that are close to one another in class ranking are likely to have overlapping memberships.

In their study of Natchez Davis, Gardner and Gardner [13] employed the same diagrammatic form to display two mode data reflecting on the earlier Newburyport hypothesis. Figure 9 shows subgroups of black males and their overlaps. In that image the men are arranged in terms of both social class and age. Both, it turned out, provided important bases for grouping.

Finally, in that same report, Davis, Gardner and Gardner also introduced an entirely different kind of social network image. Like Hobson, they had collected two mode data. Eighteen women were designated in the rows of their data matrix and fourteen social events were depicted in the columns. That matrix is reproduced here as Fig. 10.

The data shown in Fig. 10 were all collected during a single year. But, by examining the column headings, it is clear that Davis and his colleagues did not arrange the social events according to the dates upon which they took place. Instead, they listed both the events and the women who attended them in such a way that the arrangement itself suggests that these women were organized into two



Social stratification of a group of colored cliques: ages 15–60

Social Network Visualization, Methods of, Figure 9
Stratification, age and overlapping groups

groups. The two groups overlap, but for the most part they are distinct. Most of the women in the top half of the matrix attended the leftmost five events. And most of the women in the bottom half attended the rightmost five events. The middle four events apparently brought both groups of women together.

This arrangement of women and events was self-consciously produced by the authors. Davis, Gardner and Gardner were convinced that these women were organized into two groups and they presented their data matrix in a way that would illustrate that conclusion. The interesting thing is that these authors never commented explicitly about how they had rearranged the columns and rows in their matrix. They simply organized their display in a way that would make the point.

From the outset, then, four kinds of images have played important parts in the development of social network analysis. These first network graphics included drawings displaying (1) one mode undirected relations, (2) one mode directed relations, (3) two mode relations and (4) one or two mode data matrices. A few other kinds of network images have been used since then, but the four originals – particularly those based on one mode undirected and one mode directed relations – still dominate the field. In the next four sections we will examine the four original kinds of images and how their use has evolved in the social network context.

Images Based on One Mode Undirected Relations

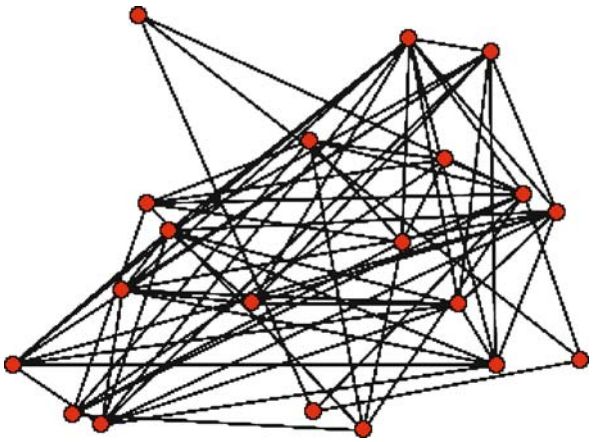
Mathematically speaking, the node and edge images introduced by Moreno and Jennings in Fig. 5 are *graphs*. A graph $G = \langle N, E \rangle$ where N is a finite set of nodes and E is a collection of pairs of nodes. In graph visualizations, a pair of nodes in E is presented as a line connecting the two nodes in question. Two nodes are called *adjacent* if there is an edge directly connecting them to each other. A graph embodies a *binary* (yes/no or on/off) relation that is *irreflexive* (no node is adjacent to itself) and *symmetric* (if a node a is adjacent to another, b, then b is adjacent to a).

A *path* is a sequence of nodes and edges, beginning with a node that has an edge connecting it to the next node in the sequence. The *length* of a path between two nodes

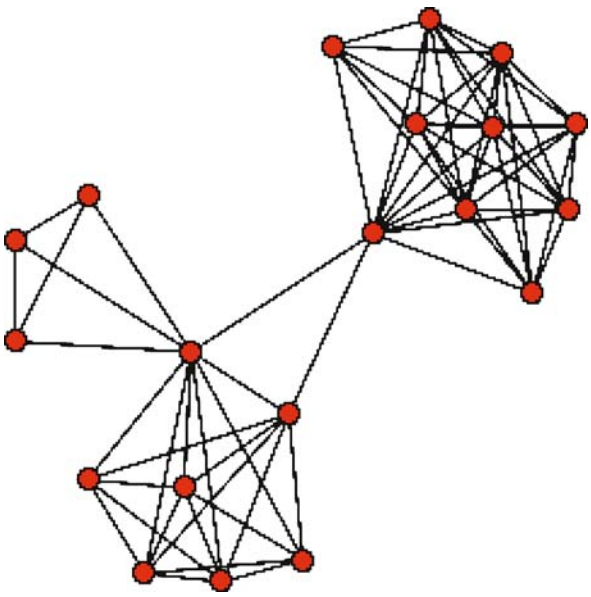
Names of Participants of Group I	Code Numbers and Dates of Social Events Reported in *Old City Herald*													
	(1) 6/27	(2) 3/2	(3) 4/12	(4) 9/26	(5) 2/25	(6) 5/19	(7) 3/15	(8) 9/16	(9) 4/8	(10) 6/10	(11) 2/23	(12) 4/7	(13) 11/21	(14) 8/3
1. Mrs. Evelyn Jefferson	X	X	X	X	X	X	....	X	X	....	....	....	....	....
2. Miss Laura Mandeville	X	X	X	....	X	X	X	X	....	....	....	....	....	....
3. Miss Theresa Anderson	....	X	X	X	X	X	X	X	X	....	....	....	....	....
4. Miss Brenda Rogers	X	....	X	X	X	X	X	X	....	....	....	....	....	....
5. Miss Charlotte McDowd	....	....	X	X	X	....	X	....	....	....	....	....	....	....
6. Miss Frances Anderson	....	....	X	....	X	X	....	X	....	....	....	....	....	....
7. Miss Eleanor Nye	....	....	....	....	X	X	X	X	....	....	....	....	....	....
8. Miss Pearl Oglethorpe	....	....	....	....	....	X	....	X	X	....	....	....	....	....
9. Miss Ruth DeSand	....	....	....	....	X	....	X	X	X	....	....	....	....	....
10. Miss Verne Sanderson	....	....	....	....	....	....	X	X	X	....	....	X	....	....
11. Miss Myra Liddell	....	....	....	....	....	....	....	X	X	X	....	X	....	....
12. Miss Katherine Rogers	....	....	....	....	....	....	....	X	X	X	....	X	X	X
13. Mrs. Sylvia Avondale	....	....	....	....	....	....	X	X	X	X	....	X	X	X
14. Mrs. Nora Fayette	....	....	....	....	....	X	X	....	X	X	X	X	X	X
15. Mrs. Helen Lloyd	....	....	....	....	....	....	X	X	....	X	X	X	....	....
16. Mrs. Dorothy Murchison	....	....	....	....	....	....	....	X	X	....	....	....	....	....
17. Mrs. Olivia Carleton	....	....	....	....	....	....	....	....	X	....	X	....	....	....
18. Mrs. Flora Price	....	....	....	....	....	....	....	....	X	....	X	....	....	....

Fig. 3.—Frequency of interparticipation of a group of women in Old City, 1936—Group I.

Social Network Visualization, Methods of, Figure 10
The Davis, Gardner and Gardner data on women's attendance at social events



Social Network Visualization, Methods of, Figure 11
Links in the network of a homeless woman I



Social Network Visualization, Methods of, Figure 12
Links in the network of a homeless woman II

is the number of edges it contains. And the shortest path connecting two nodes is called the *geodesic*.

Any two nodes in N are *connected* if there is a path from one to the other. And a whole graph G is connected if every pair of nodes in N is connected. If a path begins and ends at the same node, that path is a *cycle*. Finally, a graph is a *tree* if it is both connected and it contains no cycles.

The image in Fig. 11 is a graph. It is based on data recorded by J. Clyde Mitchell [14] on the social ties among the 19 individuals involved in the personal network of a homeless woman in Britain. I used a program called NetDraw to place the nodes representing individuals in Fig. 11 in random positions. That calls attention to the importance of the locations of points in graphic displays. Given the locations of the points in Fig. 11 it is very difficult for the viewer to see anything interesting in the patterning of this woman's network.

Compare the image in Fig. 11 with that in Fig. 12. Figure 12 was also produced using NetDraw, but this time the points were placed using a *spring embedder* [16]. A spring embedder is a computer algorithm that, in effect, places a spring of unit length between every pair of adjacent nodes and a much longer spring between nodes that are not adjacent. It starts with a random placement of nodes, then the whole apparatus is set in motion and the various springs push and pull until they reach an equilibrium.

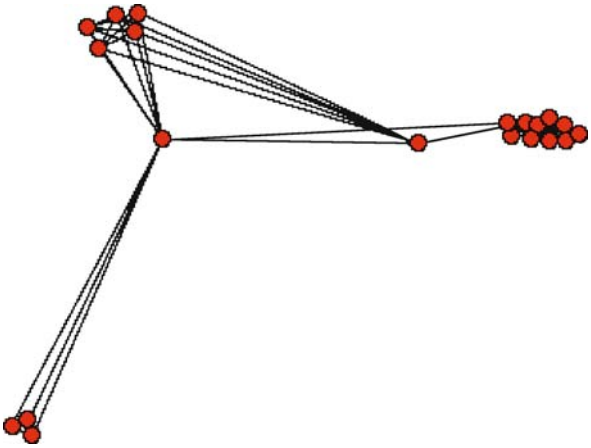
The advantage of using a spring embedder is that it does not require the investigator to make ad hoc judgments in locating nodes in a graph. It uses a standard computer algorithm to place the nodes automatically. Thus, every user will get the same result. There are several different spring embedding algorithms. And they are all examples of a more general class of computer algorithms that search for optimal locations for nodes in relatively few dimensions. This general class of search algorithms is called *multidimensional scaling* [17].

An alternative method for placing nodes automatically is grounded in algebra. It is called *singular value decomposition* [18]. Singular value decomposition is not search based. Instead, it uses matrix operations to produce a linear transformation of the data, and thus to position the nodes in one, two or three or more dimensions. There is no guarantee that it will always be effective, but often singular value decomposition provides very good placements of the nodes in few enough dimensions that visualization is possible [19]. A NetDraw image based on singular value decomposition of Mitchell's data is shown in Fig. 13.

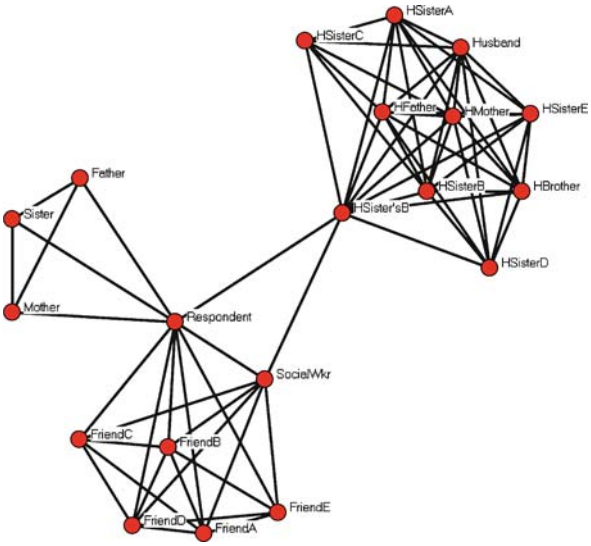
The images in Figs. 12 and 13 both show that the whole network is organized into three densely connected groups that are only loosely linked to one another. That is interesting, but it does not tell us anything about the bases for the groupings. By adding a little information, and continuing to use NetDraw, we can transform the graph of Fig. 12 into a *node labeled* graph, see Fig. 14.

Given the labels, we can identify the homeless woman, the "respondent." We can also see how her network is split up. One division includes her original family, another her friends along with her social worker, and the third contains her estranged husband and his family, her in-laws.

Mitchell's report, however, included even more details. It included estimates of the strength of the tie linking each



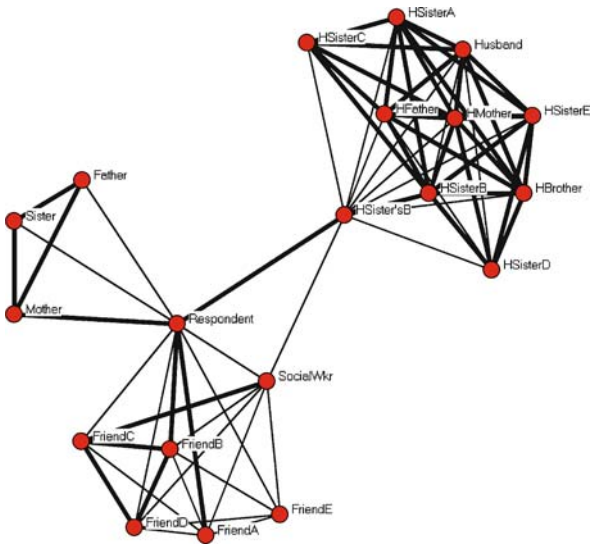
Social Network Visualization, Methods of, Figure 13
Links in the network of a homeless woman III



Social Network Visualization, Methods of, Figure 14
Links in the network of a homeless woman IV

of the pairs of individuals. He classified each tie as either strong or weak. We can embody this additional information in our NetDraw image by an adding another component to our graph. Figure 15, then, was produced using the spring embedder, it is node labeled, and, in addition, it is an *edge labeled* graph.

In Fig. 15 strong ties are indicated by wide edges. By examining their patterning, we learn that the individuals within each family are linked together mostly by strong ties, while the homeless woman's friends have fewer strong ties linking them together. This result is not surprising, but it does provide additional insight about the structural po-



Social Network Visualization, Methods of, Figure 15
Links in the network of a homeless woman V

sition of the woman in question. Clearly, it would be easier for either family to achieve consensus and provide support than it would be for the respondent's loosely connected collection of friends [20].

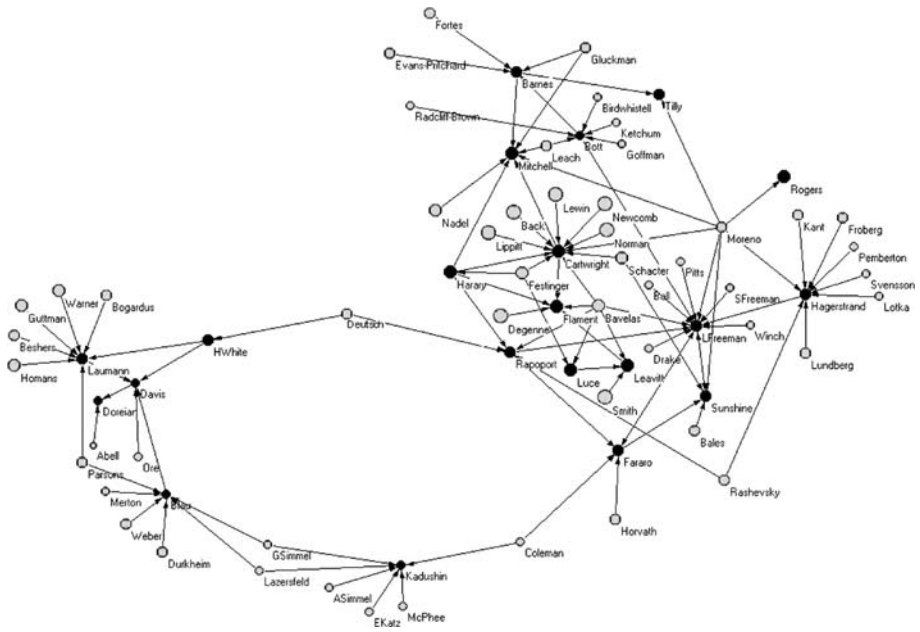
It should be clear, then, that the placement of nodes and the labeling of both nodes and edges are critical for the ability of a graph to communicate important information. Good images can provide investigators with new insights about the structural properties of the social networks they are studying. And they can, of course, help to communicate the results of social network research to outsiders.

Images Based on One Mode Directed Relations

It was obvious from the outset that these simple graphs would not permit many kinds of displays of interest to social network analysts. Even Moreno and Jennings saw the need to display the direction of choice in their sociograms. The direction of connections can be expressed using *directed graphs* or *digraphs*.

A digraph $D = \langle N, A \rangle$ where N is a finite collection of nodes and A is a set of pairs shown as *directed lines* or *arrows*. When an arrow is directed from node a to node b in a digraph, then a is the *tail* of the arrow and b is the *head*; a is the immediate *predecessor* of b and b is the immediate *successor* of a. The *outdegree* of a node is the number of arrows for which it is the tail and its *indegree* is its number for which it is the head.

In any study that involves social links that are not symmetric, digraphs provide a natural representation. Con-

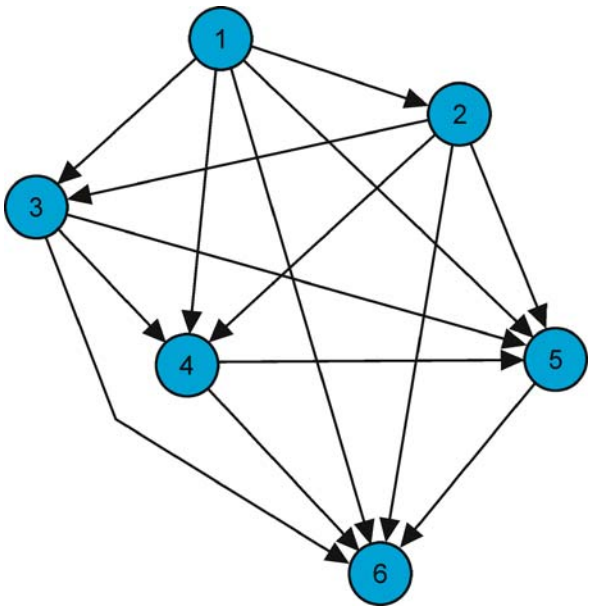


Social Network Visualization, Methods of, Figure 16
Influences on some founders of social network analysis

sider Fig. 16 that was produced by a program called Pajek [21]. In preparing a book on the development of social network analysis, I interviewed a number of the founders. Each was asked to name others who had influenced them to think in network terms. The result is a data set that obviously lacks symmetry.

My interest, however, was with clusters, or blocs, of influentials and nominees. So I placed the nodes using a spring embedder designed by Kamada and Kawai [22]. The resulting figure shows that there seem to be two fairly well defined subgroups, one on the left and one on the right. The two groups are relatively dense but they are only loosely connected together. The people on the left are almost entirely sociologists and those on the right are mostly from other fields. And from the patterning one can suspect that there was some kind of split between these two groups.

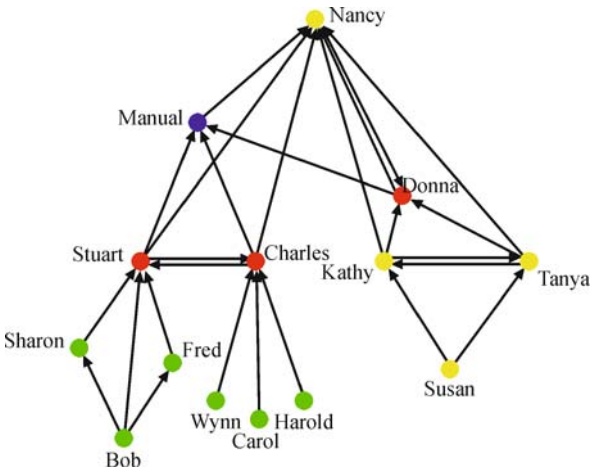
For some kinds of data the search for clusters or groups is not appropriate. For example, when we are dealing with data that should embody some sort of ordering, digraph representations are particularly important. To illustrate how digraphs can be used to display ordering, consider the data collected by Forkman and Haskell [23]. They studied several communities, each made up of six domestic hens. In five of these communities the hens formed strict pecking orders in which the top hen pecked all the others; the second pecked all but the top, and so on. Figure 17 shows a visone [24] image of the data from one of those five



Social Network Visualization, Methods of, Figure 17
Dominance among six hens

communities. There the nodes are arranged, top down, in terms of their outdegrees and the pecking order is obvious.

Often data approach, but do not achieve, a strict order. David Krackhardt [25], for example, collected data on who sought advice from whom among 14 employees in the in-



Social Network Visualization, Methods of, Figure 18
Visone image of advice seeking (from [26])

ternal auditing staff of a large company. Krackhardt's data could not be drawn with all the arrows pointing in one direction. So, in Fig. 18 he arranged the individuals in such a way that as many arrows as possible were pointing up. The viewer, then, can immediately see there is an important hierarchical element displayed by these data. From the image, it appears that Nancy is at the top of the advice chain and Bob, Wynn, Carol, Harold and Susan are at the bottom.

There is, however, an important limitation in this figure. Nancy seeks advice from Donna, Donna seeks advice from Manuel and Manuel seeks advice from Nancy. Thus, these three form a directed cycle of advice seeking. Given such a circular arrangement, no possible hierarchy among these three individuals can be established. Any order in which they were arranged would be misleading. In addition, Stuart and Charles cannot be ordered because they chose each other. The same is true for Kathy and Tanya.

The apparent ordering of nodes in Krackhardt's image was imposed by human judgment. There are computer algorithms that can automatically arrange the nodes into a hierarchical form [26]. They are, however, not as well grounded or reliable as multidimensional scaling and singular value decomposition.

Images Based on Two Mode Relations

Any time we deal with a relation that can link more than two social actors, we cannot use graphs or directed graphs. Both graphs and directed graphs can deal only with links between pairs. Two mode data, however, allow for relations that link three or more actors. So, whenever we have two mode data, like that collected by Hobson [8] or Davis, Gardner and Gardner [13] we need another way to construct images.

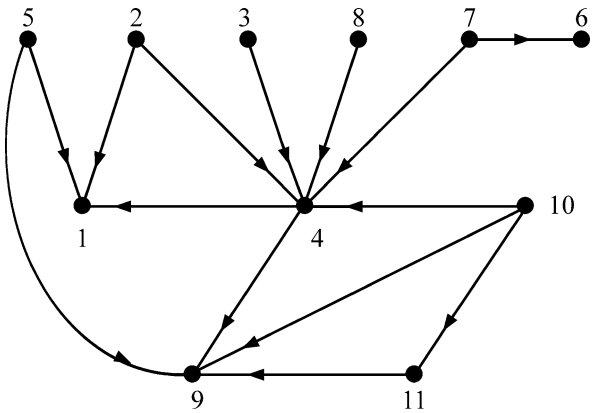
There are several ways to construct images of two mode data. I will consider three of them in the present section, hypergraphs, bipartite graphs and lattices. Then, in the next section, I will discuss the use of matrix representations for both one mode and two mode data.

Hobson [8] collected two mode data on corporations and their directors. He produced the image shown in Fig. 4 showing corporate interlocks as overlapping areas. Mathematically, images like Hobson's are *hypergraphs*. A hypergraph, $F = \langle N, H \rangle$, consists of a set of nodes N and a collection of *hyperedges*, H. While an edge in an ordinary graph connects two nodes, a hyperedge in a hypergraph may link any arbitrary subset of the nodes in N. Pictorially, hyperedges are represented as boundaries enclosing sets of nodes.

The use of hypergraphs was demonstrated in a recent report by Estrada and Rodríguez-Velázquez [27]. They began with one mode data that showed the patterning of predation among the members of eleven species in a Malaysian rain forest. Their graph, showing who preys on whom, is shown in Fig. 19.

Figure 19 shows which species preys on which other species. But if the investigator is interested, as those who study food webs often are, in defining ecological niches in terms of co-predation, Fig. 19 makes the overall pattern less than obvious. As an alternative we can build a hypergraph.

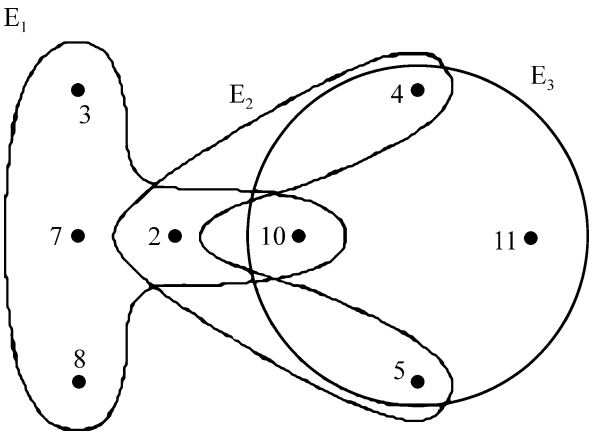
The matrix shown in Fig. 20 is based on the data in Fig. 19. It was built by considering each of the species in turn as prey. Then all of the species that share each given prey are pooled together. Species 1, 6 and 9 have no prey



Social Network Visualization, Methods of, Figure 19
Who preys on whom in a Malaysian Rain Forest

	Prey on		
	4	1	9
2	1	1	0
3	1	0	0
4	0	1	1
5	0	1	1
7	1	0	0
8	1	0	0
10	1	0	1
11	0	0	1

Social Network Visualization, Methods of, Figure 20
Two mode matrix of co-predation



Social Network Visualization, Methods of, Figure 21
Hypergraph of co-predation

in the set. And species 4, 1 and 9 are the targets of co-predation. So the new matrix is two mode. It has the three targets of co-predation as columns and the eight predators as rows.

That matrix is captured visually by the hypergraph in Fig. 19. It immediately reveals that there are three niches. The one labeled E_1 includes all the species who preyed on species 4, E_2 those who preyed on species 1 and E_3 those who preyed on 9. Thus, each edge in Fig. 21 encloses a collection of species that compete directly for at least one prey.

There are, however, other ways to picture two mode data. In a more recent study of corporate interlocks, Joel Levine [29] reported data on the board memberships of seven major American corporations. Those corporations turned out to have ten directors who appeared on two or more of their boards. Levine presented his interlock data using a *bipartite graph*. A bipartite graph, $B = \langle N, E \rangle$ where N is partitioned into two disjoint subsets, N_1 and N_2, and no edge in E has both end points in the same subset. He used singular valued decomposition to place the nodes representing both corporations and board members and produced a bipartite image similar to the one displayed in Fig. 22. I prepared that figure using NetDraw. There, the corporations are shown as red circles and the board members are blue squares. Thus, both the colors and the shapes of the nodes stress the bipartite nature of the graph.

There is still another form of graphic display, one that reveals even more structural information about a two mode data set. It is based on an algebraic procedure called *Galois lattice analysis* or *formal concept analysis* [30,31,32,33]. A Galois or formal concept lattice is defined on an object by property, matrix. Let O be a set of objects and A be a set of attributes. The binary matrix $O \times A$ indicates which objects possess which attributes.

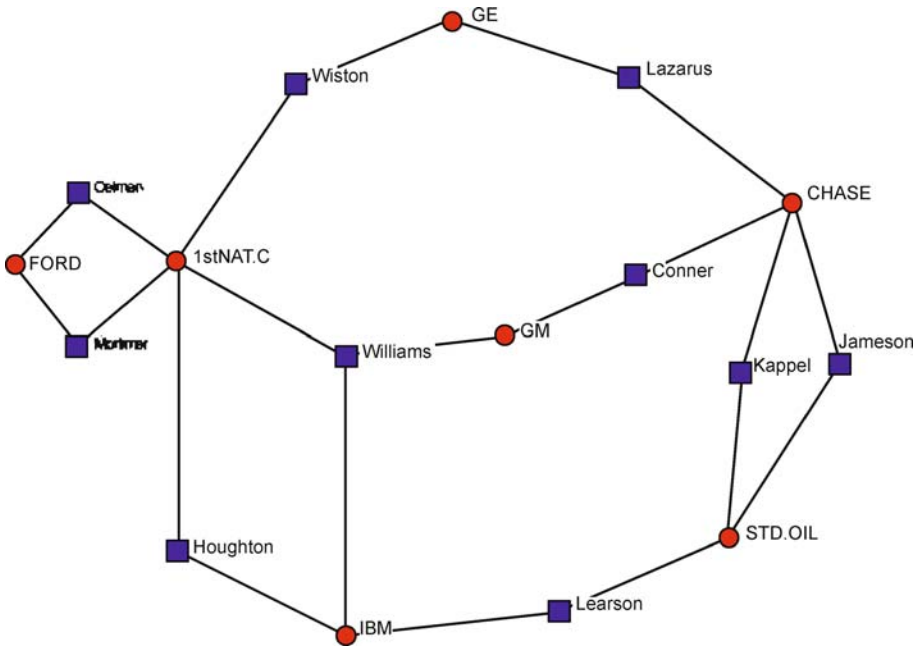
We can define a pair $\langle O_i, A_i \rangle$ such that O_i is a subset of O and A_i is a subset of A and every object in O_i has every attribute in A_i. Moreover, both O and A must be maximal. Thus, for every attribute in A that is not in A_i, there is an object in O_i that does not have that attribute. And for every object in O that is not in O_i, there is an attribute in A_i that the object lacks.

These pairs are dual and they can be partially ordered by inclusion. Given two pairs $\langle O_i, A_i \rangle$ and $\langle O_j, A_j \rangle$ we say that $\langle O_i, A_i \rangle$ is less than $\langle O_j, A_j \rangle$ when O_i is a subset of O_j or, equivalently, when A_j is a subset of A_i. Since all these pairs have unique least upper bounds and greatest lower bounds they form a dual (Galois) lattice.

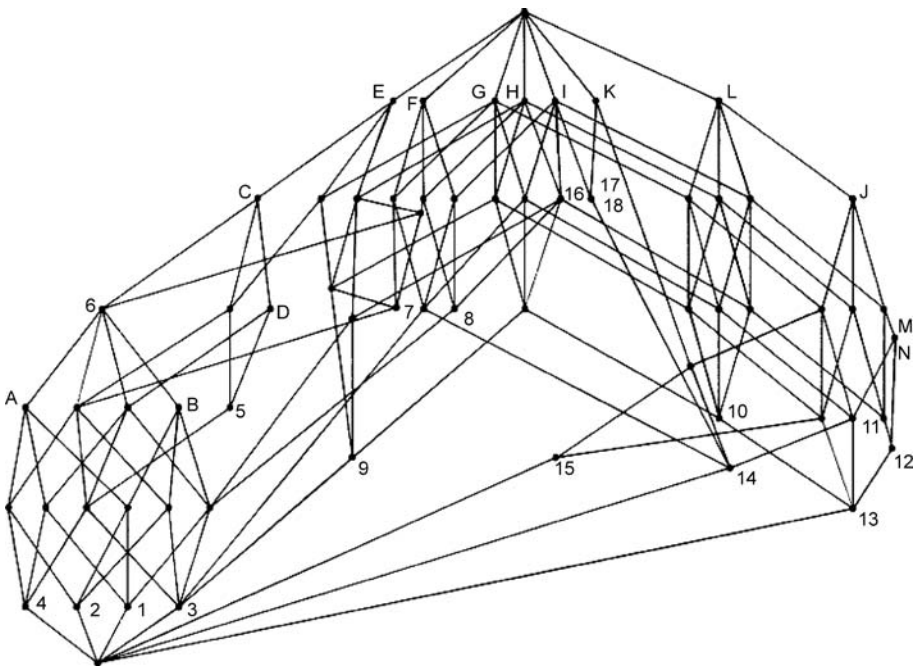
I will illustrate by considering again the woman by event data collected by Davis, Gardner and Gardner [13]. Let the women (1 through 18) be the objects and the events (A through N) be the attributes. The data, arranged into a Galois lattice by a program called GLAD, are shown in Figure into a lattice in Fig. 23.

The lattice displays the same three classes of events that define the same two groups of women that we saw in Fig. 10. But, in addition to the classes of events and groups of women, we can now see the containment structures of both events and women. To begin with, by following lines up from the bottom we can see which women attended which events. When we get to the top we hit the set of all events and, at the same time, because no woman attended all 14 events, it is also the null set of women.

The uppermost events (E–L) involved the largest sets of women. Other events are contained in the lower inter-



Social Network Visualization, Methods of, Figure 22
A NetDraw Image of Levine's interlock data as a bipartite



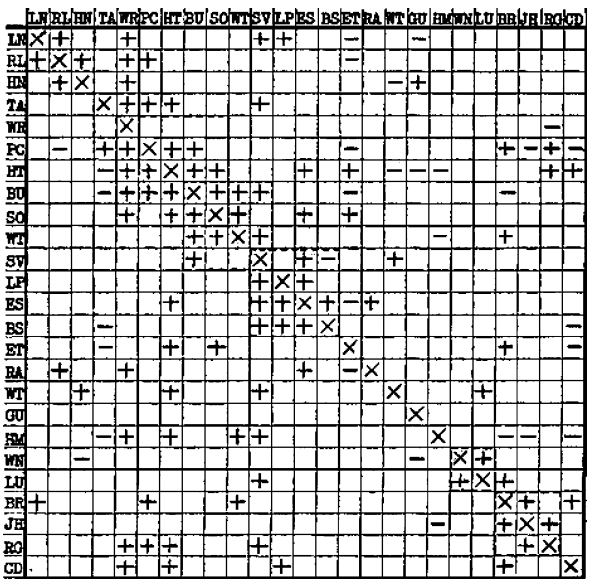
Social Network Visualization, Methods of, Figure 23
The Davis, Gardner and Gardner data as a Galois lattice

sections of these events. Event C, for example, is contained in E; everyone who attended C was present at E. And, at the next lower level, B and D are both contained in C. The events, then, can be seen as varying in their "openness".

At the same time, the figure shows the upward containment structure of the women in terms of their patterns of attendance. Because no event attracted all 18 women, the lowest point represents the set of all women as well as

the null set of events. Then, the lowest set of women (1, 2, 3, 4, 13, 14 and 15) are the "core" attendees, so to speak. The next level contains woman 9 who never attended unless woman 3 was also present, and woman 5 whose attendance depended on that of women 4 and 3. Women 6, 7, 8, 10, 11, 12, 17 and 18 are also at this second level. In some sense, these are all secondary or peripheral participants in these events. And, finally, woman 16 turns out to be a third level participant; she was extremely peripheral. Woman 16 attended events only when secondary attendees 8–12 and core attendees 1, 3 and 13 were all present. All in all, then, the image of the Galois lattice reveals a great deal about the internal structure of attendance.

In this section I have shown three ways of visualizing two mode data. All three of them, however, share one important limitation. That limitation stems from the fact that all three of them can only be used for very small data sets. As the number of cases grows, they all produce images that become increasingly difficult to read.



Social Network Visualization, Methods of, Figure 24
The Forsyth and Katz image of sociometric choices

Images Based on One or Two Mode Data Matrices

When Davis, Davis and Gardner [13] first used matrix permutation, they did so without calling attention to the process. But since that first use a number of contributors have suggested procedures explicitly designed to rearrange the rows and columns of matrices. As time has passed, the overall tendency has been to come up with more effective procedures. And, with the introduction of computers, it has become possible to manipulate ever larger matrices. Presently, there is no end in sight.

Matrix permutations, moreover can be used with either one mode or two mode data. Five years after Davis, Gardner and Gardner introduced matrix permutation in their two mode data set, Elaine Forsyth and Leo Katz [34] explicitly proposed permuting matrices as a way to uncover and display social groups in a one mode data set. They illustrated using data from one of Moreno's [10] sociometric studies. The young women in a residence hall had each been asked to name others in their hall for whom they had positive feelings and those for whom their feelings were negative. Positive choices were recorded using plus signs and negative choices were recorded as minus signs.

Forsyth and Katz adopted a brute-force procedure that involved rearranging rows and columns and redrawing the image again and again until as many of the plus signs fell as close to the principle diagonal as possible. At that point, cohesive groups become visible as clusters of plus signs around the diagonal. Their result is shown in Fig. 24.

Obviously, the Forsyth and Katz procedure was extremely cumbersome. But Beum and Brundage [35] soon came up with a systematic iterative procedure for finding groups by rearranging the rows and columns of a one mode matrix. And, by the late 1950s, when computers emerged on the scene, Coleman and MacRae [36] developed a series of Univac programs at the Operations Analysis Laboratory at the University of Chicago that were designed to uncover the groups in large networks.

An entirely different kind matrix permutation procedure was proposed by Harrison White and his students. They introduced the idea of blockmodeling [37]. In so doing, they provided a theoretical basis for reordering network data matrices, and they developed a number of algorithms for doing so.

The aim of this new thrust was to reorder the matrix in such a way that it could be partitioned to reveal two or more collections of social actors who were *not* linked by some social relation of interest. So, instead of arraying actors along the diagonal of a matrix, White et al. sought permutations that would define zero blocks – sets of actors between which there were no social links. They used their approach to examine a great many network data sets. One example is shown in Fig. 25.

The data in Fig. 25 were collected by Sampson [38] in his study of a monastery. Sampson asked each of a collection of 18 novices to report their relationships with each of the others. Figure 25 shows an 18 by 18 matrix of their responses to a question asking the novices about which oth-

	1	2	3	4	5	6	7	8	9	1	11	12	13	14	15	16	17	18
1										2						1		3
2				3									2					
3		1		3		2												
4	2	3												1				
5			1														3	2
6							3										2	1
7			2	2		3		2									1	
8												1	2			3		
9												1					3	2
1																		
11							3										1	2
12																		
13		3				2						1						
14				2									3				1	1
15			1	3									2					
16			1	3							2							
17				2		3		1										
18				3		1					2							

Social Network Visualization, Methods of, Figure 25
Sampson's data on who was a negative influence on whom

	10	5	9	6	4	11	8	12	1	2	14	15	7	16	13	3	17	18
10																		
5																1	3	2
9								1									3	2
6													3				2	1
4									2	3	1							
11													3				1	2
8								1						3	2			
12																		
1	2													1				3
2					3										2			
14					2										3		1	1
15					3										2	1		
7				3	2		2									2	1	
16					3	2										1		
13				2				1		3								
3				2	3					1								
17				3	2		1											
18				1	3	2												

Social Network Visualization, Methods of, Figure 26
White, Boorman and Breiger's partitioning of the negative influence data matrix from Sampson

0 1 1
1 0 1
1 1 0

Social Network Visualization, Methods of, Figure 27
The CONCOR reduction of the Sampson matrix

ers had negative influences on them. A response of 3 indicated a first choice. A 2 was a second choice and a 1 was a third choice. White et al. reasoned that only first and second choices represented strong responses, so they ignored the third choices and treated the entries of 1 as if they did not exist.

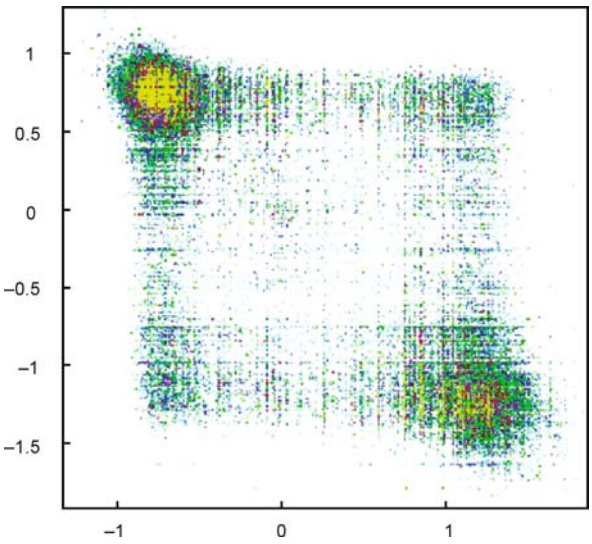
One of the several procedures procedure White et al. introduced was called CONCOR. CONCOR is a recursive procedure that begins by calculating correlations between the rows (or columns) of a network data matrix. Then correlations are calculated between the rows of the resulting correlation matrix. That procedure continues until it produces a matrix of correlations that uniformly displays values of +1 and −1. Those positive and negative values are used to partition the individuals into two subsets. The CONCOR procedure can be repeated using the data contained within each of the partitions. Thus, the original matrix can thus be refined to any desired degree.

White and his students used CONCOR on the data shown in Fig. 25 in an attempt to uncover blocks that contained only 0s. They could then use these zero blocks to reduce the complexity of the data matrix. That matrix produced three zero blocks. They are shown in Fig. 26.

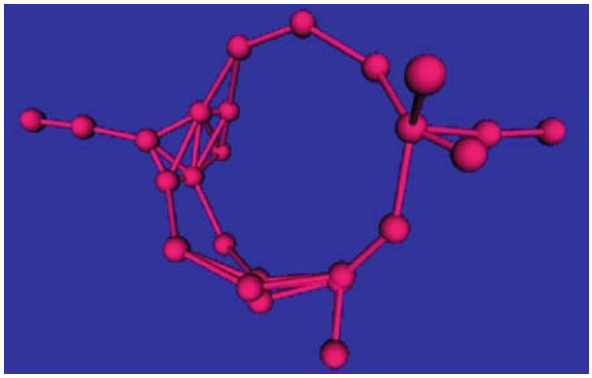
In the reduced model in Fig. 27 each cell represents one of the blocks in Fig. 26. Thus, the 18 by 18 matrix is reduced to a 3 by 3 array. The reduction is consistent with Sampson's original ethnographic description of subgroups among the novices. Moreover, its pattern of zero blocks in the principle diagonal indicates that no block member saw any fellow block member as having a negative influence. But the members of each block saw at least some of the members of both of the other blocks as negative influences. This makes sense in the light of the ongoing conflict that Sampson described in his report.

Since that time, displays based on matrix permutations have grown in size, complexity and sophistication. One particularly striking example was produced by Richards and Seary [39]. Their data were drawn from a study of participants in a needle exchange program in Baltimore, Maryland [40]. Richards and Seary examined data on 4259 individuals who picked up and returned needles at each of four exchange sites over a 30 month period. Each cell in the matrix is a record of the number of needles picked up by the individual in that row and returned by the individual in that column. About a third of all needles fall in the principle diagonal of the matrix.

Richards and Seary used the data from the largest weak component in the data set. That component involved 100 000 needle exchanges among 36 000 individuals. Richards and Seary used their program MultiNet [41] scaled the data using a form of singular value decomposition called *correspondence analysis* [18]. They used the coordinates provided by the first Eigenvector to reorder the rows and columns of the matrix. They then colored the entries in terms of frequencies. The color scale is logarithmic: gray is 1 needle, blue 2–3, green 4–7, red 8–15, ma-



Social Network Visualization, Methods of, Figure 28
The largest component in the Baltimore needle exchange data



Social Network Visualization, Methods of, Figure 29
VRML image of friendship among teens in a Dublin suburb

genta 16–31, yellow 32 and above. Their image is shown in Fig. 28.

Figure 28 dramatically illustrates the utility of images based on matrix permutation. It shows that there was not a single community of needle users in Baltimore. Instead, there were two distinct communities of individuals who regularly obtain, return and exchange needles with one another. These two relatively large communities were centered around two of the four needle exchange sites.

Future Directions

Overall, the long range the trend in visualizing social networks has been to rely on computers to do more and more of the job. First, computers used a version of singular value decomposition to locate nodes in two dimensional images [42]. Then, soon thereafter, Coleman and MacRae [35] programmed a computer both to permute rows and columns of a matrix and to print out an image of the result. And, in the early 1970s, Alba [43] wrote a program that performed calculations to place nodes and then went on to draw node and edge images of the results [44].

Since the 1970s, then, network analysts have increasingly used computers both for calculations and to draw images. And increasingly, multidimensional scaling and singular value decomposition have been used to determine locations for nodes. Moreover, when two dimensions are not enough to display network structure, three dimensional images are being produced.

When microcomputers became available it quickly became possible to produce images that gave the appearance of being three dimensional. Moreover, with the advent of color screens, color images began to be produced. Figure 29 represents data collected by Kirke [45] on social links among teenagers in a suburb of Dublin. Nodes were first located in three dimensions using multidimensional scaling. And then the virtual reality modeling language (VRML) was used to produce the appearance of three dimensions. Figure 27 was produced as a cover design for a book [46] and the colors were used simply make the image more attractive.

Colors can, however, be used to enhance the ability of an image to communicate important information. In Fig. 30 Höpner and Krempel [47] used a spring embedder and Krempel's own programs to arrange the nodes in two dimensions. The nodes represent the 100 largest German corporations in the year 2000. They used color to label both nodes and directed lines. In their image each company is represented as a node and an arrow pointing from one node to another means that the first node holds shares in the second. The size of a node indicates the number of connections to other nodes it has. Financial companies are shown as yellow nodes and industrial companies are red. Links between financial companies are yellow, those between industrial companies are red and links between financial and industrial companies are orange. Bu using color, then, his directed graph reveals a great deal of information about the organization of German industry and finance.

The image in Fig. 31 was made using a program called MAGE [48]. MAGE was written by Richardson and Richardson [49]. It is designed as a display on computer screens, and it allows the viewer to move into the picture as well as to spin and rotate it. It is useful, then, for exploring the patterning of structural data in three apparent dimensions.

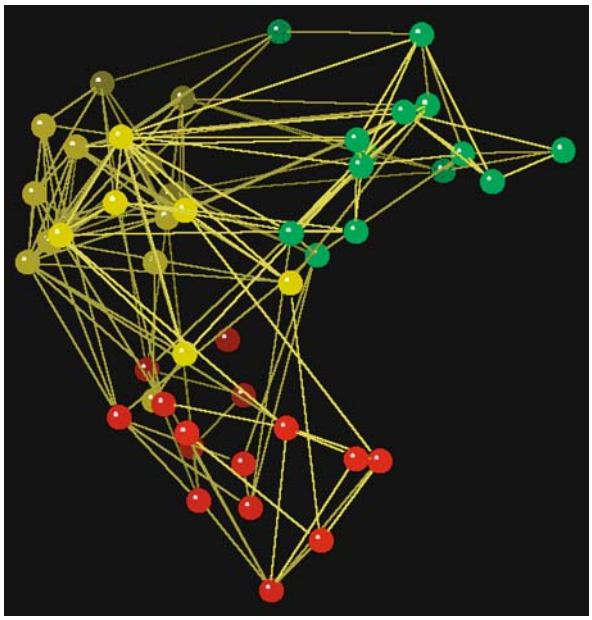
Social Network Visualization, Methods of, Figure 30
Links among German corporations in 2000

In Fig. 31 students in a university professional school program reported who their friends in their class were. Nodes were placed using a multidimensional scaling program and then they were colored according to their program in the school. It turned out that most of their friendship choices linked those that shared a program.

Richards and Seary's [50] program, Multinet, produces a wide range of graphic images. Included are images that actually can be viewed in three dimensions using *anaglyphic* glasses in which one lens is red and one is blue. Obviously, I cannot illustrate their program here, but anyone who wants to see real 3D images should explore Multinet.

The most recent development in visualizing social networks involves the production of animated graphics. As more and more process data are collected and as more process models are constructed, animated images are a natural development. A group at Stanford University has written a Java program, SoNIA [51], that makes it quite simple to produce animated node and edge and node and directed line images [52,53]. These images allow users to explore the changing structural forms generated by process data.

Overall then, in the period between Moreno's hand drawn ad hoc images and the latest animations of dynamic network processes, there has been a dramatic growth in our ability to visualize social network structure. The major contribution has come from computers. Today we can use a wide variety of readily available computer programs to both design images and to produce screen images and/or printed output.



Social Network Visualization, Methods of, Figure 31
MAGE image of friendship among classmates

But, as the job of producing images becomes easier, we must be careful not to lose our sense of why we are pro-

ducing then in the first place. From the very beginning, the important point has always been that the visual images of social networks are not produced simply to be decorative. In every case, the early images were drawn in order to dramatize some feature of social structure. Moreno produced Fig. 5 to illustrate the importance of considering the number of connections in evaluating the structural position of an individual. In Fig. 6 the number of negative ties received by one of the running backs showed, as Moreno (p. 213 in [10]) put it, "It is easy to see that when 5/RB is running with the ball he is not apt to get the maximum of cooperation in interference and blocking."

Figure 7 was a pictorial statement by Warner and Lunt that when cohesive subgroups overlap, they should not be expected to bridge wide differences in social class. Figure 8, from Davis, Gardner and Gardner, demonstrated that the Warner–Lunt hypothesis was supported by data with respect to both social class and age. And, finally, Fig. 9 illustrated the presence of cohesive groups and of the variation of different individuals in their involvement in those groups. In every case each of these early authors had a point to make, and in every case the image helped to make that point. That is the key to the effective use of visual materials in social network analysis.

In future we can expect to see continued development of computer programs designed to aid in visualizing social networks. We can look forward to continued refinement of algorithms for displaying group structure that are based on multidimensional scaling, particularly spring embedding. We can anticipate better algorithms for displaying hierarchies and approximate hierarchies. We can expect to have more powerful programs for animation. And, at the same time, we can expect to be able to produce higher quality and more refined visual displays of all sorts.

Bibliography

Primary Literature

1. Freeman LC (2004) The development of social network analysis: A study in the sociology of science. Empirical Press, Vancouver, BC
2. Schweitzer VS (1998) Words of power. Coffee Times, Fall
3. Cayley A (1857) On the theory of the analytical forms called trees. Philos Mag 13:19–30
4. Biggs NL, Lloyd EK, Wilson RJ (1977) Graph theory 1736–1936. Oxford University Press, Oxford
5. Morgan LH (1871/1997) Systems of Consanguinity and Affinity in the Human Family. University of Nebraska Press, Lincoln
6. Macfarlane A (1883) Analysis of relationships of consanguinity and affinity. J R Anthropol Inst Great Britain Ireland 12:46–63
7. Appendix to Macfarlane A (1883) Analysis of relationships of consanguinity and affinity. J R Anthropol Inst Great Britain Ireland 12:46–63
8. Hobson JA (1884) The evolution of modern capitalism; A study of machine production. Allen & Unwin, Macmillan, London, New York
9. Moreno JL (1932) Application of the group method to classification. National Committee on Prisons and Prison Labor, New York
10. Moreno JL (1934) Who shall survive? Nervous and Mental Disease Publishing Company, Washington
11. Roethlisberger FJ, Dickson WJ (1939) Management and the worker. Harvard University Press, Cambridge
12. Warner WL, Lunt PS (1941) The social life of a modern community. Yale University Press, New Haven
13. Davis A, Gardner B, Gardner MR (1941) Deep south. University of Chicago Press, Chicago
14. Mitchell JC (1994) Situational analysis and network analysis. Connections 17:16–22
15. Available at http://www.analytictech.com/downloadnd.htm
16. Eades P (1984) A heuristic for graph drawing. Congressus Nutnerantiunt 42:149–160
17. Krempel L (1999) Visualizing networks with spring embedder: Two-mode and valued data. In: Proceedings of the section of statistical graphics. American Statistical Association, Alexandria, pp 36–45
18. Weller SC, Romney AK (1990) Metric scaling: Correspondence analysis. Sage Publications, Newbury Park
19. Freeman LC (2005) Graphical techniques for exploring social network data. In: Carrington PJ, Scott J, Wasserman S (eds) Models and methods in social network analysis. Cambridge University Press, Cambridge, pp 248–269
20. Bott E (1957) Family and social network. Tavistock, London
21. Available at http://vlado.fmf.uni-lj.si/pub/networks/pajek/
22. Kamada T, Kawai S (1989) A general framework for visualizing abstract objects and relations. ACM Trans Graph 10:1–39
23. Forkman B, Haskell MJ (2004) The maintenance of stable dominance hierarchies and the pattern of aggression: Support for the suppression hypothesis. Ethology 110:737–744
24. Available at http://visone.info/download/
25. Krackhardt D (1996) Social networks and the liability of newness for managers. In: Cooper CL, Rousseau DM (eds) Trends in organizational behavior, vol 3. Wiley, New York, pp 159–173
26. Brandes U, Raab J, Wagner D (2001) Exploratory network visualisation: Simultaneous display of actor status and connections. J Soc Struct 2:4
27. Estrada E, JA Rodríguez-Velázquez. (2005) Complex networks as hypergraphs. Physica A 364:581–594
28. Available on request from its author, Vincent Duquenne (v.duquenne@wanadoo.fr)
29. Levine JH (1979) Joint-space analysis of 'pick-any' data: Analysis of choices from an unconstrained set of alternatives. Psychometrika 44:85–92
30. Wille R (1982) Restructuring lattice theory: An approach based on hierarchies of concepts. In: Rival I (ed) Ordered Sets. Reidel, Dordrecht–Boston, pp 445–470
31. Wille R (1984) Line diagrams of hierarchical concept systems. Int Classif 11:77–86
32. Duquenne V (1987) Contextual implications between attributes and some representation properties for finite lattices. In: Ganter B, Willie R, Wolff K (eds) Beiträge zur Begriffsanalyse. B. I. Wissenschaftsverlag, Berlin, pp 213–240

33. Freeman LC, White DR (1993) Using Galois lattices to represent network data. In: Marsden PV (ed) Sociological methodology. Blackwell, Oxford, pp 127–146
34. Forsyth E, Katz L (1946) A matrix approach to the analysis of sociometric data: Preliminary report. Sociometry 9:340–347
35. Beum CO, Brundage EG (1950) A method for analyzing the sociomatrix. Sociometry 13:141–145
36. Coleman JS, MacRae D (1960) Electronic processing of sociometric data for groups up to 1000 in size. Am Sociol Rev 25:722–727
37. White HC, Boorman SA, Breiger RL (1976) Social structure from multiple networks: I. Blockmodels of roles and positions. Am J Sociol 81:730–779
38. Sampson SF (1969) A noviate in a period of change: An experimental and case study of relationships. Unpublished Ph D dissertation. Department of Sociology, Cornell University
39. Richards WD, Seary AJ (2000) Cover illustration. Connections 23:1
40. Valente TW, Foreman RK, Junge B, Vlahov D (1998) Satellite exchange in the Baltimore needle exchange program. Public Health Reports 113:91–96
41. Availble by contacting Bill Richards (richards@sfu.ca)
42. Bock RD, Husain SZ (1952) Factors of the tele: A preliminary report. Sociometry 15:206–219
43. Alba R (1972) SOCK. Behav Sci 17:326–327
44. Kadushin C (1974) The american intellectual elite. Little, Brown, Boston
45. Kirke DM (1996) Collecting peer data and delineating peer networks in a complete network. Social Netw 18:333–346
46. Kirke DM (2006) Teenagers and substance use: Social networks and peer influence. Palgrave Macmillan, Basingstoke, Hampshire and New York
47. Höpner M, Krempel L (2003) The politics of the German company network. Max Planck Institute for the Study of Societies, Working Paper 03/9
48. MAGE is available at http://kinemage.biochem.duke.edu/software/mage.php
49. Richardson DC, Richardson JS (1992) The kinemage – a tool for scientific communication. Protein Sci 1:3–9
50. Richards WD, Seary AJ (2006) Multinet. http://www.sfu.ca/~richards/Multinet/Pages/multinet.htm
51. SoNIA is available at http://www.stanford.edu/group/sonia/
52. Moody J, McFarland DA, Bender-deMoll S (2005) Visualizing network dynamics. Am J Sociol 110:1206–1241
53. Bender-deMoll S, McFarland DA (2006) The art and science of dynamic network visualization. J Social Struct 7:2

Books and Reviews

Brandes U, Kenis P, Raab J, Schneider V, Wagner D (1999) Explorations into the visualization of policy networks. J Theor Polit 11:75–106
Brandes U, Corman SR (2003) Visual unrolling of network evolution and the analysis of dynamic discourse. Information Visualization 2:40–50,
Brandes U, Kenis P, Wagner D (2003) Communicating centrality in policy network drawings. IEEE Transactions on Visualization and Computer Graphics 9:241–253,
Brandes U, Wagner D (2004) Netzwerkvisualisierung. Inform Technol 46:129–134,
Brandes U, Wagner D (2004) Visone - Analysis and visualization of social networks. In: Jünger M, Mutzel P (eds) Graph drawing software. Springer-Verlag, pp 321–340
Brandes U, Kenis P, Raab J (2006) Explanation through network visualization. Methodology 2:16–23
Frank O (2000) Structural plots of multivariate binary data. J Social Struct 1:4
Freeman LC (2000) Visualizing social groups. In: Proceedings of the section on statistical graphics (1999). American Statistical Association, pp 47–54
Freeman LC (2000) Visualizing social networks. J Social Struct 1:1
Freeman LC, Webster CM, Kirke DM (1998) Exploring social structure using dynamic three-dimensional color images. Social Netw 20:109–118
Johnson JC, Krempel L (2004) Network visualization: The 'Bush Team' in Reuters News Ticker 9/11–11/15/01. J Social Struct 5:1
Klovdahl AS (1981) A note on images of networks. Social Netw 3:197–214
Klovdahl AS (1998) A picture is worth…: Interacting visually with complex network data. In: Liebrand WBG (ed) Computer modeling of dynamic social processes. Sage, London
Krempel L (2005) Visualisierung komplexer Strukturen. Grundlagen der Darstellung mehrdimensionaler Netzwerke. Campus, Frankfurt aM
Krempel L, Schnegg M (2005) About the image: Diffusion dynamics in an historical network. Structure and dynamics: eJ Anthropol Rel Sci 1:1, 10
McGrath C, Blythe J, Krackhardt D (1996) Seeing groups in graph layouts. Connections 19:22–29
McGrath C, Blythe J, Krackhardt D (1997) The effects of spatial arrangements on perceptions of graphs. Social Netw 19:223–242
McGrath C, Krackhardt D, Blythe J (2003) Visualizing complexity in networks: Seeing both the forest and the trees. Connections 25:37–47
McGrath C, Blythe J (2004) Do you see what i want you to see? The effects of motion and spatial layout on viewers' perceptions of graph structure. J Social Struct 5:2

Social Organizations with Complexity Theory: A Dramatically Different Lens for the Knowledge Economy

Russ Marion
Clemson University, Clemson, USA

Article Outline

Glossary
Definition of the Subject
Introduction
Basic Premises
Understanding Organization with Complexity
Complexity Outcomes: Creativity, Adaptability, Learning
Methodology for Studying Complex Organizations

Conclusions
Future Directions
Bibliography

Glossary

Many of these terms are commonly understood by complexity theorists; these definitions, however, describe their meaning within the context of social organizations.

Adaptive behaviors Perturbations (creative ideas, pressures, etc.) in a system that foster some observable level of phase transition. Change leading to systemic reaction.

Adaptive tension Pressures external or internal to the organization that perturb an organization thus pressuring it toward phase transitions and structural or ideational elaboration.

Administrative leadership behaviors related to such things as strategic planning, policy making, and resource acquisition and distribution.

Agent An entity that is the smallest unit of interest in a complex dynamic. An agent could be a person, idea, task, knowledge, etc.

Aggregation Coevolutionary emergence of diverse ideas, agents, etc.

Aggregation mechanisms Macro-level complexity mechanisms. Refers to the dynamic interactions of perturbations (see initiating mechanisms), amplifications (expansion and elaboration of a perturbed state), and phase transition (nonlinear, form-shift in some part of an organization).

Coevolution A process in which "reciprocal selective pressures operate to make the evolution of … (one agent in an interactive process) partially dependent on the evolution of (other agents)" [67]. Coevolution occurs when heterogeneous agents in a niche are interdependent such that they dynamically adjust to each others changes.

Commodity-based economy Refers to a manufacturing economy in which the major assets are physical commodities.

Complexity organizational knowledge Systemic knowledge that imbues organization with capacity for adaptive and creative responses; attributable to the strength and viability of the system's complex structures and to the viability of the relationship between the organization's complex structures and its bureaucratic structures.

Complexity mechanism Social mechanisms that underlie complexity dynamics.

Complexity theory In organizational sciences, the study of emergent dynamics in neural-like networks of adaptive, vision-oriented agents.

Conflicting constraints Conflict that emerges when the preferences of one agent challenges the preferences of another. In complex networks, conflicting constraints generate pressure to elaborate.

Dissipative structures Structures that emerge when far from equilibrium systems release excess energy in a phase transition.

Emergence The appearance of new structures or ideas from the actions of complex interactions.

Enabling behaviors Activities that foster conditions (e. g., interdependency, enabling rules, adaptive tension) in which complex dynamics can emerge.

Enabling rules Rules that govern interactions among adaptive agents in complex systems. Contrast with bureaucratic rules, which delimit the responsibilities of agents in a closed system bureaucracy.

Entanglement A dynamic relationship between the formal administrative forces and informal complexly adaptive, emergent forces of an organizational systems.

Equilibrium A stable, predictable relationship among agents and structures of a relationship. Related to thermodynamic concept of a low energy state.

Extreme event Instances of dramatic change that occur infrequently in social organizations; the US government's experience of the Katrina hurricane in New Orleans illustrates. Extreme events typically require rapid action and can be minimally responsive to top-down control.

Far from equilibrium Typically defined as a high energy state; defined here for organizational studies as an intensely complex, dynamic state driven by excess levels of pressure and perturbations.

Heterogeneity A diversity of skills, worldviews, preferences, beliefs, goals, (etc.) among interactive agents in a complex system.

Initiating mechanisms Micro-level complexity mechanisms. They include coevolutionary interaction and perturbations (an unexpected change in an interaction relationship; attributable to complex, interactive dynamic).

Interdependency Network conditions in which the actions of one agent are influenced by the actions of another.

Knowledge economy Refers to a production economy in which the major assets are the knowledge possessed by individuals and networks of individuals (organizational knowledge).

Meso theory Variously defined as theory that bridges macro and micro level theory; different levels of hierarchy; or different levels of analysis (individual, dyadic, group).

Multi-agent based modeling Agent based modeling procedure that analyzes different types of networks simultaneously (e. g., agent networks, task networks, etc.)

Organizational level From Jaques [40], the level of bureaucracy that lies between the upper echelon levels and the work production level; includes middle management.

Perturbations Events that disturb normal organizational interactions and generate pressures that can lead to phase transitions.

Phase transition Sudden, nonlinear form shifts attributable to the dissipation of accumulated pressure in complex systems. Phase transitions can occur at multiple levels of intensity.

Postmodernism In general, a rejection of scientific modernism. For organizational science, it represents a realization that organizational behaviors cannot be adequately expressed as mathematical relationships among variables. Complexity theory adds that organization is an ultimately unpredictable dynamic whose causal structure is based on interactions among complexity mechanisms.

Production level From Jaques [40], the level of bureaucracy responsible for line production.

Requisite complexity McKelvey and Boisot's [65] modification of Ashby's requisite variety; they maintain that viable organizations are at least as complex as their competition.

Requisite variety Ashby's [5] dictum that viable organizations have at least the same degree of flexibility as their competition.

Spaces between A proposal that creative ideas emerge from interactive dynamics. Creative emergence is generated when agents work to resolve tension; such tension is product of conflicting constraints, heterogeneous preferences, ideas, and knowledge, etc.

Social mechanism A process attributable to dynamic interactions among multiple people, variables, ideas, physical limitations, etc. Mechanisms describe "a set of interacting parts – an assembly of elements producing an effect not inherent in any one of them" p. 336 in [27], p. 74 in [37].

Strategic level From Jaques [40], the upper echelon level of bureaucracy; responsible for organizational strategy, policy making, resource acquisition and allocation, etc.

Swarm behavior Study of the dynamics of swarms in biology and in organizations. Scientists are finding that amazingly complex, "intelligent" behavior can emerge from complex behaviors that are structured by a few, simple enabling rules. Useful in organizations for improving such things as distribution efficiency or flexible response to environmental conditions.

System dynamics Network simulations in which relationships among agents are defined mathematically.

Top-down administration Administrative practices in which decisions are made by superiors to be carried out by subordinates.

Vision A teleological view of the future. Vision can be highly specific, or determinate, as when an organization projects future markets (such visions are more properly, goals; see Sharp Corp.'s strategy statement in this paper for an example). Vision statements appropriate for complex systems are indeterminate in that they do not preclude the future; for example, "This company will strive to enhance its competitive advantage by optimizing its flexibility."

Definition of the Subject

Complexity theory for organizations examines the influence of complex dynamics on (among many things) organizational structure, leadership, power and control, influence, and strategy. It is applicable not only to understanding but of practice in any organizational type whose primary commodity is knowledge and application of knowledge. The core outcomes of complex dynamics are creativity, adaptability, and learning. Complexity theory is particularly germane in what has come to be called, the knowledge economy.

This paper defines basic premises underlying complex dynamics in organizations, and argues in particular that organizational complexity is best understood as the interactions among complexity mechanisms (causal process attributable to dynamic interactions among multiple people, variables, ideas, etc.) rather than as the outcome of defined variables. We explore the relationship between bureaucracy and complexity, and define three levels of behavior in complex organizations: administrative, enabling, and adaptive. We apply these ideas along with complex premises underlying complexity to strategic leadership. The paper describes how complexity produces its core outcomes – creativity, adaptability, and learning – and argues that these outcomes are particularly pertinent to today's economy. It concludes with a discussion of unique research methodologies that can be used to study a process based on mechanisms rather than on variables.

Introduction

The science of complexity offers interesting solutions to some difficult challenges in organizations. These challenges affect nearly all economic sectors of society, from electronics to education, computer programming to automotives, banking, government, and the armed forces. They include hyper-competitive environments [38], globalization, unequal global distribution of wealth and production costs, rapid technological and social change, highly complex operations (e. g., the complexity of sophisticated software programming), and the changing nature of the workforce (less loyal, comfortable with electronic communication, demanding of flexible work arrangements). To address these challenges, practitioners and scholars recognize the need for radically new leadership paradigms [81], revised perspectives of organizational strategy [81], flexible structures [83], rapid adaptability [16], and decentralized problem-solving structures [25]. Complexity science is applicable to these challenges because, among other things, it focuses on flexible, interactive dynamics rather than top-down control, it offers unique descriptions of change in dynamic system, it reveals important clues about creativity, adaptability, and learning in organizations, and it shifts attention from central tendency variables to causality based on dynamic mechanisms [50,81].

Traditional thought in organization science was oriented toward efficient production and top-down coordination. It assumed that organizational behavior was intimately dependent upon structure provided by positional leaders, rational plans and goals, centralized coordination, organizational evolution, and vision [35]. It has explored the nature and impact of power [41], institutional forces [28], and charisma [23]. Traditional thought has, like complexity theory, described change but has assumed that organizational change and innovation must be planned, that effective action must be coordinated, and that organization must be structured in the board room.

By the 1990s, this "top-down" notion began losing credibility among organizational scholars and practitioners (e. g., [35]). The major culprit in this demise was the knowledge economy – the globalization, rapid change, and hyper-competitive environments described above. Top-down perspectives of organization had been oriented toward more stable, commodity based environments [9], and the emerging knowledge economy is oriented toward creativity, adaptability, and learning.

Complexity theorists in the organizational sciences seek relevant strategies for dealing with the knowledge economy. Complexity theory is useful for this because it does not merely improve traditional perspectives of doing organization but rather addresses different sets of behaviors than are typically addressed by organizational theorists. In particular, it focuses on informal interactive processes within an organization, on processes that emerge from informally interacting people and groups, and on productive relationships between informal and formal activities within an organization. It asks such questions as, "How can organizational decision making and interactive patterns be restructured and reorganized to increase flexibility" [49]? "How can leadership be re-conceptualized in a way that is effective for complex organizations" [70]? Or, "How do we re-focus organizational strategy for a knowledge economy" [58]?

In this paper, we look first at basic premises underlying the application of complexity theory to organizations, then look at how complexity can inform understanding of bureaucracy and strategic leadership. We examine the organizational outcomes of complex dynamics (hence the rational for applying this science), and conclude with a discussion of research methodologies that are useful for studying complex organizational behaviors.

Basic Premises

Studies have found that, like biological [30], economic [4], and physical [45] systems, emergent dynamics are readily observed among interacting people, or agents [20,68,69]. Interactive dynamics and emergence are common denominators in complexity studies, and while there are discipline-specific variations (organizational agents, for example, make learned, adaptive choices while weather systems react primarily to the interactions of physical properties), all are shaped by these defining dynamics. Agents interact and the interaction changes them (emergence).

Complexity organization researchers are interested in understanding how these dynamics operate in human systems. Lichtenstein and Plowman [51] have found, for example, that certain interactions among individuals and groups lead to perturbations, or pressure to elaborate. Consistent with Kauffman [43], these tensions pressure the system to adapt and to change. Perturbed, interactive systems also tend toward self-organization [57,62]. This phenomenon was first systematically observed in the 1940s and 1950s as informal groups [39] and was expanded by complexity theorists to include any situation in which ideas or people spontaneously, and without external motivation, act in some degree of synchrony with one another. Chiles et al. [20] identified dissipative structures – fluctuation, positive feedback, stabilization and re-

combination – in the emergence of the music industry at Branson, Missouri.

Researchers are also interested in understanding the conditions that enable emergent, self-organizing behavior in social systems [82]. *Interaction*, of course, is *sine qua non*. Schreiber, Marion, Uhl-Bien, and Carley [74], drawing from Kauffman [43], found that interactive agents must possess a requisite level of *interdependency* to enable learning outcomes (too much or too little will suppress complex dynamics; [43]). This interdependency generates *conflicting constraints*, which are important sources of perturbations in a complex system. McKelvey [64] adds that such pressure is also generated by external sources such as leadership behavior or economic pressures; he refers to this as *adaptive tension*. Uhl-Bien et al. [82] adopted this term to refer to both internal and external pressures.

Bonabeau and Meyer [10] have developed fascinating insights into the nature of swarm behavior in human organizations, and have used it to solve efficiency problems involving such things as package delivery and local surges in phone traffic [10]. Swarms are a function of a set of simple, *enabling rules*, such as "fill orders [in a distribution warehouse] until the next person in line takes over", or "drop 'digital' pheromone trails for subsequent messages to track" (in computerized instructions for phone systems).

Complex behaviors in social systems are enabled by *heterogeneity* of skills, preferences, outlooks, etc. Traditionally, organizational theorists, particularly those who study leadership, have sought to align employees to centralized goals [56]; complexity theory suggests that organizations whose employees pursue a diversity of goals (within the context of an interactive, interdependent network) will have a diversity of ideas to draw from as they seek creativity and adaptable changes.

Finally, Schreiber, Marion, Uhl-Bien, and Carley [74] found in a multi-agent based simulation that moderate levels of *vision* (or general organizational focus) are conducive to complex behaviors. High centralization of vision (in combination with high interdependency among agents) was counter-productive for complexity outcomes in their simulation, as was low levels of vision (low vision was more problematic to outcomes than was high vision).

Emergence and Phase Transitions

Emergence refers to the generation of new structures or ideas due to the actions of complex interactions. This phenomenon is important because it allows the system to rapidly produce adaptive, creative responses to environmental conditions. It occurs in social structures when heterogeneous agents or ideas act in relative synchrony with one another. The emergence of riot conditions out of the behaviors of just a few people illustrates (the emergence of a riot also illustrates that emergence is not necessarily a functional response to environmental contingencies; the notion of counter-productive emergence has been discussed by Uhl-Bien, Marion and McKelvey [82]).

One form of emergence involves the *accumulation* of different ideas or agents. Anderson [1] and Marion [55], for example, described the emergence of technology that led to the invention, in 1975, of the microcomputer. It began with disparate advances – transistors, printed boards, LEDs, etc. Engineers began using these isolated components together, as they did in the transistor radio popular in the 1960s. Accumulation took a big leap with the introduction of the handheld calculator in the late 1960s, which added such things as processors and information storage capacity. The microcomputer was the ultimate "accumulation" or emergence event.

Importantly, the components of such emergence processes accumulate because they "coevolve." Coevolution is defined by Pianka, p. 329 in [67], as a process in which "reciprocal selective pressures operate to make the evolution of either … [agent in the interactive process] partially dependent on the evolution of the other" see also McKelvey [64]. Coevolution occurs when heterogeneous agents in a niche are interdependent such that they dynamically adjust to each other's changes. It raises the important observation that an organizational species' fitness – the capacity to survive and thrive – is influenced by supporting roles from other organizational species [43].

A second form of emergence is called *phase transition*. Phase transitions are sudden, nonlinear form-shifts (changes in structure, knowledge, worldview, etc.). The sudden demise, in 1989, of the USSR illustrates this phenomenon. Phase transitions like that experienced by the USSR result from a build up of tension or pressure [32,64,71]. Haken [32], for example, illustrated phase transition in terms of slowly heated oil; eventually the heat (tension) reaches a point at which the oil precipitously transitions to a new state, observed as a gentle boil.

Phase transitions, occur at multiple levels of intensity, and a number of researchers have observed that, in complex systems, intensity of transitions plotted against their frequency describes a power law relationship ($1/f^x$); [2, 7,75]. Power law curves are steeply descending slopes that quickly bottom out into a more gently descending slope; when intensity and frequency are converted to log equivalents, the slope becomes a straight line descending left to right like a negative regression line in statistics. Power law curves describe situations in which high intensity events

occur infrequently while low intensity events occur often.

This has opened a line of exploration into what are called, extreme events, or analyzes of phenomena that are highly intense but of low frequency (the 2008 Organization Science Winter Conference was devoted to this topic). Stephenson and Bonabeau [76] concluded, for example, that bottom-up, or adaptive, leadership strategies may be more effective in extreme events such as Katrina or the Twin Tower attacks of 9/11 than are top-down leadership strategies.

Far from Equilibrium

The accrued tension described by Haken generates what Prigigine [71] calls, a far from equilibrium (FFE) state. At FFE; the system will often act spontaneously to release (or dissipate) built up tension by moving to a new structural state [64].

In organizational studies, the far from equilibrium notion is unique, for organizations have traditionally been described in terms of equilibrium structures ([12]; see Maruyama [60], for a notable exception). The idea that far from equilibrium states exist suggests an equilibrium continuum for organizations that stretches from an equilibrium state to far from equilibrium conditions. Hazy and Marion [34] argue, however, that while this may be true for some physical systems, it is not true of human systems. Human systems always function in a state of change, albeit with varying degrees of complexity, and it makes no sense to juxtapose such systems against an equilibrium end point. They conclude, then, that far from equilibrium in a social system occurs when the dynamical conditions and resulting complexity of a social system are increased beyond a certain point, not when it moves away from equilibrium.

Even so, the assumption that organization exists in a state of equilibrium is common in the traditional literature on organizations. Although it has been clarified by earlier organizational theorists to be a moving equilibrium (called homeostasis) that adjusts to environmental conditions (the influential classics of this argument are [12]; and [84]), the basic premises in organizational studies have implicitly or explicitly equated organization with adapted variations of the equilibrium view from thermodynamics. Complexity theory contradicts this notion, arguing instead that adaptive systems can only survive and thrive in a state variously called far from equilibrium [64], complexity, edge of chaos [42], or Type IV order [47]. If Hazy and Marion [34] are correct, perhaps the better terms for this are those that do not imply an equilibrium state at all.

Complexity Mechanisms

Complexity is unique in organizational studies because it is oriented around social processes rather than around variables. A variable, or measure of central tendency, is static and it reveals little about the evolutionary social processes in which the variable is embedded. Complexity seeks to understand and to describe the behaviors of those social processes, or what Davis and Marquis [27] and Hedström and Swedberg [36] and others have labeled, social mechanisms.

Mechanisms are defined as commonly observed patterns of behavior [29]; they describe "a set of interacting parts – an assembly of elements producing an effect not inherent in any one of them" p. 336 in [27]; see also p. 74 in [37]. Mechanisms are processes, or dynamic interactions among agents, ideas, and structures. Elster's [29] definition proposes that certain mechanisms can be "commonly" observed across multiple phenomena (that is, they are not necessarily unique to given conditions). They illustrate by arguing that the logistic curve (emergence begins slowly, picks up speed, then levels off) is a mechanism and that it can be observed in numerous situations (Mozart's lifetime productivity, emergence of product popularity, etc.)

Uhl-Bien and Marion [81] identify two broad, interdependent categories of universal mechanisms that they call, complexity mechanisms: These are initiating mechanisms, or interactive processes that generate perturbations, or disturbances, in a system, and aggregation mechanisms, which include accumulation mechanisms (accruing of ideas or structures) and phase transitions (described earlier as nonlinear changes that generate new order). Initiating mechanisms are more micro level while aggregation mechanisms are more macro level.

Practitioners who grasp this notion of complexity mechanisms do not perceive organization in terms of average motivation or efficiency, but rather perceive organization as sets of interacting, evolving processes. They do not ask, for example, whether satisfaction will increase if leaders engage in certain specific acts, or whether high leadership scores on a transformational scale will enhance productivity. Rather they want to know how interactive, interdependent mechanisms influence the capacity of social networks within organizations to generate unpredictable but creative outcomes. These contrasting sets of illustrative questions differ in that the former are based on scientific assumptions about linear relationships among central tendencies and the latter are based on postmodern assumptions regarding complex interactions that are indeterminate in the detail but determinate in the general

nature of their outcome. For example, we can predict dynamic creativity capacity but we cannot predict what that creativity will look like.

Understanding Organization with Complexity

We turn now to an examination of some specific applications of complexity to organizational theory and behavior. This discussion is framed around the relationship of complexity to bureaucracy and around strategic leadership.

Bureaucracy and Complexity

The discussion to this point has implicitly portrayed organization as an adaptive dynamic, or as a system driven exclusively by agents acting informally within complex networks to produce emergent outcomes. In reality, organizations are bureaucratic structures, and despite predictions by some of a post-bureaucratic economy [31,35,54], it is unrealistic to ignore this pervasive element of organization (as Weber [85], said, bureaucracy will be with us until the last shovel of coal is dug from the earth).

For this reason, complexity theorists Uhl-Bien et al. [82] have chosen to depict informal, complex dynamics as embedded within each of Jaques' (1989) three levels of bureaucracy: production, organizational, and strategic. They observe that, while bureaucratic functions differ at each level, the nature of complex dynamics differ only in the tasks they address. That is, complex behavior is similar across the system, whether observed in the board room or in the production divisions. Dynamic, complex behaviors are pervasive, and are universally driven by the enabling conditions described in the last section. Specific complex processes are entangled with one another and even across bureaucratic level. Further, the inter-influences across these dynamics is more horizontal than hierarchical: Dynamics on the shop floor – Jaques' [40] production level – can directly and quickly influence dynamics in the executive suite – Jaques' strategic level. Using complexity terminology, complex dynamics in organizations are, therefore, fractal in nature. From an organizational perspective, they represent a meso view of organization; that is, they describe the linkages that unite micro, or individual, level structures with macro, or organizational level, analysis [26]. Uhl-Bien and Marion [81] also argue that they are meso in that they link different hierarchical levels.

From this, Uhl-Bien, Marion, and McKelvey [82] propose three key dynamics within organizations (which they describe relative to leadership): administrative, adaptive, and enabling. *Administrative leadership* is related to traditional perspectives of leaders as those who generate organizational strategy, gather and distribute resources, and create policy. They differ from traditional perspectives in that their actions must serve to nurture the somewhat fragile, bottom-up (or adaptive) complex leadership processes (fragile in that they are somewhat sensitive to political power behaviors).

Adaptive leadership is defined as any interactive behavior that creates a perturbation which leads to a phase transition. The key elements are interaction, perturbation, and phase transition. Uhl-Bien et al. argue that change originates in the "spaces between" interacting agents (see also [52]); that is, it is the product of the push and shove that can occur between or among agents with heterogeneous skills or worldviews. Perturbations are resulting disturbances in a network (change, new ideas or perspectives, inventions, etc.) that reshape the interaction patterns in that network [81]. Phase transitions are nonlinear shifts within a network (these are described more fully below).

Enabling leadership, the third of Uhl-Bien et al. [82] types, refers to individuals who foster the enabling conditions described above (interaction, interdependency, heterogeneity, enabling rules, etc.) and who mediate the relationship between adaptive and administrative functions.

Emergence, Niche, and Strategic Leadership

The earlier discussion of emergence and phase transitions help us re-visualize the role of another strand in organization sciences: strategic leadership. Strategic leadership theorists study upper echelon behaviors of leader/agents responsible for positioning an organization in its environment. Traditional literature has perceived strategic leadership as a top-down, goal oriented process [13] in which the leader manages the organization's environment. For example, Sir Howard Stringer articulated the strategic goal of Sony Corporation at the 2006 International Consumer Electronics Show in Las Vegas, as:

> Sharpening our focus, shattering internal silos, streamlining our product offerings and growing ever more consumer-centric. We have analyzed our product lines and reorganized our corporate structure to become more nimble and better able to deliver champion products and a focused product line-up. [78].

The website for Sharp Corporation articulates both its philosophy and its business strategy in terms of expanding its market (sharp-world.com/corporate/info/philosophy/index.html). Strategic goals from other companies could be cited, but most will, like these, focus on the company's competitive position in its environment.

Given such assumptions, traditional strategic leadership literature has examined such things as the personality and decisions of top echelon leaders, relationships between the CEO and his or her board of directors, transformational leadership, and visionary leadership (see [13], for a summary).

More recent work by strategic leadership scholars is still concerned about position in the environment, but has focused not on upper echelon behavior but on the internal dynamics of organizations and how organizational flexibility contributes to organizational strategy [8]. Interestingly, Sharp has heeded this call: Stringer's articulation of their goal (above) includes a call for a "more nimble" firm. However, a review of the complexity literature has revealed little that has addressed this issue, despite the fact that it seems a natural subject for complexity analysis. One exception is an article by Marion and Uhl-Bien [58], which argues that flexible, complex structuring is key to organizational strategy. They propose that firms work to increase interaction, interdependency, and adaptive tension.

Marion and Uhl-Bien [58] also suggest that that traditional external-oriented strategic leadership has missed an important point. Much of this literature is, at least implicitly, about dominating the environment or beating the competition. Complexity theory argues that different organizations coevolve [43], and in doing so, they grow to depend upon one another. That is, coevolving systems aggregate and form a niche, or interdependent infrastructure upon which each part of the niche depends to some extent. Automobile makers depend upon fuel outlets and automotive repair shops, for instance – and vice versa. Automotive makers even depend on one another, for their competitive relationship generate pressures that foster structural and ideational elaboration and generate a heterogeneous set of creative solutions that foster even greater creativity. External strategy, then, may be about more than competition; it may also be about fostering effective infrastructure: Marion and Uhl-Bien called this, survival of the cooperative.

Complexity Outcomes: Creativity, Adaptability, Learning

From all this, we can generalize the outcomes of complex dynamics for organizations as creativity, adaptability, and learning. *Creativity* is defined as an ideational phase transition in which an outcome reasonably could not have been predicted based on preexisting conditions or paradigmatic assumptions. Described more basely, it is something that "no one saw coming." Since it is a phase transition, it can be a minor bit of creativity or a major insight.

The major creative insights, of course, capture our imagination and dramatically change our way of seeing things (e. g., Einstein's relativity or Schrödinger's cat). However, small creative events are more common and, arguably, just as important. Creative ideas exist within complex networks that enable them to coevolve and aggregate. That is, they become part of the emergence dynamic within an organization. Creative ideas and associated structures change, aggregate with other creative ideas or structures, and generate adaptive pressures that maintain or increase the complexity of a system. Without the pressures that creative ideas generate, an organization would move towards increased isomorphism and decreased complexity – they move to a state that some would mistakenly call, homeostasis.

Adaptability is enhanced when an organization increases its level of complexity. This translates into (among other things) increased heterogeneity of skills and preferences. Traditional organizational theory has assumed that everyone should be on the "same page" (e. g., "Leaders … are responsible for the dissemination of strategic organizational goals, as well as for convincing their constituents to effectively implement those goals" p. 626 in [6]). Complexity theorists argue that, although organizational agent should be united by a vision, dynamic heterogeneity enables organizations and agents to rapidly and effectively deal with environmental exigencies. Ashby [5] refers to this as "requisite variety;" he defines it by stating that it takes variety to defeat variety. McKelvey and Boisot [65] call it "requisite complexity," thus underscoring the dynamic, co-evolving nature of complex systems and their adaptive response capabilities.

Learning is defined for organizations in several ways. Many of us would define it as something that an individual does. Some organizational theorists [3,22,66], including some complexity scholars (e. g., [9]), argue that learning is best understood as an organizational process. The functioning of the brain exemplifies this logic (see [44]). Individual neurons "learn" only small pieces of an environmental event; it is only when those pieces are put together within a network of pieces that we began to make sense of that event. This, crudely, is what is meant by organizational learning. Organizational knowledge can be defined as a network of different but synchronized bits of individual knowledge. Organizational learning is related to the complex nature of those networks – that is, to its dynamic, changing capabilities. If, then, a system increases its complexity, it will increase its capacity to learn.

Still others propose that organizational learning is a function of the capacity of the network to disseminate information [15]. Like the complexity perspective of organizational knowledge in the previous paragraph, this perspective is intimately linked to the degree of complexity in the organizational network.

Methodology for Studying Complex Organizations

Before developing the methodological strategies that are useful for studying complex organizational system, it will be helpful to summarize a few of the assumptions that underlie complexity, all of which have been stated or implied in this manuscript. These assumptions guide our selection of methodological strategies in organizational studies.

1. Complexity science examines the dynamics of interactions among large numbers of adaptive agents [24].
2. Complex systems evolve dynamically and change nonlinearly (unpredictably; [21]).
3. Human systems do not exist in a state of equilibrium.
4. Complex systems are best described in terms of interactions among complexity mechanisms rather than as interaction among variables.
5. Complex dynamics are enabled rather than planned or controlled.

Traditional Methodologies in the Social Science

Perhaps the most common methodology used to study organizations involves field research strategies and regression statistical procedures. Field strategies permit researchers to draw conclusions in non-laboratory conditions. Regression techniques are based on assumptions of regular, predictable relationships among variables; if one finds such a relationship between variable A and variable B, the relationship was true yesterday and it will be true tomorrow [61]. Such relationships are predictable in that they assume that if you increase variable A, then a resulting change in variable B can be calculated.

While statistical methodology is certainly useful, it cannot deal with many of the conditions posed by complexity science. It is not suited, for example, for exploring dynamic aggregation in organizations, nor is it helpful for explaining nonlinear changes that occur at far from equilibrium states. In particular, its usefulness lapses when one seeks to understand interactions among, and outputs of, complexity mechanisms. Complexity mechanisms are a function of dynamic interactions among numerous agents, ideas, and even variables. Traditional statistical methodology is just not suited for this type of analysis. Statistical assumptions are better suited to equilibrium systems in which desired outcomes involve planning and control. Complexity science is not particularly interested in such outcomes.

A growing trend in organizational studies involves the use of qualitative strategies. Qualitative research requires careful observation of a social dynamic followed by thorough analysis of what was observed. Qualitative methodology is useful for observing dynamic activities and social mechanisms, and is rather widely used for that purpose [20,49,68,70]. Its drawbacks are that it is time and resource intensive and it does not allow one to freely experiment with, or freely change, a social mechanisms – one cannot readily play what – if games in qualitative analysis.

Methodologies for Complexity

The most widely used methodologies in complexity study in organizations are agent-based modeling (ABM), systems dynamics, and qualitative research. ABM replicates or mimics interactive dynamics among organizational agents. ABM is a simulation procedure in which individually programmed agents interact in a computer according to certain rules. Commonly used examples are NetLogo by Uri Wilinsky, Dynamic Network Analysis by Kathleen Carley, and Repast Agent Simulation Toolkit by Michael J. North. These simulations allow researchers to simulate interactions under a variety of conditions and to test the effects of various changes in those conditions. ABM has been used to explore such things as the dynamics that underlie personnel turnover [14], organizational design and performance under stress [19], and leadership as enabler of complex organizational functioning [73].

System dynamic analysis is similar to ABM in that it examines evolving relationships among agents. It differs in that those relationships are not governed by social rules as they are with ABM's, rather they're governed by mathematical defined relationships. Like ABM, it permits the researcher to experiment with various conditions. Sterman [77], for example, has used it to examine interactions among individual decisions and the effects of feedback on the decision process.

Both ABM and system dynamic computer programs permit researchers to base their simulations on data collected in real-world situations. Once that data is entered, however, the quality of the subsequent simulation is only as good as the quality of the rules or the equations that govern it. However the intent is not to predict specific outcomes; rather, the intent is to understand how dynamic relationships can affect outcomes. Carley's Dynamic Network Analysis, for example, permits a user to evaluate the effects of an evolving network on the capacity of the or-

ganization to diffuse knowledge. One can play with such things as the degree of interdependency among agents or the degree of centrality around vision (as Schreiber et al. did in their [74] study) to see how they influence the learning dynamic.

Qualitative research performs generally the same things that ABM and system dynamic analysis do, but it allows the researcher to continuously examine real life events rather than merely simulating those dynamics. The benefit is a closer description of reality; the drawbacks are that it is highly resource and time intensive (especially if one wants to obtain results as in-depth as are results available through simulations), and it does not generally permit much control over conditions that might influence outcomes. Interesting qualitive studies of complex dynamics have been conducted by such organizational theorists as Plowman, Baker et al. [68], who examined emergent conditions in an inner-city church, and Chiles, Meyer, and Hench [20], who studied the complex evolution of Branson, Missouri.

Conclusions

I believe that the best definition of complexity behavior is provided by Paul Cilliers [21]. To frame that definition, he argues that we often confuse complicated with complexity. A complicated system is one with many components interacting to perform complicated tasks, but all of those components maintain their original integrity. A jet plane, he said, is complicated: the components of the plane are unchanged by their relationship. Complex systems are composed of many interacting units, all of which are changed by that interaction such that they quickly loose their original form and function. Organizational theorists who study complex dynamics are seeking to observe and explain that dynamic [48,63]; to understand how organizations might be restructured to improve their capacity for complex outcomes [11]; to improve general leadership in organizations [51,57,80]; to restructure our understanding and implementation of strategic leadership [58], and to restructure how we think about reality, understanding, predictability, and scientific methodology [16,17,18,53,72].

Complexity is genuinely an interdisciplinary science, applicable equally to biology, physics, and the social sciences. Organizational sciences have experienced a number of cumbersome imports in the past: systems theory borrowed from biology [12], as did population ecology theory [33], and scientific management theory borrowed from mechanics [79]. Complexity theory, however, is not about biology or mechanics, it is about interaction dynamics – and this is common across multiple phenomena. It is also about change, not gradual change but change that surprises. It helps us understand how systems learn and adapt to their environment. It comes at a good time for organizational studies, and will help us adapt organizations to the dramatic, knowledge-oriented changes being experienced in world economies.

Future Directions

One is tempted to argue that organizational complexity theory represents a paradigm shift in the way we think about reality (in the Kuhnian sense; [46]), and there is considerable reason to do so. Complexity theorists, for example, acknowledges that the problems facing organizations are far too complex to be solved by rational thinking and planning, and that organizations must engage the interactive, complex capacity of all its employees to solve those problems [59]. Because of this, organizational complexity theory is probably one of the first truly postmodern theories of organization. Organizational complexity theory is premised on complexity mechanisms rather than on variables, thus the questions that emerge from this perspective are dramatically different from any that have been asked in the past (e. g., "how does level of interdependency interact with level of vision to influence temporal emergence of, and phase transitions in, organizational learning"). Complexity theory clearly separates itself from the notions of equilibrium that have so influenced 20th century organizational theory. It moves us away from the centrality of administration and focuses on the centrality of bottom-up, adaptive leadership. Its notions of leadership do not revolve around interpersonal relationships, as does most of leadership theory in the 20th century; rather, it revolves around interactive dynamics. Complexity theory proposes that organizational strategy should seek effective relationships, rather than competition, with other members of its niche. Simply put, complexity theory is a dramatically different way to understand organization.

Complexity theory has captured the attention of academics and practitioners. Important indications of this include the large number of articles on the subject that are appearing in scholarly journals (e. g., the August, 2007, edition of *The Leadership Quarterly* was a special issue on complexity leadership) and the fact that complexity-related articles are being published in top-tier journals (e. g., [11,68]); the increasing number of complexity-related grants in organizations (NSF began a complexity category for organizational scholars in 2007); and a significant surge in the number of practitioners who are seeking consultant relationships with complexity theorists.

A number of interesting research questions are emerging from the complexity paradigm. Andriani and McKelvey's work on extreme events could have significant implications for the way we approach Katrina-like tragedies in the future, for example. The work to develop methodologies such as agent based modeling can have tremendous impact on how we analyze organizational dynamics. The idea that organizational behavior is about the actions of mechanisms is beginning to sink in among scholars, and could dramatically realign the way we think about organization. Complexity leadership theory poses unique questions for the role of strategic leadership and challenges the very way we perceive leadership.

Arguably, there has not been a true paradigm shift in organization studies since Weber introduced his concept of bureaucracy [59]. At the risk of being overly optimistic, organizational complexity theory may be changing this.

Bibliography

1. Anderson P (1995) Microcomputer manufacturers. In: Carroll GR, Hannan MT (eds) Organizations in Industry. Oxford University Press, New York, pp 37–58
2. Andriani P, McKelvey B (2006) Extremes and scale-free dynamics in organization science: Some theory, research, statistics, and power law implications. Manuscript submitted for publication
3. Argyris C, Schön DA (1978) Organizational learning: A theory of action perspective. Addison-Wesley, Reading
4. Arthur WB (1989) The economy and complexity. In: Stein DL (ed) Lectures in the sciences of complexity, vol: 1. Addison-Wesley, Redwood City, pp 713–740
5. Ashby WR (1960) Design for a Brain, 2nd edn. Wiley, New York
6. Berson Y, Avolio BJ (2004) Transformational leadership and the dissemination of organizational goals: A case study of a telecommunication firm. Leadersh Q 15(5):625–646
7. Blau JR (1995) Art museums. In: Carroll GR, Hannan MT (eds) Organization in Industry. Oxford University Press, New York, pp 87–114
8. Boal K (2004) Strategic Leadership, Organizational Learning and Network Ties. Paper presented at the Conference on Strategic Leadership on Both Sides of the Atlantic, International Institute for Management Development, Lausanne, 2004
9. Boisot M (1998) Knowledge assets: Securing competitive advantage in the information economy. Oxford University Press, Oxford
10. Bonabeau E, Meyer C (2001) Swarm intelligence: A whole new way to think about business. Harvard Bus Rev 79(5):107–114
11. Brown SL, Eisenhardt KM (1997) The art of continuous change: Linking complexity theory and time-paced evolution in relentlessly shifting organizations. Adm Sci Q 42(1):1–34
12. Buckley W (1967) Sociology and modern systems theory. Prentice Hall, Englewood Cliffs
13. Canella AA, Monroe MJ (1997) Contrasting perspectives on strategic leaders: Toward a more realistic view of top managers. J Manag 23(3):213–230
14. Carley K (1992) Organizational learning and personnel turnover. Organ Sci 3:20–46
15. Carley K (1996) Organizational learning and personnel turnover. In: Cohen MD, Sproull LS (eds) Organizational Learning. Sage, Thousand Oaks
16. Carley K (1997) Organizational adaptation. An Oper Res 75: 25–47
17. Carley K (1999) On the evolution of social and organizational networks. In: Bacharach S, Knoke D, Andrews SB (eds) Research in the Sociology of Organizations: Networks In and Around Organizations, vol: 16. Press JAI, Greenwich, pp 3–30
18. Carley K (2002) Intra-organizational complexity and computation. In: Baum JAC (ed) The Blackwell Companion to Organizations. Blackwell, Oxford, pp 208–232
19. Carley K, Lin Z (1995) Organizational designs suited to high performance under stress. Trans IEEE Syst Man Cybern 25:221–230
20. Chiles T, Meyer A, Hench T (2004) Organizational emergence: The origin and transformation of Branson, Missouri's musical theaters. Organ Sci 15(5):499–519
21. Cilliers P (1998) Complexity and postmodernism: Understanding complex systems. Routledge, London
22. Cohen MD, Sproull LS (1996) Organizational Learning. Sage, Thousand Oaks
23. Conger JA (1999) Charismatic and transformational leadership in organizations: An insider's perspective on these developing streams of research. Leadersh Q 10(2):145–179
24. Coveney P (2003) Self-organization and complexity: A new age for theory, computation and experiment. Paper presented at the Nobel symposium on self-organization, Karolinska Institutet, Stockholm, 2003
25. Cusumano M (2001) Focusing creativity: Microsoft's "Synch and Stabilize" approach to software product development. In: Nonaka I, Nishiguchi T (eds) Knowledge emergence: Social, technical, and evolutionary dimensions of knowledge creation. Oxford University Press, Oxford, pp 111–123
26. Dansereau F, Yammarino FJ (eds) (1998) Leadership: Multiple-level approaches: Contemporary and alternative, vol: 2. Press JAI, Stamford
27. Davis GF, Marquis C (2005) Prospects for organization theory in the early twenty-first century: institutional fields and mechanisms. Organ Sci 16(4):332–343
28. DiMaggio PJ, Powell WW (1991) The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. In: Powell WW, DiMaggio PJ (eds) The new institutionalism in organizational analysis. University of Chicago Press, Chicago, pp 63–82
29. Elster J (1998) A plea for mechanisms. In: Hedström P, Swedberg R (eds) Social Mechanisms. Cambridge University Press, New York, pp 45–73
30. Goodwin B (1994) How the leopard changed its spots: The evolution of complexity. Charles Scribner's Sons, New York
31. Grey C, Garsten C (2001) Trust, Control and the Post-Bureaucracy. Organization Studies, vol: 22. Walter de Gruyter, New York, pp 229
32. Haken H (1983) Synergetics, an introduction, vol: 3. Springer-Verlag, Berlin
33. Hannan MT, Freeman J (1977) The population ecology of organizations. Am J Sociol 82:929–964
34. Hazy J, Marion R (2008) When Organizing in Extreme Situations, Is Far-From-Equilibrium (FFE) a Meaningful Concept?

Poster presented at the Organization Science Winter Conference, Squaw Creek, 2008
35. Heckscher C, Donnellon A (1994) Defining the post-bureaucratic type. New perspectives on organizational change Thousand Oaks: Sage, In: Heckscher C, Donnellon A (eds) The post-bureaucratic organization
36. Hedström P, Swedberg R (1998) Social Mechanisms: An analytical approach to social theory. Cambridge University Press, Cambridge
37. Hernes G (1998) Real virtuality. In. Hedström P, Swedberg R (eds) Social mechanisms: An analytical approach to social theory. Cambridge University Press, Cambridge, pp 74–101
38. Hitt MA (1998) Presidential Address: Twenty-first century organizations: Business firms, business schools, and the academy. Acad Manag Rev 23:218–224
39. Homans GC (1950) The human group. Harcourt, New York
40. Jaques E (1989) Requisite organization. Cason Hall, Arlington
41. Jermier JM (1998) Introduction: Critical perspectives on organizational control. Adm Sci Q 43(2):235–256
42. Kauffman SA (1993) The origins of order. Oxford University Press, New York
43. Kauffman SA (1995) At home in the universe: The search for the laws of self-organization and complexity. Oxford University Press, New York
44. Kelso JA (1995) Dynamic patterns: The self-organization of brain and behavior. Press MIT, Cambridge
45. Koschmieder EL, Davis S, Tvergaard V, Batchelor GK, Leibovich S (eds) (1993) Bernard cells and Taylor vortices–Cambridge monographs on mechanics and applied mathematics. Cambridge University Press, New York
46. Kuhn TS (1970) The structure of scientific revolutions, 2nd edn. The University of Chicago Press, Chicago
47. Langston CG (1986) Studying artificial life with cellular automata. Phys 22D:120–149
48. Lichtenstein B (2000) Dynamics of rapid growth and change: A complexity theory of entrepreneurial transitions. In: Liebcap G (ed) Advances in the study of entrepreneurship, innovation, and economic growth, vol: 6. Press JAI, Westport CT, pp 161–192
49. Lichtenstein B, Carter N, Dooley K, Gartner W (2007) Complexity dynamics of nascent entrepreneurship. J Bus Ventur 22:236–261
50. Lichtenstein B, Plowman D (2007a) Emergent leadership: A meso-model of leadership inspired by complexity. Manuscript submitted for publication
51. Lichtenstein B, Plowman D (2007b) The leadership of emergence: Meso-level leadership. Paper presented at the Festschrift for Jerry Hunt, Texas Tech, 2007
52. Lichtenstein B, Uhl-Bien M, Marion R, Seers A, Orton D, Schreiber C (2006) Leadership in Emergent Events: Exploring the Interactive Process of Leading in Complex Situations. Emerg Complex Organ 8(4):2–12
53. Lin F, Pai Y (2000) Using Multi-Agent Simulation and Learning to Design New Business Processes. Trans IEEE Syst Man Cybern Part A Syst Hum 30(3):380–384
54. Maravelias C (2003) Post-bureaucracy–control through professional freedom. J Organ Chang Manag 16(5):547
55. Marion R (1999) The edge of organization: Chaos and complexity theories of formal social organizations. Sage, Newbury Park
56. Marion R (2006) Complexity in organizations: A paradigm shift. In: Sengupta A (ed) Chaos, Nonlinearity, Complexity: The Dynamical Paradigm of Nature, vol: 206. Springer-Verlag, Berlin, pp 248–270
57. Marion R, Uhl-Bien M (2001) Leadership in complex organizations. Leadersh Q 12:389–418
58. Marion R, Uhl-Bien M (2007a) Complexity and strategic leadership. In: Hooijberg R, Hunt J, Antonakis J, Boal K, Lane N (eds) Being There Even When You Are Not: Leading Through Structures, Systems, and Processes, vol: 4. Elsevier, Amsterdam, pp 273–287
59. Marion R, Uhl-Bien M (2007b) Paradigmatic Influence and Leadership: The Perspectives of Complexity Theory and Bureaucracy Theory. In: Hazy JK, Goldstein J, Lichtenstein B (eds) Complex Systems Leadership Theory. Publishing ISCE, New York, pp 143–159
60. Maruyama M (1963) The second cybernetics: Deviation amplifying mutual causal processes. Am Sci 51:164–179
61. Maxwell JA (2004) Causal explanation, qualitative research, and scientific inquiry in education. Educ Res 33(2):3–11
62. McKelvey B (1999) Complexity theory in organization science: Seizing the promise or becoming a fad? Emerg 1(1):5–32
63. McKelvey B (2001) Energizing order-creating networks of distributed intelligence. Int J Innov Manag 5:181–212
64. McKelvey B (2008) Emergent Strategy via Complexity Leadership: Using Complexity Science and Adaptive Tension to Build Distributed Intelligence. In: Uhl M-Bien, Marion R (eds) Complexity leadership. Conceptual foundations, vol: 1. Information Age Publishing, Charlotte, pp. 225–268
65. McKelvey B, Boisot MH (2003) Transcendental organizational foresight in nonlinear contexts. Paper presented at the Conference INSEAD on Expanding Perspectives on Strategy Processes, Fontainebleau, France, 2003
66. Nonaka I, Takeuchi H (eds) (1995) The knowledge-creating company. Oxford University Press, Oxford
67. Pianka ER (1994) Evolutionary Ecology, 5th edn. HarperCollins, New York
68. Plowman D, Baker LT, Beck T, Kulkarni M, Solansky S, Travis D (2007) Radical change accidentally: The emergence and amplification of small change. Acad Manag J 50(3):515–543
69. Plowman D, Duchon D (2008) Dispelling the myths about leadership: From cybernetics to emergence. In: Uhl M-Bien, Marion R (eds) Complexity leadership. Conceptual foundations, vol: 1. Information Age Publishing, Charlotte, pp. 129–154
70. Plowman D, Solansky S, Beck T, Baker L, Kulkarni M, Travis D (2007) The role of leadership in emergent, self-organization. Leadersh Q 18:341–356
71. Prigogine I (1997) The end of certainty. The Free Press, New York
72. Schreiber C, Marion R, Uhl-Bien M, Carley K (2006) Multi-Agent Based Simulation of a Model of Complexity Leadership. Paper presented at the International Conference on Complex Systems, Boston, 2006
73. Schreiber C, Carley K (2006) Leadership style as an enabler of organizational complex functioning. Emergence Complex Organ 8(4):61–76
74. Schreiber C, Marion R, Uhl-Bien M, Carley K (2007) Multi-Agent Based Modeling of Complexity Leadership Theory. Paper presented at the North American Association for Computational Social and Organizational Sciences, Atlanta, 2007
75. Schroeder M (1991) Fractals, chaos, power laws. Freeman, New York

76. Stephenson WD, Bonabeau E (2007) Expecting the Unexpected: The Need for a Networked Terrorism and Disaster Response Strategy (Electronic Version). Homeland Security Affairs, 3 http://www.hsaj.org/?article=3.1.3
77. Sterman JD (1989) Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. Manag Sci 35(3):321–339
78. Stringer SH (2006) Entertaining the Future. International Consumer Electronics Show Las Vegas January – 5:2006 Retrieved November, 2007, http://www.sony.com/SCA/speeches/060105_stringer.shtml
79. Taylor F (1911) The principles of scientific management. Harper, New York
80. Uhl-Bien M, Marion R (2005) Complexity leadership theory. Paper presented at the Workshop on Leadership and Complexity, George Washington University
81. Uhl-Bien M, Marion R (2007) The mechanisms of emergence in complexity leadership theory: A meso-model of adaptive dynamics in organizations. Paper presented at the Festschrift honoring Jerry Hunt, Lubbock, 2007
82. Uhl-Bien M, Marion R, McKelvey B (2007) Complexity Leadership Theory: Shifting Leadership from the Industrial Age to the Knowledge Era. Leadersh Q 18:298–318
83. Volberda HW (1996) Toward the flexible form: How to remain vital in hypercompetitive environments. Organ Sci 7(4):359–374
84. von Bertalanffy L (1956) General systems theory. In: L v. Bertalanffy, Rapoport A (eds) General systems: Yearbook of the society for the advancement of general systems theory, vol: 1. George Braziller Publishers, New York, pp 1–10
85. Weber M (1947) The theory of social and economic organization (trans: Henderson AH, Parsons T). Free Press, Glencoe

Social Phenomena Simulation

Paul Davidsson[1], Harko Verhagen[2]
[1] Blekinge Institute of Technology, Ronneby, Sweden
[2] Stockholm University and Royal Institute of Technology, Stockholm, Sweden

Article Outline

Glossary
Definition of the Subject
Introduction
Why Simulate Social Phenomena?
Simulating Social Phenomena
Future Directions
Further Reading
Bibliography

Glossary

Agent (or software agent) A self-contained entity that has a state and that is situated (able to perceive and act) in an environment. In addition, agents are often assumed to be rational and autonomous.

Cellular automaton A mathematical structure modeling a set of cells that interact with their neighbors. Each cell has a set of neighbors and a state. All the cells update their values simultaneously at discrete time steps. The new state of a cell is determined by the current state of its neighbors according to a local function or rule.

Microlevel simulation A type of simulation in which the specific behaviors of specific individuals are explicitly modeled.

Definition of the Subject

Social phenomena simulation in the area of agent-based modeling and simulation concerns the emulation of the individual behavior of a group of social entities, typically including their cognition, actions, and interaction. Agent-based social simulation constitutes the intersection of three scientific fields, namely, agent-based computing, the social sciences, and computer simulation [6]. Agent-based computing is a research area mainly within computer science and includes, e. g., agent-based modeling, design, and programming. By the social sciences we here refer to a large set of different sciences that study the interaction among social entities, e. g., social psychology, management science, policy, and some areas of biology. Computer simulation concerns the study of different techniques for simulating phenomena on a computer, e. g., discrete-event, object-oriented, and equation-based simulation.

Introduction

Computer simulation consists of three main steps: (i) designing a model of an actual or theoretical system, (ii) executing the model on a computer, and (iii) analyzing the execution output. Already in the early days of computer development, simulation was used in different research areas to predict the behavior of complex systems. Such simulations were typically based on differential equations and focused on results at the aggregate level. These models of, for instance, predator-prey populations could result in fairly accurate models but were limited in the sense that the models excluded individual behavior and decision making, as well as interaction between individuals, and were based on homogeneous agents. The development of agent-based modeling offers a possible solution to this problem with its (seemingly) natural mapping onto interacting individuals with incomplete information and capabilities, no global control, decentralized data, asynchronous computing, and inclusion of hetero-

geneous agents. Agent-based simulation models also offer the possibility of studying the dynamics of the interaction processes instead of focusing on the (static) results of these processes [16,26].

Agent-based modeling can be traced back to von Neumann, who in the 1950s invented what was later termed *cellular automata*. These were used by Conway in the 1970s when he constructed the well-known *Game of Life*. It is based on very simple rules determining the life and death of the cells in a virtual world in the form of a 2-D grid. Inspired by this work, researchers developed more-refined models, often modeling the social behavior of groups of animals or artificial creatures. One example is the Boid model by Reynolds [24], which simulates coordinated animal motion such as bird flocks and fish schools. With respect to human societies, Epstein and Axtell [8] developed in the 1990s one of the first agent-based models, called Sugarscape, to explore the role of social phenomena such as seasonal migrations, pollution, sexual reproduction, combat, and transmission of disease. This work is in spirit closely related to one of the best-known and earliest examples of the use of simulation in social science, namely, the Schelling model [27], in which cellular automata were used to simulate the emergence of segregation patterns in neighborhoods based on a few simple rules expressing the preferences of the agents. Another pioneer from the 1950s worth mentioning is Barricelli [2], who to some extent used agent-based modeling for simulating biological systems.

The cellular automata models closely resemble the models used in statistical physics, which has inspired physicists to include the simulation of social phenomena in large-scale social systems in their research agenda. In this area, sometimes referred to as sociophysics, phenomena such as opinion spreading in a society and competition between languages have been studied. These models originally described the behavior of atoms and molecules, which are quite simple objects, and the macrolevel phenomena caused by their interaction (rather than by complex behavior of the individual as in the case of humans). Thus, in these models little attention is paid to individual variation and the individual decision making is rather primitively modeled. A prominent example of sociophysics is the work of Galam [10].

To sum up, we can identify two main approaches to social simulation:

- Macrolevel (or equation-based) simulation, which is typically based on mathematical models. It views the set of individuals (the population) as a structure that can be characterized by a number of variables.
- Microlevel (or agent-based) simulation, in which the specific behaviors of specific individuals are explicitly modeled. In contrast to macrolevel simulation, it views the structure as emerging from the interactions between individuals and thus exploring the standpoint that complex effects need not have complex causes.

As argued by Van Parunak et al. [21], agent-based modeling is most appropriate for domains characterized by a high degree of localization and distribution and dominated by discrete decision. Equation-based modeling, on the other hand, is most naturally applied to systems that can be modeled centrally and in which the dynamics are dominated by physical laws rather than information processing. We will here focus on agent-based models, particularly those that have a richer representation of the individual than the cellular automata and statistical physics models.

Why Simulate Social Phenomena?

Simulation of social phenomena can be done for different purposes, e. g.,

- Supporting social-theory building;
- Supporting the engineering of systems, e. g., validation, testing, etc.;
- Supporting planning, policy making, and other decision making;
- Training, in order to improve a person's skills in a certain domain.

It is possible to distinguish between four types of end users: *scientists*, who use social phenomena simulation in the research process to gain new knowledge, *policymakers*, who use it for making strategic decisions, *managers* (of systems), who use it to make operational decisions, and *other professionals*, such as architects, who use it in their daily work. We will now describe how these types of end users may use simulation of social phenomena for different purposes.

Supporting Social-Theory Building

In the context of social-theory building, agent-based simulation can be seen as an experimental method or as theories in themselves [26]. In the former case, simulations are run to test the predictions of theories, whereas in the latter case simulations in themselves are formal models of theories. Formalizing the ambiguous, natural-language-based theories of the social sciences helps to find inconsistencies and other problems, and thus contributes to theory building.

Using agent-based simulation studies as an experimental tool offers great possibilities. Many experiments with human societies are either unethical or even impossible to conduct. Experiments in silico, on the other hand, are fully possible. These can also breathe new life into the ever-present debate in sociology on the micro–macro link [1]. Agent-based models mostly focus on the emergence of macrolevel properties from the local interaction of adaptive agents that influence one another [17,26]. However, simulations in computational organization theory [4,22], for example, often try to analyze the influence of macrolevel phenomena on individuals. Using agent-based models to simulate the bidirectional relation between micro- and macrolevel concepts would provide tools to analyze the theoretical consequences of the work done by theorists such as Habermas, Giddens, and Bourdieu, to name a few [26].

Supporting the Engineering of Systems

Many new technical systems are distributed and involve complex interactions between humans and machines. The properties of agent-based simulation make it especially suitable for simulating these kinds of systems. The idea is to model the behavior of human users in terms of software agents. In particular, this seems useful in situations where it is too expensive, difficult, inconvenient, tiresome, or even impossible for real human users to test out a new technical system. Of course, also the technical system, or parts thereof, may be simulated. For instance, if the technical system includes hardware that is expensive and/or special purpose, it is natural to simulate also this part of the system when testing out the control software. An example of such a case is the testing of control systems for "intelligent buildings," where agents simulate the behavior of the people in the building [5].

Supporting Planning, Policy Making, and Other Decision Making

Here the focus is on exploring different possible future scenarios in order to choose between alternative actions. Besides this type of prediction, simulation of social phenomena may be used for analysis, i. e., to gain deeper knowledge and understanding of a certain phenomenon.

An area in which several studies of this kind have been carried out is disaster management, such as experiments concerning different roles and the efficiency of reactions to emergencies [18]. Based on individuals' observations, personal characteristics and skills, past experience and role characteristics, and social network, the agents create a plan to execute. Each agent represents a human being (acting in a particular role). The effect of adding a role (floor warden) in a fire alarm scenario upon the evacuation efficiency in an abstract environment is analyzed. In another approach, the agents are placed in an environment based on GIS (geographical information system) data, thereby tying the simulation closer to the physical reality [29]. In yet another study, real-world data were used for both the environment and the agents' internal decision-making model to analyze the effect of different insurance policies on the willingness of agents to pay for a disaster insurance policy [3].

Another application area for this type of simulation study is disease spreading. Typically, agents are used to represent human beings and the simulation model is linked to real-world geographical data. One study [32] also included agents that represent towns acting as the epicenter of disease outbreak. The town agent's behavior repertoire consisted of different containment strategies. The simulation model can be quickly adapted to local circumstances via the geographical data (given that there is data on the population as well) and is used to determine the effects of different containment strategies.

A third area where agent-based social simulation has been used to support planning and policy making is traffic and transport. An example of this is the simulation of all car travel in Switzerland during morning peak traffic [23].

Training

The main advantage of using simulation for training purposes is to be part of a real-world-like situation without real-world consequences. Especially in the military the use of simulation for training purposes is widespread. Also in medicine, where mistakes can be very expensive in terms of money and lives, the use of simulation in education is on the rise.

An early product in this area was a tool to help train police officers to manage large public gatherings such as crowds, demonstrations, and marches [31]. Another early example of agent-based simulation for training purposes is Steve [19,25]. Steve was an agent integrated with voice synthesis software and virtual reality software providing a very realistic training environment for controlling the engine room of a virtual US Navy surface ship.

An example of a more recent project is the PSI agent [15]. Whereas in most cases the simulator training is aimed at training practical skills or decision making, this work focuses on acquiring theoretical insights in the realm of psychological theory. The simulation enables students to explore psychological processes without ethical problems.

Simulating Social Phenomena

One of the first, and most simple, way of performing microlevel simulation is often called *dynamic microsimulation* [11,12]. It is used to simulate the effect of the passing of time on individuals. Data from a (preferably large) random sample from the population to be simulated is used to initially characterize the simulated individuals. Some examples of sampled features are: age, sex, employment status, income, and health status. A set of transition probabilities are used to describe how these features will change over a given time period, e. g., there is a probability that an employed person will become unemployed over the course of a year. The transition probabilities are applied to the population for each individual in turn and then repeatedly reapplied for a number of simulated time periods. Sometimes it is necessary to also model changes in the population, e. g., birth, death, and marriage. This type of simulation can be used to, e. g., predict the outcome of different social policies. However, the quality of such simulations depends on the quality of:

- the random sample, which must be representative, and
- the transition probabilities, which must be valid and complete.

In traditional microsimulation, the behavior of each individual is regarded as a "black box." The behavior is modeled in terms of probabilities and no attempt is made to justify these in terms of individual preferences, decisions, plans, etc. Also, each simulated individual is considered in isolation without regard to their interaction with others. Thus, better results may be gained if cognitive processes and communication between individuals are also simulated.

Opening the black box of individual decision making can be done in several ways. The first layer to add is often individual psychology; for instance, the so-called beliefs, desires, and intentions (BDI) model is often used. Models of individual cognition used in agent-based social simulation include the use of Soar (a computer implementation of Allen Newell's unified theory of cognition [20]), which was used in Steve (discussed in Sect. "Why Simulate Social Phenomena?").

For the simulation of social behavior the agents need to be equipped with mechanisms for reasoning at the social level (unless the social level is regarded as emerging from individual behavior and decision making). Several models have been based on theories from economics, social psychology, sociology, etc. An example of this is provided by Guye-Vuillème [13], who has developed an agent-based model for simulating human interaction in a virtual-reality environment. The model is based on sociological concepts such as roles, values, and norms and motivational theories from social psychology to simulate persons with social identities and relationships. Another example is the Consumat model [14], a metamodel combining several psychological theories on decision making in a consumer situation, used, for instance, to investigate different flood-management policies [3]. Also, nonsymbolic approaches such as neural networks have been used to model agents' decision making [18].

Future Directions

In a recent study of applications of agent-based simulation [7], it was concluded that even if agent-based simulation seems a promising approach to many problems involving the simulation of complex systems of interacting entities such as social phenomena, it seems that the full potential of the agent concept often is not utilized. For instance, most models have very primitive agent cognition, in particular if the number of agents involved is large.

Regarding future applications, Fiedrich and Burghardt [9] argue that agent-based simulation is a very promising approach to disaster management practice. In particular, agent-based social simulation in combination with sophisticated visualization techniques, such as virtual reality, in the form of "serious games," has the potential to provide very powerful training environments. In the context of military training, Stone [28] provides some interesting applications.

Further Reading

Classic books in the area of simulation of social behavior include "Growing Artificial Societies: Social Science from the Bottom Up" by Epstein and Axtell [8] and "Simulation for the Social Scientist" by Gilbert and Troitzsch [12].

More recent findings can be found in, e. g., the Journal of Artificial Societies and Social Simulation (http://jasss.soc.surrey.ac.uk) and the proceedings of, e. g., the International Workshop series on Multi-Agent-Based Simulation (MABS) (http://www.pcs.usp.br/~mabs/), the World Congress in Social Simulation (WCSS) [30], the conference of the European Social Simulation Association (ESSA) (http://www.essa.eu.org/), and the series of Agent Workshops in Chicago (http://www.agent2005.anl.gov/).

Bibliography

1. Alexander JC, Giesen B, Münch R, Smelser NJ (eds) (1987) The Micro-Macro Link. University of California Press, Berkeley
2. Barricelli NA (1957) Symbiogenetic evolution processes realized by artificial methods. Methodos 9(35–36):143–182

3. Brouwers L, Verhagen H (2004) Applying the Consumat model to flood management policies. In: Müller J, Seidel M (eds) 4th Workshop on Agent-Based Simulation. SCS, Montpellier, pp 29–34
4. Carley KM, Prietula M (eds) (1994) Computational Organization Theory. Erlbaum, Hillsdale
5. Davidsson P (2000) Multi-agent-based simulation: beyond social simulation. In: Moss S, Davidsson P (eds) Multi-Agent-Based Simulation. Lecture Notes in Computer Science, vol 1979. Springer, Berlin, pp 98–107
6. Davidsson P (2002) Agent-based social simulation: a computer science view. J Artif Soc Soc Simul 5(1)
7. Davidsson P, Holmgren J, Kyhlbäck H, Mengistu D, Persson M (2007) Applications of multi-agent-based simulation. In: Antunes L, Takadama K (eds) Multi-Agent-Based Simulation VII. Lecture Notes in Computer Science, vol 4442. Springer, Berlin
8. Epstein JM, Axtell RL (1996) Growing Artificial Societies: Social Science from the Bottom Up. MIT Press, Cambridge
9. Fiedrich F, Burghardt P (2007) Emergency response information systems: emerging trends and technologies: agent-based systems for disaster management. Commun ACM 50(3):41–42
10. Galam S (2008) Sociophysics: a review of Galam models. arXiv:0803.1800 http://arxiv.org/abs/0803.1800. Accessed 06 Aug 2008
11. Gilbert N (1994) Computer Simulation of Social Processes. Social Research Update, Issue 6, Department of Sociology, University of Surrey, UK. http://www.soc.surrey.ac.uk/sru/. Accessed 06 Aug 2008
12. Gilbert N, Troitzsch KG (2005) Simulation for the Social Scientist, 2nd edn. Open University Press, Maidenhead
13. Guye-Vuillème A (2004) Simulation of nonverbal social interaction and small groups dynamics in virtual environments. Ph D thesis, École Polytechnique Fédérale de Lausanne, No 2933
14. Janssen MA, Jager W (1999) An Integrated Approach to Simulating Behavioural Processes: A Case Study of the Lock-in of Consumption Patterns. J Artif Soc Soc Simul 2(2)
15. Künzel J, Hämmer V (2006) Simulation in university education: the artificial agent PSI as a teaching tool. Simulation 82(11):761–768
16. Lansing JS (2002) "Artificial societies" and the social sciences. Artif Life 8:279–292
17. Macy MW, Willer R (2002) From factors to actors: computational sociology and agent-based modeling. Annu Rev Sociol 28:143–166
18. Massaguer D, Balasubramanian V, Mehrotra S, Venkatasubramanian N (2006) Multi-agent simulation of disaster response. In: Proceedings of the 1st International Workshop on Agent Technology for Disaster Management 2006. Hakodate
19. Méndez G, Rickel J, de Antonio A (2003) Steve meets Jack: the integration of an intelligent tutor and a virtual environment with planning capabilities. In: Intelligent Virtual Agents. Lecture Notes on Artificial Intelligence, vol 2792. Springer, Berlin, pp 325–332
20. Newell A (1994) Unified Theories of Cognition. Harvard University Press, Cambridge
21. Parunak HVD, Savit R, Riolo RL (1998) Agent-based modeling vs. equation-based modeling: a case study and users' guide. In: Sichman JS, Conte R, Gilbert N (eds) Multi-Agent Systems and Agent-Based Simulation. Lecture Notes in Computer Science, vol 1534. Springer, Berlin, pp 10–26
22. Prietula MJ, Carley KM, Gasser L (eds) (1998) Simulating Organizations: Computational Models of Institutions and Groups. MIT Press, Cambridge
23. Raney B, Cetin N, Völlmy A, Vrtic M, Axhausen K, Nagel K (2003) An Agent-Based Microsimulation Model of Swiss Travel: First Results. J Netw Spatial Econ 3(1):23–41
24. Reynolds CW (1987) Flocks, herds, and schools: a distributed behavioral model. Comput Graph 21(4):25–34 (SIGGRAPH '87 Conference Proceedings)
25. Rickel J, Lewis Johnson W (1999) Animated agents for procedural training in virtual reality: perception, cognition, and motor control. Appl Artif Intell 13:343–382
26. Sawyer RK (2003) Artificial societies – multi-agent systems and the micro-macro link in sociological theory. Sociol Meth Res 31(3):325–363
27. Schelling TC (1971) Dynamic models of segregation. J Math Sociol 1(1):143–186
28. Stone B (2005) Serious gaming. Def Manag J (31):142–144
29. Takahashi T (2007) Agent based disaster simulation evaluation and its probability model interpretation. In: Proceedings of the 4th International Conference on Intelligent Human-Computer Systems for Crisis Response and Management, 2007, Deltft, pp 369–376
30. Takahashi S, Sallach D, Rouchier J (2007) Advancing Social Simulation – The First World Congress. Springer, Berlin
31. Williams R (1993) An agent based simulation environment for public order management training. In: Western Simulation Multiconference, Object-Oriented Simulation Conference (OOS '93), San Diego 151–156
32. Yergens D, Hiner J, Denzinger J, Noseworthy T (2006) IDESS – A multi-agent-based simulation system for rapid development of infectious disease models. Int Trans Syst Sci Appl 1(1):51–58

Social Processes, Physical Models of

FRANTIŠEK SLANINA[1,2]
[1] Institute of Physics, Academy of Sciences of the Czech Republic, Prague, Czech Republic
[2] Center for Theoretical Study, Prague, Czech Republic

Article Outline

Glossary
Definition of the Subject
Historical Introduction
Individualism Versus Cooperation
Opinion Dynamics
Bounded Confidence
Emergence of Social Classes
Future Directions
Acknowledgments
Bibliography

Glossary

Absorbing state In a stochastic process, a state from which there is no way out. Once the system reaches the absorbing state, it stays there forever.

Complete graph A graph where every pair of vertices is connected by an edge.

Configuration space The ensemble of all allowed states (configurations) a model system can reach. Any point in configuration space represents one such state.

Graph an assembly of points (vertices), some of which are connected by lines (edges).

Hypercubic lattice A generalization of a rectangular plane mesh; a graph embedded in a space of arbitrary dimensionality d. Assuming rectangular coordinates $x_1, x_2, \ldots, x_d$, the coordinates of the vertices in a hypercubic lattice are all whole numbers.

Markov process A special type of stochastic process. In a Markov process the system changes its state randomly and, most importantly, the changes of state are independent of history. A Markov process is a memoryless system.

Random walk A stochastic process describing hops of a particle to randomly chosen neighbor sites on a lattice.

Stochastic process A sequence of random events. At each time t, which may be either discrete ($t = 1, 2$, etc.) or continuous, a new random variable is introduced representing the outcome of the process in that time. Also called a random process.

Definition of the Subject

Modeling social phenomena as if they were manifestations of mutual interactions of physical objects is the ultimate goal of the reductionist approach to reality. Both the inanimate and animate worlds, including all the behavior of humans, would be traced back to the properties of atoms and molecules. This program is absolutely unrealizable, though. On the other hand, the discipline of **sociophysics** tries to bypass the brute-force approach by developing schematically effective models which aim at describing reality at a "macroscopic", rather than microscopic, level.

For example, when one wants to model the behavior of a large assembly of humans facing the necessity of choosing between two options, it is customary to neglect all details of the behavior of the people involved and describe their states by two-value quantities, such as $s = +1$ or $s = -1$; physicists call them spins. The interactions are often expressed using a cost function, which physicists call energy. The state with lowest energy is favored, but there are also external perturbations, or noise, which prevent the system from settling in that state. Physicists call the measure of the noise the temperature. The system of interacting spins at a certain temperature then serves as a model of the particular situation in human society.

It is a non-trivial question if such approach could work and if so, why. The essential ingredient for its functioning is **universality**. This notion, borrowed from statistical physics, means, that the behavior of the system does not ultimately depend on the details of that system, but only on its generic features. In plain words, for the description of the human mind, the behavior of a driver on a highway, or of a stockbroker, it is irrelevant whether the matter is composed of quarks, strings or whatever elementary particles may be discovered. The description of the world at the "macroscopic" level is detached from "microscopic" details.

It is the task of **sociophysics** to develop suitable macroscopic models that conform to observed facts about society and, at the same time, yield themselves to the methods of theoretical physics. There are non-negligible successes in this effort, yet still many areas remain unexplored. Consequently, the subject is alive and in a state of rapid progress.

Historical Introduction

It has been nearly two hundred years since the notion of social physics was introduced by the French philosopher Auguste Comte [1] in his attempt to revolutionize all thinking under the banner of positive philosophy. He himself abandoned that term later [2], but the idea of connecting social and physical phenomena in a unified frame has been around since. Later on, it somehow degenerated into a purely descriptive science, borrowing from physics nothing more than the tendency to exactitude in measurements; a prominent name to cite in this context is the Belgian statistician Adolphe Quételet. Little use stems from twopenny analogies, advocated by some later followers of Hegelian philosophy, between phase transitions, such as boiling water, and abrupt changes, mostly violent ones, in human society. Nonetheless, you may still find these considerations, based on the "quantity turns into quality" principle, in popular-philosophy brochures.

Skipping ahead a little more than a century and half since Comte, the first modern attempt to bridge the gap between social sciences and physics was marked by a memorable conference scheduled to take place in Moscow, 1 to 5 July 1974. Scientists both from the West and from the USSR were to discuss implications of physics for other fields, including social sciences and humanities. The organizing committee included people such as Kenneth Arrow, Nobel laureate for Economy and Hans

Bethe, Nobel laureate for Physics. Among the speakers was Andrei Sakharov, the well-known dissident and Nobel laureate for Peace. However, the Communist authorities stepped in, the conference was banned, most of the Russian participants were arrested, and a majority of them were exiled. However, many of the manuscripts were smuggled from the USSR to the West, and eventually the contributions were published in a proceedings volume [3]. A small sample of it appeared in [4].

In the 1980s, a very fruitful and general concept of synergetics was developed by H. Haken and his followers. The main idea consists in spontaneous emergence of coherent structures. In addition to being illustrated by examples from several branches of physical sciences (e. g. lasers) the idea was also elaborated in the context of social phenomena [5].

About the same time, the science of complexity started to gain wider popularity (and continues to do so to this day). The Santa Fe Institute in Santa Fe, USA, pioneered this research; it was here that complex system theory began to be systematically applied to economics and, more generally, to all human social behavior. Another Physics Nobel laureate, P. W. Anderson [6], has played a prominent role in promoting this path.

To finish this historical overview, the journey returns to France, where since the early 1980s the group of Serge Galam has been busy developing genuine physical models of social phenomena, also using he term sociophysics consistently [7,8,9]. Let us stop here and turn to several selected topical problems, where physically based models may help answer social scientists' questions.

Individualism Versus Cooperation

People never earn their living alone. From the darkest prehistorical past they have joined together to form gangs, stalking and hunting their prey. Without communication within the band, our ancestors could hardly have caught a mammoth, language would never have evolved and human brain capacity would probably never have exceeded that of a lemur. In short, cooperation between humans has been decisive in shaping the world around us. A tendency to cooperation is an inherent feature of human nature.

It was not until the middle of the eighteenth century or so that a different view started to spread. In this view, people are presented as selfish profit-seekers and if any cooperation is observed, it occurs *despite* the natural tendency to maximize personal gain, or as a by-product of that tendency. Under vague names such as neo-liberalism and social Darwinism, these and related ideas pervade the current thinking on human society.

On the other hand, empirically, cooperation is much more evolved now and assumes more sophisticated and complex forms than in the old times when it was considered self-evident. To detect this more complex cooperation, however strong it may be, requires external explanation. The starting point is a zero hypothesis of non-cooperation, from which cooperation emerges by some nontrivial mechanism to be discovered.

Prisoner's Dilemma Game

Games in General *Game theory* [10,11], introduced in 1940s by John von Neumann and Oskar Morgenstern provides very fertile ground on which to model the starting assumption of selfish individuals. In this model, agents are not only utterly selfish, but also absolutely rational. In a typical setup, agents meet in pairs and each of them chooses one of S possible strategies. If the first agent adopts strategy i and the second agent strategy j, then the first one gains an amount denoted A_{ij}, while the other's gain is A_{ji}. The matrix A is called a *payoff matrix*, and various types of games are distinguished according to its properties. For example, if the gain of one agent equals the loss of the other agent, so $A_{ij} + A_{ji} = 0$ for all i and j, it is a zero-sum game. If the gain of the winning party is always smaller than the loss of the adversary, the game is a negative-sum one, so it holds that $A_{ij} + A_{ji} < 0$ for all i and j.

It is assumed that the rules of the game, quantified in the matrix A, are known to both players. Therefore, they can build their strategies on rational analysis of the payoff matrix. It may happen that one strategy gives highest gain, irrespective of the action taken by the opponent. Formally, the first player's strategy, k, is such that $A_{ij} \leq A_{kj}$ for any strategy i of the first player and any strategy j of the second player. The same may also hold for the second player. Suppose that l is her best strategy, irrespective of the action of the first player. Obviously then, the first person always plays k while the second always plays l. If either player changes his or her strategy unilaterally, he or she is instantly worse off. Such a situation, if it happens, is called Nash equilibrium, after John F. Nash, a mathematician who devoted much of his career to applications of game theory in economics [12,13].

It is vital to understand that the notion of Nash equilibrium differs fundamentally from the usual equilibrium studied in various branches of physics. Whether the problem is to find the equilibrium of solid bodies on a lever or thermodynamic equilibrium in a system of steam, water and ice, the situation is formalized by finding a unique function to be minimized (energy, for example). On the contrary, Nash equilibrium means that every player maxi-

mizes his or her own function, with state of all other players fixed [14].

Introduction to Prisoner's Dilemma Collaboration has been studied using a very simple two-player game with two strategies, called the prisoner's dilemma game. Imagine that the police have arrested two accomplices for a burglary, but do not have enough evidence to prove that they actually committed the crime. The prisoners are kept well separated and the investigator offers a deal to each of them independently. Each criminal is given a promise that if he parts company with the other and confesses to the robbery they did together, he will be rewarded. If both of them confess, they will be jailed only for a short period; if one confesses and the other does not, then the first one is released and rewarded and the other gets a severe punishment; if neither of them confesses, both are released. Each must choose without knowing the other's choice.

The strategy of collaboration (C) dictates not to confess. In fact, it would be most beneficial to both suspects, if both collaborate. However, individually it is more tempting to defect (D) and confess to the police, as it prevents the situation in which the other confesses and the individual is punished alone. If both prisoners reckon in this way, the result is that both defect and must suffer some time in prison. Hence the dilemma.

The payoff matrix of this prisoner's dilemma game is characterized by four numbers, the gain if both defect, $A_{DD} = P$, or both cooperate, $A_{CC} = R$, and the gain of the defector $A_{DC} = T$ and the collaborator $A_{CD} = S$, if one of them defects and the other does not. So,

$$A_{PD} = \begin{matrix} \\ C \\ D \end{matrix} \begin{matrix} \begin{matrix} C & D \end{matrix} \\ \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{matrix} . \tag{1}$$

In order for the prisoner's dilemma to work as described above, the values must satisfy the inequalities $T > R > P > S$ and $2R > S + T$. The most studied values of the parameters are

$$T = 5\,, \quad R = 3\,, \quad P = 1\,, \quad S = 0\,. \tag{2}$$

It is easy to see that there is a single Nash equilibrium: the case in which both players defect. We arrived at this conclusion intuitively, above, and an exact check is even quicker. If the second player collaborates, the rewards of the first are R and T if he collaborates or defects, respectively; because $T > R$, defection is better. Similarly, if the second player defects, the gains for the first are S and P, and as $P > S$, it is again preferable to defect. So, defection is the optimal strategy.

Evolution of Cooperation

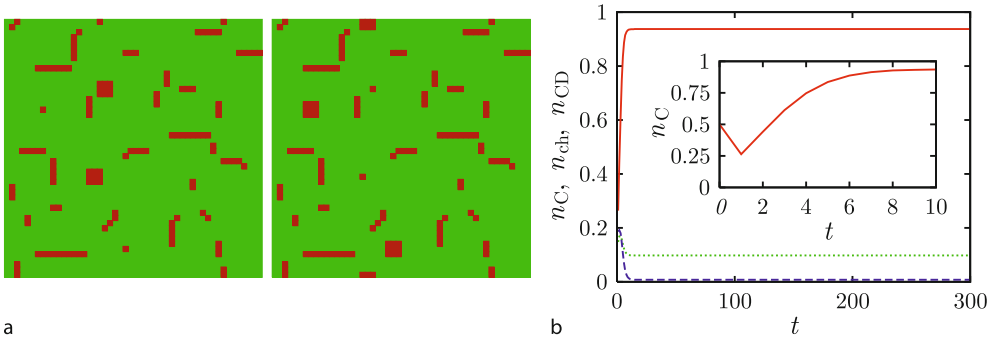
Iterated Prisoner's Dilemma The prisoner's dilemma game is rather trivial: if the players meet only once, defecting is certainly the best strategy. The Nash equilibrium excludes cooperation a priori, coinciding with the neo-liberalist view of society. What, then, is the mechanism behind the cooperation seen in reality? If there is a way for emergence of cooperation in the prisoner's dilemma game, perhaps we can see why reality sometimes follows that path.

Things become much more complex if the players face each other in an unlimited series of encounters. Robert Axelrod showed [15,16,17], in a series of computer tournaments, that complex strategies, taking into account past actions of the adversary, perform much better than simple defection. The generic and repeatedly confirmed outcome was that the strategy called Tit-For-Tat (TFT) outperformed all the rest of the strategies tried. It consists in playing cooperation at first and then repeating the last action of the other player.

Even more complicated behavior emerges if there is some level of noise. The players can make mistakes, or sometimes play at random, not obeying their usual strategies. New strategies may emerge due to mutations. Also relevant is the length of the players' memories, i. e. how many steps in the past they remember their opponents' actions. When all agents are put into a common room, where all of them they can play with all others, mixtures of strategies emerge [18,19]. The composition of the mixture depends on the length of the memory of the agents; if the memory length is unlimited, the evolution of the system proceeds indefinitely, introducing more and more complex strategies all the time.

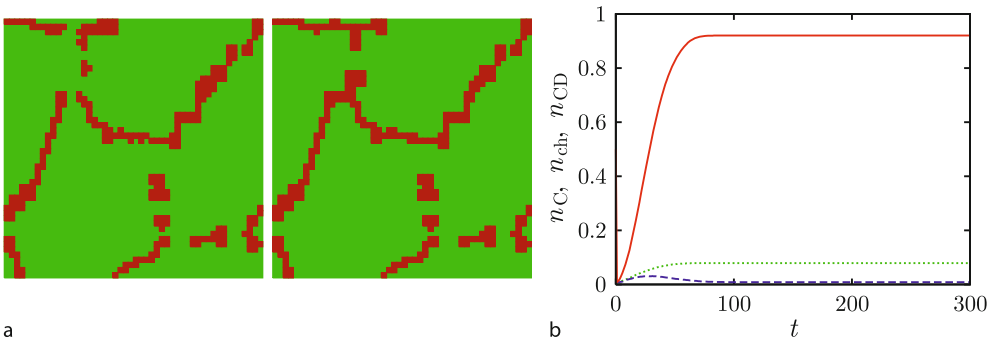
Spatial Prisoner's Dilemma The trivial prisoner's dilemma game becomes complex when the players play again and again with the same opponent. But, extending the model to cover a wider field, how it could be that among billions of people a player finds the same opponent he or she has already faced? If the two were molecules in a well-stirred container, their chance to meet repeatedly would be virtually nil. Some *compartmentalization* of the agents must be imposed, in order to apply the iterative prisoner's dilemma to a large number of players. But the influence of compartmentalization can also be studied independently, preferably in its extreme version, where immobile agents are fixed at the vertices of a network, for example a two-dimensional square lattice: that is the structure used throughout this subsection.

Agents play the usual prisoner's dilemma game with their eight neighbors. The complexity stems from their



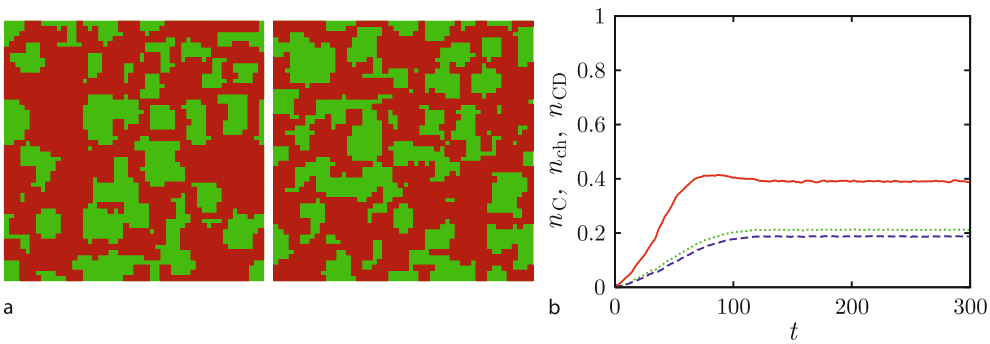
Social Processes, Physical Models of, Figure 1

Spatial prisoner's dilemma. In the *left two panels*, the configuration is drawn in times t_1, t_2, $t_1 < t_2$, on the lattice with $L = 50$. Cooperating agents are shown in *green*, the defectors in *red*. In the *right panel*, time evolution of the density of cooperators n_C (*solid line*), density of agents changing their state from one step to the next n_{ch} (*dashed line*), and density of interfaces n_{CD} (*dotted line*) are shown. System size is $L = 100$, data are averaged over 100 realizations. The incentive to defect is, $b = 1.05$. The initial condition is random spatial distribution of cooperators with density $n_C = 0.5$. In the *inset*, the detail for short times is shown



Social Processes, Physical Models of, Figure 2

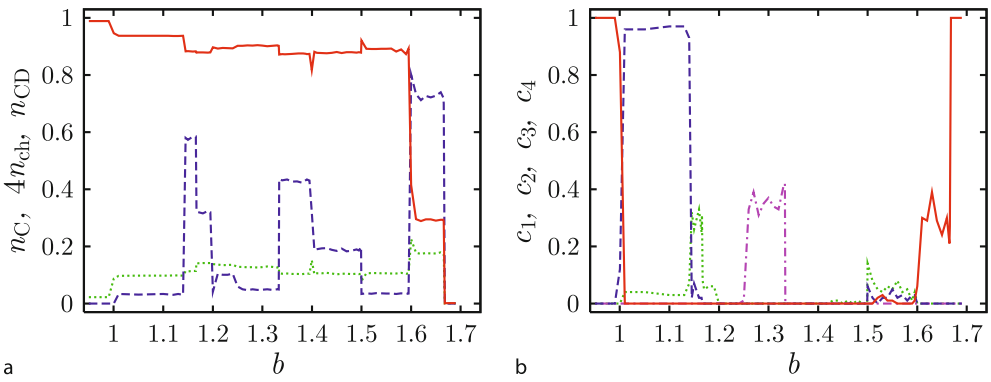
Same as Fig. 1 with $b = 1.5$



Social Processes, Physical Models of, Figure 3

Same as Fig. 1 with $b = 1.6$

ability to change their strategies from one step to the next [20]. For simplicity, let us consider only the 0-memory strategies: collaboration and defection. At each round, the agent adds her gain from the eight plays with her neighbors, who do the same in their turn. Each agent then decides upon her strategy for the next step. She looks at the gains of her neighbors and compares these numbers with her own. If any of the neighbors earns more than she, the agent adopts the current strategy of the most successful neighbor for the next step. Otherwise, the agent repeats her previous strategy. The evolution of the strategies proceeds by imitation.



Social Processes, Physical Models of, Figure 4

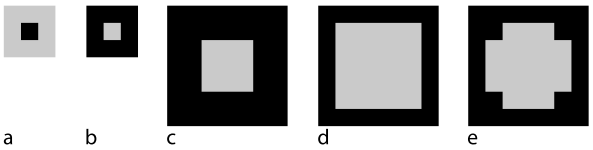
Spatial prisoner's dilemma. Various parameters of the stationary state depending on the incentive to defect *b*. All data are calculated for a lattice with $L = 100$; initial configuration is a random spatial distribution of cooperators with density $n_C = 0.5$; data are averaged over 100 realizations. The *left panel* shows the density of cooperators (*solid line*), density of sites changing their state from one step to the next (*dashed line*), and density of bonds connecting cooperators with defectors (*dotted line*). The *right panel* shows the fraction of realizations reaching static configuration (*solid line*), and fractions reaching cyclic attractors with period 2 (*dashed line*), 3 (*dot-dashed line*), and 4 (*dotted line*)

The payoff matrix can be simplified so that there is only one control parameter, the temptation to defect b. It is assumed that $T = b > 1$, $R = 1$, and $P = S = 0$. The spatial prisoner's dilemma game is then simulated, starting from a random configuration, where every agent chooses cooperation or defection with equal probability. Further evolution proceeds by deterministic parallel dynamics, as described above. In fact, the system is a cellular automaton with specific, relatively complicated, update rules. The resulting configurations are illustrated in Figs. 1 to 3. The cooperators survive in significant proportion even if the temptation for defection b is large. However, their spatial arrangement strongly depends on the value of b. Figure 1 shows the situation for the value $b = 1.05$, only slightly above the lowest limit compatible with prisoner's dilemma inequalities. Small groups of defectors survive within a large sea of cooperators. Most of the defector groups are stable or change cyclically with a short period. When the temptation is increased to $b = 1.5$, isolated islands of defectors grow into strings joined together at some places, as seen in Fig. 2. Cooperation still prevails, but only within patches encircled by defectors. The spatial structure exhibits only small variations in time. This changes when the temptation grows further. Figure 3, where $b = 1.6$, already shows more defectors than collaborators and the arrangement changes chaotically.

Systematic study of the dependence on b is summarized in Fig. 4 showing data from simulations, averaged over many realizations of the initial conditions. First, the fraction of cooperators, n_C, remains quite high, above 80%, up to about $b = 1.6$, where it drops suddenly to about 40%; only above $b = 1.7$ do the cooperators vanish. Clearly, repeated games induced by fixed spatial arrangement of the players strongly encourages cooperation. Note that the players do not follow any strategy based on observation of the past behavior of the agents. The cooperation emerges spontaneously.

Further inspection of the dependence of n_C on b reveals sudden jumps at specific values of the temptation parameter. The jumps are even more pronounced in some other parameters characterizing the stationary state, namely the concentration n_{ch} of sites changing their state from one step to the next, and the concentration of bonds (neighbor pairs) connecting a defector and a cooperator, n_{CD}, which is the measure of the density of interfaces between cooperating and defecting domains. Furthermore, some realizations end in static configurations, while other reach a cyclic attractor with short periods 2, 3, or 4. The fraction of realizations corresponding to these four types are denoted c_1, c_2, c_3, and c_4, respectively. Of course, some initial conditions may also lead to longer stationary periods or to a quasi-chaotic state with a barely identifiable periodicity. Interestingly, the quantities c_i depend on b in a very irregular fashion. For example, the period-2 states dominate in an interval from $b = 1$ to about $b = 1.14$, but are rare elsewhere. In contrast, the period-3 states occur practically only in the interval $5/4 < b < 4/3$.

To understand these features it is necessary to analyze various spatial structures produced in the dynamics [21]. For example, an isolated defector in the sea of cooperators, as shown in Fig. 5a, has 8 cooperating neighbors, so its gain is $g_D = 8p$. The neighbors themselves have 7 cooperating neighbors, resulting in gain $g_C = 7$. This means that for $p < 7/8$ the defector becomes cooperator in the next step,



Social Processes, Physical Models of, Figure 5
Several configurations of the defectors (*black*) and cooperators (*gray*) in the spatial prisoner's dilemma game on a square lattice. Their stability is discussed in the text

but for $p \geq 7/8$ it survives. Similar analysis shows that an isolated cooperator surrounded by defectors, Fig. 5b, is never stable. Indeed, the cooperator gains 0, but each of the the neighboring defectors gains p.

It is possible to proceed further to more and more complicated geometries. For example, a 2×2 square of cooperators, Fig. 5c, is much more stable than an isolated cooperator. Each of the four gains $g_C = 3$ from its three cooperating neighbors. The defecting neighbors are of two types. At the corners, they have one cooperator to exploit, so their gain is p. The defectors adjacent to the edges of the square have two cooperating neighbors, giving higher gain $g_C = 2p$. Therefore, the square persists for $p \leq 3/2$. But that is not all; the defecting neighbors see that the cooperators gain more, so they become cooperators themselves in the next step. For $p < 3/2$ the initial 2×2 square grows into a 4×4 square of cooperators, Fig. 5d, which in turn expands into 6×6 square, and so on. Cooperation spreads despite the relatively large value of the temptation to defect!

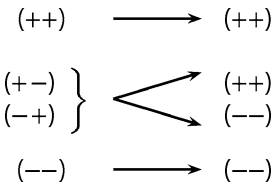
Some configurations may exhibit periodic changes. A multitude of possible generalizations can be found in the literature, either concerning update rules [22,23,24,25] or geometry of the links connecting interacting neighbors [26,27,28,29,30]. However, the general features of the spatial prisoner's dilemma game remain in force, namely the fact that repeated plays against the same agents, which are dictated by the geometry of the social network, leads naturally to coexistence of large patches of collaborators alternated with various arrangements of defectors.

To sum up, spontaneous emergence of cooperation seems to be reproduced in model situations, on condition that the game is played repeatedly. The overall picture of cooperation is typically rather complex, precluding any simplistic ideology-based conclusions.

Opinion Dynamics

Voter Model

People can make up their minds by simply looking around and picking the opinions of randomly chosen neighbors.



Social Processes, Physical Models of, Figure 6
Illustration of the dynamics of opinions in the voter model. The scheme shows how pairs of neighboring sites are updated. If the two sites have identical states, they remain unchanged. If the single-site states differ, they can become either both + or both − with equal probability. Conservation of average magnetization follows directly

That is the idea behind the stochastic process introduced in 1970s [31,32] and called the *voter model*. The voter model plays a special role among other models of opinion spreading and consensus formation, because it is exactly soluble in any spatial dimension, while showing highly non-trivial dynamics [33,34,35,36,37,38,39,40,41]. Physicists are interested in the voter model, as it can shed light on spinodal decomposition [35] and chemists use it to model catalytic reactions [36,37,42].

Definition In fact, there is a whole family of diverse voter models [43]; the exactly soluble class comprises the so-called linear voter models.

The most studied geometry is the d-dimensional hypercubic lattice Λ of linear dimension L, with periodic boundary conditions. The coordinates of the point $x \in \Lambda$ will be denoted x_α, $\alpha = 1, 2, \ldots, d$.

On each lattice site there is an agent whose state can be either $+1$ or -1. These two choices can represent a person's political preferences in a two-party system, hence the name voter model. The configuration of the entire system is described by a point in the configuration space $\sigma \in S = \{-1, +1\}^\Lambda$. The state of the site $x \in \Lambda$ is denoted $\sigma(x)$.

The dynamics of the model are very simple. In each step, one site x and its neighbor y are chosen randomly. Then x adopts the state of y, so $\sigma(x)$ is replaced by $\sigma(y)$, as illustrated in Fig. 6. This scheme also demonstrates one important property of the voter model. The average opinion, which a physicist would call magnetization, defined as $m = (1/|\Lambda|) \sum_{x \in \Lambda} \sigma(x)$ is conserved after averaging all possible realizations of the process, even though in individual realizations it may fluctuate. It is also evident that the uniform states where all sites are either $+1$ or -1 never change. These two configurations are *absorbing states* of the voter model.

In one dimension the dynamics can be easily understood. The configuration is determined by the sequence

of "domain walls", separating regions uniformly populated by $+1$ or -1. The configuration can change only by flipping the state of the sites beside the domain wall. Either the agent to the left of the wall adopts the state of the agent on the right side, or vice versa. In the former case the domain wall moves one step leftwards, in the latter case it jumps rightwards. Both possibilities have the same probability, so the domain wall performs a random walk. Moreover, when two domain walls meet, the region bordered by them disappears and the domain walls themselves annihilate. So, the one-dimensional voter model is exactly mapped onto the system of annihilating random walkers. This model is quite well understood [44]. Unfortunately, in any dimension larger than one such equivalence is no longer valid.

A slightly more formal description goes as follows: The voter model is a continuous-time Markov process σ_t which takes values in the configuration space $S = \{-1, +1\}^{\Lambda}$. For any $\sigma \in S$ denote as σ^x the state which is obtained from σ by flipping the state of site $x \in \Lambda$, so $\sigma^x(y) = (1 - 2\delta_{xy})\sigma(y)$. It is necessary to know the set of nearest neighbors of x on the lattice Λ. In the case of a d-dimensional hypercubic lattice, there are $2d$ neighbors obtained by shifting the point x by either $+1$ or -1 along the d Cartesian axes. Denote μth neighbor of x by x^μ. The transition rates describing a single flip are

$$w(\sigma, \sigma^x) = \frac{1}{2}\left[1 - \sigma(x)\frac{1}{2d}\sum_{\mu=1}^{2d}\sigma(x^\mu)\right] \tag{3}$$

while all other transition rates are zero

$$w(\sigma, \sigma') = 0\,, \quad \text{if} \quad |\{x \in \Lambda : \sigma(x) \neq \sigma'(x)\}| \neq 1\,. \tag{4}$$

The dynamics proceeds according to the master equation

$$\frac{\mathrm{d}}{\mathrm{d}t}p_t(\sigma) = \sum_{x\in\Lambda}\left[w(\sigma^x, \sigma)p_t(\sigma^x) - w(\sigma, \sigma^x)p_t(\sigma)\right]. \tag{5}$$

Solution It is possible to directly write the equation for the average state on a single site $S(x,t) \equiv \langle\sigma_t(x)\rangle = \sum_{\sigma\in S} p_t(\sigma)\sigma(x)$, starting with the master equation (5). The result is

$$\frac{\mathrm{d}}{\mathrm{d}t}S(x,t) = \Delta_x S(x,t)\,, \tag{6}$$

where the notation $\Delta_x f(x) = -f(x) + \frac{1}{2d}\sum_{\mu=1}^{2d} f(x^\mu)$ is used for the discrete Laplace operator on the d-dimensional hypercubic lattice. Equation (6) is in fact the discrete diffusion equation, describing the movement of a random walker over the lattice.

For two-site correlations $R(x-y,t) \equiv \langle\sigma_t(x)\sigma_t(y)\rangle$, the equation is

$$\frac{\mathrm{d}}{\mathrm{d}t}R(x,t) = 2\Delta_x R(x,t) \tag{7}$$

for $x \neq 0$. If the two points coincide, it holds trivially $R(0,t) = 1$ for all times $t \geq 0$. Technically, this is the boundary condition for the solution of the discrete diffusion Eq. (7). In this way it is possible to get closed equations for correlations of all orders.

Equation (6) can be solved by standard mathematical techniques, including Fourier and Laplace transforms. The initial condition is as follows: at the origin the state is $+1$, while all other sites are $+1$ or $+1$ with equal probability, so $S(x,0) = \delta_{x0}$. The solution is

$$S(x,t) = \mathrm{e}^{-t}\prod_{\alpha=1}^{d} I_{x_\alpha}\left(\frac{t}{d}\right), \tag{8}$$

where $I_\nu(z)$ is the modified Bessel function [45].

Asymptotic behavior of the Bessel function gives that, for large times, the average state of any site decays to zero as $S(x,t) \sim t^{-1/2}$, $t \to \infty$. Recall that the initial condition was $S(x,0) = \delta_{x0}$. So, the average state of the site at the origin decays to zero monotonically, while the other sites, $x \neq 0$, exhibit first an increase in $S(x,t)$ and then decay at later times. This can be understood as propagation of a diffusive wave of $+1$'s from the origin to the rest of the lattice, eventually vanishing at large times.

More information on the dynamics is contained in the two-site correlation function $R(x,t)$. Apart from the factor 2, it obeys the same equation as $S(x,t)$. However, there is an important difference. Equation (7) holds only for $x \neq 0$ and, besides the initial condition $R(x,0) = \delta_{0x}$, the solution must also satisfy the boundary condition $R(0,t) = 1$. Nevertheless, it is possible, again, to proceed by Fourier- and Laplace-transforming the Eq. (7). The boundary condition enters through the yet unknown function

$$n_{+-}(t) = \frac{1}{2}\left(1 - \frac{1}{2d}\sum_{\mu=1}^{2d} R(x - x^\mu, t)\right), \tag{9}$$

which is simply the concentration of interfaces, i. e., the fraction of the bonds connecting sites with unequal state, also called *active bonds*. This quantity is an important measure of the level of activity in the system, as changes of the configuration can occur only at sites adjacent to active bonds.

Let us now separately discuss the results in dimensions $d = 1$, $d = 2$, and $d \geq 3$. In one dimension, the density of

interfaces can be expressed in closed form

$$n_{+-}(t) = \frac{1}{2} e^{-2t} [I_0(2t) + I_1(2t)] \,. \tag{10}$$

For large times, the behavior is $n_{+-}(t) \simeq (4\pi t)^{-1/2}$ and $\lim_{t\to\infty} R(x,t) = 1$. This means that the activity measured by $n_{+-}(t)$ slowly decays to zero and eventually all the agents become fully correlated. Both features are signatures of complete ordering, which is the fate of the voter model in one dimension.

For $d = 2$ the situation is similar. Again, the density of interfaces decays to zero and the correlation function approaches 1 for large times. However, the evolution is much slower. The solution can be expressed in closed form using the so-called elliptic integrals, but let us mention here only the asymptotic behavior

$$n_{+-}(t) \simeq \frac{\pi}{2} \frac{1}{\ln(16t)} \,, \quad t \to \infty \,. \tag{11}$$

The behavior changes dramatically for $d \geq 3$, because $n_{+-}(t)$ has a positive limit for $t \to \infty$. In short, the activity never ceases and the voter model never reaches a totally ordered state. Explicit calculations show that the stationary density of interfaces becomes

$$n_{+-}(\infty) = \left(2d \int_0^\infty \left[e^{-t} I_0(t) \right]^d \mathrm{d}t \right)^{-1} \,. \tag{12}$$

The asymptotic behavior of the Bessel function guarantees that the integral converges for $d > 2$ and $n_{+-}(\infty) > 0$. Numerical values for some values of d are listed in Table 1 below. One can also find how the density of interfaces approaches its asymptotic value, $n_{+-}(t) - n_{+-}(\infty) \sim t^{1-(d/2)}$.

All the above results for the voter model tacitly assume an infinitely large lattice. In practice, e. g., in numerical simulations, the system always consists of a finite number, $N = L^d$, of sites. This implies that eventually the dynamics lead to one of the two absorbing states, even for $d \geq 3$, where the *infinite* system never orders. The time needed to reach any of the absorbing states in a particular realization of the process is called *stopping time* τ_{st}. The typical scale for the stopping time, called *consensus time* $\tau(N)$, diverges in the thermodynamics limit. It is possible to show rigorously [46] that the consensus time grows like $\tau(N) \sim N^2$ in $d = 1$, $\sim N \ln N$ in $d = 2$ and $\sim N$ for $d \geq 3$. These formulae can be understood on an intuitive level, at least in one dimension, where the dynamics is equivalent to diffusion of domain walls, annihilating upon encounter. The size of homogeneous areas grows, therefore, as the mean displacement of a random walk, $\sim t^{-1/2}$. The time needed to cover the whole system of size N is therefore $\sim N^2$. A similar consideration is also feasible in higher dimensions [36].

The asymptotics of the voter model on d-dimensional hypercubic lattice are summarized in Table 1. Where the behavior is indicated by "$\sim$", the meaning is $t \to \infty$ or $x \to \infty$ or $N \to \infty$, according to the context.

Local Majority Models

Galam Model In democracy, consensus on an issue is rarely achieved by simply waiting until one of the options pervades the whole system through pairwise contact of individuals. Instead, there are various hierarchical levels of decision making, each of which comes to a conclusion based on the principle of majority. This is supposed to lead to a state in which most people are satisfied, when the opinion of the majority of those who participate in the decision process is declared a law. If there were only a single hierarchical level, namely a public referendum (Switzerland may serve as a model example), one could be quite sure that the outcome really represents the majority opinion in the society. However, as soon as there are more levels, and decisions on lower level are passed on to levels above, there is no obvious guarantee that the opinions are not distorted or even reversed. Serge Galam devised a simple model demonstrating that the distortions may not be an exception, but rather a rule [47,48,49,50,51].

The reason for breaking the society into several hierarchical levels is that the cost of communication in a too large collective is prohibitive. It is interesting that this is also the main technical obstacle hindering the introduction of secure electronic elections. One can imagine voting for a presidential candidate from one's home computer over the Internet, but to make the protocol reliable enough to ensure proper secrecy *and* resistance to any attempt at fraud on the scale of an entire nation seems to be beyond currently available technical capabilities [52].

So, the model takes it for granted that the society is organized hierarchically and, on each level, the decision is made within a small group, say, of 3 people. Within each group, majority rule tells what opinion shall be held by the representative of the group when sent to make decision on the upper level. Now, suppose there are only two possible choices, A and B and the fraction of people with opinion A on the level l is $n_A(l)$. On the basic level, $n_A(0)$ represents the concentration of A in the whole population.

Supposing that the people are not correlated in any way, the concentrations $n_A(l)$ provide full information on the system. Essentially, this is a kind of mean-field approx-

Social Processes, Physical Models of, Table 1
Table of some properties of the voter model on a hypercubic lattice in d spatial dimensions

d	$n_{+-}(\infty)$	$n_{+-}(t) - n_{+-}(\infty)$	$R(x,\infty)$	$\tau(N)$		d	$n_{+-}(\infty)$
1	0	$\sim t^{-1/2}$	1	$\sim N^2$		3	0.329731 . . .
2	0	$\sim (\ln t)^{-1}$	1	$\sim N \ln N$		4	0.403399 . . .
≥ 3	> 0	$\sim t^{1-d/2}$	$\sim \lvert x\rvert^{2-d}$	$\sim N$		5	0.432410 . . .
						∞	1/2

imation, neglecting any social structure or network within one hierarchical level. The model of majority decisions on a complete graph, also studied in the literature [53,54], is to large extent equivalent to the Galam model.

The dynamics of the model consists in determining the fraction of A at level $n+1$ on the basis of the concentration on level n. For groups of size 3 it leads to the recurrence relation

$$n_A(l+1) = n_A^3(l) + 3\,(1 - n_A(l))\, n_A^2(l)\,. \tag{13}$$

The evolution according to this rule is depicted in Fig. 7a. There are three fixed points at values $n_{A\,\mathrm{fix}} = 0, 1$, and 1/2, which can be easily checked by insertion in (13). The two extreme values are stable, while the middle point is an unstable fixed point. This reminds us of physical systems with a phase transition, treated by renormalization group techniques [55]. In that formalism, the unstable fixed point corresponds to the critical value of a control parameter and the stable fixed points represent the various possible phases. Here there are two phases, populated uniformly by either all A or all B opinions, constituting total consensus on the issue.

The transition occurs at the symmetric point $n_{A\,\mathrm{fix}} = 1/2$, where exactly half of the people have opinion A. In this sense, the decision making is "fair" because it leads to consensus according to the initial majority. However, a seemingly minor modification makes a big difference. Imagine that the groups formed at each level are composed of an even number of persons, e. g., four. The decision must be taken as to what to do if exactly half of the group pushes for decision A but the rest for B. No majority emerges and the tie must be resolved. If there is an arbitrarily small bias towards one of the choices, say, B, the recurrence relation for the concentrations of opinion A is

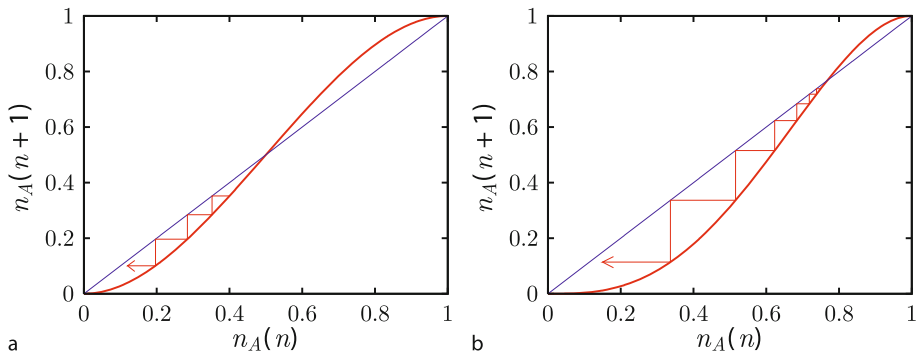
$$n_A(l+1) = n_A^4(l) + 4\,(1 - n_A(l))\, n_A^3(l)\,. \tag{14}$$

Figure 7b shows what happens. The unstable fixed point is shifted significantly to higher values of n_A, namely to

$$n_{A\,\mathrm{fix}} = \frac{1+\sqrt{13}}{6} = 0.76759\ldots \tag{15}$$

which means that opinion B may win even if it was initially in a minority!

The situation can be even more complicated if some of the people systematically vote against the majority, instead of following the crowd. Such agents are called *contrarians* [56,57]. For simplicity, take (again) groups of size 3. First, the majority option within the group in question is



Social Processes, Physical Models of, Figure 7
Galam model. In a, graph of the recurrence relation for groups of size 3. The *stairs* indicate how the iterations of the relation drive the system away from the unstable fixed point $n_{A\,\mathrm{fix}} = 1/2$. In b, an analogous scheme for groups of size 4 with infinitesimal bias in favor of option B. Note the shift of the unstable fixed point to value $n_{A\,\mathrm{fix}} = (1+\sqrt{13})/6$

found, then each member of the group flips her opinion with probability p_c, which is to be interpreted as the concentration of contrarians in the population. The recursion relation is modified to

$$n_A(l+1) = (1-p_c)\left[n_A^3(l) + 3\,(1-n_A(l))\,n_A^2(l)\right] + p_c\left[(1-n_A(l))^3 + 3(1-n_A(l))^2 n_A(l)\right] \tag{16}$$

and results in a shift of the stable fixed points away from the endpoints of the interval $[0, 1]$. This means that total consensus is never reached; the contrarians always introduce a certain level of dissidence, but generally the final decision respects the initial majority option in the society. For $p_c < 1/6$ this picture remains valid, as there are three fixed points, with the unstable one keeping its position at $n_{A\,\mathrm{fix}} = 1/2$. However, if the concentration of contrarians rises and $p_c > 1/6$, the three fixed points coalesce to a single stable fixed point at $n_{A\,\mathrm{fix}} = 1/2$. No consensus is ever reached and the distribution of opinions tend to a precisely equilibrated state of equally represented opinions A and B. Several recent cases of popular referenda or presidential elections ending in extremely narrow victories come to mind, showing that the presence of contrarians may have palpable consequences for our lives.

A natural question is, how fast is consensus approached as one climbs higher and higher up ladder of hierarchies? It is evident that as we start closer to the unstable fixed point, it takes us more time to approach one of the stable points. To get a quantitative estimate, let us turn to the simplest case described by Eq. (13) and replace the discrete level index l by a continuous variable l which can be interpreted as time elapsed during the formation of the consensus. Thus, the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}l} n_A(l) = -n_A(l)(n_A(l)-1)(2n_A(l)-1) \tag{17}$$

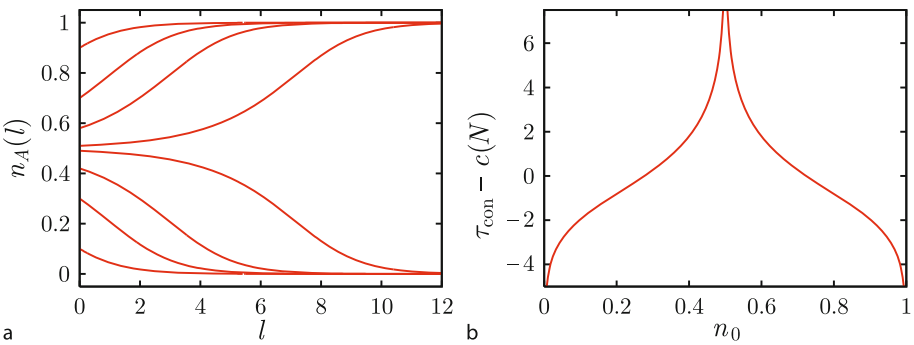
is obtained, which can be solved relatively easily. The solution for several initial conditions is shown in Fig. 8a. Of course, starting from any point $n_A(0)$ inside the interval $(0, 1)$, the time to reach either of the stable fixed points 0 or 1 is infinite. However, in reality it is not necessary to reach the fixed point exactly, because the population consists of a finite number N of people, so full consensus is reached if the procedure is stopped at a distance $1/N$ from the fixed point. The time to achieve that diverges for increasing N as $\sim \ln N$, as will be seen from the explicit calculation.

To be more precise, the stopping time t_{st} as a function of the initial concentration $n_0 \equiv n_A(0) < 1/2$ can be defined by the formulae $n_A(0) = n_0$ and $n_A(t_{\mathrm{st}}(n_0)) = 1/N$, expressing the initial and final conditions, respectively. The case $n_0 > 1/2$ differs only in the final condition, which is $n_A(t_{\mathrm{st}}(n_0)) = 1 - 1/N$. Knowing the general solution of (17) the formula is

$$t_{\mathrm{st}}(n_0) = \ln\frac{n_0(1-n_0)}{(2n_0-1)^2} + \ln\frac{(N-2)^2}{N-1}\,. \tag{18}$$

Hence the already-mentioned logarithmic divergence of the stopping time for $N \to \infty$. The N-independent part of the stopping time can be seen in Fig. 8b. Note the divergence for $n_0 \to 1/2$, which can be regarded as a sign of a certain kind of dynamic phase transition.

The Galam model has potential for extensions in various directions and, indeed, it was further generalized [58,59,60,61,62,63,64,65,66,67,68]. For us, the most physically relevant question is how the dynamics change



Social Processes, Physical Models of, Figure 8

Galam model with continuous levels, for group size 3. The evolution according to Eq. (17) is shown in a, with initial concentration of opinion *A* equal $n_0 = 0.9, 0.7, 0.58, 0.51, 0.49, 0.42, 0.3$, and 0.1 (from *top* to *bottom*). In b, the time to reach consensus in a society of *N* people is shown. In this graph, $c(N) = \ln\frac{(N-2)^2}{N-1}$

if the people are distributed on a fixed lattice or network [69,70,71]. After all, the Galam model has all the attributes of a mean-field version of some more complicated model, even though it may not be obvious which one [72]. The next section will make the point more clear.

Sznajd Model Clearly, a group of people puts stronger pressure on an individual than each member of the group taken separately. Nature designed humans so that they follow the crowd. If you see two or more people sharing an opinion on certain issue, you are tempted to join. This is the basic idea behind the model invented by Katarzyna Sznajd-Weron and Józef Sznajd (the daughter and the father) [73]. In its first version, the model was defined on a one-dimensional lattice of length N. Each site is inhabited by an agent which can be in two states, denoted $+1$ and -1, as in the voter or majority-rule models. It may correspond to people choosing between two dominant brands of a certain product in the market or voting in a two-party political system. In each step of the dynamics, a pair of neighbors is chosen randomly. If they are in the same state, say, $+1$, then the two sites adjacent to the pair adopt the same opinion $+1$, propagating the consensus outwards. Conversely, if they differ in opinion, they propagate the dissensus. The rule is shown schematically in Fig. 9.

Formally, suppose the chosen pair is in state $(\sigma(x), \sigma(x+1))$, at time t. If $\sigma(x) = \sigma(x+1)$, the neighbors of the pair are updated as $\sigma_{t+1/N}(x-1) = \sigma_{t+1/N}(x+2) = \sigma(x)$, while for $\sigma(x) \neq \sigma(x+1)$ the update rule is $\sigma_{t+1/N}(x-1) = \sigma(x+1), \sigma_{t+1/N}(x+2) = \sigma(x)$. Unlike the voter or majority rule model, there are three absorbing states. Besides the obvious "ferromagnetic" states of all $+1$ or all -1, there is the "antiferromagnetic" state where the sites in states $+1$ and -1 alternate regularly. Strictly speaking, there are two such states, one characterized by $+1$ at sites with odd coordinates and the other by $+1$ at even sites.

Looking at the rules more thoroughly, it is evident that the linear chain can be divided into two sublattices, one of them containing all odd coordinates x, the other all even sites. The state of the agents in one sublattice is never influenced by the agents in the other one. Moreover, the dynamics within one sublattice are a trivial modification of the voter model. The only difference is that in the voter model, one site induces change of state of one of its neighbors, while here the agent induces *both* of its neighbors adopt its state. Indeed, the update rule can be written as $\sigma_{t+1/N}(x-1) = \sigma(x+1)$, $\sigma_{t+1/N}(x+2) = \sigma(x)$, irrespective of the relation between $\sigma(x)$ and $\sigma(x+1)$ [74]. This observation simplifies the model to a large extent. Indeed, it is equivalent to two copies of the voter model evolving in parallel, where the only coupling between them comes from the fact that the sites to be updated are neighbors on the original lattice. The two voter models evolve so that the update occurs at the same place in both copies. However, the initial state of both sublattices is chosen randomly and independently and many properties, for example the concentrations of $+1$ in either of the sublattices, remain uncorrelated forever.

Restricting the study to one sublattice only, the solution follows the same way as was used successfully for the voter model. The flipping rates are analogous to (3), differing only in factor 2, originating from the fact that a site induces a change in state of its two neighbors. This factor means only a rescaling of time, so it is quite safe to neglect it. The result for the voter model can be translated directly to our case.

As a first application, let us look at the asymptotic states. The probability of reaching the state of all $+1$ is simply the concentration of $+1$ opinions in the initial state. This holds for both sublattices independently. From here, it is straightforward to deduce the probabilities P_+, P_- and P_{AF} that the system ends in absorbing states with all $+1$, all -1, and the antiferromagnetic state, respectively, as a function of the initial concentration of $+1$ opinions. Indeed, a state with all $+1$ means that both sublattices reached the uniform $+1$ state (and similarly for -1) while the antiferromagnetic state is obtained if one of the sublattices ended in a $+1$ state and the other in a -1 state. Hence

$$\begin{aligned} P_+(n_+) &= n_+^2 \\ P_-(n_+) &= (1-n_+)^2 \\ P_{AF}(n_+) &= 2(1-n_+)n_+ \,. \end{aligned} \tag{19}$$

The history of a single agent is also interesting. During its evolution, it may change its state several (including zero) times. The time elapsed between two subsequent changes of state is called *decision time* t_{dec}. It is natural to ask, what is the probability distribution of decision times of all agents? The answer relies on mapping the evolution of the voter model onto diffusive motion of domain walls. As explained in Sect. "Voter Model", the domain walls separating the regions of $+1$ and -1 opinions, evolve in time like annihilating random walkers. Selecting an agent at site x, its state flips if and only if the domain wall crosses the point x. The decision time is nothing other than the time between two successive visits of the random walker to the same position x, or, using language familiar to experts in random walks, the decision time is the time of first return of the random walk to the origin. Calculation of this quantity can be found in probability theory textbooks.

Model of Ref. [73]

$$\left.\begin{array}{lcl} (\cdot\,+\,+\,\cdot) & \longrightarrow & (+\,+\,+\,+) \\ (\cdot\,-\,-\,\cdot) & \longrightarrow & (-\,-\,-\,-) \\ (\cdot\,-\,+\,\cdot) & \longrightarrow & (-\,+\,-\,+) \\ (\cdot\,+\,-\,\cdot) & \longrightarrow & (-\,+\,-\,+) \end{array}\right\} \text{Sznajd model}$$

Social Processes, Physical Models of, Figure 9
Illustration of the dynamics of opinions in the the model of Sznajd–Weron and Sznajd, as it appears in the first paper [73] and as it was modified subsequently and named the "Sznajd model". The dots replace any state of the site

The decision time distribution behaves asymptotically as a power law

$$P^{>}_{\text{dec}}(t) \equiv \text{Prob}\{t_{\text{dec}} > t\} \sim t^{-1/2}\,, \quad t \to \infty \tag{20}$$

and this is exactly what was found numerically in the founding work of the Sznajds.

The solution obtained by mapping to the voter model is only the beginning of the story. Things start to be exciting again as soon as the "simplification" of the dynamic rules introduced above is made, allowing only neighbor pairs with equal states to influence their neighborhood. If the two agents in the pair do not agree in their opinions, nothing happens. In the scheme shown in Fig. 9 this corresponds to taking only first two rows as allowed updates. It is this modification that has been widely studied, and common consensus has assigned it the name Sznajd model (although calling it "Sznajds' " would perhaps do more justice to the authors).

As a first step in the analysis of the Sznajd model, it is important to note that there are again three absorbing states, all $+1$, all -1 and the antiferromagnetic one. But now the antiferromagnetic state is unstable, because randomly flipping a single site results in a nucleus of three sites in the same state, which irresistibly invades the whole system. But although unstable, the existence of such an absorbing state leaves important traces in the dynamics, as will be more clear later.

The transition rates for the underlying Markov process are

$$\begin{aligned} w(\sigma,\sigma^x) = \frac{1}{8}\Big[&\sigma(x-2)\sigma(x-1) + \sigma(x+1)\sigma(x+2) \\ &- \sigma(x)\big(\sigma(x-2) + \sigma(x-1) + \sigma(x+1) \\ &+ \sigma(x+2)\big) + 2\Big]\,. \end{aligned} \tag{21}$$

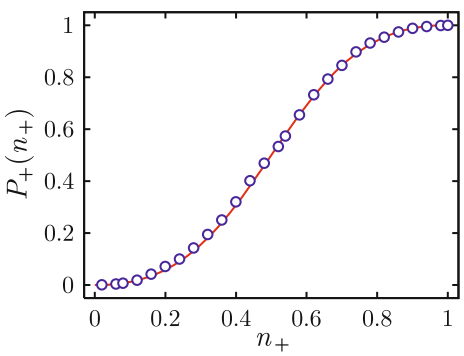
It is instructive to compare this with Eq. (3) describing voter model dynamics. The Sznajd model is a member of the family of non-linear voter models. This means that the transition rates depend non-linearly (quadratically, in fact) on the states of sites other than x.

There is a rather standard technique of approximate solution of similar systems, called the Kirkwood approximation. In our case, its use yields the probability for ending in the all $+1$ state in the form

$$P_+(n_+) = \frac{n_+^2}{(1-n_+)^2 + n_+^2}\,. \tag{22}$$

Figure 10 compares this result with numerical simulations, showing very good agreement.

Two more questions have been asked regarding the one-dimensional Sznajd model. The first is the decision time, introduced a short while ago. It was found numerically [74] that it follows the same power law (20) as in the previous case, where it resulted from underlying linear voter dynamics. For the Sznajd model, no proof is available showing that this is the exact result, but it is possible to understand this result on an intuitive level. Although the transition rates (21) correspond to a non-linear, rather than linear, voter model, the evolution of domain walls, responsible for the behavior of the decision time, remains very similar. If the domain walls are far from each other, they again perform a random walk. Non-linearity comes into play only when the domain walls come



Social Processes, Physical Models of, Figure 10
One-dimensional Sznajd model. Probability of reaching the final configuration of all sites in state $+1$, depending on the initial concentration n_+ of $+1$ sites. The *points* are numerical simulation data extracted from [75], the *line* is the analytical result (22) obtained using the Kirkwood approximation

close together, more precisely when they are one lattice spacing apart. Then, they can no longer move freely, they can make only one step towards one another, thus annihilating each other. The dynamics is again those of a diffusion-annihilation process, but the domain walls are not independent; in addition to annihilation they interact at a short distance.

The interaction is in fact a trace of the unstable antiferromagnetic absorbing state. Two walls at distance one form a small nucleus of antiferromagnetic state, thus slowing down the dynamic with respect to free annihilating random walkers. If by chance several domain walls come so close that the distance between neighboring walls is 1, they form an antiferromagnetically ordered cluster, where the walls squeezed between another two walls from both the left and right sides cannot move at all and the dynamics is hindered even more.

However, for large times, the density of domain walls is low and the influence of such antiferromagnetic islands can be neglected. Therefore, it is no big surprise that the power-law dependence of the decision time remains unaffected, at least for large times.

The second interesting question is how many agents never changed their opinion up to certain time t, or, what is the *persistence* in the dynamics of the Sznajd model? In the voter model in one dimension, there is an exact analytical solution [44] showing that the fraction of sites which retain their initial opinion at least to time t decays as $\sim t^{-3/8}$. It may seem quite surprising that the same behavior was found numerically in the one-dimensional Sznajd model [76,77]. However, applying the same arguments as above, showing that the long-time dynamics is essentially dominated by annihilating random walks of the domain walls, it follows that it is quite plausible that the long-time behavior of the persistence is governed by the same exponent 3/8 both in the voter and the Sznajd models in one dimension.

Agents organized on a line are now understood, and it is natural to ask what changes if they are arranged in a two-dimensional mesh, like the people scattered on the surface of the Earth. For example, on a square lattice, one-dimensional dynamics can be generalized in such a way that a pair of neighboring agents is chosen, and if the two have the same state, they make all their 6 nearest neighbors share their state [78,79]. Slightly more formally, let x be a random coordinate on a square lattice with $N = L^2$ sites and $y = x^\nu$ one of the 4 neighbors of x. If $\sigma_t(x) = \sigma_t(y)$, then the configuration is updated as $\sigma_{t+1/N}(x^\mu) = \sigma_{t+1/N}(y^\mu) = \sigma_t(x)$ for all $\mu \in \{1, 2, 3, 4\}$. Numerical simulations show an interesting difference from the one-dimensional case. The probability $P_+(n_+)$ has discontinuity at $n_+ = 1/2$, indicating a dynamical phase transition, while in one dimension the dependence is smooth. More strongly, the data suggest that $P_+(n_+) = 0$ for $n_+ < 1/2$ and $P_+(n_+) > 0$ in the opposite case $n_+ > 1/2$. This result is equally valid in any dimension larger than 1. In the limit of infinite dimension, the behavior should coincide with the mean-field approximation, which implies putting the Sznajd model on a complete graph [54]. Such a modification is analytically solvable, as we now show.

On a complete graph with N vertices, the configuration of the system is fully described by the number N_+ of agents in state $+1$ or by the magnetization $m = 2N_+/N - 1$. The evolution of the latter is governed by transition rates which in the limit $N \to \infty$ yield the Fokker–Planck equation

$$\frac{\partial}{\partial t} P_m(m, t) = -\frac{\partial}{\partial m}\left[\frac{1}{2}(1 - m^2) m P_m(m, t)\right] \qquad (23)$$

for the probability density $P_m(m, t)$ of the magnetization at time t. Before proceeding further, note that the average magnetization $\langle m\rangle(t) = \int m P_m(m, t)\mathrm{d}m$ satisfies the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle m\rangle(t) = \frac{1}{2}\left(\langle m\rangle(t) - \langle m^3\rangle(t)\right) , \qquad (24)$$

which has a structure similar to Eq. (17) describing the dynamics of the Galam model. Indeed, equating $m = 2n_A - 1$, identifying the level index with time, $l = t$, and making the approximation $\langle m^3\rangle = \langle m\rangle^3$, thus neglecting fluctuations, the Eqs. (17) and (24) coincide. So, the Galam and Sznajd models are solved in one shot.

It can be easily verified that the general solution of Eq. (23) has the form

$$P_m(m, t) = \frac{1}{(1 - m^2)m} f\left(\mathrm{e}^{-t/2}\frac{m}{\sqrt{1 - m^2}}\right) \qquad (25)$$

for arbitrary function $f(y)$. The latter has to be determined from initial conditions, and assuming that at the beginning the magnetization was m_0 for all realizations of the process, it occurs that $f(m/\sqrt{1 - m^2}) = (1-m^2)m\delta(m-m_0)$. Clearly, in this case the probability density consists of a single δ-function which moves, as time passes, towards one of the ends of the allowed interval for the magnetization, $m \in [-1, 1]$. The dynamics is those of a pure deterministic drift and no diffusion ever smears out the evolving probability packet, as can be seen immediately from the absence of a second-derivative term in Eq. (23). Clearly, this is due to the limit of infinite system size and for finite N there is an additional diffusive term in the Fokker–Planck equation, proportional to N^{-1}.

The deterministic nature of the evolution of opinions has profound consequences. First, the initial sharp δ-function distribution of magnetization remains sharp until the end, thus justifying the neglect of fluctuations and the replacement $\langle m^3\rangle = \langle m\rangle^3$. The fluctuations are set on only at times which diverge when $N \to \infty$ and the equivalence to the Galam model is exact in the thermodynamic limit. In a finite system of size N, the typical time $t_{\rm st}$ to reach the absorbing state, starting from magnetization m_0, can be estimated by requiring that drift brings the magnetization to the distance of order $\sim 1/N$ from either of the extremal points $m = \pm 1$. Expressed in terms of the initial concentration of the $+1$ opinion, $n_+ = (m_0 + 1)/2$, it is

$$t_{\rm st} \simeq \ln\left(\frac{(1-n_+)n_+ N}{(2n_+ - 1)^2}\right) . \tag{26}$$

So, having clarified the close relation to the Galam model, it comes as no surprise that the Eqs. (26) and (18) coincide for large N.

For any positive initial magnetization the final state is always the uniform configuration of all agents in state $+1$, and vice versa. The probability of ending in the all $+1$ state is the step function $P_+ = \theta(n_+ - 1/2)$. This is consistent with simulations of the Sznajd model in dimensions larger than 1 and confirms the existence of a dynamical phase transition at $n_+ = 1/2$. It also shows that in a society where everybody interacts with everybody, but no person changes her opinion unless she meets at least two other people who compel her to do so, the initial majority, however narrow it is, always takes all. No chance is left to minorities. Fortunately enough, reality is more complex. The next section describes model implementations of one common "complication", that the number of choices is not limited to two but can be large, or even very large.

Bounded Confidence

Axelrod Model

Definition One of the features characteristic of the way culture is shared and propagated around the globe is that similar cultures are much more prone to mutual convergence, while incompatible lifestyles often coexist side by side without visibly influencing each other. Robert Axelrod introduced a model [80] nicely describing such a situation.

In the Axelrod model, contrary to the voter or Sznajd model, the character of each of the agents is given by more than one feature. One can think of tastes regarding food, sports, music, etc. These categories represent the features. For each feature the taste can assume various values, e. g., somebody likes eating raw vegetables, spending whole days in the fitness center and listening to Mozart in the evenings, while somebody else feeds on French fries, watches football on TV and adores the pop-star of the season. If two neighbors do not agree on any feature, they are so different that they do not influence each other. Conversely, if they find at least one feature where they share the same preference, one of them may find a second feature in which they do differ and change the preference on that second feature so that it agrees with the preference of the neighboring agent. The fact that the agents do not always interact, but do so only if they have something in common, is called *bounded confidence*.

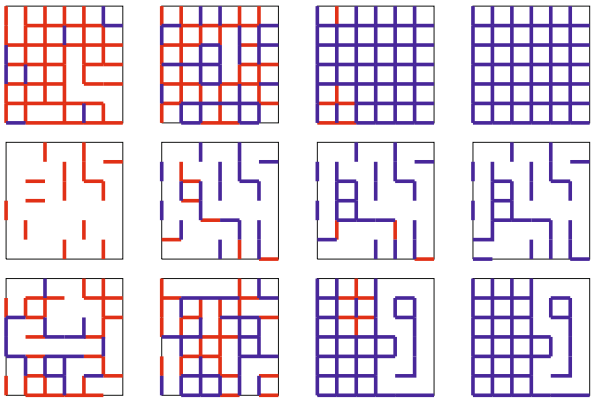
More precisely, there are $N = L^d$ agents placed again on the d-dimensional hypercubic lattice $\Lambda = \{0, 1, \ldots, L-1\}^d$ with periodic boundary conditions and endowed with $F > 1$ integers. The numbers represent the values of its F features, so the state of the agent at site $x \in \Lambda$ is described by the vector with coordinates $\sigma(x, i) \in \mathbb{Z}$, $i = 1, 2, \ldots, F$. The configuration space of the model is $\mathbb{Z}^{\Lambda\times\{1,\ldots,F\}}$ and the evolution of the configuration σ_t is a Markov process determined by transition rates

$$w(\sigma, \sigma^{x,f,a}) = \frac{1}{2d}\sum_{\mu=1}^{2d} \theta\left(A(x, x^\mu)\right) \times \left[F - A(x, x^\mu)\right]^{-1} \delta_{a,\sigma(x^\mu,f)} , \tag{27}$$

where the summation goes over the set of $2d$ neighbors x^μ of the site x. We denoted here $\sigma^{x,f,a}$ the configuration which differs from σ only in feature f, so $\sigma^{x,f,a}(x, f) = a$ and $\sigma(x, f) \neq a$. The function $A(x, y) = \sum_{g=1}^{F} \delta_{\sigma(x,g),\sigma(y,g)}$ counts the number of features on which the agents at positions x and y agree, and $\theta(x)$ is the Heaviside function, $\theta(x) = 1$ for $x > 0$ and zero otherwise. The first factor after the sum in (27) accounts for the condition that the neighbors must agree in at least one feature and the second factor is here due to the fact that in each update the agent chooses randomly among $F - A(\sigma(x), \sigma(x^\mu))$ features in which she disagrees with her neighbor.

Simulations If the two factors after the sum in (27) were absent, the transition rates would depend linearly on the states of the neighbors and a generalization of the linear voter model would emerge, with all its beautiful solubility. However, the non-linearity is there and makes the model non-trivial. Lacking suitable analytical tools, it is necessary to rely on numerical simulations.

To be precise, the rules of the dynamics of the Axelrod model implemented in simulations are as follows: On



Social Processes, Physical Models of, Figure 11
Typical evolution of the Axelrod model on a square lattice with periodic boundary conditions. In each row, snapshots of the configuration in four times $t_1 < t_2 < t_3 < t_4$, from *left* to *right*, where the last configuration is the absorbing state. The active bonds are drawn in *red*, the bonds with full consensus in *blue*, the bonds with absolute disagreement are left *white*. The initial condition is drawn from $\mathbb{Z}_q$ with uniform probability. The parameters are, from *top* to *bottom*, $F = 3$ and $q = 2$; $F = 3$ and $q = 14$; $F = 2$ and $q = 2$. The frame in *light gray* around each configuration is there only for visual convenience

a lattice of size $N = L^d$, time t is discrete and proceeds in chunks of $1/N$, so from time t to $t + 1$ there are as many elementary updates as there are sites. One update step consists of:

1. Choose a site x at random; randomly choose one if its neighbors x^μ. Count the number A of features in which the agent at x agrees with the chosen neighbor.
2. If $A = 0$, do nothing. If $0 < A < F$, randomly choose a feature f among those in which the agent and the neighbor differ. Then set the value of the feature f equal, $\sigma_{t+1/N}(x, f) = \sigma_t(x^\mu, f)$.

One of the principal quantities of interest will be the density of active bonds, which is a generalization of the density of interfaces investigated in the voter model. A bond connecting agents at sites x and y is called active if the agents differ in at least one feature and also agree in at least one feature, i. e., $0 < A(\sigma(x), \sigma(y)) < F$. Their fraction relative to the total number of bonds in the lattice will be denoted $n_A(t)$.

When the number of active bonds drops to zero, after the stopping time τ_{st}, evolution freezes and the system reaches one if its absorbing states. It means that all neighbors are either identical or absolutely incompatible. Clusters of agents which share the same values of all features can be identified, while the borders of such clusters are marked by bonds with no shared value. Even for the simplest case of two features $F = 2$ and the set of allowed values for these features constrained to only two elements, there are infinitely many possible absorbing states characterized by various cluster configurations. (For a finite lattice this number is of course finite, but grows very fast with the system size.) This leads to an important note: Speaking of culture, there might be a good many stable configurations and it is impossible (and improper) to discriminate as to which is the best one.

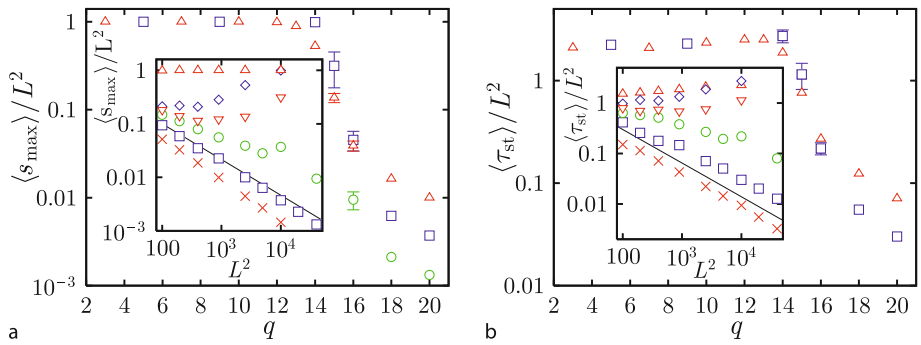
A full characterization of the set of absorbing states would be too difficult. A rough measure is the number of agents in the largest cluster $s_{\max}$, compared to the total number of agents N. When this figure is close to the total system size, it means that the absorbing state is uniform, or very close to it. If the size of the largest cluster is small, or even remains finite when $N \to \infty$, the absorbing state is "multicultural" or fragmented into many isolated islands without mutual interaction, much like the hundreds of villages in Papua-New Guinea, each of them having its own particular language.

The evolution and shape of the eventual absorbing state depends critically on the initial condition. First, note that if, at the beginning, all the values of the features belonged to certain subset $V \subset \mathbb{Z}$ of integers, all features will remain within the set V forever. The most important thing is how large the set V is. But even when V is infinite, the values can be mostly concentrated on a finite subset of V, very rarely going beyond it. A simple but reliable measure of the effective size of the set V, taking into account how often each member is actually present, is the *inverse participation ratio* (IPR). Suppose that, in the initial state, the probability of some feature of a randomly picked agent having the ith value from the set V is p_i. Then, the IPR is

$$q^{-1} = \sum_{i=1}^{|V|} p_i^2 \,. \tag{28}$$

Such a definition is motivated by a simple consideration: if set V consists of the q lowest non-negative integers, i. e., $V = \mathbb{Z}_q \equiv \{0, 1, \ldots, q-1\}$, and all members have the same probability $p_i = 1/q$, the inverse participation ratio is exactly q^{-1}. The integer q is simply the number of elements which are participating, hence the name.

Phase Transition Let us look at how the Axelrod model behaves on a two-dimensional square lattice. Recall that the linear size of the lattice is L, so there are $N = L^2$ agents. The initial condition will consist in choosing the values of all features for all agents independently from the same uniform distribution on $\mathbb{Z}_q$. The typical evolution is shown in



Social Processes, Physical Models of, Figure 12

Phase transition in Axelrod model on a square lattice of size $L \times L$, for $F = 3$, when changing the localization parameter q of the initial state. In a, relative size of the largest cluster, in b, average stopping time relative to the system size are shown. Different symbols correspond to sizes $L = 50$ (△), $L = 100$ (□), and $L = 200$ (◯). In the *insets*, dependence of the corresponding quantities on the system size L^2 are shown. Symbols correspond to parameters $q = 5$ (△), 14 (◇), 15 (▽), 16 (◯), 20 (□), and 30 (×). The *straight line* is the power $\sim L^{-4/3}$

Fig. 11. For larger F and q the absorbing state is very fragmented, but decreasing q it becomes totally uniform and a single cluster covers the entire lattice. On the other hand, for small F the absorbing state contains a few moderately-sized clusters.

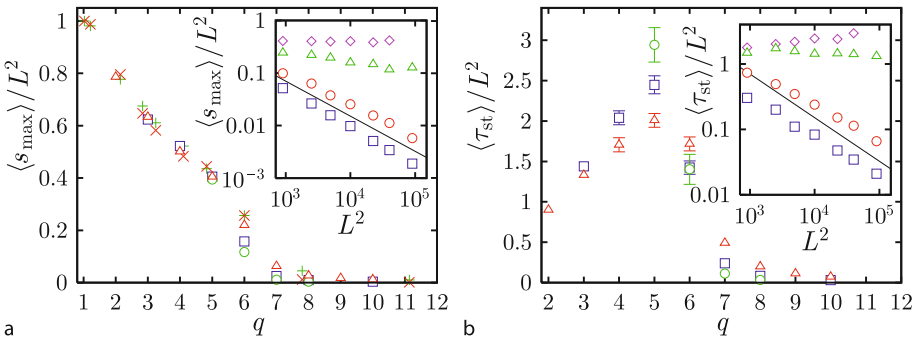
These vague observations are put on a quantitative basis in Figs. 12 and 13. Let us look first at the behavior of the average size of the largest cluster, $\langle s_{\max}\rangle$, when the parameter q is changed. The most important finding is that, at a certain value $q = q_c$, there is a phase transition separating the regime $q < q_c$, in which the maximum cluster makes up a finite fraction of the whole lattice, from the phase with $q > q_c$, where the fraction of sites within the largest cluster goes to zero for $N \to \infty$. The quantity $\langle s_{\max}\rangle/N$ can serve as an order parameter.

Another quantity which illustrates phase transition is the average stopping time, i. e., the time to reach the absorbing state. Generically, it grows with the system size as $\langle \tau_{\rm st}\rangle \sim N^{\eta}(q)$, but below the transition it is proportional to the system size, so $\eta(q) = 1$, while for $q > q_c$ it grows more slowly, with the exponent $\eta(q) \simeq 1/3$. Interestingly, comparison of the stopping time with the size of the largest cluster leads to the conclusion that, for $F = 3$, the two quantities are roughly proportional, $\langle \tau_{\rm st}\rangle \propto \langle s_{\max}\rangle$, both of them exhibiting a jump at the transition, suggesting that the clusters grow linearly in time until they reach the absorbing state. On the other hand, for $F = 2$, such proportionality is not observed. Instead, the average maximum cluster decreases continuously to zero when q_c is approached from below and at the same time the stopping time seems to diverge.

With discrete q, it is impossible to determine the position of the phase transition with high precision, but in practice it is possible to make some reasonable estimates, based on the dependence of $\langle s_{\max}\rangle/N$ on N, as shown in the insets. For $F = 3$ the critical value of the parameter q lies somewhere close to $q_c \simeq 15$, while for $F = 2$ the estimated value is $q_c \simeq 5$.

A more important difference between cases $F = 2$ and $F = 3$ than the numerical value of the transition point q_c is the very nature of the phase transition. It was already shown that the stopping time normalized to the system size blows up at the transition only for $F = 2$, remaining finite for $F = 3$. With fair trustworthiness it is possible to say that for $F = 2$ the transition is continuous, i. e., second-order, while for $F = 3$ the transition is clearly a first-order one. Although it may be dangerous to rely only on numerical data for the order parameter, the conclusion on the nature of the phase transition was also supported by results for the distribution of cluster sizes close to the transition. In Fig. 14a it can be seen that the probability of finding in the absorbing state a cluster of size larger than or equal to s decays as a power, i. e., $P^{>}_{\rm clust}(s) \sim s^{1-\alpha}$, with an obvious cutoff at large s due to the finiteness of the lattice. The value of the exponent α plays critical role. For $F = 2$ the estimate is $\alpha_{F=2} \simeq 1.6$ and for $F = 3$ it is $\alpha_{F=3} \simeq 2.65$. It has been shown that the exponent does not depend on the number of features as long as $F \geq 3$ [81]. How does the distribution of cluster sizes relate to the type of the phase transition? Denoting $N_{\rm clust}(N)$ the total number of clusters and $P_{\rm clust}(s, N) = P^{>}_{\rm clust}(s) - P^{>}_{\rm clust}(s+1)$ the fraction of clusters of size s on the lattice with N sites, the result is

$$N = N_{\rm clust}(N) \sum_{s=1}^{N} s P_{\rm clust}(s, N)\,. \tag{29}$$



Social Processes, Physical Models of, Figure 13

Phase transition in the Axelrod model on a square lattice of size $L \times L$, for $F = 2$, when changing the localization parameter q of the initial state. In a, relative size of the largest cluster, in b, average stopping time relative to the system size. In both panels, results for the uniform initial distribution is plotted for sizes $L = 50$ (△), 100 (□), 200 (◯). In the *insets*, dependence of the corresponding quantities on the system size, for $q = 5$ (◇), 6 (△), 7 (◯), and 8 (□). The *straight lines* are power laws $\sim L^{-4/3}$ In a, there are also data extracted from [81], where the initial condition was assembled from Poisson-distributed integers; sizes were $L = 50$ (+) and $L = 150$ (×)

If $P_{\rm clust}(s, N) \sim s^{-\alpha}$ and $\alpha < 2$, the sum on the right hand side diverges when $N \to \infty$. But the sum is simply the average cluster size, the quantity which plays the role of correlation length. The behavior resembles the percolation transition and the diverging correlation length is an unmistakable characteristic of a second-order phase transition.

On the other hand, if $P_{\rm clust}(s, N) \sim s^{-\alpha}$ with $\alpha > 2$, the sum converges. To keep the equality in (29) for $N \to \infty$, either the number of clusters must be proportional to the system size, or, in addition to the power-law complement, there must be an additional term in the distribution of cluster sizes, accounting for the largest cluster of size $s_{\max} \simeq N$, i. e.,

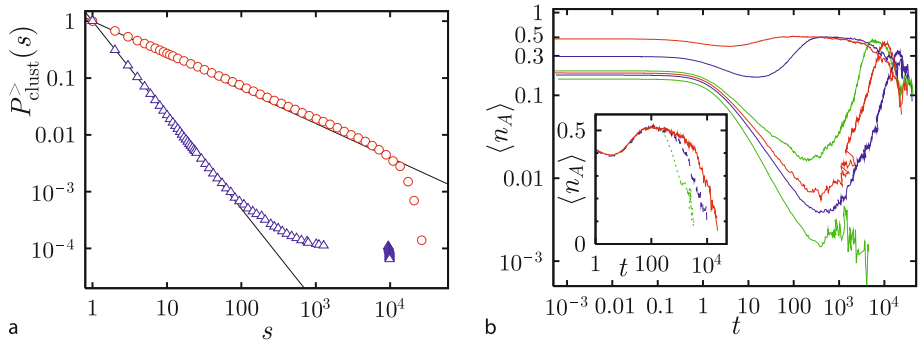
$$P_{\rm clust}(s, N) \simeq a s^{-\alpha} + b\ \delta_{s,s_{\max}}\ . \tag{30}$$

The former possibility occurs for $q > q_c$ and the latter for $q < q_c$. Indeed, the simulation data in Fig. 14a show the presence of a δ-function part in (30), with positive weight $b > 0$. Below the transition, the largest cluster spans essentially whole system, and at q_c its size drops discontinuously to a value negligible with respect to N. This is typical of first-order transitions.

These considerations are also compatible with the behavior of the stopping time as a function of q. For $F = 2$, data indicate that $\langle \tau_{\rm st}\rangle / N$ diverges as $q \to q_c$, while for $F = 3$ it remains constant. Interpreting the stopping time as the correlation time of the dynamics, the model exhibits exactly the same behaviors commonly observed in the second- and first-order phase transitions, respectively.

To conclude, the phase transition in the Axelrod model is driven by the parameter q, measuring the localization of values in the initial condition, and belongs to the class of first-order transitions, if the number of features is at least 3, while for only 2 features the transition is continuous. Interestingly, the same dependence in terms of number of components is known in the Potts model. However, one should bear in mind that the transition in the Axelrod model has a purely dynamical origin. There is no equilibrium besides the absorbing states, so the analogy with phase transitions in equilibrium statistical physics must be taken cautiously and with certain reserve.

Although the most interesting question touches the properties of the absorbing states, it is instructive, also, to see how the system approaches them. In Fig. 14b the evolution of the average number of active bonds is shown. The averages are taken only over realizations which have not yet reached the absorbing state. Around the time $t \simeq 1$, i. e., after as many updates on the computer as there are lattice sites, the density of interfaces decreases drastically. When $q > q_c$ it brings the system quickly to the absorbing state, while in the opposite case, $q < q_c$, the activity rises again and n_A increases to a value close to $n_A \simeq 0.5$. Then, a very slow evolution follows. Most often, the evolution ends due to an occasional hit into an absorbing state, but this is, rather, a finite size effect, as demonstrated in the inset in Fig. 14b. When system size increases, slow evolution is increasingly prolonged. To see the behavior at these very long times is very difficult on a computer, but it seems that the density of active bonds decays very slowly to zero, resembling the logarithmic decay of the density of interfaces in the two-dimensional voter model.



Social Processes, Physical Models of, Figure 14

Axelrod model on a square lattice. In a, distribution of cluster sizes close to the transition $q \simeq q_c$, averaged over 100 realizations. The parameters are $F = 2$, $q = 5$, $L = 200$ (◯), and $F = 3$, $q = 15$, $L = 100$ (△). The *straight lines* are power laws $\sim s^{-0.6}$ and $\sim s^{-1.65}$. In the data for $F = 3$, note the the isolated points around $s \simeq 10^4 = L^2$ indicating that the distribution is composed of a power-law part plus a δ-function located close to L^2. In b, time dependence of the number of active bonds, averaged over realizations of the process. The lattice size is $L = 100$, number of features $F = 3$ and different lines correspond to $q = 5$, 9, 14, 15, 16, and 18, from *top* to *bottom*. In the *inset*, the dynamics is shown for $F = 3$, $q = 5$, and three lattice sizes, $L = 30$ (*dotted line*), 50 (*dashed line*), and 100 (*solid line*)

Bounded Confidence with Continuous Opinions

There may be a different view on consensus formation, assuming that opinions on a certain issue can vary continuously, but a discussion on that subject and possible convergence of the opinions cannot take place unless the actual opinions are close enough. If the people differ too much, they may even avoid mutual contact completely or decide to use "different means", an often-used euphemism for killing the opponent. If, on the contrary, the difference in their opinions, measured by a continuous variable, does not exceed a certain threshold, then the difference may be further diminished by discussions and peaceful persuasion. This idea is called *bounded confidence* and a kind of it is also a key ingredient of the Axelrod model.

Linear Dynamics Modeling consensus formation using continuous variables has been around for quite a long time. Let us denote F_i as the variable describing the opinion of the individual i. The simplest way to implement the convergence of opinions, introduced by DeGroot [82], is to assume that in an upcoming period, opinions are linear combinations of the current opinions of all individuals

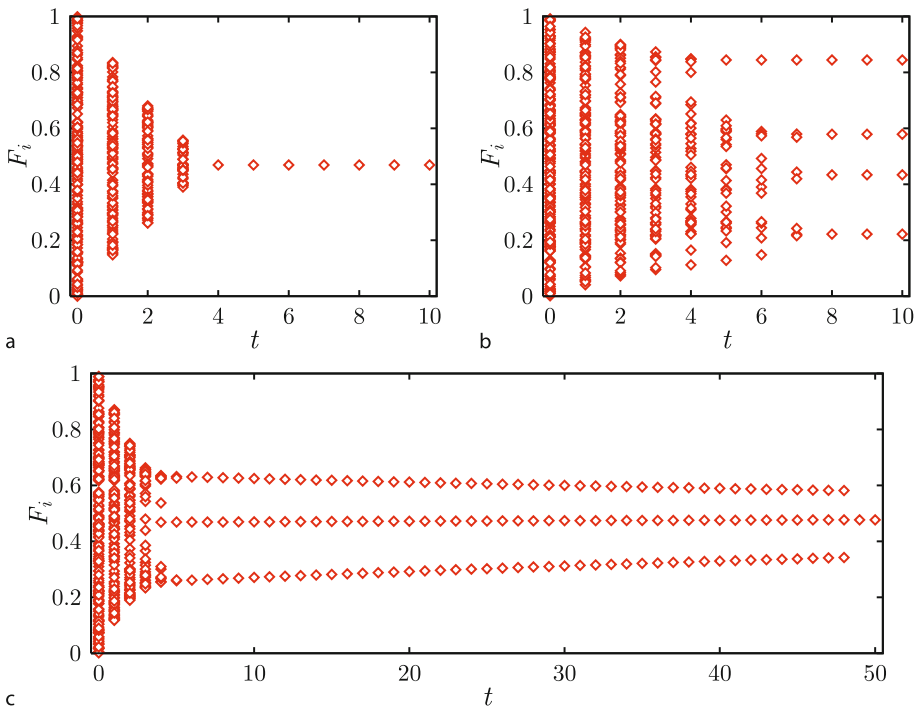
$$F_i(t+1) = \sum A_{ij} F_j(t) \,, \tag{31}$$

where A is a stochastic matrix, i. e., all its elements are non-negative, $A_{ij} \geq 0$, and the sum of all its columns is one, $\sum_i A_{ij} = 1$. This is called a *confidence matrix*. The diagonal elements are assumed to be strictly positive, $A_{ii} > 0$, which means that the individuals have at least some non-vanishing belief in their own opinions. The new values of F_i are weighted averages of previous opinions, the weights being stored in the matrix A. As such, the dynamics is very simple and is equivalent to those of a Markov chain with transition probabilities A_{ij}.

In DeGroot's model, the question of reaching a consensus or not is reduced to the study of connected components of the graph of direct interactions. When there are more of them around, each of them reaches its own consensus, independent of the others. This is quite simple. Introducing bounded confidence adds complications.

Hegselmann and Krause The new opinion of individual i takes into account others' opinions with weights A_{ij}. What if these weights are not constant, but depend on the current opinions themselves? Although some attempts in this direction have been made earlier [83], the principle of bounded confidence was first applied in the models developed by the groups of Deffuant et al. [84] and Krause and Hegselman [85,86]. Let us discuss the latter model first.

The Hegselmann–Krause (HK) model [86] considers opinion formation as a fully deterministic process, as in DeGroot's model investigated in the previous section. The new opinion of an individual is the average of current opinions of those (including the individual of concern) whose opinions lie within a fixed confidence bound $\epsilon > 0$ from the individual's own opinion.



Social Processes, Physical Models of, Figure 15
Evolution of the Hegselmann–Krause model. Each point represents one or more agents with specified value F_i. Total number of agents is $N = 200$. In a, the confidence bound is $\epsilon = 0.3$, while in b it is $\epsilon = 0.1$. c shows how the approach to consensus is slowed down when the confidence bound is close to its critical value. Here $\epsilon = 0.24$

For those pairs which do not differ more than ε in their opinions, the weights A_{ij} are uniform, so

$$A_{ij} = \begin{cases} 0 & \text{for} \quad |F_i - F_j| > \epsilon \\ \frac{1}{N_{i\epsilon}} & \text{for} \quad |F_i - F_j| \le \epsilon \,. \end{cases} \tag{32}$$

Obviously, the normalization constant is the number of individuals within the confidence bound, $N_{i\epsilon} = |\{j\colon |F_i - F_j| \le \epsilon\}|$. Note that the dependence of the confidence matrix on current opinions is very strongly non-linear. From the formal point of view, the most important thing is that the zero-pattern of the confidence matrix can also change in time. It is possible to interpret it as changing the structure of the graph of directly communicating individuals. To see if consensus is going to be reached, it is important to observe whether the graph remains connected during evolution. Clearly, once it splits into disconnected parts, it will never join together again and consensus will not be achieved.

To get an impression of how the evolution of the HK model proceeds, look at Fig. 15. For a large enough confidence bound ε, the system approaches full consensus, while lower values of ε induce several stable communicating subgroups which do not interact with each other, and the system splits into several clusters with different opinions. The number of such clusters grows as $1/\epsilon$ when the confidence bound shrinks. Numerical simulations show that the critical value ϵ_c, at which the full consensus breaks down, approaches the value about $\epsilon_c \simeq 0.2$ when the number of individuals increases. Interestingly, when the confidence bound comes close to the critical value ϵ_c, the number of steps needed to reach consensus increases, suggesting that this is indeed a kind of a dynamical phase transition. This behavior is rather robust, as it persists in various modifications of the HK model, most notably if the individuals can communicate only through a regular lattice or a random network [87,88,89,90].

Deffuant et al. Apart from the initial condition, the model of Hegselmann and Krause is fully deterministic and opinions are updated in parallel. A similar model, but with random sequential update, was introduced by G. Deffuant et al. [84].

There are, again, N individuals with opinions F_i. In each update step, two of them, say, i and j, are chosen randomly. Then, we check to see whether their opinions

differ less than (or equally to) the confidence bound ε. In the positive case, their opinions are slightly shifted towards each other

$$\left.\begin{aligned} &F_i(t+1/N) \\ &\quad = (1-\mu)F_i(t) + \mu F_j(t) \\ &F_j(t+1/N) \\ &\quad = \mu F_i(t) + (1-\mu)F_j(t) \end{aligned}\right\} \text{for} \quad |F_i(t) - F_j(t)| \le \epsilon \tag{33}$$

where μ is a parameter fixing the rate of convergence.

For very large numbers of individuals, $N \to \infty$, the dynamics can be expressed in terms of the continuous distribution of opinions $P(f,t)$, which can be written formally as $\int_0^f P(f',t)\mathrm{d}f' = \lim_{N\to\infty} \frac{1}{N}\sum_j \theta(f - F_j(t))$. The following innocent-looking rate equation is obtained [91]

$$\frac{\partial}{\partial t}P(f,t) = \int\limits_{|f_1-f_2|\le\epsilon} P(f_1,t)P(f_2,t) \times \big[\delta((1-\mu)f_1 + \mu f_2 - f) - \delta(f_1 - f)\big]\mathrm{d}f_1\mathrm{d}f_2 \,, \tag{34}$$

which exhibits fairly complex behavior. Starting from the uniform initial condition $P(f,0) = 1$ for $f \in [0,1]$, the configuration evolves into a stationary state composed of one or several δ-functions, each of them corresponding to one cluster of individuals sharing the same opinion.

To see this in a simple example, consider the case $\epsilon > 1$. All individuals interact with each other, so their opinions converge to a common limit $f = 1/2$. It is possible to see this when we use Eq. (34) to investigate the time evolution of the moments of the distribution. Obviously $\frac{\mathrm{d}}{\mathrm{d}t}\langle f\rangle = 0$, so the average opinion is independent of time, $\langle f\rangle = 1/2$. The dispersion from the mean $\langle f^2\rangle_\mathrm{c} = \langle f^2\rangle - \langle f\rangle^2$ obeys

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle f^2\rangle_\mathrm{c} = -2\mu(1-\mu)\langle f^2\rangle_\mathrm{c} \,, \tag{35}$$

which means that the dispersion decays exponentially to zero with rate $\mu(1-\mu)$. This confirms our intuition, noted earlier, that μ determines the speed of the evolution towards the stationary state.

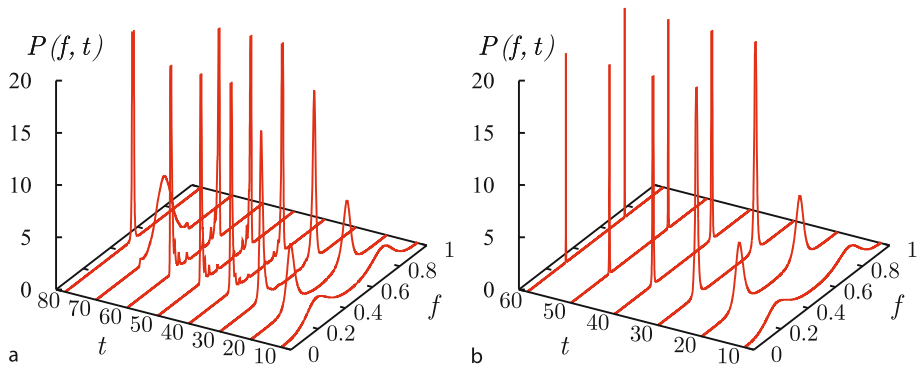
The behavior in the complementary regime of very small ε can be guessed, assuming that $P(f,t)$ does not vary too wildly at scales comparable to ε or shorter. This also means that f is assumed to be farther than ε from the extremal values 0 and 1. In this case, $P(f_{1,2},t)$ is expanded in a Taylor series and the integral on the right hand side of (34) can be performed explicitly. Finally, the following partial differential equation is obtained

$$\frac{\partial}{\partial t}P(f,t) = -\frac{\epsilon^3}{3}\mu(1-\mu)\frac{\partial^2}{\partial x^2}P^2(f,t) \,. \tag{36}$$

There is a trivial homogeneous solution $P(f,t) = C$, independent of f and t, which is, however, unstable at all length scales. To see it, it is enough to linearize (36) close to the uniform solution, $P(f,t) = C + D(f,t)$, where $|D(f,t)| \ll C$, and express the result in terms of the Fourier transform $\widetilde{D}(k,t)$ of the small perturbation. The result is $\frac{\partial}{\partial t}\widetilde{D}(k,t) = \frac{2}{3}k^2\epsilon^3 C\mu(1-\mu)\widetilde{D}(k,t)$, indicating that perturbations increase as their wavelength decreases. On the other hand, if a stationary state is reached, the structures created cannot be closer to each other than ε. This feature is lost in the derivation of Eq. (36), but nevertheless the conclusion is that, for small ε, the stationary state will be composed of regularly spaced δ-functions with period proportional to ε.

To see explicitly how the stationary state is approached for general values of the confidence bound ε, the solution of Eq. (34) should be found numerically. Examples of such evolution are shown in Fig. 16. For large enough ε, all opinions converge to the common value 1/2, as in the case $\epsilon > 1$. However, low ε produces two or more distinct peaks. Detailed analysis [91,92] reveals that a series of bifurcations occurs in the system when the confidence bound is decreased. The consensus breaks into separated groups suddenly at a critical value ϵ_{c1}, which break themselves into more groups at $\epsilon_{c2} < \epsilon_{c1}$ and so forth. The numerically found value of the first critical point is $\epsilon_{c1} = 0.269\ldots$ But how do the peaks emerge from the originally homogeneous distribution? It can be seen in Fig. 16 that they are triggered by the original inhomogeneity at the edges of the interval $[0,1]$. A structure of gradually sharpening peaks propagates towards the middle of the interval, where non-linear waves from opposite endpoints meet and form the resulting pattern.

This observation can be translated into common language by saying that the ultimate fate of the opinions in the society is largely in hands of the extremists. Indeed, individuals close to the edges of the scale of possible opinions are, from the very beginning, drawn towards the middle of the interval $[0,1]$, i. e., close to the centrist position, while those who are farther from the edges than ε keep their opinions unchanged for some finite time. The crucial question is, how far from the extreme will the former extremists move before they settle? It may happen that they will drift so close to the midpoint that the centrists will attract them all; full consensus follows. But the extremists themselves can attract enough centrists and, instead of creating one peak at the center, two (or even more) peaks oc-



Social Processes, Physical Models of, Figure 16
Evolution of the Deffuant model with $\mu = 0.5$. In a, approach to consensus for confidence bound $\epsilon = 0.27$, slightly larger than the critical value ϵ_{c1}. In a, two peaks are formed in the distribution of opinions, for $\epsilon = 0.25$

cur. The resulting society lives with polarized (or diversified) opinions. The parameter which decides the outcome is the confidence bound ε. Higher confidence leads to an overwhelming consensus among people, while lower confidence causes the society split into several non-communicating communities.

Emergence of Social Classes

People are not equal. Not only do they differ in the color of their eyes, in their ability to play chess or to run a marathon race, but even individuals with very similar talents may find themselves very different in social status. All attempts to bring more justice into such evident disequilibrium has ended in desperate or even catastrophic failure. Perhaps the best one can do is to make the membranes separating social levels as permeable as possible, so that no one is a priori disqualified. The ubiquity of social stratification in animal as well as human collectives is certainly a phenomenon which calls for an explanation, and the fact that rabbits, dogs, apes and *Homo sapiens* exhibit similar behavior suggests some common mechanisms which may not be too complicated after all, although they produce highly complex outcomes.

Bonabeau Model

A newcomer in an animal group always has to undergo some fighting before its placement in the social ladder is commonly accepted. If, on the other hand, an individual leaves the group for some prolonged period and returns back, it has to fight again, as the previously established level has faded away. These two observations motivated Bonabeau et al. [93] to introduce a model of self-organized hierarchies [93,94,95,96,97,98].

To be clear, the hierarchies are not understood in the sense of trees with a king or a marshal on the top and lesser ranks below. Instead, it means an ordering, each individual bearing a single number called strength, indicating its position among others. When two agents meet, the stronger one has higher probability to be strengthened, while the weaker is most likely pushed down even more. The strengths of the agents who do not meet at this time relax towards zero by a fixed fraction. More formally, the configuration of the system at time t is described by the collection of strengths $F_i(t)$ of the agents $i = 1, 2, \ldots, N$. In each step, a pair of agents is chosen to fight, i and j, say, resulting in a change of the strengths by ± 1 for the winner and loser, respectively. Moreover, all strengths relax to the reference (zero) level deterministically. Thus

$$\begin{aligned} F_i\left(t+\frac{1}{N}\right) &= \left(1-\frac{\mu}{N}\right) F_i(t) + \Delta_{ij} \\ F_j\left(t+\frac{1}{N}\right) &= \left(1-\frac{\mu}{N}\right) F_j(t) - \Delta_{ij} \\ F_k\left(t+\frac{1}{N}\right) &= \left(1-\frac{\mu}{N}\right) F_k(t)\,, \quad k \neq i\,, \quad k \neq j\,, \end{aligned} \tag{37}$$

with

$$\text{Prob}\left\{\Delta_{ij} = \pm 1\right\} = \frac{1}{1+\mathrm{e}^{\mp\eta(F_i(t)-F_j(t))}}\,. \tag{38}$$

The parameter η tunes the level of randomness in the dynamics, where $\eta \to \infty$ corresponds to purely deterministic outcomes, the stronger agent always beating the weaker, while $\eta = 0$ means that the strengths are increased and decreased by mere chance.

So far, the question of which pairs of agents interact and when, has not been dealt with. The most natural choice is to place the individuals on a network, or simply

on a square lattice, leaving some sites empty, and to allow the agents to diffuse along the edges of the network. When two agents happen to meet at one site, they fight. More than two agents on a site is not allowed; they behave like hard hemispheres, each site being able to accommodate one full sphere. Numerical simulations of such systems [93,94,95,96,97] show that the change in the density of agents, or, equivalently, the relaxation rate μ, induces a phase transition from a uniform state with all agents' strength close to 0 to a hierarchical state where the strengths are highly non-homogeneous. The order parameter is the dispersion of the number of fights won by the agents, i. e.,

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(w_i - \frac{1}{2}\right)^2} ,$$

where $w_i = n_i^+/(n_i^+ + n_i^-)$ and n_i^+, n_i^- are the number of encounters won and lost by the agent i, respectively. In the homogeneous phase $\sigma = 0$ while in the hierarchical phase it has a finite value.

Mean-Field Solution

Analytical study is possible in the mean-field approximation. Indeed, if diffusion is fast enough to ensure many encounters with various agents during the typical time given by the speed of relaxation of the strengths towards zero, the spatial structure of the lattice on which the diffusion takes place becomes irrelevant. In other words, at each time step, two agents are chosen at random and allowed to fight. In this case the dynamics is much simpler, and for a large system the following deterministic evolution equations for the strengths is obtained:

$$\frac{\mathrm{d}F_i}{\mathrm{d}t} = -\mu F_i + \frac{1}{N}\sum_{j=1}^{N}\frac{\sinh\eta(F_i - F_j)}{1+\cosh\eta(F_i - F_j)} . \qquad (39)$$

The stochastic term decreases as $1/N$ for large N, so it is neglected here.

It can be easily seen that the average strength $\overline{F} = \frac{1}{N}\sum_i F_i$ relaxes exponentially to 0 according to $\frac{\mathrm{d}\overline{F}}{\mathrm{d}t} = -\mu\overline{F}$. So, it is sufficient to consider only stationary states with zero mean, $\overline{F} = 0$. The simplest of them, satisfying (39), is the trivial uniform state $F_i = 0$. This may be unstable, though, and its linear stability must be investigated.

Using the notation $\frac{\mathrm{d}F_i}{\mathrm{d}t} = R_i(F_1, \ldots, F_N)$ for the expression occurring in (39), the eigenvalues of the matrix $H_{ij} = \frac{\partial}{\partial F_j}R_i(0, \ldots, 0)$ should be inspected. Thus

$$H_{ij} = \left(-\mu + \frac{\eta}{2}\right)\delta_{ij} - \frac{\eta}{2N} . \qquad (40)$$

One of the eigenvectors is uniform, $x_i = 1$, corresponding to eigenvalue $-\mu$, which is always negative. The remaining $N-1$ eigenvectors have the form $x_i = 1 - N\delta_{ik}$ for some k and they all belong to the same eigenvalue $\frac{\eta}{2} - \mu$. If the latter is positive, the uniform solution of the Eq. (39) is unstable. This happens for

$$\mu \le \mu_c \equiv \frac{\eta}{2} \qquad (41)$$

and the question naturally occurs, what are the stationary configurations beyond the critical value μ_c? Not very far beyond the critical point, linear stability analysis can offer a useful hint, but it needs to be complemented by the influence of the lowest non-linear terms. Therefore, expanding the right-hand side of (39) up to the third order in the differences $F_i - F_j$, one gets for the stationary state

$$(2\mu - \eta)F_i = \frac{\eta^3}{12N}\sum_j (F_i - F_j)^3 . \qquad (42)$$

The solution is assumed in the form

$$F_i = a\delta_{ik} - b \qquad (43)$$

suggested by the eigenvectors of H_{ij} corresponding to unstable modes. Inserting the trial solution (43) into (42) leads to a set of equations for the parameters a and b which can be solved easily

$$a = \begin{cases} 0 & \text{for} \quad \mu > \frac{\eta}{2} \\ \pm\sqrt{\frac{12N(\eta-2\mu)}{(N-2)\eta^3}} & \text{for} \quad \mu < \frac{\eta}{2} \end{cases} \qquad (44)$$

$$(2\mu - \eta)b = -\frac{1}{12N}\eta^3 a^3 . \qquad (45)$$

Hence, the order parameter σ is deduced. Note that in the stationary state, the fraction w_i of fights won by the agent i should be balanced by a relaxation of the strength. Indeed, $2w_i - 1$ is the average increase of the strength of agent i in one step, which should be equal to μF_i. So, $\sigma = \mu\sqrt{\sum_i F_i^2/N}/2$ and inserting the result (45) yields

$$\sigma = \begin{cases} 0 & \text{for} \quad \mu > \frac{\eta}{2} \\ \frac{\mu}{\eta}\sqrt{\frac{3}{N}\left(1 - \frac{2\mu}{\eta}\right)} & \text{for} \quad \mu < \frac{\eta}{2} , \\ & \mu \to \frac{\eta}{2} \quad \text{and} \quad N \gg 1 . \end{cases} \qquad (46)$$

How can this solution be interpreted? One may compare it to a situation in a society with one master and many servants, a single emperor while all the rest are servants, equal to each other in their subordination. This looks nice, but it follows from the formulae (45) that this "realistic" solution

always comes together with a mirror image of itself, a society of a single servant subject to many equal masters. This is an absurd situation, revealing something artificial in the Bonabeau model itself. Indeed, the dynamics is invariant with respect to inversion of all strengths, $F_i \to -F_i$, because no a priori advantage of being stronger was explicitly introduced. The only thing which was assumed was that the strong grow stronger and the weak grow weaker.

Now let us turn back to mathematical aspects of the model. It can be shown that a solution of type (43) exists for any value of μ and η, with only the complication that the equations for a and b become transcendental, thus analytically insoluble. This would be too simple, though. In fact, it is easy to find a stable solution in the limit $\eta \to \infty$, with μ kept finite, and see that it differs completely from (43). The expression within the sum in (39) becomes the sign function of the difference $F_i - F_j$ and the set of equations for the stationary state becomes

$$\mu F_i = \frac{1}{N} \operatorname{sign}(F_i - F_j) \tag{47}$$

and reordering the agents so that their strengths make an increasing sequence $F_1 < F_2 < \ldots < F_N$ one gets

$$F_i = \frac{1}{\mu N}(2i - N - 1)\,, \tag{48}$$

a society organized as a regular ladder of ranks. The order parameter corresponding to such a state,

$$\sigma = \frac{1}{3}\left(1 - \frac{1}{N^2}\right) \tag{49}$$

approaches a non-zero limit for $N \to \infty$, contrary to the configuration of (43) which is appropriate only in close proximity to the critical point, and whose order parameter decreases as $N^{-1/2}$, as can be seen from (44).

Despite the obvious criticism that the essential mirror symmetry $F_i \to -F_i$ is unrealistic and leads to unacceptable solutions, the Bonabeau model manifests quite well the basic idea of how the various social classes emerge. Within this framework, hierarchies are due to dynamical instabilities of the uniform state. The same general mechanism is also (at least partially) responsible for the stratification of human society. Any departure from the mean, be it positive or negative, is amplified. Any accidental misfortune sends one almost invariably even deeper. That is the whole mystery.

Future Directions

The prospects of physical modeling of social behavior are larger than what is presented here. The study of the emergence of cooperation within the framework of spatial and repeated prisoner's dilemma game is one of the classics. It may be, and has been, generalized in various ways. For example, one may wonder what happens if the agents could move through the lattice, or along the edges of the general graph representing the structure of the society. Indeed, people do not sit on the same lattice point forever. It turns out that, depending on the average velocity of the agents' movement, there is a phase transition. Less mobile agents form a phase with a macroscopic fraction of collaborators, while beyond a certain level of mobility the collaborators vanish.

Bounded confidence models pose some subtle and still unsolved problems. One of them concerns how the final state depends on the number of agents, or, stated differently, on the granularity of the distribution of opinions on the opinion axis. If the number of agents were arbitrarily large, the density could be arbitrarily small and it could happen that between any two agents there might be other agents which may mediate ultimate consensus. If that were true, the splitting of society into disparate groups would be the effect of limited size. The conclusion sounds somewhat paradoxical: small societies are more likely to be divided into non-communicating sub-cultures, while large societies turn to be more homogeneous. Another question naturally follows: how small must a social group be to allow disparate opinions to coexist? Implications for multicultural studies are evident.

The dynamics of social strata are relatively less explored and the state of the problem presented here must be considered as preliminary. The societies sketched in these models do not possess any internal structure, which would either stabilize or slow-down the emergence of stratification. Surely there is much space for future work here.

Some important areas were not covered at all, for example the whole family of physically based models of the birth and competition of languages [41,99]. Some social phenomena also closely touch economics, but these are covered by other articles in this Encyclopedia.

Acknowledgments

The original results in this article were obtained within the projects AVOZ10100520 and MSM0021620845.

Bibliography

Primary Literature

1. Comte A (1822) Plan des travaux scientifiques nécessaires pour réorganiser la société
2. Comte A (1839) Cours de philosophie positive, tome IV, 46e leçon. Bachelier Paris

3. Chigier NA, Stern EA (eds) (1975) Collective phenomena and the applicatios of physics to other fields of science. Brain Research Publications, Fayetteville
4. Callen E, Shapero D (1974) A theory of social imitation. Phys Today 27(7):23–28
5. Weidlich W (1991) Physics and social science – The approach of synergetics. Phys Rep 204:1–163
6. Anderson PW, Arrow KJ, Pines D (1988) The economy as an evolving complex system. Addison Wesley, Reading
7. Galam S (2004) Sociophysics: A personal testimony. Physica A 336:49–55
8. Galam S, Gefen Y, Shapir Y (1982) Sociophysics: A new approach of sociological collective behaviours. I. Mean-behaviour description of a strike. J Math Sociol 9:1–13
9. Galam S, Moscovici S (1991) Towards a theory of collective phenomena: Consensus and attitude changes in groups. Eur J Soc Psychol 21:49–74
10. von Neumann J, Morgenstern O (1944) Theory of games and economic behavior. Princeton University Press, Princeton
11. Vega-Redondo F (1996) Evolution, games, and economic behaviour. Oxford University Press, Oxford
12. Nash JF (1950) Equilibrium points in n-person games. Proc Natl Acad Sci USA 36:48–49
13. Nash JF (1950) The bargaining problem. Econometrica 18:155–162
14. Marsili M, Zhang YC (1997) Fluctuations around Nash equilibria in game theory. Physica A 245:181–188
15. Axelrod R (1980) Effective choice in the prisoner's dilemma. J Confl Resolut 24:3–25
16. Axelrod R (1980) More effective choice in the prisoner's dilemma. J Confl Resolut 24:379–403
17. Axelrod R, Hamilton WD (1981) The evolution of cooperation. Science 211:1390–1396
18. Lindgren K (1991) Evolutionary phenomena in simple dynamics. In: Langton CG, Taylor C, Farmer JD, Rasmussen S (eds) Artificial Life II. Addison-Wesley, Readwood City, pp 295–312
19. Lindgren K (1997) Evolutionary dynamics in game-theoretic models. In: Arthur WB, Durlauf SN, Lane DA (eds) The economy as an evolving complex system II. Perseus, Reading, pp 337–367
20. Nowak MA, May M (1992) Evolutionary games and spatial chaos. Nature 359:826–829
21. Schweitzer F, Behera L, Muhlenbein H (2002) Evolution of cooperation in a spatial prisoner's dilemma. Adv Compl Syst 5:269–299
22. Szabó G, Töke C (1998) Evolutionary prisoner's dilemma game on a square lattice. Phys Rev E 58:69–73
23. Chiappin JRN, de Oliveira MJ (1999) Emergence of cooperation among interacting individuals. Phys Rev E 59:6419–6421
24. Lim YF, Chen K, Jayaprakash C (2002) Scale-invariant behavior in a spatial game of prisoner's dilemma. Phys Rev E 65:026134
25. Szabó G, Vukov J, Szolnoki A (2005) Phase diagrams for an evolutionary prisoner's dilemma game on two-dimensional lattices. Phys Rev E 72:047107
26. Abramson G, Kuperman M (2001) Social games in a social network. Phys Rev E 63:030901(R)
27. Ebel H, Bornholdt S (2002) Evolutionary games and the emergence of complex networks. arXiv:cond-mat/0211666 (Preprint)
28. Ebel H, Bornholdt S (2002) Coevolutionary games on networks. Phys Rev E 66:056118
29. Zimmermann MG, Eguíluz VM (2005) Cooperation, social networks, and the emergence of leadership in a prisoner's dilemma with adaptive local interactions. Phys Rev E 72:056118
30. Vukov J, Szabó G, Szolnoki A (2006) Cooperation in noisy case: Prisoner's dilemma game on two types of regular random graphs. cond-mat/0603419
31. Clifford P, Sudbury A (1973) A model for spatial conflict. Biometrika 60:581–588
32. Holley RA, Liggett TM (1975) Ergodic theorems for weakly interacting infinite systems and the voter model. Ann Prob 3:643–663
33. Liggett TM (1985) Interacting particle systems. Springer, Berlin
34. Redner S (2001) A guide to first-passage processes. Cambridge University Press, Cambridge
35. Scheucher M, Spohn H (1988) A soluble kinetic model for spinodal decomposition. J Stat Phys 53:279–294
36. Krapivsky PL (1992) Kinetics of monomer-monomer surface catalytic reactions. Phys Rev A 45:1067–1072
37. Frachebourg L, Krapivsky PL (1996) Exact results for kinetics of catalytic reactions. Phys Rev E 53:R3009–R3012
38. Ben-Naim E, Frachebourg L, Krapivsky PL (1996) Coarsening and persistence in the voter model. Phys Rev E 53:3078–3087
39. Dornic I, Chaté H, Chave J, Hinrichsen H (2001) Critical coarsening without surface tension: The universality class of the voter model. Phys Rev Lett 87:045701
40. Al Hammal O, Chaté H, Dornic I, Muñoz MA (2005) Langevin description of critical phenomena with two symmetric absorbing states. Phys Rev Lett 94:230601
41. Castellano C, Fortunato S, Loreto V (2007) Statistical physics of social dynamics. arXiv:0710 3256
42. ben-Avraham D, Considine D, Meakin P, Redner S, Takayasu H (1990) Saturation transition in a monomer-monomer model of heterogeneous catalysis. J Phys A: Math Gen 23:4297–4312
43. Liggett TM (1999) Stochastic interacting systems: Contact, voter, and exclusion processes. Springer, Berlin
44. Derrida B, Hakim V, Pasquier V (1996) Exact exponent for the number of persistent spins in the zero-temperature dynamics of the one-dimensional Potts model. J Stat Phys 85:763–797
45. Gradshteyn IS, Ryzhik IM (1994) Table of integrals, series, and products, 5th edn. Academic Press, San Diego
46. Cox JT (1989) Coalescing random walks and voter model consensus times on the torus in $\mathbb{Z}^d$. Ann Prob 17:1333–1366
47. Galam S (1986) Majority rule, hierarchical structures, and democratic totalitarianism: A statistical approach. J Math Psychol 30:426–434
48. Galam S (1990) Social paradoxes of majority rule voting and renormalization group. J Stat Phys 61:943–951
49. Galam S (1999) Application of statistical physics to politics. Physica A 274:132–139
50. Galam S (2000) Real space renormalization group and totalitarian paradox of majority rule voting. Physica A 285:66–76
51. Galam S, Wonczak S (2000) Dictatorship from majority rule voting. Eur Phys J B 18:183–186
52. Schneier B (1996) Applied cryptography, 2nd edn. Wiley, New York

53. Krapivsky PL, Redner S (2003) Dynamics of majority rule in two-state interacting spin systems. Phys Rev Lett 90:238701
54. Slanina F, Lavička H (2003) Analytical results for the Sznajd model of opinion formation. Eur Phys J B 35:279–288
55. Plischke M, Bergersen B (1994) Equilibrium statistical physics. World Scientific, Singapore
56. Galam S (2004) Contrarian deterministic effects on opinion dynamics: The hung elections scenario. Physica A 333:453–460
57. Stauffer D, Sá Martins JS (2004) Simulation of Galam's contrarian opinions on percolative lattices. Physica A 334:558–565
58. Florian R, Galam S (2000) Optimizing conflicts in the formation of strategic alliances. Eur Phys J B 16:189–194
59. Galam S (2002) Minority opinion spreading in random geometry. Eur Phys J B 25:403–406
60. Galam S (2002) The September 11 attack: A percolation of individual passive support. Eur Phys J B 26:269–272
61. Galam S (2003) Modelling rumors: The no plane Pentagon french hoax case. Physica A 320:571–580
62. Galam S (2003) Global physics: From percolation to terrorism, guerilla warfare and clandestine activities. Physica A 330:139–149
63. Galam S, Mauger A (2003) On reducing terrorism power: A hint from physics. Physica A 323:695–704
64. Galam S, Vignes A (2005) Fashion, novelty and optimality: An application from Physics. Physica A 351:605–619
65. Galam S (2004) The dynamics of minority opinions in democratic debate. Physica A 336:56–62
66. Galam S (2004) Unifying local dynamics in two-state spin systems. cond-mat/0409484
67. Galam S (2005) Local dynamics vs. social mechanisms: A unifying frame. Europhys Lett 70:705–711
68. Gekle S, Peliti L, Galam S (2005) Opinion dynamics in a three-choice system. Eur Phys J B 45:569–575
69. Galam S, Chopard B, Masselot A, Droz M (1998) Competing species dynamics: Qualitative advantage versus geography. Eur Phys J B 4:529–531
70. Tessone CJ, Toral R, Amengual P, Wio HS, San Miguel M (2004) Neighborhood models of minority opinion spreading. Eur Phys J B 39:535–544
71. Galam S (2005) Heterogeneous beliefs, segregation, and extremism in the making of public opinions. Phys Rev E 71:046123
72. Sousa AO, Malarz K, Galam S (2005) Reshuffling spins with short range interactions: When sociophysics produces physical results. Int J Mod Phys C 16:1507–1517
73. Sznajd-Weron K, Sznajd J (2000) Opinion evolution in closed community. Int J Mod Phys C 11:1157–1165
74. Behera L, Schweitzer F (2003) On spatial consensus formation: Is the Sznajd model different from a voter model? cond-mat/0306576
75. Krupa S, Sznajd-Weron K (2005) Relaxation under outflow dynamics with random sequential updating. Int J Mod Phys C 16:177–1783
76. Stauffer D, de Oliveira PMC (2002) Simulation of never changed opinions in Sznajd consensus model using multi-spin coding. cond-mat/0208296
77. Stauffer D, de Oliveira PMC (2002) Persistence of opinion in the Sznajd consensus model: Computer simulation. Eur Phys J B 30:587–592
78. Stauffer D, Sousa AO, Moss de Oliveira S (2000) Generalization to square lattice of Sznajd sociophysics model. Int J Mod Phys C 11:1239–1245
79. Bernardes AT, Costa UMS, Araujo AD, Stauffer D (2001) Damage spreading, coarsening dynamics and distribution of political votes in Sznajd model on square lattice. Int J Mod Phys C 12:159–167
80. Axelrod R (1997) The dissemination of culture: A model with local convergence and global polarization. J Confl Resolut 41:203–226
81. Castellano C, Marsili M, Vespignani A (2000) Nonequilibrium phase transition in a model for social influence. Phys Rev Lett 85:3536–3539
82. DeGroot MH (1974) Reaching a consensus. J Am Stat Assoc 69:118–121
83. Chatterjee S, Seneta E (1977) Toward consensus: Some convergence theorems on repeated averaging. J Appl Prob 14:89–97
84. Deffuant G, Neau D, Amblard F, Weisbuch G (2000) Mixing beliefs among interacting agents. Adv Compl Syst 3:87–98
85. Krause U (2000) A discrete nonlinear and non-autonomous model of consensus formation. In: Elaydi S, Ladas G, Popenda J, Rakowski J (eds) Communications in difference equations. Gordon and Breach, Amsterdam, pp 227–236
86. Hegselmann R, Krause U (2002) Opinion dynamics and bounded confidence models, analysis and simulation. J Artif Soc Soc Simul 5: http://jasss.soc.surrey.ac.uk/5/3/2.html
87. Fortunato S (2004) Damage spreading and opinion dynamics on scale free networks. cond-mat/0405083
88. Fortunato S (2004) The Krause–Hegselmann consensus model with discrete opinions. Int J Mod Phys C 15:1021–1029
89. Fortunato S (2005) On the consensus threshold for the opinion dynamics of Krause–Hegselmann. Int J Mod Phys C 16:259–270
90. Pluchino A, Latora V, Rapisarda A (2005) Compromise and synchronization in opinion dynamics. physics/0510141
91. Ben-Naim E, Krapivsky PL, Redner S (2003) Bifurcations and patterns in compromise processes. Physica D 183:190
92. Weisbuch G, Deffuant G, Amblard F, Nadal JP (2001) Interacting agents and continuous opinions dynamics. cond-mat/0111494
93. Bonabeau E, Theraulaz G, Deneubourg JL (1995) Phase diagram of a model of self-organizing hierarchies. Physica A 217:373–392
94. Sousa AO, Stauffer D (2000) Reivestigation of self-organizing social hierarchies. Int J Mod Phys C 11:1063–1066
95. Stauffer D, Sá Martins JS (2003) Asymmetry in hierarchy model of Bonabeau et al. cond-mat/0308437
96. Schulze C, Stauffer D (2004) Phase diagram in Bonabeau social hierarchy model with individually different abilities. cond-mat/0405697
97. Malarz K, Stauffer D, Kułakowski K (2005) Bonabeau model on a fully connected graph. physics/0502118
98. Lacasa L, Luque B (2005) Bonabeau hierarchy models revisited. physics/0511105
99. Schulze C, Stauffer D, Wichmann S (2008) Birth, survival and death of languages by Monte Carlo simulation. Commun Comput Phys 3:271–294

Books and Reviews

Castellano C, Fortunato S, Loreto V (2007) Statistical physics of social dynamics. arXiv:0710 3256

Schweitzer F (ed) (2002) Modeling complexity in economic and social systems. World Scientific, Singapore

Weidlich W (1991) Physics and social science – The approach of synergetics. Phys Rep 204:1–163

Social Processes, Simulation Models of

KLAUS G. TROITZSCH
Universität Koblenz-Landau, Koblenz, Germany

Article Outline

Glossary
Definition of the Subject
Introduction
Social Processes and Social Systems
System Dynamics
Microanalytical Simulation
Multilevel Simulation
Multi-Agent Simulation
Tools
Future Directions
Bibliography

Glossary

Social process A social process in the current context can be defined as any series of events occurring where several human beings enter into interaction. The definition may be extended to processes in which social animals are involved.

Model A model in the current context is defined as the representation of some part of the real world, which is called a target system, where this representation is done in terms of some language, be it natural language or some formal graphic and/or computer-executable language. It represents real-world entities with nouns in natural language or objects in typical programming languages, properties of these entities with adjectives in natural language or instance variables in typical programming language, actions of these entities with verbs in natural language or methods and procedures in programming languages, and relations between such entities with relations between nouns or objects in the respective languages.

Simulation Simulation is the execution of a formal model of some process described in a programming language on a computer.

Agent An agent in the current context is usually defined as a piece of software that has some autonomy, i. e. operates without other parts of a larger computer program having direct control of its internal state and actions, that can communicate with other agents in some kind of language, i. e. without direct access to other agents' inner state, that can perceive its environment and can react on it, and that can also take the initiative, i. e. can act without a stimulus from outside and display goal-directed behavior.

System A system in the current context is a mathematical structure consisting of a set of components (composition) belonging to the system, another set of things of the same kind not belonging to the system (forming the environment of the system) and a set of relations defined on the set of things forming both composition and environment and containing at least one relation that changes the history of the involved things [6]. In other contexts, "system" may have other connotations.

Definition of the Subject

The simulation of social processes has been a topic of both the social sciences and computer science for more than five decades and goes back to the first applications of computers. This is, among others, due to the fact that the first developers of computers were not only mathematicians and people from electric engineering, but also economists and even political scientists. As one can read in biographies of Herbert A. Simon, a political scientist, [61] and John von Neumann, a mathematician and game theorist, [49] persons like these were fascinated by the idea that beyond mathematical analysis computer applications could be made fruitful for the understanding of social processes.

Whereas cross-sectional analysis in the social sciences had been supported by statistical methods even earlier, any longitudinal analysis of societies needed more advanced methods of analysis than statistics can provide.

Simulation of social processes comes in different forms, as this entry will deal with in much more detail. One usually distinguishes between pure macro-level simulations where one is only interested in what happens on one level, usually a nation, a society, or an organization and where only one entity – the whole system – is represented, from different kinds of multi-level simulations where at least a system and its components or elements are represented together with the interactions between levels and interactions among the components.

Simulation of social processes has several distinct aims. On one hand, and this is usually the primary, but most primitive aim, one is interested in the prediction of processes (this is often done before the process is properly understood and often leads to predictions that can only be invalidated by empirical research performed after the prediction). On the other hand, one is first interested to understand what is going on in a social process (i. e. to communicate to oneself what one believes are the ingredients of a social process), to make this communicable to others in order to be able to discuss one's understanding with colleagues, and to experiment with the model in order to find out whether the executable simulation model behaves like the target system it represents. Simulation thus plays a similar role as earlier kinds of models, e. g. a mathematical model in physics, which were qualified by Max Planck [48] as "a system of terms and sentences, the so-called physicist's world view, which he endows to the best of his knowledge in order that it – put in place of the real world – should send him – if ever possible – the same message as the real world would send". And a computer program written for simulation is just "a system of terms and sentences" that – even literally – is able to send messages to its user.

Introduction

This entry is organized as follows. The next section tries to give an overview of social processes that have so far been modelled for the purpose of simulation. As one will see, there is a very wide range of social processes that have been simulated with quite different simulation tools and approaches. And this section will, of course, give a more detailed definition what a social process is, and it will shortly deal with some system theories (unfortunately there are very different ones under the same name) and analyze how their adherents have used simulation to analyze what they described as systems.

Section "Microanalytical Simulation" and Sect. "Multi-level Simulation" will deal with two classical approaches to social simulation. The two classical approaches are those of system dynamics and of microanalytical simulation. Both of them have a history of more than 50 years (their learned societies celebrated their 50th anniversaries in August 2007) and have found a wide audience, with system dynamics still being an approach used from ecology to management sciences, modelling ecosystems as well as the planet as a whole, but also firms and markets, and with microanalytical simulation still being an approach used to predict consequences of tax and transfer reforms and the age structure of future generations. While system dynamics is a purely one-level, one-of-a-kind approach as it always represents exactly one undivided real-world system in terms of its properties, thus hiding its components, microanalytical simulation represents a large number of non-interacting men and women and/or households that react on the stimuli given by the experimenter who is only interested in the aggregate statistics which in turn never respond to the behavior of the micro-entities. Thus, in both of these classical approaches an important feature of social processes is in a way missing, as both of them do not actually represent interactions between human beings and, moreover, do not represent interactions between individual and group or individual and society.

Section "Multi-Agent Simulation" will deal with several contemporary approaches that have not yet made their way into the ministries or the boards of enterprises (with a few important exceptions that will be discussed). Common to all of these disparate approaches is the attempt at representing processes in which interactions among individuals and between levels play a major role. Interactions can occur in topological neighborhoods (as for instance in cellular automata) or in networks (as in many agent-based or multi-agent simulations), sometimes neither neighborhood nor network topology is represented, and models restrict themselves to the micro-macro interaction. Approaches of this type most often try to represent cognitive processes beside social processes, as at least in social systems of human beings much of the interaction is controlled by cognitive processes in the minds of participants. And it is still an open question whether a valid representation of human social system can do without taking account of cognitive processes, as it is an open question whether taking only cognitive processes into account leads to an explanation of social processes ("mind is not enough" [8], but "society is not enough" seems also to be true).

Social Processes and Social Systems

Talking about social processes makes it necessary first to define what a process is. Following Bunge [5] a process is a sequence of states a thing had or has or will have over time. It can be written as a time series whose elements denote the state of a thing at a certain point of time, the sequence of states being ordered with respect to time. Thus a process can be represented with a sequence of vectors, each vector representing the state of the thing at a certain time, and each element of each vector representing one of the properties of the thing that describe its state at this time.

When talking about social processes, what was called "thing" in the preceding paragraph (and thus following Bunge [5]) will always be a system composed of human beings (and perhaps containing several levels of subsystems) where the state of the system can and should be described both on the system level and on the component level (and perhaps on the levels defined by different subsystems). Although system dynamics describes the state of the system only on this upper level and neglects any subsystems and components, and although microanalytic simulation describes the state of the system only on the level of either households or individual (but usually not on both) and not on the aggregate level (this is done only for calculating results), both of these approaches describe social processes but in an incomplete manner.

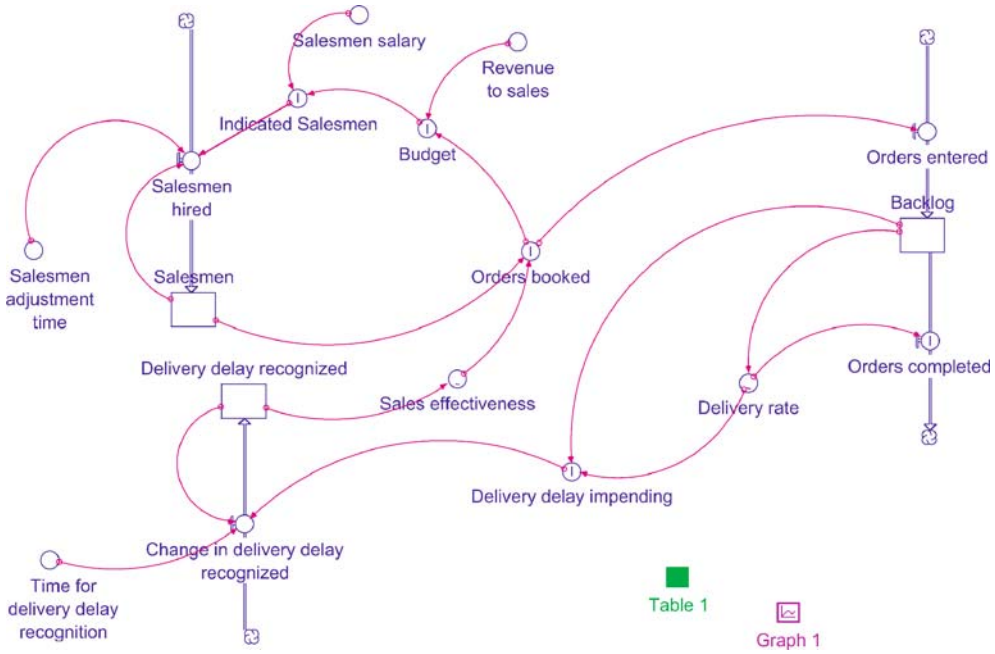
But it will not be sufficient just to describe a social process (as if videotaping it and looking at the video frame after frame), instead it is necessary for understanding and explaining why the process developed the way it developed to construct a model as a means that can create the next frame from the former frames, thus (re-) constructing the laws the process obeys. For simple things such as those elementary physics deals with this law was often represented with a differential equation, one of the classical means of modelling processes. For social systems, this representation is only rarely if ever appropriate, as the calculus does not easily lend itself to describe a population and its members at the same time. This is why system dynamics sets the members of a population (members of an organization, citizens of a country) aside and concentrates on the relatively few properties of the system as a whole, which is also done in some of the classic system theories – for Rapoport (pp. 8–9 in [52]), for instance, "the simplest mathematical description of a system – a model of a spring balance" is the "equation $L = L_0 + kW$ where L is the length of the stretched spring, W the suspended weight, L_0 the 'natural' length of the spring (when $W = 0$), and k a constant that represents the 'stretchability' of the spring." This is exactly the same view on a system as the one taken by system dynamics, where only the variable and constants attributes of the system seem to be interesting, but not the interactions between the parts of a system.

At least from the time of Coleman, but certainly even earlier, e. g. in Durckheim [14], social processes have been superficially represented with the so-called "Coleman boat" or "Coleman bath-tub", a figure [9] that mirrors the notion that processes on the level of, say, society, can be observed on this level, but it is clear that it is not society that "acts", but whatever happens on the level of society is performed by the actions of individual human beings that are in a way under the control of the society and at the same time making up society. In Coleman's example, there is an observable process in which protestant religious doctrine seems to have led to capitalism. This "single proposition breaks into three: one with an independent variable characterizing the society and a dependent variable characterizing the individual; a second with both independent and dependent variable characterizing the individual; and a third with the independent variable characterizing the individual and the dependent variable characterizing the society. Thus the proposition begins and ends in the macro level, but in between it dips to the level of the individual" [8]. The original macro proposition is usually represented with the deck of a boat, while the first of the three new proposition forms its stern, the second is the keel and the third is the bow of the boat. Thus, protestant religious doctrine introduced in a population (wherever this doctrine might originate) influences the value system of individuals, and the changed value system modifies their economic behavior; finally this economic behavior of the individuals modifies the characteristics of the population as a whole (and this latter process is often called "emergence").

The first of these inter-level processes is meanwhile known as "immergence", a term coined by Castelfranchi [7] for the process by which – in Colemans's example – an individual becomes aware of a doctrine ("cognitive emergence") and decides to act according to this doctrine; this process was already described by Durckheim [14] when he talks about "sociological phenomena [that] penetrate into us by force or at the very least by bearing down more or less heavily upon us". The other process, by which individuals change the state of the system (population, organization, …) qualitatively is often called "emergence".

System Dynamics

As superficially mentioned in earlier sections, system dynamics is a tool for modelling processes in social systems on a macro level. The approach originated in the 1950s, when Jay W. Forrester made his first steps in representing social processes in terms of differential equations (as others had done before him, e. g. Herbert A. Simon [59,60]) and found it necessary to include non-differentiable functions into his models. This was how the programming language DYNAMO was first invented, meanwhile superseded by more modern attempts at designing modelling languages with graphical user interfaces, a functional, i. e. declarative programming language, in which the programmer does not (have to) describe how the computer should perform its calculations, but only



Social Processes, Simulation Models of, Figure 1
Graphical representation of Forrester's sales growth model, using the STELLA simulation tool

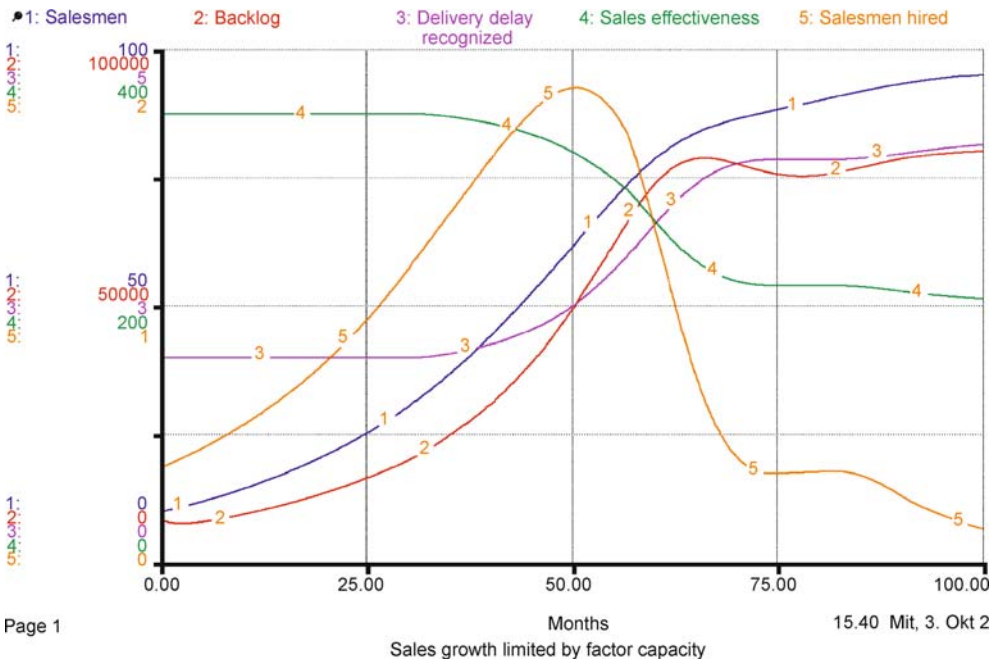
which invariants should hold over time. An invariant in this sense is an equation that holds true even when the states of a thing or system change (for instance, even if the size of population is changing over time, its growth rate could be constant for a long period, such that for all times the equation $\text{size}(t+1) = \text{size}(t) + \text{growthrate} * \text{size}(t)$ holds). System dynamics is of course much more sophisticated, but the equation in the preceding sentence is a typical representative of equations in system dynamics models. The first step in modelling is usually a graphical model showing the dependencies and feedbacks between the attributes of the modelled macro system. These attributes come in several different kinds, the two main kinds of variables being levels and rates. Levels are state variables, usually describing something like material stocks (inventory of a warehouse, size of a population etc.), whereas rates are state change variables, usually describing the material flow into a stock and from a stock per time unit. Beside these, auxiliary variables serve as abbreviations, and constants can be defined.

Among the target systems modelled with the help of system dynamics and the respective tools, there are agglomerations [18], industrial enterprises [16], ecosystems [2] as well as the world as a whole [19,42,43].

A relatively simple example of system dynamics modelling can be found in [17] where a firm with its "sales growth and saturation" is modelled and simulated. The only properties of this firm considered are the current salesperson staff, the backlog of unprocessed orders and the delivery delay recognized. These "level variables" are connected to each other via several "rate variables" – salespersons hired, change in delivery delay recognized, orders entered and orders completed – and several "auxiliary variables" – sales effectiveness, orders booked, budget, indicated salespersons, delivery rate, delivery delay impending – and a few constants – salespersons adjustment time, revenue to sales and time for delivery delay recognition. A STELLA representation of the model makes the idea behind the model sufficiently clear (see Fig. 1).

The simulation of this model shows the time series in several variables of this fictitious firm: The model starts with 10 salespersons and a high budget (stemming from a large number of orders), thus more salespersons are hired who bring more and more orders, more than can be processed such that the backlog quickly increases and the sales effectiveness decreases after some time. This is the signal for hiring less new salespersons, but the delivery delay cannot be decreased in the long run, as a still rising number of salespersons continue to bring more orders.

The graph in Fig. 2 shows some of the drawbacks of the approach: at any time (except at the start) the number of salespersons is a non-integer number, and at any time the number of newly hired persons is also a non-integer



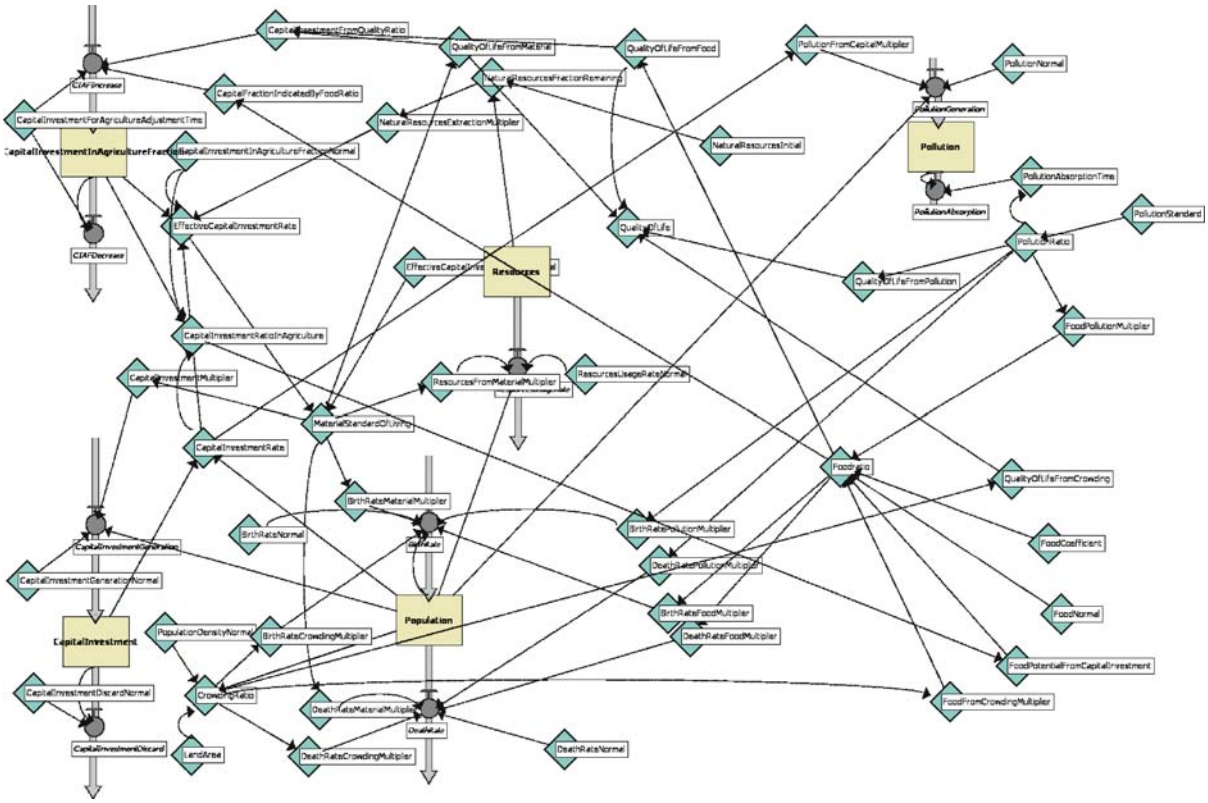
Social Processes, Simulation Models of, Figure 2
Graphical output of the sales growth STELLA model

number ranging from 0.13 to 1.85 whereas, of course, in a real firm, persons would be hired on a full-time or half-time base, but numbers such as 0.13 or 1.85 would be quite unlikely.

At the other end of the range of applications, the world models of the early 1970 can be placed, which became very famous as the Club of Rome funded and published them. Their objective was to predict the consequences of the world population growth for the use of resources and the pollution. Forrester's version [19] is replicated in a NetLogo 4.0 reimplementation shown in Fig. 3. This model has only five level variables – population, resources, pollution, capital investment (in industry) and capital investment in agriculture – which are linked together by nine rate variables and a large number of auxiliary variables and constants. The growth of the world population, for instance, is determined by two rate variables that contain the numbers of births and deaths, respectively, per time unit, and these in turn are determined by natural birth and death rates (biological constants, as it were) and several multipliers, auxiliary variables that transport the influences of causes such as pollution, food availability etc.; resources is another level variable, and this one is only determined by one rate variable, namely the decrease in resources per time unit, as Forrester (and also his successors in Meadows's research group) did not foresee that new resources might be discovered that were not known at the beginning of the simulation run (1900) or at the time when the simulation model was designed (1970).

The output of the model (see Fig. 4) shows what was predicted in the 1970s: a continued increase in the world population up to 8.6 billion in 2040, a decrease in available resources from 900 billion units (whatever these mean) in 1900 to 445 billion in 2040, and an increase in pollution from 200 million to 20 billion units (of another kind) in about 2050 – a result that was discussed worldwide in the 1970 (and perhaps this discussion was the reason why the predictions did not come out true 40 years later – although Meadows in [43] seems to have believed that his predictions were validated by the two decades after [42] appeared.

It is one of the big advantages of system dynamics models that they are easy to understand, easy to develop and to maintain and that they deliver plausible predictions that can easily be validated (or invalidated) against empirical data, though only after several decades. One of their important drawbacks is that the estimation of their input parameters is difficult if not impossible. Take as an example the estimation of a function between the pollution ratio and the pollution dependent death-rate multiplier (the pollution ratio is the actual pollution divided by the "pollution standard", which is defined as 3.6 billion pollution units – approximately the pollution in 1975). In Forrester's world model this function is programmed as



Social Processes, Simulation Models of, Figure 3
Forrester's WORLD2 model in a NetLogo 4.0 implementation

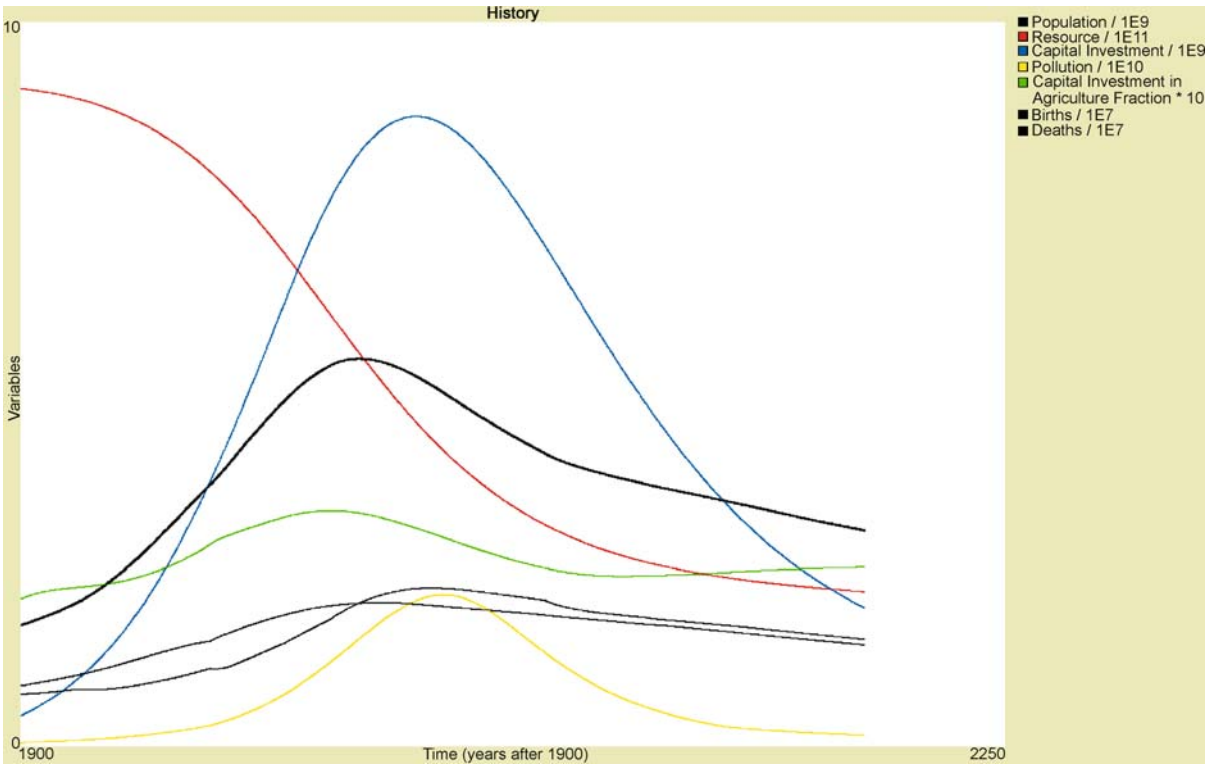
a table function (see Fig. 5), but in many cases it is quite questionable how the form of such a function was estimated.

In [42], the estimation of parameters for many functions between auxiliary variables is explained. The strategy of estimating the parameters of these functions is in most cases the same: it is usually the result of a nonlinear regression between – to take the preceding example – the pollution rate and the death rate pollution multiplier estimated from as many countries of the world as data could be found from, thus believing that synchronously found data from various places in the world were representative of data for the world as a whole collected at different points of time.

Microanalytical Simulation

Microanalytical simulation [22,47] has its origin in the necessity of predicting tax income and tax loads as well as the cost of social transfer. Every modern state needs information on how much taxes must be levied on its citizens in order to perform all the tasks of the state and how much social security contributions must be levied in order to pay retirement income, and it needs to know how many children will enter and leave the school system, how many students will enter and leave universities, how many people will enter the labor force at what time of their life and leave it at the time of their retirement, and, lastly, how many people will have near relatives who might be ready and willing to nurse them when they are in need. Whereas the age structure can be estimated even with difference equations, thus with system dynamics techniques, at least the last of the mentioned questions (which takes kinship networks into account) can only be answered on the micro level.

Microanalytical simulation always starts with a large representative sample of the target population. The sampling units are usually households (of which all members are interviewed), and the attributes sampled are at least age, gender, marital, educational and socio-economic status as well as the links to other members of the same household. This sample is then updated for every time step (usually one year) to answer the questions mentioned above for some time in the future.



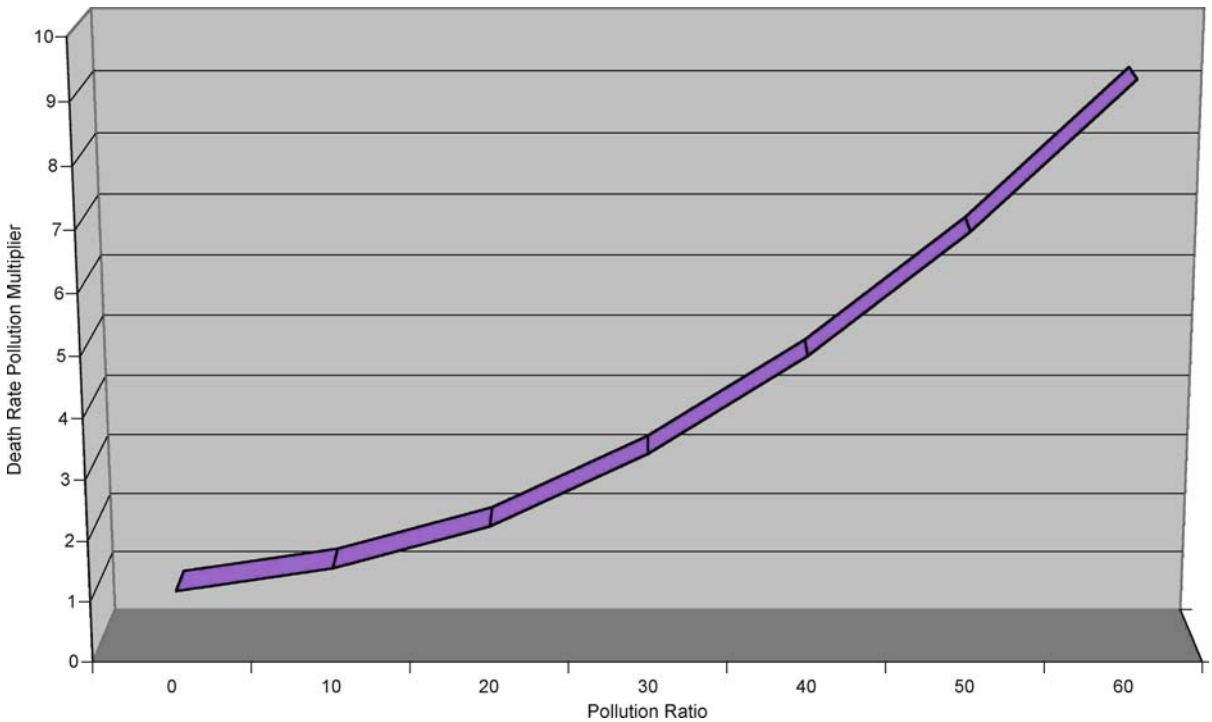
Social Processes, Simulation Models of, Figure 4
Some time series of Forrester's WORLD2 model, from a NetLogo 4.0 implementation

Microanalytical simulation has two easily distinguishable variants. They are usually labelled "static microsimulation" and "dynamic microsimulation". The former is static in so far as the attributes of the sample individuals are not changed over time, only the weights of the individual are updated in order to be representative for some macro variables where the information about the composition are taken from other (demographic or econometric) models. This static simulation is used only for short-time predictions, for instance for estimating tax loads due to revised tax regulation. Dynamic microsimulation changes all individual attributes according to known or estimated probabilities of state changes (for instance age-dependent birth, death, marriage and divorce rates or frequencies of transitions within the educational system and from the educational system into the labor force). For every time step, every micro entity has to be updated, and its link to other micro entities has to be revised, too. As compared to static microsimulation, this variant has the advantage that all state changes are endogenous and, more important, that state transition probabilities can be disaggregated to small sectors of the target and sample population (for instance, assuming different probabilities of giving birth to a first, second, third etc. child for women of different age groups, of different educational and socio-economic state), given that sufficient evidence for estimating such probabilities is available.

Early microanalytic simulation models were always programmed in general purpose programming languages from scratch, without using any particular tools. Their results were only partially published as research was mostly done for state agencies who were not always interested in having all the details of models published. Only recently, toolboxes such as UMDBS [56] and CoMicS [31] became available which lend themselves to be used in academic research and teaching, the more so as in many countries data sets of large representative samples and even panels have become available to the academic public for more than a decade now.

Multilevel Simulation

Multilevel and multi-agent simulations, different as they may be, always had their focus on interactions either among agents of the same kind or between agents of different levels or both. With levels, these approaches under-



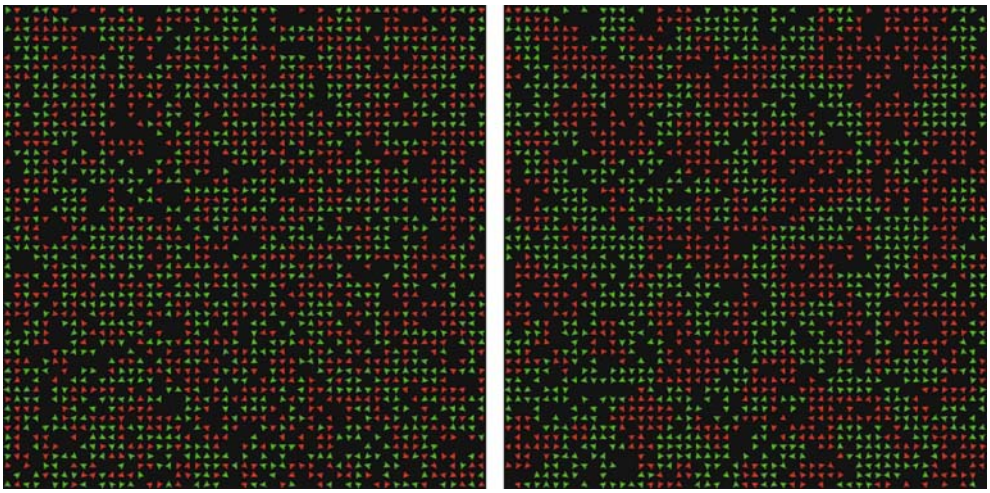
Social Processes, Simulation Models of, Figure 5
How system dynamics defines functions between its variables: the dependence of the death rate multiplier on the pollution ratio

stand different kinds of systems, where the elements of the systems of one kind are systems of the other kind [6]. To give an example: From the biologist's point of view, human beings are systems of one kind (composed of different kinds of tissues, or of cells, etc.), and on the other hand they can be seen as elements of systems called human groups, or populations, or societies etc. And what multilevel simulations are interested in is the interaction between, for instance, the group and an individual, between the macro and the micro level, thus focusing on the stern and bow of the "Coleman boat" [9]. Multi-agent simulation models are usually less focused on the inter-level interaction and more on the inter-agent interaction, but they, too, are usually interested in detecting properties emerging on the macro level.

More often than not, the label "multi1-agent simulation" is used for models in which the micro entities display only very simple behavior, have only very few states and do not communicate with other micro entities (in any reasonable sense of the word "communicate"), have no memory and make no decision. To distinguish these from multi-agent models proper (see the next section), they will here be labelled as "multilevel" models.

A first, quite well-known example may illustrate this. Thomas Schelling [57,58] designed a model in which (rather primitive) agents have to decide whether they move from one neighborhood into another, in which they can expect to be happier. Their happiness depends on the composition of their neighborhood, given that agents differ only with respect to one dichotomous property (say, their native language) – they are the happier, the more people in their neighborhood speak the same language, and there is a threshold where happiness turns into unhappiness. The world in Schelling's model is represented with a cellular automaton (see the chapter on cellular automata in this Encyclopaedia), and the neighborhood consists of the eight cells around an agent, which may be empty or inhabited by another agent. The central agent is happy if the language distribution in its neighborhood is such that a certain percentage of its neighbors speak the same language. In each time step, every agent scans its neighborhood and decides whether to move or stay, if it decides to move it tries to find another cell where the neighborhood is more attractive than in its original cell. One realization [67] of this model is shown in the two parts of Fig. 6.

It shows that the random distribution of red and green agents soon changes into a distribution with clearly distinguishable red and green clusters, part of which are even separated by uninhabited zones. The threshold between



Social Processes, Simulation Models of, Figure 6
Schelling's segregation model (*left:* initialization, *right:* state of the macro level when all individuals are happy with their neighborhoods)

happiness and unhappiness in this example is 30 per cent, i. e. an agent moves only if more than 70 per cent in its neighborhood are different from itself. Although these agents are quite tolerant with respect to strangers, clusters occur; and these clusters are larger and better separated when the tolerance threshold is higher. One should note here that the Schelling model takes only local interactions between agents into account; there is no interaction between levels if only in so far as the structure of the macro level changes due to the individual interactions, but it does not feedback into the micro level as the agents are not capable of recognizing what happens on the macro level.

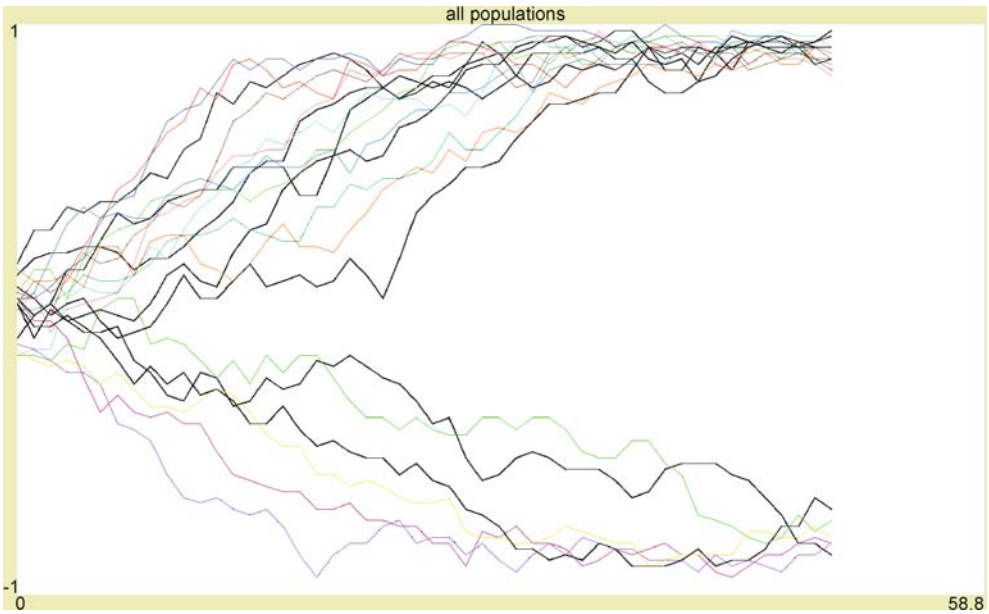
Another early approach to multilevel modelling of social processes is mainly interested in the interactions between the levels. It was originally developed by Hermann Haken [21], a German physicist who tried to apply the methods with which he had done research in laser physics to other disciplines, including sociology, and extended by his colleagues Wolfgang Weidlich and Günter Haag [66]. This approach, called synergetics by Haken, is currently better known as sociophysics or econophysics. Although the micro level entities are often also called agents by several authors, they are only particles as they only react on something like social forces or fields that are modelled in a manner very similar to the description of gravitational, electrostatic and electromagnetic forces and fields.

The original work by Weidlich and Haag [66] describes populations consisting of a large number of individuals where the individuals usually have only one discrete state variable which is either binary or can take only very few values, and the state change of the individuals depends on the macro state which is in turn determined by the individual state changes. The simplest example may make this clear: Imagine a population whose members have to decide between two options (yes and no) and whose decision depends on the perceived current majority in such a way that the probability of switching from no to yes is the higher, the larger the current yes majority is and the other way round. Then under certain assumptions of modelling the exact function between the current yes majority and the individual transition probabilities, the population as a whole shows some interesting behavior. In the example the two individual transition probabilities have the following form:

$$\mu_{\text{yes}\leftarrow\text{no}} = \nu \exp(\pi + \kappa x)$$

$$\mu_{\text{no}\leftarrow\text{yes}} = \nu \exp[-(\pi + \kappa x)]\,,$$

where ν is a flexibility parameter, π is a preference parameter, and κ is a coupling parameter, whereas x describes the current state of the population ($x = -1$ means "all no", $x = +1$ means "all yes"). The flexibility parameter only determines how likely any change between options is; the preference parameter determines whether changes from yes to no are generally more likely than those from no to yes (if it is negative) or the other way round (if it is positive); and the coupling parameter determines the strength of the influence the population exerts on the individual. It goes without saying that any individual change changes the state of the population. Figure 7 shows a typical outcome of a simulation of this very abstract social process. Each of the curves represents the history of the state of one out of twenty populations (all of them with a vanishing preference parameter and a coupling parameter of 1.4);



Social Processes, Simulation Models of, Figure 7
History of 20 populations whose members change their state with a probability that depends on the state of the population

some of the curves show that their populations soon develop into populations with a very high yes majority while others become populations with a high no probability, but none of them is ever entirely homogeneous. A more concrete social process that is likely to behave similarly is the distribution of two alternative goods that are not easily compatible, e. g. computer operating systems or videotape machines; here one could expect a high coupling as being with the majority makes the exchange of programs or videos easier. A low coupling constant results in an outcome where all populations are more or less evenly split in yes and no members – which in reality could be the case when the two alternative goods are easily compatible or substitutable, e. g. cigarette brands.

Extensions of this model were proposed by Lumsden and Wilson [41] who replaced the exponential function in the formula for the individual transition probabilities with other types of functions, to explain the co-evolution between "genes, mind, and culture". Other extensions, already given in [66], deal with several populations interacting with each other, the individual transition probabilities not only depending on the state of the members' own population, but also on the state of other populations, for instance modelling migration behavior of two populations between two regions.

Helbing [26] extended the models of his teachers, Wolfgang Weidlich and Günter Haag, in a way that he described as "quantitative sociodynamics". Much like his teachers he used closed solutions wherever possible to analyze the abstract social processes that he formalized, at the same time extending the focus not only to interactions between levels but also to interactions between individuals [24,25] which he also used to model the behavior of traffic participants, mainly pedestrians [24,27]. Here again, individual human beings are modelled as particles in a field and not endowed with the familiar human capabilities of decision making. In Helbing's pedestrian models (e. g. [27]) the pedestrians are just particles moving in a "social field", a term originally coined by Lewin [38], according to whose theory "behavioral changes are guided by so-called social fields or social forces, which has later on been put into mathematical terms" (p. 625 in [27]). These particle-pedestrians are accelerated at any time by a force which is determined by their desired directions and speeds and a repulsive force which keeps them apart from each other and from obstacles. Although the comparison between simulation results and video recordings of real pedestrian behavior shows nearly no qualitative difference, the particle-pedestrian are far from the definition of agents that is used by other authors who describe their simulations as multi-agent models.

Another kind of "social field" models was used to analyze attitude formation processes in an artificial society where the attitudes were not binary or categorical, but continuous. Models like these were inspired by Anthony Downs's early theory on party affiliation [13]. Downs as-

sumed that voters moved in a continuous attitude space (one-dimensional, from left to right) and were attracted or repelled by parties, and he believed that under certain circumstances a polarization of the electorate would emerge with high frequencies on the left and right wings and low frequencies in the middle, but he never formalized this theory. Several different formalizations were published in the sequel which tried to explain polarization without the interference of parties. One of these formalizations was inspired by empirical data which – for some German elections in the past – showed such a polarization in terms of a bi- or multimodal frequency distribution of voters over their attitude space which emerged during election campaigns and vanished shortly after the election [63]. The simulation runs started with a sample of a large numbers of model entities which obeyed an approximate bivariate normal distribution at the beginning (samples are never perfectly normal) entities then tried to move according to a function which was composed of the gradient of the frequency density distribution and some random effect. These movements changed the frequency density distribution, which in turn changed the gradient. In most simulation runs the frequency density functions became bimodal or multimodal after some time. Another representative of models inspired by Downs was presented by Guillaume Deffuant and is colleagues [12]. Here the simulated entities have a one-dimensional continuous attitude (much like in [63]), but additionally something like persuasive power, and as in [63], typically multimodal distributions emerge with modes both in the center but also at the ends of the attitude range, representing extremists in real electorates.

The approach in [12] is in some way similar to social impact theory developed by Bibb Latané [34] who also combined attitude and persuasive power, but let simulated entities interact on a grid, thus introducing topography and restricting interaction to near-neighbor interaction. The simulation models following his theory [35,36] show, similarly to the synergetic models, a survival of minorities, but also the clusters observed in Schelling's migration models (although the simulated entities in the social impact theory do not move in topological space).

Multi-Agent Simulation

An example that can easily be compared to the pedestrian models is the riot behavior model of Wander Jager and colleagues [29]. Here, agents belong to two different subpopulations (party 1, party 2) which can be imagined as the fans of two competing football clubs. Besides moving in a two-dimensional space, they can decide to fight or withdraw. In a version with police as a third subpopulation, the agents of this kind have additional behavioral options. Which action an agent takes does not depend on anything like a social field (at least not in the sense of Helbing, but perhaps in the sense of Lewin), but on their perceptions of the behavior of other agents in their neighborhood. These agents follow more or less the definition of agents that was given by Wooldridge and Jenkins [69] and which is currently the most often cited definition of agents:

An agent is a piece of software that

- has some autonomy, i. e. operates without other parts of a larger computer program having direct control of its internal state and actions,
- can communicate with other agents in some kind of language, i. e. without direct access to other agents' inner state,
- can perceive its environment and can react on it, and
- can also take the initiative, i. e. can act without a stimulus from outside and display goal-directed behavior.

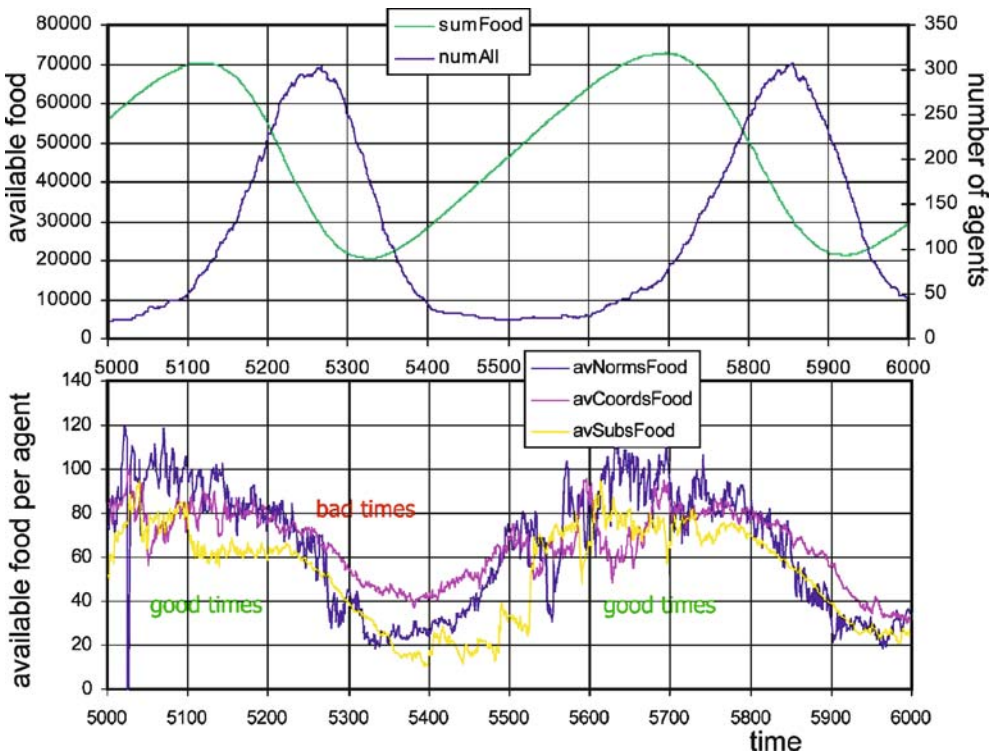
This results in several requirements: In order to be able to take the initiative (to be pro-active), agents must have goals (descriptions of desired and possible states of their neighborhood or surroundings, including themselves) and must be able to compare their perceived state of their environment to their goals. According to the outcome of such a comparison, they must decide between several possible actions which they have to evaluate, and moreover they must be able to design a plan (i. e. a course of possible actions) if there is no action that leads them directly to their goal. Another requirement can be that agents communicate with other agents, for instance asking them for their help when only co-operation enables them to achieve their goals.

Among the early forerunners of multi-agent systems in the social sciences at least two can be named where processes of voter attitude changes are modelled and simulated. Although the poor computer languages of the late 1950s and early 1960s did not allow for agents in the sense of our days, any re-implementation would nowadays be a multi-agent system with several classes of agents (representing voters, candidates, media channels, as in [1] or in the Simulmatics project supporting John F. Kennedy's election campaign [11]) as they dealt with the communications among citizens, between citizens and candidates as well as between citizens and media channels and modelled their behavior and actions in a rule-based manner.

A classical example that fulfills most of the requirements raised in [69] is Epstein's and Axtell's attempt at building "Social Science from the Bottom Up" [15], an artificial world in which agents make their living on two dif-

ferent crops which they harvest, store, consume and barter according to their needs and the available resources. An extension of this artificial world can be found in [32]. In addition to all the features of the artificial society in [15], these agents can subordinate to others if the latter offer themselves as co-ordinators and promise their subordinates to inform them about the state of a wider environment than they can perceive individually, while the subordinates in turn provide their chieftains with their local information and pay a certain tribute from what they harvest. To be able to do all this, all agents have a long-term memory in which they store information about their world and which can guide them to places in their world that promise a rich harvest of the kind of crop they actually need. This extended model shows that a world with co-ordinators and subordinates is more sustainable than a world in which all agents try to harvest on their own; it also shows that lonely agents are best off when times are getting better, whereas subordinating pays best when times are getting worse and being a co-ordinator pays best when times are worst. Figure 8 is a part of two plots written by the simulation program: the green curve in the upper graph shows the development of food available over time, the blue one shows the number of surviving agents over time, while the three curves in the lower graph show the average "income" of the three types of agents (blue. the lonely agents, magenta: the chieftains, yellow: the subordinates) – "bad times" is when the overall available food is at its minimum, "good times" is when it is at a maximum. Among others, Fig. 8 shows that the macro behavior of the process is the same as the behavior of the well-known Lotka–Volterra model [39,65], but the classical coupled differential equations of this model or the stochastic process described in [66] would not have been able to display the additional information about the fate of the three kinds of agents.

Validating a model [64] such as the one described above is nearly impossible, as the model describes an abstract social process which is very unlikely to have ever occurred in reality. Nevertheless, this model makes a prediction that is quite similar to the prediction of another agent-based simulation [52,54]: "In a society of herdsmen and farmers in Western Africa, decisions which rest on



Social Processes, Simulation Models of, Figure 8
Development of an artificial society whose members can declare themselves co-ordinators and subordinates, respectively (graphs produced by the simulation model described in [32])

friendship networks ('friend-priority' decisions) proved to be much more effective then decisions which were made on pure cost deliberations ('cost priority' decisions)." See p. 189 in [52]. The papers quoted here are at the same time representatives of a style of modelling and simulating social processes which is known as "participative simulation" and usually used to discuss models with stakeholders who work with the simulation models and give hints at improving and amending models to adapt them better to their personal experience. This style of modelling and simulation is being used more and more often in consultancy to give stakeholders an opportunity to work with models on their own, instead having to believe what the consultant presents in the end.

Tools

Computer simulation always needs some tools to be run. In the early times of computer simulation, simulation programs were written in general purpose programming languages (like FORTRAN), but over the decades special simulation toolboxes were designed and implemented, first by researchers in fields such as organization science, production management (particularly for discrete event and queuing systems) or physics and engineering (particularly for numerical solutions of differential equations). Both of these kinds of tools were also used by social scientists, though only rarely. DYNAMO was the first simulation tool designed for and used by economists and social scientists [50], but only for system dynamics models. It was later on superseded by STELLA [53] with the same purpose, but a graphical user interface (see Figs. 1 and 2). There are several simulation tools for cellular automata (see the entry on cellular automata for more details). Multi-level simulation models are often written in general purpose programming languages (but see [44] describing a language – MIMOSE – explicitly designed for this kind of models). Multi-agent simulation proper has been developed with the help of toolboxes such as SWARM [30], RePAST [45,46], MASON [40], CORMAS [37] and NetLogo [68] (the latter was also used to produce the majority of graphics in this entry). All of these are freeware and capable of supporting most social simulation techniques (with the exception of microanalytic simulation, as this always needs a large database connected to the simulation tool). With the exception of NetLogo, which has a simulation language of its own, the actual description of agent behavior has to be done in object-oriented languages such as Smalltalk, Objective C or Java (for more detailed reviews of these toolboxes see [51,62]).

Future Directions

In a way, "agents cover all the world" [4] in that multi-agent systems can be used for all simulation purposes, as agents can always be programmed in a way that they behave as continuous or discrete models, can be activated according to event scheduling or synchronously or in a round-robin manner, can use rule bases as well as stochastic state transitions, and all these kinds of agents can even be nested into each other, thus supporting a wider range of applications than any of the classical simulation approaches. This leads to a third aspect of complexity (after the complexity of domains and the complexity of time): agent-based models can encompass several different approaches, both from a technical and implementation point of view, but also from the disciplines making use of simulation (for instance, disciplines such as neurophysiology, cognitive psychology, social psychology and sociology can combine their contributions into a deeply structured simulation model).

It goes too far to say that all this would not be possible in a non-agents world (as everything is programmable in Assembler), but the examples above will have made clear that models combining aspects from even neighboring disciplines as the ones enumerated in this entry are only understandable and communicable when they come in a modular form that is typical for agent-based models – and the same applies to ecological models where disciplines from physics and biochemistry to population biology and economics would play their co-operative roles.

Although agent communication languages (see the special issue of Autonomous Agents and Multi-Agent Systems, vol. 14 no. 2) such as KQML [33] have been developed for a long time (back to 1993), it is still an open question how agents in a simulation model could develop a communication means on their own and/or extend their communication tool to be able to refer to a changing environment (see the special issue of Autonomous Agents and Multi-Agent Systems, vol. 14 no. 1). Hutchins and Hazlehurst [28] made a first step into the field of the emergence of a lexicon, but their agents were only able to agree on names of things (patterns) they saw. The NEWTIES project [20], ambitious as it is, aims at creating an artificial society that develops its own culture and will also need to define agent capabilities that allow them to develop something like a language although it is still questionable whether the experimenters will be able to understand what their artificial agents talk about. The same objective is aimed at in the current EMIL project (Emergence in the Loop [3,10]) which attempts at creating an agent society in which norms emerge as agents observe each

other and draw conclusions about which behavioral feature is desirable and which is misdemeanor in the eyes of other agents – which, as in the case of language emergence, makes it necessary that agents can make abstractions and generalizations from what they observe in order that ambiguities are resolved.

This means that multi-agent models in the proper sense – where software agents model cognitive capacities of human beings – will be the aim of any future development in the simulation of social processes, although even these models will still contain features of the classical approaches, as even human beings display behavior that can be categorized as reflex acts and instinctive and which can be modelled with models borrowed from particle kinematics in which human beings are represented by stochastic automata. On the other hand it should be clear that only part of human behavior in very large groups can be explained this way, as an important part of this behavior consists in deliberate and conscious actions that are determined by cognition.

Bibliography

Primary Literature

1. Abelson RP, Bernstein A (1963) A Computer Simulation of Community Referendum Controversies. Pub Opin Q 27(1):93–122
2. Anderson JM (1973) The Eutrophication of Lakes. In: Meadows D, Meadows D (eds) Toward Global Equilibrium. Wright Allen, Cambridge, pp 171–140
3. Andrighetto G, Campenni M, Conte R, Paolucci M (2007) On the Immergence of Norms: a Normative Agent Architecture. In: Proceedings of AAAI Symposium, Social and Organizational Aspects of Intelligence, Washington
4. Brassel KH, Möhring M, Schumacher E, Troitzsch KG (1997) Can Agents Cover All the World? In: Conte R, Hegselmann R, Terna P (eds) Simulating Social Phenomena. Springer, Berlin, pp 55–72
5. Bunge M (1977) Ontology I: The Furniture of the World. Treatise of Basic Philisophy, vol 3. Reidel, Dordrecht
6. Bunge M (1979) Ontology II: A World of Systems. Treatise of Basic Philosophy, vol 4. Reidel, Dordrecht
7. Castelfranchi C (1998) Simulating with Cognitive Agents: The Importance of Cognitive Emergence. In: Sichman JS, Conte R, Gilbert N (eds) Multi-Agent Systems and Agent-Based Simulation. Springer, Berlin, pp 26–44
8. Castelfranchi C, Conte R (1994) Mind is not Enough. The Precognitive Bases of Social Interaction. In: Doran JE, Gilbert N (eds) Simulating Societies: The Computer Simulation of Social Phenomena. UCL Press, London, pp 267–286
9. Coleman JS (1990) Foundations of Social Theory. The Belknap Press of Harvard University Press, Cambridge
10. Conte R, Andrighetto G, Campenni M, Paolucci M (2007) Emergent and Immergent Effects in Complex Social Systems. In: Proceedings of AAAI Symposium, Social and Organizational Aspects of Intelligence, Washington
11. de Sola Pool I, Abelson RP (1962) The Simulmatics Project. Pub Opin Q 25:167–183
12. Deffuant G, Amblard F, Weisbuch G, Faure T (2002) How can Extremism Prevail? A study based on the relative agreement interaction model. J Artif Soc Soc Simul 5(4) http://www.soc.surrey.ac.uk/JASSS/5/4/1.html
13. Downs A (1957) An Economic Theory of Democracy. Harper, New York
14. Durckheim E (1982) The rules of the sociological method (trans: Halls WD). The Free Press, New York
15. Epstein JM, Axtell R (1996) Growing Artificial Societies – Social Science from the Bottom Up. MIT Press, Cambridge
16. Forrester JW (1961) Industrial Dynamics. MIT Press, Cambridge
17. Forrester JW (1968) Principles of Systems. MIT Press, Cambridge
18. Forrester JW (1969) Urban Dynamics. MIT Press, Cambridge
19. Forrester JW (1971) World Dynamics. MIT Press, Cambridge
20. Gilbert N, den Besten M, Botovics A, Craenen BGW, Divina F, Eiben AE, Griffioen R, Hévízi G, Lõrincz A, Paechter B, Schuster S, Schut M,C Tzolov C, Vogt P, Yang Lu (2006) Emerging Artificial Societies Through Learning. J Artif Soc Soc Simul 9(2) http://www.soc.surrey.ac.uk/JASSS/9/2/9.html
21. Haken H (1978) Synergetics. An Introduction. In: Nonequilibrium Phase Transitions and Self-Organization in Physics, Chemistry and Biology, 2nd edn. Springer, Berlin
22. Harding A (ed) (1996) Microsimulation and Public Policy. Contributions to Economic Analysis, vol 232. Elsevier, Amsterdam
23. Helbing D (1991) A Mathematical Model for the Behavior of Pedestrians. Behav Sci 36:298–310
24. Helbing D (1991/1992) A Mathematical Model for Behavioral Changes by Pair Interactions and Its Relation to Game Theory. Angew Soz 17:179–194
25. Helbing D (1992) A Mathematical Model for Attitude Formation by Pair Interactions. Behav Sci 37:190–214
26. Helbing D (1994) Quantitative Sociodynamics. Stochastic Methods and Models of Social Interaction Processes. Kluwer, Dordrecht
27. Helbing D, Johansson A (2007) Quantitative Agent-Based Modeling of Human Interactions in Space and Time. In: Amblard F (ed) Proceedings of the 4th Conference of the European Social Simulation Association (ESSA'07), Toulouse, 10–14 Sept, 2007, pp 623–637
28. Hutchins E, Hazlehurst B (1995) How to invent a lexicon: the development of shared symbols in interaction. In: Gilbert N, Conte R (eds) Artificial Societies. The Computer Simulation of Social Life. UCL Press, London, pp 157–189
29. Jager W, Popping R, Sande HVD (2001) Clustering and fighting in two-party crowds: Simulating the approach-avoidance conflict. J Artif Soc Soc Simul 4(3) http://www.soc.surrey.ac.uk/JASSS/4/3/7.html
30. Johnson P, Lancaster A (2001) Swarm User Guide. http://www.swarm.org/swarmdocs-2.1.1/userbook/userbook.html
31. Klein C, Fuchs D, Berger P, Hassenpflug P (2006) MicS. Agentenbasierte Mikrosimulation, Project Thesis (Studienarbeit), Universität Koblenz-Landau, Koblenz
32. König A, Möhring M, Troitzsch KG (2003) Agents, Hierarchies and Sustainability. In: Billari F, Fürnkranz-Prskawetz A (eds) Agent-Based Computational Demography. Springer, Berlin, pp 197–210
33. Labrou Y, Finin T (1997) A Proposal for a new KQML Specification, Computer Science and Electrical Engineering Depart-

ment (CSEE). University of Maryland Baltimore County (UMBC), Baltimore
34. Latané B (1981) The psychology of social impact. Amer Psychol 36:343–356
35. Latané B (1996) Dynamic Social Impact. Robust Predictions from Simple Theory. In: Hegselmann R, Mueller U, Troitzsch KG (eds) Modelling and Simulation in the Social Sciences from a Philosophy of Science Point of View. Kluwer, Dordrecht, pp 287–310
36. Latané B, Nowak A (1994) Attitudes as Catastrophes: From Dimensions to Categories with Increasing Involvement. In: Vallacher R, Nowak A (eds) Dynamical Systems in Social Psychology. Academic Press, San Diego, pp 219–250
37. LePage C, Bousquet F, Bakam I, Bah A, Baron C (2000) CORMAS: A multiagent simulation toolkit to model natural and and social dynamics at multiple scales. In: Proceedings of the Workshop "The ecology of scales", Wageningen
38. Lewin K (1951) Field Theory in Social Science. Harper, New York
39. Lotka AJ (1925) Elements of Physical Biology. Williams & Wilkins, Baltimore
40. Luke S, Cioffi-Revilla C, Panait L, Sullivan K (2004) MASON: A New Multi-Agent Simulation Toolkit. In: Proceedings of the 2004 SwarmFest Workshop, Santa Fe
41. Lumsden CJ, Wilson EO (1981) Genes, Mind, and Culture. The Coevolutionary Process. Harvard University Press, Cambridge
42. Meadows D et al (1974) Dynamics of Growth in a Finite World. MIT Press, Cambridge
43. Meadows D et al (1992) Beyond the Limits. MIT Press, Cambridge
44. Möhring M (1996) Social Science Multilevel Simulation with MIMOSE. In: Troitzsch KG, Mueller U, Gilbert N, Doran JE (eds) Social Science Microsimulation. Springer, Berlin, pp 123–137
45. North MJ, Howe TR, Collier NT, Vos RJ (2005) The Repast Simphony Runtime System. In: Agent 2005 Conference on Generative Social Processes, Models and Mechanisms. Argonne National Laboratory, Argonne
46. North MJ, Collier NT, Vos JR (2006) Experiences Creating Three Implementations of the Repast Agent Modeling Toolkit. ACM Trans Model Comp Simul 16(1):1–25
47. Orcutt GH, Merz J, Quinke H (1986) Microanalytic simulation models to support social and financial policy. Information Research and Resource Reports, vol 7. North-Holland, Amsterdam
48. Planck M (1930) Positivismus und reale Außenwelt. In: Planck M (1949) Vorträge und Erinnerungen. Hirzel, Stuttgart, pp 228–245
49. Poundstone W (1992) Prisoners' Dilemma. John von Neumann, Game Theory, and the Puzzle of the Bomb. Oxford UP, Oxford
50. Pugh III AL (1976) DYNAMO User's Manual. MIT Press, Cambridge
51. Railsback SF, Lytinen SL, Jackson SK (2006) Agent-based Simulation Platforms: Review and Development Recommendations. Simulation 82(9):609–623
52. Rapoport A (1986) General Systems Theory. Essential Concepts & Applications. Abacus, Tunbridge Wells and Cambridge
53. Richmond B (2001) An introduction to systems thinking: STELLA software. ISEE Systems, Lebanon
54. Rouchier J, Bousqet F, Barreteau O, LePage C, Bonnefoy JL (2000) Multi-Agent Modelling and Renewable Resources Issues: The Relevance of Shared Representations for Interacting Agents. In: Moss S, Davidsson P (eds) Multi-Agent-Based Simulation. Springer, Berlin, pp 181–197
55. Rouchier J, Bousquet F, Requier-Desjardins M, Antona M (2001) A multi-agent model for describing transhumance in North Cameroon: comparison of different rationality to develop a routine. J Econo Dyn Control 25:527–559
56. Sauerbier T (2002) UMDBS – A New Tool for Dynamic Microsimulation. J Artif Soc Soc Simul 5:2–5 http://jasss.soc.surrey.ac.uk/5/2/5.html
57. Schelling TC (1971) Dynamic Models of Segragation. J Math Soc 1:143–186
58. Schelling TC (1978) Micromotives and Macrobehavior. Norton, New York
59. Simon HA (1957) A Formal Theory of Interactions in a Social Group. Amer Soc Rev 17:202–211
60. Simon HA (1957) Models of Man, Social and Rational. Math Essays Ration Hum Behav Soc Setting. Wiley, New York
61. Simon HA (1996) Models of My Life. MIT Press, Cambridge
62. Tobias R, Hofmann C (2004) Evaluation of free Java-libraries for social-scientific agent based simulation. J Artif Soc Soc Simul 7 http://jasss.soc.surrey.ac.uk/7/1/6.html
63. Troitzsch KG (1998) Multilevel Process Modeling in the Social Sciences: Mathematical Analysis and Computer Simulation. In: Liebrand WBG et al (eds) Computer Modeling of Social Processes. Sage, London, pp 20–36
64. Troitzsch KG (2004) Validating Simulation Models. In: Horton G (ed) Networked Simulations and Simulated Networks. 18th European Simulation Multiconference. SCS, Erlangen, San Diego, pp 265–270
65. Volterra V (1926) Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. Allt Accad Naz Lincei 6(2):31–113
66. Weidlich W, Haag G (1983) Concepts and Models of a Quantitative Sociology. The Dynamics of Interacting Populations. Springer, Berlin
67. Wilensky U (1998) NetLogo Segregation model. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston. http://ccl.northwestern.edu/netlogo/models/Segregation
68. Wilensky U (2007) NetLogo 4.0 User Manual. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston. http://ccl.northwestern.edu/netlogo/docs/NetLogo%20User%20Manual.pdf
69. Wooldridge M, Jennings NR (1995) Intelligent agents: theory and practice. Knowl Eng Rev, 10:115–152

Books and Reviews

Ahrweiler P, Gilbert N (1998) Computer Simulations in Science and Technology Studies. Springer, Berlin
Amblard F (2007) Interdisciplinary Approaches to the Simulation of Social Phenomena. In: Proceedings of the 4th Conference of the European Social Simulation Association. IRIT, Toulouse
Casti JL (1997) Would-Be Worlds. How Simulation Is Changing the Frontier of Science. Wiley, New York
Coelho H, Espinasse B (2004) 5th Workshop on Agent-Based Simulation. SCS, Erlangen
Conte R, Hegselmann R, Terna P (eds) (1997) Simulating Social Phenomena. Springer, Berlin
Doran JE, Gilbert N (1994) Simulating Societies: The Computer Simulation of Social Phenomena. UCL Press, London

Edmunds B, Hernández C, Troitzsch KG (2008) Social Simulation. Technologies, Advances, and New Discoveries. Information Science Reference. Hershey, New York
Gilbert N, Conte R (eds) (1995) Artificial Societies. The Computer Simulation of Social Life. UCL Press, London
Gilbert N, Troitzsch KG (2005) Simulation for the Social Scientist. 2nd edn. Open University Press, Maidenhead
Hegselmann R, Mueller U, Troitzsch KG (eds) (1996) Modelling and Simulation in the Social Sciences from a Philosophy of Science Point of View. Kluwer, Dordrecht
Liebrand WBG et al (1998) Computer Modeling of Social Processes. Sage, London
Moss S, Davidsson P (2000) Multi-Agent-Based Simulation. Lecture Notes in Artificial Intelligence, vol 1979. Springer, Berlin
Müller JP, Seidel MM (2003) 4th Workshop on Agent-Based Simulation. SCS, Erlangen
Sichman JS, Conte R, Gilbert N (1998) Multi-Agent Systems and Agent-Based Simulation. Springer, Berlin
Simon HA (1977) Models of Discovery and Other Topics in the Methods of Science. Reidel, Dordrecht
Suleiman R, Troitzsch KG, Gilbert N, Mueller U (1998) Social Science Microsimulation. Tools for Modeling, Parameter Optimization and Sensitivity Analysis. Springer, Heidelberg
Sun R (2006) Cognition and Multi-Agent Interaction. From Cognitive Modeling to Social Simulation. Cambridge UP, New York
Troitzsch KG (2005) Representing Social Reality. In: Pre-Proceedings of the Third Conference of the European Social Simulation Association. Fölbach, Koblenz
Troitzsch KG, Mueller U, Gilbert N, Doran JE (eds) (1996) Social Science Microsimulation. Springer, Berlin
Vallacher R, Nowak A (eds) (1994) Dynamical Systems in Social Psychology. Academic Press, San Diego

Social Psychology, Applications of Complexity to

Robin R. Vallacher
Florida Atlantic University, Boca Raton, USA

Article Outline

Glossary
Definition of the Subject
Introduction
Intrinsic Dynamics
Attractors in Psychological Systems
Dynamical Minimalism
The Dynamics of Social Influence
Dynamics of Interpersonal Coordination
Future Directions
Bibliography

Glossary

Dynamical system A dynamical system is a set of interconnected elements that change due to their mutual influences. A change in each element depends on the nature of the influences from other elements. Due to these mutual influences, the system as a whole evolves in time. Because the state of every element is the effect of the states of all the relevant elements in the previous moment, and because the state of every element is one of the causes determining the state of the other elements in the next moment, a dynamical system is characterized by bi-directional causality.

Intrinsic dynamics In a dynamical system, the current state of the system is directly caused by the preceding state of the system. More precisely, the state of each element at a given moment is caused by the states of other elements in the previous moment and by its own previous state. This chain of cause and effect iterated over time gives rise to internally-generated or "intrinsic" dynamics.

Self-organization In a dynamical system consisting of inter-connected elements, the state of each element adjusts to the current state of other elements to which it is connected. Because of this mutual influence, higher order units develop that represent the organization of the basic elements. No higher-order agent is required for such order to emerge. Hence, the process is referred to as self-organization.

Emergence Emergence occurs when the individual elements of a dynamical system achieve organization by means of their mutual influence. The development of the higher level state is said to be emergent because this state was not inherent in the properties of the lower-level elements and because the higher-level state was not imposed on the system from forces outside the system. The higher-order properties that result from the mutual adjustment among lower-level elements provide coordination for the lower-level elements. Emergence thus provides for substantial growth in the complexity of a system's processes and properties. Because of emergence, very complex systems can often be described by very simple models.

Attractor An attractor represents a subset of a system's phase space to which the system evolves over time and which resists forces that would perturb this temporal trajectory. This subset can consist of a region of nearby states or it can represent two (or more) regions among which the system oscillates over time in a periodic, quasi-periodic, or chaotic manner. In a system governed by attractor dynamics, a relatively wide range of starting points (initial states) will eventually converge on a much smaller set of states or on a pattern of change between states defining the attractor. External influences may push the system out of the attrac-

tor, but over time the system will return to this equilibrium.

Cellular automata Cellular automata are programmable dynamical systems. In this approach, a finite set of elements is specified, each of which can adopt a finite number of discrete states. The elements are arranged in a specific spatial configuration that usually takes the form of a two-dimensional lattice or grid. The location of each element on this grid specifies the element's neighborhood. The elements evolve in discrete units of time, such that the state of an element at $t + 1$ depends on the states of the neighboring elements at time t. The dynamics of cellular automata depend on the nature of the updating rule and on the format of the grid dictating the neighborhood structure.

Definition of the Subject

The seminal insights that launched the field of social psychology over a century ago have stood the test of time. Such early scholars as James [41], Cooley [21], Mead [65], Lewin [59], and Asch [4] emphasized the multiplicity of interacting forces operating in individual minds and in social groups and the potential for sustained patterns of change resulting from such complexity. The inherent complexity and dynamism of human experience, however, proved problematic for mainstream social psychology during much of the 20th century. The canonical paradigm in this period reduced dynamics to a one-step process involving a purported cause, operationalized as an independent variable at an arbitrary time 1, and its effect, operationalized as a dependent variable that was assessed at an arbitrary time 2. The complexity of experience, with multiple variables interacting over time to generate a stream of thought or action, was difficult to investigate with available statistical techniques (e. g., correlation, regression, analysis of variance). Social psychology thus fostered an image of human experience that emphasized simplicity and stability at the expense of complexity and dynamism.

Contemporary social psychology shows signs of returning to the deep intuitions concerning human experience articulated by the field's founding fathers. The renewed appreciation for complexity and dynamism was made possible by developments in the understanding of nonlinear dynamical systems in the 1970s and 1980s, and the application of these developments to social processes within the last two decades. Computer simulations capture the complexity of social processes and document the emergence of higher-order properties from the interaction of basic elements in a mental or social system. Innovative means of collecting and analyzing time-series data provide rigorous insight into the intrinsic dynamics of mental, affective, behavioral, and interpersonal processes. And formal models enable researchers to identify the parameters that are critical for understanding the dynamism and complexity of social psychological phenomena. In short, the contemporary emphasis on complexity and dynamics provides a new paradigm for the field, one that holds promise for advancing the social psychology as a precise science while preserving the basic insights that launched the field over a century ago.

Introduction

Despite its ubiquity, dynamism is a poor candidate for theory construction. Theories are expressed in terms of invariant properties representing stable "signals" that are obscured by the "noise" associated with personal, interpersonal, and societal processes. The stream of consciousness may be an accurate depiction of subjective experience, for example, but its turbulent nature seems inconsistent with fundamental properties that transcend particular individuals and their moment-to-moment mental states. To identify regularities and invariant properties, social psychologists commonly ignore the ever-changing undercurrent of mind and action, focusing instead on those elements of thought and behavior that admit to stability and structure. Accordingly, they typically emphasize higher order units of mental and behavioral phenomena (e. g., traits, schemata, global evaluations, norms) that presumably lend stability and coherence to subjective experience. The turbulent flow of lower level elements (thoughts, feelings, movements) is effectively rendered irrelevant to "true" understanding and meaningful prediction.

The gulf between the reality of experiential turbulence and the focus on stability in theory construction is unnecessary. Dynamism and structure represent complementary aspects of experience that together provide the basis for cognitive, emotional, and behavioral accommodation. The basic notion is that the flow of lower-level elements gives rise to higher order structures, which in turn constrain the dynamics of lower-level elements. This reciprocal causal relationship between lower-level and higher-level units of experience is central to *dynamical social psychology* [74,106] a recently developed paradigm that represents an adaptation of the concepts, principles, and tools developed within dynamical systems and complexity science.

It is not feasible to describe every version of the dynamical perspective, nor every application of this general approach. This may have been possible in the

1990s [26,31,35,45,102] when the dynamical perspective was a promissory note rather than an established paradigm. Since that time, there has been a proliferation of many innovative research strategies—some providing formal models implemented in computer simulations, others offering empirical means for investigating the dynamics of personal, interpersonal, and societal processes—that have been used to investigate a wide variety of phenomena. The aim herein is to highlight the crucial elements of the dynamical perspective that find expression in otherwise distinct theories, research strategies, and topical agendas.

The first two sections (Sects. "Intrinsic Dynamics" and "Attractors in Psychological Systems") describe basic concepts from the study of nonlinear dynamical systems that are directly relevant to social psychology. Section "Dynamical Minimalism" describes dynamical minimalism [72], an approach that provides a workable entrée into the nature and expression of dynamic processes at different levels of social reality. This approach is illustrated in the next two sections with respect to two research agendas—one emphasizing the emergence of group-level properties from the self-organization of individual agents (Sect. "The Dynamics of Social Influence") the other exploring the tendency of individuals to coordinate their behavioral and mental dynamics in service of forming dyads and social groups (Sect. "Dynamics of Interpersonal Coordination"). The final section ("Future Directions") reflects on the trajectory of dynamical social psychology thus far and offers caveats concerning the relevance of this approach to the unique features of human experience.

Intrinsic Dynamics

People's mental and emotional states, overt behavior, and social relations evolve and change in the absence of external influence. The ubiquity of intrinsic dynamics is apparent at different levels of social reality, from basic intrapersonal processes to macro-level societal phenomena. At the level of the mind, the temporal pattern of cognitive and affective elements in the stream of thought [41] often provides a more accurate depiction of a person's mental make-up than do the summary aspects of the person's mental process (e. g., overall attitude, final decision) that are more often the focus of investigation [103]. For example, research has shown that simply thinking about an attitude object (e. g., another person) in the absence of external influence or new information tends to promote more extreme (polarized) evaluations of the object over time [97].

Research on social judgment has shown that internally generated thoughts and feelings about a target person often reflect elaborate but identifiable patterns of change that convey important information. A judgment that is neutral when average over time, for instance, can have very different meanings and implications, depending on the intrinsic dynamics of the judgment process [109]. When neutrality represents little variation in evaluation occurring on a relatively slow time-scale, the summary judgment might well reflect a truly neutral sentiment. If neutrality reflects oscillation between highly positive and highly negative judgments on a rapid timescale, however, the summary judgment signifies heightened involvement and ambivalence rather than neutrality per se.

Intrinsic dynamics also characterize personal action. Actions have a hierarchical structure, in that the performance of an action entails the coordinated interplay of more basic actions or sub-acts. "Going to work," for example, may involve getting dressed, leaving the house, driving a car, parking the car, and entering a building. Each of the lower-level acts, in turn, can be decomposed into yet more basic lower-level elements. "Driving," for instance, consists of starting the car, turning the steering wheel, making turns, and braking. Each level in an action hierarchy is associated with a different time scale, with increasingly lower-level acts taking place in correspondingly shorter intervals of time [69]. "Going to work" unfolds on a longer time scale than does "driving," for example, and the time scale for "driving" is longer than that for each instance of "turning the steering wheel." There is evidence that the embedded time scales in an action hierarchy often have a fractal structure [69].

The intrinsic dynamics of action, in turn, span the levels of action in an overall action hierarchy. What may appear to be a continual succession of momentary movements when defined in low-level, mechanistic terms can take on the appearance of switching between qualitatively different actions, each occurring on a longer time scale, when defined in higher-level terms. Research on action identification [107] has identified several factors that determine the level at which an action is regulated. This line of work has also demonstrated that people reliably differ in their default level of action identification across many action domains [108]. Individuals who tend to think about their actions in lower-level terms are predisposed to the emergence of higher-level action understanding. Emergence occurs when such individuals are exposed to cues (e. g., feedback from other people) that suggest higher-level meaning for the actions [114] or when they privately reflect on their actions, allowing the lower-level action elements to self-organize into a higher-level act identity that provides subjective integration for these elements [104].

Social interaction has also been investigated with respect to intrinsic dynamics. For example, research has focused on the interpersonal coordination of relatively low-level actions, such as speaking (e. g. [20,23]) and limb movement (e. g. [6,48,69,101]). In one approach, two individuals are asked to swing their legs while sitting down across from one another [6]. One person swings his or her legs in time to a metronome and the other person tries to match those movements. This research has revealed two forms of coordination: *in-phase*, with the individuals swinging their legs in unison, and *anti-phase*, with the individuals swinging their legs with the same frequency but in the opposite direction. Individuals can maintain anti-phase coordination only up to a certain frequency of movement, at which point they switch to in-phase coordination. When the frequency is subsequently decreased, at some value they are able to coordinate anti-phase again, but this tempo is significantly lower than the point at which they originally started to coordinate in-phase. Such *hysteresis* indicates that movement coordination can be analyzed as a nonlinear dynamical system [38,48]. This line of research has identified modes of coordination more complex than in-phase and anti-phase [5,87,101].

Interpersonal dynamics are not confined to the coordination of speech and motor movements, but also include the temporal coordination of higher-level actions (e. g., plans, goals) and internal states (e. g., moods, judgments). Although this topic has not been heavily researched, there is evidence that the quality of a social relationship is reflected in partners' ability to coordinate in-phase with respect to their respective higher-level actions, opinions, and feelings (e. g. [5,36,64,82,83,100]). The ebb and flow of feelings, information exchange, and action conveys deeper insight into the nature of a relationship than do global indices such as the average sentiment, the amount of information exchanged, or the summary action tendencies. Colloquially, people who feel positively about one another are said to "be in synch" or "on the same wavelength" with respect to their internal states.

At the societal level, tracking the temporal trajectory associated with the emergence of norms and public opinion provides greater insight into the society's future make-up and likely response to external threat than simply knowing what the societal norms and opinions are at a single point in time [76]. When norms and opinions develop gradually over a long period of time, for example, the society displays resistance to external threats or even to new information that might promote better economic conditions. But societal change in political and economic ideology can occur in a rapid, nonlinear manner (e. g. [75,84]), with a temporal trajectory that resembles phase transitions in physical systems [58]. Societies that undergo such nonlinear transitions are vulnerable to later rebounds of the earlier ideologies and are highly responsive to threats and new information, and they can experience a period of sustained oscillation between conflicting worldviews [75].

Attractors in Psychological Systems

Psychological systems are characterized by intrinsic dynamics, but they also demonstrate stability and resistance to change. Each day, people encounter vast amounts of information relevant to social judgment and interpersonal relations, much of which is mutually contradictory. People nonetheless manage to form and maintain coherent patterns of thought and behavior in their social lives. The partners to a romantic relationship, for example, experience a wide variety of thoughts and feelings about one another, but over time each partner's mental state tends to converge on positive sentiment toward the other. Thus, despite the ever-changing nature of intrapersonal and interpersonal experience, people's mental, emotional, and behavioral states tend to converge on relatively narrow sets of specific states or on patterns of change between specific states. These states or patterns of change are referred to as *attractors*.

When a system is at its attractor, it tends to maintain that state despite potentially destabilizing forces and influences. An external influence may move the system to another state, but the system will return fairly quickly to one of its attractors. Several social psychological phenomena imply the existence of an attractor. For example, self-regulation is defined in terms of resistance to temptations and distractions, impulse control, and the maintenance of states representing salient personal standards and values [14,15,105]. In the same fashion, self-esteem maintenance [98], self-verification [95], and psychological reactance [8] all reflect the tendency of mental systems to converge on a particular state (e. g., a level of self-esteem) and to resist influences that threaten to dislodge the person's judgments and beliefs from that state.

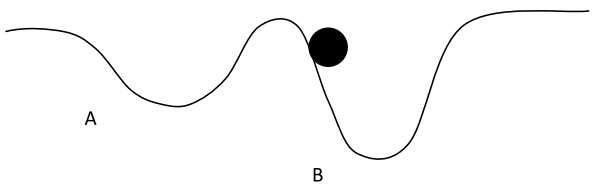
Three basic types of attractors are commonly distinguished: fixed-point attractors, periodic (including multiperiodic) attractors, and deterministic chaos (cf. [25,73,89]). Each type is likely to have relevance for understanding different intrapersonal and interpersonal processes, although work designed to establish such relevance is in its nascent stage. A system can be characterized by another set of points that have the opposite effect of attractors. Termed *repellors*, they represent unstable equilibria of the system, so that the smallest departure from their exact values will result in the system rapidly escaping

from the region surrounding that value. Whereas attractors may be described as states that the system "seeks" over long periods of time, repellors reflect states that the system selectively "avoids." The repellor concept has been used to describe how neural networks avoid particular memories [105].

Fixed-Point Attractors

A *fixed-point attractor* is similar to the notion of equilibrium or homeostasis [12,66]. It represents the case in which the state of the system converges to a stable value. A desired end-state or goal [14,15,104,106] for example, can be described as a fixed-point attractor. This is apparent when a person maintains a belief, an interpersonal evaluation, or an action tendency despite forces or sources of information that challenge these tendencies. Attractors, however, are not limited to goals, intentions, or other desired states. Thus, a person might display a pattern of antagonistic behavior in his or her social relations, despite efforts to avoid behaving in this manner. Similarly, a person with low self-esteem may initially embrace flattering feedback from someone, but over time the person may discount or reinterpret this feedback, displaying instead a pattern of self-evaluative thought that converges on a negative state [96]. In an inter-group context, meanwhile, warring factions may display conciliatory gestures when prompted to do so, but revert to a pattern of antagonistic thought and behavior when the outside interventions are relaxed [19]. In short, when a system's dynamics are governed by a fixed-point attractor, the system will consistently evolve to a particular state, even if this state is not hedonically pleasant, and will return to this state despite being perturbed by forces that might promote a more pleasant state.

A psychological system may have more than one attractor, with each corresponding to a distinct equilibrium. Which attractor governs a system's dynamics in a particular instance depends on the starting values of the system's evolution. The set of initial states that converge on each attractor represents the *basin of attraction* for that attractor. For a person or a group characterized by multiple fixed-point attractors, then, the process in question can display different equilibrium tendencies, each associated with a distinct basin of attraction. Within each basin, different initial states will eventually converge on the same stable value. However, even a slight change in the system's initial state will promote a large change in the system's trajectory if this change represents a state that falls just outside the original basin of attraction and within a basin for a different attractor. For example, in conflict situations there are typically two dominant responses, one corresponding to aggression and one corresponding to conciliation. Slight differences in the circumstances surrounding the conflict can thus promote dramatically different behaviors, with no option for a response that integrates the two tendencies (e. g. [19]).



A dynamical system with two fixed-point attractors (A and B)

Social Psychology, Applications of Complexity to, Figure 1
A dynamical system with two fixed-point attractors [106]

The attractor concept can be captured in a simple metaphor. Figure 1 shows a ball on a hilly landscape. The ball represents the current state of the system and the valleys (A and B) represent different fixed-point attractors. The evolution of the system toward an attractor is characterized by the ball rolling down a hill and coming to rest in the bottom of a valley. Each attractor in Fig. 1 has a basin of attraction, represented by the width of the valley. The basin of attraction for Attractor A in Fig. 1 is wider than the basin for Attractor B, which means that a wider range of states will evolve toward Attractor A. Attractors can also vary in their respective strength, which is represented by the corresponding depth of the two valleys. Attractor B, then, is stronger than Attractor A. Hence, it is more difficult for external influence to change the current state of the system when the system is within the basin of attraction for Attractor B as opposed to Attractor A.

The existence of multiple fixed-point attractors in a system captures the intuition that people may have different (even mutually contradictory) goals, self-concepts, recurring emotional states, and patterns of social behavior. Someone may have two or more standards for self-regulation, for example, with each providing for action guidance under different sets of conditions. The person's behavior may reflect an affiliation standard under one set of conditions, but reflect an achievement standard under a different set of conditions. In similar fashion, a person may have multiple self-views (e. g. [37,60,61,93]), each representing an integrated and stable way of thinking about him or herself. One self-view is likely to become salient when a specific set of self-relevant information is made salient by virtue of context or role expectations. Characteristic patterns of emotional response, meanwhile, can be viewed as distinct attractors, with each attractor reflecting an "emotional Gestalt" that provides coherence for a person's cog-

nitive-affective dynamics under a unique set of conditions and constraints [57,99]. More generally, apparent inconsistency in personality can be viewed as the existence of multiple attractors, each associated with a different basin of attraction for thought, feeling, and action [82,86,92]. One set of conditions might promote a trajectory that evolves toward dominance and competition, for instance, whereas another might promote empathy and cooperation.

The strength of an attractor and the size of its basin of attraction are potentially independent, and different combinations of these basic properties may have unique implications for psychological processes [82,83]. Note again the two attractors in Fig. 1. Attractor A is relatively weak but has a wide basin of attraction. Thus, a relatively small force may change the state of the system (i. e., move the ball up the gradual slope), but the system is likely to return to the attractor (i. e., it will roll back into the valley) even if these changes are relatively large. In contrast, Attractor B has a narrower basin of attraction but is relatively strong. A great deal of force is required to promote even a slight impact on the system (i. e., move the ball up the steep slope), but if this effect is achieved, the system will not return to the attractor (i. e., it will escape the valley).

As an example, consider a romantic couple that has two attractors: a strong attractor associated with positive feelings and a weak attractor associated with negative feelings. Assume the couple has a wider range of attraction for positive feelings than for negative feelings. The partners are likely to evolve toward positive feelings about one another if they begin an interaction within a broad range of affective states (e. g., neutral to very positive), but they may end up feeling negative if they begin an interaction within a different (more restricted) range of affective states (e. g., mildly to highly negative). A broader range of initial states are likely to promote a communication trajectory that results in an exchange of warm sentiments as opposed to critical comments. If, however, the couple typically starts out with negative feelings, the negative attractor, despite having a narrow basin, may dictate the trajectory for feelings expressed in the couple's interactions. It is conceivable, of course, that the couple has a wider basin of attraction for negative feelings, so that anything short of a highly positive initial state will dissolve into a negatively toned exchange (see [33,34]).

Latent Attractors

A system may have multiple attractors, but when the system is at one of them, the others may not be visible to observers, or perhaps not even to the actors themselves. The existence of these potential states might not even be suspected. Such *latent attractors* may be highly important in the long run, though, because they define which states are possible for the system when conditions change. By identifying possibilities for a system that have yet to be experienced, the concept of latent attractor goes beyond the traditional notion of equilibrium (e. g. [2]). Critical changes in a system might not be reflected in the manifest state of the system, but rather in the creation or destruction of a latent attractor representing a potential state that is currently invisible.

The implications of latent attractors have been explored in the context of inter-group characterized by seemingly intractable conflict [18,19,85]. Factors such as objectification, dehumanization, and stereotyping of outgroup members are preconditions for the development of intractable conflict [17,22] but their immediate impact may not be apparent. Rather, these factors may create a latent attractor to which the system can abruptly switch in response to a provocation that is relatively minor, even trivial. By the same token, efforts at conflict resolution that seem fruitless in the short run may have the effect of creating a latent positive attractor for inter-group relations, thus establishing a potential relationship to which the groups can switch if other conditions permit. In this case, the existence of a latent positive attractor might promote a rapid de-escalation of conflict, even between groups with a long history of and seemingly intractable conflict.

Periodic Attractors

Rather than converging on a stable value, systems can display sustained rhythmic or oscillatory behavior. Such a temporal pattern is referred to as a periodic or limit-cycle attractor. Periodicity is obviously associated with many biological phenomena, such as circadian rhythms and menstrual cycles [30], but this dynamic tendency may also underlie certain psychological phenomena [32]. Moods have been shown to have a periodic structure, for example, often corresponding to a weekly cycle (e. g. [9,50,51]). Research on the intrinsic dynamics of both social judgment [109] and self-evaluation [110], meanwhile, has demonstrated that the stream of thought often oscillates between positive and negative assessments, sometimes in accordance with remarkably fast time scales. Periodic structure also characterizes human action [69] and is a feature of social interaction as well (e. g. [6,32,70]).

Distinguishing a periodic attractor from the existence of multiple fixed-point attractors can prove difficult, since the system in both cases displays movement between different states over time. The difference centers on the reg-

ularity of the movement between states and the role of external factors in producing such movement. A periodic attractor reflects a repetitive temporal pattern, such that the values of the dynamical variable repeat after a time $T, x_i(t) = (t + T)$, where T is the period of motion. The state of the system undergoes constant change even in the absence of noise or external influence. To qualify as a periodic attractor, then, a pattern of change must represent a pattern on which the system converges, and to which it returns after small perturbations. In a daily activity cycle, for example, a sleepless night might temporarily disrupt the pattern (e. g., oversleeping the next few days), but eventually the pattern will be restored.

A system with fixed-point attractors, in contrast, tends to stabilize on a particular state or set of states. Such attractors capture all trajectories within their respective basins, so a disturbance, noise, or an external influence is necessary to move the system from one stable state to another. A person with self-regulatory standards for both compromise and confrontation, for example, will display one of these tendencies as long as the action context is within the basin of attraction for that tendency. If the attractors differ in the size of their respective basins, and if contexts are avoided that attract the person's mental, emotional, or behavioral state toward the smaller basin, the person may behave for long periods of time in line with the stronger attractor. In similar fashion, a romantic couple may have fixed-point attractors for both positive and negative affective states, but whether they display periodic movement between them will depend on the starting conditions associated with their interactions. Even if the couple oscillates between positive and negative states, each state provides at least temporary stability. In periodic evolution, stability is not provided by any particular state, but rather by the pattern of changes between states.

The distinction between fixed-point and periodic attractors was observed in a study investigating the temporal trajectories of affective states on the part of bipolar depressive individuals [42]. Time-series analysis of mood and other symptoms revealed that many of these patients oscillated between a normal and a depressed state. However, patients whose temporal dynamics did not reflect fixed-point attractor tendencies were at highest risk for suicide and were hospitalized more often for their depression. These risks were low for individuals whose moods oscillated around a single attractor, even one corresponding to a depressed state, and for those whose moods switched between two distinct attractors reflecting a normal state and a depressed state. These results imply an interesting connection between attractor dynamics and self-regulatory tendencies. Self-regulation implies stabilization with respect to some states and de-stabilization of other states. The stable states reflect fixed-point attractors for a person's mental and emotional dynamics. The lack of fixed-point attractors for one's internal state signals a breakdown in the capacity for self-regulation.

Deterministic Chaos

The best known phenomena concerning nonlinear dynamical systems centers on *deterministic chaos* (cf. [89]). Many researchers and scholars, especially those from fields other than mathematics and physics, commonly discuss the primary insights from the work on nonlinear dynamics as chaos theory. When investigating a chaotic system, anything less than infinite precision in the knowledge of a system at one point in time can undermine prediction of the system's future states. This decoupling between determinism and practical predictability happens because all initial inaccuracies are amplified by the system's intrinsic dynamics, so that the inaccuracies grow exponentially over time. After some time, exponential growth assures that the size of the error will exceed the possible range of states of the system's behavior.

Chaos clearly represents a possibility in nonlinear dynamical systems [31,73], and has been demonstrated in many biological and physical phenomena. In principle, then, human thought and behavior may sometimes follow a chaotic trajectory. Despite this potential, though, unequivocal evidence for deterministic chaos in human thought and behavior remains to be documented. Human dynamics always contain some degree of randomness and human behavior is often unpredictable. It can be quite difficult, however, to determine the degree to which such unpredictability reflects deterministic chaos, the stochastic nature of the laws governing human nature, or the multitude of influences unaccounted for by measurement that can be treated as noise.

Dynamical Minimalism

In canonical social psychology, the complexity of human thought and behavior is typically assumed to reflect complex interactions among a large number of variables in the traditional approach to theory construction. The approach of *dynamical minimalism* [72], in contrast, assumes that complex properties can result from simple rules specifying the interactions among very simple elements. The emergence of complexity occurs when the elements are nonlinear and interact over time [39]. For example, complex cognitive phenomena (e. g., pattern recognition, error correction), have been observed in a simple network of binary elements, where each element reacts to the input it

receives from other elements [40]. In this spirit, dynamical minimalism attempts to identify the simplest set of assumptions capable of producing a phenomenon of interest. In formal models, this is equivalent to identifying the simplest mathematical rules to express what is known about a phenomenon. The goal of dynamical minimalism, then, is to achieve parsimony in theory construction without stripping the phenomenon of its subtlety and nuance.

This approach provides a new perspective on the relation between micro and macro levels of description. Reductionism is commonly assumed in mainstream models, such that the rules observed at one level of description correspond to the rules operating at another level. In effect, the properties at a macro level of description are reduced to the properties of elements at a micro level. One might, for example, explain the relation between poverty and crime in a social system by reducing this relation to the relation between frustration and crime at the level of individuals in the society. Dynamical models, however, do not assume isomorphism among levels of description. To the contrary, the rules specifying the interaction among a system's elements are likely to generate very different rules at higher levels of system behavior (see [24] for an early appreciation of this idea).

The emergence of new properties at a macro level is illustrated in the *society of self* model of self-structure [81]. This model assumes very simple rules by which self-relevant information is integrated in forming a self-concept. Each element of information (e. g., an episodic memory, a physical feature, a self-perceived trait) adopts the prevailing valence (positive vs. negative) of related elements. This simple rule, when iterated over time, generated interesting but largely unexpected consequences at the global level of self-understanding. The self-structure became differentiated into locally coherent regions (e. g., social roles, areas of competence), each of which displayed resistance to discrepant information (e. g., negative social feedback). Global self-esteem and high self-concept certainty also emerged at the macro level from the simple rule of influence among elements at the micro level.

Emergence seems to represent a paradox for theory construction. How can knowledge of the lower-level elements provide an explanation of the higher-level properties if properties at a macro level cannot be derived from properties of the system's lower level elements? The role of computer simulations in dynamical minimalism helps to resolve this paradox. With computer models, one can specify the properties of system elements and the rules of interaction among these elements. As the elements interact in accordance with these rules over time, dynamics appear at the system level that were not assumed or programmed for the elements themselves. With computer simulation, then, a theory constructed at a basic level of psychological reality (e. g., moment-to-moment thought process, dyadic social interaction) can be tested at a higher level of psychological reality (e. g., social judgment, group norms).

Computer simulations play another important role in dynamical minimalism. The basic elements comprising a system are often uninteresting and trivial, and the interactions among them may have only minor impact on the systems' global properties. But some properties of system elements may have an important impact on the system's higher-order properties as the elements interact over time. It may not be obvious which properties are trivial and which are essential for the emergence process. Dynamical minimalism makes this distinction and thus constructs a model that incorporates only the properties that are critical for the emergence of macro level phenomena. Computer simulations enable one to systematically vary the assumptions regarding the properties of elements and their interactions, and then observe which assumptions promote meaningful changes at the macro level. Those properties that have trivial consequences at the macro level are eliminated from the model. In short, computer simulations enable one to distill the minimal set of components necessary to capture the essence of a phenomenon.

Computer simulations cannot substitute for empirical verification of a theoretical model. They are critical in identifying the properties that are central to the model and investigating the consequences of these properties for the functioning of the system in question. Knowledge of these consequences then provides the basis for framing hypotheses to be investigated in empirical research. The relationship between computer simulations and empirical research work in the other direction as well. Empirical studies can refine a model, which is then implemented in computer simulations. The results of the simulations, in turn, may generate new hypotheses to be tested in subsequent empirical research. The reciprocal feedback loops among theory, computer simulation, and empirical research is central to the approach of dynamical minimalism.

The Dynamics of Social Influence

Social influence refers to any change in an individual's thoughts, feelings, or behavior that occurs as the result of the real or imagined presence of others [3]. It is widely considered to be the core process of social experience [111] and is manifest in a wide array of phenomena, including obedience to authority, conformity, imitation and modeling, bystander intervention in emergencies, social loaf-

ing, stage fright, persuasion, and groupthink. There is evidence, however, that the essence of social influence reflects the interplay of three basic factors: the number, strength, and immediacy of the sources of influence [52]. Empirical research has shown that the magnitude and nature of social influence represents a multiplicative function of these sources.

Nowak et al. [76] modeled the dynamics of social influence using cellular automata. In this approach, each individual is represented with three properties: an opinion on a topic, a degree of persuasive strength, and a position in a social space. Individuals are commonly assumed to have one of two opinions on an issue (e. g., pro vs. con). The group consists of n individuals located on a two-dimensional grid, with each cell corresponding to an individual (see Fig. 2). The color of each cell specifies that individual's current opinion (light gray denotes pro, dark gray denotes con), whereas the height of the cell represents the individual's strength (e. g., expertise, confidence, charisma). Each individual discusses the issue with other group members ascertains the degree of support for each position. As a result of these assessments, each individual adopts the opinion that is most prevalent. The following formula expresses the strength of influence of each opinion:

$$I_i = \left(\sum_{1}^{N} \left(\frac{s_j}{d_{ij}^2} \right)^2 \right)^{1/2} ,$$

where I_i denotes total influence, s_j represents the strength of each individual, and d_{ij} represents the distance between individuals i and j. The opinions of those who are closest to the individual and have the greatest strength are weighted most heavily by the individual. An individual's own position is also considered and is weighted most heavily because of its immediacy (0 distance). Influence grows with the square root of the number of people exerting influence.

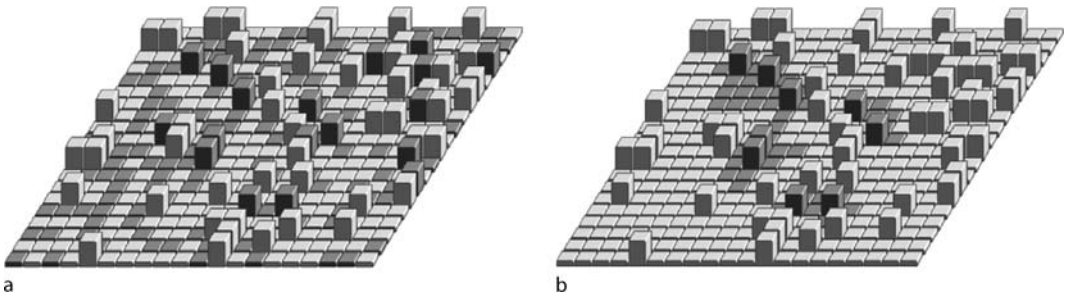
In each round of the simulations, one individual is chosen at random and influence is computed for each opinion in the group. A simple updating rule is employed: the individual changes his or her opinion to match the prevailing opinion if the resultant strength of this opinion position is greater than the strength of the individual's current position. This basic process is performed for each individual in the group. This procedure is repeated until there are no further changes in opinion. This typically involves several rounds of simulation, because an individual who had previously changed his or her position to match that of his or her neighbors may revert to the original position if the neighbors change their opinions.

Figure 2 depicts representative results of the computer simulations. In the initial configuration (Fig. 2a), there is a majority of 60% (light gray) and a minority of 40% (dark gray), with the majority and minority opinions randomly distributed in social space. The majority and minority groups have the same relative proportions of strong and weak members (tall vs. short cells). Figure 2b shows that an equilibrium is reached after six rounds of simulated discussion. Although the majority has grown (to 90%) at the expense of the minority (now 10%), the minority opinion has survived by forming clusters of like-minded people. Note that these clusters are primarily formed around strong individuals.

These two group-level outcomes—polarization and clustering—are routinely observed in computer simulations of this process [55]. Each is reminiscent of well-documented social processes. Research on group dynamics (e. g. [68]), for example, has shown that the average attitude in a group becomes polarized in the direction of the prevailing attitude as a result of group discussion. In the model, polarization reflects the greater influence of the majority opinion. In the initial (random) configuration (Fig. 2a) the average proportion of neighbors holding a particular opinion (pro or con) reflects the proportion of this opinion in the total group. The average group member, in other words, is surrounded by more majority than minority members, so more minority members are converted to the majority position than vice versa. However, some majority members are converted to the minority position because they happen to be located close to an especially strong minority member or because more minority members happen to be at this location.

Clustering is also a pervasive feature of social life, having been documented for such diverse facets of social life as attitudes, political beliefs, religions, clothing fashions, and farming techniques. It has been show, for example, that attitudes tend to cluster in residential neighborhoods [28]. When opinions are distributed randomly, the sampling of opinions through social interaction provides a realistic portrayal of the distribution of opinions in the larger society. However, when opinions are clustered, the same sampling process produces a very biased result because the opinions of one's nearby neighbors are weighted the most heavily. The prevalence of one's own opinion is therefore likely to be over-estimated. Opinions that are in the minority in global terms, then, can form a local majority. This process enables individuals with a minority opinion to maintain this opinion in the belief that it actually represents a majority position.

In the spirit of dynamical minimalism, this model focuses on the minimal set of processes responsible for the emergence of group-level properties that are invariant across diverse areas of topical interest. In so doing, it pro-



Social Psychology, Applications of Complexity to, Figure 2
a Initial distribution of opinions in the simulated group. b Final equilibrium of opinions in the simulated group [106]

vides a platform for investigating how group-level properties emerge for different domains of social functioning. This potential has been developed in recent years. Kenrick and his colleagues, for example, have simulated the emergence of cultural norms from the mutual influence among individuals with conflicting strategies (decision-rules) pertaining to fundamental human goals (e. g. [49]). These goals, reflecting domains of adaptive functioning confronted by human groups throughout history [10,11,29], include self-protection, coalition formation, status-seeking, mate choice, relationship maintenance, and offspring care.

In their computer simulations, individuals with different decision-rules regarding a particular domain interacted with one another over time, and the group-level consequences of these interactions were observed. One series of simulations investigated how individual differences in decision-rules for cooperation versus competition affected community level propensities for such behavior. Another series explored the mutual impact of male and female mating strategies on the emergence of societal-level norms regarding mating. Clustering was observed in both cases, with an initial random configuration of decision rules giving way over time to local communities characterized by coherent norms regarding the behavior in question. In the mating simulations, for example, some communities were characterized by relatively unrestricted mating strategies (reflecting the decision-rule preferred by males), whereas other communities developed norms sanctioning restricted mating strategies (reflecting the decision-rule preferred by females).

Dynamics of Interpersonal Coordination

The dynamical account of social influence provides a concise description of how the state (e. g., attitude) of a single individual depends on the state of other individuals. Because many psychological processes reflect intrinsic dynamics, however, individuals can be conceptualized as displaying patterns of change rather than as a set of states. In this view, social influence (and social interaction generally) can be investigated as the coordination over time of individual dynamics.

The most basic form of coordination is positive correlation or in-phase relation. In social interaction, this occurs when the overt behaviors, attitudes, or emotions of one person induce similar behaviors or states in the other person at the same time (e. g. [16]). This basic form of coordination is reflected in such familiar phenomena as imitation, mimicry, and empathy [62]. Other forms of coordination can be identified, however, with counterparts in different contexts for social interaction [62,69]. Turn taking in conversation represents negative correlation or anti-phase relation between individuals in their respective talking and listening (when one person speaks, the other is silent). Negative correlation also characterizes antagonistic relationships, in that the sadness or despair of one person induces satisfaction or happiness in the other person and vice versa. More complex forms of synchronization can also be identified that reflect nonlinear relationships and higher-order interactions between the partners' respective behaviors and internal states (cf. [74,82]).

A Model of Synchronization Dynamics

This perspective provides the foundation for a recently developed model of *synchronization* [82,83], a phenomenon that characterizes coupled dynamical systems [44,91]. Because positive correlation represents the most fundamental and common form of coordination, it has provided the primary focus to date in this class of models. The Nowak et al. [82,83] model assumes that each individual attempts to achieve synchronization by adjusting his or her internal state or overt behavior in response to the state or behavior of the individual with whom he or she is interacting. Individuals in social interaction, in other words, modify their respective thoughts, feelings, or action tendencies to promote positive correlation over time in these features

of experience. The synchronization of individuals' dynamics results in a higher order system with its own dynamic properties.

Coupled logistic are used to model interpersonal synchronization [74,81,82,83]. The dynamics of each individual are represented with a logistic equation, which is the simplest dynamical system capable of displaying complex (e. g., chaotic) behavior [27,89]. However, the behavior of each person not only depends on his or her preceding state but also to a certain extent on the preceding state of the other person. The coupling of individuals' dynamics is specified in the following equation:

$$x_1(t+1) = \frac{r_1 x_1(t)(1 - x_1(t)) + \alpha r_2 x_2(t)(1 - x_2(t))}{1 + \alpha}$$
$$x_2(t+1) = \frac{r_2 x_2(t)(1 - x_2(t)) + \alpha r_1 x_1(t)(1 - x_1(t))}{1 + \alpha} .$$

The dynamical variable (x) represents the intensity of behavior, and the control parameter, r, corresponds to internal states (e. g., personality traits, moods, values) that shape the person's pattern of behavior (i. e., changes in x over time). To the value of the dynamical variable representing one individual's behavior x_1, one adds a fraction α of the value of the dynamical variable representing the other individual's behavior x_2. The magnitude of α represents the strength of coupling and can be viewed as the degree of mutual influence characterizing the interaction. Depending on the social context, it might reflect the intensity of communication or the degree of mutual imitation. When α is 0, there is no coupling (e. g., no influence or communication) on the behavior level, whereas when α is 1, each individual's behavior is determined equally by his or her preceding behavior and the influence of the other individual. Intermediate values of α represent moderate values of coupling in the relationship.

Modeling Behavioral Synchronization

The respective control parameters of two individuals are unlikely to be identical when they first interact with one another. And the degree of influence between individuals differs across interactions and relationships. To represent this variability in social reality, Nowak, Vallacher [74] systematically varied the similarity of partners' control parameters (r), representing their internal states, and their degree of coupling (α), representing their mutual influence (e. g., communication, imitation). Each simulation began with a random value of x for each individual, drawn from a uniform distribution that varied from 0 to 1. They let the simulations run for 300 steps, allowing each system to converge on its pattern of intrinsic dynamics and both systems to synchronize. For the next 500 simulation steps, they recorded the values of x for each system and the degree of synchronization of the two systems.

The results demonstrated that behavioral synchronization increased both with increases in α and in the similarity in r. At each degree of coupling, synchronization increased with greater similarity in partner's internal states. Likewise, at each level of similarity in internal states, synchronization increased with increased strength of coupling. These results have straightforward implications. If two people have similar control parameters (internal states), relatively little coupling (e. g., influence, communication, mutual reinforcement, monitoring) is necessary for them to achieve a high degree of synchronization in their behavior. If the partners have different internal states, on the other hand, high mutual influence is required to maintain the same level of synchronization. In a close relationship, this suggests that constant and intense communication may be a sign that the partners are not well coordinated with respect to relevant internal states (e. g., temperament, desires values). But when the partners are similar with respect to such internal states, they can devote their energy to common pursuits rather than to constant clarification, monitoring, and other forms of influence.

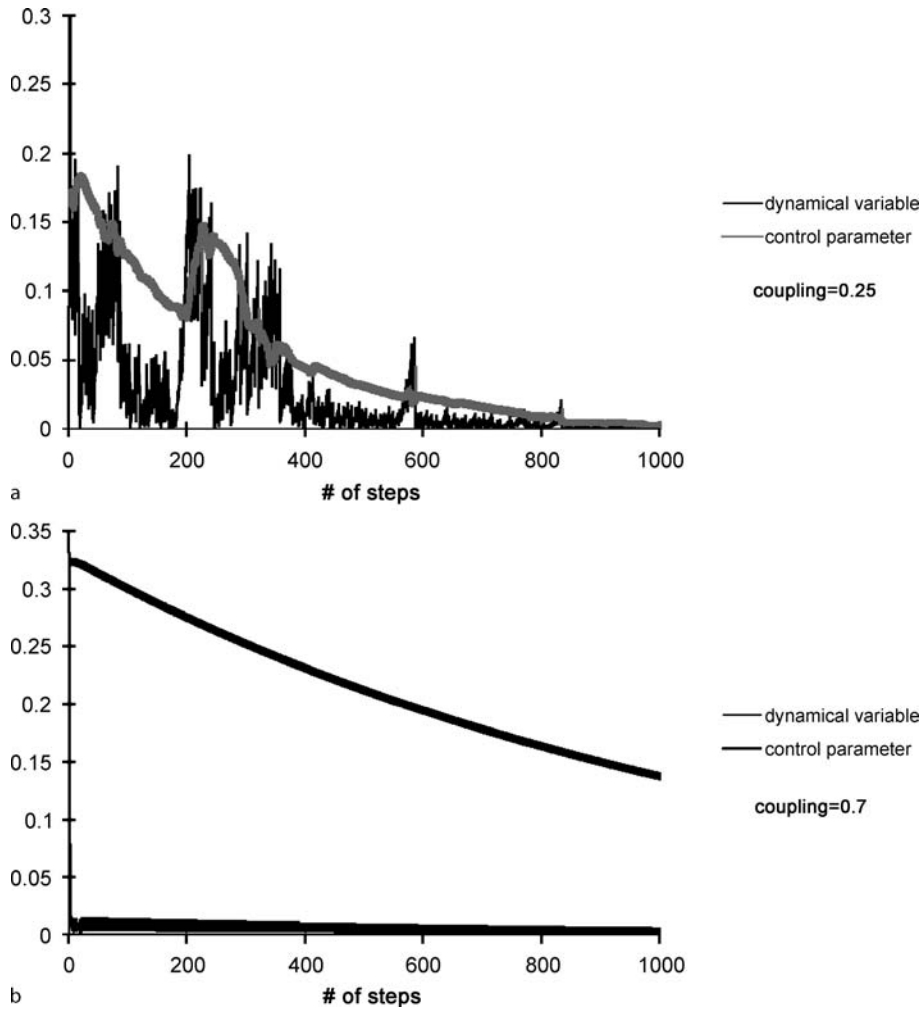
Modeling Internal Synchronization

Some internal parameters, such as moods and emotional states, vary considerably across time and settings. If a particular context induces a common mood or emotion (e. g., joy or sadness) in a dyad or group, interpersonal synchronization is easy to achieve. Other internal states, however, demonstrate greater stability and less likely to vary across different contexts. Attitudes, values, and personality traits, in particular, are commonly considered to be enduring (cross-situational) features of a person's psychological make-up. Yet, even these internal parameters admit to variability and even modification. The Nowak et al. [82,83] model has been used to understand and investigate the nature of such change. The core assumption is that individuals are motivated to achieve coordination in their internal parameters [13,62]. To satisfy this motive, individuals vary their internal parameters in a direction that leads to increasing synchronization. When synchronization is achieved, the value of the resultant control parameter is engraved as an attractor for that internal state. In essence, people are assumed to develop stable internal states through social synchronization (e. g. [112]).

To model this process, Nowak et al. simply assumed that the value of each individual's control parameter drifts somewhat in the direction of the value of the other indi-

vidual's control parameter on each simulation step. How quickly the respective control parameters begin to match depends on the size of the initial discrepancy between them and on the rate at which they drift. This process occurs with each individual knowing the exact value of the other individual's control parameter. This is an important feature of the model, since experimental research has shown that people's internal states are often difficult to infer [43,47,71,113]. The model assumes only that each individual remembers the other individual's most recent behaviors (i. e., the most recent values of x) as well as his or her own most recent behaviors. Each individual compares his or her own behavior with that of the other individual, and then adjusts his or her internal parameter in order to promote increased similarity with the other person's behavior pattern, until a match is achieved [115]. Thus, if the other individual's behavior is more complex than the individual's own pattern of behavior, the individual slightly increases the value of his or her own control parameter. Conversely, the individual slightly decreases the value of his or her own control parameter if the other individual's behavior is less complex than his or her own. In short, interacting individuals estimate one another's internal state by monitoring the evolution of one another's behavior.

Nowak et al. [82,83] ran simulations to investigate the convergence of behavior and internal states under both relatively weak and relatively strong coupling ($\alpha = .25$ and .7, respectively). The y-axis in Fig. 3 corresponds to the magnitude of difference between the two systems in their behavior or internal states, and the x-axis corre-



Social Psychology, Applications of Complexity to, Figure 3
a Convergence of behavior and internal states under weak coupling b Convergence of behavior and internal states under strong coupling [82]

sponds to time, as reflected in the number of iterations. Consider first the results observed when the coupling was weak (Fig. 3a). The behavior of the two systems converged in a relatively slow and non-linear manner. The control parameters also showed a clear tendency towards convergence. When a match in internal states was achieved, moreover, full synchronization of behavior was also obtained. Compare these results with those observed under strong coupling (Fig. 3b). Note that although there was immediate synchronization of behavior, the control parameters failed to synchronize, even after 1000 iterations.

The differential results obtained for weak and strong coupling may have noteworthy implications for interpersonal relations. First, even for people with very different internal parameters, strong coupling tends to promote full synchronization of behavior. Once synchronization is achieved, the two people may be totally unaware that their internal states are different. This suggests that if the coupling were to be reduced in magnitude (or removed altogether), the dynamics of the two individuals would immediately diverge. Hence, people who employ very strong influence (e. g., reinforcement, monitoring) to obtain coordination of behavior may effectively hinder synchronization of their respective internal parameters. In more general terms, there may be an optimal level of influence and control over one another's behavior in interpersonal relations [111]. When influence is too weak, synchronization between individuals may fail to develop. But when influence is very strong, it can prevent the development of a relationship based on mutual understanding and empathy. Intermediate levels of mutual influence, then, may be most effective for the development of synchronization on a deep level. Stated differently, the most advantageous degree of coupling (e. g., influence) is the minimal amount necessary to achieve synchronization.

These implications are consistent with extensive research in social psychology suggesting that behavior attributed to external causes is less likely to promote psychological change than is behavior attributed to internal causes. For example, people are resistant to changing their preferences and attitudes if they believe that their behavior is in response to direct orders, rewards, threats, and other external influences [7,56]. Salient external influences, in fact, may activate mechanisms to counter the influences, creating an internal state that is opposite of the intended effect of the influence [8,111].

Future Directions

The application of complexity science to social psychology makes for an ironic discipline. The tools of complexity and dynamical systems are ideally suited to capture the subtlety and uniqueness of human experience, yet they are grounded in concepts and methods that provide meaningful integration with the natural sciences. Disciplines such as mathematics, physics, and chemistry have affirmed the significance of intrinsic dynamics, nonlinear phenomena, self-organization, and complexity. In recent years, psychologists have become increasingly cognizant that these features of systems in nature have counterparts in mental, interpersonal, and collective experience. It would be odd for a contemporary social psychologist to discount the potential for emergence or to ignore the role of computer simulations and time series in illuminating how minds, groups, and societies work.

A word of caution is in order, however. Social reality is not the same as physical reality. Unlike atoms, individuals are not interchangeable, and groups are more than self-organized ensembles of simple particles. People have values and beliefs, universal concerns and idiosyncratic tendencies, and moments of self-reflection and sudden impulse. One of the basic rules of human psychology, moreover, is the capacity for reflecting on one's operating rules and attempting to override them. People exist in a symbolically constructed world and do not respond in a reflexive way to objective reality. These unique features of human experience cannot be chalked up to the recognition that people are dynamic and complex. The task of dynamical social psychology is thus more daunting than discovering the differential equations that govern interpersonal and collective dynamics. Models derived from complexity and dynamical systems provide a foundational science for the discipline and much remains to be discovered within this new paradigm. Ultimately, though, the properties that separate us from other systems in nature must be incorporated into theoretical models. It is an ironic testament to the complexity perspective that a coherent theory of social psychology should be assembled from elements that are both universal within nature and unique to human experience.

Bibliography

Primary Literature

1. Abelson RP (1979) Social clusters and opinion clusters. In: Holland PW, Leinhardt S (eds) Perspectives in social network research. Academic Press, New York, pp 239–256
2. Abraham RH, Shaw CD (1992) Dynamics, the geometry of behavior, 2nd edn. Addison-Wesley, Reading
3. Allport GW (1968) The historical background of modern social psychology. In: Lindzey GA, Aronson E (eds) The Handbook of Social Psychology, vol 1. Addison-Wesley, Reading

4. Asch SE (1946) Forming impressions of personalities. J Abnorm Soc Psychol 41:258–290
5. Baron RM, Amazeen PM, Beek PJ (1994) Local and global dynamics of social relations. In: Vallacher RR, Nowak A (eds) Dynamical systems in social psychology. Academic Press, San Diego, pp 111–138
6. Beek PJ, Hopkins B (1992) Four requirements for a dynamical systems approach to the development of social coordination. Huma Mov Sci 11:425–442
7. Bem DJ (1967) Self-perception: An alternative interpretation of cognitive dissonance phenomena. Psychol Rev 74:183–200
8. Brehm SS, Brehm JW (1981) Psychological reactance: A theory of freedom and control. Academic Press, New York
9. Brown KW, Moskowitz DS (1998) Dynamic stability of behavior: The rhythms of our interpersonal lives. J Personal 66:105–134
10. Bugental DB (2000) Acquisition of the algorithms of social life: A domain-base approach. Psychol Bull 126:187–219
11. Buss DM (1999) Evolutionary psychology: The new science of the mind. Allyn Bacon, Boston
12. Cannon WB (1932) The wisdom of the body. Norton, New York
13. Caporeal LR, Baron RM (1997) Groups as the mind's natural environment. In: Simpson JA (ed) Evolutionary social psychology. Lawrence Erlbaum Associates, Hillsdale, pp 317–344
14. Carver CS, Scheier MF (1999) Themes and issues in the self-regulation of behavior. In: Wyer RS Jr (ed) Advances in social cognition, vol 12. Erlbaum, Mahwah, pp 1–105
15. Carver CS, Scheier MF (2002) Control processes and self-organization as complementary principles underlying behavior. Personal Soc Psychol Rev 6:304–315
16. Chartrand TL, Bargh JA (1999) The chameleon effect: The perception-action link and social interaction. J Pers Soc Psychol 76:893–910
17. Coleman PT (2003) Characteristics of protracted, intractable conflict: Towards the development of a meta-framework - I. First paper in a three-paper series. Peace Confl: J Peace Psychol 9:1–37
18. Coleman PT, Bui-Wrzosinska L, Vallacher RR, Nowak A (2006) Protracted conflicts as dynamical systems. In: Schneider AK, Honeyman C (eds) The negotiator's fieldbook: The desk reference for the experience negotiator. American Bar Association Books, Chicago, pp 61–74
19. Coleman PT, Vallacher RR, Nowak A, Bui-Wrzosinska L (2007) Intractable conflict as an attractor: A dynamical systems approach to conflict escalation and intractability. Am Behav Sci 50:1454–1475
20. Condon WS, Ogston WD (1967) A segmentation of behavior. J Psychiatr Res 5:221–235
21. Cooley CH (1902) Human nature and the social order. Scribner, New York
22. Deutsch M (1973) The resolution of conflict: Constructive and destructive processes. Yale University Press, New Haven
23. Dittman AT, Llewellyn LG (1969) Body movement and speech rhythm in social conversation. J Personal Soc Psychol 11:98–106
24. Durkheim E (1938) The rules of sociological method. University of Chicago Press, Chicago
25. Eckmann JP, Ruelle D (1985) Ergodic theory of chaos and strange attractors. Rev Mod Phys 57:617–656
26. Eiser JR (1994) Attitudes, chaos, and the connectionist mind. Blackwell, Oxford
27. Feigenbaum MJ (1978) Quantitative universality for a class of nonlinear transformations. J Stat Phys 19:25–52
28. Festinger L, Schachter S, Back K (1950) Social pressures in informal groups. Stanford University Press, Stanford
29. Fiske AP (1992) The four elementary forms of sociality: Framework for a unified theory of social relations. Psychol Rev 99:689–723
30. Glass L, Mackey MC (1988) From clocks to chaos: The rhythms of life. Princeton University Press, Princeton
31. Goldstein J (1996) Causality and emergence in chaos and complexity theories. In: Sulis W, Combs A (eds) Nonlinear dynamics and human behavior. World Scientific, Singapore, pp 161–190
32. Gottman JM (1979) Detecting cyclicity in social interaction. Psychol Bull 86:338–348
33. Gottman J, Swanson C, Swanson K (2002) A general systems theory of marriage: Nonlinear difference equation modeling of marital interaction. Personal Soc Psychol Rev 4:326–340
34. Gottman JM, Murray JD, Swanson CC, Tyson R, Swanson KR (2002) The Mathematics of Marriage. MIT Press, Cambridge
35. Guastello SJ (1995) Chaos, catastrophe, and human affairs: Applications of nonlinear dynamics to work, organizations, and social evolution. Erlbaum, Mahwah
36. Guastello SJ, Pincus D, Gunderson PR (2006) Electrodermal arousal between participants in a conversation: Nonlinear dynamics and linkage effects. Nonlinear Dyn Psychol Life Sci 10:365–399
37. Higgins ET (1987) Self-discrepancy: A theory relating self and affect. Psychol Rev 94:319–340
38. Hock HS, Kelso JAS, Schoner G (1993) Bistability and hysteresis in the organization of apparent motion pattern. J Exp Psychol Hum Percept Perform 19:63–80
39. Holland JH (1995) Emergence: From chaos to order. Addison-Wesley, Reading
40. Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences USA 79:2554-2558
41. James W (1890) Principles of psychology. Holt, New York
42. Johnson SL, Nowak A (2002) Dynamical patterns in bipolar depression. Personal Soc Psychol Rev 6:380–387
43. Jones EE, Davis KE (1965) From acts to dispositions: The attribution process in person perception. In: Berkowitz L (ed) Advances in experimental social psychology, vol 2. Academic Press, New York, pp 220–266
44. Kaneko K (ed) (1993) Theory and applications of coupled map lattices. World Scientific, Singapore
45. Kaplowitz SA, Fink EL (1992) Dynamics of attitude change. In: Levine RL, Fitzgerald HE (eds) Analysis of dynamic psychological systems, vol 2. Plenum, New York, pp 341–369
46. Kelley HH (1967) Attribution in social psychology. Neb Symp Motiv 15:419–422
47. Kelley HH (1971) Attribution in social interaction. Morristown
48. Kelso JAS (1995) Dynamic patterns: The self-organization of brain and behavior. The MIT Press, Cambridge
49. Kenrick DT, Maner JK, Butner J, Li NP, Becker DV, Schaller M (2002) Dynamical evolutionary psychology: Mapping the domains of the new interactionist paradigm. Personal Soc Psychol Rev 6:347–356

50. Larsen RJ (1987) The stability of mood variability: A spectral analytic approach to daily mood assessments. J Personal Soc Psychol 52:1195–1204
51. Larsen RJ, Kasimatis M (1990) Individual differences in entrainment of mood to the weekly calendar. J Personal Soc Psychol 58:164–171
52. Latané B (1981) The psychology of social impact. Am Psychol 36:343–356
53. Latané B, Nowak A (1994) Attitudes as catastrophes: From dimensions to categories with increasing involvement. In: Vallacher RR, Nowak A (eds) Dynamical systems in social psychology. Academic Press, San Diego, pp 219–249
54. Latané B, Nowak A (1997) The causes of polarization and clustering in social groups. Prog Commun Sci 13:43–75
55. Latané B, Nowak A, Liu J (1994) Measuring emergent social phenomena: dynamism, polarization and clustering as order parameters of social systems. Behav Sci 39:1–24
56. Lepper MR, Greene D (eds) (1978) The hidden costs of reward. Erlbaum, Hillsdale
57. Lewis MD (2005) Bridging emotion theory and neurobiology through dynamic systems modeling. Behav Brain Sci 28:169–194
58. Lewenstein M, Nowak A, Latané B (1993) Statistical mechanics of social impact. Phys Rev 45:703–716
59. Lewin K (1936) Principles of topological psychology. McGraw-Hill, New York
60. Linville PW (1985) Self-complexity and affective extremity: Don't put all your eggs in one cognitive basket. Soc Cogn 3:94–120
61. Markus H, Nurius P (1986) Possible selves. Am Psychol 41:954–969
62. Marsh KL, Richardson MJ, Baron RM (2006) Contrasting approaches to perceiving and acting with others. Ecol Psychol 18:1–37
63. Mazanov J, Byrne DG (2006) A cusp catastrophe model analysis of changes in adolescent substance use: Assessment of behavioural intention as a bifurcation variable. Nonlinear Dyn Psychol Life Sci 10:445–470
64. McGrath JE, Kelley JR (1986) Time and human interaction: Toward a psychology of time. Guilford Publications, New York
65. Mead GH (1934) Mind, self, and society. University of Chicago Press, Chicago
66. Miller NE (1944) Experimental studies of conflict. In: Hunt JM (ed) Personality and the behavior disorders. Ronald, New York
67. Moscovici S, Lage E, Naffrechoux M (1969) Influence of a consistent minority on responses of a majority in a color perception task. Sociometry 32:365–379
68. Myers DG, Lamm H (1976) The group polarization phenomenon. Psychol Bull 83:602–627
69. Newtson D (1994) The perception and coupling of behavior waves. In: Vallacher RR, Nowak A (eds) Dynamical systems in social psychology. Academic Press, San Diego, pp 139–167
70. Nezlek JB (1993) The stability of social interaction. J Personal Soc Psychol 65:930–941
71. Nisbett R, Ross L (1980) Human inference: Stategies and shortcomings of social judgment. Prentice-Hall, Englewood Cliffs
72. Nowak A (2004) Dynamical minimalism: Why less is more in psychology. Personality and Social Psychology Review 8:183-192
73. Nowak A, Lewenstein M (1994) Dynamical systems: A tool for social psychology? In: Vallacher RR, Nowak A (eds) Dynamical systems in social psychology. Academic Press, San Diego, pp 17–53
74. Nowak A, Vallacher RR (1998) Dynamical social psychology. Guilford, New York
75. Nowak A, Vallacher RR (2001) Societal transition: Toward a dynamical model of social change. In: Wosinska W, Cialdini RB, Barrett DW, Reykowski J (eds) The practice of social influence in multiple cultures. Lawrence Erlbaum, Mahwah, pp 151–171
76. Nowak A, Szamrej J, Latané B (1990) From private attitude to public opinion: A dynamic theory of social impact. Psychol Rev 97:362–376
77. Nowak A, Lewenstein M, Szamrej J (1993) Social transitions occur through bubbles. Sci Am Pol vers 12:16–25
78. Nowak A, Urbaniak J, Zienkowski L (1994) Clustering processes in economic transition. RECESS Res Bull 3:43–61
79. Nowak A, Lewenstein M, Frejlak P (1996) Dynamics of public opinion and social change. In: Hegselman R, Pietgen HO (eds) Modeling social dynamics: Order, chaos, and complexity. Helbin, Vienna, pp 54–78
80. Nowak A, Vallacher RR, Borkowski W (2000) Modeling the temporal coordination of behavior and internal states. In: Ballot G, Weisbuch G (eds) Applications of simulation to the social sciences. Hermes Science Publications, Oxford, pp 67–86
81. Nowak A, Vallacher RR, Tesser A, Borkowski W (2000) Society of self: The emergence of collective properties in self-structure. Psychol Rev 107:39–61
82. Nowak A, Vallacher RR, Zochowski M (2002) The emergence of personality: Personal stability through interpersonal synchronization. In: Cervone D, Mischel W (eds) Advances in personality science. Guilford Press, New York, pp 292–331
83. Nowak A, Vallacher RR, Zochowski M (2005) The emergence of personality: Dynamic foundations of individual variation. Dev Rev 25:351–385
84. Nowak A, Vallacher RR, Kus M, Urbaniak J (2005) The dynamics of societal transition: Modeling nonlinear change in the Polish economic system. Int J Sociol 35:65–88
85. Nowak A, Vallacher RR, Bui-Wrzosinska L, Coleman PT (2007) Attracted to conflict: A dynamical perspective on malignant social relations. In: Golec A, Skarzynska K (eds) Understanding social change: Political psychology in Poland. Nova Science Publishers, Haauppague
86. Read SJ, Miller LC (2002) Virtual personalities: A neural network model of personality. Personal Soc Psychol Rev 6:357–369
87. Rosenblum LD, Turvey MT (1988) Maintenance tendency in coordinated rhythmic movements: Relative fluctuations and phase. Neuroscience 27:289–300
88. Schachter S (1951) Deviation, rejection and communication. J Abnorm Soc Psychol 46:199–207
89. Schuster HG (1984) Deterministic chaos. Physik, Vienna
90. Sherif M (1936) The psychology of social norms. Harper, New York
91. Shinbrot T (1994) Synchronization of coupled maps and stable windows. Phys Rev E 50:3230–3233
92. Shoda Y, LeeTiernan S, Mischel W (2002) Personality as a dynamical system: Emergence of stability and distinctiveness

from intra- and interpersonal interactions. Personal Soc Psychol Revi 6:316–325
93. Showers CJ (1995) The evaluative organization of self-knowledge. In: Kernis MH (ed) Efficacy, agency, and self-esteem. Plenum, New York, pp 101-120
94. Simon D, Holyoak KJ (2002) Structural dynamics of cognition: From consistency theories to constraint satisfaction. Personal Soc Psychol Rev 6:283–294
95. Swann WB Jr (1990) To be adored or to be known? The interplay of self-enhancement and self-verification. In: Higgins ET, Sorrentino RM (eds) Handbook of motivation and cognition: Foundations of social behavior, vol 2. Guilford, New York, pp 408–448
96. Swann WB, Hixon JG, Stein-Seroussi A, Gilbert D (1990) The fleeting gleam of praise: Cognitive processes underlying behavioral reactions to self-relevant feedback. J Personal Soc Psychol 59:17–26
97. Tesser A (1978) Self-generated attitude change. In: Berkowitz L (ed) Advances in experimental social psychology, vol 11. Academic Press, New York, pp 85–117
98. Tesser A, Martin L, Cornell D (1996) On the substitutability of self-protective mechanisms. In: Gollwitze PM, Bargh JA (eds) The psychology of action. Guilford Publications, New York, pp 48–68
99. Thagard P, Nerb J (2002) Emotional Gestalts: Appraisal, change, and the dynamics of affect. Personal Soc Psychol Rev 6:274–282
100. Tickle-Degnen L, Rosenthal R (1987) Group rapport and nonverbal behavior. Rev Personal Soc Psychol 9:113–136
101. Turvey MT (1990) Coordination. Am Psychol 4:938–953
102. Vallacher RR, Nowak A (eds) (1994) Dynamical systems in social psychology. Academic Press, San Diego
103. Vallacher RR, Nowak A (1994) The stream of social judgment. In: Vallacher RR, Nowak A (eds) Dynamical systems in social psychology. Academic Press, San Diego, pp 251–277
104. Vallacher RR, Nowak A (1997) The emergence of dynamical social psychology. Psychol Inq 8:73–99
105. Vallacher RR, Nowak A (1999) The dynamics of self-regulation. In: Wyer RS Jr (ed) Advances in self-regulation, vol 12. Lawrence Erlbaum, Mahwah, pp 241–259
106. Vallacher RR, Nowak A (2007) Dynamical social psychology: Finding order in the flow of human experience. In: Kruglanski AW, Higgins ET (eds) Social psychology: Handbook of basic principles, 2nd edn. Guilford Publications, New York, pp 734–758
107. Vallacher RR, Wegner DM (1987) What do people think they're doing? Action identification and human behavior. Psychol Rev 94:1–15
108. Vallacher RR, Wegner DM (1989) Levels of personal agency: Individual variation in action identification. J Personal Soc Psychol 57:660–671
109. Vallacher RR, Nowak A, Kaufman J (1994) Intrinsic dynamics of social judgment. J Personal Soc Psychol 66:20–34
110. Vallacher RR, Nowak A, Froehlich M, Rockloff M (2002) The dynamics of self-evaluation. Personal Soc Psychol Rev 6:370–379
111. Vallacher RR, Nowak A, Miller ME (2003) Social influence and group dynamics. In: Weiner I, Millon T, Lerner MJ (eds) Handbook of psychology, vol 5. Personality and social psychology. Wiley, New York, pp 383–417
112. Vallacher RR, Nowak A, Zochowski M (2007) Dynamics of social coordination: The synchronization of internal states in close relationships. In: Hauf P, Forsterling F (eds) Making minds: The shaping of human minds through social context. John Benjamins Publishing Company, Amsterdam, pp 31–46
113. Wegner DM, Vallacher RR (1977) Implicit psychology: An introduction to social cognition. Oxford University Press, New York
114. Wegner DM, Vallacher RR, Kiersted G, Dizadji D (1986) Action identification in the emergence of social behavior. Soc Cogn 4:18–38
115. Zochowski M, Liebovitch LS (1999) Self-organizing dynamics of coupled map systems. Phys Rev E 59:2830

Books and Reviews

Guastello S, Koopmans M, Pincus D (eds) (2008) Chaos and Complexity in Psychology: The Theory of Nonlinear Dynamical Systems. Cambridge University Press, New York
Hegselman R, Pietgen HO (eds) (1996) Modeling social dynamics: Order, chaos, and complexity. Helbin, Vienna
Lewin K (1936) Principles of topological psychology. McGraw-Hill, New York
Liebrand W, Nowak A, Hegselman R (eds) (1998) Computer modeling and the study of dynamic social processes. Sage, New York
Read SJ, Miller LC (eds) (1998) Connectionist models of social reasoning and social behavior. Erlbaum, Mahwah
Smith ER (1996) What do connectionism and social psychology offer each other? J Personal Soc Psychol 70:893–912
Vallacher RR, Read SJ, Nowak A (eds) (2002) The dynamical perspective in social psychology. Personal Soc Psychol Rev, Special Issue 6:264–387

Soft Computing, Introduction to

Janusz Kacprzyk
Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

As science progresses, scholars and researchers continue to develop models of various phenomena and problems. These models have become more and more formal, notably mathematical, and hence have relied more and more on computing.

The initial period has been characterized by computing meant as well-defined algorithmic procedures implemented by using computing architectures, tools and techniques based on traditional binary (yes – no or 0 – 1) logic. This logic has also been a basis for all kinds of approaches within the field of artificial intelligence, a field of science and technology that has been vividly developed since the mid-1900s and aimed at devising machines capable of performing human-specific tasks that require

intelligence. Symbolic computation has become a main paradigm within artificial intelligence, and the role of numerical computing has been neglected to a large extent.

As promises of artificial research gurus that machines will be able to perform virtually all sophisticated human-specific tasks and practically replace humans in a short time have failed, it has become more and more evident that sticking to symbolic computations only is a mistake. Clearly, there may be tasks for which symbolic computations are adequate, but there may be tasks, presumably even more, when numerical computations are indispensable.

Even if this necessity of a synergistic use of symbolic and numerical computations in the construction of intelligent systems has become obvious, people have been, for a long time, not fully convinced that, since the human being is a crucial element in all kinds of research, analysis and construction of broadly perceived intelligent systems, computational techniques and processes involved therein should reflect human-specific features.

In our context, the main feature of the human element in computation is that natural language is the only fully natural means of articulation and communication of a human being, and natural language is plagued by inherent vagueness, ambiguity, imprecision, etc. These kinds of information imperfections are out of scope of conventional mathematical tools based on – for instance – probabilistic or statistical tools.

The necessity of developing a new framework which would differ from the traditional "hard" computing based on traditional mathematics, i. e., basically on binary logic, and the dichotomy of true and false, has been recognized quite early. However, it was Lotfi A. Zadeh from the University of California at Berkeley who first pronounced that necessity – in the context of the analysis of complex animate systems, notably human-centered – in a more comprehensive way in the early 1960s. In 1965 he introduced fuzzy sets theory, presumably the first comprehensible, yet simple and efficient calculus of imprecision. A fuzzy set is a class to which elements can belong to a degree, from full membership, through all intermediate values, to full non-membership. Clearly, this has provided a constructive means for the representation of imprecisely specified concepts, relations, reasoning schemes, etc. Zadeh then proceeded to the formulation of possibility theory, and finally based his ideas on a more direct link to natural language which has culminated in his computing with words and perceptions.

The introduction of a fuzzy set, followed by numerous applications of fuzzy logic-based tools and techniques, notably for control, ranging from the control of domestic appliances, automatic transmissions in cars, to the control of industrial installations, has certainly been very important for the proliferation of fuzzy logic, maybe even more generally for non-conventional computational tools and techniques.

As a next step, Zadeh has advocated a wider view of those new computational tools and techniques and coined the term of *soft computing*. Basically, soft computing may be viewed as a departure from conventional (hard and "precise") computing in that it is tolerant of imprecision, uncertainty, partial truth, and approximate relationships, approximately satisfied properties, etc.

The main principle underlying soft computing is to be able to exploit a tolerance for imprecision, uncertainty, partial truth, and approximation to attain tractable, robust and computationally inexpensive modeling and solution tools and techniques. A natural connection to neural computing and evolutionary computing has more recently been recognized.

It is usually assumed that the principal components of soft computing are fuzzy logic (complemented with rough sets theory which may be viewed to address related problems of imprecise information but understood and dealt with in a different way), neural computation, and evolutionary computation. These areas are sometimes complemented with machine learning and probabilistic reasoning, with the latter subsuming belief networks, chaos theory and parts of learning theory. Moreover, for a long time, rough sets have been considered to be a relevant part of soft computing.

However, it should be noticed that soft computing is not a confluence of more or less independent fields, but is meant as a synergistic combination of various methodologies and methods in which each one contributes in a synergistic and complementary way to the analysis and solution of a problem.

In this section we will present the main components of soft computing. The concepts of synergy and complementarity mentioned above will provide a pivotal perspective.

We start with a brief introduction to fuzzy sets and systems (see J. Kacprzyk, ► Fuzzy Sets Theory, Foundations of). First, we define the concept of a fuzzy set followed by the definitions of its main properties. The idea of a fuzzy relation is then introduced, and shown to be a convenient apparatus for devising approximate reasoning schemes. The concept of a linguistic variable is shown to be a basis for a natural-language-based approach to modeling and reasoning. Fuzzy arithmetic is presented making it possible to use imprecisely specified (fuzzy) numbers in models and algorithms. A universal problem of decision making

is dealt with under fuzzy goals, constraints, etc., in a static and dynamic context.

Though fuzzy sets in the traditional Zadeh sense do provide a simple yet efficient means for the representation and handling of imprecision, some natural extensions of the basic concept of a fuzzy set have appeared. One of the most relevant and popular is the concept of a type 2 fuzzy set in which the degree of membership itself, which is a real number from the unit interval in the traditional fuzzy set, is now a fuzzy set defined in the unit interval. This natural extension has implied relevant extensions of fuzzy modeling and reasoning tools and techniques which are presented in detail, and potentials for applications are underlined (see R.I. John and J.M. Mendel, ► Fuzzy Logic, Type-2 and Uncertainty).

As in any formal model in which, as is true in virtually all cases in reality, multiple aspects, points of view and entities have to be accounted for, a proper aggregation is of a paramount importance. A comprehensive account of various aggregation operators is provided, both from a strictly mathematical point of view and from a more constructive, application oriented one (see V. Torra, ► Aggregation Operators and Soft Computing).

The discussion of the foundations of fuzzy logic is completed with a brief account of possibility theory initiated by Zadeh in the late 1970s as an extension of fuzzy sets theory and fuzzy logic, a mathematical theory for dealing with certain types of imperfect (uncertain) information in a way that is an alternative to probability theory. Main concepts related to possibility theory are presented, and a historical perspective is provided (see D. Dubois and H. Prade, ► Possibility Theory).

The comprehensive presentation of foundational concepts and properties related to fuzzy logic is then followed by sections presenting some more relevant areas which can be of use in the analysis of many systems and in the solution of many problems in diverse areas of science and technology.

First, fuzzy optimization and fuzzy mathematical programming are outlined starting with basic concepts of optimality, feasibility, etc., under fuzzy information. Then, various classes of fuzzy optimization and fuzzy mathematical programming problem classes are briefly outlined and basic solution concepts and algorithms are presented. Some examples of applications are mentioned (see W. Lodwick and E. Untiedt, ► Fuzzy Optimization).

The topic of statistics with imprecise data is considered next, a very interesting topic from a conceptual point of view, and practically relevant. First, a sound motivation is presented by pointing out that, in reality, much available data is imprecise, and so they cannot be used by powerful and well founded, yet too "hard" traditional statistical tools and techniques. New concepts are formulated and extensions of well known statistical means under an imprecise information are presented. Some applications are outlined (see M.A. Gil and O. Hryniewicz, ► Statistics with Imprecise Data).

The next papers are concerned with the two other key elements of soft computing: neural computation and evolutionary computation. To emphasize their synergy and complementarity, these tools have been presented in a hybrid context, that is as a presentation of new fields in which the confluence of fuzzy, neural and evolutionary tools yields a new quality, and results in the emergence of a new class of tools and techniques which have already changed the landscape of soft computing.

We start with the neuro-fuzzy systems, which refer to a combination of (artificial) neural networks and fuzzy logic, and may be viewed as a main, fundamental step into the development of hybrid intelligent systems (see L. Rutkowski, K. Cpałka, R. Nowicki, A. Pokropińska, R. Scherer, ► Neuro-fuzzy Systems). Basically, neuro-fuzzy systems try to combine transparent, human-like reasoning types of fuzzy rule-based systems with a parallel, connectionist type of neural networks. Therefore, they are universal approximators with an ability to yield interpretable IF-THEN rules, bridging the gap between accuracy and interpretability. A broad coverage of main architectures and implementations is presented.

Though neuro-fuzzy systems have been a considerable step forward towards in developing modeling tools of a better expressive power, effectiveness and efficiency, some natural extensions have been proposed that take advantage of the increasing popularity of evolutionary computation. It has been shown theoretically and experimentally that the incorporation in broadly perceived fuzzy systems, including neuron-fuzzy systems, of a learning mechanism which can make it possible to adjust the systems' structure and parameters may greatly enhance performance. The use of evolutionary computation has led in this context to the concept of an evolving fuzzy system that has been considered a promising alternative for some time (see P. Angelov, ► Evolving Fuzzy Systems).

The above mentioned hybridization techniques, which basically boil down to architectures and algorithms being a synergistic combination of fuzzy systems, notably fuzzy rule based systems, neural networks and evolutionary computation, in particular evolutionary learning, can be further extended when applied to real world problems which always provide much inspiration. Applications for systems modeling and control are of a universal importance for many fields of science and technology and hence

are dealt with in detail (see O. Castillo and P. Melin, ► Hybrid Soft Computing Models for Systems Modeling and Control).

The chapters briefly described above have been related to fuzzy sets and some of their hybridizations by including elements of neural networks and evolutionary computation to attain some "superadditivity" of the strengths and power of the particular single tools and techniques.

Rough sets theory, introduced by the late Zdzislaw Pawlak in the early 1980s, has been for some time considered a very promising tool for dealing with imperfect information, notably in data analysis and decision making.

Basically, a rough set is a formal approximation of a conventional crisp set in that it is represented as a pair of two traditional crisp sets which constitute the lower and the upper approximation of the original. These can be viewed as sets of elements that possibly and surely belong to the original set. Rough sets theory is presented in much detail, first in a more traditional, pure setting in Pawlak's spirit. Various concepts are presented, relations between them are formulated and proved, and their possible applications are outlined, notably in the relevant areas of data mining and knowledge discovery. Some extensions of the basic concept of a rough set are also presented (see J. Peters, A. Skowron and J. Stepaniuk, ► Rough Sets: Foundations and Perspectives).

Next, a "meta-problem" in science, that is decision making, is dealt with in the context of rough sets. In addition to the basic definition of a rough set, which provides much insight into the formalization and solution of various classes of decision making problems, a novel idea of a dominance-based rough set is presented. Its use to derive new, more human consistent solution concepts in decision making and data analysis is proposed (see R. Słowiński, S. Greco and B. Matarazzo, ► Rough Sets in Decision Making).

To summarize, we have tried to present the vast and diversified area of soft computing in a comprehensive, illustrative and constructive way. The basic perspective assumed may be viewed as presenting soft computing as a significant paradigm shift in computing in that it tries to exploit the fact that the human mind, unlike the computer we have now, has a remarkable ability to store and process information which is predominantly imprecise, uncertain and granular.

We have started with some basic fundamental concepts and properties, and then have devoted much space to the presentation of what is characteristic for soft computing: a synergistic, complementary use of various tools and techniques to solve problems in an effective and efficient way. This has led to various types of hybridization in which tools from fuzzy systems, rough sets, neural networks and evolutionary computation have been employed. It seems that soft computing should play a more and more important role in the years to come, in which more and more emphasis should be put on human-consistent and human-centric tools and techniques, and on hybrid systems.

Software Architectures for Autonomy

Robert A. Touchton[1], Carl D. Crane III[2]
[1] Honeywell International, Phoenix, USA
[2] University of Florida, Gainesville, USA

Article Outline

Glossary
Definition of the Subject
Introduction
Joint Architecture for Unmanned Systems (JAUS)
NIST 4D/RCS
Service Oriented Architecture/Component Oriented Architecture
Distributed Architecture for Mobile Navigation
Situation Assessment
Planning and Decision Making
Knowledge Representation
Adaptive Planning Framework
Future Directions
Bibliography

Glossary

Reference architecture (RA) A reference architecture is a framework containing valuable implementation guidance for meeting the requirements of an enterprise.

Automated guided vehicle (AGV) A vehicle that can be programmed to automatically drive to designated points and perform preprogrammed functions.

JAUS The Joint Architecture for Unmanned Systems (JAUS) Reference Architecture defines a set of reusable components and their interfaces. The architecture emphasizes vehicle platform independence, mission isolation, computer hardware independence, and technology independence.

NIST 4D/RCS The acronym NIST 4D/RCS refers to the National Institute of Standards and Technology Real-time Control System. 4D/RCS describes in detail the functions and associated interfaces necessary to provide sensory processing, world modeling, knowledge

management, cost/benefit analysis, and behavior generation.

Adaptive planning framework The Adaptive Planning Framework addresses dynamic situation assessment, behavior management, and decision-making. It incorporates a three-stage process of 1) understanding the current situation, 2) understanding the suitability and viability of the available behaviors in light of that situation, and 3) providing the capability to autonomously make and execute behavior-related decisions, all in real-time.

Service oriented architecture (SOA) A Service Oriented Architecture maintains a strictly enforced standardized interface among entities. A standardized messaging construct enables one entity to request a service from another entity and for that service provider to send its response.

Definition of the Subject

Developing an autonomous system, such as an autonomous ground vehicle, that can operate and maneuver in an unstructured environment is a complicated task. One of the most daunting issues facing autonomous vehicle researchers is how to best exploit sensor and other information discovered during the execution of a plan. If the system takes too long to deliberate on the possible meanings and implications of this newfound data and knowledge, the vehicle may well have progressed beyond the point where it can benefit from it. Indeed, it may now be sitting atop the unforeseen obstacle that spawned the influx of new information that was being processed.

The execution of specific autonomous behaviors is becoming reasonably well understood, such as "waypoint-following with obstacle detection", though improvements and breakthroughs in these areas continue. However, the autonomous selection of which behavior(s) should be invoked, and in what sequence and by what method, is in need of movement in the state of the practice. Advanced ways of thinking about, organizing, and applying situational knowledge to macro-level planning and decision-making are needed by the autonomous robotics community in order to achieve the full potential of the field. This article discusses various approaches and various reference architectures that have been developed to address these problems.

Introduction

Three predominant standards are discussed in the following sections, i. e. the Joint Architecture for Unmanned Systems (JAUS) Reference Architecture, the National Institute of Standards and Technology (NIST) 4D/RCS (Real-time Control System), and the Service Oriented Architecture. A fourth architecture referred to as the Distributed Architecture for Mobile Navigation (DAMN) is also discussed although this architecture is not widely adopted. Lastly, the Adaptive Planning Framework approach that was developed at the University of Florida to address the problem of decision making is introduced.

Joint Architecture for Unmanned Systems (JAUS)

The Joint Architecture for Unmanned Systems (JAUS) Reference Architecture, Version 3.2 [18] defines a set of reusable components and their interfaces. In order to ensure that the architecture will be applicable to the entire domain of unmanned mobile systems, the following four characteristics have been considered by the JAUS Working Group in the creation of the Reference Architecture:

1. Vehicle platform independence. In order for JAUS components to be interoperable, no assumptions about the underlying vehicle or its means of propulsion are made.
2. Mission isolation. The JAUS components can typically be assembled such that a variety of missions can be supported.
3. Computer hardware independence. No assumption of or requirement of particular computer hardware is made. This allows for future adaptability and enhancement as new computer hardware becomes available.
4. Technology independence. This is similar to the computer hardware independence, but focuses more on the technical approach rather than the computer hardware. For example, many approaches could be used to determine vehicle position and orientation. No one approach, such as GPS, inertial dead reckoning, or landmark-based navigation for example, is specified.

As currently defined, JAUS Reference Architekture (RA) establishes a pre-defined set of standard, yet flexible, components that provide a menu of capabilities that can be drawn from to design an unmanned system. Components are divided into five categories:

- Command and control components
- Communications components
- Platform components
- Manipulator components
- Environmental sensor components

The RA also defines a standardized messaging construct (header and content) that enables JAUS components to exchange information in an efficient and robust

fashion. The messaging approach first defines the content and usage of a standardized JAUS Header. It then prescribes the legal JAUS data types that can be incorporated into a message. Then it defines each JAUS message.

The Adaptive Planning Framework Reference Implementation was developed within such components and using such messages as defined by JAUS. Specifically, the Reference Implementation for the Adaptive Planning Framework was cast in an operational JAUS-compliant vehicle. However, the concepts and ideas that make up the Adaptive Planning Framework are not tied to nor specifically depend on JAUS. This enables other organizations to implement the framework in an alternative architecture or in future evolutions of the JAUS Reference Architecture itself.

NIST 4D/RCS

The National Institute of Standards and Technology (NIST) has been working for over two decades on establishing a standardized approach to the intelligent control of unmanned vehicle systems. The most comprehensive summary of their approach is given in NISTIR 6910, 4D/RCS: A Reference Model Architecture [27]. The 4D/RCS architecture is itself derived from NIST RCS, a domain-independent architecture developed by NIST a decade plus earlier (see [26] for a good overview of the generic RCS methodology). 4D/RCS goes on to specialize RCS to the domain of intelligent vehicle systems for military use.

4D/RCS focuses on ways to ensure that military missions involving unmanned vehicles can be analyzed, decomposed, distributed, planned, and executed in an intelligent, effective, efficient, and coordinated fashion. 4D/RCS describes in detail the functions and associated interfaces necessary to provide sensory processing, world modeling, knowledge management, cost/benefit analysis, and behavior generation. Of particular interest is its hierarchical treatment of time, providing a temporally layered set of eight planning/execution regimes (see Fig. 1). For example, it suggests that a vehicle Subsystem Planner (Level 3) ought to execute at ~1–5 Hz with a 5 second, 50 meter planning horizon at a 40 cm grid map resolution, while a Section Planner (Level 5) might need to re-plan every 50 seconds, with a 10 minute, 5 km planning horizon at a 40 m grid map resolution.

A challenge posed by 4D/RCS is that their hierarchy of nodes calls for each node to possess a complete set of functional capabilities (i. e., World Model, Value Judgment, Behavior Generation, etc.), scaled and scoped to its level of operation in the hierarchy. The partitioning, decomposition, and distribution of the Adaptive Planning Framework Specialists and Decision Brokers across a 4D/RCS hierarchy will be a completely new research area. Of greater concern is that 4D/RCS puts a great deal of power and functionality into their World Model, including prediction and simulation.

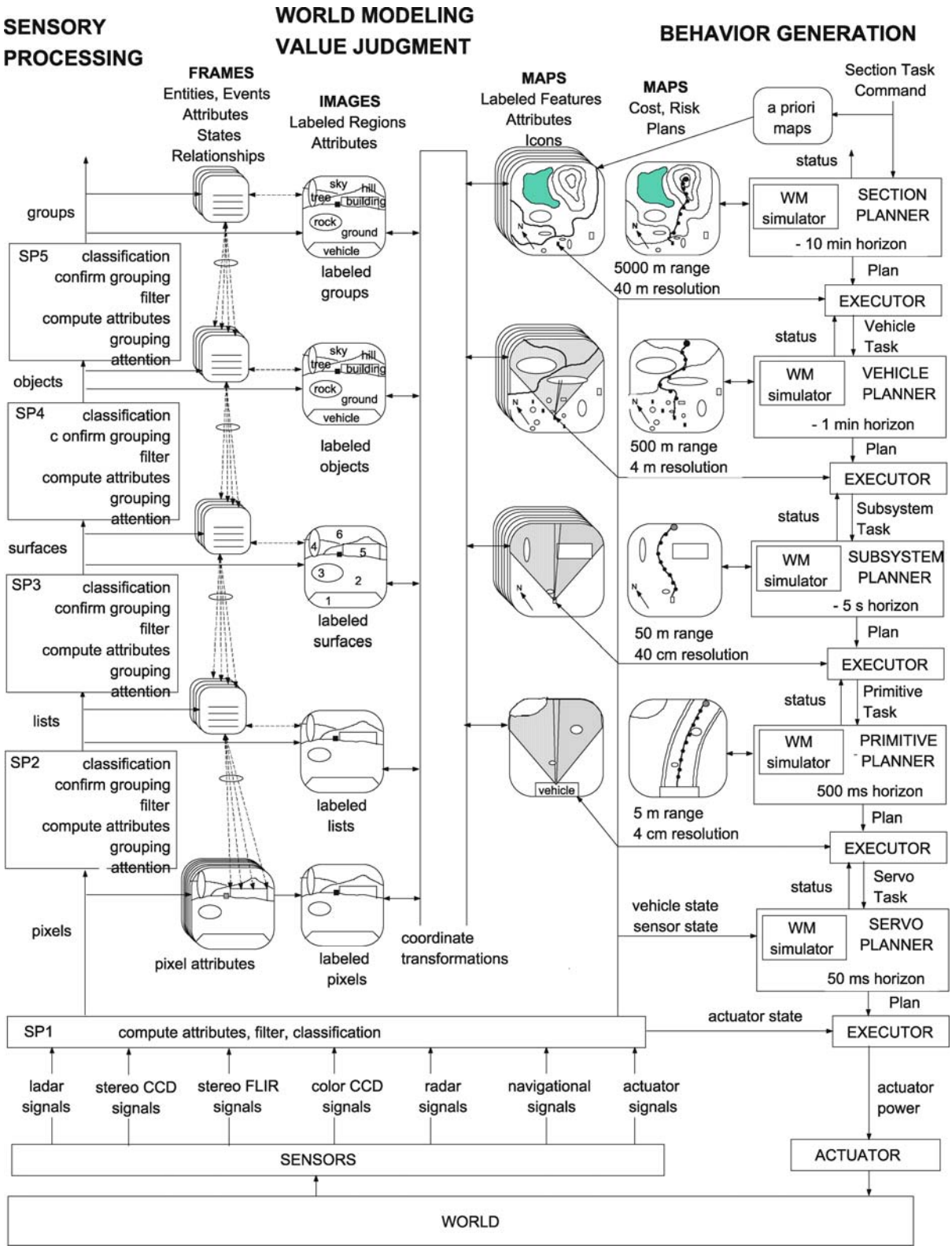
Service Oriented Architecture/ Component Oriented Architecture

Several facets of the Information Technology (IT) sector have been working to establish standards that support software interoperability across diverse organizations under the moniker of Service Oriented Architecture or SOA. SOA enables loose coupling among diverse software entities across a common network. This is accomplished by maintaining a strictly enforced standardized interface among the entities and a standardized messaging construct that enables one entity to request a service from another entity and for that service provider to send its response. This rapidly emerging standard is of interest here because the JAUS Working Group has begun a transition to a SOA-style architecture.

The most mature of these efforts is sponsored by the World Wide Web Consortium (W3C), which relies heavily upon SOA as the foundation of its Web Services initiative and, therefore, is leading the way in its maturation and adoption as a standard. Web Services extend the SOA concept to address anonymous entities that can discover one another and engage one another's services autonomously over the World Wide Web. They have published a treatise on the Web Services Architecture that includes an excellent overview of SOA in Sect. 3.1 of [45]. They go on in that section to outline some of the pitfalls of a SOA, such as network reliability and latency, lack of shared memory between service provider and consumer (i. e., everything that must be conveyed from one entity to another must be done explicitly via message content, and side effects of receipt of a message must be well understood and agreed upon), concurrency mismatches, and so on.

Industry has also taken a strong role in promoting SOA as a de facto standard. IBM (http://www-128.ibm.com/developerworks/webservices/standards/), Sun Microsystems (http://java.sun.com/developer/technicalArticles/WebServices/soa2/SOATerms.html#soaterms), and Microsoft (http://msdn.microsoft.com/architecture/soa/default.aspx?pull=/library/en-us/dnmaj/html/aj1soa.asp), to name three, have all embraced the notion.

Academia has also been active in this arena. IEEE Computer Society has formed a Technical Committee on Services Computing (http://tab.computer.org/tcsc/



Software Architectures for Autonomy, Figure 1
Excerpt from NIST PowerPoint Presentation (Source: http://www.isd.mel.nist.gov/projects/rcs/presentationhui/sld019.htm)

link.htm), and ACM has been actively including SOA topics in many of their conferences and symposiums.

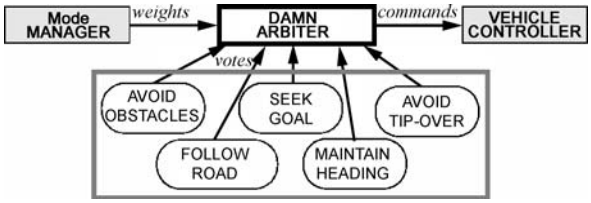
A closely related predecessor to SOA is component-based architecture (COA), which differs primarily in its stronger predisposition of what services a software entity ("component") will provide and less standardization of how components communicate with each other. In other words, COA does not worry so much about a component performing a single task (as in SOA) as long as the multiple services provided by a given component, and the interface for executing those services, are well documented. The emphasis is on providing a good platform for problem decomposition and loose coupling among components, with less emphasis on component interoperability. Aksit [1] provides an excellent compilation of articles on the topic of COA, especially Chap. 3, "Component–Based Architecting for Distributed Real-Time Systems", which, in turn, includes a detailed example of using a COA to devise a Car Navigation System (page 85). All in all, SOA can be considered a maturation, and perhaps specialization, of COA.

Distributed Architecture for Mobile Navigation

The Distributed Architecture for Mobile Navigation (DAMN) was originally published as a Ph.D. Dissertation [35] and, while not as widely adopted as the architectures discussed above, it has provided many useful insights. Even though the scope of DAMN is limited to navigation and obstacle avoidance, its distributed approach, its support of hybrid planning and implementation styles, its blend of centralized and decentralized processing, and its thoughtful treatment of salient challenges to real-time decision-making all make it worthy of elaboration here.

The basic premise behind DAMN is that centralized arbitration of distributed decision-making processes provides a reasonable and useful balance between the demands for real-time responsiveness and the challenges brought about by the asynchronous, latency-filled, heterogeneous, uncertain environment encountered by an autonomous ground vehicle. As in the other architectures discussed, DAMN provides a modular, extensible, and interoperable framework for supporting the generation and arbitration of sensor data, behaviors, and commands to the mobility platform, controllers, and actuators. This notion is shown schematically in Fig. 2, where sensor data and high-level commands have been bundled with the assortment of behaviors depicted.

The treatise goes on to present and analyze alternative structuring of the placement of arbitration (e. g., sensor vs. command vs. effect) and to explore various action selection schemes. A detailed presentation of the DAMN implementation on a CMU Navlab AGV and the experimental results achieved provides further insights into the merits and shortcomings of the architecture. Another major contribution of that research was the application of utility theory to the behavior arbitration process, as further discussed later in this article.



Software Architectures for Autonomy, Figure 2
DAMN Arbiter and Behaviors (Source: [35], page 9, Figure 1–2)

Situation Assessment

The situation assessment domain of interest to the topic of software architectures for autonomy is that of an unmanned system understanding its surroundings and status at a higher, more abstract level than that provided directly by its perception systems. In reviewing the literature, one must filter the use of the term when used in the context of the design of manned combat systems; such references often address such topics as own and enemy radars, missile tracking, and weapon lethality. Most such references are in the context of providing situational assessment for a human [16,48], such as pilot support on board a combat aircraft. Of interest here, however, is the applicability to unmanned systems, wherein the raw sensor and signal data is processed into more general situational conclusions, usually as a result of some form of inference or deduction. For clarity, the term "situation assessment" when used in this document will refer to this latter connotation. This domain is sometimes mentioned in the literature as "situational awareness" and could be referring to either of the connotations discussed above.

Work at USC [49] described the use of templates and patterns to provide situation assessment in virtual humans. They demonstrate a way to use situation assessment to improve decision-making by allowing the software system to better focus its attention (i. e., computing resources) with the goal of improved utilization of onboard resources.

Of particular interest is the work underway at NIST. They are working in several areas that address situation assessment. One has to do with incorporating situation assessment feedback to human operators of robotic devices [40]. While their emphasis is on the human-machine

interface, there are insights to be gained from the situation identification and classification schemes that they developed. An even more relevant front is their work on using 4D/RCS to control on-road robotic vehicles. There are both formal papers [38] and materials and presentations available on the NIST web site (see Fig. 3) that demonstrate ways to incorporate situation assessment notions into the 4D/RCS architecture.

Weiss, Philipps and To et al. [46] present a capability that could be adapted into a Traffic Specialist. It provides situation classification and prediction for an assortment of expressway-related conditions (such as same/different lane assignment for other vehicles that are detected), and states (Such as approach rate {approach | approach with distance warning | approach with collision warning}). Similarly, [13] introduce material that could form a Collision Avoidance Specialist that could manage interactions with moving obstacles using such notions as "time to collision", "time to brake", and "time to disappear". Another area of interest is vehicle self-awareness and work such as [33] sheds light on how a Vehicle Health Specialist might be devised.

Finally, it should be noted that much of the discussion of situation assessment in the literature was secondary to a broader discussion and is, thus, of most use in providing insights into possible nomenclature and classification. References such as these are discussed in the Knowledge Representation section rather than here.

Planning and Decision Making

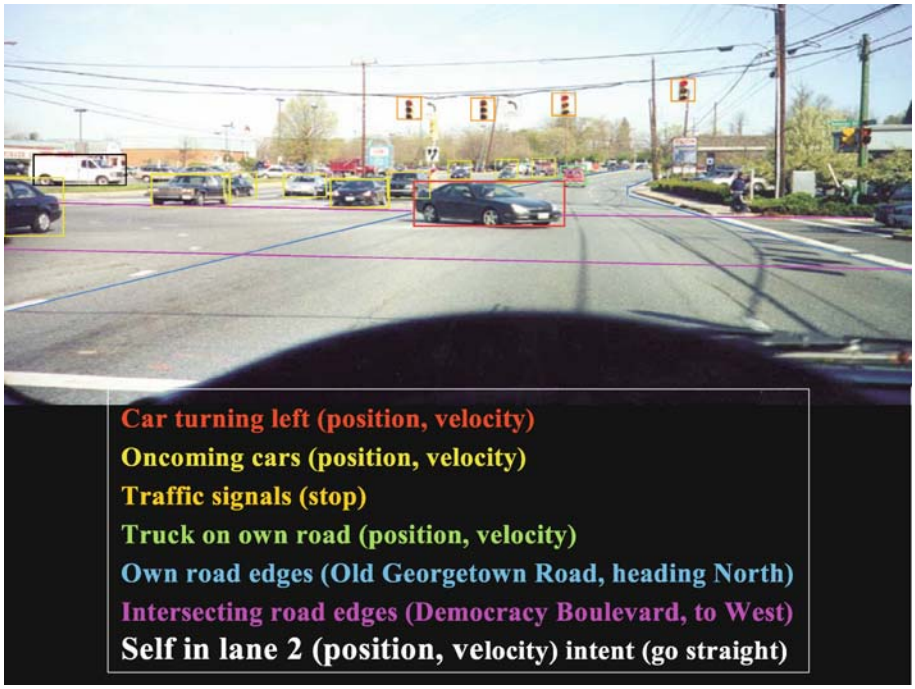
Since the scope of this topic is so broad, its treatment here will be, first of all, limited to the domain of real-time planning and decision-making on an AGV and then further organized as an assortment of "views". The notion of planning refers to the orchestration of executable behaviors to achieve a goal (e. g., find a series of waypoints that will take the vehicle to a desired goal, then drive the vehicle to those waypoints while avoiding obstacles, obeying driving rules and maintaining stability), as well as the low-level planning conducted by a given behavior (e. g., finding an obstacle-free path towards the next waypoint within the perception horizon of the vehicle). The following list of behavior primitives are but a sampling of those gleaned from the literature:

- Seek Goal
- Avoid Obstacles
- Follow Road
- Respond to Blocked Path
- Explore
- Wander
- Maintain Stability
- Seek Target/Intruder
- Intercept Target
- Mark Location

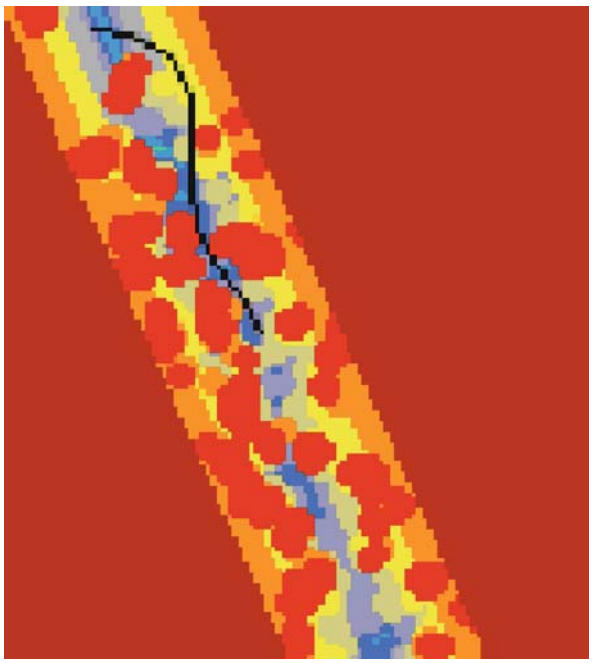
Viewed as a Sense-Plan-Act Problem

This is perhaps the most fundamental view of autonomous control of a mobile robot and one into which many autonomous robotic implementations can be cast. The notion is to neutralize uncertainties in the robot's perception of its world, its understanding of its own state, and the effects of its own actions by indirectly "closing the loop" through the continuous gathering of feedback from its environment while executing its plan [25]. Since it is anticipated that the plan itself will be divided into a sequence of steps, the idea is that the results of executing the initial steps can be observed and compared with expected results. If expectations are not being adequately met (in essence, forming an "error" signal), then the subsequent steps can be adjusted accordingly, or an entirely new plan can be published. It is presumed that the robot will have an ability to store its perception and state knowledge in some form of a world model, which can, in turn, be used by the planner.

This design style best describes the autonomous control used on the NAVIGATOR [8]. The four environmental sensors publish their findings, in the form of a traversability grid, to a sensor arbiter. Two additional pseudo-sensors each publish a traversability grid to the sensor arbiter denoting the a priori route boundary and a priori path plan. The sensor arbiter then fuses these inputs and publishes to the receding horizon planner a comprehensive traversability grid, which represents a localized view of a world model. The receding horizon planner uses an A* search algorithm and a simple vehicle model to iteratively produce viable plans that achieve a desired goal state and to choose the one that minimizes traversability cost. The goal itself is based on the a priori path plan and is replaced with a new goal once the vehicle nears it. The planning search that occurs has as its only objective the publishing of an instantaneous wrench command (steering, throttle, brake) to the vehicle's primitive driver, whose job is to execute that wrench as actuator positions. Thus, every cycle of the planner produces a new wrench command. Since every component in the chain executes at a nominal rate of 20 Hz, a new "plan" (as manifested in the instantaneous wrench command) is always being issued, thus providing a responsive behavior, with some deliberation on how that behavior is generated. Figure 4 shows a snapshot of an arbitrated traversability grid and the instantaneous plan. Red, orange, and yellow indicate lessen-



Software Architectures for Autonomy, Figure 3
Excerpt from NIST PowerPoint Presentation (Source: http://www.isd.mel.nist.gov/projects/rcs/presentationhui/sld061.htm)



Software Architectures for Autonomy, Figure 4
Example Traversability Grid taken from the NAVIGATOR while in a Cluttered Roadway

ing severity of obstacles, gray and blue indicate improving degrees of smoothness of terrain. The instantaneous plan is indicated in black.

One difference in the NAVIGATOR's implementation of the Sense-Plan-Act paradigm is that, by encapsulating the a priori plan into a pseudo-sensor whose findings compete with those of the other sensors, the conventional aspects of planning provide only "suggestions" for a preferred action, rather than forcing the vehicle onto a defined course. Although implemented quite differently, this notion is in concert with the findings of Payton, Rosenblatt and Keirsey [29], who go on to note that "In general, internalized plans should be conceived as representations that allow the raw results of search in an abstract state space to be made available as advice to continuous real-time decision-making processes".

There are many good examples of robotic systems that have implemented some fashion of the Sense-Plan-Act paradigm. Most have to do with navigation and obstacle avoidance, such as Batavia and Nourbakhsh [6]. Examples of this paradigm applied to other aspects of robotic planning and decision-making are much harder to find.

Viewed as a Subsumption Problem

The notion of empowering a mobile robot to operate without any centralized control was first introduced by Rod-

ney Brooks as he devised a self-managing, layered control scheme dubbed the "subsumption architecture" [7]. By decomposing a robot's control system into layers of task-achieving behaviors, control is governed by the dominant layer in play at an instance in time, which, in turn, "subsumes" the behaviors of the lower level layers. For example, let "Wander" be considered a level 1 behavior and "Explore" a level 2 behavior. Since Explore is the higher level, it will self-determine whether exploring is an appropriate behavior under current circumstances. If so, then it will alter the Wander behavior to be not random, but to fulfill its wishes to visit new areas. If not, then it will allow the Wander behavior to proceed without any alteration. This notion is extrapolated across all possible behaviors. This style of planning and decision-making is often referred to as "reactive". The resultant behavior of the robot is referred to as "emergent" since it is likely that the observed behavior is some extemporaneous blend of the possible behaviors that the robot could execute.

This approach to planning and decision-making has a dedicated following and is especially appealing for multi-agent and swarm applications. For example, the subsumption architecture and reactive behavior play a major (though not exclusive) role in the design of robots at the Idaho National Lab (see www.inl.gov/adaptiverobotics).

The differences between these first two views can be captured by the relative importance placed on each of the three components of the Sense-Plan-Act view. For example, a purely reactive system does almost no local planning since every stimulus anticipated during the sensing stage has a prescribed behavioral action, thus relegating the planning stage to simply resolving action conflicts when more than one stimulus is perceived or queuing and dispensing actions when one stimulus invokes multiple actions (i. e., managing the subsumption process). Conversely, a deliberative system will have a large emphasis on the planning stage, attempting to formulate a new plan that incorporates newly sensed information along with any changes in state of the vehicle or its mission while simulating the effect of alternative actions on the quality and viability of the plan. The juxtaposition of the Sense-Plan-Act view's emphasis on deep planning through possibly time-consuming deliberation and the Subsumption view's potentially unpredictable, but fast, reaction to stimuli, explains why researchers are still seeking other, hybrid or blended, planning and decision-making styles.

Viewed as a Decision Theory Problem

Another rich area of exploration is how classical decision theory might be applied to the AGV domain. For example, Karacapilidis and Papadias [20] describe how argumentation can be automated and used to support collaborative decision-making. Their ideas for automating argumentation constructs include evaluations such as "Scintilla of Evidence", "Beyond Reasonable Doubt", and "Preponderance of Evidence".

Rauenbusch and Grosz [32] and others speak of devising explicit "Plan Trees" whose nodes encapsulate the desired action/behavior, associated constraints, and contextual applicability and whose structure models the desired decision-making outcomes. The search through the tree is conducted using some measure of cost or value such that the correct path through the tree delivers the correct series of actions/behaviors.

Hoffman and Yates [15] present a synopsis of what has become known as the "three-step decision-making process". In this paper, they report that most, if not all decisions can be modeled as a cascading set of three-step activities. One of the models specifically referenced for use in process control is 1) Situation Assessment, 2) Planning, and 3) Commitment to a course of action. Each of these steps may be expanded into another three-step decision-making process, such as deciding which situational conditions are present or relevant, or whether to keep or abandon a committed action.

A final realm under decision theory is known as Hierarchical Task Network (HTN) planning [9] provide an overview of this concept and cite the seminal works that have contributed to it on the way to introducing a formalism of HTN planning semantics. The basic premise of HTN planning is to iteratively decompose tasks until primitive tasks are reached (defined as tasks that cannot be further decomposed and that are actionable). These primitive tasks are assembled into a network of increasingly abstract tasks allowing a planning algorithm to select a high-level task, recursively expand its children until its primitive tasks are reached. Some expansions may be constrained based on the current situation, thus pruning the search when compared with an unconstrained expansion of the network. Each reachable path from the high-level tasks to the primitive tasks becomes a candidate plan. While this exploration of the HTN is taking place, the candidates are being evaluated by so-called "critics" so that any arising conflicts can be identified and the winning candidate declared. Because of its deep reasoning, HTN-based planning is not typically used for real-time applications.

Viewed as a Behavior Arbitration Problem

The concept of Behavior Arbitration was introduced as part of the Distributed Architecture for Mobile Navigation

(DAMN) [35] as a key ingredient for achieving its goal of balancing centralized and decentralized design styles. All (decentralized) behavior generators submit their control output (referred to as a "vote") and the (centralized) DAMN Arbiter fuses their votes into a single command set to the vehicle.

"Utility fusion", which uses traditional utility theory to provide an alternative to command fusion, is another concept that evolved from DAMN [36]. This notion requires each behavior generator to submit a probabilistic utility estimate along with its vote, thus enabling a "utility arbiter" to compute the Maximum Expected Utility and use it to select the optimal behavior.

Viewed as an Action Selection Problem

Action Selection is another way to view planning and decision-making on an AGV. In this view, the mobile robot is tasked with selecting the most appropriate action based on the current situation, in spite of inaccurate, incomplete, and possibly unforeseen information. For this to happen, there must exist some catalog or repository of possible actions from which to select and the criteria upon which to base a selection decision. Pirjanian [30] provides an excellent overview of ten varying approaches to the action selection problem. In this treatise, he summarizes each (including DAMN), then compares and contrasts them in terms of eight criteria, including planning vs. reactivity, synchronous vs. asynchronous, hierarchy vs. no hierarchy, and knowledge representation which all have a direct bearing on the current research.

NIST has also developed an approach to action selection via its hierarchical planning and control scheme [21, 22]. The scheme enables the system to plan at different rates at each level, with the scope of planning fixed for each level. For example, high-level goal planning might take place at a lower resolution and update rate, but would cover a larger expanse than say planning for obstacle avoidance. The plans themselves are broken into a tree or graph of subgoals and subtasks (task decomposition itself is discussed under Knowledge Representation) and the actions are selected, executed, and monitored in accordance with the defined planning levels. The planning levels are chosen to be consistent with the time, duty cycle, and range horizon parameters established in the 4D/RCS architecture. For example, AGV mobility planning is broken into four levels: Servo, Prim(itive), Autonomous Mobility, and Vehicle System. Balakirsky and Lacaze [2] elaborate how planning, in the form of Value Judgment and Behavior Generation, takes place for the Vehicle System planning level.

Viewed as an Adaptive Planning Problem

Note that in some literature, "adaptive" is used to mean that the system "learns" from its experience, thus improving its performance over time, whereas the connotation used here is that the system alters its plan based on new, situational information that has been provided by upstream knowledge and data processing. Thus, while the possibility of actually changing the a priori behaviors from which to choose through learning should not be ruled out for future generations of the Adaptive Planning Framework, it is certainly not the emphasis or the motivation for using the term "adaptive" in its moniker. The genesis of adaptive planning as used here was a search to improve the performance of (manual) military mission planning through the use of expert systems, such as the Adaptive Mission Planning System in [41]. The quest continues as military planners seek to reduce 24-month planning cycles down to a year or less for complex deployments and even less for Crisis Action Planning [14]. In fact, their definition, "Adaptive Planning is the systematic, on-demand creation and revision of executable plans, with up-to-date options, as circumstances require", could suffice for the work conducted here as long as its transition to an autonomous, real-time setting is understood.

The need to alter a plan already in progress can have a number of causes, including insufficient time for completion, ineffective results, changes in the situation, and receipt of a new objective to name a few. The Artificial Intelligence community has driven related work in this area, but application to mobile autonomous robotics has not been at the forefront. For example, [12] presents an excellent treatise of an adaptive planning architecture based on the premise that an "agent dynamically constructs explicit control plans to guide its choices among situation-triggered behaviors". To accomplish this, she identified and explored five areas where an intelligent system might require adaptive behavior, depending on the situation encountered:

1. Perception Strategy – Adapt to information requirements and resource limitations
2. Control Mode – Adapt to goal-based constraints and environmental uncertainty
3. Reasoning Tasks – Adapt to perceived and inferred conditions
4. Reasoning Methods – Adapt to available information and current performance criteria
5. Meta-Control Strategy – Adapt to dynamic configurations of demands and opportunities

As an example of more recent work that does focus on mobile robotics in a real-time setting, Hassan, Simo and Crespo [11] offer a behavior-based architecture that will adapt to temporal constraints by allowing itself to utilize more deliberative techniques when time is available, but moving towards more reactive behaviors when time is at a premium. They also introduce the notion of adjusting the quality of service that a given element might deliver based on the situation encountered. For example, this approach might allow the system to attempt to achieve its goal with a "rough" plan if a "complete" plan could not be delivered in a timely enough manner. Musliner [23] and his Adaptive Mission Planner, provides another view on how to empower an autonomous system to alter its plans based on temporal constraints and in light of changing environments, objectives, and system capabilities. That work built upon his earlier efforts to devise the Cooperative Intelligent Real-time Control Architecture (CIRCA) [24], which provides formalisms on how to represent tasks and decisions in a LISP setting. While CIRCA has not been applied in the mobile robotic domain (making its suitability to support an AGV unknown), there are insights to be gained from this work. Finally, NIST has incorporated an element of adaptive planning in their recent work on autonomous on-road driving as part of 4D/RCS. For example, Balakirsky and Scrapper [3] discuss an expert system and knowledge representation scheme that support adaptive planning for autonomous lane and speed management.

Knowledge Representation

In this section, related work on Knowledge Representation relevant to the domain of AGVs is explored. Knowledge Representation refers to the schemas and constructs used to document, standardize, normalize, and utilize the entities within the domain of interest. It must capture the semantics and meanings of the relationships among the entities, as well as their names, descriptions, attributes, and the method or reasoning mechanism for determining their current state or value.

Sources of such domain knowledge include technical documents, specifications, training manuals, etc. (many of which can be accessed via the web). Example knowledge sources include a table of Autonomous Mobility Situation Coverage Requirements from Demo III requirements analysis [34], a Functional Taxonomy chart for an AGV from a TACOM (the U. S. Army's Tank-Automotive COMmand) PowerPoint presentation [31], and by drawing analogies from human military operations as found in the Army Universal Task List [44]. Remaining knowledge gaps must be filled in by interviews of subject matter experts or perhaps empirically through experimentation.

By far, the most work in knowledge representation for intelligent vehicles has been done by NIST. Thus, this section will conclude with an extended example, demonstrating their approach to representing knowledge about situational conditions, states and events, planning, and behaviors within the 4D/RCS context.

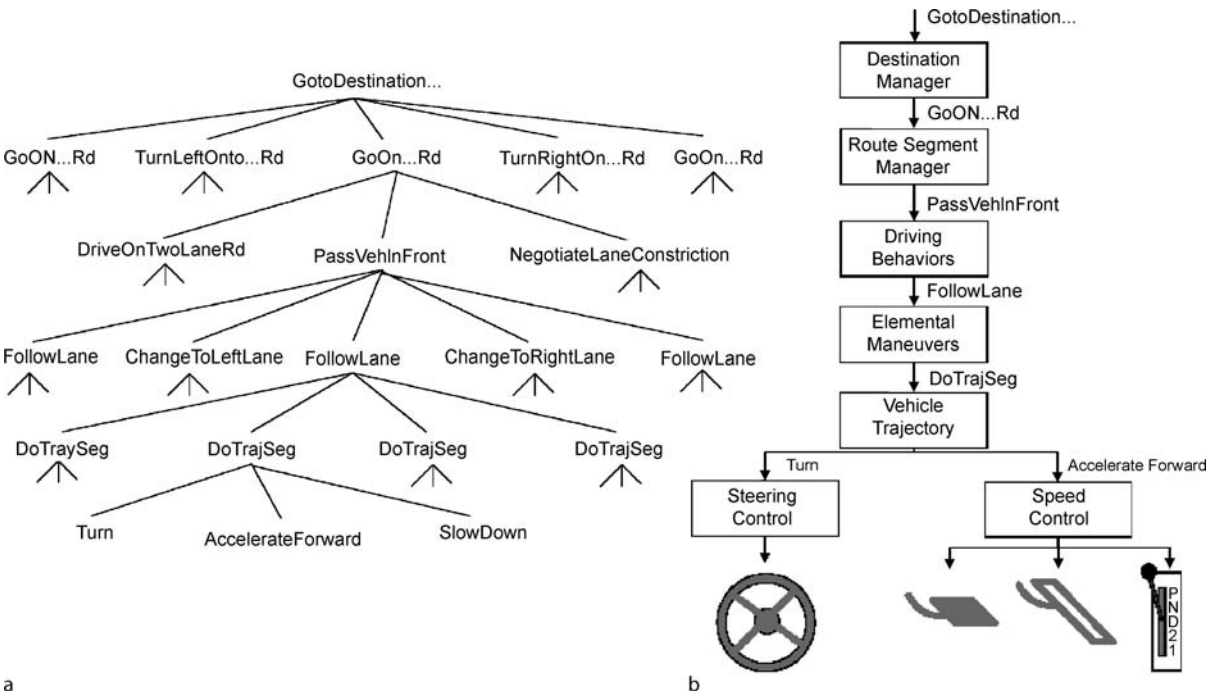
Lexicons, Taxonomies, and Ontologies

One technique for knowledge representation is to progress from a lexicon (a domain-specific dictionary of terms), to a taxonomy (a logical ordering and categorization of those terms), to an ontology (an explicit specification of those terms along with the semantics and relationships among them). One on-line dictionary defines ontology as follows:

> An explicit formal specification of how to represent the objects, concepts and other entities that are assumed to exist in some area of interest and the relationships that hold among them … Definitions associate the names of entities in the universe of discourse (e. g. classes, relations, functions or other objects) with human-readable text describing what the names mean and formal axioms that constrain the interpretation and well-formed use of these terms … The hierarchical structuring of knowledge about things by subcategorizing them according to their essential (or at least relevant and/or cognitive) qualities [17].

Much work is in progress attempting to build general purpose, or even "common sense" ontologies that would be useful to all domains. The most famous of these is the OpenCyc project (http://www.opencyc.org/), which, with 47,000 concepts and 306,000 assertions about them to date, is well on its way to achieving its vision to become "the world's largest and most complete general knowledge base and commonsense reasoning engine". Another initiative of interest is the DARPA Agent Markup Language (DAML) project (http://www.daml.org/), which was sponsored by DARPA to create an xml extension that provides (among other things) a rich and suitable language for the creation of general-purpose ontologies (282 distinct ontologies had been created using DAML by the time the program funding was terminated in 2006 and the work absorbed by W3C).

For the AGV domain, the scope of the knowledge that must be represented via the techniques discussed in this subsection is still quite broad. Situational knowledge spans from urban to highway to off-road environments,



Software Architectures for Autonomy, Figure 5
NIST Knowledge representation schemes for on-road driving. (Source: [4,5], Figures 3 and 4)

potential obstacles and hazards that might be present, traffic rules and driving best practices, and so on. Planning and behavior knowledge encompasses a wide variety of missions and tasks, whether they are high-level (conduct search and rescue operation), tactical (pass the vehicle), elementary (change lane), behaviors (avoid obstacles, maintain stability), planning rules and processes, and so on. Even knowledge about "self" or "ego" must be represented, such as capabilities, limitations and constraints, or current status. There is a major initiative under way at NIST, sponsored by TACOM, to develop an Intelligent Systems Ontology. Although still a work in progress, this intelligent-vehicle-specific ontology is expected to provide a standard set of domain concepts, their attributes and their interrelationships, delivered in a fashion that facilitates knowledge capture and reuse [39]. This ontology is beginning to gain traction as it makes its way into the AGV navigation planning community outside of NIST [37].

World Model Knowledge Store (WMKS)

Another dimension of knowledge representation is how data, information, and knowledge are stored. Whether it is provided a priori, or it is perceived, inferred, or received by the AGV, there must be a place and a format for storing, accessing, and analyzing it. Such data, information, and knowledge are often referred to as the "world model" and the place where they are stored as the "knowledge store". The breadth and sophistication of the world model knowledge store for a given AGV design will vary widely, depending on its degree of autonomy, the scope of its behavior, the complexity of its design, etc.

Situation Assessment findings, which must also be managed, fit into what some communities refer to as "meta-knowledge", i. e., knowledge about the knowledge. For example, while pumping out its perception data, a sensor could independently assess and report on its own confidence in its findings and its own health, and perhaps even declare that its own results should not be used right now (say, due to a camera white-out).

Although not always so, the knowledge store is usually persistent, using either a relational database or an object-oriented knowledge based system. Since much of the information stored is of a geo-spatial nature, the knowledge store often includes geo-spatial extensions for explicitly representing GIS and topographical data, polygonal objects, etc. Another consideration is whether the WMKS contents are stored in a central location, accessible by all AGV modules (sometimes referred to as a "blackboard architecture") or each module maintains a subset of the WMKS containing just the content it needs, with data, information, and knowledge marshaled among the AGV

modules on an as-needed/as-requested basis (sometimes referred to as a "publish/subscribe architecture").

Knowledge Representation at NIST

NIST advocates task decomposition as a key knowledge representation technique to support the hierarchical control strategy emphasized in its 4D/RCS architecture and has published widely on various ways to accomplish it [4,5]. This technique for representing the actionable elements that could be assembled to create a plan strives to break high-level tasks (e. g., a mission objective) into distinct hierarchical levels and also to identify multiple subtasks at a given level. Figure 5a–5b shows an example of how the "GoToDestination" task is decomposed into a "planning graph" that ultimately leads to a specific wrench command to the vehicle. The system must know (or be able to infer) the state of each node in the tree along with the cost of each arc in order for the associated control module to formulate the appropriate plan. Extending the example in Fig. 2–5b, a Destination Manager has determined that staying on the current road is appropriate and a Route Segment Manager has decided that passing the vehicle in front of it is the most desirable way to reach the destination. A Driving Behaviors module knows that its own vehicle has already changed into the passing lane and has further determined that the best thing to do right now is to stay in that lane, while a low-level Elemental Maneuvers module has found a wrench that ought to produce the requested outcome. Each Manager or module manages its own situational understanding either from direct sensory input or from its own local subset of the World Model Knowledge Store. Naturally, there are other tree elements and control modules that address following distance, speed, and so on, in addition to non-mobility-related tasks, such as payload management, communications, etc.

Once a plan is devised and approved, its elements must be executed by invoking one or more actions or behaviors, or perhaps by unleashing an entire subsystem to take over low-level control of the vehicle. NIST advocates the use of State Tables to represent the action decision-making knowledge [4]. A State Table is crafted for each node in the Task Decomposition Tree containing the rules that the control module is to use for mapping node inputs (states or situations) to allowable output actions.

To trigger the appropriate and desired state response, the matching situation must be known. The NIST approach to this is to determine and store the cascading precursor situational knowledge as a collection of "world states", but, in conformance to the 4D/RCS architecture, only that subset relevant to a given module. The lane-changing example concludes with a glimpse of the dozens of situational findings that lead up to the finding of interest ("ConditionsGoodToPass").

Adaptive Planning Framework

The Adaptive Planning Framework was developed by the authors at the University of Florida in order to address the requirements of the DARPA Urban Challenge that was held in November 2007. Figure 6 shows the University of Florida Navigator vehicle on which the Adaptive Planning Framework was implemented for the DARPA Urban Challenge. In the Adaptive Planning Framework, the system is assumed to be able to operate in a finite number of behavior modes. These behavior modes govern how the vehicle operates under various driving conditions. The framework is predominantly used to make intelligent decisions pertaining to these behaviors. The framework is scalable to systems of varying complexity and size and is compatible with existing architectures such as JAUS RA-3.2, NIST 4D/RCS, and others. The Adaptive Planning Framework is composed of three principle elements tasked with assessing the situation, determining the suitability and viability of all possible solutions, and executing the most suitable of all recommended solutions.

Situation Assessment Specialist

Dynamic environment information, originating from any array of sensors is monitored and managed by the Situation Assessment Specialists. Each specialist design is tailored to the sensor or collection of sensors whose data it analyzes. These specialists can, but are not required to, "live" on the same computing node that directly receives the sensor input. While the inputs to the specialist can come from any data source, the output or "finding" must adhere to specific guidelines outlined by the framework. Findings can be in the form of conditions, state, or events. A condition may have a value of present or absent only. All conditions are by default absent and must be proven present at each iteration. A finding classified as a state can only exhibit one of many a priori states. The event category is reserved for findings whose occurrence at some point in time is of significance even after the initial finding has passed. Once the findings have been generated the information is disseminated to all other components that might need it.

An example of a situation assessment specialist would be a software component whose sole function was to determine if it is safe to move to the adjacent lane. This component would monitor sensor data and reach a Boolean

Software Architectures for Autonomy, Figure 6
Team Gator Nation NaviGATOR

conclusion that would be stored as metadata for use by other processes. A second example would be a software component whose sole function was to determine if it is 'legal' to move to an adjacent lane. Here 'legal' is defined as not crossing a yellow line or not changing lanes when approaching an intersection.

Behavior Specialist

The findings rendered by the Situation Assessment Specialists are consumed by the behavior specialists. There is a one-to-one mapping of each behavior with a behavior specialist. The role of the specialist is to monitor the findings and evaluate the suitability of its behavior under the current perceived operating conditions. As with the specialist findings, the default recommendation is unsuitable and must be proven appropriate at every iteration of the program to ensure truth of the results and operating safety. A specialist does not possess the ability to activate or deactivate its associated behavior; such authority is only given to the Decision Broker.

For the DARPA Urban Challenge problem, the vehicle was programmed with six behavior modes. The corresponding behavior specialist constantly evaluates the appropriateness of its behavior. The six behavior modes were:

1. Roadway Navigation. The Roadway Navigation behavior is the primary driving behavior deriving commands to be sent to the vehicle actuators while the objective is lane following. This behavior allows the vehicle to navigate the roadway within the lines of its desired lane and maintain a safe following distance behind any vehicles ahead.
2. Change Lane Maneuver. The change lane maneuver is used in passing situations or in cases where the vehicle must change lanes in a multi-lane road in order to pass through a mission goal point. The behavior constrains the vehicle to remain within the lane boundaries of the new lane.
3. Reverse Direction. This behavior is called whenever it is determined that the current lane is blocked and there is no alternate clear lane available for passing. It is also applicable in cases where the vehicle has entered a 'dead end' road that it must "escape" to reach a mission goal point.
4. Intersection Traversal. The intersection traversal behavior is applicable when the vehicle enters the vicinity of an intersection. This is one of the most complicated behavior modes in that the system must rely on a series of situation assessment specialists to safely navigate the intersection. This behavior mode must handle queuing, stopping at the stop line, determining right of way, and ultimately traveling through the intersection while avoiding other vehicles.
5. Open Area Navigation. Open area navigation is a behavior that is only needed in special circumstances. This behavior allows the vehicle to move towards a goal location without striking any object, while avoiding any

rough terrain. This is in effect the only behavior mode that was required in the 2005 DARPA Urban Challenge. It is useful in the Urban Challenge when the vehicle is in an open area such as a parking lot prior to performing an actual parking maneuver.

6. Parking Lot. This behavior must deal with the problems that arise in the parking lot scenario where precise motion is necessary. When the vehicle approaches the vicinity of an assigned parking space, precise path planning will be initiated to align the vehicle as required. Situation assessment specialists will be monitoring the near surroundings of the vehicle to center the vehicle in its parking space while avoiding any static or dynamic objects.

Decision Broker

At the highest level of the framework lies the Decision Broker. Its role is to monitor all Behavior Specialist recommendations. It assumes ultimate authority over how the vehicle will operate while in autonomous mode. Like the other entities within the framework, the Decision Broker can base its conclusions on not only the recommendations and findings of other specialists, but it may also look at data from any other pertinent source. The author's implementation of the Adaptive Planning Framework centralized all the Decision Broker functionality within the JAUS Subsystem Commander component. The framework architecture employs an asynchronous, iterative, forward chaining reasoning approach to decision making.

Future Directions

JAUS is migrating to the Society of Automotive Engineering (SAE) Standard AS-4, Unmanned Systems. The transition to the SAE offered opportunity to further advance the standard through the introduction of more formalized specifications and methodologies.

The AS-4A Architecture Framework subcommittee has prepared document AIR 5665, Architecture Framework for Unmanned Systems. This document describes the concepts, capabilities and interoperability concerns of unmanned systems, and lays the foundation for all subsequent Aerospace Standards to be released by AS-4.

The AS-4B Network Environment subcommittee is preparing the document AIR5645, "JAUS Transport Considerations". This document examines aspects of JAUS, aspects of communications media, and aspects of the domain of unmanned systems, explaining how these affect the design of transport mechanisms for JAUS messaging. The committee has also prepared an initial draft version of AS5669, "JAUS Transport Specification".

The AS-4C Information Modeling and Definition subcommittee has prepared a working draft of the document AS5684, JAUS Service Interface Definition Language. This specification allows for the unambiguous definition of JAUS services. AS5684 is specified using Relax NG Compact notation, a machine-readable language, allowing for the creation of various tools to aid in the unmanned system design process. The AS-4C Subcommittee has also begun work on AS5710, JAUS Service Set. AS5710 is a collection of standard unmanned system services. Each service is defined using the JAUS Service Interface Definition Language. Current work on this document includes defining a discovery service and transport service, and transitioning the JAUS Reference Architecture 3.2 components to services.

Bibliography

1. Aksit M (2002) Software architectures and component technology. Kluwer, Boston
2. Balakirsky S, Lacaze A (2000) World modeling and behavior generation for autonomous ground vehicle. In: Proceedings of the 2000 IEEE International Conference on Robotics & Automation. San Francisco, pp 1201–1206
3. Balakirsky S, Scrapper C (2004) Knowledge representation and planning for on-road driving. Robotics Auton Syst 49(2004):57–66
4. Barbera A, Albus J, Messina E, Schlenoff C, Horst J (2004) How task analysis can be used to derive and organize the knowledge for the control of autonomous vehicles. In: Proceedings of the 2004 AAAI Spring Symposium Series on Knowledge Representation and Ontology for Autonomous Systems. Palo Alto, pp 67–78
5. Barbera A, Messina E, Huang H-M, Schlenoff C, Balakirsky S (2004) Software engineering for intelligent control systems. Künstliche Intell 3(4):22–26
6. Batavia PH, Nourbakhsh I (2000) Path planning for the cye personal robot. In: Proceedings of International Conference on Intelligent Robots and Systems, 2000 (IROS 2000). Takamatsu, Japan, 15–20
7. Brooks RA (1986) A robust layered control system for a mobile robot. IEEE J Robotics Autom 2(1):14–23
8. Crane CD, Armstrong DG, Touchton R, Galluzzo T, Solanki S, Lee J, Kent D, Ahmed M, Montane R, Ridgeway S, Velat S, Garcia G, Griffis M, Gray S, Washburn J, Routson G (2006) Team CIMAR's NaviGATOR: An unmanned ground vehicle for application to the 2005 DARPA grand challenge. J Field Robotics 23(8):599–623
9. Erol K, Hendler J, Nau D (1994) Semantics for hierarchical task-network planning. CS-TR-3239, UMIACS-TR-94-31, ISR-TR-95-9, University of Maryland, Institute for Advanced Computer Studies College Park, MD
10. Franklin S, Graesser A (1996) Is it an agent, or just a program?: A taxonomy for autonomous agents. In: Third International

Workshop on Agent Theories, Architectures, and Languages. Budapest, pp 21–35
11. Hassan H, Simo J, Crespo A (2001) Flexible real-time mobile robotic architecture based on behavioural models. Eng Appl Artif Intell 14(5):685–702
12. Hayes-Roth B (1995) An architecture for adaptive intelligent systems. Artif Intell 72:329–365
13. Hillenbrand J, Kroschel K, Schmid V (2005) Situation assessment algorithm for a collision prevention assistant. In: Intelligent Vehicle Symposium, 2005. Las Vegas, pp 459–465
14. Hoffman H (2004) Adaptive planning overview. Presented at Military Operations Research Society, Capabilities-Based Planning: The Road Ahead. The Military Operations Research Society (http://www.mors.org/), Alexandria
15. Hoffman RR, Yates JF (2005) Decision(?)Making(?). IEEE Intell Syst 20(4):76–83
16. Howard C, Stumptner M (2005) Probabilistic reasoning techniques for situation assessments. In: Third International Conference on Information Technology and Applications, 2005 (ICITA '05). Sydney, pp 383–386
17. Howe D (2005) The Free Online Dictionary of Computing. http://foldoc.org/foldoc.cgi?query=ontology. Accessed Oct 2006
18. JAUS (2005) JAUS Reference Architecture, version 3.2. JAUS Working Group, (http://www.jauswg.org/)
19. JAUS-OPC (2005) OPC 2.75 Interface Control Document, version 1.0, Payload Interface section. JAUS Working Group – OCU and Payloads Committee
20. Karacapilidis N, Papadias D (2001) Computer supported argumentation and collaborative decision making: the HERMES system. Inf Syst 26(4):259–277
21. Lacaze A (2002) Hierarchical planning algorithms. Presented at SPIE 16th Annual International Symposium on Aerospace/Defense Sensing, Simulation and Controls, Orlando
22. Murphy K, Abrams M, Balakirsky S, Coombs D, Hong T, Legowik S, Chang T, Lacaze A (2000) Intelligent control for unmanned vehicles. Presented at World Automation Conference (WAC 2000), Maui
23. Musliner DJ (2001) Imposing real-time constraints on self-adaptive controller synthesis. Lect Notes Comput Sci 1936:143–160
24. Musliner DJ, Durfee EH, Shin KG (1995) World modeling for the dynamic construction of real-time control plans. Artif Intell 74(1):83–127
25. Nilsson NJ (1998) Artificial intelligence: a new synthesis. Morgan Kaufmann, San Francisco
26. NIST (National Institute of Standards and Technology) (1992) A real-time control system methodology for developing intelligent control systems. NISTIR 4936. National Institute of Standards and Technology, Gaithersburg
27. NIST (National Institute of Standards and Technology) (2002) 4D/RCS: A Reference Model Architecture for Unmanned Vehicle System, Version 2.0. NISTIR 6910. National Institute of Standards and Technology, Gaithersburg
28. Panzarasa P, Jennings NR, Norman TJ (2002) Formalising collaborative decision-making and practical reasoning in multi-agent systems. J Logic Comput 12(1):55–117
29. Payton DW, Rosenblatt JK, Keirsey DM (1990) Plan guided reaction. IEEE Trans Syst, Man, Cybernetics 20(6):1370–1382
30. Pirjanian P (1997) An Overview of System Architecture for Action Selection in Mobile Robotics. Laboratory of Image Analysis, Aalborg University, Aalborg
31. Pritchett W (2002) A domain framework for intelligent ground combat vehicle systems. Presented at Software Technology Conference (STC 2002). TACOM (Tank-automotive & Armaments COMmand), Salt Lake City
32. Rauenbusch TW, Grosz BJ (2003) A decision making procedure for collaborative planning. Presented at AAMAS'03 – Second International Joint Conference on Autonomous Agents and Multi-Agent Systems, Melbourne
33. Reichard KM, Crow EC (2005) Intelligent self-situational awareness for unmanned and robotic platforms. AUVSI Unmanned Systems Conference, Baltimore
34. Robotic Systems Technology (1998) Demo III experiemental unmanned vehicle (XUV) program autonomous mobility requirements analysis. prepared for TACOM
35. Rosenblatt JK (1997) DAMN: A distributed architecture for mobile navigation. Dissertation, Carnegie Mellon
36. Rosenblatt JK (2000) Optimal selection of uncertain actions by maximizing expected utility. Auton Robots 9(1):17–25
37. Schlenoff C, Balakirsky S, Uschold M, Provine R, Smith S (2003) Using ontologies to aid navigation planning in autonomous vehicles. Knowl Eng Rev 18(3):243–255
38. Schlenoff C, Madhavan R, Barbera T (2004) A hierarchical, multi-resolutional moving object prediction approach for autonomous on-road driving. 2004 ICRA Conference, New Orleans, pp 1956–1961
39. Schlenoff C, Washington R, Barbera T, Manteuffel C (2005) A standard intelligent system ontology. Proceedings of SPIE – Unmanned Ground Vehicle Technology VII, Orlando, pp 46–56
40. Scholtz J, Antonishek B, Young J (2004) Evaluation of a human-robot interface: Development of a situational awareness methodology. National Institute of Standards and Technology, last accessed Oct 2006
41. Seares CDF (1987) Adaptive mission planning: Squeezing out greater combat capability. Air University Review, http://www.airpower.maxwell.af.mil/airchronicles/aureview/1987/seares2.html. Accessed Oct 2006
42. Touchton RA (1988) Reactor emergency action level monitor. In: Majumdar MC, Majumdar D, Sackett JI (eds) Artificial Intelligence and Other Innovative Computer Applications in the Nuclear Industry. Plenum, New York, pp 189–197
43. Touchton RA, Gunter AD, Wilson KM, Eldredge II TD, Weaver ME (1988) Reactor emergency action level monitor vol 1: REALM Technical Report. NP-5719. Electric Power Research Institute, Palo Alto
44. US Army (2003) Army Universal Task List. FM 7-15, Washington
45. W3C (2004) Web Services Architecture. World Wide Web Consortium, http://www.w3.org/TR/2004/NOTE-ws-arch-20040211/. Accessed Oct 2006
46. Weiss K, Philipps H, To TB, Kirchner A (2005) Environmental perception and situation assessment for an advanced highway assistant. In: 2005 IEEE Intelligent Vehicles Symposium Proceedings. Las Vegas, pp 472–477
47. Winston P, Horn B (1989) LISP: Third Edition. Addison-Wesley, Reading
48. Yanco HA, Drury J (2004) Where am I? Acquiring situation awareness using a remote robot platform. In: 2004 IEEE International Conference on Systems, Man and Cybernetics, The Hague, pp 2835–2840

49. Zhang W, Hill RW (2000) A template-based and pattern-driven approach to situation awareness and assessment in virtual humans. In: Fourth International Conference on Autonomous Agents, Barcelona, pp 116–123

Soliton Perturbation

Ji-Huan He
Modern Textile Institute, Donghua University, Shanghai, China

Article Outline

Glossary
Definition of the Subject
Introduction
Methods for Soliton Solutions
Future Directions
Bibliography

Glossary

Soliton A soliton is a nonlinear pulse-like wave that can exist in some nonlinear systems. The isolated wave can propagate without dispersing its energy over a large region of space; collision of two solitons leads to unchanged forms, solitons also exhibit particlelike properties.

Soliton perturbation theory The soliton perturbation theory is used to study the solitons that are governed by the various nonlinear equations in presence of the perturbation terms.

Homotopy perturbation method The homotopy perturbation method is a useful tool to the search for solitons without the requirement of presence of small perturbations. In this method, a homotopy is constructed with a homotopy parameter, p. When $p = 0$, it becomes a nonlinear wave equation such as a KdV equation with a known soliton solution; when $p = 1$, it turns out to be the original nonlinear equation. To change p from zero to unity, one must only change from a trial soliton to the solved soliton.

Variational iteration method The variational iteration method is a new method for obtaining soliton-type solutions of various nonlinear wave equations. The method begins with a soliton-type solution with some unknown parameters which can be determined after few iterations. The iteration formulation is constructed by a general Lagrange multiplier which can be identified optimally via variational theory.

Exp-function method The exp-function method is a new method for searching for both soliton-type solutions and periodic solutions of nonlinear systems. The method assumes that the solutions can be expressed in arbitrary forms of the exp-function.

Definition of the Subject

The soliton is a kind of nonlinear wave. There are many equations of mathematical physics which have solutions of the soliton type. The first observation of this kind of wave was made in 1834 by John Scott Russell [1]. In 1895, the famous KdV equation, which possesses soliton solutions, was obtained by D. J. Korteweg and H. de Vries [2], who established a mathematical basis for the study of various solitary phenomena.

From a modern perspective, the soliton is used as a constructive element to formulate the complex dynamical behavior of wave systems throughout science: from hydrodynamics to nonlinear optics, from plasmas to shock waves, from tornados to the Great Red Spot of Jupiter, from traffic flow to the Internet, from Tsunamis to turbulence [3]. More recently, solitary waves are of key importance in the quantum fields: on extremely small scales and at very high observational resolution equivalent to a very high energy, space–time resembles a stormy ocean and particles and their interactions have soliton-type solutions [4].

Introduction

The soliton was first discovered in 1834 by John Scott Russell, who observed that a canal boat stopping suddenly gave rise to a solitary wave which traveled down the canal for several miles, without breaking up or losing strength. Russell named this phenomenon the ‘soliton’.

In a highly informative as well as entertaining article [1] J.S. Russell gave an engaging historical account of the important scientific observation:

> *I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses, when the boat suddenly stopped – not so the mass of water in the channel which it had put in motion; it accumulated round the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth and well-defined heap of water, which continued its course along the channel apparently without change of form or diminution of speed. I followed it on horseback, and overtook it still rolling on at*

a rate of some eight or nine miles an hour, preserving its original figure some thirty feet long and a foot to a foot and a half in height. Its height gradually diminished, and after a chase of one or two miles I lost it in the windings of the channel. Such, in the month of August 1834, was my first chance interview with that singular and beautiful phenomenon which I have called the Wave of Translation.

His ideas did not earn attention until 1965 when N.J. Zabusky and M.D. Kruskal began to use a finite difference approach to the study of KdV equation [5], and various analytical methods also led to a complete understanding of Solitons, especially the inverse scattering transform proposed by Gardner, Greene, Kruskal, and Miura [6] in 1967. The significance of Russell's discovery was then fully appreciated. It was discovered that many phenomena in physics, electronics and biology can be described by the mathematical and physical theory of the 'Soliton'.

The particle-like properties of solitons [7] also caught much attention, and were proposed as models for elementary particles [8]. More recently it has been realized that some of the quantum fields which are used to describe particles and their interactions also have solutions of the soliton type [9].

Methods for Soliton Solutions

The investigation of soliton solutions of nonlinear evolution equations plays an important role in the study of nonlinear physical phenomena. There are many analytical approaches to the search for soliton solutions, such as soliton perturbation, tanh-function method, projective approach, F-expansion method, and others [10,11,12,13,14].

Soliton Perturbation

We consider the following perturbed nonlinear evolution equation [15,16]

$$u_T + N(u) = \varepsilon R(u), \quad 0 < \varepsilon \ll 1 . \tag{1}$$

When $\varepsilon = 0$, we have the un-perturbed equation

$$u_T + N(u) = 0 , \tag{2}$$

which is assumed to have a soliton solution.

When $\varepsilon \neq 0$, but $0 < \varepsilon \ll 1$, we can use perturbation theory [15,16], and look for approximate solutions of Eq. (1), which are close to the soliton solutions of Eq. (2). Using multiple time scales (a slow time τ and a fast time t, such that $\partial_T = \partial_t + \varepsilon\partial_\tau$), we assume that the soliton solution can be expressed in the form

$$u(x, T) = u_0(\xi, \tau) + \varepsilon u_1(\xi, \tau, t) + \varepsilon^2 u_2(\xi, \tau, t) + \cdots \tag{3}$$

where $\xi = x - ct$, and τ is a slow time and t is a fast time.

Substituting Eq. (3) into Eq. (1) and then equating like-powers of ε, we can obtain a series of linear equations for u_i $(i = 0, 1, 2, 3, \ldots)$.

In most cases the nonlinear term $R(u)$ in Eq. (1) plays an import role in understanding various solitary phenomena, and the coefficient ε is not limited to a "small parameter".

Variational Approach

Recently, variational theory and homotopy technology have been successfully applied to the search for soliton solutions [17,18] without requiring the small parameter assumption. Both variational and homotopy technologies can lead to an extremely simple and elementary, but rigorous, derivation of soliton solutions.

Considering the KdV equation

$$\frac{\partial u}{\partial t} - 6u\frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3} = 0 , \tag{4}$$

we seek its traveling wave solutions in the following frame

$$u(x, t) = U(\xi) \quad v(x, t) = V(\xi), \quad \xi = x - ct , \tag{5}$$

where c is angular frequency. Substituting Eq. (5) into Eq. (4) yields

$$-cu' - 6uu' + u''' = 0 , \tag{6}$$

where a prime denotes the differential with respect to ξ.

Integrating Eq. (6) yields the result

$$-cu - 3u^2 + u'' = 0 . \tag{7}$$

By the semi-inverse method [19], the following variational formulation is established

$$J = \int_0^\infty \left(\frac{1}{2}cu^2 + u^3 + \frac{1}{2}\left(\frac{\mathrm{d}u}{\mathrm{d}\xi}\right)^2 \right) \mathrm{d}\xi . \tag{8}$$

The semi-inverse method is a powerful mathematical tool to the search for variational formulae for real-life physical problems.

By the Ritz method, we search for a solitary wave solution in the form

$$u = p\,\mathrm{sech}^2(q\xi) , \tag{9}$$

where p and q are constants to be further determined.

Substituting Eq. (9) into Eq. (8) results in

$$J = \int_0^\infty \left[\frac{1}{2}cp^2\,\mathrm{sech}^4(q\xi) + p^3\,\mathrm{sech}^6(q\xi) + \frac{1}{2}(4p^2q^2\,\mathrm{sech}^4(q\xi)\tanh^2(q\xi)\right]\mathrm{d}\xi$$
$$= \frac{cp^2}{2q}\int_0^\infty \mathrm{sech}^4(z)\mathrm{d}z + \frac{p^3}{q}\int_0^\infty \mathrm{sech}^6(z)\mathrm{d}z + 2p^2q\int_0^\infty \left\{\mathrm{sech}^4(z)\tanh^4(z)\right\}\mathrm{d}z$$
$$= \frac{cp^2}{3q} + \frac{8p^3}{15q} + \frac{4p^2q}{15}\,. \tag{10}$$

Making J stationary with respect to p and q results in

$$\frac{\partial J}{\partial p} = \frac{2cp}{3q} + \frac{24p^2}{15q} + \frac{8pq}{15} = 0\,, \tag{11}$$
$$\frac{\partial J}{\partial q} = -\frac{cp^2}{3q^2} - \frac{8p^3}{15q^2} + \frac{4p^2}{15} = 0\,, \tag{12}$$

or simplifying

$$5c + 12p + 4q^2 = 0\,, \tag{13}$$
$$-5c - 8p + 4q^2 = 0\,. \tag{14}$$

From Eqs. (13) and (14), we can easily obtain the following relations:

$$p = -\frac{1}{2}c, \quad q = \sqrt{\frac{c}{4}}\,. \tag{15}$$

So the solitary wave solution can be approximated as

$$u = -\frac{c}{2}\,\mathrm{sech}^2\sqrt{\frac{c}{4}}\,(x - ct - \xi_0)\,, \tag{16}$$

which is the exact solitary wave solution of KdV equation (4).

The preceding analysis has the virtue of utter simplicity. The suggested variational approach can be readily applied to the search for solitary wave solutions of other nonlinear problems, and the present example can be used as paradigms for many other applications in searching for solitary wave solutions of real-life physics problems.

Variational Iteration Method

The variational iteration method [20] is an alternative approach to soliton solutions without the requirement of establishing a variational formulation for the discussed problems [17,21,22,23,24]. As an illustrating example, we consider the $K(3,1)$ equation in the form [17]:

$$u_t + u^2u_x + u_{xxx} = 0\,. \tag{17}$$

According to the variational iteration method, its iteration formulation can be constructed as follows

$$u_{n+1}(x,t) = u_n(x,t) - \int_0^t \left\{(u_n)_t + u_n^2(u_n)_x + (u_n)_{xxx}\right\}\mathrm{d}t\,. \tag{18}$$

To search for its compacton-like solution, we assume the solution has the form

$$u_0(x,t) = \frac{a\sin^2(kx+wt)}{b + c\sin^2(kx+wt)}\,, \tag{19}$$

where a, b, k, and w are unknown constants further to be determined after few iterations [17].

Homotopy Perturbation Method

The homotopy perturbation method [25] provides a simple mathematical tool for searching for soliton solutions without any small perturbation [18,26]. Considering the following nonlinear equation

$$\frac{\partial u}{\partial t} + au\frac{\partial u}{\partial x} + b\frac{\partial^3 u}{\partial x^3} + N(u) = 0, \quad a > 0,\ b > 0\,, \tag{20}$$

we can construct a homotopy in the form

$$(1-p)\left\{\frac{\partial u}{\partial t} + 6u\frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3}\right\} + p\left\{\frac{\partial u}{\partial t} + au\frac{\partial u}{\partial x} + b\frac{\partial^3 u}{\partial x^3} + N(u)\right\} = 0\,. \tag{21}$$

When $p = 0$, we have

$$\frac{\partial u}{\partial t} + 6u\frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3} = 0\,, \tag{22}$$

a well-known KdV equation whose soliton solution is known. When $p = 1$, Eq. (21) turns out to be the original equation. According to the homotopy perturbation method, we assume

$$u = u_0 + pu_1 + p^2u_2 + \cdots \tag{23}$$

Substituting Eq. (23) into Eq. (21), and proceeding with the same process as the traditional perturbation method does, we can easily solve u_0, u_1 and other components. The solution can be expressed finally in the form

$$u = u_0 + u_1 + u_2 + \cdots \tag{24}$$

The homotopy perturbation method always stops before the second iteration, so the solution can be expressed as

$$u = u_0 + u_1 \tag{25}$$

for most cases.

Parameter-Expansion Method

The parameter-expansion method [27,28,29,30] does not require one to construct a homotopy. To illustrate its solution procedure, we re-write Eq. (20) in the form

$$\frac{\partial u}{\partial t} + au\frac{\partial u}{\partial x} + b\frac{\partial^3 u}{\partial x^3} + 1 \cdot N(u) = 0\,. \tag{26}$$

Supposing that the parameters a, b, and 1 can be expressed in the forms

$$a = a_0 + pa_1 + p^2 a_2 + \cdots \tag{27}$$

$$b = b_0 + pb_1 + p^2 b_2 + \cdots \tag{28}$$

$$1 = pc_1 + p^2 c_2 + \cdots \tag{29}$$

where p is a bookkeeping parameter, $p = 1$. Substituting Eqs. (23), (27), (28) and (29) into Eq. (26) and proceeding the same way as the perturbation method, we can easily obtain the needed solution.

Exp-function Method

The exp-function method [31,32,33] provides us with a straightforward and concise approach to obtaining generalized solitonary solutions and periodic solutions and the solution procedure, with the help of Matlab or Mathematica, is utterly simple. Consider a general nonlinear partial differential equation of the form

$$F(u, u_x, u_y, u_z, u_t, u_{xx}, u_{yy}, u_{zz}, u_{tt}, u_{xy}, u_{xt}, u_{yt}, \ldots) = 0\,. \tag{30}$$

Using a transformation

$$\eta = ax + by + cz + dt\,, \tag{31}$$

we can re-write Eq. (30) in the form of the following nonlinear ordinary differential equation:

$$G(u, u', u'', u''', \ldots) = 0\,, \tag{32}$$

where a prime denotes a derivation with respect to η.

According to the exp-function method, the traveling wave solutions can be expressed in the form

$$u(\eta) = \frac{\sum_{n=-k}^{l} a_n \exp(n\eta)}{\sum_{m=-i}^{j} b_m \exp(m\eta)}\,, \tag{33}$$

where i, j, k, and l are positive integer which could be freely chosen, a_n and b_m are unknown constants to be determined. The solution procedure is illustrated in [32].

Future Directions

It is interesting to point out the connection of catastrophe theory to loop soliton chaos, and finally to chaotic Cantorian spacetime [34,35].

El Naschie [34,35] studied the Eguchi–Hanson gravitational instanton solution and its interpretation by 't Hooft in the context of a quantum gravitational Hilbert space, as an event and a possible solitonic "extended" particle. Transferring a certain solitonic solution of Einstein's field equations in Euclidean "real" space–time to the mathematical infinitely-dimensional Hilbert space, it is possible to observe a new non-standard process by which a definite mass can be assigned to massless particles. Thus by invoking Einstein's gravity in "solitonic" gauge theory and vice versa, an alternative explanation for how massless particles acquire mass is found, which is also in harmony with the basic structure of our standard model as it stands at present [34,35].

Bibliography

Primary Literature

1. Russell JS (1844) Report on Waves. Fourteenth Meeting of the British Association for the Advancement of Science, John Murray, London, pp 311–390
2. Korteweg DJ, De Vires G (1895) On the change of form of long waves advancing in a rectangular channel, and a new type of long stationary wave. Phil Mag Ser 539:422–443
3. Eilbeck C (2007) John Scott Russell and the solitary wave. Heriot-Watt University, Edinburgh http://www.ma.hw.ac.uk/~chris/scott_russell.html
4. El Naschie MS (2004) A review of E infinity theory and the mass spectrum of high energy particle physics. Chaos Solit Fract 19(1):209–236
5. Zabusky NJ, Kruskal MD (1965) Interaction of 'Solitons' in a collisionless plasma and the recurrence of initial states. Phys Rev Lett 15:240–243
6. Gardner CS, Greene JM, Kruskal MD, Miura RM (1967) Method for solving the KdV equation. Phys Rev Lett 19:1095–1097
7. Bode M, Liehr AW, Schenk CP et al (2002) Interaction of dissipative solitons: particle-like behavior of localized structures in a three-component reaction-diffusion system. Physica D (1–2):45–66
8. Braun HB, Kulda J, Roessli B et al (2005) Emergence of soliton chirality in a quantum antiferromagnet. Nat Phys 1(3):159–163
9. Ahufinger V, Mebrahtu A, Corbalan R et al (2007) Quantum switches and quantum memories for matter-wave lattice solitons. New J Phys 9:4
10. Ye JF, Zheng CL, Xie LS (2006) Exact solutions and localized excitations of general Nizhnik–Novikov–Veselov system in (2+1)-dimensions via a projective approach. Int J Nonlinear Sci Num Simul 7(2):203–208
11. Bogning JR, Tchakoutio-Nguetcho AS, Kofane TC (2005) Gap solitons coexisting with bright soliton in nonlinear fiber arrays.

Int J Nonlinear Sci Num Simul 6(4):371–385; Abdusalam HA (2005) On an improved complex tanh-function method. Int J Nonlinear Sci Num Simul 6(2):99–106
12. El-Sabbagh MF, Ali AT (2005) New exact solutions for (3+1)-dimensional Kadomtsev–Petviashvili equation and generalized (2+1)-dimensional Boussinesq equation. Int J Nonlinear Sci Num Simul 6(2):151–162
13. Shen JW, Xu W (2004) Bifurcations of smooth and non-smooth travelling wave solutions of the Degasperis-Procesi equation. Int J Nonlinear Sci Num Simul 5(4):397–402
14. Sheng Z (2007) Further improved F-expansion method and new exact solutions of Kadomstev–Petviashvili equation. Chaos Solit. Fract 32(4):1375–1383
15. Yu H, Yan J (2006) Direct approach of perturbation theory for kink solitons. Phys Lett A 351(1–2):97–100
16. Herman RL (2005) Exploring the connection between quasistationary and squared eigenfunction expansion techniques in soliton perturbation theory. Nonlinear Anal 63(5–7): e2473–e2482
17. He JH, Wu XH (2006) Construction of solitary solution and compacton-like solution by variational iteration method. Chaos Solit Fract 29(1):108–113
18. He JH (2005) Application of homotopy perturbation method to nonlinear wave equations. Chaos Solit Fract 26(3): 695–700
19. He JH (2004) Variational principles for some nonlinear partial differential equations with variable coefficients. Chaos Solit Fract 19(4):847–851
20. He JH (1999) Variational iteration method – a kind of non-linear analytical technique: Some examples. Int J Non-Linear Mech 34(4):699–708
21. Abulwafa EM, Abdou MA, Mahmoud AA (2007) Nonlinear fluid flows in pipe-like domain problem using variational-iteration method. Chaos Solit Fract 32(4):1384–1397
22. Inc M (2007) Exact and numerical solitons with compact support for nonlinear dispersive K(*m*,*p*) equations by the variational iteration method. Phys A 375(2):447–456
23. Soliman AA (2006) A numerical simulation and explicit solutions of KdV-Burgers' and Lax's seventh-order KdV equations. Chaos Solit Fract 29(2):294–302
24. Abdou MA, Soliman AA (2005) Variational iteration method for solving Burger's and coupled Burger's equations. J Comput Appl Math 181(2):245–251
25. He JH (2000) A coupling method of a homotopy technique and a perturbation technique for non-linear problems. Int J Non-Linear Mech 35(1):37–43
26. Ganji DD, Rafei M (2006) Solitary wave solutions for a generalized Hirota-Satsuma coupled KdV equation by homotopy perturbation method. Phys Lett A 356(2):131–137
27. Shou DH, He JH (2007) Application of Parameter-expanding Method to Strongly Nonlinear Oscillators. Int J Nonlinear Sci Numer Simul 8:113–116
28. He JH (2001) Bookkeeping parameter in perturbation methods. Int J Nonlinear Sci Numer Simul 2:257–264
29. He JH (2002) Modified Lindstedt-Poincare methods for some strongly non-linear oscillations. Part I: expansion of a constant. Int J Non-Linear Mech 37:309–314
30. Xu L (2007) He's parameter-expanding methods for strongly nonlinear oscillators. J Comput Appl Math 207(1):148–157
31. He J-H, Wu X-H (2006) Exp-function method for nonlinear wave equations. Chaos Solit Fract 30(3):700–708
32. He J-H, Abdou MA (2007) New periodic solutions for nonlinear evolution equations using Exp-function method. Chaos Solit Fract 34(5):1421–1429
33. Wu X-H, He J-H (2008) EXP-function method and its application to nonlinear equations. Chaos Solit Fract 38(3):903–910
34. El Naschie MS (2004) Gravitational instanton in Hilbert space and the mass of high energy elementary particles. Chaos Solit Fract 20(5):917–923
35. El Naschie MS (2004) How gravitational instanton could solve the mass problem of the standard model of high energy particle physics. Chaos Solit Fract 21(1):249–260

Books and Reviews

He JH (2006) Some Asymptotic Methods for Strongly Nonlinear Equations. Int J Mod Phys B 20(10):1141–1199; 20(18): 2561–2568
He JH (2006) Non-perturbative methods for strongly nonlinear problems. dissertation.de-Verlag im Internet, Berlin
Drazin PG, Johnson RS (1989) Solitons: An Introduction. Cambridge University Press, Cambridge

Solitons and Compactons

Ji-Huan He[1], Shun-dong Zhu[2]
[1] Modern Textile Institute, Donghua University, Shanghai, China
[2] Department of Science, Zhejiang Lishui University, Lishui, China

Article Outline

Glossary
Definition of the Subject
Introduction
Solitons
Compactons
Generalized Solitons and Compacton-like Solutions
Future Directions
Cross References
Bibliography

Glossary

Soliton A soliton is a stable pulse-like wave that can exist in some nonlinear systems. The soliton, after a collision with another soliton, eventually emerges unscathed.

Compacton A compacton is a special solitary traveling wave that, unlike a soliton, does not have exponential tails.

Generalized soliton A generalized soliton is a soliton with some free parameters. Generally a generalized soliton can be expressed by exponential functions.

Compacton-like solution A compcton-like solution is a special wave solution which can be expressed by the squares of sinusoidal or cosinoidal functions.

Definition of the Subject

Soliton and compacton are two kinds of nonlinear waves. They play an indispensable and vital role in all ramifications of science and technology, and are used as constructive elements to formulate the complex dynamical behavior of wave systems throughout science: from hydrodynamics to nonlinear optics, from plasmas to shock waves, from tornados to the Great Red Spot of Jupiter, from traffic flow to Internet, from Tsunamis to turbulence. More recently, soliton and compacton are of key importance in the quantum fields and nanotechnology especially in nanohydrodynamics.

Introduction

Solitary waves were first observed by John Scott Russell in 1895, and were studied by D. J. Korteweg and H. de Vries in 1895 . Compactons are special solitons with finite wavelength. It was Philip Rosenau and his colleagues who first found compactons in 1993. Please refer to "▶ Soliton Perturbation" for detailed information.

Solitons

A soliton is a special solitary traveling wave that after a collision with another soliton eventually emerges unscathed. Solitons are solutions of partial differential equations that model phenomena like water waves or waves along a weakly anharmonic mass-spring chain.

The Korteweg–de Vries (KdV) equation is the generic model for the study of nonlinear waves in fluid dynamics, plasma and elastic media. KdV equation is one of the most fundamental equations in nature and plays a pivotal role in nonlinear phenomena. We consider the KdV equation in the form

$$\frac{\partial u}{\partial t} + 6u\frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3} = 0\,. \tag{1}$$

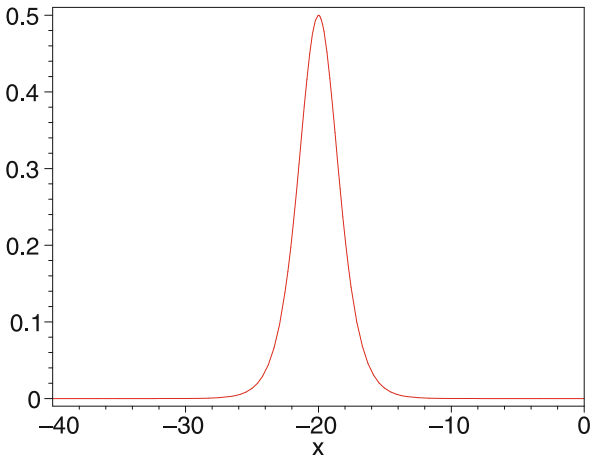
Its solitary traveling wave solution can be solved as

$$u(x,t) = \tfrac{1}{2}c\,\mathrm{sech}^2\left\{\tfrac{1}{2}c^{1/2}(x-ct)\right\}\,. \tag{2}$$

The bell-like solution as illustrated in Fig. 1 is called a soliton.

We re-write Eq. (2) in an equivalently form:

$$u(\xi) = p\,\mathrm{sec}\,h^2(q\xi) = \frac{4p}{\mathrm{e}^{2q\xi} + \mathrm{e}^{-2q\xi} + 2}\,. \tag{3}$$



Solitons and Compactons, Figure 1
Bell-like solitary wave

where $u(x,t) = u(\xi), \xi = x - ct, c$ is the wave velocity.

It is obvious that

$$\lim_{\xi\to\infty} u(\xi) = 0 \quad \text{and} \quad \lim_{\xi\to-\infty} u(\xi) = 0\,. \tag{4}$$

The soliton has exponential tails, which are the basic character of solitary waves. The soliton obeys a superposition-like principle: solitons passing through one another emerge unmodified, see Fig. 2.

Compactons

Now consider a modified version of KdV equation in the form

$$u_t + (u^2)_x + (u^2)_{xxx} = 0\,. \tag{5}$$

Introducing a complex variable ξ defined as $\xi = x - ct$, where c is the velocity of traveling wave, integrating once, we have

$$-cu + u^2 + (u^2)_{\xi\xi} = D\,, \tag{6}$$

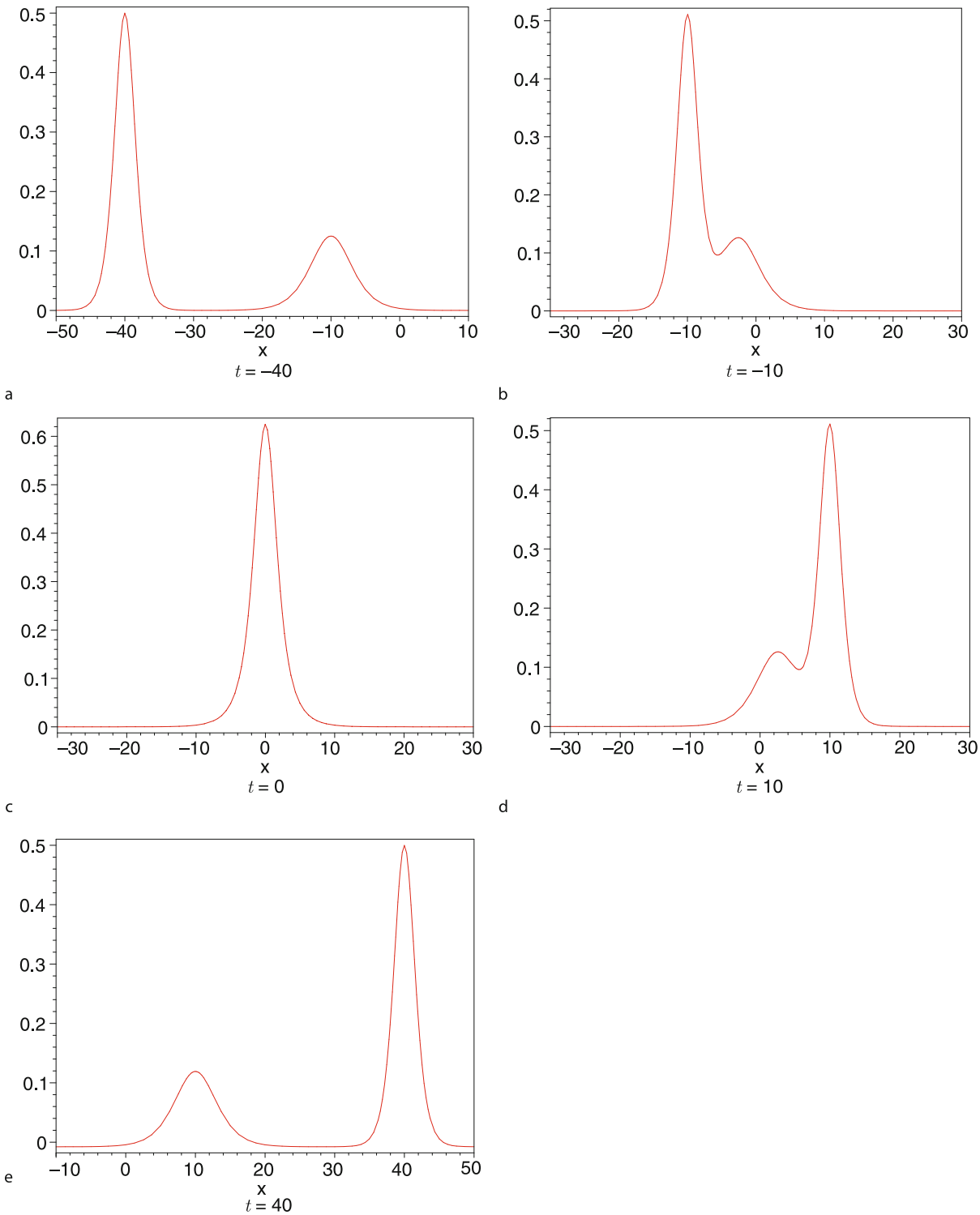
where D is an integral constant.

To avoid singular solutions, we set $D = 0$. We re-write Eq. (6) in the form

$$v_{\xi\xi} + v - cv^{1/2} = 0\,, \tag{7}$$

where $u^2 = v$.

In case $c = 0$, we have periodic solution: $v(\xi) = A\cos\xi + B\sin\xi$. Periodic solution of nonlinear oscillators can be approximated by sinusoidal function. It helps understanding if an equation can be classified as oscillatory by direct inspection of its terms.



Solitons and Compactons, Figure 2
Collision of two solitary waves

We consider two common order differential equations whose exact solutions are important for physical understanding:

$$u'' - k^2 u = 0\,, \tag{8}$$

and

$$u'' + \omega^2 u = 0\,. \tag{9}$$

Both equations have linear terms with constant coefficients.

The crucial difference between these two very simple equations is the sign of the coefficient of u in the second term. This determines whether the solutions are exponential or oscillatory. The general solution of Eq. (8) is

$$u = Ae^{kt} + Be^{-kt}\,. \tag{10}$$

The second Eq. (9) has a positive coefficient of u, and in this case the general solution reads

$$u = A\cos\omega t + B\sin\omega t\,. \tag{11}$$

This solution describes an oscillation at the angular velocity ω.

Equation (7) behaves sometimes like an oscillator when $1 - cv^{-1/2} > 0$, i. e., $u = v^{1/2}$ has a periodic solution, we assume v can be expressed in the form

$$v = u^2 = A^2\cos^4\omega\xi\,. \tag{12}$$

Substituting Eq. (12) into Eq. (7) results in

$$12A^2\omega^2\cos^2\omega\xi - 16A^2\omega^2\cos^4\omega\xi + A^2\cos^4\omega\xi - cA\cos^2\omega\xi = 0\,. \tag{13}$$

We, therefore, have

$$\begin{aligned} 12A^2\omega^2 - cA &= 0 \\ -16A^2\omega^2 + A^2 &= 0\,. \end{aligned} \tag{14}$$

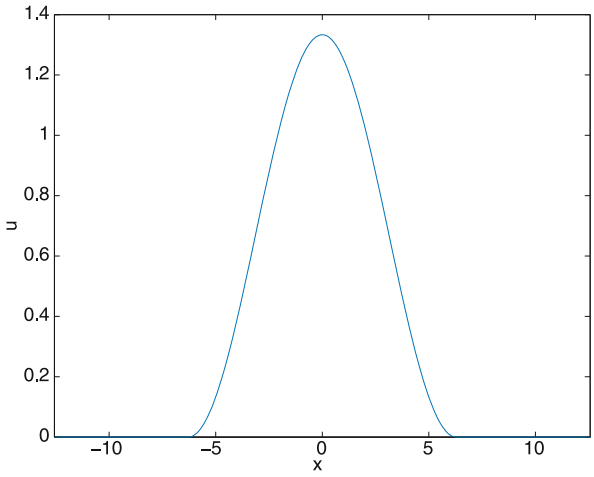
Solving the above system, Eq. (14), yields

$$\omega = \frac{1}{4}\,, \quad A = \frac{4}{3}c\,. \tag{15}$$

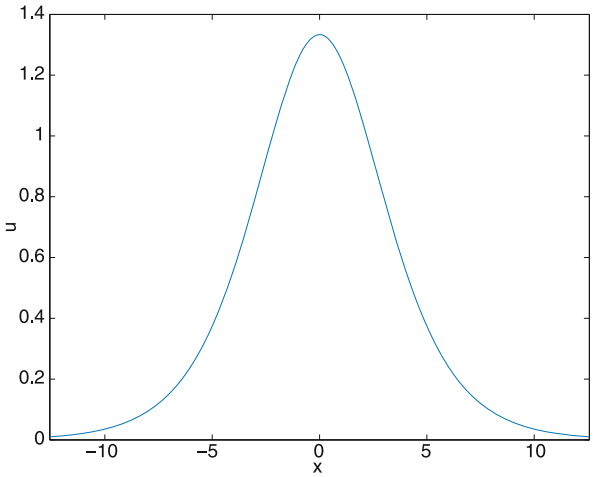
We obtain the solution in the form

$$u = v^{1/2} = \frac{4c}{3}\cos^2\left[\frac{1}{4}(x - ct)\right]\,. \tag{16}$$

By a careful inspection, v can tend to a very small value or even zero, as a result, $1 - cv^{-1/2}$ tends to negative infinite,



Solitons and Compactons, Figure 3
Compaton wave without tails



Solitons and Compactons, Figure 4
Solitary wave with two tails

and Eq. (7) behaves like Eq. (8) with $k \to \infty$, the exponential tails vanish completely at the edge of the bell-shape (see Fig. 3):

$$u = \begin{cases} \frac{4c}{3}\cos^2\left[\frac{1}{4}(x - ct)\right]\,, & |x - ct| \le 2\pi \\ 0\,, & \text{otherwise.} \end{cases} \tag{17}$$

This is a compact wave. Unlike solitons (Fig. 4), compacton does not have exponential tails (Fig. 3).

Generalized Solitons and Compacton-like Solutions

Solitary solutions have tails, which can be best expressed by exponential functions. We can assume that a solitary

solution can be expressed in the following general form

$$u(\eta) = \frac{\sum_{n=-c}^{d} a_n \exp(n\eta)}{\sum_{m=-p}^{q} b_m \exp(m\eta)}, \tag{18}$$

where c, d, p, and q are positive integers which are unknown to be further determined, a_n and b_m are unknown constants. The unknown constants can be easily determined using Matlab, the method is called the Exp-function method.

We consider the modified KdV equation in the form:

$$u_t + u^2 u_x + u_{xxx} = 0\,. \tag{19}$$

Using a transformation: $u(x,t) = u(\xi), \xi = kx + \omega t$, we have

$$\omega u' + k u^2 u' + k^3 u''' = 0\,, \tag{20}$$

where prime denotes the differential with respect to ξ.

We suppose that the solution of Eq. (20) can be expressed as

$$u(\xi) = \frac{a_c \exp(c\xi) + \cdots + a_{-d}\exp(-d\xi)}{b_p \exp(p\xi) + \cdots + b_{-q}\exp(-q\xi)}\,. \tag{21}$$

To determine values of c, d, p and q, we balance the linear term of highest order in Eq. (20) with the highest order nonlinear term. According to the homogeneous balance principle, we obtain the result $c = p$ and $d = q$. For simplicity, we set $c = p = 1$ and $d = q = 1$, so Eq. (21) reduces to

$$u(\xi) = \frac{a_1 \exp(\xi) + a_0 + a_{-1}\exp(-\xi)}{\exp(p\xi) + b_0 + b_{-1}\exp(-\xi)}\,. \tag{22}$$

Substituting Eq. (22) into Eq. (20), and by the help of Matlab, clearing the denominator and setting the coefficients of power terms like $\exp(j\xi)$, $j = 1, 2, \cdots$ to zero yield a system of algebraic equations, solving the obtained system, we obtain the following exact solutions:

$$\begin{cases} a_0 = a_1 b_0 + \dfrac{3k^2 b_0}{a_1}\,, & a_{-1} = \dfrac{b_0^2\left(3k^2 + 2a_1^2\right)}{8a_1}\,, \\ b_{-1} = \dfrac{b_0^2\left(3k^2 + 2a_1^2\right)}{8a_1^2}\,, & \omega = -ka_1^2 - k^3\,, \end{cases} \tag{23}$$

where a_1 and b_0 are free parameters, which depends upon the initial conditions and/or boundary conditions. The property that stability may depend on initial/boundary conditions is characteristic only for nonlinear systems.

The relationship between wave speed and frequency is

$$\omega = -ka_1^2 - k^3\,. \tag{24}$$

Note that the value of a_1 is determined from the initial/boundary conditions, so frequency or wave speed may not independent of initial/boundary conditions.

Then, the closed form solution of Eq. (19) reads

$$\begin{aligned} &u(x,t) \\ &= \frac{\left(a_1 \exp[kx - (ka_1^2 + k^3)t] + a_1 b_0 \right.}{\left. + \frac{3k^2 b_0}{a_1} + \frac{b_0^2(3k^2+2a_1^2)}{8a_1}\exp[-kx + (ka_1^2 + k^3)t]\right)} \\ &\quad \frac{\left(\exp[kx - (ka_1^2 + k^3)t] + b_0 \right.}{\left. + \frac{b_0^2(3k^2+2a_1^2)}{8a_1^2}\exp[-kx + (ka_1^2 + k^3)t]\right)} \\ &= a_1 + \frac{\frac{3k^2 b_0}{8}}{\left(\exp[kx - (ka_1^2 + k^3)t] + b_0 \right.} \\ &\quad \frac{}{\left. + \frac{b_0^2(3k^2+2a_1^2)}{8a_1^2}\exp[-kx + (ka_1^2 + k^3)t]\right)}\,. \end{aligned} \tag{25}$$

Generally a_1, b_0 and k are real numbers, and the obtained solution, Eq. (25), is a generalized soliton solution.

If we choose $k = 1, a_1 = 1, b_0 = \sqrt{8/5}$, Eq. (25) becomes

$$u(x,t) = 1 + \frac{3\sqrt{1/40}}{\exp[x - 2t] + \sqrt{8/5} + \exp[-x + 2t]}\,. \tag{26}$$

The bell-like solution is illustrated in Fig. 5.

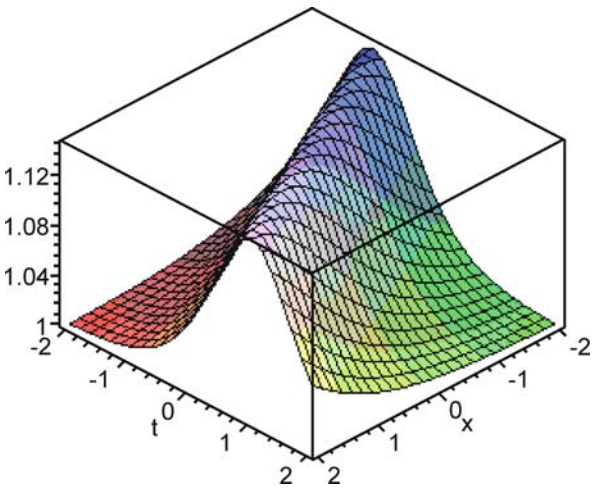
In case k is an imaginary number, the obtained solitary solution can be converted into periodic solution or compact-like solution. We write $k = iK$, Eq. (25) becomes

$$\begin{aligned} &u(x,t) = a_1 \\ &+ \frac{-\frac{3K^2 b_0}{8}}{\left((1+p)\exp[Kx - (Ka_1^2 - K^3)t] + b_0 \right.} \\ &\quad \frac{}{\left. + i(1-p)\frac{b_0^2(3k^2+2a_1^2)}{8a_1^2}\exp[-Kx + (Ka_1^2 - K^3)t]\right)}\,, \end{aligned} \tag{27}$$

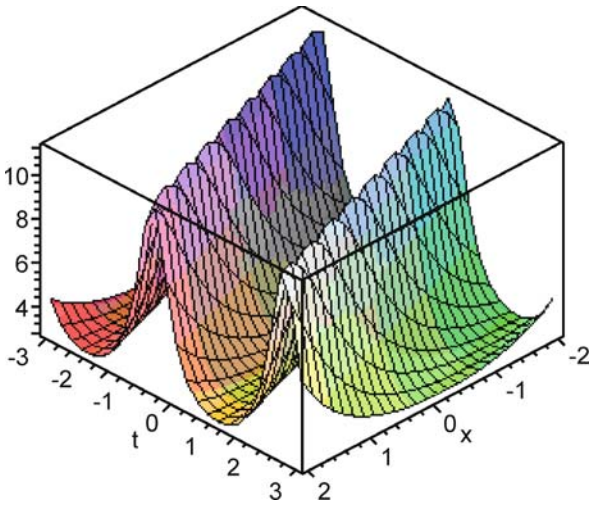
where $p = \frac{b_0^2(-3K^2+2a_1^2)}{8a_1^2}$.

If we search for a periodic solution or compact-like solution, the imaginary part in the denominator of Eq. (27) must be zero, that requires that

$$1 - p = 1 - \frac{b_0^2(-3K^2 + 2a_1^2)}{8a_1^2} = 0\,. \tag{28}$$



Solitons and Compactons, Figure 5
Propagation of a solution with respect to time



Solitons and Compactons, Figure 6
Periodic solution

Solving b_0 from Eq. (28), we obtain

$$b_0 = \pm\sqrt{\frac{8}{-3K^2 + 2a_1^2}}\,. \tag{29}$$

Substituting Eq. (29) into Eq. (27) results in a periodic solution, which reads

$$u(x,t) = a_1 + \frac{\pm 3K^2\sqrt{\frac{2}{-3K^2+2a_1^2}}}{\cos[Kx - (Ka_1^2 - K^3)t] \pm \sqrt{\frac{2}{-3K^2+2a_1^2}}} \tag{30}$$

or a generalized compact-like solution:

$$u(x,t) = \begin{cases} a_1 + \dfrac{\pm 3K^2\sqrt{\frac{2}{-3K^2+2a_1^2}}}{\left(\cos[Kx-(Ka_1^2-K^3)t] \pm \sqrt{\frac{2}{-3K^2+2a_1^2}}\right)}\,, \\ a_1 + 3K^2\,, \quad \text{otherwise} \end{cases}$$
$$\left|Kx - \left(Ka_1^2 - K^3\right)t\right| \le \frac{\pi}{2} \tag{31}$$

where a_1 and K are free parameters, and it requires that $2a_1^2 > 3K^2$. If we choose $k = 1, a_1 = 1, b_0 = \sqrt{8/5}$, Eq. (30) becomes

$$u(x,t) = 1 + \frac{3\sqrt{2}}{\cos[x - 3t] + \sqrt{2}}\,. \tag{32}$$

The periodic solution is illustrated in Fig. 6.

Now we give an heuristical explanation of why Eq. (19) behaves sometimes periodically and sometimes compaton-like.

We re-written Eq. (20) in form

$$u'' + \frac{\omega}{k^3}u + \frac{1}{3k^2}u^3 = 0. \tag{33}$$

It is a well-known Duffing equation with a periodic solution for all $\omega > 0$ and $k > 0$.

Actually in our study ω can be negative, we re-write Eq. (33) in the form

$$u'' - \frac{\omega}{k^3}u + \frac{1}{3k^2}u^3 = 0\,, \quad \omega > 0\,. \tag{34}$$

This equation, however, has not always a periodic solution. We use the parameter-expansion method to find its period and the condition to be an oscillator. In order to carry out a straightforward expansion like that in the classical perturbation method, we need to introduce a parameter, λ, because none appear explicitly in this equation. To this end, we seek an expansion in the form

$$u = u_0 + \lambda u_1 + \lambda^2 u_2 + \lambda^3 u_3 + \cdots\,. \tag{35}$$

The parameter λ is used as a bookkeeping device and is set equal to unity.

The coefficients of the linear term and nonlinear term can be, respectively, expanded in a similar way:

$$-\frac{\omega}{k^3} = \Omega^2 + m_1\lambda + m_2\lambda^2 + \ldots \tag{36}$$

$$\frac{1}{3k^2} = n_1\lambda + n_2\lambda^2 + \ldots\,, \tag{37}$$

where m_i and n_i are unknown constants to be further determined.

Interpretation of why such expansions work well is given by [1].

Substituting Eqs.(35)–(37) to (34), we have

$$\begin{aligned}&\left(u_0+\lambda u_1+\lambda^2u_2+\dots\right)''\\&+\left(\Omega^2+m_1\lambda+m_2\lambda^2+\dots\right)\left(u_0+\lambda u_1+\lambda^2u_2+\dots\right)\\&+\left(n_1\lambda+n_2\lambda^2+\dots\right)\cdot\left(u_0+\lambda u_1+\lambda^2u_2+\dots\right)^3=0\end{aligned}\tag{38}$$

and equating coefficients of like powers of λ, we obtain

Coefficient of λ^0

$$u_0''+\Omega^2u_0=0\,.\tag{39}$$

Coefficient of λ^1

$$u_1''+\Omega^2u_1+m_1u_0+n_0u_0^3=0\,.\tag{40}$$

The solution of Eq. (39) is

$$u_0=A\cos\Omega t\,.\tag{41}$$

Substituting u_0 into (40) gives

$$\begin{aligned}u_1''+\Omega^2u_1+A(m_1+\tfrac{3}{4}n_0A^2)\cos\Omega t\\+\tfrac{1}{4}n_0A^3\cos3\Omega t=0\,.\end{aligned}\tag{42}$$

No secular term in u_1 requires that

$$m_1+\frac{3}{4}n_0A^2=0\quad\text{or}\quad A=0\,.\tag{43}$$

If the first-order approximate solution is searched for, then we have

$$-\frac{\omega}{k^3}=\Omega^2+m_1\tag{44}$$

$$\frac{1}{3k^2}=n_1\,.\tag{45}$$

We finally obtain the following relationship

$$\Omega^2=-\frac{\omega}{k^3}+\frac{1}{4k^2}A^2\,.\tag{46}$$

To behave like an oscillator requires that

$$-\frac{\omega}{k^3}+\frac{1}{4k^2}A^2>0\tag{47}$$

or

$$\frac{\omega}{k}<\frac{1}{4}A^2\,.\tag{48}$$

The amplitude A may strongly depend upon initial/boundary conditions which may determine the wave type of a nonlinear equation.

Now we approximate Eq. (34) in the form

$$u''+\frac{1}{k^2}\left(-\frac{\omega}{k}+\frac{A^2\cos^2\Omega t}{3}\right)u=0\,.\tag{49}$$

In case $|\Omega t|\to\pi/2$, the above equation behaves exponentially, resulting in a compact-like wave as discussed above.

Future Directions

It is interesting to identify the conditions for a nonlinear equation to have solitary, or periodic, or compacton-like solutions. In most open literature, many papers on soliton and compacton are focused themselves on a special solution with either a soliton or a compacton without considering the initial/boundary conditions, which might be vital important for its actual wave type.

Solitons and compactons for difference-differential equations (e. g. Lotka–Volterra-like problems) have been caught much attention due to the fact that discrete space-time may be the most radical and logical viewpoint of reality (refer to E-infinity theory detailed concept). For small scales, e. g., nano scales, the continuum assumption becomes invalid, and difference equations have to be used for space variables.

Fractional differential model is another compromise between the discrete and the continuum, and can best describe solitons and compactons.

Many interesting phenomena arise in nanohydrodynamics recently, such as remarkably excellent thermal and electric conductivity, and extremely extraordinary fast flow in nanotubes. Consider a single compacton wave along a nanotube, and its wavelength is as same as the diameter of the nanotubes, under such a case, almost no energy is lost during the transportation, resulting in extremely extraordinary fast flow in the nanotubes.

The physical understanding of the transformation $k=iK$ is also worth further studying.

Cross References

► Soliton Perturbation

Bibliography

Primary Literature

1. He JH (2006) New interpretation of homotopy perturbation method. Int J Mod Phys 20(18):2561–2568

Some Famous Papers on Solitons and Compactons

2. Burger S, Bongs K, Dettmer S et al (1999) Dark solitons in Bose–Einstein condensates. Phys Rev Lett 83(25):5198–5201
3. Denschlag J, Simsarian JE, Feder DL et al (2000) Generating solitons by phase engineering of a Bose–Einstein condensate. Science 287(5450):97–101
4. Diakonov D, Petrov V, Polyakov M (1997) Exotic anti-decuplet of baryons: prediction from chiral solitons. Z Phys A 359(3):305–314
5. Duff MJ, Khuri RR, Lu JX (1995) Sting Solitons. Phys Rep 259(4–5):213–326
6. Eisenberg HS, Silberberg Y, Morandotti R et al (1998) Discrete spatial optical solitons in waveguide arrays. Phys Rev Lett 81(16):3383–3386
7. Fleischer JW, Segev M, Efremidis NK, et al (2003) Observation of two-dimensional discrete solitons in optically induced nonlinear photonic lattices. Nature 422(6928):147–150
8. He JH (2005) Application of homotopy perturbation method to nonlinear wave equations. Chaos Soliton Fract 26(3):695–700
9. He JH, Wu XH (2006) Construction of solitary solution and compacton-like solution by variational iteration method. Chaos Soliton Fract 29(1):108–113
10. Khaykovich L, Schreck F, Ferrari G et al (2002) Formation of a matter-wave bright soliton. Science 296(5571):1290–1293
11. Rosenau P (1997) On nonanalytic solitary waves formed by a nonlinear dispersion. Phys Lett A 230(5–6):305–318
12. Rosenau P (2000) Compact and noncompact dispersive patterns. Phys Lett A 275(3):193–203
13. Strecker KE, Partridge GB, Truscott AG et al (2002) Formation and propagation of matter-wave soliton trains. Nature 417(6885):150–153
14. Torruellas WE, Wang Z, Hagan DJ et al (1995) Observation of 2-dimensional spatial solitary waves in a quadratic medium. Phys Rev Lett 74(25):5036–5039

Review Article

15. He JH (2006) Some asymptotic methods for strongly nonlinear equations. Int J Mod Phys B 20(10):1141–1199 and 20(18):2561–2568

Exp-function Method

16. He JH, Wu XH (2006) Construction of solitary solution and compacton-like solution by variational iteration method. Chaos Soliton Fract 29(1):108–113
17. He JH, Wu XH (2006) Exp-function method for nonlinear wave equations. Chaos Soliton Fract 30(3):700–708
18. Zhu SD (2007) Exp-function method for the Hybrid–Lattice system. Int J Nonlinear Sci 8(3):461–464
19. Zhu SD (2007) Exp-function method for the discrete mKdV lattice. Int J Nonlinear Sci 8(3):465–468

Parameter-Expansion Method

20. He JH (2001) Bookkeeping parameter in perturbation methods. Int J Nonlinear Sci Numer Simul 2(3):257–264
21. He JH (2002) Modified Lindstedt–Poincare methods for some strongly non-linear oscillations Part I: Expansion of a constant. Int J Nonlinear Mech 37(2):309–314
22. He JH (2006) Non-perturbative methods for strongly nonlinear problems. dissertation.de-Verlag im Internet GmbH, Berlin
23. Shou DH, He JH (2007) Application of parameter-expanding method to strongly nonlinear oscillators. Int J Nonlinear Sci Numer Simul 8(1):121–124
24. Xu L (2007) Application of He's parameter-expansion method to an oscillation of a mass attached to a stretched elastic wire. Phys Lett A 368(3–4):259–262
25. Xu L (2007) Determination of limit cycle by He's parameter-expanding method for strongly nonlinear oscillators. J Sound Vib 302(1–2):178–184

Nanohydrodynamics and Nano-effect

26. He JH, Wan Y-Q, Xu L (2007) Nano-effects, quantum-like properties in electrospun nanofibers. Chaos Solitons Fract 33(1), 26–37
27. Majumder M, Chopra N, Andrews R et al (2005) Nanoscale hydrodynamics – Enhanced flow in carbon nanotubes. Nature 438(7064):44–44

E-Infinity Theory

28. El Naschie MS (2007) A review of applications and results of E-infinity theory. Int J Nonlinear Sci Numer Simul 8(1):11–20
29. El Naschie MS (2007) Deterministic quantum mechanics versus classical mechanical indeterminism. Int J Nonlinear Sci Numer Simul 8(1):5–10

Fractional-Order Differential Equations

30. Draganescu GE (2006) Application of a variational iteration method to linear and nonlinear viscoelastic models with fractional derivatives. J Math Phys 47(8):082902
31. He JH (1998) Approximate analytical solution for seepage flow with fractional derivatives in porous media. Comput Methods Appl Mech Eng 167(1–2):57–68
32. Momani S, Odibat Z (2007) Homotopy perturbation method for nonlinear partial differential equations of fractional order. Phys Lett A 365(5–6):345–350
33. Odibat ZM, Momani S (2006) Application of variational iteration method to Nonlinear differential equations of fractional order. Int J Nonlinear Sci Numer Simul 7(1):27–34
34. Wang Q (2007) Homotopy perturbation method for fractional KdV equation. Appl Math Comput 190(2):1795–1802

Solitons: Historical and Physical Introduction

François Marin
Laboratoire Ondes et Milieux Complexes, Fre CNRS 3102, Le Havre Cedex, France

Article Outline

Glossary
Definition of the Subject

Introduction
Historical Discovery of Solitons
Physical Properties of Solitons and Associated Applications
Mathematical Methods Suitable for the Study of Solitons
Future Directions
Bibliography

Glossary

Breaking waves As waves increase in height through the shoaling process, the crest of the wave tends to speed up relative to the rest of the wave. Waves break when the speed of the crest exceeds the speed of the advance of the wave as a whole.

Crystal lattice A geometric arrangement of the points in space at which the atoms, molecules, or ions of a crystal occur.

Deep water Water sufficiently deep that surface waves are little affected by the ocean bottom. Water deeper than one-half the surface wave length is considered deepwater.

Fluxon Quantum of magnetic flux.

Freak waves Single waves which result from a local focusing of wave energy. They are of considerable danger to mariners because of their unexpected nature.

Geostrophic adjustment The process by which an unbalanced atmospheric flow field is modified to geostrophic equilibrium, generally by a mutual adjustment of the atmospheric wind and pressure fields depending on the initial horizontal scale of the disturbance.

Geostrophic equilibrium A state of motion of an inviscid fluid in which the horizontal Coriolis force exactly balances the horizontal pressure force at all points of the field.

Hydraulic jump A sudden turbulent rise in water level, such as often occurs at the foot of a spillway when the velocity of rapidly flowing water is instantaneously slowed.

Katabatic wind Most widely used in mountain meteorology to denote a downslope flow driven by cooling at the slope surface during periods of light larger-scale winds.

Lightning Lightning is a transient, high-current electric discharge.

Plasma Hot, ionized gas.

Shallow water Water depths less than or equal to one half of the wavelength of a wave.

Solitary wave Localized wave that propagates along one space direction only, with undeformed shape.

Soliton Large-amplitude pulse of permanent form whose shape and speed are not altered by collision with other solitary waves, the exact solution of a nonlinear equation.

Spillway A feature in a dam allowing excess water to pass without overtopping the dam.

Thermocline A layer in which the temperature decreases significantly (relative to the layers above and below) with depth.

Synoptic scale Used with respect to weather systems ranging in size from several hundred kilometers to several thousand kilometers.

Thunder The sound emitted by rapidly expanding gases along the channel of a lightning discharge.

Thunderstorm In general, a local storm, invariably produced by a cumulonimbus cloud and always accompanied by lightning and thunder, usually with strong gusts of wind, heavy rain, and sometimes with hail.

Tidal bore Tidal wave that propagates up a relatively shallow and sloping estuary or river, in a solitary wave form. The leading edge presents an abrupt rise in level, sometimes with continuous breaking and often immediately followed by several large undulations. The tidal bore is usually associated with high tidal range and a sharp narrowing and shoaling at the entrance. Also called pororoca (Brazilian) and mascaret (French).

Troposphere The portion of the atmosphere from the earth's surface to the tropopause, that is the lowest 10–20 km of the atmosphere.

Tsunami Long period ocean wave generated by an earthquake or a volcanic explosion.

Definition of the Subject

The interest in nonlinear physics has grown significantly over the last fifty years. Although numerous nonlinear processes had been previously identified the mathematic tools of nonlinear physics had not yet been developed. The available tools were linear, and nonlinearities were avoided or treated as perturbations of linear theories. The solitary water wave, experimentally discovered in 1834 by John Scott Russell, led to numerous discussions. This hump-shape localized wave that propagates along one space-direction with undeformed shape has spectacular stability properties. John Scott Russell carried out many experiments to obtain the properties of this wave. The theories which were based on linear approaches concluded that this kind of wave could not exist. The controversy was resolved by J. Boussinesq [5] and by Lord Rayleigh [64] who showed that if dissipation is neglected, the increase in local wave velocity associated with finite amplitude is balanced by the decrease associated with dispersion, leading