K

Knowledge Discovery: Clustering

PAVEL BERKHIN[1], INDERJIT S. DHILLON[2]
[1] eBay Inc., San Jose, USA
[2] University of Texas, Austin, USA

Article Outline

Glossary
Definition of the Subject
Introduction
Partitioning Relocation Clustering
Hierarchical Clustering
Spectral Clustering
Graph Clustering
Clusters as Dense Connected Components
Clustering Transactional Data
Co-clustering
Scalability
Other Aspects of Clustering
Future Directions
Bibliography

Glossary

BIRCH A scalable clustering solution that first assembles an in-memory balanced tree of micro-clusters (called clustering features) containing sufficient statistics for the data.

CACTUS Clustering algorithm for transactional data based on the idea of co-occurrence.

CHAMELEON Hierarchical clustering algorithm using a very fine partitioning of a sparsified connectivity graph followed by agglomeration phase.

Cluster A subset of the data that consists of similar objects.

CLIQUE Clustering algorithm for high-dimensional data that tries to find high density low-dimensional segments using an inductive process similar to the Apriori algorithm.

Co-clustering Clustering methodology that along with grouping points, also groups attributes that have similar distributions among data points.

Connectivity matrix A matrix of similarities or dissimilarities between data points, which gives pairwise connectivity information that is used in agglomerative, spectral and graph clustering algorithms.

CURE Important scalable hierarchical clustering algorithm that uses a fixed number of points as cluster representatives.

DBSCAN Density-based partitioning that utilizes a definition of density-connectivity and a core point (a point whose ϵ-neighborhood has sufficiently many points).

Density-based partitioning Clustering algorithm that tries to identify clusters with dense connected components of arbitrary shape.

GRACLUS A super-fast graph clustering algorithm that optimizes weighted graph cuts.

Graph clustering Clustering of graph nodes; instead of a concept of distance, a concept of connectivity based on graph edges is used.

Grid methods Clustering algorithms that find relatively high populated segments in an underlying attribute space and then assemble clusters from adjacent segments.

Hierarchical clustering Represents the data in the form of a tree dendrogram whose leaves correspond to individual points and nodes to clusters of different granularities.

***k*-means** Partitioning relocation clustering algorithm used in many applications that deals with numerical data; represents a cluster with its centroid or mean.

***k*-medoid methods** Partitioning relocation algorithm (e. g. CLARANS) that represents a cluster by one of its points, called medoid.

Linkage metric Metric used in hierarchical clustering in conjunction with the Lance-Williams updating formula to compute similarity between two subsets.

Probabilistic clustering Clustering algorithm that associates with each cluster a particular probability distribution whose parameters are fitted by the algorithm.

Spectral clustering Clustering algorithm that uses eigenvectors or singular vectors for data or graph clustering.

Definition of the Subject

Data that we find in scientific and business applications usually does not fit a particular parameterized probability distribution. In other words, the data is *complex*. Knowledge discovery starts with exploration of this complexity in order to find inconsistencies, artifacts, errors, etc in the data. After data is cleaned, it is usually still extremely complex. Descriptive data mining deals with comprehending and reducing this complexity. Clustering is a premier methodology in descriptive unsupervised data mining.

Clustering is the division of N data points $X = \{x_1, x_2, \dots, x_N\}$ into k disjoint groups C_i. Each group, called *cluster*, is required to consist of points that are similar to one another and dissimilar to points in other groups:

$$X = C_1 \bigcup \cdots \bigcup C_k \bigcup C_{\text{outliers}}\,, C_i \bigcap C_j = 0\,, \quad i \neq j\,.$$

A cluster could represent an important subset of the data such as a galaxy in astronomical data or a segment of customers in marketing applications. Clustering is important as a fundamental technology to reduce data complexity and to find data patterns in an unsupervised fashion. It is universally used as a first technology of choice in data exploration.

Introduction

Classic clustering algorithms have existed for a long time. Contemporary clustering faces many challenges, such as (a) sheer size of modern data sets, (b) objects with many attributes, (c) attributes of different types, and (d) unstructured data or data of complex structure. These challenges have led to the emergence of powerful and broadly applicable clustering methods. First, we should ask ourselves: does a set of axioms exists that would result in a consistent clustering framework? A sobering answer was given by Jon Kleinberg [49]: under natural assumptions, clustering cannot be axiomatized. General references on clustering include [7,25,28,33,39,40,43,44,47,62]. From a statistical standpoint, clustering relates to a traditional multivariate statistical estimation.

System complexity provides another fruitful way of looking at clustering. As a result of clustering, data complexity is reduced to a small number of clusters. More precisely, to transmit data we can transmit (1) k cluster "prototypes"; (2) each point's cluster ID and a relatively short encoding that describes deviation of a point from its cluster "prototype". This connection to data compression in used in image processing (*vector quantization* [32]). Other clustering applications include scientific data analysis (astronomy), biochemistry and medicine, information retrieval and text mining, spatial database applications, sequence and heterogeneous data analysis, web applications, marketing, user segmentation, fraud detection, and many others.

We use the following notation. Dataset X consists of data points (*objects*, *instances*, *cases*, etc.) $x_i = (x_{i1}, \cdots, x_{id})$, $i = 1 : N$, in d-dimensional attribute space A, $x_{il} \in A_l$, $l = 1 : d$. This point-by-attribute data format conceptually corresponds to a $N \times d$ matrix and is used by the majority of algorithms. A simplest subset in A, a *segment*, is a direct Cartesian product of sub-ranges. Some clustering algorithms indeed use segments as building blocks for clusters. Data in other formats, such as variable length sequences and heterogeneous data, are not uncommon.

Partitioning Relocation Clustering

If you know nothing about clustering, the k-means algorithm is what you should start with and many people never go beyond it in their practice.

The name comes from representing each of the k clusters C_j by the mean (or weighted average) c_j of its points, the so-called *centroid*. While this representation does not work well with categorical attributes, it makes sense from a geometric and statistical perspective for numerical attributes. K-Means tries to minimize the objective function equal to the sum of the squares of L_2-norm errors between the points and the corresponding centroids:

$$E(C) = \sum_{j=1:k} \sum_{x_i \in C_j} \left\| x_i - c_j \right\|^2 .$$

The above can be thought of as the within-cluster variance. It turns out that the total data variance can be expressed as the sum of the within-cluster variance and the between-cluster variance:

$$\sum_{i=1}^{N} \|x_i - c\|^2 = \sum_{j=1:k} \sum_{x_i \in C_j} \left\| x_i - c_j \right\|^2 + \sum_{j=1:k} n_j \left\| c_j - c \right\|^2 ,$$

where $n_j = |C_j|$. Thus while k-means explicitly tries to minimize the within-cluster variance, it implicitly maximizes the between-cluster variance. By examining the k-means objective carefully, it is easy to show that k-means is restricted to separating clusters by linear separators, i. e., by hyperplanes.

Up to a constant, $E(C)$ can be recognized as the negative of the log-likelihood for a normally distributed mixture model with uniform variance (points in C_j are distributed as $N(c_j, 1)$). Therefore, the k-means algorithm is related to a general probabilistic framework. This suggests generalizations: fit not only means, but variances, or hyper-ellipsoidal clusters.

The squared L_2-distance has many unique properties. For example, $E(C)$ equals the sum of pair-wise errors within all the clusters:

$$E'(C) = \frac{1}{2} \sum_{j=1:k} \sum_{x_i, y_i \in C_j} \|x_i - y_i\|^2 .$$

Other dissimilarity measures can also be used with a k-means like algorithm. Given a dissimilarity measure $d(x, y)$, the representative vector of a cluster can be defined as:

$$z_j = \text{argmin}_z \sum_{x_i \in C_j} d(x_i, z) .$$

For the squared L_2 distance, z_j simply equals the mean vector c_j. It turns out that the same result is true, i. e., the mean vector is the representative vector for a much larger class of dissimilarity measures called Bregman divergences [4]. These divergences are not always symmetric and they do not obey the triangle inequality but they have many other desirable properties.

The exact optimum for $E(C)$ cannot be computed, but two versions of k-means iterative optimization converging to a local minimum are known. The first version is similar to the EM algorithm. It makes two-step major iterations: (1) reassign points to their nearest centroids; (2) recompute centroids of newly assembled groups. Iterations continue until a stopping criterion is achieved. The result is independent of data ordering, and straightforward parallelization can be applied:

Initialize centroids c_i, $i = 1 : k$
Until convergence is achieved do
  for each $x \in X$ ** Step 1: Reassign points
    $I(x) = \text{argmin}\{\|c_i - x\|\}$
  for each $1 = 1 : k$ ** Step 2: Recompute centroids
    $c_i = mean\{x \in X : I(x) = i\}$

The second version, tries to readjust centroids as soon as reassignment happens. It is not obvious that it is computationally feasible, but it is: in fact, the computational complexity of both versions is the same. This second version can actually result in a better optimum, but it depends on the ordering of points, and is somewhat more difficult to implement:

Initialize centroids c_i, $i = 1 : k$
Until convergence is achieved do
  for each $x \in X$ ** Iterate over all points
    let $x \in C_i$, *and let* $d_i > 0$ *denote the change in* $E(C)$ *on deleting* x *from* C_i
    for each $j \neq i$ *let* $-d_j$ *be the change in* $E(C)$ *on adding* x *to* C_j
    $j = \text{argmax}\{d_i - d_j\}$
    if $d_i - d_j > 0$ *reassign* x *to cluster* C_j *and recompute* c_i *and* c_j

The popularity of the k-means algorithm is well deserved: it is easily understood, easily implemented, and based on the firm foundation of analysis of variances. It also has shortcomings:

- During the reassignment stage a cluster can become empty or unbalanced
- The computed local optimum may be quantitatively and qualitatively much worse than the global optimum
- Initialization of centroids is crucial (see [10] for suggestions)
- Choice of k is unclear
- The process is sensitive to outliers
- Algorithm lacks scalability
- Only numerical attributes are covered in a straightforward manner

K-means is the most popular example of a family of clustering algorithms called *Partitioning Relocation*. Algorithms of this family try to reassign points from one potential cluster to another in order to achieve some objective through iterative optimization. For example, in *Probabilistic Clustering* the data is represented as a mixture of k models whose parameters we want to reconstruct. Each model is expressed as a probability distribution – it turns out that a rich family of probability distributions, namely, the exponential family is in one-to-one correspondence with Bregman divergences [4], that were discussed earlier. The exponential family includes multivariate Gaussians, the Poisson, Bernoulli and exponential distributions.

Clusters discovered by probabilistic clustering are conveniently *interpretable*.

We assume that data points are generated (a) by randomly picking a model j (cluster) with probability τ_j, $j = 1 : k$, and (b) by drawing a point x from a corresponding distribution. Maximization of log-likelihood $\log(L)$, where

$$L = \Pr(X|C) = \prod_{i=1:N} \sum_{j=1:k} \tau_j \Pr\left(x_i | C_j\right) ,$$

is achieved through *Expectation-Maximization* (EM) iterations, similar to used in k-means: (E) recompute membership probabilities, (M) estimate model parameters (maximize likelihood). Advantages of this method are:

- Points of complex structure can be handled (heterogeneous data, dynamic sequences) by appropriate probabilistic modeling
- Iterations can be stopped and resumed; intermediate model is available
- Clusters have conceptual meaning
- Number of parameters, in particular k, can be addressed within the Bayesian framework

An example of probabilistic clustering is the algorithm AUTOCLASS [15] that covers a broad variety of probability distributions. including Bernoulli, Poisson, Gaussian, and log-normal distributions. Beyond fitting a particular fixed mixture model, AUTOCLASS extends the search to different models and different values of k.

Another *Partitioning Relocation* family, namely k-medoids methods, represents a cluster by one of its points called a *medoid*. This is an easy solution: any attribute type can be handled, no guess on a probability distributions is required, clusters are subsets of points close to respective medoids, and the objective function is defined as the averaged dissimilarity measure between a point and its medoid.

An example is the CLARANS algorithm (Clustering Large Applications based upon RANdomized Search) [56] that deals with spatial databases. In CLARANS a search over a graph whose nodes are subsets of k medoids (points) is performed. Two nodes are connected by an edge if they differ by exactly one medoid. CLARANS is extended to large databases in [26]; this extension relies heavily on data indexing.

Hierarchical Clustering

Hierarchical clustering combines data points into clusters, those clusters into larger clusters, and so forth, creating a hierarchy. A tree representing this hierarchy of clusters is known as a *dendrogram*. Individual data points are the tree leaves. A dendrogram allows the exploration of data at different levels of granularity. An *agglomerative* hierarchical clustering algorithm starts with one-point (singleton) clusters and recursively merges two or more of the most similar clusters. A *divisive* hierarchical clustering starts with a single cluster containing all data points and recursively splits that cluster into appropriate sub-clusters. The process may be terminated when a stopping criterion (frequently, the requested number k of clusters) is achieved. The advantages of hierarchical clustering include flexibility regarding the level of granularity, ease of handling any form of similarity or distance, and applicability to any attribute type. On the negative side, hierarchical clustering faces the difficulty of choosing the right stopping criteria and most hierarchical algorithms do not revisit (intermediate) clusters once they are constructed.

To explore the topic further, consider agglomerative clustering. An $N \times N$ matrix of distances or similarities between points, called *connectivity* matrix, is used to find closest singleton data points to merge together. To merge or split subsets of points rather than individual points, the distance between individual points has to be generalized to the distance between subsets. Such a derived proximity measure is called a *linkage metric*. It has a significant impact on hierarchical algorithms, because it reflects a particular concept of *closeness* and *connectivity*. Important inter-cluster linkage metrics include *single link*, *average link*, and *complete link*. The linkage metric between two subsets of nodes is computed by applying an aggregator operation Op to pairs of dissimilarities between nodes in the first subset C_1 and nodes in the second subset C_2:

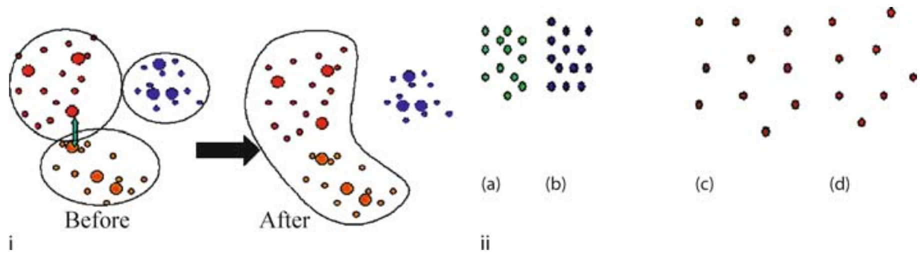
$$d(C_1, C_2) = Op\,\{d(x,y),\ x \in C_1,\ y \in C_2\} .$$

Examples of Op include minimum (single link, algorithm SLINK [61], $Op = \min$), average (average link, Voorhees' method [64], $Op = \text{Avr}$), or maximum (complete link, algorithm CLINK [19], $Op = \max$). All these linkage metrics can be derived from the Lance–Williams updating formula [51],

$$\begin{aligned} d\left(C_i \bigcup C_j, C_k\right) &= a(i)d(C_i, C_k) + a(j)d(C_j, C_k) \\ &\quad + b \cdot d(C_i, C_j) + c\left|d(C_i, C_k) - d(C_j, C_k)\right| \end{aligned}$$

with coefficients a, b, c depending on a particular linkage metric. This formula helps to avoid actual computing of all pairwise dissimilarities that makes a definition computationally feasible, see [18].

Numerous connections to graph theory exist. SLINK, for example, is related to the problem of finding the Euclidean minimal spanning tree [67] and has $O(N^2)$ complexity. More importantly, when the $N \times N$ connectivity matrix is sparsified, graph methods directly dealing with the connectivity graph can be used. In particular, the hierarchical divisive MST (Minimum Spanning Tree) algorithm is based on partitioning the graph [44].



Knowledge Discovery: Clustering, Figure 1
Agglomeration of clusters of arbitrary shapes. i Algorithm CURE ii Algorithm CHAMELEON

Guha et al. [36] introduced the hierarchical agglomerative clustering algorithm CURE (Clustering Using REpresentatives). This algorithm has a number of novel and important features. CURE is the first hierarchical clustering algorithm that was designed with a scalability requirement in mind. It also takes special steps to handle outliers and to provide labeling in the assignment stage. In CURE a cluster is represented by a fixed number, c, of points scattered around it. Selecting representatives scattered around a cluster allows to cover non-spherical shapes. The distance between two clusters used in the agglomerative process is the minimum of distances between two scattered representatives and combines ideas of single and average link closeness. Agglomeration continues until the requested number k of clusters is achieved. CURE employs one additional trick: the original selected scattered points are shrunk to the geometric centroid of the cluster by a user-specified factor α. Shrinkage decreases the impact of outliers; outliers happen to be located further from the cluster centroid than the other scattered representatives. CURE is capable of finding clusters of different shapes and sizes. Figure 1(i) illustrates agglomeration in CURE. Three clusters, each with three representatives, are shown before and after the merge and shrinkage. The two closest representatives are connected. The algorithm CURE works with numerical attributes (particularly, low dimensional spatial data). It is complemented by the algorithm ROCK that handles categorical attributes.

The hierarchical algorithm CHAMELEON, [45], uses a sparsified connectivity graph G: edges corresponding to the K most similar points to any given point are preserved, and the rest are pruned. CHAMELEON performs both the steps of partitioning and agglomeration. It first partitions the data in small tight clusters by using the HMETIS library. Then it agglomerates these small micro-clusters taking into account local measures of connectivity and closeness. The CHAMELEON algorithm does not depend on assumptions about the data model, and has been shown to find clusters of different shapes, densities, and sizes in two-dimensional space. CHAMELEON has complexity $O(Nm + N\log(N) + m^2\log(m))$, where m is the number of micro-clusters built during the first partitioning phase. Figure 1ii clarifies the difference between CHAMELEON and CURE – it shows a choice of four clusters (a)–(d) for a merge. While CURE would merge clusters (a) and (b), CHAMELEON makes the intuitively better choice of merging (c) and (d).

Finally, k-way partitions, when available, naturally lead to divisive hierarchical algorithms. For example, k-means can be used to first divide the data into k clusters, and then it can be recursively applied to each of the k-clusters to yield a hierarchical divisive partitioning. The singular value decomposition (SVD) yields a spectral hierarchical divisive clustering method for document collections called PDDP (Principal Direction Divisive Partitioning) [9]. PDDP is an algorithm that uses the SVD to repeatedly bisect the data into two clusters. In our notation, point x_i is a document, its lth attribute corresponds to a word (*index term*), and x_{il} is a measure (e. g. TF-IDF) of the frequency of term l in document x_i. PDDP begins with the SVD of the matrix

$$(X - \bar{x}\mathrm{e}^{\mathrm{T}}), \quad \text{where} \quad \bar{x} = \frac{1}{N}\sum_{i=1:N} x_i, \mathrm{e} = (1, \ldots, 1)^{\mathrm{T}}.$$

PDDP bisects data in Euclidean space by a hyperplane that passes through the data centroid $\bar{x}$ and is orthogonal to the singular vector with the largest singular value.

Spectral Clustering

As mentioned above, the PDDP method uses spectral information, namely, singular vectors for data clustering. In general, the methods of spectral clustering use such spectral information for clustering. Spectral clustering can be viewed as first constructing an $N \times N$ *connectivity matrix* (explicitly or implicitly) between all data points. As an example, the (i, j) entry of the connectivity matrix may be formed to be $\mathrm{e}^{-\|x_i - x_j\|^2/2\sigma^2}$, or $(x_i - \bar{x})^{\mathrm{T}}(x_j - \bar{x})$ (as done implicitly in PDDP). Then the eigenvectors of the connectivity matrix are computed, typically only the leading few.

From these eigenvectors, the clustering of the data can be extracted in a myriad of ways, as discussed below.

Although spectral clustering was introduced to the data mining and machine learning communities recently [55], it has had a long history in graph clustering problems that arise in a variety of applications described in Sect. "Graph Clustering". The connection to graph clustering is not a surprise since the entries of the connectivity matrix may be viewed as weights of edges between vertices that correspond to data points.

There are three main computational issues in spectral clustering. Construction of the connectivity matrix might take time quadratic in the number of data items, which might be prohibitive for large-scale applications. The connectivity matrix is large in size if the number of data points is large, and can often be sparsified, i. e., many of the entries can be thresholded to zero. Computation of the eigenvectors of such a large, sparse symmetric matrix, is done by typically invoking the Lanczos algorithm [57], which can again be a computational bottleneck if many eigenvectors are desired. A PDDP type recursive spectral bisection algorithm is more effective. An important issue in spectral clustering is to obtain the clusters once the eigenvectors are computed. If r eigenvectors $v_1, v_2, \dots, v_r$ of an $N \times N$ connectivity matrix are computed, then the ith components $v_1(i), v_2(i), \dots, v_r(i)$ may be viewed as the reduced dimensional representation of the ith data point. In this case, a simple way to obtain clusters from the eigenvectors is to run k-means on the reduced dimensional representations of the data points [55]. More complicated methods for obtaining the clusters can be used that perform better [68].

Graph Clustering

Graph clustering, also called graph partitioning, is applicable when the data is presented in the form of a graph, for example the link structure of the web, or a social network. The graph clustering problem is to partition or cluster the nodes of the given graph, such that the connectivity between partitions is minimized. The most popular measure of connectivity is the sum of the crossing edges or *cut* between the partitions. The edges of the graph have weights that reflect similarities between the vertices, and the cut is defined to be the sum of the weights of the crossing edges. The minimum cut problem is solvable in polynomial time, however the graph clustering problem has additional (explicit or implicit) constraints on the sizes of the partitions. For example, the partitions might be constrained to be equal in size [48], or a weighted cut objective, such as ratio-cut [14] or normalized-cut [60] might need to be minimized. These constraints make the graph clustering problem NP-hard.

Graph clustering has been employed in many applications, for example, in circuit layout, partitioning the workload among processors for parallel processing, image segmentation, and of course, in the analysis of networks that arise in data mining. Since the problem is NP-hard, several heuristics are employed to try and solve this important problem. Some of the early successful approaches include the greedy search heuristic of Kernighan and Lin [48]. Another approach to graph clustering is based on the idea of graph flows. A survey of this research is presented in [52].

It turns out that the graph clustering objective may be written as a quadratic programming objective with discrete constraints on the variables. However, if the variables are relaxed to be real-valued, then eigenvectors of a symmetric matrix called the Graph Laplacian can be shown to exactly solve the relaxed quadratic problem. Consider the adjacency matrix A of an undirected graph which is defined to be a matrix with entries $A(i, j) = 1$ if there is an edge between vertices i and j and zero otherwise. The Graph Laplacian L equals $D - A$, where D is a diagonal matrix whose ith diagonal entry equals the sum of all entries in the ith row of A. Hence the sum of each row of L is zero, which implies that the vector of all 1's is an eigenvector of L with eigenvalue 0. The eigenvector of L corresponding to the next smallest eigenvalue is often called the Fiedler vector [29] and since it is orthogonal to the all 1's eigenvector, it has positive as well as negative entries. The graph bi-partition can now be obtained by placing the vertices with positive entries in the Fiedler vector in one partition, while the ones with negative entries are placed in the other partition. Such spectral methods have a long and rich history in graph clustering, dating back to the early 1970s [23,29,37], and have been used to optimize various weighted graph clustering objectives, such as ratio cut in circuit layout [14] and normalized cut in image segmentation [60]. One of the reasons for the success of spectral methods is that these methods provide a globally optimal solution to the relaxed problem, and thus provide a good global heuristic to solve the actual graph clustering problem.

Spectral graph clustering can be computationally expensive, especially when many eigenvectors of a large, sparse matrix need to be computed in order to directly give a k-way cut. An alternative class of algorithms is made up of multilevel graph clustering algorithms, which are attractive, efficient and powerful alternatives. In multilevel algorithms, the input graph is repeatedly coarsened level by level until only a small number of nodes remain. An initial clustering is performed on the coarsened graph, and then this clustering is refined as the graph is uncoarsened level by level. These methods are extremely fast and give high-quality partitions. However, earlier multilevel methods, such as Metis [46] and Chaco [41], force clusters to be of nearly equal size, and are all based on optimizing the Kernighan-Lin objective [48]. In graph clustering for data mining problems, there is no reason why clusters should be of the same size. A recently developed multilevel graph clustering method, called GRACLUS (GRAph CLUStering) [22] removes the restriction of equal cluster sizes. In fact, this multilevel algorithm is able to optimize for several weighted spectral clustering objectives, such as ratio cut and normalized cut, without having to compute any eigenvectors. GRACLUS exploits a mathematical equivalence between two objectives – the seemingly different clustering objectives of weighted graph cuts, and weighted kernel k-means (an enhanced version of k-means) are shown to be equivalent in [22]. Using this equivalence, GRACLUS employs the weighted kernel k-means algorithm during the refinement phase to directly optimize various weighted graph cuts. Furthermore, this multilevel algorithm does not require any extra memory, which is in contrast to spectral methods that compute k eigenvectors for k-way cuts and require $O(Nk)$ storage, where N is the number of vertices. Thus multilevel algorithms are scalable to much larger data sets than standard spectral methods, and offer a state-of-the-art solution to the graph clustering problem.

Clusters as Dense Connected Components

In this section, a dataset X will be divided (with the exception of outliers) into dense connected components that will serve as its natural clusters. But how can the concepts of *density* and *connectivity* be defined for discrete data? Three approaches to answering this question exist: *Density-Based Partitioning*, *Density Functions*, and *Grid Methods*.

The first approach, *Density-Based Partitioning*, pins these concepts to a particular point $x \in X \subset A$.

Density and connectivity are defined in terms of a point's nearest neighbors. A cluster grows in any direction that density leads it to. Therefore, density-based algorithms are capable of discovering clusters of arbitrary shapes and they provide a natural protection against outliers. Figure 2 illustrates some cluster shapes that present problems for partitioning relocation clustering (e. g., k-means), but are handled properly by density-based algorithms. Finally, density-based algorithms are scalable.

These outstanding properties come along with certain inconveniences:

Knowledge Discovery: Clustering, Figure 2
Irregular shapes

- A dense cluster consisting of two adjacent areas with significantly different densities (both higher than a threshold) is not very informative
- Clusters lack interpretability
- Density-based partitioning primarily works with low-dimensional spatial data (though generalizations exist).

To explain the last point, note that scalability is achieved through usage of some sort of data index (such as an R^*-tree). Index construction is a topic of active research. Classic indices are effective only with low-dimensional data, since index complexity increases exponentially with data dimension. An excellent introduction to density-based methods is contained in [39].

As a representative of its class, consider the algorithm DBSCAN (Density Based Spatial Clustering of Applications with Noise) [27]. Two input parameters ϵ and MinPts are used to introduce:

1. An ϵ-*neighborhood* $N_\epsilon(x) = \{y \in X \mid \mathrm{dist}(x, y) \le \epsilon\}$ of the point x,
2. A *core object*, which is a point with $|N_\epsilon(x)| \ge \mathrm{MinPts}$,
3. A notion of a point y *density-reachable* from a core object x (a sequence of core objects between x and y exists such that each object in the sequence belongs to an ϵ-neighborhood of its predecessor),
4. A definition of *density-connectivity* between two points x and y (they should be density-reachable from a common core object).

Density-connectivity is an equivalence relation. All the points reachable from core objects can be factorized into maximal connected components serving as clusters. The core points are *internal* points. The non-core points inside a cluster represent its *boundary*. The points not connected to any core point are declared to be outliers, and are not covered by any cluster. For low-dimensional spatial data, the theoretical complexity of DBSCAN is $O\left(N \log(N)\right)$. Experiments confirm a slightly super-linear runtime.

Selection of two parameters ϵ and MinPts presents a problem. Besides, no choice allows the fitting of data that has different densities in different localities as frequently happens. Algorithms OPTICS (Ordering Points To Identify the Clustering Structure) [3] and DBCLASD (Distribution Based Clustering of Large Spatial Databases) [66] rectify DBSCAN to address these issues.

The second approaches to definition of density and connectivity is based on *Density Functions* and is due to Hinneburg and Keim [42]. They proposed the algorithm DENCLUE (DENsity-based CLUstEring). In this approach definitions are no longer pinned to a point $x \in X$, but to a point in the underlying attribute space $x \in A$. Data affects density indirectly through so-called density functions.

DENCLUE uses a *density function*

$$f^D(x) = \sum_{y \in D(x)} f(x, y)$$

that is the superposition of several *influence functions*, for example, $f(x, y) = \theta\left(\|x - y\| / \sigma\right)$ (equals to one, if the distance between x and y is less than or equal to σ) or Gaussian influence functions $f(x, y) = \exp(-\|x - y\|^2/\sigma^2)$. This provides a high level of generality: the first example leads to DBSCAN, and the second to k-means clustering. Both examples depend on the parameter σ.

Restricting the summation to $D(x) = \{y \in X : \|x - y\| < k\sigma\} \subset X$ enables a practical implementation. DENCLUE concentrates on local maxima of density function called *density-attractors* and uses a gradient hill-climbing technique to find them. In addition to *center-defined* clusters, *arbitrary-shape* clusters are defined as unions of local shapes along sequences of neighbors whose local densities are no less than a prescribed threshold.

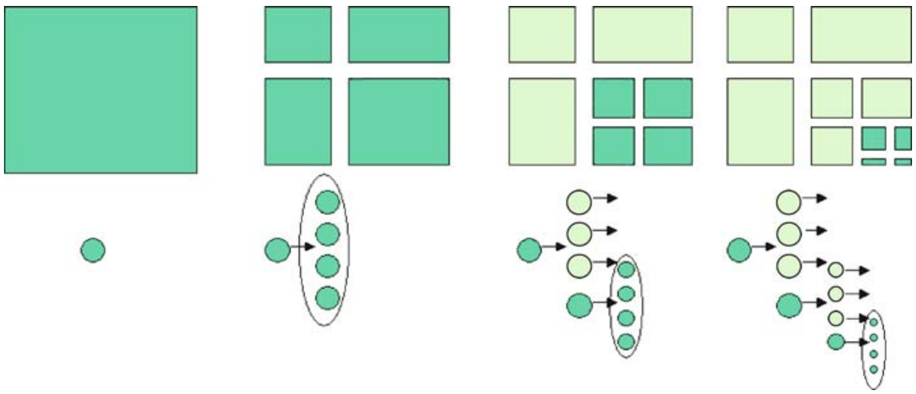
Applications include high dimensional multimedia and molecular biology data. While no clustering algorithm could have less than $O(N)$ complexity, the runtime of DENCLUE scales with N sub-linearly! The explanation is that though all the points are fetched, the bulk of the analysis in the clustering stage involves only points in highly populated areas.

The third approach, *Grid-Based Methods*, suggests exploiting structure of the underlying attribute space A and partitioning it rather than data X. Data partitioning is induced by a point's membership in segments resulting from space partitioning, while space partitioning is based on grid-characteristics accumulated from input data. One advantage of this indirect handling (data $\rightarrow$ grid-data $\rightarrow$ space-partitioning $\rightarrow$ data-partitioning) is that it is independent of the data ordering and it works with numerical and categorical attributes (actually numerical attributes are binned).

To limit the amount of computations, multi-rectangular segments (direct Cartesian product of individual attribute sub-ranges) are considered. The elementary segment whose sides correspond to single-bins is called a *unit*. To some extent, the grid-based methodology reflects a technical point of view: it contains both partitioning and hierarchical algorithms and is used as an intermediate step in many clustering algorithms including the just described algorithm DENCLUE (that uses grids at its initial stage), and the important algorithm CLIQUE presented later.

Few other examples are considered below.

The algorithm BANG [58] summarizes data over the segments. The segments are stored in a grid-directory incorporating different scales. Adjacent segments are considered neighbors. If a common face has maximum dimension they are called nearest neighbors. More generally, neighbors of degree (dimension of a common face) between 0 and $d - 1$ can be defined. The density of a segment is defined as a ratio between the number of points in



Knowledge Discovery: Clustering, Figure 3
Algorithm STING

that segment and its volume. From the grid-directory, a hierarchical clustering (a dendrogram) is calculated directly.

The algorithm STING (STatistical INformation Grid-based method) [65] works with numerical attributes (spatial data) and is designed to facilitate "region oriented" queries. STING assembles summary statistics in a hierarchical tree of nodes that are grid-cells. Figure 3 illustrates the proliferation of cells in 2D space and the construction of the corresponding tree. Each cell has four (default) children.

We want to briefly mention two other algorithms that work with numerical data and are based on certain mathematical ideas. The algorithm WaveCluster [59] supports a multi-resolution technique. It uses *wavelet* transforms and has several salient properties: (1) it can work with relatively high-dimensional data, (2) it successfully handles outliers, and (3) it has $O(N)$ complexity. The second algorithm, FC (Fractal Clustering) [6], utilizes another mathematical concept, *Hausdorff Fractal Dimension* (HFD). Its implementation relies on the hierarchy of grids: FC scans the full data incrementally attempting to add an incoming point to a cluster so as to minimally increase its HFD. The FC algorithm has its pros: (1) incremental structure, (2) suspendable nature, (3) ability to discover clusters of irregular shapes, (4) $O(N)$ complexity. Its cons: dependency on (1) data ordering, (2) on cluster initialization, and (3) on some input parameters.

Clustering Transactional Data

In this section we consider clustering of categorical data consisting of transactions, finite sets of elements (items) from a common item universe. Market basket data is typical transaction data. Co-occurrence of elements is the major idea in clustering transactional data. For example, one well-known similarity measure between two transactions is the Jaccard coefficient $\mathrm{sim}(T_1, T_2) = |T_1 \bigcap T_2| / |T_1 \bigcup T_2|$. Unfortunately, two random transactions rarely have elements in common.

We start with the algorithm CACTUS (Clustering Categorical Data Using Summaries), [30], that exemplifies the idea of co-occurrence. Values a, b of two different attributes are *strongly connected* if the number of transactions containing both a and b is larger than expected under an independence assumption by a user-defined margin α. This definition is extended to subsets A, B of two different attributes (each value pair $a \in A, b \in B$ has to be strongly connected), to segments (each 2D projection is strongly connected), and to the similarity of a pair of values of a single attribute via connectivity to other attributes. A cluster is defined as the maximal strongly connected segment (a Cartesian product of attribute subsets) having at least α times more elements than expected from the segment under the attribute independence assumption. Only two scans of data are required by this fast and scalable algorithm.

Han et al. exploited another idea [38] especially relevant when the item universe is large: (1) cluster items in item subsets C_j (*association rules* and *hyper-graph* machineries are used) and then (2) assign transactions T to a subset with highest similarity $\left|T \bigcap C_j\right| / \left|C_j\right|$.

Finally, consider the elegant algorithm STIRR (Sieving Through Iterated Reinforcement) [34] that deals with co-occurrence phenomenon for d-dimensional transactional records, *tuples* (e. g. tables of car sales with d fields). With each value (e. g. *Honda*) of one of d fields (e. g. *Manufacturer*) we associate a node v. Nodes v_1 and v_2 are close when many tuples containing them have a lot of co-occurring values (e. g., the highest sale month *August*). To formalize this concept, *STIRR* introduces a functional analysis concept of a *dynamical system*: with each node it associates a weight w_v and a transformation $w' = \Phi(w)$ over defined weights. This transformation computes w'_v as a sum of some *recombination* of weights w_b corresponding to the other $d-1$ values b over all tuples containing v (the overall weight is then renormalized). A fixed point $w = \Phi(w)$ can be achieved in several iterations. In addition to the idea of co-occurrence encapsulated in a weight propagation process, *STIRR* uses ideas of *spectral clustering* to perform the actual partitioning using non-principal eigenvectors.

Co-clustering

To explain the idea, consider text clustering, where each document is represented by a bag of its term frequencies (e. g. TF-IDF). In our usual case-by-attribute representation such textual data will result in a very large and sparse rectangular matrix with as many rows as number of documents and as many columns as number of terms, which is typically very high. Meanwhile, some terms are very similar to each other in the sense that they are used in the same subsets of documents. It would be handy to group such words together. Not only would a grouping of terms reduce dimensionality, but it would also eliminate certain randomness in preferential usage of a term from the same group by a particular document. In reality, this process is no different than merging documents that have similar term distributions. This is why the described idea is known as *co-clustering*.

For now, let us assume that we are dealing with a large sparse rectangular numerical non-negative data X. One approach to co-clustering in text mining is based on minimal cuts partitioning for bipartite graphs using SVD [20]. Another popular approach is based on *distributional* or *informational* clustering. After a row normalization (nonnegative elements sum to one), a data point can be viewed as a probability distribution (e. g. term distribution within a document). The same is true about every column attribute. Two attributes (two columns in matrix X) with exactly the same probability distributions are identical for the purpose of data mining, and so, one can be deleted. Attributes having distributions with small Kullback-Leibler (KL) distance (a common statistical measure to compare probability distributions) can still be grouped together without much impact on the information contained in the data. In addition, a natural derived attribute, the mixed distribution (a normalized sum of two columns), is now available to represent the group. This process can be generalized to more than two attributes. The grouping simplifies the original matrix X into the compressed form $\bar{X}$. In fact, our observation about impact on information can be quantified: information reduction $R = I(X) - I(\bar{X})$, where $I(X)$ is the mutual information contained in X[17], and is exactly equal to a sum of KL distances between original attributes assembled into a group and total group distribution (similarly to the objective function $E(C)$ used before in Sect. "Partitioning Relocation Clustering".

Now it is obvious that the intuitive idea of grouping similar attributes corresponds exactly to a minimization of information loss resulting in data compression. Moreover, R is symmetric with respect to columns and rows. So both dimensions can be gradually agglomerated under control of a single measure.

The above idea serves as a framework to several *co-clustering* developments, such as the *SimplifyRelation* algorithm [8] and *Information-Theoretic Co-clustering method* [21]. A natural question arises – how should co-clustering be done when the data matrix cannot be viewed as a probability distribution, for example, gene expression data that contains positive as well as negative entries. Recently, the theory and practice of co-clustering has been significantly advanced by extension to generalized distance measures known as *Bregman divergences* [5]. This advancement follows by using the important notion of *Bregman Information* and associating a co-clustering with a matrix approximation.

Scalability

Scalability challenges to clustering are much more severe than to predictive mining, with respect to computing time, to memory, and to the clustering concept itself in case of high dimensionality. We only reflect on a few of the most influential ideas without any references to many others. The first idea is to *squash* data into some summaries (*sufficient statistics*) [24] that, in the case of clustering, was first implemented in the BIRCH algorithm (Balanced Iterative Reduction and Clustering using Hierarchies) [70]. Many other developments (e. g. [11,31]) followed in its footsteps.

BIRCH pre-processes data into small tight micro-clusters, called *Cluster Features* (CF) accumulating zero, first, and second moments of CF members. CF are viewed as leaves in a height-balanced tree that resides in memory. They can later be used to build clusters via the algorithm of choice, for example hierarchical or k-means. The emphasis is shifted to a pre-processing phase of building a high-quality CF tree that is controlled by some parameters (as *branching factor*). When a tree reaches the assigned memory size, it is rebuilt with somewhat less tight CF. Outliers are saved in an auxiliary file. The overall complexity of BIRCH is $O(N)$ and it only takes one or two scans of data (which is not assumed to reside in memory).

Many algorithms (e. g. CLARANS) use old-fashioned sampling without rigorous statistical reasoning. *Sampling* was advanced to a new level with introduction of Hoeffding or Chernoff bounds [54]. In a nutshell, independently of the distribution for a real-valued random variable Y, $0 \le Y \le R$, the average of its n independent observations lies within ϵ of the actual mean

$$\left| \bar{Y} - \frac{1}{n} \sum_{j=1:n} Y_j \right| \le \epsilon$$

with probability $1 - \delta$ as soon as $\epsilon = \sqrt{R^2 \ln(1/\delta)/2n}$.

For example, these bounds are used in the clustering algorithm CURE [36] mentioned previously.

Another challenge is related to a situation when the dimensionality d is high.

Clustering in high-dimensional spaces presents two difficulties:

(1) *irrelevant attributes* – while in predictive mining they are easily discarded, irrelevant attributes can render the clustering task hopeless; their inspection in high dimension is nothing but easy;
(2) *curse of dimensionality* – in high-dimensional space, distance to a furthest point and to a closest point becomes on average undistinguishable; this is purely geometric phenomenon making overall proximity clustering very suspicious.

Some algorithms adapt to high dimensionality better than others. For example, the algorithm CACTUS adjusts well because it defines a cluster only in terms of a cluster's 2D projections. However, as a rule of thumb, the performance of classic clustering algorithms can degrade gradually with dimension.

A proposed cure against *the curse of dimensionality* is *subspace clustering*. It is exemplified by the algorithm CLIQUE (Clustering In QUEst) [2]. CLIQUE works with high-dimensional numerical data and combines many ideas. It starts with the units (elementary rectangular cells) in low-dimensional subspaces. Only units whose densities exceed a threshold τ are retained (this is similar to grid-based clustering). A bottom-up approach of finding such units is applied, starting from units of dimension one. An inductive step to move from a dimension $q-1$ to a dimension q involves a self-joint over common $q-2$ faces. This is similar to the Apriori-reasoning in association rules. All the subspaces are sorted by their coverage and lesser-covered subspaces are pruned. A cut point between retained and pruned subspaces is selected based on the Minimum Description Length (MDL) principle. A cluster is defined as a maximal set of connected dense units. Effectively, CLIQUE selects several subspaces, each representing different perspective. The result is a series of cluster systems in different subspaces. Unfortunately, the result is not a partition, since the systems overlap.

There have been many follow-up improvements to subspace clustering. On a theoretical side, a more transparent criterion for subspace selection based on the concept of *entropy* was suggested in [16]. A more advanced inductive unit generation that, in addition, utilizes adaptive grids is used in the algorithm MAFIA (Merging of Adaptive Finite Intervals) [35]. MAFIA starts with one data pass to construct *adaptive grids* in each dimension. Many (1000) bins are used to compute histograms by reading blocks of data into memory. The bins are then merged to come up with a smaller number of adaptive variable-size bins than CLIQUE. The algorithm uses a parameter α, called the cluster dominance factor, to select bins that are α-times more densely populated than average. These variable-size bins are $q=1$ candidate dense units (CDUs). Then MAFIA proceeds recursively to higher dimensions (every time a data scan is involved). Unlike CLIQUE, when constructing a new q-CDU, MAFIA tries two $(q-1)$-CDUs as soon as they share any (not only the first dimensions) $(q-2)$-face. This creates an order of magnitude more candidates. The algorithm PROCLUS (PROjected CLUstering) [1] explores pairs consisting of a data subset $C \subset X$ and a subspace in an attribute space A. A subset-subspace pair is called a projected cluster, if a projection of C onto the corresponding subspace is a tight cluster. Unlike with CLIQUE, projected clusters do not overlap.

Other Aspects of Clustering

Due to limited space, we have only touched on major clustering approaches. Now without any details we want to present a reader with a brief account of a few other clustering techniques.

The first important direction of thought is assessing clustering results. Different measures (e. g. *Silhouette* and *Partition* coefficients) are available, but not very useful. Usually assessment is done based purely on the objectives of the application. Another important research thread tries to answer the question of what is the optimal number k of clusters to build. Different statistical criteria have been developed such as BIC, AIC, MDL, and others. From the system complexity angle, clustering provides simplified data description. Given k cluster "prototypes" (e. g. centroids), data can be very roughly described by assigning each data point its cluster ID. Such very economical date compression is lossy, but in principle residuals between data points and their cluster centroids are tightly distributed and so takes relatively small amount of information to transmit. The more clusters k we introduce, the smaller the residuals will be, in the extreme each point coinciding with cluster centroid. However, more and more cluster centroids are needed to be transmitted upfront. So, the optimal k may be found from such considerations.

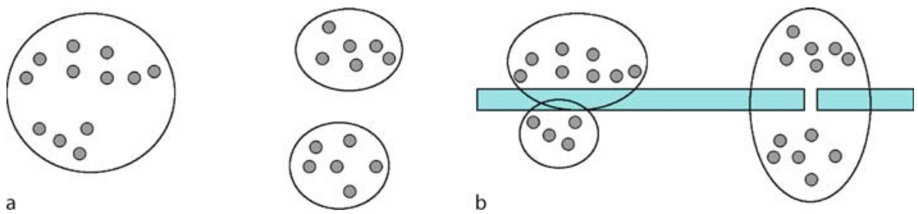
In real-world applications customers are rarely interested in unconstrained solutions. Clusters are frequently subjected to some problem-specific limitations that make them suitable for particular business actions. Finding clusters satisfying certain limitations is the subject of active research [40].

Consider, for example, clustering algorithm COD (Clustering with Obstructed Distance) [63]. The problem is easily illustrated by Fig. 4, where we show the difference in constructing three clusters in the absence of any obstacle (left) and in the presence of a river with a bridge (right).

Soft cluster assignments (assigning a point to several clusters with some probabilities) make sense if the k-means objective function is modified to incorporate "fuzzy errors" (similar to EM)

$$E'(C) = \sum_{i=1:N} \sum_{j=1:k} \left\| x_i - c_j \right\|^2 \omega_{ij}^2$$

The probabilities ω are defined based on Gaussian models. This makes the objective function differentiable with respect to means and allows application of general *gradient decent* methods presented in [53] in a context of



Knowledge Discovery: Clustering, Figure 4
Clustering with Obstructed Distance (COD) clustering. a No obstacles b River with the bridge

vector quantization. Another soft-assignment leads to so-called *harmonic means* [69]. Another iterative approach is used in *Self-Organizing Map* (SOM) [50].

Future Directions

Clustering is an important and well-studied problem. As discussed earlier, many algorithms exist that are applicable in different scenario. However, the theory of clustering is far from satisfactory. Even a simple and often used algorithm such as k-means does not have a performance guarantee. There have been some efforts to modify k-means, but none of these modified algorithms are simultaneously efficient and provably optimal. Thus a very important future direction of research is to develop a sound theory of clustering, that includes algorithms that have satisfactory performance guarantees. There have been a few efforts to axiomatize clustering, such as [49], however more such efforts are needed.

For the practitioner too, the variety of clustering algorithms is bewildering. An immensely beneficial tool will be an expert system for clustering. That is, a system that queries the user about various aspects of the data, the target application and the target clustering and then suggests appropriate clustering formulations and algorithms. Ideally, such a system would also provide a justification (relevant features) for the suggested algorithm. Besides an expert system, high quality public domain software would be invaluable.

Bibliography

1. Aggarwal CC, Wolf JL, Yu PS, Procopiuc C, Park JS (1999) Fast algorithms for projected clustering. In: Proceedings ACM SIGMOD International Conference on Management of Data, Philadelphia, pp 61–72
2. Agrawal R, Gehrke J, Gunopulos D, Raghavan P (1998) Automatic subspace clustering of high dimensional data for data mining applications. In: Proceedings of ACM SIGMOD, Seattle, pp 94–105
3. Ankerst M, Breunig MM, Kriegel HP, Sander J (1999) Optics: Ordering points to identify the clustering structure. In: Proceedings ACM SIGMOD International Conference on Management of Data, Philadelphia, pp 49–60
4. Banerjee A, Merugu S, Dhillon I, Ghosh J (2005) Clustering with Bregman divergences. J Mach Learn Res (JMLR) 6:1705–1749
5. Banerjee A, Dhillon I, Ghosh J, Merugu S, Modha D (2007) A generalized maximum entropy approach to Bregman co-clustering and matrix approximation. J Mach Learn Res(JMLR) 8:1919–1986
6. Barbara D, Chen P (2000) Using the fractal dimension to cluster datasets. In: Proceedings of the 6th ACM SIGKDD, Boston, pp 260–264
7. Berkhin P (2005) A survey of clustering data mining techniques. In: Kogan J, Nicholas C, Teboulle M (eds) Grouping Multidimensional Data. Springer
8. Berkhin P, Becher JD (2002) Learning simple relations: Theory and applications. In: Proceeding Second SIAM International Conference on Data Mining, Arlington, pp 420–436
9. Boley DL (1998) Principal direction divisive partitioning. Data Min Knowl Discov 2(4):325–344
10. Bradley P, Fayyad U (1998) Refining initial points for k-means clustering. In: Shavlik J (ed) Proceedings of the Fifteenth International Conference on Machine Learning (ICML). AAAI Press, San Francisco, pp 91–99
11. Bradley P, Fayyad U, Reina C (1998) Scaling clustering algorithms to large databases. In: Proceedings Fourth International Conference on Knowledge Discovery and Data Mining. AAAI Press, San Francisco
12. Cadez I, Gaffney S, Smyth P (2000) A general probabilistic framework for clustering individuals. Tech Rep UCI-ICS 00-09. University of California, Irvine
13. Cadez I, Smyth P, Mannila H (2001) Probabilistic modeling of transactional data with applications to profiling, visualization, and prediction. In: Proceedings of the 7th ACM SIGKDD, San Francisco, pp 37–46
14. Chan P, Schlag M, Zien J (1994) Spectral k-way ratio cut partitioning. IEEE Trans CAD-Integr Circuits Syst 13:1088–1096
15. Cheeseman P, Stutz J (1996) Bayesian classification (autoclass): Theory and results. In: Fayyad UM, Piatetsky-Shapiro G, Smyth P, Uthurusamy R (eds) Advances in Knowledge Discovery and Data Mining. AAAI/MIT Press, pp 153–180
16. Cheng C, Fu A, Zhang Y (1999) Entropy-based subspace clustering for mining numerical data. In: Proceedings of the 5th ACM SIGKDD, San Diego, pp 84–93
17. Cover TM, Thomas JA (1991) Elements of Information Theory. Wiley, New York
18. Day W, Edelsbrunner H (1984) Efficient algorithms for agglomerative hierarchical clustering methods. J Classif 1(7):7–24

19. Defays D (1977) An efficient algorithm for a complete link method. Comp J 20:364–366
20. Dhillon IS (2001) Co-clustering documents and words using bipartite spectral graph partitioning. In: Proceedings of the The Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD-2001). Also appears as: UT CS Technical Report # TR 2001-05, March, 1999
21. Dhillon IS, Mallela S, Kumar R (2003) A divisive information-theoretic feature clustering algorithm for text classification. J Mach Learn Res (JMLR): Var Feature Sel Special Issue 3:1265–1287
22. Dhillon IS, Guan Y, Kulis B (2007) Weighted graph cuts without eigenvectors: A multilevel approach. IEEE Trans Pattern Anal Mach Intell 29(11):1944–1957
23. Donath WE, Hoffman AJ (1973) Lower bounds for the partitioning of graphs. IBM J Res Dev 17:422–425
24. DuMouchel W, Volinsky C, Johnson T, Cortes C, Pregibon D (1999) Squashing flat files flatter. In: Proceedings of the 5th ACM SIGKDD, San Diego, pp 6–15
25. Dubes R (1993) Cluster analysis and related issues. In: Chen C, Pau L, Wang P (eds) Handbook of Pattern Recognition and Computer Vision. World Scientific, River Edge, pp 3–32
26. Ester M, Kriegel HP, Xu X (1995) A database interface for clustering in large spatial databases. In: Proceedings of the 1st ACM SIGKDD, Montreal, pp 94–99
27. Ester M, Kriegel HP, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the 2nd ACM SIGKDD, Portland, pp 226–231
28. Everitt B (1993) Cluster Analysis, 3rd edn. Edward Arnold, London
29. Fiedler M (1975) A property of eigenvectors of non-negative symmetric matrices and its application to graph theory. Czech Math J 25:619–633
30. Ganti V, Gehrke J, Ramakrishnan R (1999) CACTUS-clustering categorical data using summaries. In: Proceedings of the 5th ACM SIGKDD, San Diego, pp 73–83
31. Ganti V, Ramakrishnan R, Gehrke J, Powell A, French J (1999) Clustering large datasets in arbitrary metric spaces. In: Proceedings of the 15th ICDE, Sydney, pp 502–511
32. Gersho A, Gray RM (1992) Vector quantization and signal compression. Kluwer, Dordrecht
33. Ghosh J (2003) Scalable clustering methods for data mining. In: Ye N (ed) Handbook of Data Mining. Lawrence Ealbaum Assoc, pp 247–277
34. Gibson D, Kleinberg J, Raghavan P (1998) Clustering categorical data: An approach based on dynamic systems. In: Proceedings of the 24th International Conference on Very Large Databases, New York, pp 311–323
35. Goil S, Nagesh H, Choudhary A (1999) MAFIA: Efficient and scalable subspace clustering for very large data sets. Tech Rep CPDC-TR-9906-010. Northwestern University, Evanston
36. Guha S, Rastogi R, Shim K (1998) CURE: A clustering algorithm for large databases. ACM SIGMOD International Conference on Management of Data, pp 73–84
37. Hall KM (1970) A r-dimensional quadratic placement algorithm. Manag Sci 11(3):219–229
38. Han EH, Karypis G, Kumar V, Mobasher B (1997) Clustering based on association rule hypergraphs. In: SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery, (SIGMOD-DMKD'97)
39. Han J, Kamber M (2001) Data Mining: Concepts and Techniques. Morgan Kaufmann, San Mateo
40. Han J, Kamber M, Tung AKH (2001) Spatial clustering methods in data mining: A survey. In: Miller H, Han J (eds) Geographic Data Mining and Knowledge Discovery. Taylor and Francis
41. Hendrickson B, Leland R (1993) A multilevel algorithm for partitioning graphs. Tech Rep SAND93-1301, Sandia National Laboratories, Albuquerque
42. Hinneburg A, Keim D (1998) An efficient approach to clustering large multimedia databases with noise. In: Proceedings of the 4th ACM SIGKDD, New York, pp 58–65
43. Jain A, Murty M, Flynn PJ (1999) Data clustering: A review. ACM Comp Surv 31(3):264–323
44. Jain AK, Dubes RC (1988) Algorithms for Clustering Data. Prentice-Hall, Englewood Cliffs
45. Karypis G, Han EH, Kumar V (1999) CHAMELEON: Hierarchical clustering using dynamic modeling. IEEE Comput 32(8):68–75
46. Karypis G, Kumar V (1999) A fast and high quality multilevel scheme for partitioning irregular graphs. SIAM J Sci Comput 20(1):359–392
47. Kaufman L, Rousseeuw P (1990) Finding Groups in Data: An Introduction to Cluster Analysis. Wiley, New York
48. Kernighan B, Lin S (1970) An efficient heuristic procedure for partitioning graphs. Bell Syst Tech J 49(2):291–307
49. Kleinberg J (2002) An impossibility theorem for clustering. In: Proc Neural Info. Processing Systems (NIPS 15)
50. Kohonen T (1990) The self-organizing map. Proc IEEE 9:1464–1479
51. Lance G, Williams W (1967) A general theory of classification sorting strategies. Comp J 9:373–386
52. Leighton T, Rao S (1999) Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. J ACM 46(6):787–832
53. Marroquin J, Girosi F (1993) Some extensions of the k-means algorithm for image segmentation and pattern classification. Tech Rep AI Memo, MIT, Cambridge, p 1390
54. Motwani R, Raghavan P (1995) Randomized Algorithms. Cambridge University Press, Cambridge
55. Ng A, Jordan M, Weiss Y (2001) On spectral clustering: Analysis and an algorithm. In: Proc Neural Info Proc Syst (NIPS), Vancouver
56. Ng R, Han J (1994) Efficient and effective clustering methods for spatial data mining. In: Proc of the 20th Int Conf on Very Large Data Bases (VLDB), Santiago, pp 144–155
57. Parlett B (1997) The Symmetric Eigenvalue Problem, 2nd edn. SIAM, Philadelphia
58. Schikuta E, Erhart M (1997) The BANG-clustering system: grid-based data analysis. In: Proceeding of Advances in Intelligent Data Analysis, Reasoning about Data, 2nd International Symposium, London, pp 513–524
59. Sheikholeslami G, Chatterjee S, Zhang A (1998) Wavecluster: A multi-resolution clustering approach for very large spatial databases. In: Proceedings of the 24th Conference on VLDB, New York, pp 428–439
60. Shi J, Malik J (2000) Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell 22(8):888–905
61. Sibson R (1973) SLINK: An optimally efficient algorithm for the single link cluster method. Comput J 16:30–34
62. Spath H (1980) Cluster Analysis Algorithms. Ellis Horwood, Chichester

63. Tung A, Hou J, Han J (2001) Spatial clustering in the presence of obstacles. In: Proceedings of the 17th ICDE, Heidelberg, pp 359–367
64. Voorhees EM (1986) Implementing agglomerative hierarchical clustering algorithms for use in document retrieval. Inf Proc Manag 22(6):465–476
65. Wang W, Yang J, Muntz R (1997) STING: a statistical information grid approach to spatial data mining. In: Proceedings of the 23rd Conference on VLDB, Athens, pp 186–195
66. Xu X, Ester M, Kriegel HP, Sander J (1998) A distribution-based clustering algorithm for mining large spatial datasets. In: Proceedings of the 14th ICDE, Orlando, pp 324–331
67. Yao A (1982) On constructing minimum spanning trees in k-dimensional space and related problems. SIAM J Comput 11(4):721–736
68. Yu SX, Shi J (2003) Multiclass spectral clustering. In: International Conference on Computer Vision, Nice
69. Zhang B (2001) Generalized k-harmonic means – dynamic weighting of data in unsupervised learning. In: Proceedings of the 1st SIAM International Conference on Data Mining, Chicago
70. Zhang T, Ramakrishnan R, Livny M (1997) BIRCH: A new data clustering algorithm and its applications. J Data Mining Knowl Discov 1(2):141–182

Kolmogorov–Arnold–Moser (KAM) Theory

LUIGI CHIERCHIA
Dipartimento di Matematica, Università "Roma Tre", Roma, Italy

Article Outline

Glossary
Definition of the Subject
Introduction
Kolmogorov Theorem
Arnold's Scheme
The Differentiable Case: Moser's Theorem
Future Directions
A The Classical Implicit Function Theorem
B Complementary Notes
Bibliography

Glossary

Action-angle variables A particular set of variables $(y, x) = ((y_1, \dots, y_d), (x_1, \dots, x_d))$, x_i ("angles") defined modulus 2π, particularly suited to describe the general behavior of an integrable system.

Fast convergent (Newton) method Super-exponential algorithms, mimicking Newton's method of tangents, used to solve differential problems involving small divisors.

Hamiltonian dynamics The dynamics generated by a Hamiltonian system on a symplectic manifold, i. e., on an even-dimensional manifold endowed with a symplectic structure.

Hamiltonian system A time reversible, conservative (without dissipation or expansion) dynamical system, which generalizes classical mechanical systems (solutions of Newton's equation $m_i \ddot{x}_i = f_i(x)$, with $1 \le i \le d$ and $f = (f_1, \dots, f_d)$ a conservative force field); they are described by the flow of differential equations (i. e., the time t map associating to an initial condition, the solution of the initial value problem at time t) on a symplectic manifold and, locally, look like the flow associated with the system of differential equation $\dot{p} = -H_q(p, q)$, $\dot{q} = H_p(p, q)$ where $p = (p_1, \dots, p_d)$, $q = (q_1, \dots, q_d)$.

Integrable Hamiltonian systems A very special class of Hamiltonian systems, whose orbits are described by linear flows on the standard d-torus: $(y, x) \to (y, x + \omega t)$ where (y, x) are action-angle variables and t is time; the ω_i's are called the "frequencies" of the orbit.

Invariant tori Manifolds diffeomorphic to tori invariant for the flow of a differential equation (especially of Hamiltonian differential equations); establishing the existence of tori invariant for Hamiltonian flows is the main object of KAM theory.

KAM Acronym from the names of Kolmogorov (Andrey Nikolaevich Kolmogorov, 1903–1987), Arnold (Vladimir Igorevich Arnold, 1937) and Moser (Jürgen K. Moser, 1928–1999), whose results, in the 1950's and 1960's, in Hamiltonian dynamics, gave rise to the theory presented in this article.

Nearly-integrable Hamiltonian systems Hamiltonian systems which are small perturbations of an integrable system and which, in general, exhibit a much richer dynamics than the integrable limit. Nevertheless, KAM theory asserts that, under suitable assumptions, the majority (in the measurable sense) of the initial data of a nearly-integrable system behaves as in the integrable limit.

Quasi-periodic motions Trajectories (solutions of a system of differential equations), which are conjugate to linear flow on tori.

Small divisors/denominators Arbitrary small combinations of the form $\omega \cdot k := \sum_{j=1}^{d} \omega_i k_i$ with $\omega = (\omega_1, \dots, \omega_d) \in \mathbb{R}^d$ a real vector and $k \in \mathbb{Z}^d$ an integer vector different from zero; these combinations arise in the denominators of certain expansions appearing in the perturbation theory of Hamiltonian systems, making (when $d > 1$) convergent arguments very delicate. Physically, small divisors are related to "resonances", which are a typical feature of conservative systems.

Stability The property of orbits of having certain properties similar to a reference limit; more specifically, in the context of KAM theory, stability is normally referred to as the property of action variables of staying close to their initial values.

Symplectic structure A mathematical structure (a differentiable, non-degenerate, closed 2-form) apt to describe, in an abstract setting, the main geometrical features of conservative differential equations arising in mechanics.

Definition of the Subject

KAM theory is a mathematical, quantitative theory which has as its primary object the persistence, under small (Hamiltonian) perturbations, of typical trajectories of integrable Hamiltonian systems. In integrable systems with bounded motions, the typical trajectory is quasi-periodic, i. e., may be described through the linear flow $x \in \mathbb{T}^d \to x + \omega t \in \mathbb{T}^d$ where $\mathbb{T}^d$ denotes the standard d-dimensional torus (see Sect. "Introduction" below), t is time, and $\omega = (\omega_1, \dots, \omega_d) \in \mathbb{R}^d$ is the set of frequencies of the trajectory (if $d = 1$, $2\pi/\omega$ is the *period* of the motion).

The main motivation for KAM theory is related to stability questions arising in celestial mechanics which were addressed by astronomers and mathematicians such as Kepler, Newton, Lagrange, Liouville, Delaunay, Weierstrass, and, from a more modern point of view, Poincaré, Birkhoff, Siegel, …

The major breakthrough in this context, was due to Kolmogorov in 1954, followed by the fundamental work of Arnold and Moser in the early 1960s, who were able to overcome the formidable technical problem related to the appearance, in perturbative formulae, of arbitrarily small divisors[1]. Small divisors make the use of classical analytical tools (such as the standard Implicit Function Theorem, fixed point theorems, etc.) impossible and could be controlled only through a "fast convergent method" of Newton-type[2], which allowed, in view of the super-exponential rate of convergence, counterbalancing the divergences introduced by small divisors.

Actually, the main bulk of KAM theory is a set of *techniques* based, as mentioned, on fast convergent methods, and solving various questions in Hamiltonian (or generalizations of Hamiltonian) dynamics. By now, there are excellent reviews of KAM theory – especially Sect. 6.3 of [6] and [60] – which should complement the reading of this article, whose main objective is not to review but rather to explain the main fundamental ideas of KAM theory. To do this, we re-examine, in modern language, the main ideas introduced, respectively, by the founders of KAM theory, namely Kolmogorov (in Sect. "Kolmogorov Theorem"), Arnold (in Sect. "Arnold's Scheme") and Moser (Sect. "The Differentiable Case: Moser's Theorem").

In Sect. "Future Directions" we briefly and informally describe a few developments and applications of KAM theory: this section is by no means exhaustive and is meant to give a non technical, short introduction to some of the most important (in our opinion) extensions of the original contributions; for more detailed and complete reviews we recommend the above mentioned articles Sect. 6.3 of [6] and [60].

Appendix A contains a quantitative version of the classical Implicit Function Theorem.

A set of technical notes (such as notes 17, 18, 19, 21, 24, 26, 29, 30, 31, 34, 39), which the reader not particularly interested in technical mathematical arguments may skip, are collected in Appendix B and complete the mathematical expositions. Appendix B also includes several other complementary notes, which contain either standard material or further references or side comments.

Introduction

In this article we will be concerned with Hamiltonian flows on a symplectic manifold $(\mathcal{M}, dy \wedge dx)$; for general information, see, e. g., [5] or Sect. 1.3 of[6]. Notation, main definitions and a few important properties are listed in the following items.

(a) As symplectic manifold ("phase space") we shall consider $\mathcal{M} := B \times \mathbb{T}^d$ with $d \geq 2$ (the case $d = 1$ is trivial for the questions addressed in this article) where: B is an open, connected, bounded set in $\mathbb{R}^d$; $\mathbb{T}^d := \mathbb{R}^d/(2\pi\mathbb{Z}^d)$ is the standard flat d-dimensional torus with periods[3] 2π

(b) $dy \wedge dx := \sum_{i=1}^d dy_i \wedge dx_i$, $(y \in B,\ x \in \mathbb{T}^d)$ is the standard symplectic form[4]

(c) Given a real-analytic (or smooth) function $H\colon \mathcal{M} \to \mathbb{R}$, the *Hamiltonian flow governed by* H is the one-parameter family of diffeomorphisms $\phi_H^t\colon \mathcal{M} \to \mathcal{M}$, which to $z \in \mathcal{M}$ associates the solution at time t of the differential equation[5]

$$\dot z = J_{2d} \nabla H(z)\,, \quad z(0) = z\,, \tag{1}$$

where $\dot z = dz/dt$, J_{2d} is the standard symplectic $(2d \times 2d)$-matrix

$$J_{2d} = \begin{pmatrix} 0 & -\mathbb{1}_d \\ \mathbb{1}_d & 0 \end{pmatrix}\,,$$

$\mathbb{1}_d$ denotes the unit $(d \times d)$-matrix, 0 denotes a $(d \times d)$ block of zeros, and ∇ denotes gradient; in the symplectic coordinates $(y, x) \in B \times \mathbb{T}^d$, equations (1) reads

$$\begin{cases} \dot{y} = -H_x(y, x) \\ \dot{x} = H_y(y, x) \end{cases}, \qquad \begin{cases} y(0) = y \\ x(0) = x \end{cases} \tag{2}$$

Clearly, the flow ϕ_H^t is defined until $y(t)$ eventually reaches the border of B.

Equations (1) and (2) are called the *Hamilton's equations* with *Hamiltonian H*; usually, the symplectic (or "conjugate") variables (y, x) are called *action-angle* variables[6]; the number d (= half of the dimension of the phase space) is also referred to as "the number of degrees of freedom[7]".

The Hamiltonian H is constant over trajectories $\phi_H^t(z)$, as it follows immediately by differentiating $t \to H(\phi_H^t(z))$. The constant value $E = H(\phi_H^t(z))$ is called the energy of the trajectory $\phi_H^t(z)$.

Hamilton equations are left invariant by *symplectic* (or "canonical") change of variables, i. e., by diffeomorphisms on $\mathcal{M}$ which preserve the 2-form $\mathrm{d}y \wedge \mathrm{d}x$; i. e., if $\phi\colon (y, x) \in \mathcal{M} \to (\eta, \xi) = \phi(y, x) \in \mathcal{M}$ is a diffeomorphism such that $\mathrm{d}\eta \wedge \mathrm{d}\xi = \mathrm{d}y \wedge \mathrm{d}x$, then

$$\phi \circ \phi_H^t \circ \phi^{-1} = \phi_{H \circ \phi^{-1}}^t \, . \tag{3}$$

An equivalent condition for a map ϕ to be symplectic is that its Jacobian ϕ' is a *symplectic matrix*, i. e.,

$$\phi'^T J_{2d} \phi' = J_{2d} \tag{4}$$

where J_{2d} is the standard symplectic matrix introduced above and the superscript T denotes matrix transposition.

By a (generalization of a) theorem of Liouville, the Hamiltonian flow is symplectic, i. e., the map $(y, x) \to (\eta, \xi) = \phi_H^t(y, x)$ is symplectic for any H and any t; see Corollary 1.8, [6].

A classical way of producing symplectic transformations is by means of *generating functions*. For example, if $g(\eta, x)$ is a smooth function of $2d$ variables with

$$\det \frac{\partial^2 g}{\partial \eta \partial x} \neq 0 \, ,$$

then, by the Implicit Function Theorem (IFT; see [36] or Sect. "A The Classical Implicit Function Theorem" below), the map $\phi\colon (y, x) \to (\eta, \xi)$ defined implicitly by the relations

$$y = \frac{\partial g}{\partial x} \, , \qquad \xi = \frac{\partial g}{\partial \eta} \, ,$$

yields a local symplectic diffeomorphism; in such a case, g is called the generating function of the transformation ϕ. For example, the function $\eta \cdot x$ is the generating function of the identity map.

For general information about symplectic changes of coordinates, generating functions and, in general, about symplectic structures we refer the reader to [5] or [6].

(d) A solution $z(t) = (y(t), x(t))$ of (2) is a *maximal quasi-periodic solution* with frequency vector $\omega = (\omega_1, \ldots, \omega_d) \in \mathbb{R}^d$ if ω is a rationally-independent vector, i. e.,

$$\exists n \in \mathbb{Z}^d \ \text{ s.t. } \ \omega \cdot n := \sum_{i=1}^{d} \omega_i n_i = 0 \implies n = 0 \, , \tag{5}$$

and if there exist smooth (periodic) functions $v, u\colon \mathbb{T}^d \to \mathbb{R}^d$ such that[8]

$$\begin{cases} y(t) = v(\omega t) \\ x(t) = \omega t + u(\omega t) \, . \end{cases} \tag{6}$$

(e) Let ω, u and v be as in the preceding item and let U and ϕ denote, respectively, the maps

$$\begin{cases} U\colon \theta \in \mathbb{T}^d \to U(\theta) := \theta + u(\theta) \in \mathbb{T}^d \\ \phi\colon \theta \in \mathbb{T}^d \to \phi(\theta) := (v(\theta), U(\theta)) \in \mathcal{M} \end{cases}$$

If U is a smooth diffeomorphism of $\mathbb{T}^d$ (so that, in particular[9] $\det U_\theta \neq 0$) then ϕ is an embedding of $\mathbb{T}^d$ into $\mathcal{M}$ and the set

$$\mathcal{T}_\omega = \mathcal{T}_\omega^d := \phi(\mathbb{T}^d) \tag{7}$$

is an embedded d-torus invariant for ϕ_H^t and on which the motion is conjugated to the linear (Kronecker) flow $\theta \to \theta + \omega t$, i. e.,

$$\phi^{-1} \circ \phi_H^t \circ \phi(\theta) = \theta + \omega t \, , \quad \forall \theta \in \mathbb{T}^d \, . \tag{8}$$

Furthermore, the invariant torus $\mathcal{T}_\omega$ is a graph over $\mathbb{T}^d$ and is *Lagrangian*, i. e., ($\mathcal{T}_\omega$ has dimension d and) the restriction of the symplectic form $\mathrm{d}y \wedge \mathrm{d}x$ on $\mathcal{T}_\omega$ vanishes[10].

(f) In KAM theory a major role is played by the numerical properties of the frequencies ω. A typical assumption is that ω is a (homogeneously) *Diophantine vector*: $\omega \in \mathbb{R}^d$ is called Diophantine or (κ, τ)-Diophantine if, for some constants $0 < \kappa \leq \min_i |\omega_i|$ and

$\tau \geq d-1$, it verifies the following inequalities:

$$|\omega \cdot n| \geq \frac{\kappa}{|n|^{\tau}}\,, \quad \forall n \in \mathbb{Z}^d \setminus \{0\}\,, \tag{9}$$

(normally, for integer vectors n, $|n|$ denotes $|n_1|+\cdots+|n_d|$, but other norms may well be used). We shall refer to κ and τ as the Diophantine constants of ω. The set of Diophantine numbers in $\mathbb{R}^d$ with constants κ and τ will be denoted by $\mathcal{D}^d_{\kappa,\tau}$; the union over all $\kappa > 0$ of $\mathcal{D}^d_{\kappa,\tau}$ will be denoted by $\mathcal{D}^d_{\tau}$ and the union over all $\tau \geq d-1$ of $\mathcal{D}^d_{\tau}$ will be denoted by $\mathcal{D}^d$. Basic facts about these sets are[11]: if $\tau < d-1$ then $\mathcal{D}^d_\tau = \emptyset$; if $\tau > d-1$ then the Lebesgue measure of $\mathbb{R}^d \setminus \mathcal{D}^d_\tau$ is zero; if $\tau = d-1$, the Lebesgue measure of $\mathcal{D}^d_\tau$ is zero but its intersection with any open set has the cardinality of $\mathbb{R}$.

(g) The tori $\mathcal{T}_\omega$ defined in (e) with $\omega \in \mathcal{D}^d$ will be called *maximal KAM tori* for H.

(h) A Hamiltonian function $(\eta, \xi) \in \mathcal{M} \to H(\eta, \xi)$ having a maximal KAM torus (or, more generally, a maximal invariant torus as in (e) with ω rationally independent) $\mathcal{T}_\omega$, can be put into the form[12]

$$\begin{aligned} &K(y,x) := E + \omega \cdot y + Q(y,x)\,, \\ &\text{with } \partial_y^\alpha Q(0,x) = 0, \ \forall \alpha \in \mathbb{N}^d, \ |\alpha| \leq 1\,; \end{aligned} \tag{10}$$

compare, e. g., Sect. 1 of [59]; in the variables (y, x), the torus $\mathcal{T}_\omega$ is simply given by $\{y = 0\} \times \mathbb{T}^d$ and E is its (constant) energy. A Hamiltonian in the form (10) is said to be in *Kolmogorov normal form.*

If

$$\det\langle \partial_y^2 Q(0,\cdot)\rangle \neq 0\,, \tag{11}$$

(where the brackets denote an average over $\mathbb{T}^d$ and ∂_y^2 the Hessian with respect to the y-variables) we shall say that the Kolmogorov normal form K in (10) is *non-degenerate*; similarly, we shall say that the KAM torus $\mathcal{T}_\omega$ for H is non-degenerate if H can be put in a non-degenerate Kolmogorov normal form.

Remark 1

(i) A classical theorem by H. Weyl says that the flow

$$\theta \in \mathbb{T}^d \to \theta + \omega t \in \mathbb{T}^d\,, \quad t \in \mathbb{R}$$

is dense (ergodic) in $\mathbb{T}^d$ if and only if $\omega \in \mathbb{R}^d$ is rationally independent (compare [6], Theorem 5.4 or Sect. 1.4 of [33]). Thus, trajectories on KAM tori fill them densely (i. e., pass in any neighborhood of any point).

(ii) In view of the preceding remark, it is easy to see that if ω is rationally independent, $(y(t), x(t))$ in (6) is a solution of (2) if and only if the functions v and u satisfy the following quasi-linear system of PDEs on $\mathbb{T}^d$:

$$\begin{cases} D_\omega v = -H_x(v(\theta), \theta + u(\theta)) \\ \omega + D_\omega u = H_y(v(\theta), \theta + u(\theta)) \end{cases} \tag{12}$$

where D_ω denotes the directional derivative $\omega \cdot \partial_\theta = \sum_{i=1}^d \omega_i \frac{\partial}{\partial \theta_i}$.

(iii) Probably, the main motivation for studying quasi-periodic solutions of Hamiltonian systems on $\mathbb{R}^d \times \mathbb{T}^d$ comes from perturbation theory of *nearly-integrable* Hamiltonian systems: a completely integrable system may be described by a Hamiltonian system on $\mathcal{M} := B(y_0, r) \times \mathbb{T}^d \subset \mathbb{R}^d \times \mathbb{T}^d$ with Hamiltonian $H = K(y)$ (compare Theorem 5.8, [6]); here $B(y_0, r)$ denotes the open ball $\{y \in \mathbb{R}^d : |y - y_0| < r\}$ centered at $y_0 \in \mathbb{R}^d$. In such a case the Hamiltonian flow is simply

$$\begin{aligned} &\phi_K^t(y,x) = \big(y, x + \omega(y) t\big)\,, \\ &\qquad \omega(y) := K_y(y) := \frac{\partial K}{\partial y}(y)\,. \end{aligned} \tag{13}$$

Thus, if the "frequency map" $y \in B \to \omega(y)$ is a diffeomorphism (which is guaranteed if $\det K_{yy}(y_0) \neq 0$, for some $y_0 \in B$ and B is small enough), in view of (f), for almost all initial data, the trajectories (13) belong to maximal KAM tori $\{y\} \times \mathbb{T}^d$ with $\omega(y) \in \mathcal{D}^d$.

The main content of (classical) KAM theory, in our language, is that, *if the frequency map $\omega = K_y$ of a (real-analytic) integrable Hamiltonian $K(y)$ is a diffeomorphism, KAM tori persist under small (smooth enough) perturbations of K*; compare Remark 7–(iv) below.

The study of the dynamics generated by the flow of a one-parameter family of Hamiltonians of the form

$$K(y) + \varepsilon P(y, x; \varepsilon)\,, \quad 0 < \varepsilon \ll 1\,, \tag{14}$$

was called by H. Poincaré *le problème général de la dynamique*, to which he dedicated a large part of his monumental *Méthodes Nouvelles de la Mécanique Céleste* [49].

(iv) A big chapter in KAM theory, strongly motivated by applications to PDEs with Hamiltonian structure (such as nonlinear wave equation, Schrödinger equation, KdV, etc.), is concerned with quasi-periodic solutions with $1 \le n < d$ frequencies, i. e., solutions of (2) of the form

$$\begin{cases} y(t) = v(\omega t) \\ x(t) = U(\omega t)\,, \end{cases} \tag{15}$$

where $v\colon \mathbb{T}^n \to \mathbb{R}^d$, $U\colon \mathbb{T}^n \to \mathbb{T}^d$ are smooth functions, $\omega \in \mathbb{R}^n$ is a rationally independent n-vector. Also in this case, if the map U is a diffeomorphism onto its image, the set

$$\mathcal{T}^n_\omega := \{(y, x) \in \mathcal{M}\colon \\ y = v(\theta)\,,\ x = U(\theta)\,,\ \theta \in \mathbb{T}^n\} \tag{16}$$

defines an invariant n-torus on which the flow ϕ^t_H acts by the linear translation $\theta \to \theta + \omega t$. Such tori are normally referred to as *lower dimensional tori*.

Even though this article will be mainly focused on "classical KAM theory" and on maximal KAM tori, we will briefly discuss lower dimensional tori in Sect. "Future Directions".

Kolmogorov Theorem

In the 1954 International Congress of Mathematicians, in Amsterdam, A.N. Kolmogorov announced the following fundamental (for the terminology, see (f), (g) and (h) above).

Theorem 1 (Kolmogorov [35]) *Consider a one-parameter family of real-analytic Hamiltonian functions on $\mathcal{M} := B(0, r) \times \mathbb{T}^d$ given by*

$$H := K + \varepsilon P \quad (\varepsilon \in \mathbb{R})\,, \tag{17}$$

where: (i) K is a non-degenerate Kolmogorov normal form (10)–(11); (ii) $\omega \in \mathcal{D}^d$ is Diophantine. Then, there exists $\varepsilon_0 > 0$ and for any $|\varepsilon| \le \varepsilon_0$ a real-analytic symplectic transformation $\phi_*\colon \mathcal{M}_* := B(0, r_*) \times \mathbb{T}^d \to \mathcal{M}$, for some $0 < r_* < r$, putting H in non-degenerate Kolmogorov normal form, $H \circ \phi_* = K_*$, with $K_* := E_* + \omega \cdot y' + Q_*(y', x')$. Furthermore[13], $\|\phi_* - \mathrm{id}\|_{C^1(\mathcal{M}_*)}$, $|E_* - E|$, and $\|Q_* - Q\|_{C^1(\mathcal{M}_*)}$ are small with ε.

Remark 2

(i) From Theorem 1 it follows that the torus

$$\mathcal{T}_{\omega,\varepsilon} := \phi_*(0, \mathbb{T}^d)$$

is a maximal non-degenerate KAM torus for H and the H-flow on $\mathcal{T}_{\omega,\varepsilon}$ is analytically conjugated (by ϕ_*) to the translation $x' \to x' + \omega t$ with the *same frequency* vector of $\mathcal{T}_{\omega,0} := \{0\} \times \mathbb{T}^d$, while the energy of $\mathcal{T}_{\omega,\varepsilon}$, namely E_*, is in general different from the energy E of $\mathcal{T}_{\omega,0}$. The idea of keeping the frequency fixed is a key idea introduced by Kolmogorov and its importance will be made clear in the analysis of the proof.

(ii) In fact, *the dependence upon ε is analytic* and therefore the torus $\mathcal{T}_{\omega,\varepsilon}$ is an analytic deformation of the unperturbed torus $\mathcal{T}_{\omega,0}$ (which is invariant for K); see Remark 7–(iii) below.

(iii) Actually, Kolmogorov not only stated the above result but gave also a precise outline of its proof, which is based on a *fast convergent "Newton" scheme*, as we shall see below; compare also [17].

The map ϕ_* is obtained as

$$\phi_* = \lim_{j\to\infty} \phi_1 \circ \cdots \circ \phi_j\,,$$

where the ϕ_j's are (ε-dependent) symplectic transformations of $\mathcal{M}$ closer and closer to the identity. It is enough to describe the construction of ϕ_1; ϕ_2 is then obtained by replacing $H_0 := H$ with $H_1 = H \circ \phi_1$ and so on.

We proceed to analyze the scheme of Kolmogorov's proof, which will be divided into three main steps.

Step 1: Kolmogorov Transformation

The map ϕ_1 is close to the identity and is generated by

$$g(y', x) := y' \cdot x + \varepsilon\left(b \cdot x + s(x) + y' \cdot a(x)\right)$$

where s and a are (respectively, scalar and vector-valued) real-analytic functions on $\mathbb{T}^d$ with zero average and $b \in \mathbb{R}^d$: setting

$$\begin{aligned} \beta_0 &= \beta_0(x) := b + s_x\,, \\ A &= A(x) := a_x \quad \text{and} \\ \beta &= \beta(y', x) := \beta_0 + Ay'\,, \end{aligned} \tag{18}$$

($s_x = \partial_x s = (s_{x_1}, \ldots, s_{x_d})$ and a_x denotes the matrix $(a_x)_{ij} := \frac{\partial a_i}{\partial x_j}$) ϕ_1 is implicitly defined by

$$\begin{cases} y = y' + \varepsilon\beta(y', x) := y' + \varepsilon(\beta_0(x) + A(x)y') \\ x' = x + \varepsilon a(x)\,. \end{cases} \tag{19}$$

Thus, for ε small, $x \in \mathbb{T}^d \to x + \varepsilon a(x) \in \mathbb{T}^d$ defines a diffeomorphism of $\mathbb{T}^d$ with inverse

$$x = \varphi(x') := x' + \varepsilon\alpha(x'; \varepsilon)\,, \tag{20}$$

for a suitable real-analytic function α, and ϕ_1 is explicitly given by

$$\phi_1 : (y', x') \to \begin{cases} y = y' + \varepsilon\beta\left(y', \varphi(x')\right) \\ x = \varphi(x') . \end{cases} \tag{21}$$

Remark 3

(i) Kolmogorov transformation ϕ_1 is actually the composition of two "elementary" symplectic transformations: $\phi_1 = \phi_1^{(1)} \circ \phi_1^{(2)}$ where $\phi_1^{(2)} : (y', x') \to (\eta, \xi)$ is the symplectic lift of the $\mathbb{T}^d$-diffeomorphism given by $x' = \xi + \varepsilon a(\xi)$ (i. e., $\phi_1^{(2)}$ is the symplectic map generated by $y' \cdot \xi + \varepsilon y' \cdot a(\xi)$), while $\phi_1^{(1)} : (\eta, \xi) \to (y, x)$ is the angle-dependent action translation generated by $\eta \cdot x + \varepsilon(b \cdot x + s(x))$; $\phi_1^{(2)}$ acts in the "angle direction" and will be needed to straighten out the flow up to order $O(\varepsilon^2)$, while $\phi_1^{(1)}$ acts in the "action direction" and will be needed to keep the frequency of the torus fixed.

(ii) The inverse of ϕ_1 has the form

$$(y, x) \to \begin{cases} y' = M(x)y + c(x) \\ x' = \phi(x) \end{cases} \tag{22}$$

with M a $(d \times d)$-invertible matrix and ϕ a diffeomorphism of $\mathbb{T}^d$ (in the present case $M = (\mathbb{1}_d + \varepsilon A(x))^{-1} = \mathbb{1}_d + O(\varepsilon)$ and $\phi = \mathrm{id} + \varepsilon a$) and it is easy to see that the symplectic diffeomorphisms of the form (22) form a subgroup of the symplectic diffeomorphisms, which we shall call *the group of Kolmogorov transformations*.

Determination of Kolmogorov transformation Following Kolmogorov, we now try to determine b, s and a so that the "new Hamiltonian" (better: "the Hamiltonian in the new symplectic variables") takes the form

$$H_1 := H \circ \phi_1 = K_1 + \varepsilon^2 P_1 , \tag{23}$$

with K_1 in the Kolmogorov normal form

$$K_1 = E_1 + \omega \cdot y' + Q_1(y', x') , \quad Q_1 = O(|y'|^2) . \tag{24}$$

To proceed we insert $y = y' + \varepsilon\beta(y', x)$ into H and, after some elementary algebra and using Taylor formula, we find[14]

$$\begin{aligned} H(y' + \varepsilon\beta, x) = E + \omega \cdot y' + Q(y', x) \\ + \varepsilon Q'(y', x) + \varepsilon F'(y', x) + \varepsilon^2 P'(y', x) \end{aligned} \tag{25}$$

where, defining

$$\begin{cases} Q^{(1)} := Q_y(y', x) \cdot (a_x y') \\ Q^{(2)} := [Q_y(y', x) - Q_{yy}(0, x)y'] \cdot \beta_0 \\ \qquad = \displaystyle\int_0^1 (1 - t) Q_{yyy}(ty', x) y' \cdot y' \cdot \beta_0 \mathrm{d}t \\ Q^{(3)} := P(y', x) - P(0, x) - P_y(0, x) y' \\ \qquad = \displaystyle\int_0^1 (1 - t) P_{yy}(ty', x) y' \cdot y' \mathrm{d}t \\ P^{(1)} := \dfrac{1}{\varepsilon^2} [Q(y' + \varepsilon\beta, x) - Q(y', x) \\ \qquad - \varepsilon Q_y(y', x) \cdot \beta] \\ \qquad = \displaystyle\int_0^1 (1 - t) Q_{yy}(y' + t\varepsilon\beta, x) \beta \cdot \beta \mathrm{d}t \\ P^{(2)} := \dfrac{1}{\varepsilon} [P(y' + \varepsilon\beta, x) - P(y', x)] \\ \qquad = \displaystyle\int_0^1 P_y(y' + t\varepsilon\beta, x) \cdot \beta \mathrm{d}t , \end{cases} \tag{26}$$

(recall that $Q_y(0, x) = 0$) and denoting the ω-directional derivative

$$D_\omega := \sum_{j=1}^{d} \omega_j \frac{\partial}{\partial x_j}$$

one sees that $Q' = Q'(y', x)$, $F' = F'(y', x)$ and $P' = P'(y', x)$ are given by, respectively

$$\begin{cases} Q'(y', x) := Q^{(1)} + Q^{(2)} + Q^{(3)} = O(|y'|^2) \\ F'(y', x) := \omega \cdot b + D_\omega s + P(0, x) \\ \qquad + \left\{D_\omega a + Q_{yy}(0, x)\beta_0 + P_y(0, x)\right\} \cdot y' \\ P' := P^{(1)} + P^{(2)} , \end{cases} \tag{27}$$

where $D_\omega a$ is the vector function with kth entry $\sum_{j=1}^d \omega_j \frac{\partial a_k}{\partial x_j}$; $D_\omega a \cdot y' = \omega \cdot (a_x y') = \sum_{j,k=1}^d \omega_j \frac{\partial a_k}{\partial x_j} y'_k$; recall, also, that $Q = O(|y|^2)$ so that $Q_y = O(y)$ and $Q' = O(|y'|^2)$.

Notice that, as an intermediate step, we are considering H as a function of mixed variables y' and x (and this causes no problem, as it will be clear along the proof).

Thus, recalling that x is related to x' by the (y'-independent) diffeomorphism $x = x' + \varepsilon\alpha(x'; \varepsilon)$ in (21), we see that in order to achieve relations (23)–(24), we have to determine b, s and a so that

$$F'(y', x) = \text{const} . \tag{28}$$

Remark 4

(i) F' is a first degree polynomial in y' so that (28) is equivalent to

$$\begin{cases} \omega \cdot b + D_\omega s + P(0,x) = \text{const} , \\ D_\omega a + Q_{yy}(0,x)\beta_0 + P_y(0,x) = 0 . \end{cases} \tag{29}$$

Indeed, the second equation is necessary to keep the torus frequency fixed and equal to ω (which, as we shall see in more detail later, is a key ingredient introduced by Kolmogorov).

(ii) In solving (28) or (29), we shall encounter differential equations of the form

$$D_\omega u = f , \tag{30}$$

for some given function f real-analytic on $\mathbb{T}^d$. Taking the average over $\mathbb{T}^d$ shows that $\langle f \rangle = 0$, and we see that (30) can be solved only if f has vanishing mean value

$$\langle f \rangle = f_0 = 0 ;$$

in such a case, expanding in Fourier series[15], one sees that (30) is equivalent to

$$\sum_{\substack{n\in\mathbb{Z}^d \\ n\neq 0}} i\omega \cdot n u_n e^{in\cdot x} = \sum_{\substack{n\in\mathbb{Z}^d \\ n\neq 0}} f_n e^{in\cdot x} , \tag{31}$$

so that the solutions of (30) are given by

$$u = u_0 + \sum_{\substack{n\in\mathbb{Z}^d \\ n\neq 0}} \frac{f_n}{i\omega \cdot n} e^{in\cdot x} , \tag{32}$$

for an arbitrary u_0. Recall that for a continuous functionf over $\mathbb{T}^d$ to be analytic it is necessary and sufficient that its Fourier coefficients f_n decay exponentially fast in n, i. e., that there exist positive constants M and ξ such that

$$|f_n| \le M e^{-\xi|n|} , \quad \forall n . \tag{33}$$

Now, since $\omega \in \mathcal{D}^d_{\kappa,\tau}$ one has that (for $n \neq 0$)

$$\frac{1}{|\omega \cdot n|} \le \frac{|n|^\tau}{\kappa} \tag{34}$$

and one sees that if f is analytic so is u in (32) (although the decay constants of u will be different from those of f; see below)

Summarizing, *if f is real-analytic on $\mathbb{T}^d$ and has vanishing mean value f_0, then there exists a unique real-analytic solution of (30) with vanishing mean value, which is given by*

$$D_\omega^{-1} f := \sum_{\substack{n\in\mathbb{Z}^d \\ n\neq 0}} \frac{f_n}{i\omega \cdot n} e^{in\cdot x} ; \tag{35}$$

all other solutions of (30) are obtained by adding an arbitrary constant to $D_\omega^{-1} f$ as in (32) with u_0 arbitrary.

Taking the average of the first relation in (29), we may the determine the value of the constant denoted const, namely,

$$\text{const} = \omega \cdot b + P_0(0) := \omega \cdot b + \langle P(0,\cdot) \rangle . \tag{36}$$

Thus, by (ii) of Remark 4, we see see that

$$s = -D_\omega^{-1}\left(P(0,x) - P_0(0)\right) = -\sum_{\substack{n\in\mathbb{Z}^d \\ n\neq 0}} \frac{P_n(0)}{i\omega \cdot n} e^{in\cdot x} , \tag{37}$$

where $P_n(0)$ denote the Fourier coefficients of $x \to P(0,x)$; indeed s is determined only up to a constant by the relation in (29) but *we select the zero-average solution*. Thus, s has been completely determined.

To solve the second (vector) equation in (29) we first have to require that the left hand side (l.h.s.) has vanishing mean value, i. e., recalling that $\beta_0 = b + s_x$ (see (18)), we must have

$$\langle Q_{yy}(0,\cdot)\rangle b + \langle Q_{yy}(0,\cdot) s_x \rangle + \langle P_y(0,\cdot)\rangle = 0 . \tag{38}$$

In view of (11) this relation is equivalent to

$$b = -\langle Q_{yy}(0,\cdot)\rangle^{-1}\left(\langle Q_{yy}(0,\cdot) s_x\rangle + \langle P_y(0,\cdot)\rangle\right) , \tag{39}$$

which uniquely determines b. Thus β_0 is completely determined and the l.h.s. of the second equation in (29) has average zero; thus its unique zero-average solution (again zero-average of a is required as a normalization condition) is given by

$$a = -D_\omega^{-1}\left(Q_{yy}(0,x)\beta_0 + P_y(0,x)\right) . \tag{40}$$

Finally, if $\varphi(x') = x' + \varepsilon\alpha(x';\varepsilon)$ is the inverse diffeomorphism of $x \to x + \varepsilon a(x)$ (compare (20)), then, by Taylor's formula,

$$Q(y',\varphi(x')) = Q(y',x') + \varepsilon \int_0^1 Q_x(y', x' + \varepsilon\alpha t)\cdot \alpha \, dt .$$

In conclusion, we have

Proposition 1 *If ϕ_1 is defined in (19)–(18) with s, b and a given in (37), (39) and (40) respectively, then (23) holds with*

$$\begin{cases} E_1 := E + \varepsilon\widetilde{E} \\ \quad \widetilde{E} := \omega\cdot b + P_0(0) \\ Q_1(y',x') := Q(y',x') + \varepsilon\widetilde{Q}(y',x') \\ \quad \widetilde{Q} := \int_0^1 Q_x(y',x'+t\varepsilon\alpha)\cdot\alpha dt + Q'(y',\varphi(x')) \\ P_1(y',x') := P'(y',\varphi(x')) \end{cases} \tag{41}$$

with Q' and P' defined in (26), (27) and φ in (20).

Remark 5 The main technical problem is now transparent: because of the appearance of the *small divisors* $\omega\cdot n$ (which may become arbitrarily small), the solution $D_\omega^{-1}f$ is *less regular* than f so that the approximation scheme cannot work on a fixed function space. To overcome this fundamental problem – which even Poincaré was unable to solve notwithstanding his enormous efforts (see, e. g., [49]) – three ingredients are necessary:

(i) To set up a Newton scheme: this step has just been performed and it has been summarized in the above Proposition 1; such schemes have the following features: they are "quadratic" and, furthermore, after one step one has reproduced the initial situation (i. e., the form of H_1 in (23) has the same properties of H_0). It is important to notice that the new perturbation $\varepsilon^2 P_1$ is proportional to the *square* ε; thus, if one could iterate j times, at the jth step, would find

$$H_j = H_{j-1}\circ\phi_j = K_j + \varepsilon^{2^j}P_j\,. \tag{42}$$

The appearance of the exponential of the exponential of ε justifies the term "super-convergence" used, sometimes, in connection with Newton schemes.

(ii) One needs to introduce a *scale of Banach function spaces* $\{\mathcal{B}_\xi : \xi > 0\}$ with the property that $\mathcal{B}_{\xi'} \subset \mathcal{B}_\xi$ when $\xi < \xi'$: the generating functions ϕ_j will belong to $\mathcal{B}_{\xi_j}$ for a suitable decreasing sequence ξ_j;

(iii) One needs to control the small divisors at each step and this is granted by Kolmogorov's idea of keeping the frequency *fixed* in the normal form so that one can systematically use the Diophantine estimate (9).

Kolmogorov in his paper very neatly explained steps (i) and (iii) but did not provide the details for step (ii); in this regard he added: "*Only the use of condition* (9) *for proving the convergence of the recursions, ϕ_j, to the analytic limit for the recursion ϕ_* is somewhat more subtle*". In the next paragraph we shall introduce classical Banach spaces and discuss the needed straightforward estimates.

Step 2: Estimates

For $\xi \le 1$, we denote by $\mathcal{B}_\xi$ the space of function $f\colon B(0,\xi)\times\mathbb{T}^d \to \mathbb{R}$ analytic on

$$W_\xi := D(0,\xi)\times\mathbb{T}^d_\xi\,, \tag{43}$$

where

$$D(0,\xi) := \{y\in\mathbb{C}^d : |y| < \xi\} \quad \text{and} \\ \mathbb{T}^d_\xi := \{x\in\mathbb{C}^d : |\mathrm{Im}x_j| < \xi\}/(2\pi\mathbb{Z}^d) \tag{44}$$

with finite sup-norm

$$\|f\|_\xi := \sup_{D(0,\xi)\times\mathbb{T}^d_\xi} |f|\,, \tag{45}$$

(in other words, $\mathbb{T}^d_\xi$ denotes the complex points x with real parts $\mathrm{Re}x_j$ defined modulus 2π and imaginary part $\mathrm{Im}x_j$ with absolute value less than ξ).

The following properties are elementary:

(P1) $\mathcal{B}_\xi$ equipped with the $\|\cdot\|_\xi$ norm is a Banach space;

(P2) $\mathcal{B}_{\xi'} \subset \mathcal{B}_\xi$ when $\xi < \xi'$ and $\|f\|_\xi \le \|f\|_{\xi'}$ for any $f\in\mathcal{B}_{\xi'}$;

(P3) if $f\in\mathcal{B}_\xi$, and $f_n(y)$ denotes the n-Fourier coefficient of the periodic function $x\to f(y,x)$, then

$$|f_n(y)| \le \|f\|_\xi \mathrm{e}^{-|n|\xi}\,, \quad \forall n\in\mathbb{Z}^d,\ \forall y\in D(0,\xi)\,. \tag{46}$$

Another elementary property, which together with **(P3)** may found in any book of complex variables (e. g., [1]), is the following "Cauchy estimate" (which is based on Cauchy's integral formula):

(P4) let $f\in\mathcal{B}_\xi$ and let $p\in\mathbb{N}$ then there exists a constant $B_p = B_p(d) \ge 1$ such that, for any multi-index $(\alpha,\beta)\in\mathbb{N}^d\times\mathbb{N}^d$ with $|\alpha|+|\beta|\le p$ (as above for integer vectors α, $|\alpha| = \sum_j|\alpha_j|$) and for any $0\le\xi'<\xi$ one has

$$\|\partial_y^\alpha\partial_x^\beta f\|_{\xi'} \le B_p\|f\|_\xi(\xi-\xi')^{-(|\alpha|+|\beta|)}\,. \tag{47}$$

Finally, we shall need estimates on $D_\omega^{-1}f$, i. e., on solutions of (30):

(P5) Assume that $x\to f(x)\in\mathcal{B}_\xi$ has a zero average (all above definitions may be easily adapted to functions depending only on x); assume that $\omega\in\mathcal{D}^d_{\kappa,\tau}$ (recall Sect. "Introduction", point (f)), and let $p\in\mathbb{N}$. Then, there exist constants $\bar{B}_p = \bar{B}_p(d,\tau)\ge 1$ and $k_p = k_p(d,\tau)\ge 1$ such that, for any multi-index $\beta\in\mathbb{N}^d$ with $|\beta|\le p$ and for any $0\le\xi'<\xi$ one has

$$\|\partial_x^\beta D_\omega^{-1}f\|_{\xi'} \le \bar{B}_p\frac{\|f\|_\xi}{\kappa}(\xi-\xi')^{-k_p}\,. \tag{48}$$

Remark 6

(i) A proof of (48) is easily obtained observing that by (35) and (46), calling $\delta := \xi - \xi'$, one has

$$\begin{aligned}\|\partial_x^\beta D_\omega^{-1} f\|_{\xi'} &\le \sum_{\substack{n\in\mathbb{Z}^d\\ n\neq 0}} \frac{|n|^{|\beta|}|f_n|}{|\omega\cdot n|} \mathrm{e}^{\xi'|n|}\\ &\le \|f\|_\xi \sum_{\substack{n\in\mathbb{Z}^d\\ n\neq 0}} \frac{|n|^{|\beta|+\tau}}{\kappa}\mathrm{e}^{-\delta|n|}\\ &= \frac{\|f\|_\xi}{\kappa}\delta^{-(|\beta|+\tau+d)}\\ &\quad\cdot \sum_{\substack{n\in\mathbb{Z}^d\\ n\neq 0}} [\delta|n|]^{|\beta|+\tau}\mathrm{e}^{-\delta|n|}\delta^d\\ &\le \mathrm{const}\frac{\|f\|_\xi}{\kappa}(\xi-\xi')^{-(|\beta|+\tau+d)}\,,\end{aligned}$$

where the last estimate comes from approximating the sum with the Riemann integral

$$\int_{\mathbb{R}^d} |y|^{|\beta|+\tau}\mathrm{e}^{-|y|}\mathrm{d}y\,.$$

More surprising (and much more subtle) is that (48) holds with $k_p = |\beta|+\tau$; such an estimate has been obtained by Rüssmann [54,55]. For other explicit estimates see, e. g., [11] or [12].

(ii) If $|\beta| > 0$ it is not necessary to assume that $\langle f\rangle = 0$.

(iii) Other norms may be used (and, sometimes, are more useful); for example, rather popular are Fourier norms

$$\|f\|'_\xi := \sum_{n\in\mathbb{Z}^d} |f_n|\mathrm{e}^{\xi|n|}\,; \tag{49}$$

see, e. g., [13] and references therein.

By the hypotheses of Theorem 1 it follows that there exist $0<\xi\le 1$, $\kappa>0$ and $\tau \ge d-1$ such that $H\in\mathcal{B}_\xi$ and $\omega\in\mathcal{D}^d_{\kappa,\tau}$. Denote

$$T := \langle Q_{yy}(0,\cdot)\rangle^{-1}\,,\quad M := \|P\|_\xi\,. \tag{50}$$

and let $C>1$ be a constant such that[16]

$$|E|, |\omega|, \|Q\|_\xi, \|T\| < C \tag{51}$$

(i. e., each term on the l.h.s. is bounded by the r.h.s.); finally, fix

$$0<\delta<\xi \quad\text{and define } \bar\xi := \xi - \frac{2}{3}\delta\,,\ \xi' := \xi-\delta\,. \tag{52}$$

The parameter ξ' will be the size of the domain of analyticity of the new symplectic variables (y',x'), domain on which we shall bound the Hamiltonian $H_1 = H\circ\phi_1$, while $\bar\xi$ is an intermediate domain where we shall bound various functions of y' and x.

By **(P4)** *and* **(P5)**, *it follows that there exist constants $\bar c = \bar c(d,\tau,\kappa) > 1$, $\bar\mu\in\mathbb{Z}_+$ and $\bar\nu = \bar\nu(d,\tau) > 1$ such that*[17]

$$\begin{cases} \|s_x\|_{\bar\xi}, |b|, |\widetilde{E}|, \|a\|_{\bar\xi}, \|a_x\|_{\bar\xi}, \|\beta_0\|_{\bar\xi}, \|\beta\|_{\bar\xi},\\ \qquad \|Q'\|_{\bar\xi}, \|\partial^2_{y'}Q'(0,\cdot)\|_0 \le \bar c C^{\bar\mu}\delta^{-\bar\nu}M =: \bar L\,,\\ \|P'\|_{\bar\xi} \le \bar c C^{\bar\mu}\delta^{-\bar\nu}M^2 = \bar L M\,. \end{cases} \tag{53}$$

The estimate in the first line of (53) allows us to construct, for ε small enough, the symplectic transformation ϕ_1, whose main properties are collected in the following

Lemma 1 *If $|\varepsilon|\le\varepsilon_0$ and ε_0 satisfies*

$$\varepsilon_0\bar L \le \frac{\delta}{3}\,, \tag{54}$$

then the map $\psi_\varepsilon(x) := x+\varepsilon a(x)$ has an analytic inverse $\varphi(x') = x'+\varepsilon\alpha(x';\varepsilon)$ such that, for all $|\varepsilon|<\varepsilon_0$,

$$\|\alpha\|_{\xi'} \le \bar L \quad \textit{and} \quad \varphi = \mathrm{id}+\varepsilon\alpha\colon \mathbb{T}^d_{\xi'} \to \mathbb{T}^d_{\bar\xi}\,. \tag{55}$$

Furthermore, for any $(y',x)\in W_{\bar\xi}$, $|y'+\varepsilon\beta(y',x)|<\xi$, so that

$$\begin{aligned}&\phi_1 = \big(y'+\varepsilon\beta(y',\varphi(x')),\varphi(x')\big) : W_{\xi'}\to W_\xi\,, \quad \textit{and}\\ &\|\phi_1 - \mathrm{id}\|_{\xi'} \le |\varepsilon|\bar L\,;\end{aligned} \tag{56}$$

finally, the matrix $\mathbb{1}_d + \varepsilon a_x$ is, for any $x\in\mathbb{T}^d_{\bar\xi}$, invertible with inverse $\mathbb{1}_d+\varepsilon S(x;\varepsilon)$ satisfying

$$\|S\|_{\bar\xi} \le \frac{\|a_x\|_{\bar\xi}}{1-|\varepsilon|\|a_x\|_{\bar\xi}} < \frac{3}{2}\bar L\,, \tag{57}$$

so that ϕ_1 defines a symplectic diffeomorphism.

The simple proof[18] of this statement is based upon standard tools in mathematical analysis such as the contraction mapping theorem or the inversion of close-to-identity matrices by Neumann series (see, e. g., [36]).

From the Lemma and the definition of P_1 in (41), it follows immediately that

$$\|P_1\|_{\xi'} \le \bar L\,. \tag{58}$$

Next, by the same technique used to derive (53), one can easily check that

$$\|\widetilde{Q}\|_{\xi'}\,,\ \ 2C^2\|\partial^2_{y'}\widetilde{Q}(0,\cdot)\|_0 \le cC^\mu\delta^{-\nu}M = L\,, \tag{59}$$

for suitable constants $c \geq \bar{c}$, $\bar{\mu} \geq \mu$, $\bar{\nu} \geq \nu$ (the factor $2C^2$ has been introduced for later convenience; notice also that $L \geq \bar{L}$). But, then, if

$$\varepsilon_0 L := \varepsilon_0 c C^{\mu} \delta^{-\nu} M \leq \frac{\delta}{3}\,, \tag{60}$$

there follows that[19] $\|\widetilde{T}\| \leq L$; this bound, together with (53), (59), (56), and (58), shows that

$$\begin{cases} |\tilde{E}|, \|\widetilde{Q}\|_{\xi'}, \|\widetilde{T}\|, \|\phi_1 - \mathrm{id}\|_{\xi'} \leq L \\ \|P_1\|_{\xi'} \leq LM\,; \end{cases} \tag{61}$$

provided (60) holds (notice that (60) implies (54)).

One step of the iteration has been concluded and the needed estimates obtained. The idea is to iterate the construction infinitely many times, as we proceed to describe.

Step 3: Iteration and Convergence

In order to iterate Kolmogorov's construction, analyzed in Step 2, so as to construct a sequence of symplectic transformations

$$\phi_j \colon W_{\xi_{j+1}} \to W_{\xi_j}\,, \tag{62}$$

closer and closer to the identity, and such that (42) hold, the first thing to do is to choose the sequence ξ_j: this sequence must be convergent, so that $\delta_j = \xi_j - \xi_{j+1}$ has to go to zero rather quickly. Inverse powers of δ_j (which, at the jth step will play the role of δ in the previous paragraph) appear in the smallness conditions (see, e. g., (54)): this "divergence" will, however, be beaten by the super-fast decay of ε^{2^j}.

Fix $0 < \xi_* < \xi$ (ξ_* will be the domain of analyticity of ϕ_* and K_* in Theorem 1 and, for $j \geq 0$, let

$$\begin{cases} \xi_0 := \xi \\ \delta_0 := \dfrac{\xi - \xi_*}{2} \end{cases} \qquad \begin{cases} \delta_j := \dfrac{\delta_0}{2^j} \\ \xi_{j+1} := \xi_j - \delta_j = \xi_* + \dfrac{\delta_0}{2^j} \end{cases} \tag{63}$$

and observe that $\xi_j \downarrow \xi_*$. With this choice[20], Kolmogorov's algorithm can be iterated infinitely many times, provided ε_0 is small enough. To be more precise, let c, μ and ν be as in (59), and define

$$C := 2 \max\left\{|E|, |\omega|, \|Q\|_{\xi}, \|T\|, 1\right\}\,. \tag{64}$$

Smallness Assumption: *Assume that $|\varepsilon| \leq \varepsilon_0$ and that ε_0 satisfies*

$$\varepsilon_0 D \|P\|_{\xi} \leq 1 \qquad \text{where } D := 3c\delta_0^{-(\nu+1)} C^{\mu}\,, \quad B := 2^{\nu+1}\,; \tag{65}$$

notice that the constant C in (64) satisfies (51) and that (65) implies (54). Then the following claim holds.

Claim C: *Under condition (65) one can iteratively construct a sequence of Kolmogorov symplectic maps ϕ_j as in (62) so that (42) holds in such a way that $\varepsilon^{2^j} P_j$, $\Phi_j := \phi_1 \circ \phi_2 \circ \cdots \circ \phi_j$, E_j, K_j, Q_j converge uniformly on W_{ξ_*} to, respectively, 0, ϕ_*, E_*, K_*, Q_*, which are real-analytic on W_{ξ_*} and $H \circ \phi_* = K_* = E_* + \omega \cdot y + Q_*$ with $Q_* = O(|y|^2)$. Furthermore, the following estimates hold for any $|\varepsilon| \leq \varepsilon_0$ and for any $i \geq 0$:*

$$|\varepsilon|^{2^i} M_i := |\varepsilon|^{2^i} \|P_i\|_{\xi_i} \leq \frac{(|\varepsilon| DBM)^{2^i}}{DB^{i+1}}\,, \tag{66}$$

$$\|\phi_* - \mathrm{id}\|_{\xi_*}, |E - E_*|, \|Q - Q_*\|_{\xi_*}, \|T - T_*\| \leq |\varepsilon| DBM\,, \tag{67}$$

where $T_* := \langle \partial_y^2 Q_*(0,\cdot)\rangle^{-1}$, showing that K_* is non-degenerate.

Remark 7

(i) From Claim **C** Kolmogorov Theorem 1 follows at once. In fact we have proven the following quantitative statement: *Let $\omega \in \mathcal{D}^d_{\kappa,\tau}$ with $\tau \geq d-1$ and $0 < \kappa < 1$; let Q and P be real-analytic on $W_\xi = D^d(0,\xi) \times \mathbb{T}^d_\xi$ for some $0 < \xi \leq 1$ and let $0 < \theta < 1$; let T and C be as in, respectively, (50) and (64). There exist $c_* = c_*(d,\tau,\kappa,\theta) > 1$ and positive integers $\sigma = \sigma(d,\tau)$, μ such that if*

$$|\varepsilon| \leq \varepsilon_* := \frac{\xi^{\sigma}}{c_* \|P\|_{\xi} C^{\mu}} \tag{68}$$

then one can construct a near-to-identity Kolmogorov transformation (Remark 3–(ii)) $\phi_* \colon W_{\theta\xi} \to W_\xi$ such that the thesis of Theorem 1 holds together with the estimates

$$\|\phi_* - \mathrm{id}\|_{\theta\xi}, |E - E_*|, \|Q - Q_*\|_{\theta\xi}, \|T - T_*\| \leq \frac{|\varepsilon|}{\varepsilon_*} = |\varepsilon| c_* \|P\|_{\xi} C^{\mu} \xi^{-\sigma}\,. \tag{69}$$

(The correspondence with the above constants being: $\xi_* = \theta\xi, \delta_0 = \xi(1-\theta)/2, \sigma = \nu + 1, D = 3c(2/(1-\theta))^{\nu+1} C^{\mu}, c_* = 3c(4/(1-\theta))^{\nu+1}$).

(ii) From Cauchy estimates and (67), it follows that $\|\phi_* - \mathrm{id}\|_{C^p}$ and $\|Q - Q_*\|_{C^p}$ are small for any p (small in $|\varepsilon|$ but not uniformly in[21] p).

(iii) All estimates are uniform in ε, therefore, from Weierstrass theorem (compare note 18) it follows that ϕ_* and K_* are *analytic in ε in the complex ball of radius* ε_0.

Power series expansions in ε were very popular in the nineteenth and twentieth centuries[22], however convergence of the formal ε-power series of quasi-periodic solutions was proved for the first time only in the 1960s thanks to KAM theory [45]. Some of this matter is briefly discussed in Sect. "Future Directions" below.

(iv) **The Nearly-Integrable Case** In [35] it is pointed out that Kolmogorov's Theorem easily yields the existence of many KAM tori for nearly-integrable systems (14) for $|\varepsilon|$ small enough, provided K is nondegenerate in the sense that

$$\det K_{yy}(y_0) \neq 0 \,. \tag{70}$$

In fact, without loss of generality we may assume that $\omega := H_0'$ is a diffeomorphism on $B(y_0, 2r)$ and $\det K_{yy}(y) \neq 0$ for all $y \in B(y_0, 2r)$. Furthermore, letting $B = B(y_0, r)$, fixing $\tau > d-1$ and denoting by ℓ_d the Lebesgue measure on $\mathbb{R}^d$, from the remark in note 11 and from the fact that ω is a diffeomorphism, there follows that there exists a constant $c_\#$ depending only on d, τ and r such that

$$\ell_d(\omega(B) \backslash \mathcal{D}^d_{\kappa,\tau}), \ell_d(\{y \in B \colon \omega(y) \notin \mathcal{D}^d_{\kappa,\tau}\}) < c_\# \kappa \,. \tag{71}$$

Now, let $B_{\kappa,\tau} := \{y \in B \colon \omega(y) \in \mathcal{D}^d_{\kappa,\tau}\}$ (which, by (71) has Lebesgue measure $\ell_d(B_{\kappa,\tau}) \geq \ell_d(B) - c_\#\kappa$), then for any $\bar{y} \in B_{\kappa,\tau}$ we can make the trivial symplectic change of variables $y \to \bar{y} + y$, $x \to x$ so that K can be written as in (10) with

$$E := K(\bar{y}) \,, \quad \omega := K_y(\bar{y}) \,,$$
$$Q(y, x) = Q(y) := K(y) - K(\bar{y}) - K_y(\bar{y}) \cdot y \,,$$

(where, for ease of notation, we did not change names to the new symplectic variables) and $P(\bar{y} + y, x)$ replacing (with a slight abuse of notation) $P(y, x)$. By Taylor's formula, $Q = O(|y|^2)$ and, furthermore (since $Q(y, x) = Q(y)$, $\langle \partial_y^2 Q(0, x)\rangle = Q_{yy}(0) = K_{yy}(\bar{y})$, which is invertible according to out hypotheses. Thus K is Kolmogorov non-degenerate and Theorem 1 can be applied yielding, for $|\varepsilon| < \varepsilon_0$, a KAM torus $\mathcal{T}_{\omega,\varepsilon}$, with $\omega = K_y(\bar{y})$, for each $\bar{y} \in B_{\kappa,\tau}$. Notice that the measure of initial phase points, which, perturbed, give rise to KAM tori, has a small complementary bounded by $c_\#\kappa$ (see (71)).

(v) In the nearly-integrable setting described in the preceding point, the union of KAM tori is usually called the **Kolmogorov set**. It is not difficult to check that the dependence upon $\bar{y}$ of the Kolmogorov transformation ϕ_* is Lipschitz[23], implying that the measure of the complementary of Kolmogorov set itself is also bounded by $\hat{c}_\#\kappa$ with a constant $\hat{c}_\#$ depending only on d, τ and r.
Indeed, the estimate on the measure of the Kolmogorov set can be made more quantitative (i. e., one can see how such an estimate depends upon ε as $\varepsilon \to 0$). In fact, revisiting the estimates discussed in **Step 2** above one sees easily that the constant c defined in (53) has the form[24]

$$c = \hat{c}\kappa^{-4} \,. \tag{72}$$

where $\hat{c} = \hat{c}(d, \tau)$ depends only on d and τ (here the Diophantine constant κ is assumed, without loss of generality, to be smaller than one). Thus the smallness condition (65) reads $\varepsilon_0 \kappa^{-4} \bar{D} \leq 1$ with some constant $\bar{D}$ independent of κ: such condition is satisfied by choosing $\kappa = (\bar{D}\varepsilon_0)^{1/4}$ and since $\hat{c}_\#\kappa$ was an upper bound on the complementary of Kolmogorov set, we see that *the set of phase points which do not lie on KAM tori may be bounded by a constant times* $\sqrt[4]{\varepsilon_0}$. Actually, it turns that this bound is not optimal, as we shall see in the next section: see Remark 10.

(vi) The proof of claim **C** follows easily by induction on the number j of the iterative steps[25].

Arnold's Scheme

The first detailed proof of Kolmogorov Theorem, in the context of nearly-integrable Hamiltonian systems (compare Remark 1–(iii)), was given by V.I. Arnold in 1963.

Theorem 2 (Arnold [2]) *Consider a one-parameter family of nearly-integrable Hamiltonians*

$$H(y, x; \varepsilon) := K(y) + \varepsilon P(y, x) \quad (\varepsilon \in \mathbb{R}) \tag{73}$$

with K and P real-analytic on $\mathcal{M} := B(y_0, r) \times \mathbb{T}^d$ (endowed with the standard symplectic form $dy \wedge dx$) satisfying

$$K_y(y_0) = \omega \in \mathcal{D}^d_{\kappa,\tau} \,, \quad \det K_{yy}(y_0) \neq 0 \,. \tag{74}$$

Then, if ε is small enough, there exists a real-analytic embedding

$$\phi \colon \theta \in \mathbb{T}^d \to \mathcal{M} \tag{75}$$

close to the trivial embedding (y_0, id), such that the d-torus

$$\mathcal{T}_{\omega,\varepsilon} := \phi\left(\mathbb{T}^d\right) \tag{76}$$

is invariant for H and

$$\phi_H^t \circ \phi(\theta) = \phi(\theta + \omega t) , \tag{77}$$

showing that such a torus is a non-degenerate KAM torus for H.

Remark 8

(i) The above Theorem is a corollary of Kolmogorov Theorem 1 as discussed in Remark 7–(iv).

(ii) Arnold's proof of the above Theorem *is not based* upon Kolmogorov's scheme and is rather different in spirit – although still based on a Newton method – and introduces several interesting technical ideas.

(iii) Indeed, the iteration scheme of Arnold is more classical and, from the algebraic point of view, easier than Kolmogorov's, but the estimates involved are somewhat more delicate and introduce a logarithmic correction, so that, in fact, the smallness parameter will be

$$\epsilon := |\varepsilon|(\log|\varepsilon|^{-1})^{\rho} \tag{78}$$

(for some constant $\rho = \rho(d, \tau) \geq 1$) rather than $|\varepsilon|$ as in Kolmogorov's scheme; see, also, Remark 9–(iii) and (iv) below.

Arnold's Scheme

Without loss of generality, one may assume that K and P have analytic and bounded extension to $W_{r,\xi}(y_0) := D(y_0, r) \times \mathbb{T}^d_\xi$ for some $\xi > 0$, where, as above, $D(y_0, r)$ denotes the complex ball of center y_0 and radius r. We remark that, in what follows, the analyticity domains of actions and angles play a different role

The Hamiltonian H in (73) admits, for $\varepsilon = 0$ the (KAM) invariant torus $\mathcal{T}_{\omega,0} = \{y_0\} \times \mathbb{T}^d$ on which the K-flow is given by $x \to x + \omega t$. Arnold's basic idea is to find a symplectic transformation

$$\phi_1 \colon W_1 := D(y_1, r_1) \times \mathbb{T}^d_{\xi_1} \to W_0 := D(y_0, r) \times \mathbb{T}^d_{\xi} , \tag{79}$$

so that $W_1 \subset W_0$ and

$$\begin{cases} H_1 := H \circ \phi_1 = K_1 + \varepsilon^2 P_1 , \quad K_1 = K_1(y) , \\ \partial_y K_1(y_1) = \omega , \quad \det \partial_y^2 K_1(y_1) \neq 0 \end{cases} \tag{80}$$

(with abuse of notation we denote here the new symplectic variables with the same name of the original variables; as above, dependence on ε will, often, not be explicitly indicated). In this way the initial set up is reconstructed and, for ε small enough, one can iterate the scheme so as to build a sequence of symplectic transformations

$$\phi_j \colon W_j := D(y_j, r_j) \times \mathbb{T}^d_{\xi_j} \to W_{j-1} \tag{81}$$

so that

$$\begin{cases} H_j := H_{j-1} \circ \phi_j = K_j + \varepsilon^{2^j} P_j , \quad K_j = K_j(y) , \\ \partial_y K_j(y_j) = \omega , \quad \det \partial_y^2 K_j(y_j) \neq 0 . \end{cases} \tag{82}$$

Arnold's transformations, as in Kolmogorov's case, are closer and closer to the identity, and the limit

$$\begin{aligned} \phi(\theta) &:= \lim_{j\to\infty} \Phi_j(y_j, \theta) , \\ \Phi_j &:= \phi_1 \circ \cdots \circ \phi_j \colon W_j \to W_0 , \end{aligned} \tag{83}$$

defines a real-analytic embedding of $\mathbb{T}^d$ into the phase space $B(y_0, r) \times \mathbb{T}^d$, which is close to the trivial embedding (y_0, id); furthermore, the torus

$$\mathcal{T}_{\omega,\varepsilon} := \phi(\mathbb{T}^d) = \lim_{j\to\infty} \Phi_j(y_j, \mathbb{T}^d) \tag{84}$$

is invariant for H and (77) holds as announced in Theorem 2. Relation (77) follows from the following argument. The radius r_j will turn out to tend to 0 but in a much slower way than $\varepsilon^{2^j} P_j$. This fact, together with the rapid convergence of the symplectic transformation Φ_j in (83) implies

$$\begin{aligned} \phi_H^t \circ \phi(\theta) &= \lim_{j\to\infty} \phi_H^t\left(\Phi_j(y_j, \theta)\right) \\ &= \lim_{j\to\infty} \Phi_j \circ \phi_{H_j}^t(y_j, \theta) \\ &= \lim_{j\to\infty} \Phi_j(y_j, \theta + \omega t) \\ &= \phi(\theta + \omega t) \end{aligned} \tag{85}$$

where: the first equality is just smooth dependence upon initial data of the flow ϕ_H^t together with (83); the second equality is (3); the third equality is due to the fact that (see (82)) $\phi_{H_j}^t(y_j, \theta) = \phi_{K_j}^t(y_j, \theta) + O(\varepsilon^{2^j}\|P_j\|) = (y_j, \theta + \omega t) + O(\varepsilon^{2^j}\|P_j\|)$ and $O(\varepsilon^{2^j}\|P_j\|)$ goes very rapidly to zero; the fourth equality is again (83).

Arnold's Transformation

Let us look for a near-to-the-identity transformation ϕ_1 so that the first line of (80) holds; this transformation will be determined by a generating function of the form

$$y' \cdot x + \varepsilon g(y', x) , \quad \begin{cases} y = y' + \varepsilon g_x(y', x) \\ x' = x + \varepsilon g_{y'}(y', x) . \end{cases} \tag{86}$$

Inserting $y = y' + \varepsilon g_x(y', x)$ into H, one finds

$$H(y' + \varepsilon g_x, x) = K(y') + \varepsilon \left[K_y(y') \cdot g_x + P(y', x)\right] + \varepsilon^2 \left(P^{(1)} + P^{(2)}\right) \quad (87)$$

with (compare (26))

$$\begin{aligned} P^{(1)} &:= \frac{1}{\varepsilon^2}[K(y' + \varepsilon g_x) - K(y') - \varepsilon K_y(y') \cdot g_x] \\ &= \int_0^1 (1-t) K_{yy}(y' + t\varepsilon g_x, x)\, g_x \cdot g_x \mathrm{d}t \\ P^{(2)} &:= \frac{1}{\varepsilon}[P(y' + \varepsilon g_x, x) - P(y', x)] \\ &= \int_0^1 P_y(y' + t\varepsilon g_x, x) \cdot g_x \mathrm{d}t\,. \end{aligned} \quad (88)$$

Remark 9

(i) The (naive) idea is to try determine g so that

$$K_y(y') \cdot g_x + P(y', x) = \text{function of } y' \text{ only}, \quad (89)$$

however, such a relation is impossible to achieve. First of all, by taking the x-average of both sides of (89) one sees that the "function of y' only" has to be the mean of $P(y', \cdot)$, i. e., the zero-Fourier coefficient $P_0(y')$, so that the *formal* solution of (89), is (by Fourier expansion)

$$\begin{cases} g = \sum_{n \neq 0} \dfrac{-P_n(y')}{iK_y(y') \cdot n} \mathrm{e}^{in\cdot x}\,, \\ K_y(y') \cdot g_x + P(y', x) = P_0(y')\,. \end{cases} \quad (90)$$

But, (at difference with Kolmogorov's scheme) the frequency $K_y(y')$ is *a function of the action* y' and since, by the Inverse Function Theorem (Appendix "A The Classical Implicit Function Theorem"), $y \to K_y(y)$ is a local diffeomorphism, it follows that, in any neighborhood of y_0, there are points y such that $K_y(y) \cdot n = 0$ for some[26] $n \in \mathbb{Z}^d$. Thus, in any neighborhood of y_0, some divisors in (90) will actually vanish and, therefore, *an analytic solution g cannot exist*[27].

(ii) On the other hand, since $K_y(y_0)$ is rationally independent, it is clearly possible (simply by continuity) to control a finite number of divisors in a suitable neighborhood of y_0, more precisely, for any $N \in \mathbb{N}$ one can find $\bar{r} > 0$ such that

$$K_y(y) \cdot n \neq 0\,, \quad \forall y \in D(y_0, \bar{r})\,, \quad \forall 0 < |n| \leq N\,; \quad (91)$$

the important quantitative aspects will be shortly discussed below.

(iii) Relation (89) is also one of the main "identity" in *Averaging Theory* and is related to the so-called *Hamilton–Jacobi equation*. Arnold's proof makes such a theory rigorous and shows how a Newton method can be built upon it in order to establish the existence of invariant tori. In a sense, Arnold's approach is more classical than Kolmogorov's.

(iv) When (for a given y and n) it occurs that $K_y(y)\cdot n = 0$ one speaks of an (exact) *resonance*. As mentioned at the end of point (i), in the general case, *resonances are dense*. This represents the main problem in Hamiltonian perturbation theory and is a typical feature of *conservative systems*. For generalities on Averaging Theory, Hamilton–Jacobi equation, resonances etc. see, e. g., [5] or Sects. 6.1 and 6.2 of [6].

The key (simple!) idea of Arnold is to split the perturbation into two terms

$$P = \hat{P} + \check{P} \quad \text{where} \quad \begin{cases} \hat{P} := \sum_{|n| \leq N} P_n(y) \mathrm{e}^{in\cdot x} \\ \check{P} := \sum_{|n| > N} P_n(y) \mathrm{e}^{in\cdot x} \end{cases} \quad (92)$$

choosing N so that

$$\check{P} = O(\varepsilon) \quad (93)$$

(this is possible because of the fast decay of the Fourier coefficients of P; compare (33)). Then, for $\varepsilon \neq 0$, (87) can be rewritten as follows

$$H(y' + \varepsilon g_x, x) = K(y') + \varepsilon \left[K_y(y') \cdot g_x + \hat{P}(y', x)\right] + \varepsilon^2 \left(P^{(1)} + P^{(2)} + P^{(3)}\right) \quad (94)$$

with $P^{(1)}$ and $P^{(2)}$ as in (88) and

$$P^{(3)}(y', x) := \frac{1}{\varepsilon} \check{P}(y', x)\,. \quad (95)$$

Thus, letting[28]

$$g = \sum_{0 < |n| \leq N} \frac{-P_n(y')}{iK_y(y') \cdot n} \mathrm{e}^{in\cdot x}\,, \quad (96)$$

one gets

$$H(y' + \varepsilon g_x, x) = K_1(y') + \varepsilon^2 P'(y', x) \quad (97)$$

where

$$\begin{aligned} K_1(y') &:= K(y') + \varepsilon P_0(y')\,, \\ P'(y', x) &:= P^{(1)} + P^{(2)} + P^{(3)}\,. \end{aligned} \quad (98)$$

Now, by the IFT (Appendix "A The Classical Implicit Function Theorem"), for ε small enough, the map $x \to x + g_{y'}(y', x)$ can be inverted with a real-analytic map of the form

$$\varphi(y', x'; \varepsilon) := x' + \varepsilon\alpha(y', x'; \varepsilon) \tag{99}$$

so that Arnold's symplectic transformation is given by

$$\phi_1 \colon (y', x') \to \begin{cases} y = y' + \varepsilon g_x\left(y', \varphi(y', x'; \varepsilon)\right) \\ x = \varphi(y', x'; \varepsilon) \\ \quad = x' + \varepsilon\alpha(y', x', ; \varepsilon) \end{cases} \tag{100}$$

(compare (21)). To finish the construction, observe that, from the IFT (see Appendix "A The Classical Implicit Function Theorem" and the quantitative discussion below) it follows that there exists a (unique) point $y_1 \in B(y_0, \bar{r})$ so that the second line of (80) holds, provided ε is small enough.

In conclusion, the analogue of Proposition 1 holds, describing Arnold's scheme:

Proposition 2 *If ϕ_1 is defined in (100) with g given in (96) (with N so that (93) holds) and φ given in (99), then (80) holds with K_1 as in (98) and $P_1(y', x') := P'(y', \varphi(y', x'))$ with P' defined in (98), (95) and (88).*

Estimates and Convergence

If f is a real-analytic function with analytic extension to $W_{r,\xi}$, we denote, for any $r' \le r$ and $\xi' \le \xi$,

$$\|f\|_{r',\xi'} := \sup_{W_{r',\xi'}(y_0)} |f(y, x)| ; \tag{101}$$

furthermore, we define

$$T := K_{yy}(y_0)^{-1} , \quad M := \|P\|_{r,\xi} , \tag{102}$$

and assume (without loss of generality)

$$\kappa < 1 , \quad r < 1 , \quad \xi \le 1 , \\ \max\{1, \|K_y\|_r, \|K_{yy}\|_r, \|T\|\} < C , \tag{103}$$

for a suitable constant C (which, as above, will not change during the iteration).

We begin by discussing how N and $\bar{r}$ depend upon ε. From the exponential decay of the Fourier coefficients (33), it follows that, choosing

$$N := 5\delta^{-1}\lambda , \quad \text{where } \lambda := \log|\varepsilon|^{-1} , \tag{104}$$

then

$$\|\check{P}\|_{r,\xi-\frac{\delta}{2}} \le |\varepsilon| M \tag{105}$$

provided

$$|\varepsilon| \le \text{const}\,\delta \tag{106}$$

for a suitable[29] const $=$ const(d).

The second key inequality concerns the control of the small divisors $K_y(y') \cdot n$ appearing in the definition of g (see (96)), in a neighborhood $D(y_0, \bar{r})$ of y_0: this will determine the size of $\bar{r}$.

Recalling that $K_y(y_0) = \omega \in \mathcal{D}^d_{\tau,\kappa}$, by Taylor's formula and (9), one finds, for any $0 < |n| \le N$ and any $y' \in D(y_0, \bar{r})$,

$$\begin{aligned} |K_y(y') \cdot n| &= \left|\omega \cdot n + \left(K_y(y') - K_y(y_0)\right) \cdot n\right| \\ &\ge |\omega \cdot n| \left(1 - \frac{\|K_{yy}\|_r}{|\omega \cdot n|}|n|\bar{r}\right) \\ &\ge \frac{\kappa}{|n|^\tau}\left(1 - \frac{C}{\kappa}|n|^{\tau+1}\bar{r}\right) \\ &\ge \frac{\kappa}{|n|^\tau}\left(1 - \frac{C}{\kappa}N^{\tau+1}\bar{r}\right) \\ &\ge \frac{1}{2}\frac{\kappa}{|n|^\tau} , \end{aligned} \tag{107}$$

provided $\bar{r} \le r$ satisfies also

$$\bar{r} \le \frac{\kappa}{2CN^{\tau+1}} \overset{(104)}{=} \frac{\kappa}{2 \cdot 5^{\tau+1} C(\delta^{-1}\lambda)^{\tau+1}} . \tag{108}$$

Equation (107) allows us to easily control Arnold's generating function g. For example:

$$\begin{aligned} \|g_x\|_{\bar{r},\xi-\frac{\delta}{2}} &= \sup_{D(y_0,\bar{r})\times\mathbb{T}^d_{\xi-\frac{\delta}{2}}} \left| \sum_{0<|n|\le N} \frac{nP_n(y')}{K_y(y') \cdot n} e^{in\cdot x} \right| \\ &\le \sum_{0<|n|\le N} \frac{\sup_{D(y_0,r)} |P_n(y')|}{|K_y(y') \cdot n|} |n| e^{(\xi-\frac{\delta}{2})|n|} \\ &\le \sum_{n\in\mathbb{Z}^d} M \frac{2|n|^{\tau+1}}{\kappa} e^{-\frac{\delta}{2}|n|} \\ &\le \text{const}\frac{M}{\kappa}\delta^{-(\tau+1+d)} , \end{aligned} \tag{109}$$

where "const" denotes a constant depending on d and τ only; compare also Remark 6–(i).

Let us now discuss, from a quantitative point of view, how to choose the new "center" of the action variables y_1, which is determined by the requirements in (80). Assuming that

$$\bar{r} \le \frac{r}{2} \tag{110}$$

(allowing the use of Cauchy estimates for y-derivatives of K or P in $D(y_0, \bar r)$), it is not difficult to see that the quantitative IFT of Appendix "A The Classical Implicit Function Theorem" implies that there exists a unique $y_1 \in D(y_0, \bar r)$ such that (80) holds and, furthermore

$$|y_1 - y_0| \le 4CMr^{-1}|\varepsilon| \,, \tag{111}$$

and

$$\begin{aligned} \partial_y^2 K_1(y_1) &:= K_{yy}(y_1) + \varepsilon \partial_y^2 P_0(y_1) \\ &=: T^{-1}(\mathbb{1}_d + A) \end{aligned} \tag{112}$$

with a matrix A satisfying

$$\|A\| \le 10C^3 M|\varepsilon| \le \frac{1}{2} \tag{113}$$

provided[30]

$$\begin{cases} 8C^2 \dfrac{\bar r}{r} \le 1 \,, \\ 8CM\bar r^{-2}|\varepsilon| \le 1 \end{cases} \tag{114}$$

Equation (113) shows that $\partial_y^2 K(y_1)$ is invertible (Neumann series) and that[31]

$$\partial_y^2 K_1(y_1)^{-1} = T + \varepsilon \widetilde T \,, \quad \|\widetilde T\| \le 20C^3 M \,. \tag{115}$$

Finally, notice that the second conditions in (114) and (111) imply that $|y_1 - y_0| < \bar r/2$ so that

$$D(y_1, \bar r/2) \subset D(y_0, \bar r) \,. \tag{116}$$

Now, all the estimating tools are set up and, writing

$$\begin{aligned} K_1 &:= K + \varepsilon \widetilde K = K + \varepsilon P_0(y') \,, \\ y_1 &:= y_0 + \varepsilon \tilde y \,, \end{aligned} \tag{117}$$

one can easily prove (along the lines that led to (53)) the following estimates, where as in Sect. "Kolmogorov Theorem", $\bar\xi := \xi - \frac{2}{3}\delta$ and $\bar r$ is as above:

$$\begin{cases} \dfrac{\|g_x\|_{\bar r, \bar\xi}}{r}, \|g_y\|_{\bar r,\bar\xi}, \|\widetilde K_y\|_{\bar r}, \|\widetilde K_{yy}\|, |\tilde y|, \|\widetilde T\| \\ \qquad \le c\kappa^{-2} C^\mu \delta^{-\nu} \lambda^\rho M =: L \,, \\ \|P'\|_{\bar r, \bar\xi} \le c\kappa^{-2} C^\mu \delta^{-\nu} \lambda^\rho M^2 = LM \,, \end{cases} \tag{118}$$

where $c = c(d, \tau) > 1$, $\mu \in \mathbb{Z}_+$, ν and ρ are positive integers depending on d and τ. Now, by[32] Lemma 1 and (118), one has that map $x \to x + \varepsilon g_y(y', x)$ has, for any $y' \in D_{\bar r}(y_0)$, an analytic inverse $\varphi = x' + \varepsilon\alpha(x'; y', \varepsilon) =: \varphi(y', x')$ on $\mathbb{T}^d_{\bar\xi - \frac{\delta}{3}}$ provided (54), with $\bar L$ replaced by L in (118), holds, in which case (55) holds (for any $|\varepsilon| \le \varepsilon_0$ and any $y' \in D_r(y_0)$). Furthermore, under the above hypothesis, it follows that[33]

$$\begin{cases} \phi_1 := \big(y' + \varepsilon g_x(y', \varphi(y', x')), \varphi(y', x')\big) : \\ \qquad W_{\bar r/2, \xi - \delta}(y_1) \to W_{r,\xi}(y_0) \\ \|\phi_1 - \mathrm{id}\|_{\bar r/2, \xi - \delta} \le |\varepsilon| L \,. \end{cases} \tag{119}$$

Finally, letting $P_1(y', x') := P'(y', \varphi(y', x'))$ one sees that P_1 is real-analytic on $W_{\bar r/2, \bar\xi - \delta}(y_1)$ and bounded on that domain by

$$\|P_1\|_{\bar r/2, \xi - \delta} \le LM \,. \tag{120}$$

In order to iterate the above construction, we fix $0 < \xi_* < \xi$ and set

$$\begin{aligned} C &:= 2\max\{1, \|K_y\|_r, \|K_{yy}\|_r, \|T\|\} \,, \\ \gamma &:= 3C \,, \\ \delta_0 &:= \frac{(\gamma - 1)(\xi - \xi_*)}{\gamma} \,; \end{aligned} \tag{121}$$

ξ_j and δ_j as in (63) *but with* δ_0 *as in* (121); we also define, for any $j \ge 0$,

$$\begin{aligned} \lambda_j &:= 2^j \lambda = \log \varepsilon_0^{-2^j} \,, \\ r_j &:= \frac{\kappa}{4 \cdot 5^{\tau+1} C(\delta_j^{-1}\lambda_j)^{\tau+1}} \,; \end{aligned} \tag{122}$$

(this part is adapted from **Step 3** in Sect. "Kolmogorov Theorem"; see, in particular, (103)). With such choices it is not difficult to check that the iterative construction may be carried out infinitely many times yielding, as a byproduct, Theorem 2 with ϕ real-analytic on $\mathbb{T}^d_{\xi_*}$, provided $|\varepsilon| \le \varepsilon_0$ with ε_0 satisfying[34]

$$\begin{cases} \varepsilon_0 \le \mathrm{e}^{-\beta} & \text{with} \quad \beta := \dfrac{\delta_0}{5}\left(\dfrac{\kappa}{Cr}\right)^{\frac{1}{\tau+1}} \\ \varepsilon_0 DB\|P\|_\xi \le 1 & \text{with} \quad D := 3c\kappa^{-2}\delta_0^{-(\nu+1)} C^\mu \,, \\ & \qquad\quad B := \gamma^{\nu+1}(\log\varepsilon_0^{-1})^\rho \,. \end{cases} \tag{123}$$

Remark 10 Notice that the power of κ^{-1} (the inverse of the Diophantine constant) in the second smallness condition in (123) is two, which implies (compare Remark 7–(v)) that the measure of the complement of the Kolmogorov set may be bounded by a constant times $\sqrt{\varepsilon_0}$. This bound is optimal as the trivial example $(y_1^2 + y_2^2)/2 +$

$\varepsilon\cos(x_1)$ shows: the Hamiltonian is integrable and the phase portrait shows that the separatrices of the pendulum $y_1^2/2 + \varepsilon\cos x_1$ bound a region of area $\sqrt{|\varepsilon|}$ with no KAM tori (as the librational curves within such region are not graphs over the angles).

The Differentiable Case: Moser's Theorem

J.K. Moser, in 1962, proved a perturbation (KAM) Theorem, in the framework of area-preserving twist mappings of an annulus[35] $[0,1]\times\mathbb{S}^1$, for integrable analytic systems perturbed by a C^k perturbation, [42] and [43]. Moser's original setup corresponded to the Hamiltonian case with $d=2$ and the required smoothness was C^k with $k=333$. Later, this number was brought down to 5 by H. Rüssmann, [53].

Moser's original approach, similarly to the approach that led J. Nash to prove its theorem on the smooth embedding problem of compact Riemannian manifolds [48], is based on a smoothing technique (via convolutions), which re-introduces at each step of the Newton iteration a certain number of derivatives which one loses in the inversion of the small divisor operator.

The technique, which we shall describe here, is again due to Moser [46] but is rather different from the original one and is based on a quantitative analytic KAM Theorem (in the style of statement in Remark 7–(i) above) in conjunction with a characterization of differentiable functions in terms of functions, which are real-analytic on smaller and smaller complex strips; see [44] and, for an abstract functional approach, [65], [66]. By the way, this approach, suitably refined, leads to optimal differentiability assumptions (i. e., the Hamiltonian may be assumed to be C^ℓ with $\ell > 2d$); see, [50] and the beautiful exposition [59], which inspires the presentation reported here.

Let us consider a Hamiltonian $H = K + \varepsilon P$ (as in (17)) with K a real-analytic Kolmogorov normal form as in (10) with $\omega\in\mathcal{D}^d_{\kappa,\tau}$ and Q real-analytic; P is assumed to be a $C^\ell(\mathbb{R}^d,\mathbb{T}^d)$ function with $\ell=\ell(d,\tau)$ to be specified later[36].

Remark 11 The analytic KAM theorem, we shall refer to is the quantitative Kolmogorov Theorem as stated in Remark 7–(i) above, with (69) strengthened by including in the left hand side of (69) also[37] $\|\partial(\phi_*-\mathrm{id})\|_{\theta\xi}$ and $\|\partial(Q-Q_*)\|_{\theta\xi}$ (where "∂" denotes, here, "Jacobian" with regard to (y,x) for $(\phi_*-\mathrm{id})$ and "gradient" for $(Q-Q_*)$).

The analytic characterization of differentiable functions, suitable for our purposes, is explained in the following two lemmata[38]

Lemma 2 (Jackson, Moser, Zehnder) *Let $f\in C^l(\mathbb{R}^d)$ with $l>0$. Then, for any $\xi>0$ there exists a real-analytic function $f\colon X^d_\xi := \{x\in\mathbb{C}^d : |\mathrm{Im}\,x_j| < \xi\}\to\mathbb{C}$ such that*

$$\begin{cases} \sup\limits_{X^d_\xi} |f_\xi| \le c\|f\|_{C^0}\,, \\ \sup\limits_{X^d_{\xi'}} |f_\xi - f_{\xi'}| \le c\|f\|_{C^l}\,\xi^l\,, \quad \forall\, 0<\xi'<\xi\,, \end{cases} \tag{124}$$

where $c=c(d,l)$ is a suitable constant; if f is periodic in some variable x_j, so is f_ξ.

Lemma 3 (Bernstein, Moser) *Let $l\in\mathbb{R}_+\backslash\mathbb{Z}$ and $\xi_j := 1/2^j$. Let $f_0=0$ and let, for any $j\ge 1$, f_j be real analytic functions on $X^d_j := \{x\in\mathbb{C}^d : |\mathrm{Im}\,x_j| < 2^{-j}\}$ such that*

$$\sup_{X^d_j} |f_j - f_{j-1}| \le A 2^{-jl} \tag{125}$$

for some constant A. Then, f_j tends uniformly on $\mathbb{R}^d$ to a function $f\in C^l(\mathbb{R}^d)$ such that, for a suitable constant $C=C(d,l)>0$,

$$\|f\|_{C^l(\mathbb{R}^d)} \le CA\,. \tag{126}$$

Finally, if the f_i's are periodic in some variable x_j then so is f.

Now, denote by $X_\xi = X^d_\xi\times\mathbb{T}^d\subset\mathbb{C}^{2d}$ and define (compare Lemma 2)

$$P^j := P_{\xi_j}\,,\quad \xi_j := \frac{1}{2^j}\,. \tag{127}$$

Claim M: *If $|\varepsilon|$ is small enough and if $\ell>\sigma+1$, then there exists a sequence of Kolmogorov symplectic transformations $\{\Phi_j\}_{j\ge 0}$, $|\varepsilon|$-close to the identity, and a sequence of Kolmogorov normal forms K_j such that*

$$H_j\circ\Phi_j = K_{j+1} \quad\text{on}\quad W_{\xi_{j+1}} \tag{128}$$

where

$$\begin{aligned} &H_j := K+\varepsilon P^j \\ &\Phi_0 = \phi_0 \quad\text{and}\quad \Phi_j := \Phi_{j-1}\circ\phi_j\,,(j\ge 1) \\ &\phi_j\colon W_{\xi_{j+1}}\to W_{\alpha\xi_j}\,,\ \ \Phi_{j-1}\colon W_{\alpha\xi_j}\to X_{\xi_j}\,, \\ &\qquad j\ge 1 \ \text{and}\ \alpha := \frac{1}{\sqrt{2}}\,, \\ &\quad \sup_{x\in\mathbb{T}^d_{\xi_{j+1}}} |\Phi_j(0,x)-\Phi_{j-1}(0,x)| \le \mathrm{const}|\varepsilon| 2^{-(\ell-\sigma)j}\,. \end{aligned} \tag{129}$$

The proof of Claim **M** follows easily by induction[39] from Kolmogorov's Theorem (compare Remark 11) and Lemma 2.

From Claim **M** and Lemma 3 (applied to $f(x) = \Phi_j(0,x) - \Phi_0(0,x)$ and $l = \ell - \sigma$, which may be assumed not integer) it then follows that $\Phi_j(0,x)$ converges in the C^1 norm to a C^1 function $\phi\colon \mathbb{T}^d \to \mathbb{R}^d \times \mathbb{T}^d$, which is ε-close to the identity, and, because of (128),

$$\begin{aligned}\phi(x+\omega t) &= \lim \Phi_j(0, x+\omega t)\\ &= \lim \phi_H^t \circ \Phi_j(0,x) = \phi_H^t \circ \phi(x)\end{aligned} \tag{130}$$

showing that $\phi(\mathbb{T}^d)$ is a C^1 KAM torus for H (note that the map ϕ is close to the trivial embedding $x \to (y,x)$).

Future Directions

In this section we review in a schematic and informal way some of the most important developments, applications and possible future directions of KAM theory. For exhaustive surveys we refer to [9], Sect. 6.3 of [6] or [60].

1. *Structure of the Kolmogorov set and Whitney smoothness*
The Kolmogorov set (i. e., the union of KAM tori), in nearly-integrable systems, tends to fill up (in measure) the whole phase space as the strength of the perturbation goes to zero (compare Remark 7–(v) and Remark 10). A natural question is: what is the global geometry of KAM tori?
It turns out that KAM tori smoothly interpolate in the following sense. *For ε small enough, there exists a C^∞ symplectic diffeomorphism ϕ_* of the phase space $\mathcal{M} = B \times \mathbb{T}^d$ of the nearly-integrable, non-degenerate Hamiltonians $H = K(y) + \varepsilon P(y,x)$ and a Cantor set $C_* \subset B$ such that, for each $y' \in C_*$, the set $\phi_*^{-1}(\{y'\} \times \mathbb{T}^d)$ is a KAM torus for H*}; in other words, the Kolmogorov set is a smooth, symplectic deformation of the fiber bundle $C_* \times \mathbb{T}^d$. Still another way of describing this result is that *there exists a smooth function $K_*\colon B \to \mathbb{R}$ such that $(K + \varepsilon P) \circ \phi_*$ and K_* agree, together with their derivatives, on $C_* \times \mathbb{T}^d$*: we may, thus, say that, in general, nearly-integrable Hamiltonian systems are integrable on Cantor sets of relative big measure.
Functions defined on closed sets which admit C^k extensions are called *Whitney smooth*; compare [64], where H. Whitney gives a sufficient condition, based on Taylor uniform approximations, for a function to be Whitney C^k.
The proof of the above result – given, independently, in [50] and [19] in, respectively, the differentiable and the analytic case – follows easily from the following lemma[40]:

Lemma 4 *Let $C \subset \mathbb{R}^d$ a closed set and let $\{f_j\}$, $f_0 = 0$, be a sequence of functions analytic on $W_j := \cup_{y \in C} D(y, r_j)$. Assume that $\sum_{j \ge 1} \sup_{W_j} |f_j - f_{j-1}| r_j^{-k} < \infty$. Then, f_j converges uniformly to a function f, which is C^k in the sense of Whitney on C.*

Actually, the dependence upon the angles x' of ϕ_* is analytic and it is only the dependence upon $y' \in C_*$ which is Whitney smooth ("anisotropic differentiability", compare Sect. 2 in [50]).
For more information and a systematic use of Whitney differentiability, see [9].

2. *Power series expansions*
KAM tori $\mathcal{T}_{\omega,\varepsilon} = \phi_\varepsilon(\mathbb{T}^d)$ of nearly-integrable Hamiltonians correspond to quasi-periodic trajectories $z(t;\theta,\varepsilon) = \phi_\varepsilon(\theta + \omega t) = \phi_H^t(z(0;\theta,\varepsilon))$; compare items (d) and (e) of Sect. "Introduction" and Remark 2–(i) above. While the *actual* existence of such quasi-periodic motions was proven, for the first time, only thanks to KAM theory, the *formal* existence, in terms of formal ε-power series[41] was well known in the nineteenth century to mathematicians and astronomers (such as Newcombe, Lindstedt and, especially, Poincaré; compare [49], vol. II). Indeed, formal power solutions of nearly-integrable Hamiltonian equations are not difficult to construct (see, e. g., Sect. 7.1 of [12]) but *direct proofs* of the convergence of the series, i. e., proofs not based on Moser's "indirect" argument recalled in Remark 7–(iii) but, rather, based upon direct estimates on the kth ε-expansion coefficient, are quite difficult and were carried out only in the late eighties by H. Eliasson [27]. The difficulty is due to the fact that, in order to prove the convergence of the Taylor–Fourier expansion of such series, one has to recognize compensations among huge terms with different signs[42]. After Eliasson's breakthrough based upon a semi-direct method (compare the "Postscript 1996" at p. 33 of [27]), fully direct proofs were published in 1994 in [30] and [18].

3. *Non-degeneracy assumptions*
Kolmogorov's non-degeneracy assumption (70) can be generalized in various ways. First of all, Arnold pointed out in [2] that the condition

$$\det \begin{pmatrix} K_{yy} & K_y \\ K_y & 0 \end{pmatrix} \neq 0\,, \tag{131}$$

(this is a $(d+1) \times (d+1)$ symmetric matrix where last column and last row are given by the $(d+1)$-vector $(K_y, 0)$) which is independent from condition (70), is also sufficient to construct KAM tori. Indeed, (131) may be used to construct *iso-energetic* KAM tori, i. e., tori on a *fixed energy level*[43] E.

More recently, Rüssmann [57] (see, also, [58]), using results of Diophantine approximations on manifolds due to Pyartly [52], formulated the following condition (the "Rüssmann non-degeneracy condition"), which is essentially necessary and sufficient for the existence of a positive measure set of KAM tori in nearly-integrable Hamiltonian systems: *the image* $\omega(B) \subset \mathbb{R}^d$ *of the unperturbed frequency map* $y \to \omega(y) := K_y(y)$ *does not lie in any hyperplane passing through the origin*. We simply add that one of the prices that one has to pay to obtain these beautiful general results is that one cannot fix the frequency ahead of time.

For a thorough discussion of this topic, see Sect. 2 of [60].

4. *Some physical applications*

We now mention a short (and non-exhaustive) list of important physical application of KAM theory. For more information, see Sect. 6.3.9 of [6] and references therein.

4.1. *Perturbation of classical integrable systems*

As mentioned above (Remark 1–(iii)), one of the main original motivations of KAM theory is the perturbation theory for nearly-integrable Hamiltonian systems. Among the most famous classical integrable systems we recall: one-degree-of freedom systems; Keplerian two-body problem, geodesic motion on ellipsoids; rotations of a heavy rigid body with a fixed point (for special values of the parameters: Euler's, Lagrange's, Kovalevskaya's and Goryachev–Chaplygin's cases); Calogero–Moser's system of particles; see, Sect. 5 of [6] and [47].

A first step, in order to apply KAM theory to such classical systems, is to explicitly construct action-angle variables and to determine their analyticity properties, which is in itself a technically non-trivial problem. A second problem which arises, especially in Celestial Mechanics, is that the integrable (transformed) Hamiltonian governing the system may be highly degenerate (*proper degeneracies* – see Sect. 6.3.3, B of [6]), as is the important case of the planetary n-body problem. Indeed, the first complete proof of the existence of a positive measure set of invariant tori[44] for the planetary $(n+1)$ problem (one body with mass 1 and n bodies with masses smaller than ε) has been published only in 2004 [29]. For recent reviews on this topic, see [16].

4.2. *Topological trapping in low dimensions*

The general 2-degree-of-freedom nearly-integrable Hamiltonian exhibits a kind of particularly strong stability: the phase space is 4-dimensional and the energy levels are 3-dimensional; thus KAM tori (which are two-dimensional and which are guaranteed, under condition (131), by the iso-energetic KAM theorem) *separate* the energy levels and orbits lying between two KAM tori will remain forever trapped in the invariant region. In particular the evolution of the action variables stays forever close to the initial position ("total stability").

This observation is originally due to Arnold [2]; for recent applications to the stability of three-body problems in celestial mechanics see [13] and item **4.4** below.

In higher dimension this topological trapping is no longer available, and in principle nearby any point in phase space it may pass an orbit whose action variables undergo a displacement of order one ("Arnold's diffusion"). A rigorous complete proof of this conjecture is still missing[45].

4.3. *Spectral Theory of Schrödinger operators*

KAM methods have been applied also very successfully to the spectral analysis of the one-dimensional Schrödinger (or "Sturm–Liouville") operator on the real line $\mathbb{R}$

$$L := -\frac{d^2}{dt^2} + v(t) , \quad t \in \mathbb{R} . \tag{132}$$

If the "potential" v is bounded then there exists a unique self-adjoint operator on the real Hilbert space $\mathcal{L}^2(\mathbb{R})$ (the space of Lebesgue square-integrable functions on $\mathbb{R}$) which extends L above on C_0^2 (the space of twice differentiable functions with compact support). The problem is then to study the spectrum $\sigma(L)$ of L; for generalities, see [23].

If v is periodic, then $\sigma(L)$ is a continuous band spectrum, as it follows immediately from Floquet theory [23]. Much more complicated is the situation for quasi-periodic potentials $v(t) := V(\omega t) = V(\omega_1 t, \dots, \omega_n t)$, where V is a (say) real-analytic function on $\mathbb{T}^n$, since small-divisor problems appear and the spectrum can be nowhere dense. For a beautiful classical exposition, see [47], where, in particular, interesting connections with mechanics are discussed[46]; for deep developments of generalization of Floquet theory to quasi-periodic Schrödinger operators ("reducibility"), see [26] and [7].

4.4. *Physical stability estimates and break-down thresholds*

KAM Theory is perturbative and works if the parameter ε measuring the strength of the perturbation is small enough. It is therefore a fundamental question: *how small ε has to be in order for KAM results to hold*. The first concrete applications were extremely discouraging: in 1966, the French astronomer M. Hénon [32] pointed out that Moser's theorem applied to the restricted three-body problem (i. e., the motion of an asteroid under the gravitational influence of two unperturbed primary bodies revolving on a given Keplerian ellipse) yields existence of invariant tori if the mass ratio of the primaries is less than[47] 10^{-52}. Since then, much progress has been made and very recently, in [13], it has been shown via a computer-assisted proof[48], that, for a restricted-three body model of a subsystem of the Solar system (namely, Sun, Jupiter and Asteroid Victoria), KAM tori exist for the "actual" physical values (in that model the Jupiter/Sun mass ratio is about 10^{-3}) and, in this mathematical model – thanks to the trapping mechanism described in item **4.2** above – they trap the actual motion of the subsystem.

From a more theoretical point of view, we notice that, (compare Remark 2–(ii)) KAM tori (with a fixed Diophantine frequency) are analytic in ε; on the other hand, it is known, at least in lower dimensional settings (such as twist maps), that above a certain critical value KAM tori (curves) cannot exist ([39]). Therefore, there must exist a critical value $\varepsilon_c(\omega)$ ("breakdown threshold") such that, for $0 \le \varepsilon < \varepsilon_c\omega$, the KAM torus (curve) $\mathcal{T}_{\omega,\varepsilon}$ exists, while for $\varepsilon > \varepsilon_c(\omega)$ does not. The mathematical mechanism for the breakdown of KAM tori is far from being understood; for a brief review and references on this topic, see, e. g., Sect. 1.4 in [13].

5. *Lower dimensional tori*

In this item we consider (very) briefly, the existence of quasi-periodic solutions with a number of frequencies smaller than the number of degrees of freedom[49]. Such solutions span *lower dimensional* (non Lagrangian) *tori*. Certainly, this is one of the most important topics in modern KAM theory, not only in view of applications to classical problems, but especially in view of extensions to infinite dimensional systems, namely PDEs (Partial Differential Equations) with a Hamiltonian structure; see, item **6** below. For a recent, exhaustive review on lower dimensional tori (in finite dimensions), we refer the reader to [60].

In 1965 V.K. Melnikov [41] stated a precise result concerning the persistence of *stable* (or "elliptic") lower dimensional tori; the hypotheses of such results are, now, commonly referred to as "Melnikov conditions". However, a proof of Melnikov's statement was given only later by Moser [45] for the case $n = d - 1$ and, in the general case, by H. Eliasson in [25] and, independently, by S.B. Kuksin [37]. The *unstable* ("partially hyperbolic") case (i. e., the case for which the lower dimensional tori are linearly unstable and lie in the intersection of stable and unstable Lagrangian manifolds) is simpler and a complete perturbation theory was already given in [45], [31] and [66] (roughly speaking, the normal frequencies to the torus do not resonate with the inner (or "proper") frequencies associated with quasi-periodic motion). Since then, Melnikov conditions have been significantly weakened and much technical progress has been made; see [60], Sects. 5, 6 and 7, and references therein.

To illustrate a typical situation, let us consider a Hamiltonian system with $d = n + m$ degrees of freedom, governed by a Hamiltonian function of the form

$$H(y, x, v, u; \xi) = K(y, v, u; \xi) + \varepsilon P(y, x, v, u; \xi), \quad (133)$$

where $(y, x) \in \mathbb{T}^n \times \mathbb{R}^n$, $(v, u) \in \mathbb{R}^{2m}$ are pairs of standard symplectic coordinates and ξ is a real parameter running over a compact set $\Pi \subset \mathbb{R}^n$ of positive Lebesgue measure[50]; K is a Hamiltonian admitting the n-torus

$$\mathcal{T}_0^n(\xi) := \{y = 0\} \times \mathbb{T}^n \times \{v = u = 0\}\,, \quad \xi \in \Pi\,,$$

as invariant linearly stable invariant torus and is assumed to be in the normal form:

$$K = E(\xi) + \omega(\xi) \cdot y + \frac{1}{2} \sum_{j=1}^{m} \Omega_j(\xi)(u_j^2 + v_j^2)\,. \quad (134)$$

The ϕ_K^t flow decouples in the linear flow $x \in \mathbb{T}^n \to x + \omega(\xi)t$ times the motion of m (decoupled) harmonic oscillators with characteristic frequencies $\Omega_j(\xi)$ (sometimes referred to as *normal frequencies*). Melnikov's conditions (in the form proposed in [51]) reads as follows: assume that ω is a Lipschitz homeomorphism; let $\Pi_{k,l}$ denote the "resonant parameter set" $\{\xi \in \Pi : \omega(\xi) \cdot k + \Omega \cdot (\xi) = 0\}$ and assume

$$\begin{cases} \Omega_i(\xi) > 0\,, \quad \Omega_i(\xi) \neq \Omega_j(\xi)\,, \quad \forall \xi \in \Pi\,, \forall i \neq j \\ \text{meas } \Pi_{k,l} = 0\,, \quad \forall k \in \mathbb{Z}^n \backslash \{0\}\,, \quad \forall l \in \mathbb{Z}^m : |l| \le 2\,. \end{cases} \quad (135)$$

Under these assumptions and if $|\varepsilon|$ is small enough, there exists a (Cantor) subset of parameters $\Pi_* \subset \Pi$ of posi-

tive Lebesgue measure such that, to each $\xi \in \Pi_*$, there corresponds a n-dimensional, linearly stable H-invariant torus $\mathcal{T}^n_\varepsilon(\xi)$ on which the H flow is analytically conjugated to $x \to x + \omega_*(\xi)t$ where ω_* is a Lipschitz homeomorphism of Π_* assuming Diophantine values and close to ω.

This formulation has been borrowed from [51], to which we refer for the proof; for the differentiable analog, see [22].

Remark 12 The small-divisor problems arising in the perturbation theory of the above lower dimensional tori are of the form

$$\omega \cdot k - l \cdot \Omega\,, \quad |l| \le 2\,, \quad |k| + |l| \ne 0\,, \tag{136}$$

where one has to regard the normal frequency Ω as *functions of the inner frequencies* ω and, at first sight, one has – in J. Moser words – a *lack-of-parameter* problem. To overcome this intrinsic difficulty, one has to give up full control of the inner frequencies and construct, iteratively, n-dimensional sets (corresponding to smaller and smaller sets of ξ-parameters) on which the small divisors are controlled; for more motivations and informal explanations on lower dimensional small divisor problems, see, Sects. 5, 6 and 7 of [60].

6. *Infinite dimensional systems*

As mentioned above, the most important recent developments of KAM theory, besides the full applications to classical n-body problems mentioned above, is the successful extension to infinite dimensional settings, so as to deal with certain classes of partial differential equations carrying a Hamiltonian structure. As a typical example, we mention the non-linear wave equation of the form

$$\begin{aligned} &u_{tt} - u_{xx} + V(x)u = f(u)\,,\\ &\qquad f(u) = O(u^2)\,,\quad 0 < x < 1\,,\quad t \in \mathbb{R}\,. \end{aligned} \tag{137}$$

These extensions allowed, in the pioneering paper [63], establishing the existence of small-amplitude quasi-periodic solutions for (137), subject to Dirichlet or Neumann boundary conditions (on a finite interval for odd and analytic nonlinearities f); the technically more difficult periodic boundary condition case was considered later; compare [38] and references therein.

A technical discussion of these topics goes far beyond the scope of the present article and, for different equations, techniques and details, we refer the reader to the review article [38].

A The Classical Implicit Function Theorem

Here we discuss the classical Implicit Function Theorem for complex functions from a quantitative point of view. The following Theorem is a simple consequence of the Contraction Lemma, which asserts that a contraction Φ on a closed, non-empty metric space[51] X has a unique fixed point, which is obtained as $\lim_{j\to\infty} \Phi^j(u_0)$ for any[52] $u_0 \in X$. As above, $D^n(y_0, r)$ denotes the ball in $\mathbb{C}^n$ of center y_0 and radius r.

Theorem 3 (Implicit Function Theorem) *Let*

$$\begin{aligned} F\colon (y,x) \in D^n(y_0,r) \times D^m(x_0,s) \subset \mathbb{C}^{n+m}\\ \to F(y,x) \in \mathbb{C}^n \end{aligned}$$

be continuous with continuous Jacobian matrix F_y; assume that $F_y(y_0,x_0)$ is invertible and denote by T its inverse; assume also that

$$\begin{aligned} \sup_{D(y_0,r)\times D(x_0,s)} \|\mathbb{1}_n - TF_y(y,x)\| \le \frac{1}{2}\,,\\ \sup_{D(x_0,s)} |F(y_0,x)| \le \frac{r}{2\|T\|}\,. \end{aligned} \tag{138}$$

Then, all solutions $(y,x) \in D(y_0,r)\times D(x_0,s)$ of $F(y,x) = 0$ are given by the graph of a unique continuous function $g\colon D(x_0,s) \to D(y_0,r)$ satisfying, in particular,

$$\sup_{D(x_0,s)} |g| \le 2\|T\| \sup_{D(x_0,s)} |F(y_0,\cdot)|\,. \tag{139}$$

Proof Let $X = C(D^m(x_0,s), D^n(y_0,r))$ be the closed ball of continuous function from $D^m(x_0,s)$ to $D^n(y_0,r)$ with respect to the sup-norm $\|\cdot\|$ (X is a non-empty metric space with distance $d(u,v) := \|u - v\|$) and denote $\Phi(y;x) := y - TF(y,x)$. Then, $u \to \Phi(u) := \Phi(u,\cdot)$ maps $C(D^m(x_0,s))$ into $C(\mathbb{C}^m)$ and, since $\partial_y\Phi = \mathbb{1}_n - TF_y(y,x)$, from the first relation in (138), it follows that $u \to \Phi(u)$ is a contraction. Furthermore, for any $u \in C(D^m(x_0,s), D^n(y_0,r))$,

$$\begin{aligned} |\Phi(u) - y_0| &\le |\Phi(u) - \Phi(y_0)| + |\Phi(y_0) - y_0|\\ &\le \frac{1}{2}\|u - y_0\| + \|T\|\|F(y_0,x)\|\\ &\le \frac{1}{2}r + \|T\|\frac{r}{2\|T\|} = r\,, \end{aligned}$$

showing that $\Phi\colon X \to X$. Thus, by the Contraction Lemma, there exists a unique $g \in X$ such that $\Phi(g) = g$, which is equivalent to $F(g,x) = 0\ \forall x$. If $F(y_1,x_1) = 0$ for some $(y_1,x_1) \in D(y_0,r) \times D(x_0,s)$, it follows that $|y_1 - g(x_1)| = |\Phi(y_1;x_1) - \Phi(g(x_1),x_1)| \le \alpha|y_1 - g(x_1)|$, which implies that $y_1 = g(x_1)$ and that all solutions of $F = 0$ in $D(y_0, r) \times D(x_0, s)$ coincide with the graph og g. Finally, (139) follows by observing that $\|g - y_0\| = \|\Phi(g) - y_0\| \le \|\Phi(g) - \Phi(y_0)\| + \|\Phi(y_0) - y_0\| \le \frac{1}{2}\|g - y_0\| + \|T\|\|F(y_0, \cdot)\|$, finishing the proof. □

Additions:

(i) If F is periodic in x or/and real on reals, then (by uniqueness) so is g;
(ii) If F is analytic, then so is g (Weierstrass Theorem, since g is attained as uniform limit of analytic functions);
(iii) The factors 1/2 appearing in the right-hand sides of (138) may be replaced by, respectively, α and β for any positive α and β such that $\alpha + \beta = 1$.

Taking $n = m$ and $F(y, x) = f(y) - x$ for a given $C^1(D(y_0, r), \mathbb{C}^n)$ function, one obtains the

Theorem 4 (Inverse Function Theorem) *Let $f\colon y \in D^n(y_0, r) \to \mathbb{C}^n$ be a C^1 function with invertible Jacobian $f_y(y_0)$ and assume that*

$$\sup_{D(y_0,r)} \|\mathbb{1}_n - T f_y\| \le \frac{1}{2}, \quad T := f_y(y_0)^{-1}, \tag{140}$$

then there exists a unique C^1 function $g\colon D(x_0, s) \to D(y_0, r)$ with $x_0 := f(y_0)$ and $s := r/(2\|T\|)$ such that $f \circ g(x) = \mathrm{id} = g \circ f$.

Additions analogous to the above also hold in this case.

B Complementary Notes

1 Actually, the first instance of a small divisor problem solved analytically is the linearization of the germs of analytic functions and is due to C.L. Siegel [61].

2 The well-known Newton's tangent scheme is an algorithm, which allows us to find roots (zeros) of a smooth function f in a region where the derivative f' is bounded away from zero. More precisely, if x_n is an "approximate solution" of $f(x) = 0$, i. e., $f(x_n) := \varepsilon_n$ is small, then the next approximation provided by Newton's tangent scheme is $x_{n+1} := x_n - f(x_n)/f'(x_n)$ [which is the intersection with x-axis of the tangent to the graph of f passing through $(x_n, f(x_n))$] and, in view of the definition of ε_n and Taylor's formula, one has that $\varepsilon_{n+1} := f(x_{n+1}) = \frac{1}{2} f''(\xi_n){\varepsilon_n}^2/(f'(x_n))^2$ (for a suitable ξ_n) so that $\varepsilon_{n+1} = O(\varepsilon_n^2) = O(\varepsilon_1^{2^n})$ and, in the iteration, x_n will converge (at a super-exponential rate) to a root $\bar{x}$ of f. This type of extremely fast convergence will be typical in the analyzes considered in the present article.

3 The elements of $\mathbb{T}^d$ are equivalence classes $x = \bar{x} + 2\pi\mathbb{Z}^d$ with $\bar{x} \in \mathbb{R}^d$. If $x = \bar{x} + 2\pi\mathbb{Z}^d$ and $y = \bar{y} + 2\pi\mathbb{Z}^d$ are elements of $\mathbb{T}^d$, then their distance $d(x, y)$ is given by $\min_{n\in\mathbb{Z}^d} |\bar{x} - \bar{y} + 2\pi n|$ where $|\cdot|$ denotes the standard euclidean norm in $\mathbb{R}^n$; a smooth (analytic) function on $\mathbb{T}^d$ may be viewed as ("identified with") a smooth (analytic) function on $\mathbb{R}^d$ with period 2π in each variable. The torus $\mathbb{T}^d$ endowed with the above metric is a real-analytic, compact manifold. For more information, see [62].

4 A symplectic form on an (even dimensional) manifold is a closed, non-degenerate differential 2-form. The symplectic form $\alpha = \mathrm{d}y \wedge \mathrm{d}x$ is actually *exact* symplectic, meaning that $\alpha = \mathrm{d}(\sum_{i=1} y_i \mathrm{d}x_i)$. For general information see [5].

5 For general facts about the theory of ODE (such as Picard theorem, smooth dependence upon initial data, existence times, ...) see, e. g., [23].

6 This terminology is due to that fact the the x_j are "adimensional" angles, while analyzing the physical dimensions of the quantities appearing in Hamilton's equations one sees that $\dim(y) \times \dim(x) = \dim H \times \dim(t)$ so that y has the dimension of an energy (the Hamiltonian) times the dimension of time, i. e., by definition, the dimension of an action.

7 This terminology is due to the fact that a classical mechanical system of d particles of masses $m_i > 0$ and subject to a potential $V(q)$ with $q \in A \subset \mathbb{R}^d$ is governed by a Hamiltonian of the form $\sum_{j=1}^d p_j^2/2m_j + V(q)$ and d may be interpreted as the (minimal) number of coordinates necessary to physically describe the system.

8 To be precise, (6) should be written as $y(t) = v(\pi_{\mathbb{T}^d}(\omega t))$, $x(t) = \pi_{\mathbb{T}^d}(\omega t + u(\pi_{\mathbb{T}^d}(\omega t)))$ where $\pi_{\mathbb{T}^d}$ denotes the standard projection of $\mathbb{R}^d$ onto $\mathbb{T}^d$, however we normally omit such a projection.

9 As standard, U_θ denotes the $(d \times d)$ Jacobian matrix with entries $(\partial U_i)/(\partial\theta_j) = \delta_{ij} + (\partial u_i)/(\partial\theta_j)$.

10 For generalities, see [5]; in particular, a Lagrangian manifold $L \subset \mathcal{M}$ which is a graph over $\mathbb{T}^d$ admits a "generating function", i. e., there exists a smooth function $g\colon \mathbb{T}^d \to \mathbb{R}$ such that $L = \{(y, x)\colon y = g_x(x), x \in \mathbb{T}^d\}$.

11 Compare [54] and references therein. We remark that, if $B(\omega_0, r)$ denote the ball in $\mathbb{R}^d$ of radius r centered at ω_0 and fix $\tau > d - 1$, then one can prove that the Lebesgue measure of $B(y_0, r)\backslash\mathcal{D}^d_{\kappa,\tau}$ can be bounded by $c_d \kappa r^{d-1}$ for a suitable constant c_d depending only on d; for a simple proof, see, e.g, [21].

12 The sentence "can be put into the form" means "there exists a symplectic diffeomorphism $\phi\colon (y, x) \in \mathcal{M} \to$

$(\eta, \xi) \in \mathcal{M}$ such that $H \circ \phi$ has the form (10)"; for multi-indices α, $|\alpha| = \alpha_1 + \cdots + \alpha_d$ and $\partial_y^\alpha = \partial_{y_1}^{\alpha_1} \cdots \partial_{y_d}^{\alpha_d}$; the vanishing of the derivatives of a function $f(y)$ up to order k in the origin will also be indicated through the expression $f = O(|y|^{k+1})$.

13 **Notation:** If A is an open set and $p \in \mathbb{N}$, then the C^p-norm of a function $f\colon x \in A \to f(x)$ is defined as $\|f\|_{C^p(A)}\colon \sup_{|\alpha| \le p} \sup_A |\partial_x^\alpha f|$.

14 **Notation:** If f is a scalar function f_y is a d-vector; f_{yy} the Hessian matrix $(f_{y_i y_j})$; f_{yyy} the symmetric 3-tensor of third derivatives acting as follows: $f_{yyy} a \cdot b \cdot c := \sum_{i,j,k=1}^d (\partial^3 f)/(\partial y_i \partial y_j \partial y_k) a_i b_j c_k$.

15 **Notation:** If f is (a regular enough) function over $\mathbb{T}^d$, its Fourier coefficients are defined as $f_n := \int_{\mathbb{T}^d} f(x) \mathrm{e}^{-in\cdot x} \mathrm{d}x/(2\pi)^d$; where, as usual, $i = \sqrt{-1}$ denotes imaginary unit; for general information about Fourier series see, e. g., [34].

16 The choice of norms on finite dimensional spaces ($\mathbb{R}^d$, $\mathbb{C}^d$, space of matrices, tensors, etc.) is not particularly relevant for the analysis in this article (since changing norms will change d-depending constants); however for matrices, tensors (and, in general, linear operators), it is convenient to work with the "operator norm", i. e., the norm defined as $\|L\| = \sup_{u \ne 0} \|Lu\|/\|u\|$, so that $\|Lu\| \le \|L\|\|u\|$, an estimate, which will be constantly be used; for a general discussion on norms, see, e. g., [36].

17 As an example, let us work out the first two estimates, i. e., the estimates on $\|s_x\|_{\bar\xi}$ and $|b|$: actually these estimates will be given on a larger intermediate domain, namely, $W_{\xi-\delta/3}$, allowing to give the remaining bounds on the smaller domain $W_{\bar\xi}$ (recall that W_s denotes the complex domain $D(0,s) \times \mathbb{T}_s^d$). Let $f(x) := P(0,x) - \langle P(0,\cdot)\rangle$. By definition of $\|\cdot\|_\xi$ and M, it follows that $\|f\|_\xi \le \|P(0,x)\|_\xi + \|\langle P(0,\cdot)\rangle\|_\xi \le 2M$. By **(P5)** with $p = 1$ and $\xi' = \xi - \delta/3$, one gets

$$\|s_x\|_{\xi - \frac{\delta}{3}} \le \bar B_1 \frac{2M}{\kappa} 3^{k_1} \delta^{-k_1} ,$$

which is of the form (53), provided $\bar c \ge (\bar B_1 2 \cdot 3^{k_1})/\kappa$ and $\bar\nu \ge k_1$. To estimate b, we need to bound first $|Q_{yy}(0,x)|$ and $|P_y(0,x)|$ for real x. To do this we can use Cauchy estimate: by **(P4)** with $p = 2$ and, respectively, $p = 1$, and $\xi' = 0$, we get

$$\|Q_{yy}(0,\cdot)\|_0 \le m B_2 C \xi^{-2} \le m B_2 C \delta^{-2} , \quad \text{and}$$
$$\|P_y(0,x)\|_0 \le m B_1 M \delta^{-1} ,$$

where $m = m(d) \ge 1$ is a constant which depend on the choice of the norms, (recall also that $\delta < \xi$). Putting these bounds together, one gets that $|b|$ can be bounded by the r.h.s. of (53) provided $\bar c \ge m(B_2 \bar B_1 2 \cdot 3^{k_1} \kappa^{-1} + B_1)$, $\mu \ge 2$ and $\bar\nu \ge k_1 + 2$. The other bounds in (53) follow easily along the same lines.

18 We sketch here the **proof of Lemma 1**. The defining relation $\psi_\varepsilon \circ \varphi = \mathrm{id}$ implies that $\alpha(x') = -a(x' + \varepsilon\alpha(x'))$, where $\alpha(x')$ is short for $\alpha(x';\varepsilon)$ and that equation is a fixed point equation for the non-linear operator $f\colon u \to f(u) := -a(\mathrm{id} + \varepsilon u)$. To find a fixed point for this equation one can use a standard contraction Lemma (see [36]). Let Y denote the closed ball (with respect to the sup-norm) of continuous functions $u\colon \mathbb{T}_{\xi'}^d \to \mathbb{C}^d$ such that $\|u\|_{\xi'} \le \bar L$. By (54), $|\mathrm{Im}(x' + \varepsilon u(x'))| < \xi' + \varepsilon_0 \bar L < \xi' + \delta/3 = \bar\xi$, for any $u \in Y$, and any $x' \in \mathbb{T}_{\xi'}^d$; thus, $\|f(u)\|_{\xi',\varepsilon_*} \le \|a\|_{\bar\xi} \le \bar L$ by (53), so that $f\colon Y \to Y$; notice that, in particular, this means that f sends periodic functions into periodic functions. Moreover, (54) implies also that f is a contraction: if $u, v \in Y$, then, by the mean value theorem, $|f(u) - f(v)| \le \bar L |\varepsilon| |u - v|$ (with a suitable choice of norms), so that, by taking the sup-norm, one has $\|f(u) - f(v)\|_{\xi'} < \varepsilon_0 \bar L \|u - v\|_{\xi'} < \frac{1}{3}\|u - v\|_{\xi'}$ showing that f is a contraction. Thus, there exists a unique $\alpha \in Y$ such that $f(\alpha) = \alpha$. Furthermore, recalling that the fixed point is achieved as the uniform limit $\lim_{n\to\infty} f^n(0)$ $(0 \in Y)$ and since $f(0) = -a$ is analytic, so is $f^n(0)$ for any n and, hence, by Weierstrass Theorem on the uniform limit of analytic function (see [1]), the limit α itself is analytic. In conclusion, $\varphi \in \mathcal{B}_{\xi'}$ and (55) holds.

Next, for $(y', x) \in W_{\bar\xi}$, by (53), one has $|y' + \varepsilon\beta(y', x)| < \bar\xi + \varepsilon_0 \bar L < \bar\xi + \delta/3 = \xi$ so that (56) holds. Furthermore, since $\|\varepsilon a_x\|_{\bar\xi} < \varepsilon_0 \bar L < 1/3$ the matrix $\mathbb{1}_d + \varepsilon a_x$ is invertible with inverse given by the "Neumann series" $(\mathbb{1}_d + \varepsilon a_x)^{-1} = \mathbb{1}_d + \sum_{k=1}^\infty (-1)^k (\varepsilon a_x)^k =: \mathbb{1}_d + \varepsilon S(x;\varepsilon)$, so that (57) holds. The proof is finished.

19 From (59), it follows immediately that $\langle \partial_{y'}^2 Q_1(0,\cdot)\rangle = \langle \partial_y^2 Q(0,\cdot)\rangle + \varepsilon \langle \partial_{y'}^2 \widetilde{Q}(0,\cdot)\rangle = T^{-1}(\mathbb{1}_d + \varepsilon T \langle \partial_{y'}^2 \widetilde{Q}(0,\cdot)\rangle) =: T^{-1}(\mathbb{1}_d + \varepsilon R)$ and, in view of (51) and (59), we see that $\|R\| < L/(2C)$. Therefore, by (60), $\varepsilon_0 \|R\| < 1/6 < 1/2$, implying that $(1 + \varepsilon R)$ is invertible and $(\mathbb{1}_d + \varepsilon R)^{-1} = \mathbb{1}_d + \sum_{k=1}^\infty (-1)^k \varepsilon^k R^k =: 1 + \varepsilon D$ with $\|D\| \le \|R\|/(1 - |\varepsilon|\|R\|) < L/C$. In conclusion, $T_1 = (1 + \varepsilon R)^{-1} T = T + \varepsilon D T =: T + \varepsilon \widetilde{T}$, $\|\widetilde{T}\| \le \|D\| C \le (L/C) C = L$.

20 Actually, there is quite some freedom in choosing the sequence $\{\xi_j\}$ provided the convergence is not too fast; for general discussion, see, [56], or, also, [10] and [14].

21 In fact, denoting by B_* the real d-ball centered at 0 and of radius $\theta\xi_*$ for $\theta \in (0,1)$, from Cauchy estimate (47) with $\xi = \xi_*$ and $\xi' = \theta\xi_*$, one has $\|\phi_* -$

$\mathrm{id}\|_{C^p(B_*\times\mathbb{T}^d)} = \sup_{B_*\times\mathbb{T}^d}\sup_{|\alpha|+|\beta|\le p}|\partial_y^\alpha\partial_x^\beta(\phi_* - \mathrm{id})| \le \sup_{|\alpha|+|\beta|\le p}\|\partial_y^\alpha\partial_x^\beta(\phi_* - \mathrm{id})\|_{\theta\xi_*} \le B_p\|\phi_* - \mathrm{id}\|_{\xi_*}1/(\theta\xi_*)^p \le \mathrm{const}_p|\varepsilon|$ with $\mathrm{const}_p := B_p DBM1/(\theta\xi_*)^p$. An identical estimate holds for $\|Q_* - Q\|_{C^p(B_*\times\mathbb{T}^d)}$.

22 Also very recently ε-power series expansions have been shown to be a very powerful tool; compare [13].

23 A function $f\colon A\subset\mathbb{R}^n\to\mathbb{R}^n$ is Lipschitz on A if there exists a constant ("Lipschitz constant") $L>0$ such that $|f(x)-f(y)|\le L|x-y|$ for all $x,y\in A$. For a general discussion on how Lebesgue measure changes under Lipschitz mappings, see, e. g., [28]. In fact, the dependence of ϕ_* on $\bar{y}$ is much more regular, compare Remark 11.

24 In fact, notice that inverse powers of κ appear through (48) (inversion of the operator D_ω), therefore one sees that the terms in the first line of (53) may be replaced by $\tilde{c}\kappa^{-2}$ (in defining a one has to apply the operator D_ω^{-1} twice) but then in $P^{(1)}$ (see (26)) there appears $\|\beta\|^2$, so that the constant c in the second line of (53) has the form (72); since $\kappa<1$, one can replace in (53) c with $\hat{c}\kappa^{-4}$ as claimed.

25 **Proof of Claim C** Let $H_0 := H$, $E_0 := E$, $Q_0 := Q$, $K_0 := K$, $P_0 := P$, $\xi_0 := \xi$ and let us assume (*inductive hypothesis*) that we can iterate the Kolmogorov transformation j times obtaining j symplectic transformations $\phi_{i+1}\colon W_{\xi_{i+1}}\to W_{\xi_i}$, for $0\le i\le j-1$, and j Hamiltonians $H_{i+1} = H_i\circ\phi_{i+1} = K_i + \varepsilon^{2^i}P_i$ real-analytic on W_{ξ_i} such that

$$\begin{aligned}&|\omega|,|E_i|,\|Q_i\|_{\xi_i},\|T_i\|<C\,,\\&|\varepsilon|^{2^i}L_i := |\varepsilon|^{2^i}cC^\mu\delta_0^{-\nu}2^{\nu i}M_i\le\frac{\delta_i}{3}\,,\\&\qquad\forall\,0\le i\le j-1\,.\end{aligned}\qquad(*)$$

By $(*)$, Kolmogorov iteration (**Step 2**) can be applied to H_i and therefore all the bounds described in paragraph **Step 2** hold (having replaced $H,E,\dots,\xi,\delta,H',E',\dots,\xi'$ with, respectively, $H_i,E_i,\dots,\xi_i,\delta_i,H_{i+1},E_{i+1},\dots,\xi_{i+1}$); in particular (see (61)) one has, for $0\le i\le j-1$ (and for any $|\varepsilon|\le\varepsilon_0$),

$$\begin{cases}|E_{i+1}|\le|E_i|+|\varepsilon|^{2^i}L_i\,,\\ \quad\|Q_{i+1}\|_{\xi_{i+1}}\le\|Q_i\|_{\xi_i}+|\varepsilon|^{2^i}L_i\,,\\ \quad\|\phi_{i+1}-\mathrm{id}\|_{\xi_{i+1}}\le|\varepsilon|^{2^i}L_i\\ M_{i+1}\le M_iL_i\end{cases}\qquad(C.1)$$

Observe that the definition of D, B and L_I, $|\varepsilon|^{2^j}L_j(3C\delta_j^{-1}) =: DB^j|\varepsilon|^{2^j}M_j$, so that $L_i<DB^iM_i$, thus by the second line in $(C.1)$, for any $0\le i\le j-1$, $|\varepsilon|^{2^{i+1}}M_{i+1} < DB_i(M_i|\varepsilon|^{2^i})^2$, which iterated, yields (66) for $0\le i\le j$. Next, we show that, thanks to (65), $(*)$ holds also for $i=j$ (and this means that Kolmogorov's step can be iterated an infinite number of times). In fact, by $(*)$ and the definition of C in (64): $|E_j|\le|E|+\sum_{i=0}^{j-1}\varepsilon_0^{2^i}L_i\le|E|+\frac{1}{3}\sum_{i\ge0}\delta_i < |E|+\frac{1}{6}\sum_{i\ge1}2^{-i}<|E|+1<C$. The bounds for $\|Q_i\|$ and $\|T_i\|$ are proven in an identical manner. Now, by $(66)_{i=j}$ and (65), $|\varepsilon|^{2^j}L_j(3\delta_j^{-1}) = DB^j|\varepsilon|^{2^j}M_j\le DB^j(DB\varepsilon_0M)^{2^j}/(DB^{j+1})\le 1/B<1$, which implies the second inequality in $(*)$ with $i=j$; the proof of the induction is finished and one can construct an *infinite sequence* of Kolmogorov transformations satisfying $(*)$, $(C.1)$ and (66) *for all* $i\ge0$. To check (67), we observe that $|\varepsilon|^{2^i}L_i = \delta_0/(3\cdot2^i)DB^i|\varepsilon|^{2^i}M_i\le(1/2^{i+1})(|\varepsilon|DBM)^{2^i}\le(|\varepsilon|DBM/2)^{i+1}$ and therefore $\sum_{i\ge0}|\varepsilon|^{2^i}L_i\le\sum_{i\ge1}(|\varepsilon|DBM/2)^i\le|\varepsilon|DBM$. Thus, $\|Q-Q_*\|_{\xi_*}\le\sum_{i\ge0}\|\tilde{Q}_i\|_{\xi_i}\le|\varepsilon|^{2^i}L_i\le|\varepsilon|DBM$; and analogously for $|E-E_*|$ and $\|T-T_*\|$. To estimate $\|\phi_*-\mathrm{id}\|_{\xi_*}$, observe that $\|\Phi_i-\mathrm{id}\|_{\xi_i}\le\|\Phi_{i-1}\circ\phi_i-\phi_i\|_{\xi_i}+\|\phi_i-\mathrm{id}\|_{\xi_i}\le\|\Phi_{i-1}-\mathrm{id}\|_{\xi_{i-1}}+|\varepsilon|^{2^i}L_i$, which iterated yields $\|\Phi_i-\mathrm{id}\|_{\xi_i}\le\sum_{k=0}^i|\varepsilon|^{2^k}L_k\le|\varepsilon|DBM$: taking the limit over i completes the proof of (67) and the proof of Claim **C**.

26 In fact, observe: (i) given any integer vector $0\ne n\in\mathbb{Z}^d$ with $d\ge2$, one can find $0\ne m\in\mathbb{Z}^d$ such $n\cdot m=0$; (ii) the set $\{tn\colon t>0 \text{ and } n\in\mathbb{Z}^d\}$ is dense in $\mathbb{R}^d$; (iii) if U is a neighborhood of y_0, then $K_y(U)$ is a neighborhood of $\omega=K_y(y_0)$. Thus, by (ii) and (iii), in $K_y(U)$ there are infinitely many points of the form tn with $t>0$ and $n\in\mathbb{Z}^d$ to which correspond points $y(t,n)\in U$ such that $K_y(y(t,n)) = tn$ and for any of such points one can find, by (i), $m\in\mathbb{Z}$ such that $m\cdot n=0$, whence $K_y(y(t,n))\cdot m = tn\cdot m=0$.

27 This fact was well known to Poincaré, who based on the above argument his non-existence proof of integral of motions in the general situation; compare Sect. 7.1.1, [6].

28 Compare (90) but observe, that, since $\hat{P}$ is a trigonometric polynomial, in view of Remark 9–(ii), g in (96) defines a real-analytic function on $D(y_0,\bar{r})\times\mathbb{T}^d_{\xi'}$ with a suitable $\bar{r}=\bar{r}(\varepsilon)$ and $\xi'<\xi$. Clearly it is important to see explicitly how the various quantities depend upon ε; this is shortly discussed after Proposition 2.

29 In fact: $\|\check{P}\|_{r,\xi-\delta/2}\le M\sum_{|n|>N}\mathrm{e}^{-|n|\delta/2}\le M\mathrm{e}^{-(\delta/4)N}\sum_{|n|>N}\mathrm{e}^{-|n|\delta/4}\le M\mathrm{e}^{-(\delta/4)N}\sum_{|n|>0}\mathrm{e}^{-|n|\delta/4}\le\mathrm{const}\,M\mathrm{e}^{-(\delta/4)N}\delta^{-d}\le|\varepsilon|M$ if (106) holds and N is taken as in (104).

30 Apply the IFT of Appendix "A The Classical Implicit Function Theorem" to $F(y,\eta) := K_y(y) + \eta\partial_y P_0(y) - K_y(y_0)$ defined on $D^d(y_0,\bar r)\times D^1(0,|\varepsilon|)$: using the mean value theorem, Cauchy estimates and (114), $\|\mathbb{1}_d - TF_y\| \le \|\mathbb{1}_d - TK_{yy}\| + |\varepsilon|\|\partial_y^2 P_0\| \le \|T\|\|K_{yyy}\|\bar r + \|T\||\varepsilon|\|\partial_y^2 P_0\| \le C^2 2\bar r/r + C|\varepsilon|4/r^2 M \le \frac14 + \frac18 < \frac12$; also: $2\|T\|\|F(y_0,\eta)\| = 2\|T\||\eta|\partial_y P_0(y_0)\| < 2C|\varepsilon|M2/r \le 2CM\bar r^{-1}|\varepsilon| < \frac14\bar r$ (where last inequality is due to (114)), showing that conditions (138) are fulfilled. Equation (111) comes from (139) and (113) follows easily by repeating the above estimates.

31 Recall note 18 and notice that $(\mathbb{1}_d + A)^{-1} = \mathbb{1}_d + D$ with $\|D\| \le \|A\|/(1-\|A\|) \le 2\|A\| \le 20C^3M|\varepsilon|$, where last two inequalities are due to (113).

32 Lemma 1 can be immediately extended to the y'-dependent case (which appear as a dummy parameter) as far as the estimates are uniform in y' (which is the case).

33 By (118) and (54), $|\varepsilon|\|g_x\|_{\bar r,\bar\xi} \le |\varepsilon|rL \le r/2$ so that, by (116), if $y'\in D_{\bar r/2}(y_1)$, then $y'+\varepsilon g_x(y',\varphi(y',x'))\in D_r(y_0)$.

34 The first requirement in (123) is equivalent to require that $r_0\le r$, which implies that if $\bar r$ is *defined* as the r.h.s. of (108), then $\bar r\le r/2$ as required in (110). Next, the first requirement in (114) at the $(j+1)$th step of the iteration translates into $16C^2 r_{j+1}/r_j \le 1$, which is satisfied, since, by definition, $r_{j+1}/r_j = (1/(2\gamma))^{\tau+1} \le (1/(2\gamma))^2 = 1/(36C^2) < 1/(16C^2)$. The second condition in (114), which at the $(j+1)$th step, reads $2CM_j r_{j+1}^{-2}|\varepsilon|^{2^j}$ is implied by $|\varepsilon|^{2^j}L_j \le \delta_j/(3C)$ (corresponding to (54)), which, in turn, is easily controlled along the lines explained in note 25.

35 An area-preserving twist mappings of an annulus $A = [0,1]\times\mathbb{S}^1$, $(\mathbb{S}^1 = \mathbb{T}^1)$, is a symplectic diffeomorphism $f=(f_1,f_2)\colon (y,x)\in A \to f(y,x)\in A$, leaving invariant the boundary circles of A and satisfying the twist condition $\partial_y f_2 > 0$ (i. e., f twists clockwise radial segments). The theory of area preserving maps, which was started by Poincaré (who introduced such maps as section of the dynamics of Hamiltonian systems with two degrees of freedom), is, in a sense, the simplest nontrivial Hamiltonian context. After Poincaré the theory of area-preserving maps became, in itself, a very rich and interesting field of Dynamical Systems leading to very deep and important results due to Herman, Yoccoz, Aubry, Mather, etc; for generalities and references, see, e. g., [33].

36 It is not necessary to assume that K is real-analytic, but it simplify a little bit the exposition. In our case, we shall see that ℓ is related to the number σ in (66). We recall the definition of Hölder norms: If $\ell = \ell_0 + \mu$ with $\ell_0\in\mathbb{Z}_+$ and $\mu\in(0,1)$, then $\|f\|_{C^\ell} := \|f\|_{C^\ell} + \sup_{|\alpha|=\ell_0}\sup_{0<|x-y|<1}|\partial^\alpha f(x)-\partial^\alpha f(y)|/|x-y|^\mu$; $C^\ell(\mathbb{R}^d)$ denotes the Banach space of functions with finite C^ℓ norm.

37 To obtain these new estimates, one can, first replace θ by $\sqrt{\theta}$ and then use the remark in the note 21 with $p=1$; clearly the constant σ has to be increased by one unit with respect to the constant σ appearing in (69).

38 For general references and discussions about Lemma 2 and 3, see, [44] and [65]; an elementary detailed proof can be found, also, in [15].

39 **Proof of Claim M** The first step of the induction consists in constructing $\Phi_0 = \phi_0$: this follows from Kolmogorov's Theorem (i. e., Remark 7–(i) and Remark 11) with $\xi = \xi_1 = 1/2$ (assume, for simplicity, that Q is analytic on W_1 and note that $|\varepsilon|\|P^0\|_{\xi_1} \le |\varepsilon|\|P\|_{C^0}$ by the first inequality in (124)). Now, assume that (128) and (129) holds together with $C_i < 4C$ and $\|\partial(\Phi_i - \mathrm{id})\|_{\alpha\xi_{i+1}} < (\sqrt2 - 1)$ for $0\le i\le j$ ($C_0 = C$ and C_i are as in (64) for, respectively, $K_0 := K$ and K_i). To determine ϕ_{j+1}, observe that, by (128), one has $H_{j+1}\circ\Phi_{j+1} = (K_{j+1}+\varepsilon P_{j+1})\circ\phi_{j+1}$ where $P_j := (P^{j+1} - P^j)\circ\Phi_j$, which is real-analytic on $W_{\alpha\xi_{j+1}}$; thus we may apply Kolmogorov's Theorem to $K_{j+1}+\varepsilon P_{j+1}$ with $\xi = \alpha\xi_{j+1}$ and $\theta = \alpha$; in fact, by the second inequality in (124), $\|P_{j+1}\|_{\alpha\xi_{j+1}} \le \|P^{j-1} - P^j\|_{X_{j+1}} \le c\|P\|_{C^\ell}\xi_{j+1}^\ell$ and the smallness condition (66) becomes $|\varepsilon|D\xi_{j+1}^{\ell-\sigma}$ (with $D := c_* c\|P\|_{C^\ell}(4C)^b 2^{\sigma/2}$), which is clearly satisfied for $|\varepsilon| < D^{-1}$. Thus, ϕ_{j+1} has been determined and (notice that $\alpha^2\xi_{j+1} = \xi_{j+1}/2 = \xi_{j+2}$) $\|\phi_{j+1}-\mathrm{id}\|_{\xi_{j+2}}$, $\partial(\|\phi_{j+1}-\mathrm{id})\|_{\xi_{j+2}} \le |\varepsilon|D\xi_{j+1}$. Let us now check the domain constraint $\Phi_j\colon W_{\alpha\xi_{j+1}} \to X_{\xi_{j+1}}$. By the inductive assumptions and the real-analyticity of Φ_j, one has that, for $z\in W_{\alpha\xi_{j+1}}$, $|\mathrm{Im}\Phi_j(z)| = |\mathrm{Im}(\Phi_j(z) - \Phi_j(\mathrm{Re}z))| \le |\Phi_j(z) - \Phi_j(\mathrm{Re}z)| \le \|\partial\Phi_j\|_{\alpha\xi_{j+1}}|\mathrm{Im}z| \le (1+\|\partial(\Phi_i - \mathrm{id})\|_{\alpha\xi_{i+1}})\alpha\xi_{j+1} < \sqrt2\alpha\xi_{j+1} = \xi_{j+1}$ so that $\Phi_j\colon W_{\alpha\xi_{j+1}} \to X_{\xi_{j+1}}$. The remaining inductive assumptions in (129) with j replaced by $j+1$ are easily checked by arguments similar to those used in the induction proof of Claim **C** above.

40 See, e. g., the Proposition at page 58 of [14] with $g_j = f_j - f_{j-1}$. In fact, the lemma applies to the Hamiltonians H_j and to the symplectic map ϕ_j in (82) in Arnold's scheme with W_j in (81) and taking $C = C_* := \{y' = \lim_{j\to\infty} y_j(\omega)\colon \omega\in B\cap K_y^{-1}(\mathcal{D}^d_{\kappa,\tau})\}$ and $y_j(\omega) := y_j$ is as in (82).

41 A formal ε-power series quasi-periodic trajectory, with rationally-independent frequency ω, for a nearly-integrable Hamiltonian $H(y,x;\varepsilon) := K(y) + \varepsilon P(y,x)$

is, by definition, a sequence of functions $\{z_k\} := (\{v_k\}, \{u_k\})$, real-analytic on $\mathbb{T}^d$ and such that $D_\omega z_k = J_{2d}\pi_k(\nabla H(\sum_{j=0}^{k-1}\varepsilon^j z_j))$ where $\pi_k(\cdot) := \frac{1}{k!}\partial_\varepsilon^k(\cdot)|_{\varepsilon=0}$; compare Remark 1–(ii) above.

42 In fact, Poincaré was not at all convinced of the convergence of such series: see chapter XIII, n° 149, entitled "Divergence des séries de M. Lindstedt", of his book [49].

43 Equation (70) guarantees that the map from y in the $(d-1)$-dimensional manifold $\{K = E\}$ to the $(d-1)$-dimensional real projective space $\{\omega_1 : \omega_2 : \cdots : \omega_d\} \subset \mathbb{RP}^{d-1}$ (where $\omega_i = K_{y_i}$) is a diffeomorphism. For a detailed proof of the "iso-energetic KAM Theorem", see, e. g., [24].

44 Actually, it is not known if such tori are KAM tori in the sense of the definitions given above!

45 The first example of a nearly-integrable system (with two parameters) exhibiting Arnold's diffusion(in a certain region of phase space) was given by Arnold in [4]; a theory for "a priori unstable systems" (i. e., the case in which the integrable system carries also a partially hyperbolic structure) has been worked out in [20] and in recent years a lot of literature has been devoted to study the "a priori unstable" case and to to try to attack the general problem (see, e. g., Sect. 6.3.4 of [6] for a discussion and further references). We mention that J. Mather has recently announced a complete proof of the conjecture in a general case [40].

46 Here, we mention briefly a different and very elementary connection with classical mechanics. To study the spectrum $\sigma(L)$ (L as above with a quasi-periodic potential $V(\omega_1 t, \dots, \omega_n t)$) one looks at the equation $\ddot{q} = (V(\omega t) - \lambda)q$, which is the q-flow of the Hamiltonian ϕ_H^t $H = H(p, q, I, \varphi; \lambda) := p^2/2 + [\lambda - V(\varphi)]q^2/2$ where $(p, q) \in \mathbb{R}^2$ and $(I, \varphi) \in \mathbb{R}^n \times \mathbb{T}^n$ w.r.t. the standard form $\mathrm{d}p \wedge \mathrm{d}q + \mathrm{d}I \wedge \mathrm{d}\varphi$ and λ is regarded as a parameter. Notice that $\dot{\varphi} = \omega$ so that $\varphi = \varphi_0 + \omega t$ and that the (p, q) decouples from the I-flow, which is, then, trivially determined one the (p, q) flow is known. Now, the action-angle variables(J, θ) for the harmonic oscillator $p^2/2 + \lambda q^2/2$ are given by $J = r^2/\sqrt{\lambda}$ and (r, θ) are polar coordinates in the $(p, \sqrt{\lambda}q)$-plane; in such variables, H takes the form $H = \omega \cdot I + \sqrt{\lambda}J - V(\varphi)/\sqrt{\lambda}\sin^2\theta$. Now, if, for example V is small, this Hamiltonian is seen to be a perturbation of $(n+1)$ harmonic oscillator with frequencies $(\omega, \sqrt{\lambda})$ and it is remarkable that one can provide a KAM scheme, which preserves the linear-in-action structure of this Hamiltonian and selects the (Cantor) set of values of the frequency $\alpha = \sqrt{\lambda}$ for which the KAM scheme can be carried out so as to conjugate H to a Hamiltonian of the form $\omega \cdot I + \alpha J$, proving the existence of (generalized) quasi-periodic eigen-functions. For more details along these lines, see [14].

47 The value 10^{-52} is about the proton-Sun mass ratio: the mass of the Sun is about $1.991 \cdot 10^{30}$ kg, while the mass of a proton is about $1.672 \cdot 10^{-21}$ kg, so that (mass of a proton)/(mass of the Sun) $\simeq 8.4 \cdot 10^{-52}$.

48 "Computer-assisted proofs" are mathematical proofs, which use the computers to give rigorous upper and lower bounds on chains of long calculations by means of so-called "interval arithmetic"; see, e. g., Appendix C of [13] and references therein.

49 Simple examples of such orbits are equilibria and periodic orbits: in such cases there are no small-divisor problems and existence was already established by Poincaré by means of the standard Implicit Function Theorem; see [49], Volume I, chapter III.

50 Typically, ξ may indicate an initial datum y_0 and y the distance from such point or (equivalently, if the system is non-degenerate in the classical Kolmogorov sense) $\xi \to \omega(\xi)$ might be simply the identity, which amounts to consider the unperturbed frequencies as parameter; the approach followed here is that in [51], where, most interestingly, m is allowed to be ∞.

51 I. e., a map $\Phi\colon X \to X$ for which $\exists 0 < \alpha < 1$ such that $d(\Phi(u), \Phi(v)) \le \alpha d(u, v)$, $\forall u, v \in X$, $d(\cdot,\cdot)$ denoting the metric on X; for generalities on metric spaces, see, e. g., [36].

52 $\Phi^j = \Phi \circ \cdots \circ \Phi$ j-times. In fact, let $u_j := \Phi^j(u_0)$ and notice that, for each $j \ge 1$ $d(u_{j+1}, u_j) \le \alpha d(u_j, u_{j-1}) \le \alpha^j d(u_1, u_0) =: \alpha^j\beta$, so that, for each $j, h \ge 1$, $d(u_{j+h}, u_j) \le \sum_{i=0}^{h-1} d(u_{j+i+1}, u_{j+i}) \le \sum_{i=0}^{h-1}\alpha^{j+i}\beta \le \alpha^j\beta/(1-\alpha)$, showing that $\{u_j\}$ is a Cauchy sequence. Uniqueness is obvious.

Bibliography

Primary Literature

1. Ahlfors LV (1978) Complex analysis. An introduction to the theory of analytic functions of one complex variable, 3rd edn. International Series in Pure and Applied Mathematics. McGraw–Hill, New York, xi+331 pp
2. Arnold VI (1963) Proof of a Theorem by Kolmogorov AN on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian. Russ Math Surv 18:13–40
3. Arnold VI (1963) Small divisor problems in classical and Celestial Mechanics. Russ Math Surv 18:85–191
4. Arnold VI (1964) Instability of dynamical systems with many degrees of freedom (Russian). Dokl Akad Nauk SSSR 156:9–12
5. Arnold VI (1974) Mathematical methods of classical mechanics series: graduate texts in mathematics, vol 60. Originally published by Nauka, Moscow, 1974. 2nd edn. 1989. Corr. 4th printing, 1997, pp XVI,509

6. Arnold VI, Kozlov VV, Neishtadt AI (ed) (2006) Mathematical Aspects of Classical and Celestial Mechanics, Dynamical Systems III, 3rd edn. Series: Encyclopedia of Mathematical Sciences, vol 3. Springer, Berlin
7. Avila A, Krikorian R (2006) Reducibility or nonuniform hyperbolicity for quasiperiodic Schrödinger cocycles. Ann Math 164:911–940
8. Broer HW (2004) KAM theory: the legacy of AN Kolmogorov's 1954 paper. Comment on: "The general theory of dynamic systems and classical mechanics". (French in: Proceedings of the International Congress of Mathematicians, Amsterdam, 1954, vol 1, pp 315–333, Erven P, Noordhoff NV, Groningen, 1957). Bull Amer Math Soc (N.S.) 41(4):507–521 (electronic)
9. Broer HW, Huitema GB, Sevryuk MB (1996) Quasi-periodic motions in families of dynamical systems. Order amidst chaos. Lecture Notes in Mathematics, vol 1645. Springer, Berlin
10. Celletti A, Chierchia L (1987) Rigorous estimates for a computer-assisted KAM theory. J Math Phys 28:2078–2086
11. Celletti A, Chierchia L (1988) Construction of analytic KAM surfaces and effective stability bounds. Commun Math Phys 118:119–161
12. Celletti A, Chierchia L (1995) A Constructive Theory of Lagrangian Tori and Computer-assisted Applications. In: Jones CKRT, Kirchgraber U, Walther HO (eds) Dynamics Reported, vol 4 (New Series). Springer, Berlin, pp 60–129
13. Celletti A, Chierchia L (2007) KAM stability and celestial mechanics. Mem Amer Math Soc, 187(878), viii+134 pp
14. Chierchia L (1986) Quasi-periodic Schrödinger operators in one dimension, Absolutely continuous spectra, Bloch waves and integrable Hamiltonian systems. Quaderni del Consiglio Nazionale delle Ricerche, Firenze, Italia, 127 pp http://www.mat.uniroma3.it/users/chierchia/REPRINTS/CNR86.pdf
15. Chierchia L (2003) KAM Lectures. In: Dynamical Systems. Part I: Hamiltonian Systems and Celestial Mechanics. Pubblicazioni della Classe di Scienze, Scuola Normale Superiore, Pisa, pp 1–56. Centro di Ricerca Matematica "Ennio De Giorgi"
16. Chierchia L (2006) KAM theory and Celestial Mechanics. In: Françoise J-P, Naber GL, Tsou ST (eds) Encyclopedia of Mathematical Physics, vol 3. Elsevier, Oxford, pp 189–199. ISBN 978-0-1251-2666-3
17. Chierchia L (2008) A.N. Kolmogorov's 1954 paper on nearly-integrable Hamiltonian systems. A comment on: On conservation of conditionally periodic motions for a small change in Hamilton's function [Dokl Akad Nauk SSSR (N.S.) 98:527–530 (1954)]. Regul Chaotic Dyn 13(2):130–139
18. Chierchia L, Falcolini C (1994) A direct proof of a theorem by Kolmogorov in Hamiltonian systems. Ann Sc Norm Su Pisa IV(XXI):541–593
19. Chierchia L, Gallavotti G (1982) Smooth prime integrals for quasi-integrable Hamiltonian systems. Il Nuovo Cimento 67(B)277–295
20. Chierchia L, Gallavotti G (1994) Drift and diffusion in phase space. Ann Inst Henri Poincaré, Phys Théor 60:1–144; see also Erratum, Ann Inst Henri Poincaré, Phys Théor 68(1):135 (1998)
21. Chierchia L, Perfetti P (1995) Second order Hamiltonian equations on $\mathbb{T}^\infty$ and almost-periodic solutions. J Differ Equ 116(1):172–201
22. Chierchia L, Qian D (2004) Moser's theorem for lower dimensional tori. J Differ Equ 206:55–93
23. Coddington EA, Levinson N (1955) Theory of Ordinary Differential Equations. McGraw–Hill, New York. 14+429 pp
24. Delshams A, Gutierrez P (1996) Effective stability and KAM theory. J Differ Equ 128(2):415–490
25. Eliasson L (1988) Perturbations of stable invariant tori for Hamiltonian systems. Ann Scuola Norm Sup Pisa Cl Sci 15:115–147
26. Eliasson LH (1992) Floquet solutions for the 1-dimensional quasi-periodic Schrödinger equation. Comm Math Phys 146:447–482
27. Eliasson LH (1996) Absolutely convergent series expansion for quasi-periodic motions. Math Phys Electron J 2(Paper 4):33 (electronic); see also: Report 2–88, Dept. of Math., University of Stockholm (1988)
28. Evans LC, Gariepy RF (1992) Measure theory and fine properties of functions. Studies in Advanced Mathematics. CRC Press, Boca Raton, pp viii+268
29. Féjoz J (2004) Démonstration du 'théorème d'Arnold' sur la stabilité du système planétaire (d'après Herman). (French) Ergod Theor Dyn Syst 24(5):1521–1582
30. Gallavotti G (1994) Twistless KAM tori, quasi flat homoclinic intersections and other cancellations in the perturbation series of certain completely integrable Hamiltonian systems. A review. Rev Math Phys 6:343–411
31. Graff S (1974) On the continuation of stable invariant tori for Hamiltonian systems, J Differ Equ 15:1–69
32. Hénon M (1966) Explorationes numérique du problème restreint IV: Masses egales, orbites non periodique. Bull Astronom 3(1, fasc. 2):49–66
33. Katok A, Hasselblatt B (1995) Introduction to the modern theory of dynamical systems. In: Encyclopedia of Mathematics and its Applications, vol 54, Cambridge University Press, Cambridge, xviii+802 pp
34. Katznelson Y (2004) An introduction to harmonic analysis, 3rd edn. Cambridge Mathematical Library. Cambridge University Press, Cambridge, pp xviii+314 pp
35. Kolmogorov AN (1954) On the conservation of conditionally periodic motions under small perturbation of the Hamiltonian. Dokl Akad Nauk SSR 98:527–530
36. Kolmogorov AN, Fomin SV (1999) Elements of the Theory of Functions and Functional Analysis. Courier Dover Publications, 288 pp
37. Kuksin SB (1988) Perturbation theory of conditionally periodic solutions of infinite-dimensional Hamiltonian systems and its applications to the Korteweg–de Vries equation. (Russian) Mat Sb (N.S.) 136,178(3):396–412,431; translation in Math USSR-Sb 64(2):397–413
38. Kuksin SB (2004) Fifteen years of KAM for PDE. Geometry, topology, and mathematical physics. Am Math Soc Transl Ser 212(2):237–258
39. Mather J (1984) Nonexistence of invariant circles. Erg Theor Dyn Syst 4:301–309
40. Mather JN (2003) Arnold diffusion. I. Announcement of results. (Russian) Sovrem Mat Fundam Napravl 2:116–130 (electronic); translation in J Math Sci 124(5):5275–5289
41. Melnikov VK (1965) On certain cases of conservation of almost periodic motions with a small change of the Hamiltonian function. (Russian) Dokl Akad Nauk SSSR 165:1245–1248
42. Moser JK (1961) A new technique for the construction of solutions of nonlinear differential equations. Proc Nat Acad Sci USA 47:1824–1831

43. Moser JK (1962) On invariant curves of area-preserving mappings of an annulus. Nach Akad Wiss Göttingen, Math Phys Kl II 1:1–20
44. Moser J (1966) A rapidly convergent iteration method and non-linear partial differential equations. Ann Scuola Norm Sup Pisa 20:499–535
45. Moser J (1967) Convergent series expansions for quasi-periodic motions. Math Ann 169:136–176
46. Moser J (1970) On the construction of almost periodic solutions for ordinary differential equations. In: Proc. Internat. Conf. on Functional Analysis and Related Topics, Tokyo, 1969. University of Tokyo Press, Tokyo, pp 60–67
47. Moser J (1983) Integrable Hamiltonian systems and spectral theory, Lezioni Fermiane. [Fermi Lectures]. Scuola Normale Superiore, Pisa, iv+85 pp
48. Nash J (1956) The imbedding problem for Riemannian manifolds. Ann Math 63(2):20–63
49. Poincaré H (1892) Les méthodes nouvelles de la Mécanique céleste. Tome 1: Gauthier–Villars, Paris, 1892, 385 pp; tome 2: Gauthier–Villars, Paris 1893, viii+479 pp; tome 3: Gauthier–Villars, Paris, 1899, 414 pp
50. Pöschel J (1982) Integrability of Hamiltonian sytems on Cantor sets. Comm Pure Appl Math 35:653–695
51. Pöschel J (1996) A KAM-theorem for some nonlinear PDEs Ann Scuola Norm Sup Pisa Cl Sci 23:119–148
52. Pyartli AS (1969) Diophantine approximations of submanifolds of a Euclidean space. Funktsional Anal Prilozhen 3(4):59–62 (Russian). English translation in: Funct Anal Appl 3:303–306
53. Rüssmann H (1970) Kleine Nenner. I. Über invariante Kurven differenzierbarer Abbildungen eines Kreisringes. (German) Nachr Akad Wiss Göttingen Math-Phys Kl II 1970:67–105
54. Rüssmann H (1975) On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. In: Moser J (ed) Dynamical Systems, Theory and Applications. Lecture Notes in Physics vol 38. Springer, Berlin, pp 598–624
55. Rüssmann H (1976) Note on sums containing small divisors. Comm Pure Appl Math 29(6):755–758
56. Rüssmann H (1980) On the one-dimensional Schrödinger equation with a quasiperiodic potential. In: Nonlinear dynamics, Internat. Conf., New York, 1979, Ann New York Acad Sci, vol 357. New York Acad Sci, New York, pp 90–107
57. Rüssmann H (1989) Nondegeneracy in the perturbation theory of integrable dynamical systems, Number theory and dynamical systems (York, 1987). London Math Soc. Lecture Note Ser, vol 134. Cambridge University Press, Cambridge, pp 5–18
58. Rüssmann H (2001) Invariant tori in non-degenerate nearly integrable Hamiltonian systems. Regul Chaot Dyn 6(2):119–204
59. Salamon D (2004) The Kolmogorov–Arnold–Moser theorem. Math Phys Electron J 10(Paper 3):37 (electronic)
60. Sevryuk MB (2003) The Classical KAM Theory and the dawn of the twenty-first century. Mosc Math J 3(3):1113–1144
61. Siegel CL (1942) Iteration of analytic functions. Ann Math 43:607–612
62. Spivak M (1979) A Comprehensive Introduction to Differential Geometry, vol 1, 2nd edn. Publish or Perish Inc, Wilmington
63. Wayne CE (1990) Periodic and quasi-periodic solutions for nonlinear wave equations via KAM theory. Commun Math Phys 127:479–528
64. Whitney H (1934) Analytic extensions of differentiable functions defined in closed sets. Trans Amer Math Soc 36(1):63–89
65. Zehnder E (1975) Generalized implicit function theorems with applications to some small divisor problems, vol I. Comm Pure Appl Math 28:91–14
66. Zehnder E (1976) Generalized implicit function theorems with applications to some small divisor problems, vol II. Comm Pure Appl Math 29(1):49–111

Books and Reviews

Arnold VI, Novikov SP (eds) (1985) Dynamical Systems IV, Symplectic Geometry and its Applications. In: Encyclopaedia of Mathematical Sciences, vol 4, Volume package Dynamical Systems. Original Russian edition published by VINITI, Moscow. 2nd expanded and rev. edn., 2001

Berretti A, Celletti A, Chierchia L, Falcolini C (1992) Natural Boundaries for Area-Preserving Twist Maps. J Stat Phys 66:1613–1630

Bost JB (1984/1985) Tores invariants des systèmes dynamiques hamiltoniens. Séminaire Bourbaki, exp. 639

Bourgain J (1997) On Melnikov's persistency problem. Math Res Lett 4(4):445–458

Broer HW, Sevryuk MB KAM theory: Quasi-periodicity in dynamical systems. In: Handbook of Dynamical Systems. Elsevier, to appear

Celletti A, Chierchia L (2006) KAM tori for *N*-body problems: a brief history. Celest Mech Dynam Astronom 95(1–4):117–139

Chierchia L, Falcolini C (1996) Compensations in small divisor problems. Comm Math Phys 175:135–160

Gallavotti G (1983) The Elements of Mechanics. Springer, New York

Giorgilli A, Locatelli U (1997) On classical series expansion for quasi-periodic motions. MPEJ 3(5):1–25

Herman M (1986) Sur les courbes invariantes pour les difféomorphismes de l'anneau. Astérisque 2:248

Herman M (1987) Recent results and some open questions on Siegel's linearization theorems of germs of complex analytic diffeomorphisms of $\mathbf{C}^n$ near a fixed point. In: VIIIth International Congress on Mathematical Physics, Marseille, 1986. World Sci. Publishing, Singapore, pp 138–184

Herman M (1989) Inégalités a priori pour des tores Lagrangiens invariants par des difféomorphismes symplectiques. Publ Math IHÈS 70:47–101

Hofer H, Zehnder E (1995) Symplectic invariants and Hamiltonian dynamics. The Floer memorial volume. Progr. Math., 133. Birkhäuser, Basel, pp 525–544

Kappeler T, Pöschel J (2003) KdV & KAM. Springer, Berlin, XIII,279 pp

Koch H (1999) A renormalization group for Hamiltonians, with applications to KAM tori. Ergod Theor Dyn Syst 19:475–521

Kuksin SB (1993) Nearly integrable infinite dimensional Hamiltonian systems. Lecture Notes in Mathematics, vol 1556. Springer, Berlin

Lanford OE III (1984) Computer assisted proofs in analysis. Phys A 124:465–470

Laskar J (1996) Large scale chaos and marginal stability in the solar system. Celest Mech Dyn Astronom 64:115–162

de la Llave R (2001) A tutorial on KAM theory. Smooth ergodic theory and its applications. In: Proc. Sympos. Pure Math., vol 69, Seattle, WA, 1999. Amer Math Soc, Providence, pp 175–292

de la Llave R, González A, Jorba À, Villanueva J (2005) KAM theory without action-angle variables. Nonlinearity 18(2):855–895

de la Llave R, Rana D (1990) Accurate strategies for small divisor problems. Bull Amer Math Soc (N.S.) 22:(1)85–90
Lazutkin VF (1993) KAM theory and semiclassical approximations to eigenfunctions. Springer, Berlin
MacKay RS (1985) Transition to chaos for area-preserving maps. Lect Notes Phys 247:390–454
Meyer KR, Schmidt D (eds) (1991) Computer Aided Proofs in Analysis. Springer, Berlin
Moser J (1973) Stable and random motions in dynamical systems. With special emphasis on celestial mechanics. Hermann Weyl Lectures, the Institute for Advanced Study, Princeton, N. J Annals of Mathematics Studies, No. 77. Princeton University Press, Princeton; University of Tokyo Press, Tokyo
Pöschel J (2001) A lecture on the classical KAM theorem. Smooth ergodic theory and its applications. In: Proc Sympos Pure Math, vol 69 Seattle,WA, 1999. Amer Math Soc, Providence, pp 707–732
Salamon D, Zehnder E (1989) KAM theory in configuration space. Comment Math Helvetici 64:84–132
Sevryuk MB (1995) KAM-stable Hamiltonians. J Dyn Control Syst 1(3):351–366
Sevryuk MB (1999) The lack-of-parameters problem in the KAM theory revisited, Hamiltonian systems with three or more degrees of freedom (S'Agaró, 1995). NATO Adv Sci Inst Ser C Math Phys Sci, vol 533. Kluwer, Dordrecht, pp 568–572
Siegel CL, Moser JK (1971) Lectures on Celestial Mechanics. Springer, Berlin
Weinstein A (1979) Lectures on symplectic manifolds, 2nd edn., CBMS Regional Conference Series in Mathematics, vol 29. American Mathematical Society, Providence
Xu J, You J, Qiu Q (1997) Invariant tori for nearly integrable Hamiltonian systems with degeneracy. Math. Z 226(3):375–387
Yoccoz J-C (1992) Travaux de Herman sur les tores invariants. Séminaire Bourbaki, 1991/92, Exp. 754. Astérisque 206:311–344
Yoccoz J-C (1995) Théorème de Siegel, pôlynomes quadratiques et nombres de Brjuno. Astérisque 231:3–88

Korteweg–de Vries Equation (KdV), Different Analytical Methods for Solving the

YU-JIE REN[1,2,3]
[1] Department of Mathematics, Shantou University, Shantou, People's Republic of China
[2] Department of Mathematics and Physics, Dalian Polytechnic University, Dalian, People's Republic of China
[3] Beijing Institute of Applied Physics and Computational Mathematics, Beijing, People's Republic of China

Article Outline

Glossary
Definition of the Subject
Introduction
The Generalized Hyperbolic Function–Bäcklund Transformation Method and Its Application in the $(2+1)$-Dimensional KdV Equation
The Generalized F-expansion Method and Its Application in Another $(2+1)$-Dimensional KdV Equation
The Generalized Algebra Method and Its Application in $(1+1)$-Dimensional Generalized Variable – Coefficient KdV Equation
A New Exp-N Solitary-Like Method and Its Application in the $(1+1)$-Dimensional Generalized KdV Equation
The Exp-Bäcklund Transformation Method and Its Application in $(1+1)$-Dimensional KdV Equation
Future Directions
Acknowledgment
Bibliography

Glossary

Analytical method for solving an equation The method for obtaining the exact solutions of an equation.

Solitons The special solitary waves which retain their original shapes and speeds after collision and exhibit only a small overall phase shift.

$\boldsymbol{\varepsilon}$ ± 1

$\boldsymbol{t}$ the time

$(\boldsymbol{x}, \boldsymbol{y})$ Cartesian coordinates of a point

$\boldsymbol{\partial_x^{-1}}$ an indefinite integrate operator $\int \mathrm{d}x$.

Definition of the Subject

This paper presents the analytical methods for obtaining the exact solutions of the Korteweg–de Vries Equation (KdV equation).

The KdV equation and its exact solutions can describe and explain many physical problems. In addition, it is a typical, relatively simple and classical equation among the many nonlinear equations in physics. Much of the literature of nonlinear equation theory customarily uses solving soliton solutions of the KdV equation as an example to introduce the nonlinear theory, method and character of soliton solutions. During the last five decades, the construction of exact solution for a wide class of nonlinear equations has been an exciting and extremely active area of research. This includes the most famous nonlinear example of the KdV equation.

In the family of the KdV equations, a well known KdV equation is expressed in its simplest form as [1,2,3,4]:

$$u_t(x,t) + \alpha u(x,t) u_x(x,t) + \beta u_{xxx}(x,t) = 0\,. \tag{1}$$

where the coefficient of the nonlinear term α and the coefficient of the dispersive term β are independent of x and t.

The KdV equation was derived in 1885 by the two scientists, Korteweg and de Vries [1], to describe long wave propagation on shallow water. Its importance comes from its many applications in weakly nonlinear and weakly dispersive physical systems and its interesting mathematical properties. It can be considered as a paradigm in nonlinear science. The equation is so powerful it has been used to simulate the red spots on Jupiter. Moreover, it describes non-linear waves in rotating fluids [5], the giant ocean waves known as Tsunami, and other aspects of solitary waves in plasma. The giant internal waves in the interior of the ocean arising from temperature differences which may destroy marine vessels can be also described by the powerful KdV equation. It possesses infinitely many generalized symmetries and infinitely many integrals of motion [6], and can be viewed as a completely integrable Hamiltonian system and solved by the inverse scattering transform. It also provides resources for studying integrability of nonlinear differential (and difference) equations, serving as a typical model of integrable equations. More remarkable is that various physically important solutions to the KdV equation can be presented explicitly in a simple way, among which are solitons, rational solutions, positons and negatons.

The KdV equation has been widely investigated and developed continually in recent decades. Several generalizations of the KdV equation have found applications in many areas, including quantum field theory, plasma physics, solid-state physics, liquid-gas bubble mixtures, and anharmonic crystals. For instance, the classical generalized Korteweg–de Vries equation (gKdV) has the following form [7,8,9]

$$u_t(x,t) + \alpha u^p(x,t) u_x(x,t) + \beta u_{xxx}(x,t) = 0\,. \quad (2)$$

where the coefficient of the nonlinear term α and the coefficient of the dispersive term β are independent of x and t.

The more generalized form of the KdV equation is defined by parameters (l, p) as follows [10,11,12,13,14]

$$\begin{aligned} u_t(x,t) = u^{l-2}(x,t)u_x(x,t) + \alpha[2u^p(x,t)u_{xxx}(x,t) \\ + 4pu^{p-1}(x,t)u_x(x,t)u_{xx}(x,t) \\ + p(p-1)u^{p-2}(x,t)u_x^3(x,t)]\,. \quad (3) \end{aligned}$$

Propagation of weakly nonlinear long waves in an inhomogeneous waveguide is governed by a variable-coefficient KdV equation of the form [15]

$$u_t(x,t) + 6u(x,t)u_x(x,t) + B(t)u_{xxx}(x,t) = 0\,. \quad (4)$$

where $u(x,t)$ is the wave amplitude, t the propagation coordinate, x the temporal variable and $B(t)$ is the local dispersion coefficient.

The applicability of the variable-coefficient KdV equation (4) arises in many areas of physics as, for example, the description of the propagation of gravity-capillary and interfacial-capillary waves, internal waves and Rossby waves [15]. In order to study the propagation of weakly nonlinear, weakly dispersive waves in the inhomogeneous media, Eq. (4) is rewritten as follows [16]

$$u_t(x,t) + 6A(t)u(x,t)u_x(x,t) + B(t)u_{xxx}(x,t) = 0\,. \quad (5)$$

which includes a variable nonlinearity coefficient $A(t)$.

In studying long waves, tides, solitary waves, and related phenomena, one is led to an equation of the form [17]

$$\begin{aligned} u_t(x,t) + (m+1)(m+2)u^m(x,t)u_x(x,t) \\ + u_{xxx}(x,t) = f(x,t)\,. \quad (6) \end{aligned}$$

where $f(x,t)$ is a given function and $m = 1, 2, \ldots$, with $u(x,t), u_x(x,t), u_{xx}(x,t) \to 0$ as $|x| \to +\infty$. This equation is referred to as a generalized Korteweg–de Vries equation in [15].

The extended Korteweg–de Vries (eKdV) equation [18]

$$\begin{aligned} u_t(x,t) + \alpha u(x,t)u_x(x,t) + \beta u^2(x,t)u_x(x,t) \\ + \delta u_{xxx}(x,t) = 0\,. \quad (7) \end{aligned}$$

incorporates both quadratic and cubic nonlinearities, and serves as a useful model in the propagation of long waves in physical oceanography.

A system of two coupled Korteweg–de Vries equations in [4] is given by the following equation:

$$\begin{aligned} u_{it}(x,t) + \sum_{h,k} A_i^{hk} u_h(x,t) u_{kx}(x,t) \\ + \sum_k A_i^k u_{kxxx}(x,t) = 0\,, \quad i,k,h = 1,2\,. \quad (8) \end{aligned}$$

The $(2+1)$-dimensional KdV equation is presented by the following equation [19]:

$$\begin{cases} u_t(x,y,t) - 3v(x,y,t)u_x(x,y,t) \\ \quad -3v_x(x,y,t)u(x,y,t) + u_{xxx}(x,y,t) = 0\,, \\ u_x(x,y,t) - v_y(x,y,t) = 0\,. \end{cases} \quad (9)$$

Another $(2+1)$-dimensional KdV equation in [20] is written the following form

$$\begin{aligned} &u_t(x,y,t) - 4u(x,y,t)u_y(x,y,t) \\ &\quad - 4u_x(x,y,t)\partial_x^{-1}u_y(x,y,t) - u_{xxy}(x,y,t) = 0 \,. \end{aligned} \tag{10}$$

where ∂_x^{-1} is an indefinite integrate operator $\int \mathrm{d}x$, possessing some interesting coherent structures.

In addition, the complex coupled KdV equation in [49] is written the following form

$$\begin{cases} u_t(x,t) = \frac{1}{2}[u_{xxx}(x,t) - 6u(x,t)u_x(x,t)] \\ \qquad\qquad\qquad + 3(\mid v(x,t) \mid^2)_x \,, \\ v_t(x,t) = -v_{xxx}(x,t) + 3u(x,t)v_x(x,t) \,. \end{cases} \tag{11}$$

Introduction

It is well known that nonlinear complex physical phenomena are related to nonlinear partial differential equations (NLPDEs), which are involved in many fields from physics to biology, chemistry, mechanics, etc. The investigation of exact solutions of NLPDEs as mathematical models of phenomena can help one to better understand a variety of phenomena. In the past several decades, many analytical methods for obtaining exact solutions of NLPDEs have been presented, such as the inverse scattering method, Hirota's bilinear method [21], the Bäcklund transformation [22], the Painleve expansion [23], the tanh function method [24], the sine-cosine method [25], the homogenous balance method [26], the homotopy perturbation method [27], the variational method [28], asymptotic methods [29], non-perturbative methods [30], the exp-function method [31], the Adomian Pade approximation [32], the Jacobi elliptic function expansion method [33], the F-expansion method [34], the Weierstrass semi-rational expansion method [35], the unified rational expansion method [36], the algebraic method [37, 50,51], the auxiliary equation method [38], and so on.

In recent years, based on the ideas of unification methods, algorithm realization and mechanization for solving NLPDEs, we have improved and presented some analytical methods for obtaining exact solutions of NLPDEs, such as the generalized hyperbolic function – the Bäcklund transformation method [39], the method for constructing higher order and higher dimension Bäcklund transformation [40], the generalized hyperbolic function – the Riccati method [40], the generalized F-expansion method [41, 42], the extended generalized algebraic method [43,44], the Exp-Bäcklund method [45,46], the Exp-N soliton-like method [47,48], the extended sine-cosine method [52], and so on. The present article is motivated by the desire to introduce and make use of our works published in [39, 40,41,42,43,44,45,46,47,48] to construct more general exact solutions, which contain not only the results obtained by using the methods [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15, 16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34, 35,36,37,38,49,50,51] and Ablowitz and Clarkson 1991, Dodd et al. 1984, Drazin and Johnson 1989, Fan 2004, Guo 1996, Liu and Liu 2000 Miura 1976, and Miurs 1978, but also a series of new and more general exact solutions. For illustration, we apply our analytical methods above to the different types of the KdV equations, and successfully obtain many new and more general exact solutions.

The remainder of this article is organized as follows. In Sect. "The Generalized Hyperbolic Function–Bäcklund Transformation Method and Its Application in the $(2+1)$-Dimensional KdV Equation", in order to construct the unification solutions of the hyperbolic function solution and exponential function solution of NLPDEs, we first find a new theory of generalized hyperbolic functions which includes two new definitions of generalized hyperbolic functions and generalized hyperbolic function transformations and their properties. Then we present a new higher order and higher dimension Bäcklund transformation method and a new generalized hyperbolic function – Bäcklund transformation method. The validity of the methods are tested by their application in the $(2+1)$-dimensional KdV equation (9). In Sect. "The Generalized F-expansion Method and Its Application in Another $(2+1)$-Dimensional KdV Equation", we present a new and more general formal transformation and a new generalized F-expansion method and apply them to another $(2+1)$-dimensional KdV equation (10). In Sect. "The Generalized Algebra Method and Its Application in $(1+1)$-Dimensional Generalized Variable – Coefficient KdV Equation", in order to develop the algebraic method [50,51] for constructing the traveling wave solutions of NLPDEs, we first present two new and general transformations, two new theorems and their proofs, by using Maple, and a new mechanization method to find the exact solutions of a first-order nonlinear ordinary differential equations with any degree. The validity of the method is tested by its application in the first-order nonlinear ordinary differential equation with six degrees. Next, we present a new and more general transformation and a new generalized algebra method and apply them to the $(1+1)$-dimensional generalized variable – coefficient KdV equation (5). In Sect. "A New Exp-N Solitary-like Method and Its Application in the $(1+1)$-Dimensional Generalized KdV Equation", in order to develop the Exp-function method [31], we present two new generic transformations, a new Exp-N solitary-like method, and its algorithm. In addition, we apply our methods to construct new exact solutions of the $(1+1)$-dimensional classical generalized KdV (gKdV) equation (2); In Sect. "The the Exp-Bäcklund Transformation Method and Its Application in $(1+1)$-Dimensional KdV Equation", a new Exp-Bäcklund transformation method and its algorithm is presented to find more exact solutions of NLPDEs. We choose the $(1+1)$-dimensional KdV equation (1) to illustrate the effectiveness and convenience of our algorithm. As a result, we obtain more general exact solutions including non-traveling wave solutions and traveling wave solutions. In addition, the long-term behavior of the non-traveling wave solutions and traveling wave solutions are illustrated by some Figures. In Sect. "Future Directions", future directions for Solving NLPDEs are given.

The Generalized Hyperbolic Function–Bäcklund Transformation Method and Its Application in the (2 + 1)-Dimensional KdV Equation

In this section, in order to construct the unification solutions of the hyperbolic function solution and the exponential function solution of NLPDEs, we establish a new theory of generalized hyperbolic functions, which includes two new definitions of generalized hyperbolic functions and generalized hyperbolic function transformations and their properties that we first presented in [39,40]. Then we present a new higher order and higher dimension Bäcklund transformation method and a new generalized hyperbolic function-Bäcklund transformation method that we first presented in [39,40]. With the aid of symbolic computation, we choose the $(2+1)$-dimensional Korteweg–de Vries equations to illustrate the validity and advantages of the methods. As a result, many new and more general exact non-traveling waves are obtained.

The Definition and Properties of Generalized Hyperbolic Functions

In [39], we first defined the following new functions which named generalized hyperbolic functions and studied the properties of these functions for constructing new exact solutions of NLPDEs.

Definition 1 Suppose that ξ is an independent variable, p, q and k are constants.

The **generalized hyperbolic sine function** is

$$\sinh_{pqk}(\xi) = \frac{pe^{k\xi} - qe^{-k\xi}}{2} \tag{12}$$

generalized hyperbolic cosine function is

$$\cosh_{pqk}(\xi) = \frac{pe^{k\xi} + qe^{-k\xi}}{2} \tag{13}$$

generalized hyperbolic tangent function is

$$\tanh_{pqk}(\xi) = \frac{pe^{k\xi} - qe^{-k\xi}}{pe^{k\xi} + qe^{-k\xi}} \tag{14}$$

generalized hyperbolic cotangent function is

$$\coth_{pqk}(\xi) = \frac{pe^{k\xi} + qe^{-k\xi}}{pe^{k\xi} - qe^{-k\xi}} \tag{15}$$

generalized hyperbolic secant function is

$$\operatorname{sech}_{pqk}(\xi) = \frac{2}{pe^{k\xi} + qe^{-k\xi}} \tag{16}$$

generalized hyperbolic cosecant function is

$$\operatorname{csch}_{pqk}(\xi) = \frac{2}{pe^{k\xi} - qe^{-k\xi}} \tag{17}$$

the above six kinds of functions are said **generalized hyperbolic functions**.

Thus we can prove the following theory of generalized hyperbolic functions on the basis of Definition 1.

Theorem 1 *The generalized hyperbolic functions satisfy the following relations:*

$$\cosh^2_{pqk}(\xi) - \sinh^2_{pqk}(\xi) = pq\,, \tag{18}$$

$$1 - \tanh^2_{pqk}(\xi) = pq \cdot \operatorname{sech}^2_{pqk}(\xi)\,, \tag{19}$$

$$1 - \coth^2_{pqk}(\xi) = -pq \cdot \operatorname{csch}^2_{pqk}(\xi)\,, \tag{20}$$

$$\operatorname{sech}_{pqk}(\xi) = \frac{1}{\cosh_{pqk}(\xi)}\,, \tag{21}$$

$$\operatorname{csch}_{pqk}(\xi) = \frac{1}{\sinh_{pqk}(\xi)}\,, \tag{22}$$

$$\tanh_{pqk}(\xi) = \frac{\sinh_{pqk}(\xi)}{\cosh_{pqk}(\xi)}\,, \tag{23}$$

$$\coth_{pqk}(\xi) = \frac{\cosh_{pqk}(\xi)}{\sinh_{pqk}(\xi)}\,, \tag{24}$$

$$\tanh_{pqk}(\xi) = \frac{1}{\coth_{pqk}(\xi)}\,, \tag{25}$$

$$\sinh_{pqk}(-\xi) = -\sinh_{qpk}(\xi)\,, \tag{26}$$

$$\cosh_{pqk}(-\xi) = \cosh_{qpk}(\xi)\,, \tag{27}$$

$$\tanh_{pqk}(-\xi) = -\tanh_{qpk}(\xi)\,, \tag{28}$$

$$\coth_{pqk}(-\xi) = -\coth_{qpk}(\xi)\,, \tag{29}$$

$$\operatorname{sech}_{pqk}(-\xi) = \operatorname{sech}_{qpk}(\xi)\,, \tag{30}$$

$$\operatorname{csch}_{pqk}(-\xi) = -\operatorname{csch}_{qpk}(\xi)\,, \tag{31}$$

$$\sinh_{11k}(\xi-\eta) = \frac{1}{pq}[\sinh_{qpk}(\xi)\cdot\cosh_{pqk}(\eta) - \cosh_{pqk}(\xi)\cdot\sinh_{qpk}(\eta)]\,, \tag{32}$$

$$\sinh_{p^2q^2k}(\xi+\eta) = \sinh_{qpk}(\xi)\cdot\cosh_{pqk}(\eta) + \cosh_{pqk}(\xi)\cdot\sinh_{qpk}(\eta)\,, \tag{33}$$

$$\cosh_{p^2q^2k}(\xi+\eta) = \cosh_{qpk}(\xi)\cdot\cosh_{pqk}(\eta) + \sinh_{pqk}(\xi)\cdot\sinh_{qpk}(\eta)\,, \tag{34}$$

$$\cosh_{11k}(\xi-\eta) = \frac{1}{pq}[\cosh_{qpk}(\xi)\cdot\cosh_{pqk}(\eta) - \sinh_{pqk}(\xi)\cdot\sinh_{qpk}(\eta)]\,, \tag{35}$$

$$\sinh_{q^2p^2k}(2\xi) = 2\sinh_{pqk}(\xi)\cdot\cosh_{pqk}(\xi)\,, \tag{36}$$

$$\cosh_{p^2q^2k}(2\xi) = \sinh^2_{qpk}(\xi) + \cosh^2_{qpk}(\xi)\,. \tag{37}$$

$$\frac{\tanh_{qpk}(\xi) + \tanh_{qpk}(\eta)}{1-\tanh_{qpk}(\xi)\tanh_{qpk}(\eta)} = \frac{\sinh_{p^2q^2k}(\xi+\eta)}{p\cdot q\cdot\cosh_{11k}(\xi-\eta)}\,, \tag{38}$$

$$\frac{\tanh_{qpk}(\xi) - \tanh_{qpk}(\eta)}{1+\tanh_{qpk}(\xi)\tanh_{qpk}(\eta)} = \frac{p\cdot q\cdot\sinh_{11k}(\xi-\eta)}{\cosh_{p^2q^2k}(\xi+\eta)}. \tag{39}$$

For simplicity, only a sample of these are proved here.

Proof By (18) and (21), we obtain

$$\begin{aligned} 1-\tanh^2_{pqk}(\xi) &= 1-\frac{\sinh^2_{pqk}(\xi)}{\cosh^2_{pqk}(\xi)} \\ &= \frac{\cosh^2_{pqk}(\xi)-\sinh^2_{pqk}(\xi)}{\cosh^2_{pqk}(\xi)} \\ &= \frac{pq}{\cosh^2_{pqk}(\xi)} = pq\cdot\ \operatorname{sech}^2_{pqk}(\xi)\,. \end{aligned}$$

Similarly, we can prove other conclusions in Theorem 1.□

Theorem 2 *The derivative formulae of generalized hyperbolic functions are as follows:*

$$\frac{d(\sinh_{pqk}(\xi))}{d\xi} = k\cdot\cosh_{pqk}(\xi)\,, \tag{40}$$

$$\frac{d(\cosh_{pqk}(\xi))}{d\xi} = k\cdot\sinh_{pqk}(\xi)\,, \tag{41}$$

$$\frac{d(\tanh_{pqk}(\xi))}{d\xi} = kpq\cdot\operatorname{sech}^2_{pqk}(\xi)\,, \tag{42}$$

$$\frac{d(\coth_{pqk}(\xi))}{d\xi} = -kpq\cdot\operatorname{csch}^2_{pqk}(\xi)\,, \tag{43}$$

$$\frac{d(\operatorname{sech}_{pqk}(\xi))}{d\xi} = -k\cdot\operatorname{sech}_{pqk}(\xi)\cdot\tanh_{pqk}(\xi)\,, \tag{44}$$

$$\frac{d(\operatorname{csch}_{pqk}(\xi))}{d\xi} = -k\cdot\operatorname{csch}_{pqk}(\xi)\cdot\coth_{pqk}(\xi)\,. \tag{45}$$

Proof According to (40) and (41), we can get

$$\begin{aligned} \frac{d(\tanh_{pqk}(\xi))}{d\xi} &= \left(\frac{\sinh_{pqk}(\xi)}{\cosh_{pqk}(\xi)}\right)' \\ &= \frac{(\sinh_{pqk}(\xi))'\cosh_{pqk}(\xi) - \sinh_{pqk}(\xi)(\cosh_{pqk}(\xi))'}{\cosh_{pqk}{}^2(\xi)} \\ &= \frac{k\cdot\cosh_{pqk}(\xi)\cdot\cosh_{pqk}(\xi) - \sinh_{pqk}(\xi)\cdot(k\cdot\sinh_{pqk}(\xi))}{\cosh^2_{pqk}(\xi)} \\ &= kpq\cdot\operatorname{sech}^2_{pqk}(\xi)\,. \end{aligned}$$

□

Similarly, we can prove other differential coefficient formulae in Theorem 2.

Remark 1 We see that when $p=1, q=1, k=1$ in (12)–(17), new generalized hyperbolic functions $\sinh_{pqk}(\xi)$, $\cosh_{pqk}(\xi), \tanh_{pqk}(\xi), \coth_{pqk}(\xi), \operatorname{sech}_{pqk}(\xi)$ and $\operatorname{csch}_{pqk}(\xi)$ degenerate as hyperbolic functions $\sinh(\xi)$, $\cosh(\xi), \tanh(\xi), \coth(\xi), \operatorname{sech}(\xi)$ and $\operatorname{csch}(\xi)$, respectively. In addition, when $p=0$ or $q=0$ in (12)–(17), $\sinh_{pqk}(\xi), \cosh_{pqk}(\xi), \operatorname{sech}_{pqk}(\xi), \operatorname{csch}_{pqk}(\xi)$, $\tanh_{pqk}(\xi)$ and $\coth_{pqk}(\xi)$ degenerate as exponential functions $\frac{1}{2}pe^{k\xi}, \pm\frac{1}{2}qe^{-k\xi}, 2pe^{-k\xi}, \pm 2qe^{k\xi}$ and ± 1, respectively.

A New Higher Order and Higher Dimension Bäcklund Transformation Method to Construct an Auto–Bäcklund Transformation of the (2 + 1)-Dimensional KdV Equation

In this section, we obtain an auto-Bäcklund transformation of the following (2 + 1)-dimensional KdV equation by making use of the method for constructing higher order and higher dimension Bäcklund transformations which we presented in [39,40] via symbolic computation.

The $(2+1)$-dimensional KdV equations [19] take the form

$$\begin{cases} u_t(x,y,t) - 3v(x,y,t)u_x(x,y,t) \\ \qquad -3v_x(x,y,t)u(x,y,t) + u_{xxx}(x,y,t) = 0\,, \\ u_x(x,y,t) - v_y(x,y,t) = 0\,. \end{cases} \tag{46}$$

The auto-Bäcklund transformations of Eqs. (46) have the following forms:

$$\begin{cases} u(x,y,t) = \sum_{j_1=0}^{m_1} u_{j_1}(x,y,t) f^{j_1-m_1}(x,y,t)\,, \\ v(x,y,t) = \sum_{j_2=0}^{m_2} v_{j_2}(x,y,t) f^{j_2-m_2}(x,y,t)\,, \end{cases} \tag{47}$$

where $f(x,y,t), u_{j_1}(x,y,t), j_1 = 0,1,2,\ldots,m_1-1$ and $v_{j_2}(x,y,t), j_2 = 0,1,2,\ldots,m_2-1$ are all differential functions to be determined later, and $u_{m_1}(x,y,t), v_{m_2}(x,y,t)$ are the trivial seed solutions of Eqs. (46).

By balancing the highest order linear term and nonlinear terms in Eqs. (46), we get $m_1 = 2$ and $m_2 = 2$, and (47) has the following formal

$$\begin{cases} u(x,y,t) = u_0(x,y,t)f^{-2}(x,y,t) \\ \qquad +u_1(x,y,t)f^{-1}(x,y,t) + u_2(x,y,t)\,, \\ v(x,y,t) = v_0(x,y,t)f^{-2}(x,y,t) \\ \qquad +v_1(x,y,t)f^{-1}(x,y,t) + v_2(x,y,t)\,, \end{cases} \tag{48}$$

We take the trivial seed solutions as

$$u_2 = u_2(x,y,t), v_2 = v_2(x,y,t)\,. \tag{49}$$

With the aid of Maple symbolic computation software, substituting (48) and (49) into (46), and collecting all terms with $f^{-i}(x,y,t), i = 0,1,2,\ldots$, we obtain

$$\big(-24u_0 f_x^3 + 12u_0 v_0 f_x\big) f^{-5} + \big(12u_1 f_x^3 + 36 f_x^2 f_y f_{x^2} + 18v_1 f_x^2 f_y + 24 f_x^3 f_{xy}\big) f^{-4} + \cdots = 0\,, \tag{50}$$

$$\big(-2u_0 f_x + 2v_0 f_y\big) f^{-3} + \big(u_{0x} - u_1 f_x + v_1 f_y - v_{0y}\big) f^{-2} + \big(u_{1x} - v_{1y}\big) f^{-1} + u_{2x} - v_{2y} = 0\,, \tag{51}$$

Setting the coefficient of f^{-5} in (50) and f^{-3} in (51) to be zero, we obtain a differential equation

$$-24u_0 f_x^3 + 12u_0 v_0 f_x = 0\,, \tag{52}$$

$$-2u_0 f_x + 2v_0 f_y = 0\,, \tag{53}$$

which has the solution

$$u_0 = 2f_x f_y, \quad v_0 = 2f_x^2\,. \tag{54}$$

Setting the coefficients of f^{-4} in (50) and f^{-2} in (51) to be zero, we obtain a differential equation

$$12u_1 f_x^3 + 36 f_x^2 f_y f_{x^2} + 18v_1 f_x^2 f_y + 24 f_x^3 f_{xy} = 0\,, \tag{55}$$

$$u_{0x} - u_1 f_x + v_1 f_y - v_{0y} = 0\,, \tag{56}$$

from which we can get the following expression:

$$u_1 = -2f_{yx}, \quad v_1 = -2f_{xx}\,. \tag{57}$$

By (48), (49), (54) and (57), we obtain an auto-Bäcklund transformation of Eqs. (46)

$$\begin{aligned} u &= -2\partial_{yx}\ln(f(x,y,t)) + u_2(x,y,t)\,, \\ v &= -2\partial_{xx}\ln(f(x,y,t)) + v_2(x,y,t)\,, \end{aligned} \tag{58}$$

where $u_2(x,y,t), v_2(x,y,t)$ are the trivial seed solutions of Eqs. (46), $f = f(x,y,t)$ satisfies the Eqs. (52), (53), (55), (56) and the following equations

$$\begin{aligned} &-u_0 v_{1x} - u_1 v_{0x} + 2u_1 v_1 f_x + 2u_2 v_0 f_x - 2/3 u_0 f_t \\ &- 2/3 u_0 f_{x^3} + 2u_1 f_x f_{x^2} - 2u_{0x} f_{x^2} \\ &+ 2u_{1x} f_x^2 - 2u_{0x^2} f_x - v_0 u_{1x} - v_1 u_{0x} \\ &+ 2v_2 u_0 f_x = 0\,, \\ &u_{0x^3} - u_1 f_{x^3} - 3u_0 v_{2x} - 3u_1 v_{1x} - 3u_2 v_{0x} \\ &+ 3u_2 v_1 f_x - 3u_{1x} f_{x^2} - u_1 f_t - 3v_0 u_{2x} - 3v_1 u_{1x} \\ &- 3v_2 u_{0x} + 3v_2 u_1 f_x + u_{0t} - 3u_{1x^2} f_x = 0\,. \end{aligned} \tag{59}$$

The Generalized Hyperbolic Function–Bäcklund Transformation Method and Its Application in the (2 + 1)-Dimensional KdV Equation

We first give the definition of a generalized hyperbolic function transformation [40] as follows:

Definition 2 If a transformation includes the generalized hyperbolic functions, then the transformation is defined to be a generalized hyperbolic function transformation.

In this section, we will use the generalized hyperbolic function–Bäcklund transformation method to seek new exact solutions of Eqs. (46). We take $f(x,y,t)$ in (48), (49), (54), (57) and (59) as being of a new form which is the following generalized hyperbolic function transformation

$$f(x,y,t) = H(t) + \sum_{i=1}^{n} K_i(t)\cdot F_i(\xi_i(x,y,t))\cdot G_i(\eta_i(x,y,t))\,, \tag{60}$$

where $F_i(\xi_i)$ and $G_i(\eta_i)$ may take any two generalized hyperbolic functions among $\sinh_{pqk}(\xi)$, $\cosh_{pqk}(\xi)$,

$\tanh_{pqk}(\xi), \coth_{pqk}(\xi), \mathrm{sech}_{pqk}(\xi)$and $\mathrm{csch}_{pqk}(\xi)$. And then (60) has twenty–one types of general forms. For example

$$f(x,y,t) = H(t) + \sum_{i=1}^{n} K_i(t)\cdot \mathrm{sech}_{p_{1i}q_{1i}k_{1i}}(\xi_i(x,y,t)) \cdot \tanh_{p_{2i}q_{2i}k_{2i}}(\eta_i(x,y,t)) \ , \quad (61)$$

$$f(x,y,t) = H(t)+\sum_{i=1}^{n} K_i(t)\cdot \mathrm{csch}^2_{p_iq_ik_i}(\xi_i(x,y,t)), \quad (62)$$

We see that when $p = 1, q = 1, k = 1$ in (12)–(12), new generalized hyperbolic function $\sinh_{pqk}(\xi), \cosh_{pqk}(\xi), \tanh_{pqk}(\xi), \coth_{pqk}(\xi), \mathrm{sech}_{pqk}(\xi)$ and $\mathrm{csch}_{pqk}(\xi)$ degenerate as hyperbolic functions $\sinh(\xi), cosh(\xi), \tanh(\xi), \coth(\xi), \mathrm{sech}(\xi)$ and $\mathrm{csch}(\xi)$ respectively. Therefore (3.49) includes twenty-one types of hyperbolic function forms. For example

$$f(x,y,t) = H(t) + \sum_{i=1}^{n} K_i(t)\cdot \sinh(\xi_i(x,y,t)) \cdot \mathrm{sech}(\eta_i(x,y,t)) \ , \quad (63)$$

and thirty-six types of the omnibus forms of generalized hyperbolic functions and hyperbolic functions. For example

$$f(x,y,t) = H(t) + \sum_{i=1}^{n} K_i(t)\cdot \mathrm{sech}(\xi_i(x,y,t)) \cdot \mathrm{csch}_{p_{2i}q_{2i}k_{2i}}(\eta_i(x,y,t)) \ , \quad (64)$$

In addition, when $p = 0$ or $q = 0$ in (12)–(17), $\sinh_{pqk}(\xi), \cosh_{pqk}(\xi), \mathrm{sech}_{pqk}(\xi), \mathrm{csch}_{pqk}(\xi), \tanh_{pqk}(\xi)$ and $\coth_{pqk}(\xi)$ degenerate as exponential functions $\frac{1}{2}pe^{k\xi}$, $\pm\frac{1}{2}qe^{-k\xi}, 2pe^{-k\xi}, \pm 2qe^{k\xi}$ and ± 1 respectively.

For example, we take $F_i(\xi_i(x,y,t)) = \tanh_{p_{1i}q_{1i}k_{1i}}(\xi_i), G_i(\eta_i(x,y,t)) = \sinh_{p_{2i}q_{2i}k_{2i}}(\eta_i)$. Then (60) has the new form

$$f(x,y,t) = H(t) + \sum_{i=1}^{n} K_i(t)\tanh_{p_1q_1k_1}(\xi_i) \mathrm{sech}_{p_2q_2k_2}(\eta_i) \ , \quad (65)$$

where $\xi_i = \alpha_i(t)x+\beta_i(t)y+r_i(t), \eta_i = \mu_i(t)x+\nu_i(t)y+l_i(t), (i = 1,2,\ldots,n)$, and $H(t), K_i(t), \alpha_i(t), \beta_i(t), r_i(t), \mu_i(t), \nu_i(t)$ and $l_i(t)$ are the functions of t, $p_{1i}, q_{1i}, k_{1i}, p_{2i}, q_{2i}, k_{2i}$ are the constants, and $\sum_{i=1}^{n} K_i^2(t) \neq 0 (i = 0,1,2,\ldots,n)$.

We take the initial solution of Eqs. (46) as $u_2(x,y,t) = u_2(t)$, and $v_2(x,y,t) = v_2(t)$ in (58) for convenience. So based on the Bäcklund transformation (58), the generalized hyperbolic function transformation (65), and

$$\frac{d^2(\tanh_{pqk}(\xi))}{d\xi^2} = -2k^2pq\cdot\mathrm{sech}^2_{pqk}(\xi)\cdot\tanh_{pqk}(\xi), \quad (66)$$

$$\frac{d^2(\mathrm{sech}_{pqk}(\xi))}{d\xi^2} = k^2\cdot\mathrm{sech}_{pqk}(\xi)\cdot(2\tanh^2_{pqk}(\xi)-1), \quad (67)$$

we can get the following new multi-soliton-like solution which we call a generalized hyperbolic function solution of Eq. (46):

$$\begin{aligned} u = {} & \frac{2\sum_{i=1}^{n} p_1q_1k_1k_2K_i(t)\nu_i(t)\alpha_i(t)\tanh_{p_2q_2k_2}(\eta_i)\,\mathrm{sech}^2_{p_1q_1k_1}(\xi_i)\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)}{H(t)+\sum_{i=1}^{n}K_i(t)\tanh_{p_1q_1k_1}(\xi_i)\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)} \\ & + \frac{2\sum_{i=1}^{n} k_2^2K_i(t)\mu_i(t)\nu_i(t)\tanh_{p_1q_1k_1}(\xi_i)\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)(1-2\tanh^2_{p_2q_2k_2}(\eta_i))}{H(t)+\sum_{i=1}^{n}K_i(t)\tanh_{p_1q_1k_1}(\xi_i)\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)} \\ & + \frac{2\sum_{i=1}^{n} p_1q_1k_1k_2K_i(t)\beta_i(t)\mu_i(t)\,\mathrm{sech}^2_{p_1q_1k_1}(\xi_i)\tanh_{p_2q_2k_2}(\eta_i)\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)}{H(t)+\sum_{i=1}^{n}K_i(t)\tanh_{p_1q_1k_1}(\xi_i)\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)} \\ & \frac{2\sum_{i=1}^{n} K_i(t)(p_1q_1k_1\beta_i(t)\,\mathrm{sech}^2_{p_1q_1k_1}(\xi_i)-k_2\nu_i(t)\tanh_{p_1q_1k_1}(\xi_i)\tanh_{p_2q_2k_2}(\eta_i))\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)}{H(t)+\sum_{i=1}^{n}K_i(t)\tanh_{p_1q_1k_1}(\xi_i)\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)} \\ & \frac{\sum_{i=1}^{n} K_i(t)(p_1q_1k_1\alpha_i(t)\,\mathrm{sech}^2_{p_1q_1k_1}(\xi_i)-k_2\mu_i(t)\tanh_{p_1q_1k_1}(\xi_i)\tanh_{p_2q_2k_2}(\eta_i))\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)}{H(t)+\sum_{i=1}^{n}K_i(t)\tanh_{p_1q_1k_1}(\xi_i)\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)} \\ & + u_2 \ , \end{aligned} \quad (68)$$

$$\begin{aligned} v = {} & \frac{\sum_{i=1}^{n} 4p_1q_1k_1\alpha_i(t)K_i(t)\,\mathrm{sech}^2_{p_1q_1k_1}(\xi_i)\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)(k_1\alpha_i(t)\tanh_{p_1q_1k_1}(\xi_i)+k_2\mu_i(t)\tanh_{p_2q_2k_2}(\eta_i))}{H(t)+\sum_{i=1}^{n}K_i(t)\tanh_{p_1q_1k_1}(\xi_i)\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)} \\ & + \frac{2\sum_{i=1}^{n} K_i(t)k_2^2(\mu_i(t))^2\tanh_{p_1q_1k_1}(\xi_i)\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)(1-2\tanh^2_{p_2q_2k_2}(\eta_i))}{H(t)+\sum_{i=1}^{n}K_i(t)\tanh_{p_1q_1k_1}(\xi_i)\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)} \\ & \frac{\left(\sum_{i=1}^{n} K_i(t)\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)(p_1q_1k_1\alpha_i(t)\,\mathrm{sech}^2_{p_1q_1k_1}(\xi_i)-k_2\mu_i(t)\tanh_{p_1q_1k_1}(\xi_i)\tanh_{p_2q_2k_2}(\eta_i))\right)^2}{\left(H(t)+\sum_{i=1}^{n}K_i(t)\tanh_{p_1q_1k_1}(\xi_i)\,\mathrm{sech}_{p_2q_2k_2}(\eta_i)\right)^2} \\ & + v_2 \ , \end{aligned} \quad (69)$$

where $\xi_i = \alpha_i(t)x + \beta_i(t)y + r_i(t)$, $\eta_i = \mu_i(t)x + \nu_i(t)y + l_i(t)$.

For example, when we take $n = 2, H, K_i, p_i, q_i, k_i, \alpha_i, \beta_i$ and r_i, $i = 1, 2$ are all constants, then (65) becomes

$$\begin{aligned} f(x,y,t) &= H + K_1 \tanh_{p_1 q_1 k_1}(\alpha_1 x + \beta_1 y + r_1) \\ &\cdot \operatorname{sech}_{p_1 q_1 k_1}(\alpha_1 x + \beta_1 y + r_1) + K_2 \tanh_{p_2 q_2 k_2} \\ &\cdot (\alpha_2 x + \beta_2 y + r_2) \operatorname{sech}_{p_2 q_2 k_2}(\alpha_2 x + \beta_2 y + r_2)\,, \end{aligned} \tag{70}$$

With the help of symbolic computation, substituting (70) and $u_2(x,y,t) = u_2(t)$, $v_2(x,y,t) = v_2(t)$ into (52), (53), (55), (56) and (59), we can get $H, K_i, p_i, q_i, k_i, \alpha_i, \beta_i$ and r_i, $i = 1, 2$ respectively as follows:

Case 1

$$\begin{aligned} &u_2 = C_1\,, \\ &v_2 = \frac{C_1({\beta_2}^2{\alpha_1}^2 + {\beta_1}^2{\alpha_2}^2 + \beta_1\alpha_2\beta_2\alpha_1)}{3\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)}\,, \\ &p_1 = p_1\,, \quad p_2 = p_2\,, \quad q_1 = 0\,, \quad q_2 = 0\,, \\ &k_2 = \varepsilon\frac{\sqrt{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)C_1}(2\beta_1\alpha_2 + \beta_2\alpha_1)}{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)\alpha_2} \quad \text{or} \\ &k_2 = \varepsilon\frac{\sqrt{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)C_1}(2\beta_1\alpha_2 + \beta_2\alpha_1)}{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)\alpha_2}\,, \\ &k_1 = \varepsilon\frac{\sqrt{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)C_1}(\beta_1\alpha_2 + 2\beta_2\alpha_1)}{(\beta_1{\beta_2}^2\alpha_1 + {\beta_1}^2\beta_2\alpha_2)\alpha_1}\,, \\ &H = H\,, \quad \alpha_1 = \alpha_1\,, \quad \alpha_2 = \alpha_2\,, \\ &\beta_1 = \beta_1\,, \quad \beta_2 = \beta_2\,, \quad \gamma_1 = \gamma_1\,, \quad \gamma_2 = \gamma_2\,, \\ &K_1 = K_1\,, \quad K_2 = K_2\,. \end{aligned} \tag{71}$$

Case 2

$$\begin{aligned} &p_2 = p_2\,, \quad u_2 = C_1\,, \\ &v_2 = \frac{C_1({\beta_2}^2{\alpha_1}^2 + {\beta_1}^2{\alpha_2}^2 + \beta_1\alpha_2\beta_2\alpha_1)}{3\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)}\,, \\ &p_1 = p_1\,, \quad q_1 = 0\,, \\ &k_1 = \varepsilon\frac{\sqrt{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)C_1}(\beta_1\alpha_2 + 2\beta_2\alpha_1)}{(\beta_1{\beta_2}^2\alpha_1 + {\beta_1}^2\beta_2\alpha_2)\alpha_1}\,, \\ &q_2 = q_2\,, \quad k_2 = 0\,, \quad H = H\,, \quad \alpha_1 = \alpha_1\,, \quad \alpha_2 = \alpha_2\,, \\ &\beta_1 = \beta_1\,, \quad \beta_2 = \beta_2\,, \quad \gamma_1 = \gamma_1\,, \quad \gamma_2 = \gamma_2\,, \\ &K_1 = K_1\,, \quad K_2 = K_2\,. \end{aligned} \tag{72}$$

Case 3

$$\begin{aligned} &p_2 = 0\,, \quad u_2 = C_1\,, \\ &v_2 = \frac{C_1({\beta_2}^2{\alpha_1}^2 + {\beta_1}^2{\alpha_2}^2 + \beta_1\alpha_2\beta_2\alpha_1)}{3\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)}\,, \\ &p_1 = p_1\,, \quad q_1 = 0\,, \quad K_1 = K_1\,, \quad K_2 = K_2\,, \\ &k_1 = \varepsilon\frac{\sqrt{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)C_1}(\beta_1\alpha_2 + 2\beta_2\alpha_1)}{(\beta_1{\beta_2}^2\alpha_1 + {\beta_1}^2\beta_2\alpha_2)\alpha_1}\,, \\ &q_2 = q_2\,, \quad k_2 = 0\,, \quad H = H\,, \quad \alpha_1 = \alpha_1\,, \quad \alpha_2 = \alpha_2\,, \\ &\beta_1 = \beta_1\,, \quad \beta_2 = \beta_2\,, \quad \gamma_1 = \gamma_1\,, \quad \gamma_2 = \gamma_2\,. \end{aligned} \tag{73}$$

Case 4

$$\begin{aligned} &k_2 = \varepsilon\frac{\sqrt{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)C_1}(2\beta_1\alpha_2 + \beta_2\alpha_1)}{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)\alpha_2} \quad \text{or} \\ &k_2 = -\varepsilon\frac{\sqrt{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)C_1}(2\beta_1\alpha_2 + \beta_2\alpha_1)}{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)\alpha_2}\,, \\ &p_2 = 0\,, \quad u_2 = C_1\,, \\ &v_2 = \frac{C_1({\beta_2}^2{\alpha_1}^2 + {\beta_1}^2{\alpha_2}^2 + \beta_1\alpha_2\beta_2\alpha_1)}{3\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)}\,, \\ &p_1 = p_1\,, \quad q_1 = 0\,, \\ &k_1 = \varepsilon\frac{\sqrt{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)C_1}(\beta_1\alpha_2 + 2\beta_2\alpha_1)}{(\beta_1{\beta_2}^2\alpha_1 + {\beta_1}^2\beta_2\alpha_2)\alpha_1}\,, \\ &q_2 = q_2\,, \quad H = H\,, \quad \alpha_1 = \alpha_1\,, \quad \alpha_2 = \alpha_2\,, \\ &\beta_1 = \beta_1\,, \quad \beta_2 = \beta_2\,, \quad \gamma_1 = \gamma_1\,, \quad \gamma_2 = \gamma_2\,, \\ &K_1 = K_1\,, \quad K_2 = K_2\,. \end{aligned} \tag{74}$$

Case 5

$$\begin{aligned} &k_2 = \varepsilon\frac{\sqrt{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)C_1}(2\beta_1\alpha_2 + \beta_2\alpha_1)}{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)\alpha_2} \quad \text{or} \\ &k_2 = -\varepsilon\frac{\sqrt{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)C_1}(2\beta_1\alpha_2 + \beta_2\alpha_1)}{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)\alpha_2}\,, \\ &v_2 = \frac{C_1({\beta_2}^2{\alpha_1}^2 + {\beta_1}^2{\alpha_2}^2 + \beta_1\alpha_2\beta_2\alpha_1)}{3\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)}\,, \\ &k_1 = -\varepsilon\frac{\sqrt{\beta_1\beta_2(\beta_1\alpha_2 + \beta_2\alpha_1)C_1}(\beta_1\alpha_2 + 2\beta_2\alpha_1)}{(\beta_1{\beta_2}^2\alpha_1 + {\beta_1}^2\beta_2\alpha_2)\alpha_1}\,, \\ &H = H\,, \quad \alpha_1 = \alpha_1\,, \quad \alpha_2 = \alpha_2\,, \\ &\beta_1 = \beta_1\,, \quad \beta_2 = \beta_2\,, \quad \gamma_1 = \gamma_1\,, \quad \gamma_2 = \gamma_2\,, \\ &q_2 = 0\,, \quad p_1 = 0\,, \quad q_1 = q_1\,, \\ &u_2 = C_1\,, \quad p_2 = p_2\,, \\ &K_1 = K_1\,, \quad K_2 = K_2\,. \end{aligned} \tag{75}$$

Case 6

$$p_1 = 0\,, \quad k_2 = k_2\,, \;\; k_1 = k_1\,, \quad v_2 = v_2\,,$$
$$u_2 = C_1\,, \quad p_2 = p_2\,, \quad q_1 = 0\,, \;\; q_2 = q_2\,,$$
$$H = H\,, \;\; \alpha_1 = \alpha_1\,, \;\; \alpha_2 = \alpha_2\,, \tag{76}$$
$$\beta_1 = \beta_1\,, \;\; \beta_2 = \beta_2\,, \;\; \gamma_1 = \gamma_1\,, \;\; \gamma_2 = \gamma_2\,,$$
$$K_1 = K_1\,, \;\; K_2 = K_2\,.$$

Case 7

$$p_2 = 0\,, \quad q_2 = 0\,, \quad p_1 = 0\,, \;\; q_1 = q_1\,, \quad v_2 = C_1\,,$$
$$k_2 = \varepsilon\frac{\sqrt{3C_1}}{\alpha_2} \;\text{ or }\; k_2 = -\varepsilon\frac{\sqrt{3C_1}}{\alpha_2}\,, \quad k_1 = \varepsilon\frac{\sqrt{3C_1}}{\alpha_1}\,,$$
$$u_2 = 0\,, \;\; H = H\,, \;\; \alpha_1 = \alpha_1\,, \;\; \alpha_2 = \alpha_2\,,$$
$$\beta_1 = \beta_1\,, \;\; \beta_2 = \beta_2\,, \;\; \gamma_1 = \gamma_1\,, \;\; \gamma_2 = \gamma_2\,,$$
$$K_1 = K_1\,, \;\; K_2 = K_2\,. \tag{77}$$

Case 8

$$v_2 = C_2\,, \quad p_2 = 0\,, \quad k_1 = \varepsilon\frac{\sqrt{3\beta_1(C_2\beta_1 + C_1\alpha_1)}}{\beta_1\alpha_1} \;\text{ or}$$
$$k_1 = -\varepsilon\frac{\sqrt{3\beta_1(C_2\beta_1 + C_1\alpha_1)}}{\beta_1\alpha_1}\,, \;\; q_1 = q_1\,, \quad u_2 = C_1\,,$$
$$k_2 = \varepsilon\frac{\sqrt{3\beta_2(C_2\beta_2 + C_1\alpha_2)}}{\beta_2\alpha_2}\,, \quad q_2 = 0\,, \quad p_1 = 0\,,$$
$$H = H\,, \;\; \alpha_1 = \alpha_1\,, \;\; \alpha_2 = \alpha_2\,,$$
$$\beta_1 = \beta_1\,, \;\; \beta_2 = \beta_2\,, \;\; \gamma_1 = \gamma_1\,, \;\; \gamma_2 = \gamma_2\,,$$
$$K_1 = K_1\,, \;\; K_2 = K_2\,. \tag{78}$$

Case 9

$$v_2 = C_2\,, \quad p_2 = 0\,, \;\; p_1 = 0\,, \quad q_1 = q_1\,,$$
$$k_1 = 0\,, \;\; k_2 = \varepsilon\frac{\sqrt{3\beta_2(C_2\beta_2 + C_1\alpha_2)}}{\beta_2\alpha_2}\,,$$
$$u_2 = C_1\,, \quad q_2 = q_2\,, \tag{79}$$
$$H = H\,, \;\; \alpha_1 = \alpha_1\,, \;\; \alpha_2 = \alpha_2\,,$$
$$\beta_1 = \beta_1\,, \;\; \beta_2 = \beta_2\,, \;\; \gamma_1 = \gamma_1\,, \;\; \gamma_2 = \gamma_2\,,$$
$$K_1 = K_1\,, \;\; K_2 = K_2\,.$$

Case 10

$$v_2 = C_2\,, \quad p_2 = 0\,, \quad k_1 = \varepsilon\frac{\sqrt{3\beta_1(C_2\beta_1 + C_1\alpha_1)}}{\beta_1\alpha_1}\,,$$
$$p_1 = 0\,, \quad q_1 = q_1\,, \quad k_2 = 0\,, \quad u_2 = C_1\,,$$
$$q_2 = q_2\,, \;\; H = H\,, \;\; \alpha_1 = \alpha_1\,, \;\; \alpha_2 = \alpha_2\,,$$
$$\beta_1 = \beta_1\,, \;\; \beta_2 = \beta_2\,, \;\; \gamma_1 = \gamma_1\,, \;\; \gamma_2 = \gamma_2\,,$$
$$K_1 = K_1\,, \;\; K_2 = K_2\,. \tag{80}$$

Case 11

$$p_2 = p_2\,, \quad q_2 = 0\,, \quad p_1 = 0\,, \quad q_1 = q_1\,,$$
$$k_2 = \varepsilon\frac{\sqrt{3C_1}}{\alpha_2}\,, \quad v_2 = C_1 \;\text{ or }\; k_2 = -\varepsilon\frac{\sqrt{3C_1}}{\alpha_2}\,,$$
$$v_2 = C_1\,, \;\; k_1 = \varepsilon\frac{\sqrt{3C_1}}{\alpha_1}\,, \quad u_2 = 0\,, \tag{81}$$
$$H = H\,, \;\; \alpha_1 = \alpha_1\,, \;\; \alpha_2 = \alpha_2\,,$$
$$\beta_1 = \beta_1\,, \;\; \beta_2 = \beta_2\,, \;\; \gamma_1 = \gamma_1\,, \;\; \gamma_2 = \gamma_2\,,$$
$$K_1 = K_1\,, \;\; K_2 = K_2.$$

Case 12

$$p_2 = p_2\,, \quad q_2 = 0\,, \quad p_1 = 0\,, \quad q_1 = q_1\,,$$
$$k_2 = \varepsilon\frac{\sqrt{3C_1}}{\alpha_2} \;\text{ or }\; k_2 = -\varepsilon\frac{\sqrt{3C_1}}{\alpha_2}\,, \quad v_2 = C_1\,,$$
$$k_1 = \varepsilon\frac{\sqrt{3C_1}}{\alpha_1}\,, \tag{82}$$
$$u_2 = 0\,, \;\; K_1 = K_1\,, \;\; K_2 = K_2\,,$$
$$H = H\,, \;\; \alpha_1 = \alpha_1\,, \;\; \alpha_2 = \alpha_2\,,$$
$$\beta_1 = \beta_1\,, \;\; \beta_2 = \beta_2\,, \;\; \gamma_1 = \gamma_1\,, \;\; \gamma_2 = \gamma_2\,.$$

Case 13

$$v_2 = C_2\,, \quad p_2 = 0\,, \quad k_1 = \varepsilon\frac{\sqrt{3\beta_1(C_2\beta_1 + C_1\alpha_1)}}{\beta_1\alpha_1}\,,$$
$$q_2 = 0\,, \quad p_1 = 0\,, \quad q_1 = q_1\,, \quad k_2 = k_2\,, \quad u_2 = C_1\,,$$
$$H = H\,, \;\; \alpha_1 = \alpha_1\,, \;\; \alpha_2 = \alpha_2\,,$$
$$\beta_1 = \beta_1\,, \;\; \beta_2 = \beta_2\,, \;\; \gamma_1 = \gamma_1\,, \;\; \gamma_2 = \gamma_2\,,$$
$$K_1 = K_1\,, \;\; K_2 = K_2\,. \tag{83}$$

Case 14

$$p_2 = 0\,, \quad v_2 = \varepsilon\frac{C_1\alpha_2}{\beta_2}\,, \quad k_1 = k_1\,, \quad q_2 = 0\,,$$
$$p_1 = 0\,, \quad q_1 = q_1\,, \quad k_2 = 0\,, \quad u_2 = C_1\,,$$
$$H = H\,, \;\; \alpha_1 = \alpha_1\,, \;\; \alpha_2 = \alpha_2\,, \tag{84}$$
$$\beta_1 = \beta_1\,, \;\; \beta_2 = \beta_2\,, \;\; \gamma_1 = \gamma_1\,, \;\; \gamma_2 = \gamma_2\,,$$
$$K_1 = K_1\,, \;\; K_2 = K_2\,.$$

Case 15

$$v_2 = C_2\,, \quad p_2 = 0\,, \quad k_2 = \varepsilon\frac{\sqrt{3\beta_2(C_2\beta_2 + C_1\alpha_2)}}{\beta_2\alpha_2}\,,$$
$$k_1 = k_1\,, \quad q_2 = 0\,, \quad p_1 = 0\,, \quad q_1 = q_1\,,$$
$$u_2 = C_1\,, \;\; H = H\,, \;\; \alpha_1 = \alpha_1\,, \;\; \alpha_2 = \alpha_2\,,$$
$$\beta_1 = \beta_1\,, \;\; \beta_2 = \beta_2\,, \;\; \gamma_1 = \gamma_1\,, \;\; \gamma_2 = \gamma_2\,,$$
$$K_1 = K_1\,, \;\; K_2 = K_2\,. \tag{85}$$

Case 16

$$\begin{aligned}&p_2 = 0\,, \quad k_1 = k_1\,, \quad q_2 = 0\,, \quad p_1 = 0\,,\\ &q_1 = q_1\,, \quad v_2 = v_2\,, \quad k_2 = k_2\,, \quad u_2 = C_1\,,\\ &H = H\,, \ \alpha_1 = \alpha_1\,, \ \alpha_2 = \alpha_2\,,\\ &\beta_1 = \beta_1\,, \ \beta_2 = \beta_2\,, \ \gamma_1 = \gamma_1\,, \ \gamma_2 = \gamma_2\,,\\ &K_1 = K_1\,, \ K_2 = K_2\,.\end{aligned} \tag{86}$$

Case 17

$$\begin{aligned}&p_2 = 0\,, \quad v_2 = \varepsilon\frac{C_1\alpha_2}{\beta_2} \ \text{or} \ v_2 = -\varepsilon\frac{C_1\alpha_2}{\beta_2}\,,\\ &k_2 = 0\,, \ k_1 = \varepsilon\frac{\sqrt{-3\beta_2\beta_1 C_1(\beta_1\alpha_2 - \beta_2\alpha_1)}}{\beta_2\beta_1\alpha_1}\,,\\ &q_1 = 0\,, \quad p_1 = p_1\,, \ u_2 = C_1\,, \ q_2 = q_2\,,\\ &H = H\,, \ \alpha_1 = \alpha_1\,, \ \alpha_2 = \alpha_2\,,\\ &\beta_1 = \beta_1\,, \ \beta_2 = \beta_2\,, \ \gamma_1 = \gamma_1\,, \ \gamma_2 = \gamma_2\,,\\ &K_1 = K_1\,, \ K_2 = K_2\,.\end{aligned} \tag{87}$$

Case 18

$$\begin{aligned}&p_2 = p_2\,, \quad v_2 = \varepsilon\frac{C_1\alpha_2}{\beta_2} \ \text{or} \ v_2 = -\varepsilon\frac{C_1\alpha_2}{\beta_2}\,,\\ &k_2 = 0\,, \ k_1 = \frac{\sqrt{-3\beta_2\beta_1 C_1}(\beta_1\alpha_2 - \beta_2\alpha_1)}{\beta_2\beta_1\alpha_1}\,,\\ &q_1 = 0\,, \quad p_1 = p_1\,, \ u_2 = C_1\,, \quad q_2 = q_2\,,\\ &H = H\,, \ \alpha_1 = \alpha_1\,, \ \alpha_2 = \alpha_2\,,\\ &\beta_1 = \beta_1\,, \ \beta_2 = \beta_2\,, \ \gamma_1 = \gamma_1\,, \ \gamma_2 = \gamma_2\,,\\ &K_1 = K_1\,, \ K_2 = K_2\,.\end{aligned} \tag{88}$$

Case 19

$$\begin{aligned}&p_2 = 0\,, \quad q_1 = q_1\,, \quad k_1 = 0\,, \quad v_2 = \varepsilon\frac{C_1\alpha_1}{\beta_1} \ \text{or}\\ &v_2 = -\varepsilon\frac{C_1\alpha_1}{\beta_1}\,,\\ &k_2 = -\varepsilon\frac{\sqrt{3\beta_2\beta_1 C_1(\beta_1\alpha_2 - \beta_2\alpha_1)}}{\beta_2\beta_1\alpha_2}\,, \quad p_1 = p_1\,,\\ &u_2 = C_1\,, \quad q_2 = q_2\,,\\ &H = H\,, \ \alpha_1 = \alpha_1\,, \ \alpha_2 = \alpha_2\,,\\ &\beta_1 = \beta_1\,, \ \beta_2 = \beta_2\,, \ \gamma_1 = \gamma_1\,, \ \gamma_2 = \gamma_2\,,\\ &K_1 = K_1\,, \ K_2 = K_2\,.\end{aligned} \tag{89}$$

Case 20

$$\begin{aligned}&p_2 = 0\,, \quad v_2 = C_2\,, \quad k_2 = \varepsilon\frac{\sqrt{3\beta_2(C_2\beta_2 + C_1\alpha_2)}}{\beta_2\alpha_2}\\ &\text{or} \ k_2 = -\varepsilon\frac{\sqrt{3\beta_2(C_2\beta_2 + C_1\alpha_2)}}{\beta_2\alpha_2}\,,\\ &k_1 = \varepsilon\frac{\sqrt{3\beta_1(C_2\beta_1 + C_1\alpha_1)}}{\beta_1\alpha_1}\,, \quad q_2 = 0\,, \ q_1 = 0\,,\\ &p_1 = p_1\,, \quad u_2 = C_1\,,\\ &H = H\,, \ \alpha_1 = \alpha_1\,, \ \alpha_2 = \alpha_2\,,\\ &\beta_1 = \beta_1\,, \ \beta_2 = \beta_2\,, \ \gamma_1 = \gamma_1\,, \ \gamma_2 = \gamma_2\,,\\ &K_1 = K_1\,, \ K_2 = K_2\,.\end{aligned} \tag{90}$$

Case 21

$$\begin{aligned}&v_2 = C_1\,, \quad k_2 = \varepsilon\frac{\sqrt{3C_1}}{\alpha_2} \ \text{or} \ k_2 = -\varepsilon\frac{\sqrt{3C_1}}{\alpha_2}\,,\\ &k_1 = -\varepsilon\frac{\sqrt{3C_1}}{\alpha_1}\,, \quad u_2 = 0\,, \quad p_2 = 0\,, \quad q_1 = 0\,,\\ &p_1 = p_1\,, q_2 = q_2\,,\\ &H = H\,, \ \alpha_1 = \alpha_1\,, \ \alpha_2 = \alpha_2\,,\\ &\beta_1 = \beta_1\,, \ \beta_2 = \beta_2\,, \ \gamma_1 = \gamma_1\,, \ \gamma_2 = \gamma_2\,,\\ &K_1 = K_1\,, \ K_2 = K_2\,.\end{aligned} \tag{91}$$

Case 22

$$\begin{aligned}&p_2 = 0\,, \quad q_1 = q_1\,, \quad k_2 = \varepsilon\frac{\sqrt{3\beta_2(C_2\beta_2 + C_1\alpha_2)}}{\beta_2\alpha_2}\\ &\text{or} \ k_2 = -\varepsilon\frac{\sqrt{3\beta_2(C_2\beta_2 + C_1\alpha_2)}}{\beta_2\alpha_2}\,,\\ &k_1 = \varepsilon\frac{\sqrt{3}\sqrt{\beta_1(C_2\beta_1 + C_1\alpha_1)}}{\beta_1\alpha_1}\,, \quad v_2 = C_2\,,\\ &q_2 = 0\,, \quad p_1 = p_1\,, \quad u_2 = C_1\,,\\ &H = H\,, \ \alpha_1 = \alpha_1\,, \ \alpha_2 = \alpha_2\,,\\ &\beta_1 = \beta_1\,, \ \beta_2 = \beta_2\,, \ \gamma_1 = \gamma_1\,, \ \gamma_2 = \gamma_2\,,\\ &K_1 = K_1\,, \ K_2 = K_2\,.\end{aligned} \tag{92}$$

Case 23

$$\begin{aligned}&p_2 = 0\,, \quad q_1 = q_1\,, \quad k_2 = \varepsilon\frac{\sqrt{3\beta_2(C_2\beta_2 + C_1\alpha_2)}}{\beta_2\alpha_2}\,,\\ &v_2 = C_2\,, \quad k_1 = 0\,, \quad p_1 = p_1\,, \quad u_2 = C_1\,,\\ &q_2 = q_2\,, \ H = H\,, \ \alpha_1 = \alpha_1\,, \ \alpha_2 = \alpha_2\,,\\ &\beta_1 = \beta_1\,, \ \beta_2 = \beta_2\,, \ \gamma_1 = \gamma_1\,, \ \gamma_2 = \gamma_2\,,\\ &K_1 = K_1\,, \ K_2 = K_2\,.\end{aligned} \tag{93}$$

Case 24

$$\begin{aligned}
&p_2 = p_2\,, \quad v_2 = C_1\,, \quad k_1 = \varepsilon\frac{\sqrt{3C_1}}{\alpha_1} \quad \text{or}\\
&k_1 = -\varepsilon\frac{\sqrt{3C_1}}{\alpha_1}\,, \quad u_2 = 0\,, \quad q_2 = 0\,, \quad q_1 = 0\,,\\
&p_1 = p_1\,, \quad k_2 = -\varepsilon\frac{\sqrt{3C_1}}{\alpha_2}\,,\\
&H = H\,, \ \alpha_1 = \alpha_1\,, \ \alpha_2 = \alpha_2\,,\\
&\beta_1 = \beta_1\,, \ \beta_2 = \beta_2\,, \ \gamma_1 = \gamma_1\,, \ \gamma_2 = \gamma_2\,,\\
&K_1 = K_1\,, \ K_2 = K_2\,.
\end{aligned} \tag{94}$$

Case 25

$$\begin{aligned}
&p_2 = 0\,, \quad k_2 = k_2\,, \ k_1 = k_1\,, \quad v_2 = v_2\,,\\
&q_1 = q_1\,, \quad q_2 = 0\,, \quad p_1 = p_1\,, \quad u_2 = C_1\,,\\
&H = H\,, \ \alpha_1 = \alpha_1\,, \ \alpha_2 = \alpha_2\,,\\
&\beta_1 = \beta_1\,, \ \beta_2 = \beta_2\,, \ \gamma_1 = \gamma_1\,, \ \gamma_2 = \gamma_2\,,\\
&K_1 = K_1\,, \ K_2 = K_2\,.
\end{aligned} \tag{95}$$

We get the following new solution (called generalized hyperbolic function solution) of Eq. (46):

$$\begin{aligned}
u_1 = {}& 2\frac{\begin{array}{c}K_1k_1\beta_1(t)\operatorname{sech}_{p_1q_1k_1}(\xi)(1-2p_1q_1\\ \operatorname{sech}^2_{p_1q_1k_1}(\xi)) + K_2k_2\beta_2(t)\operatorname{sech}_{p_2q_2k_2}(\eta)\\ (1-p_2q_2(1+p_2q_2)\operatorname{sech}^2_{p_2q_2k_2}(\eta))\end{array}}{\begin{array}{c}H + K_1\tanh_{p_1q_1k_1}(\xi)\operatorname{sech}_{p_1q_1k_1}(\xi)\\ +K_2\tanh_{p_2q_2k_2}(\eta)\operatorname{sech}_{p_2q_2k_2}(\eta)\end{array}}\\
&\frac{\begin{array}{c}K_1k_1\alpha_1(t)\operatorname{sech}_{p_1q_1k_1}(\xi)(1-2p_1q_1\operatorname{sech}^2_{p_1q_1k_1}(\xi))\\ +K_2k_2\alpha_2(t)\operatorname{sech}_{p_2q_2k_2}(\eta)(1-p_2q_2(1+p_2q_2)\\ \operatorname{sech}^2_{p_2q_2k_2}(\eta))\end{array}}{\begin{array}{c}H + K_1\tanh_{p_1q_1k_1}(\xi)\operatorname{sech}_{p_1q_1k_1}(\xi)\\ +K_2\tanh_{p_2q_2k_2}(\eta)\operatorname{sech}_{p_2q_2k_2}(\eta)\end{array}}\\
&-2\frac{\begin{array}{c}{k_1}^2K_1\alpha_1(t)\beta_1(t)\tanh_{p_1q_1k_1}(\xi)\operatorname{sech}_{p_1q_1k_1}(\xi)\\ (1-6p_1q_1\operatorname{sech}^2_{p_1q_1k_1}(\xi))\end{array}}{\begin{array}{c}H + K_1\tanh_{p_1q_1k_1}(\xi)\operatorname{sech}_{p_1q_1k_1}(\xi)\\ +K_2\tanh_{p_2q_2k_2}(\eta)\operatorname{sech}_{p_2q_2k_2}(\eta)\end{array}}\\
&-2\frac{\begin{array}{c}{k_2}^2K_2\alpha_2(t)\beta_2(t)\tanh_{p_2q_2k_2}(\eta)\operatorname{sech}_{p_2q_2k_2}(\eta)\\ (1-4p_2q_2(1+p_2q_2)\operatorname{sech}^2_{p_2q_2k_2}(\eta))\end{array}}{\begin{array}{c}H + K_1\tanh_{p_1q_1k_1}(\xi)\operatorname{sech}_{p_1q_1k_1}(\xi)\\ +K_2\tanh_{p_2q_2k_2}(\eta)\operatorname{sech}_{p_2q_2k_2}(\eta)\end{array}} + u_2\,,
\end{aligned} \tag{96}$$

$$\begin{aligned}
v_1 = {}& 2\frac{\begin{array}{c}(K_1k_1\alpha_1(t)\operatorname{sech}_{p_1q_1k_1}(\xi)(1-2p_1q_1\\ \operatorname{sech}^2_{p_1q_1k_1}(\xi)) + K_2k_2\alpha_2(t)\operatorname{sech}_{p_2q_2k_2}(\eta)\\ (1-p_2q_2(1+p_2q_2)\operatorname{sech}^2_{p_2q_2k_2}(\eta)))^2\end{array}}{\begin{array}{c}(H + K_1\tanh_{p_1q_1k_1}(\xi)\operatorname{sech}_{p_1q_1k_1}(\xi)\\ +K_2\tanh_{p_2q_2k_2}(\eta)\operatorname{sech}_{p_2q_2k_2}(\eta))^2\end{array}}\\
&-2\frac{\begin{array}{c}{k_1}^2K_1(\alpha_1(t))^2\tanh_{p_1q_1k_1}(\xi)\operatorname{sech}_{p_1q_1k_1}(\xi)\\ (1-6p_1q_1\operatorname{sech}^2_{p_1q_1k_1}(\xi))\end{array}}{\begin{array}{c}H + K_1\tanh_{p_1q_1k_1}(\xi)\operatorname{sech}_{p_1q_1k_1}(\xi)\\ +K_2\tanh_{p_2q_2k_2}(\eta)\operatorname{sech}_{p_2q_2k_2}(\eta)\end{array}}\\
&-2\frac{\begin{array}{c}K_2{k_2}^2(\alpha_2(t))^2\tanh_{p_2q_2k_2}(\eta)\operatorname{sech}_{p_2q_2k_2}(\eta)\\ (1-2p_2q_2(2+p_2q_2)\operatorname{sech}^2_{p_2q_2k_2}(\eta))\end{array}}{\begin{array}{c}H + K_1\tanh_{p_1q_1k_1}(\xi)\operatorname{sech}_{p_1q_1k_1}(\xi)\\ +K_2\tanh_{p_2q_2k_2}(\eta)\operatorname{sech}_{p_2q_2k_2}(\eta)\end{array}} + v_2\,,
\end{aligned} \tag{97}$$

where $\xi = \alpha_1(t)x + \beta_1(t)y + r_1(t)$, $\eta = \alpha_2(t)x + \beta_2(t)y + r_2(t)$, here $H, K_i, p_i, q_i, k_i, \alpha_i, \beta_i$ and $r_i, i = 1,2$ satisfy (71)–(95), respectively.

The Generalized F-expansion Method and Its Application in Another (2 + 1)-Dimensional KdV Equation

In the section, we will introduce the main steps of the generalized F-expansion method for constructing exact solutions of NLPDEs which we first presented in [40,41,42]. Then we will make use of the method to find new exact solutions of the (2 + 1)-dimensional KdV equation.

Summary of the Generalized F-expansion Method

In the following we would like to outline the main steps of our general method (called the generalized F-expansion method) in [40,41,42].

Step 1 For a given nonlinear partial differential equation system with some physical fields $u_i(t, x_1, x_2, \ldots, x_m)$, $(i = 1, 2, \ldots, n)$ in $m + 1$ independent variables $t, x_1, x_2, \ldots, x_m$,

$$\begin{cases}
F_1(u_1, \ldots, u_n, u_{1t}, \ldots, u_{nt}, u_{1x_1}, \ldots, u_{nx_m},\\
\qquad u_{1tx_1}, \ldots, u_{ntx_m}, \ldots) = 0\,,\\
F_2(u_1, \ldots, u_n, u_{1t}, \ldots, u_{nt}, u_{1x_1}, \ldots, u_{nx_m},\\
\qquad u_{1tx_1}, \ldots, u_{ntx_m}, \ldots) = 0\,,\\
\ldots\ldots\ldots\ldots,\\
F_n(u_1, \ldots, u_n, u_{1t}, \ldots, u_{nt}, u_{1x_1}, \ldots, u_{nx_m},\\
\qquad u_{1tx_1}, \ldots, u_{ntx_m}, \ldots) = 0\,.
\end{cases} \tag{98}$$

We first consider the following general formal solutions of the Eqs. (98)

$$u_i(t, x_1, x_2, \ldots, x_m) = u_i(\omega), \quad \omega = \omega(x), \quad (i = 1, 2, \ldots, n), \tag{99}$$

where $x = (x_1, x_2, x_3, \ldots, x_m, t)$ and $\omega(x)$ is a function to be determined later. For example, when $n = 2$, we may take $\omega(x) = p(x_2, t)x_1 + q(x_2, t)$ for convenience, here $p(x_2, t)$ and $q(x_2, t)$ are functions to be determined later.

Step 2 We introduce new and more general formal transformations of the Eq. (98), if available, in the forms which we first presented in [42]:

$$\begin{cases} u_1(x) = a_{1_0}(x) + \sum_{i_1=1}^{k_1} (f(\omega(x)))^{i_1-1}\Big(a_{i_1}(x) \\ \quad f(\omega(x)) + b_{i_1}(x)g(\omega(x)) + \frac{c_{i_1}(x)}{f(\omega(x))} + \frac{d_{i_1}(x)}{g(\omega(x))}\Big), \\ u_2(x) = a_{2_0}(x) + \sum_{i_2=1}^{k_2} (f(\omega(x)))^{i_2-1}\Big(a_{i_2}(x) \\ \quad f(\omega(x)) + b_{i_2}(x)g(\omega(x)) + \frac{c_{i_2}(x)}{f(\omega(x))} + \frac{d_{i_2}(x)}{g(\omega(x))}\Big), \\ \ldots\ldots\ldots\ldots, \\ u_n(x) = a_{n_0}(x) + \sum_{i_n=1}^{k_n} (f(\omega(x)))^{i_n-1}\Big(a_{i_n}(x) \\ \quad f(\omega(x)) + b_{i_n}(x)g(\omega(x)) + \frac{c_{i_n}(x)}{f(\omega(x))} + \frac{d_{i_n}(x)}{g(\omega(x))}\Big). \end{cases} \tag{100}$$

where $a_{i_j}^2(x) + b_{i_j}^2(x) + c_{i_j}^2(x) + d_{i_j}^2(x) \neq 0$, $a_{j_0}(x), a_{i_j}(x), b_{i_j}(x), c_{i_j}(x), d_{i_j}(x) (j = 1, 2, \ldots, n, i_j = 1, 2, \ldots, k_j)$ and $\omega(x)$ are all functions to be determined later, $k_j (j = 1, 2, \ldots, n)$ is an integer which is determined by balancing the highest order derivative terms with the nonlinear terms in the given Eqs. (98), and the new variables $f(\omega) = f(\omega(x))$ and $g(\omega) = g(\omega(x))$ satisfy the following relations:

$$\begin{cases} f'^2(\omega) = l_1 f^4(\omega) + m_1 f^2(\omega) + n_1 \\ g'^2(\omega) = l_2 g^4(\omega) + m_2 g^2(\omega) + n_2 \\ g^2(\omega) = \frac{l_1 f^2(\omega)}{l_2} + \frac{m_1 - m_2}{3l_2} \\ n_1 = \frac{{m_1}^2 - {m_2}^2 + 3l_2 n_2}{3l_1} \end{cases} \tag{101}$$

Step 3 Determine $k_j (j = 1, 2, \ldots, n)$ in (100) by balancing the highest nonlinear terms and the highest-order partial differential terms in the given Eq. (98). If k_j is a nonnegative integer, then we first make the transformation $u_j(x) = v_j^{k_j}(x)$.

Step 4 Substitute (100) into the Eq. (98) with (101) and collecting coefficients of polynomials of $f(\omega), g(\omega)$, $\sqrt{l_1 f^4(\omega) + m_1 f^2(\omega) + n_1}$, with the aid of Maple, then setting each coefficient to zero to get a set of over-determined partial differential equations with respect to $l_1, l_2, m_1, m_2, n_1, n_2, a_{j_0}(x), a_{i_j}(x), b_{i_j}(x), c_{i_j}(x), d_{i_j}(x) (j = 1, 2, \ldots, n, i_j = 1, 2, \ldots, k_j)$ and $\omega(x)$.

Step 5 Solving the over-determined partial differential equations with Maple, we then can determine $a_{j_0}(x), a_{i_j}(x), b_{i_j}(x), c_{i_j}(x), d_{i_j}(x) (j = 1, 2, \ldots, n, i_j = 1, 2, \ldots, k_j)$ and $\omega(x)$.

Step 6 Selecting the proper value of parameters $l_1, l_2, m_1, m_2, n_1, n_2$ to determine the corresponding the solutions of the Jacobi elliptic functions $f(\omega)$ and $g(\omega)$ in (101). The relations between the parameters and their corresponding Jacobi elliptic functions are known and are given in the following Table 1.

Remark 2 Relations between values of $l_1, m_1, n_1, l_2, m_2, n_2$ and corresponding $f(\omega), g(\omega)$ in ODEs (101) are in the following Table 1.

Step 7 By using the results obtained in the above step, we can derive a series of generalized solutions, such as different kinds of Jacobi elliptic function solutions in terms of Remark 2. Finally, substituting $a_{j_0}(x), a_{i_j}(x), b_{i_j}(x), c_{i_j}(x), d_{i_j}(x) (j = 1, 2, \ldots, n, i_j = 1, 2, \ldots, k_j)$ and $\omega(x)$ into the generalized solutions with the corresponding solutions of $f(\omega), g(\omega)$, we can get the the new exact solutions of the given Eq. (98).

Remark 3 As is known to all, when $k \to 1$, the Jacobi elliptic functions can degenerate as hyperbolic functions, and when $k \to 0$, the Jacobi elliptic functions degenerate as trigonometric functions. So by this method we can obtain many other new exact solutions to Eq. (98).

Remark 4 The main advantages of the generalized F-expansion method are simpler, more powerful, and more convenient than the Jacobi elliptic function expansion method and the F-expansion method. First of all, with the aid of nonlinear ordinary differential equations (ODEs) (101), one only needs to calculate the functions $f(\omega)$ and $g(\omega)$ which are the solutions of the ODEs (101), instead of calculating the Jacobi elliptic functions one by one. Secondly, the values of the coefficients $l_1, m_1, n_1, l_2, m_2, n_2$ of the ODEs (101) can be selected so that the corresponding solutions of the coupled functions $f(\omega)$ and $g(\omega)$ are Jacobi elliptic functions in the above Table 1. The relations between the coefficients and the corresponding Jacobi elliptic function solutions are known and are given in Remark 2. Thus, in terms of Remark 2, one can simul-

Korteweg–de Vries Equation (KdV), Different Analytical Methods for Solving the, Table 1
The ODEs (101) and Jacobi Elliptic Functions Relation between values of l_i, m_i, n_i $(i = 1, 2)$ and corresponding $f(\omega)$ and $g(\omega)$ in ODEs (101)

l_1	m_1	n_1	l_2	m_2	n_2	$f(\omega)$	$g(\omega)$
k^2	$-1-k^2$	1	$-k^2$	$2k^2-1$	$1-k^2$	$sn(\omega)$	$cn(\omega)$
k^2	$-1-k^2$	1	-1	$2-k^2$	$-1+k^2$	$sn(\omega)$	$dn(\omega)$
$-k^2$	$2k^2-1$	$1-k^2$	-1	$2-k^2$	k^2-1	$cn(\omega)$	$dn(\omega)$
$1-k^2$	$2k^2-1$	$-k^2$	$1-k^2$	$2-k^2$	1	$nc(\omega)$	$sc(\omega)$
1	$2-k^2$	$1-k^2$	1	$2k^2-1$	$-k^2(1-k^2)$	$cs(\omega)$	$ds(\omega)$
1	$-1-k^2$	k^2	$1-k^2$	$2k^2-1$	$-k^2$	$dc(\omega)$	$nc(\omega)$
1	$-1-k^2$	k^2	1	$2-k^2$	$1-k^2$	$ns(\omega)$	$cs(\omega)$
1	$-1-k^2$	k^2	1	$2k^2-1$	$-k^2(1-k^2)$	$ns(\omega)$	$ds(\omega)$
k^2	$-1-k^2$	1	$-1+k^2$	$2-k^2$	-1	$cd(\omega)$	$nd(\omega)$
$-k^2(1-k^2)$	$2k^2-1$	1	$-1+k^2$	$2-k^2$	-1	$sd(\omega)$	$nd(\omega)$
$-k^2(1-k^2)$	$2k^2-1$	1	k^2	$-1-k^2$	1	$sd(\omega)$	$cd(\omega)$
$1-k^2$	$2-k^2$	1	1	$-1-k^2$	k^2	$sc(\omega)$	$dc(\omega)$

taneously obtain more periodic wave solutions expressed by various Jacobi elliptic functions. Thirdly, we present a new transformations (100) that are more general than the transformations of the Jacobi elliptic function expansion method and the F-expansion method.

The Generalized F-expansion Method to Find the Exact Solutions of the $(2+1)$-Dimensional KdV Equation

In this section, we will make use of our method [40,41,42] and symbolic computation to find new exact solutions of the $(2+1)$-dimension KdV equation.

The $(2+1)$-dimension KdV equation in [20] is written in the following form

$$u_t(x,y,t) - 4u(x,y,t)u_y(x,y,t) - 4u_x(x,y,t)\partial_x^{-1} u_y(x,y,t) - u_{xxy}(x,y,t) = 0\,. \quad (102)$$

where ∂_x^{-1} is an indefinite integrate operator $\int \mathrm{d}x$, possessing some interesting coherent structures.

We discuss the similar solution of (102) in its potential form $(u = v_x)$,

$$v_{xt}(x,y,t) - v_{xxxy}(x,y,t) - 4v_x(x,y,t)v_{xy}(x,y,t) - 4v_{xx}(x,y,t)v_y(x,y,t) = 0\,. \quad (103)$$

By balancing the highest nonlinear terms and the highest-order partial derivative terms in (103), we suppose (103) to have the following formal solution:

$$v = c_0(y,t) + c_1(y,t)g(\omega) + c_2(y,t)f(\omega) + \frac{c_3(y,t)}{g(\omega)} + \frac{c_4(y,t)}{f(\omega)}\,, \quad (104)$$

where $\omega = \alpha(y,t)x + p(y,t) + q(t), \alpha(y,t), p(y,t), q(t)$, $c_0(y,t)$, $c_1(y,t)$, $c_2(y,t)$, $c_3(y,t)$ and $c_4(y,t)$ are all functions to be determined later. And the new variables $f(\omega)$ and $g(\omega)$ satisfy the relation (101).

Substitute (104) into the Eq. (103) with (101). Collect coefficients of polynomials of $f(\omega), g(\omega)$, $\sqrt{l_1 f^4(\omega) + m_1 f^2(\omega) + n_1}$, with the aid of Maple, then set each coefficient to zero to get a set of over-determined partial differential equations with respect to l_1, l_2, m_1, m_2, $n_1, n_2, \alpha(y,t), p(y,t), c_i(y,t), (i = 0,1,2,3,4)$ and $q(t)$ as follows:

$$\begin{aligned}
&120(\alpha(y,t))^2 c_4(y,t)c_1(y,t)\left(\frac{\partial}{\partial y}\alpha(y,t)\right) x {m_1}^4 {m_2}^2 \\
&+ 168(\alpha(y,t))^2 c_4(y,t) l_2 c_3(y,t)\left(\frac{\partial}{\partial y}p(y,t)\right) {m_1}^4 m_2 \\
&+ 120(\alpha(y,t))^2 c_4(y,t)c_1(y,t)(\frac{\partial}{\partial y}p(y,t)) {m_2}^2 {m_1}^4 \\
&+ 48(\alpha(y,t))^2 c_4(y,t) l_2 c_3(y,t)\left(\frac{\partial}{\partial y}\alpha(y,t)\right) x {m_1}^2 {m_2}^3 \\
&- 80(\alpha(y,t))^2 c_4(y,t)c_1(y,t)(\frac{\partial}{\partial y}\alpha(y,t)) x {m_2}^3 {m_1}^3 \\
&- 72(\alpha(y,t))^2 c_4(y,t)c_1(y,t)\left(\frac{\partial}{\partial y}p(y,t)\right) {m_1}^5 m_2 \\
&+ 24(\alpha(y,t))^2 c_4(y,t)c_1(y,t)\left(\frac{\partial}{\partial y}\alpha(y,t)\right) x {m_2}^5 m_1 \\
&+ \cdots = 0\,, \\
&23328 c_2(y,t)(\alpha(y,t))^3 l_2 {l_1}^5 \left(\frac{\partial}{\partial y}p(y,t)\right) m_2 \\
&+ 1458 c_2(y,t)\alpha(y,t) l_2 {l_1}^5 \frac{\mathrm{d}}{\mathrm{d}t}q(t) - 37908 c_2(y,t)
\end{aligned}$$

$$(\alpha(y,t))^3 l_2 l_1^5 \left(\frac{\partial}{\partial y} p(y,t)\right) m_1 + 23328 c_2(y,t)$$
$$(\alpha(y,t))^3 l_2 l_1^5 \left(\frac{\partial}{\partial y} \alpha(y,t)\right) x m_2 + 1458 c_2(y,t)$$
$$\alpha(y,t) l_2 l_1^5 \left(\frac{\partial}{\partial t} \alpha(y,t)\right) x$$
$$- 5832 (\alpha(y,t))^2 \left(\frac{\partial}{\partial y} c_0(y,t)\right) l_2 l_1^5 c_2(y,t)$$
$$- 37908 c_2(y,t)(\alpha(y,t))^3 l_2 l_1^5 \left(\frac{\partial}{\partial y} \alpha(y,t)\right) x m_1$$
$$+ \cdots = 0$$
$$- 60 (\alpha(y,t))^3 c_1(y,t) m_1^3 \left(\frac{\partial}{\partial y} \alpha(y,t)\right) x m_2^3 l_1$$
$$- 540 (\alpha(y,t))^3 c_1(y,t) l_2 n_2 \left(\frac{\partial}{\partial y} p(y,t)\right) m_2^4 l_1$$
$$+ 2160 (\alpha(y,t))^3 c_1(y,t) l_2 n_2 \left(\frac{\partial}{\partial y} \alpha(y,t)\right) x m_1 m_2^3 l_1$$
$$- 3240 (\alpha(y,t))^3 c_1(y,t) l_2 n_2 \left(\frac{\partial}{\partial y} \alpha(y,t)\right) x m_1^2 m_2^2 l_1$$
$$+ 2160 (\alpha(y,t))^3 c_1(y,t) l_2 n_2 \left(\frac{\partial}{\partial y} \alpha(y,t)\right) x m_1^3 m_2 l_1$$
$$- 1296 c_3(y,t)(\alpha(y,t))^3 l_2^2 n_2 \left(\frac{\partial}{\partial y} \alpha(y,t)\right) x l_1 m_2^3$$
$$+ 180 c_3(y,t)(\alpha(y,t))^3 l_2 l_1 m_1^3 \left(\frac{\partial}{\partial y} \alpha(y,t)\right) x m_2^2$$
$$+ \cdots = 0\,,$$
$$\ldots \tag{105}$$

Because there are so many over-determined partial differential equations, only a few of them are shown here for convenience. With the aid of Maple symbolic computation software, we solve the over-determined partial differential equations, and we get the following solutions of $\alpha(y,t)$, $p(y,t)$, $q(t)$, $c_0(y,t)$, $c_1(y,t)$, $c_2(y,t)$, $c_3(y,t)$ and $c_4(y,t)$:

Case 1

$$c_4(y,t) = \frac{C_1}{F_1(t)}\,, \quad p(y,t) = F_2(t)\,, \quad c_3(y,t) = \frac{C_2}{F_1(t)}\,,$$
$$\alpha(y,t) = F_1(t)\,, \quad c_2(y,t) = \frac{C_3}{F_1(t)}\,, \quad c_1(y,t) = \frac{C_4}{F_1(t)}\,,$$
$$q(t) = \int 4F_1(t)F_3(t) - \left(\frac{\mathrm{d}}{\mathrm{d}t} F_1(t)\right) x - \frac{\mathrm{d}}{\mathrm{d}t} F_2(t)\mathrm{d}t + C_5\,, \quad c_0(y,t) = F_3(t) y + F_4(t)\,, \tag{106}$$

where C_i, $i = 1,2,3,4,5$ are arbitrary constants and F_i, $i = 1,2,3,4$ are all arbitrary functions with respect to variable t.

Case 2

$$c_2(y,t) = -\frac{48 C_1 l_1}{F_1(t)(m_1 - m_2)}\,,$$
$$q(t) = \int 4F_1(t)F_3(t) - \left(\frac{\mathrm{d}}{\mathrm{d}t} F_1(t)\right) x - \frac{\mathrm{d}}{\mathrm{d}t} F_2(t)\mathrm{d}t + C_4\,,$$
$$\alpha(y,t) = F_1(t)\,, \quad c_1(y,t) = \frac{C_3}{F_1(t)}\,,$$
$$c_4(y,t) = \frac{C_1}{F_1(t)}\,, \quad p(y,t) = F_2(t)\,,$$
$$c_3(y,t) = \frac{C_2}{F_1(t)}\,, \quad c_0(y,t) = F_3(t) y + F_4(t)\,, \tag{107}$$

where C_i, $i = 1,2,3,4$ are arbitrary constants and F_i, $i = 1,2,3,4$ are all arbitrary functions with respect to variable t.

Case 3

$$c_2(y,t) = -\frac{48 C_1 l_1}{F_1(t)(m_1 - m_2)}\,, \quad c_4(y,t) = \frac{C_1}{F_1(t)}\,,$$
$$p(y,t) = F_2(t)\,, \quad c_3(y,t) = \frac{C_2}{F_1(t)}\,, \quad \alpha(y,t) = F_1(t)\,,$$
$$c_0(y,t) = F_3(t) y + F_4(t)\,, \quad c_1(y,t) = 0\,,$$
$$q(t) = \int 4F_1(t)F_3(t) - \left(\frac{\mathrm{d}}{\mathrm{d}t} F_1(t)\right) x - \frac{\mathrm{d}}{\mathrm{d}t} F_2(t)\mathrm{d}t + C_3\,, \tag{108}$$

where C_i, $i = 1,2,3,4$ are arbitrary constants and F_i, $i = 1,2,3,4$ are all arbitrary functions with respect to variable t.

Case 4

$$\begin{aligned} &c_3(y,t) = c_3(y,t)\,, \quad c_4(y,t) = c_4(y,t)\,,\\ &q(t) = q(t)\,, \quad c_0(y,t) = c_0(y,t)\,,\\ &p(y,t) = p(y,t)\,, \quad c_1(y,t) = c_1(y,t)\,,\\ &c_2(y,t) = c_2(y,t)\,, \quad \alpha(y,t) = 0\,, \end{aligned} \tag{109}$$

where C_1 is an arbitrary constant and $c_3(y,t)$, $c_4(y,t)$, $q(t)$, $c_0(y,t)$, $p(y,t)$, $c_1(y,t)$, $c_2(y,t)$ are all arbitrary functions with respect to variable y and t.

Select the proper value of parameters l_1, l_2, m_1, m_2, n_1, n_2 to determine the corresponding Jacobi elliptic functions $f(\omega)$ and $g(\omega)$ in (101). The relations between the parameters and the corresponding Jacobi elliptic function are known and given in Table 1, above.

By using the results obtained in the above step, we can derive a series of generalized solutions, such as different kinds of Jacobi elliptic functions solutions. Finally, by substituting (106)–(109) into the generalized solutions with the corresponding solutions of $f(\omega)$ and $g(\omega)$, respectively, we can get the the new exact solutions of the given Eqs. (103).

Type 1 If we select $l_1 = k^2, m_1 = -(1+k^2), n_1 = 1, l_2 = -k^2, m_2 = (2k^2-1), n_2 = (1-k^2)$, corresponding to (101), we can get $f = sn(\omega), g = cn(\omega)$.

Thus we get the following Jacobi elliptic function solutions of the equation (103):

$$v_1(x,y,t) = c_0(y,t) + c_1(y,t)cn(\omega) + c_2(y,t)sn(\omega) + \frac{c_3(y,t)}{cn(\omega)} + \frac{c_4(y,t)}{sn(\omega)}, \quad (110)$$

where $\omega = \alpha(y,t)x + p(y,t) + q(t), \alpha(y,t), p(y,t), q(t), c_0(y,t), c_1(y,t), c_2(y,t), c_3(y,t)$ and $c_4(y,t)$ satisfy (106)–(109) respectively.

For example, when we select $\alpha(y,t), p(y,t), q(t), c_0(y,t), c_1(y,t), c_2(y,t), c_3(y,t)$ and $c_4(y,t)$ to satisfy (106), namely Case 1, we can easily get the following Jacobi elliptic function solution of Eq. (103):

$$v_2(x,y,t) = F_3(t)y + F_4(t) + \frac{C_4 cn(\omega)}{F_1(t)} + \frac{C_3 sn(\omega)}{F_1(t)} + \frac{C_2}{F_1(t)cn(\omega)} + \frac{C_1}{F_1(t)sn(\omega)}, \quad (111)$$

where $F_i, i = 1,2,3,4$ are arbitrary functions of t, $C_i(i = 1,2,3,4,5)$ are arbitrary constants, and

$$\omega = F_1(t)x + F_2(t) + \int 4F_1(t)F_3(t) - \left(\frac{\mathrm{d}}{\mathrm{d}t}F_1(t)\right)x - \frac{\mathrm{d}}{\mathrm{d}t}F_2(t)\mathrm{d}t + C_5 .$$

When $k \to 1, sn(\omega) \to \tanh(\omega), cn(\omega) \to \mathrm{sech}(\omega)$, so we get degenerative soliton-like solutions from the solution (111):

$$v_3(x,y,t) = F_3(t)y + F_4(t) + \frac{C_4\,\mathrm{sech}(\omega)}{F_1(t)} + \frac{C_3\tanh(\omega)}{F_1(t)} + \frac{C_2}{F_1(t)\,\mathrm{sech}(\omega)} + \frac{C_1}{F_1(t)\tanh(\omega)} . \quad (112)$$

When $k \to 0, sn(\omega) \to \sin(\omega), cn(\omega) \to \cos(\omega)$, so we get the degenerative trigonometric function solutions from the solution (111):

$$v_4(x,y,t) = F_3(t)y + F_4(t) + \frac{C_4\cos(\omega)}{F_1(t)} + \frac{C_3\sin(\omega)}{F_1(t)} + \frac{C_2}{F_1(t)\cos(\omega)} + \frac{C_1}{F_1(t)\sin(\omega)} . \quad (113)$$

So from Case 1, we can obtain Jacobi elliptic function solutions, degenerative soliton-like solutions and trigonometric function solutions of Eq. (103). We can also get many other solutions if we make use of Case 2–Case 4. Therefore, through selecting $l_1 = k^2, m_1 = -(1+k^2), n_1 = 1, l_2 = -k^2, m_2 = (2k^2-1), n_2 = (1-k^2)$, we can get families of new exact solutions of Eq. (103).

Type 2 If we select $l_1 = -k^2, m_1 = (2k^2-1), n_1 = 1-k^2, l_2 = -1, m_2 = (2-k^2), n_2 = k^2-1$, corresponding to (101), we can get $f = cn(\omega), g = dn(\omega)$.

So we get the following Jacobi elliptic function solutions of Eq. (103):

$$v_5(x,y,t) = c_0(y,t) + c_1(y,t)dn(\omega) + c_2(y,t)cn(\omega) + \frac{c_3(y,t)}{dn(\omega)} + \frac{c_4(y,t)}{cn(\omega)}, \quad (114)$$

where $\omega = \alpha(y,t)x + p(y,t) + q(t), \alpha(y,t), p(y,t), q(t), c_0(y,t), c_1(y,t), c_2(y,t), c_3(y,t)$ and $c_4(y,t)$ satisfy (106)–(109), respectively.

For example, when we select $\alpha(y,t), p(y,t), q(t), c_0(y,t), c_1(y,t), c_2(y,t), c_3(y,t)$ and $c_4(y,t)$ to satisfy (107), namely Case 2, we can easily get the following Jacobi elliptic function solution of Eq. (103):

$$v_6(x,y,t) = F_3(t)y + F_4(t) + \frac{C_3 dn(\omega)}{F_1(t)} - 48\frac{C_1 l_1 cn(\omega)}{F_1(t)(m_1-m_2)} + \frac{C_2}{F_1(t)dn(\omega)} + \frac{C_1}{F_1(t)cn(\omega)} . \quad (115)$$

where $F_i(t), i = 1,2,3,4$ are arbitrary functions of t, $C_i(i = 1,2,3,4)$ are arbitrary constants, and

$$\omega = F_1(t)x + F_2(t) + \int 4F_1(t)F_3(t) - \left(\frac{\mathrm{d}}{\mathrm{d}t}F_1(t)\right)x - \frac{\mathrm{d}}{\mathrm{d}t}F_2(t)\mathrm{d}t + C_4 .$$

When $k \to 1, cn(\omega) \to \mathrm{sech}(\omega), dn(\omega) \to \mathrm{sech}(\omega)$, so we get a degenerative soliton-like solution from the solution (115):

$$v_7(x,y,t) = F_3(t)y + F_4(t) + \frac{C_3\,\mathrm{sech}(\omega)}{F_1(t)} - 48\frac{C_1 l_1\,\mathrm{sech}(\omega)}{F_1(t)(m_1-m_2)} + \frac{C_1+C_2}{F_1(t)\,\mathrm{sech}(\omega)}\,, \quad (116)$$

When $k \to 0$, $cn(\omega) \to cos(\omega)$, $dn(\omega) \to 1$, so we get the degenerative trigonometric function solutions from the solution (115):

$$v_8 = F_3(t)y + F_4(t) + \frac{C_2+C_3}{F_1(t)} - 48\frac{C_1 l_1 \cos(\omega)}{F_1(t)(m_1-m_2)} + \frac{C_1}{F_1(t)\cos(\omega)}\,. \quad (117)$$

where $F_i(t)$, $i=1,2,3,4$ are arbitrary functions of t and $C_i (i=1,2,3,4)$ are arbitrary constants.

So from Case 2, we we can get the Jacobi elliptic function solutions, degenerative soliton-like solutions and trigonometric function solutions of Eq. (103). We can also get some other solutions if we make use of other cases. Therefore, through selecting $l_1 = -k^2$, $m_1 = (2k^2-1)$, $n_1 = 1-k^2$, $l_2 = -1$, $m_2 = (2-k^2)$, $n_2 = k^2-1$, we can get families of new exact solutions of Eq. (103).

Type 3 If we select $l_1 = 1$, $m_1 = (2-k^2)$, $n_1 = 1-k^2$, $l_2 = 1$, $m_2 = (2k^2-1)$, $n_2 = -k^2(1-k^2)$, corresponding to (101), we can get $f = cs(\omega)$, $g = ds(\omega)$.

So we get the following Jacobi elliptic function solutions of Eq. (103):

$$v_9(x,y,t) = c_0(y,t) + c_1(y,t)ds(\omega) + c_2(y,t)cs(\omega) + \frac{c_3(y,t)}{ds(\omega)} + \frac{c_4(y,t)}{cs(\omega)}\,, \quad (118)$$

where $\omega = \alpha(y,t)x + p(y,t) + q(t)$, $\alpha(y,t)$, $p(y,t)$, $q(t)$, $c_0(y,t)$, $c_1(y,t)$, $c_2(y,t)$, $c_3(y,t)$ and $c_4(y,t)$ satisfy (106)–(109), respectively.

For example, when we select $\alpha(y,t)$, $p(y,t)$, $q(t)$, $c_0(y,t)$, $c_1(y,t)$, $c_2(y,t)$, $c_3(y,t)$ and $c_4(y,t)$ to satisfy (108), namely Case 3, we can easily get the following Jacobi elliptic function solution of Eq. (103):

$$v_{10}(x,y,t) = F_3(t)y + F_4(t) - 48\frac{C_1 l_1 cs(\omega)}{F_1(t)(m_1-m_2)} + \frac{C_2}{F_1(t)ds(\omega)} + \frac{C_1}{F_1(t)cs(\omega)}\,, \quad (119)$$

where $F_i(t)$, $i=1,2,3,4$ are arbitrary functions of t, $C_i(i=1,2,3,4)$ are arbitrary constants, and

$$\omega = F_1(t)x + F_2(t)q(t) = \int 4F_1(t)F_3(t) - \left(\frac{\mathrm{d}}{\mathrm{d}t}F_1(t)\right)x - \frac{\mathrm{d}}{\mathrm{d}t}F_2(t)\mathrm{d}t + C_3\,.$$

When $k \to 1$, $cs(\omega) \to \mathrm{sech}(\omega)\coth(\omega)$, $ds(\omega) \to \mathrm{sech}(\omega)\coth(\omega)$, so we get a degenerative soliton-like solution from the solution (119):

$$v_{11}(x,y,t) = F_3(t)y + F_4(t) - 48\frac{C_1 l_1\,\mathrm{sech}(\omega)\coth(\omega)}{F_1(t)(m_1-m_2)} + \frac{C_1+C_2}{F_1(t)\,\mathrm{sech}(\omega)\coth(\omega)}\,. \quad (120)$$

When $k \to 0$, $cs(\omega) \to cot(\omega)$, $ds(\omega) \to \csc(\omega)$, so we get the triangular function solution from the solution (119):

$$v_{12}(x,y,t) = F_3(t)y + F_4(t) - 48\frac{C_1 l_1 \cot(\omega)}{F_1(t)(m_1-m_2)} + \frac{C_2}{F_1(t)\csc(\omega)} + \frac{C_1}{F_1(t)\cot(\omega)}\,. \quad (121)$$

So from Case 3, we can get Jacobi elliptic function solutions, degenerative soliton-like solutions and trigonometric function solutions of Eq. (103). We can also get many other solutions if we make use of other cases. Therefore, through selecting $l_1 = 1$, $m_1 = (2-k^2)$, $n_1 = 1-k^2$, $l_2 = 1$, $m_2 = (2k^2-1)$, $n_2 = -k^2(1-k^2)$, we can get families of new exact solutions of Eq. (103).

Type 4 If we select $l_1 = 1-k^2$, $m_1 = (2k^2-1)$, $n_1 = -k^2$, $l_2 = (1-k^2)$, $m_2 = (2-k^2)$, $n_2 = 1$, corresponding to (101), we can get $f = nc(\omega)$, $g = sc(\omega)$.

So we get the following Jacobi elliptic function solutions of Eq. (103):

$$v_{13}(x,y,t) = c_0(y,t) + c_1(y,t)sc(\omega) + c_2(y,t)nc(\omega) + \frac{c_3(y,t)}{sc(\omega)} + \frac{c_4(y,t)}{nc(\omega)}\,, \quad (122)$$

where $\omega = \alpha(y,t)x + p(y,t) + q(t)$, $\alpha(y,t)$, $p(y,t)$, $q(t)$, $c_0(y,t)$, $c_1(y,t)$, $c_2(y,t)$, $c_3(y,t)$ and $c_4(y,t)$ satisfy (106)–(109), respectively.

So, in the same way, we can get many other Jacobi elliptic function solutions of Eq. (103). When $k \to 1$, or $k \to 0$, we can also get families of new exact solutions of Eq. (103).

Type 5 If we select $l_1 = k^2$, $m_1 = -(1+k^2)$, $n_1 = 1$, $l_2 = -1$, $m_2 = (2-k^2)$, $n_2 = -(1-k^2)$, corresponding to (101), we can get $f = sn(\omega)$, $g = dn(\omega)$.

So we get the following Jacobi elliptic function solutions of Eq. (103):

$$v_{14}(x,y,t) = c_0(y,t) + c_1(y,t)dn(\omega) + c_2(y,t)sn(\omega) + \frac{c_3(y,t)}{dn(\omega)} + \frac{c_4(y,t)}{sn(\omega)}\,, \quad (123)$$

where $\omega = \alpha(y,t)x + p(y,t) + q(t), \alpha(y,t), p(y,t), q(t), c_0(y,t), c_1(y,t), c_2(y,t), c_3(y,t)$ and $c_4(y,t)$ satisfy (106)–(109), respectively.

So, in the same way, we can get many other Jacobi elliptic function solutions of Eq. (103). When $k \to 1$, or $k \to 0$, we can also get families of new exact solutions of Eq. (103).

Type 6 If we select $l_1 = 1, m_1 = -(1+k^2), n_1 = k^2, l_2 = 1-k^2, m_2 = (2k^2-1), n_2 = -k^2$, corresponding to (101), we can get $f = dc(\omega), g = nc(\omega)$.

So we get the following Jacobi elliptic function solutions of Eq. (103):

$$\begin{aligned} v_{15}(x,y,t) &= c_0(y,t) + c_1(y,t)nc(\omega) + c_2(y,t)dc(\omega) \\ &\quad + \frac{c_3(y,t)}{nc(\omega)} + \frac{c_4(y,t)}{dc(\omega)}, \end{aligned} \tag{124}$$

where $\omega = \alpha(y,t)x + p(y,t) + q(t), \alpha(y,t), p(y,t), q(t), c_0(y,t), c_1(y,t), c_2(y,t), c_3(y,t)$ and $c_4(y,t)$ satisfy (106)–(109), respectively.

So, in the same way, we can get many other Jacobi elliptic function solutions of Eq. (103). When $k \to 1$, or $k \to 0$, we can also get families of new exact solutions of Eq. (103).

Type 7 If we select $l_1 = 1, m_1 = -(1+k^2), n_1 = k^2, l_2 = 1, m_2 = 2-k^2, n_2 = 1-k^2$, corresponding to (101), we can get $f = ns(\omega), g = cs(\omega)$.

So we get the following Jacobi elliptic function solutions:

$$\begin{aligned} v_{16}(x,y,t) &= c_0(y,t) + c_1(y,t)cs(\omega) + c_2(y,t)ns(\omega) \\ &\quad + \frac{c_3(y,t)}{cs(\omega)} + \frac{c_4(y,t)}{ns(\omega)}, \end{aligned} \tag{125}$$

where $\omega = \alpha(y,t)x + p(y,t) + q(t), \alpha(y,t), p(y,t), q(t), c_0(y,t), c_1(y,t), c_2(y,t), c_3(y,t)$ and $c_4(y,t)$ satisfy (106)–(109), respectively.

So, in the same way, we can get many other Jacobi elliptic function solutions of Eq. (103). When $k \to 1$, or $k \to 0$, we can also get families of new exact solutions of Eq. (103).

Type 8 If we select $l_1 = 1, m_1 = -(1+k^2), n_1 = k^2, l_2 = 1, m_2 = 2k^2-1, n_2 = -k^2(1-k^2)$, corresponding to (101), we can get $f = ns(\omega), g = ds(\omega)$.

So we get the following Jacobi elliptic function solutions:

$$\begin{aligned} v_{17}(x,y,t) &= c_0(y,t) + c_1(y,t)ds(\omega) + c_2(y,t)ns(\omega) \\ &\quad + \frac{c_3(y,t)}{ds(\omega)} + \frac{c_4(y,t)}{ns(\omega)}, \end{aligned} \tag{126}$$

where $\omega = \alpha(y,t)x + p(y,t) + q(t), \alpha(y,t), p(y,t), q(t), c_0(y,t), c_1(y,t), c_2(y,t), c_3(y,t)$ and $c_4(y,t)$ satisfy (106)–(109), respectively.

So, in the same way, we can get many other Jacobi elliptic function solutions of Eq. (103). When $k \to 1$, or $k \to 0$, we can also get families of new exact solutions of Eq. (103).

Type 9 If we select $l_1 = k^2, m_1 = -(1+k^2), n_1 = 1, l_2 = -(1-k^2), m_2 = 2-k^2, n_2 = -1$, corresponding to (101), we can get $f = cd(\omega), g = nd(\omega)$.

So we get the following Jacobi elliptic function solutions of Eq. (103):

$$\begin{aligned} v_{18}(x,y,t) &= c_0(y,t) + c_1(y,t)nd(\omega) + c_2(y,t)cd(\omega) \\ &\quad + \frac{c_3(y,t)}{nd(\omega)} + \frac{c_4(y,t)}{cd(\omega)}, \end{aligned} \tag{127}$$

where $\omega = \alpha(y,t)x + p(y,t) + q(t), \alpha(y,t), p(y,t), q(t), c_0(y,t), c_1(y,t), c_2(y,t), c_3(y,t)$ and $c_4(y,t)$ satisfy (106)–(109), respectively.

So, in the same way, we can get many other Jacobi elliptic function solutions of Eq. (103). When $k \to 1$, or $k \to 0$, we can also get families of new exact solutions of Eq. (103).

Type 10 If we select $l_1 = -k^2(1-k^2), m_1 = 2k^2-1, n_1 = 1, l_2 = -(1-k^2), m_2 = 2-k^2, n_2 = -1$, corresponding to (101), we can get $f = sd(\omega), g = nd(\omega)$.

So we get the following Jacobi elliptic function solutions of Eq. (103):

$$\begin{aligned} v_{19}(x,y,t) &= c_0(y,t) + c_1(y,t)nd(\omega) + c_2(y,t)sd(\omega) \\ &\quad + \frac{c_3(y,t)}{nd(\omega)} + \frac{c_4(y,t)}{sd(\omega)}, \end{aligned} \tag{128}$$

where $\omega = \alpha(y,t)x + p(y,t) + q(t), \alpha(y,t), p(y,t), q(t), c_0(y,t), c_1(y,t), c_2(y,t), c_3(y,t)$ and $c_4\left(y,t\right)$ satisfy (106)–(109), respectively.

So, in the same way, we can get many other Jacobi elliptic function solutions of Eq. (103). When $k \to 1$, or $k \to 0$, we can also get families of new exact solutions of Eq. (103).

Type 11 If we select $l_1 = -k^2(1-k^2), m_1 = 2k^2 - 1, n_1 = 1, l_2 = k^2, m_2 = -(1+k^2), n_2 = 1$, corresponding to (101), we can get $f = sd(\omega), g = cd(\omega)$.

So we get the following Jacobi elliptic function solutions of Eq. (103):

$$v_{20}(x,y,t) = c_0(y,t) + c_1(y,t)cd(\omega) + c_2(y,t)sd(\omega) + \frac{c_3(y,t)}{cd(\omega)} + \frac{c_4(y,t)}{sd(\omega)}, \quad (129)$$

where $\omega = \alpha(y,t)x + p(y,t) + q(t), \alpha(y,t), p(y,t), q(t), c_0(y,t), c_1(y,t), c_2(y,t), c_3(y,t)$ and $c_4(y,t)$ satisfy (106)–(109), respectively.

So, in the same way, we can get many other Jacobi elliptic function solutions of Eq. (103). When $k \to 1$, or $k \to 0$, we can also get families of new exact solutions of Eqs. (103).

Type 12 If we select $l_1 = 1 - k^2, m_1 = 2 - k^2, n_1 = 1, l_2 = 1, m_2 = -(1+k^2), n_2 = k^2$, corresponding to (101), we can get $f = sc(\omega), g = dc(\omega)$.

So we get the following Jacobi elliptic function solutions:

$$v_{21}(x,y,t) = c_0(y,t) + c_1(y,t)dc(\omega) + c_2(y,t)sc(\omega) + \frac{c_3(y,t)}{dc(\omega)} + \frac{c_4(y,t)}{sc(\omega)}, \quad (130)$$

where $\omega = \alpha(y,t)x + p(y,t) + q(t), \alpha(y,t), p(y,t), q(t), c_0(y,t), c_1(y,t), c_2(y,t), c_3(y,t)$ and $c_4(y,t)$ satisfy (106)–(109), respectively.

So, in the same way, we can get many other Jacobi elliptic function solutions of Eq. (103). When $k \to 1$, or $k \to 0$, we can also get families of new exact solutions of Eq. (103).

Remark 5 From the details above, it is very easy to see that we have gotten many new exact solutions of the $(2+1)$-dimensional KdV equation (103). The solutions we get include Jacobi elliptic doubly periodic function solutions, soliton-like solutions and triangular function solutions. These solutions are more general than the solutions which the extended F-expansion method gets, and they are more abundant than the solutions in [34]. Besides, our method is more convenient than the method in [34]. Among the solutions, the arbitrary functions imply that these solutions have rich local structures. In this paper, we have provided some figures to describe the character of the new exact solutions of Eq. (103).

The Generalized Algebra Method and Its Application in (1 + 1)-Dimensional Generalized Variable – Coefficient KdV Equation

In this section, with the aid of the Maple symbolic computation system, we will develop the algebraic methods [50,51] for constructing the traveling wave solutions of NLPDEs and present the following new methods and theorems:

1. A new and general transformation, a new theorem and its proof by using Maple, are presented in [40].
2. A new mechanization method to find the exact solutions of a first-order nonlinear ordinary differential equation with any degree. The validity and reliability of the method are tested by its application to the first-order nonlinear ordinary differential equation with six degrees, eight degrees, ten degrees, and twelve degrees in [40,43,44].
3. A general transformation, a new generalized algebraic method and their algorithms are suggested based on a nonlinear ordinary differential equation with any degree. The $(1+1)$-Dimensional Generalized Variable – Coefficient KdV Equations are chosen to illustrate our algorithm so that more families of new exact solutions are obtained, which contain both non-traveling and traveling wave solutions in [40,43,44].

Recently, much research work has been concentrated on the various extensions and applications of the algebraic method with computerized symbolic computation [50,51]. The algebraic method can obtain many types of traveling wave solutions based on a nonlinear ordinary differential equation with four degrees and the following theorem:

Theorem 3 *The following nonlinear ordinary differential equation with four degrees*

$$\frac{d}{d\xi}\phi(\xi) = \sqrt{c_0 + c_1\phi(\xi) + c_2\phi^2(\xi) + c_3\phi^3(\xi) + c_4\phi^4(\xi)}, \quad c_i = constant, \quad i = 0,1,2,3,4 \quad (131)$$

admits many kinds of fundamental solutions, some of which are listed as follows.

Case 1 If $c_0 = c_1 = c_3 = 0$, Eq. (131) admits: a bell shaped solitary wave solution

$$\phi(\xi) = \sqrt{-\frac{c_2}{c_4}}\,\mathrm{sech}\left(\sqrt{c_2}\xi\right), \quad c_2 > 0, \quad c_4 < 0, \quad (132)$$

a triangular type solution

$$\phi(\xi) = \sqrt{-\frac{c_2}{c_4}} sec\left(\sqrt{-c_2}\xi\right), \quad c_2 < 0, \quad c_4 > 0, \tag{133}$$

a rational polynomial type solution

$$\phi(\xi) = -\frac{1}{\sqrt{c_4}\xi}, \quad c_2 = 0, \quad c_4 > 0. \tag{134}$$

Case 2 If $c_0 = d_2^2/(4c_4)$, $c_1 = c_3 = 0$, Eq. (131) admits a kink shaped solitary wave solution

$$\phi(\xi) = \sqrt{-\frac{c_2}{2c_4}} \tanh\left(\sqrt{-\frac{c_2}{2}}\xi\right), \quad c_2 < 0, \quad c_4 > 0, \tag{135}$$

a triangular type solution

$$\phi(\xi) = \sqrt{\frac{c_2}{2c_4}} \tan\left(\sqrt{\frac{c_2}{2}}\xi\right), \quad c_2 > 0, \quad c_4 > 0, \tag{136}$$

a rational polynomial type solution

$$\phi(\xi) = -\frac{1}{\sqrt{c_4}\xi}, \quad c_2 = 0, \quad c_4 > 0. \tag{137}$$

Case 3 If $c_1 = c_3 = 0$, Eq. (131) admits three Jacobian elliptic functions type solutions

$$\phi(\xi) = \sqrt{-\frac{c_2 m^2}{c_4(2m^2-1)}} cn\left(\sqrt{\frac{c_2}{2m^2-1}}\xi\right), \quad c_0 = \frac{d_2^2 m^2(m^2-1)}{c_4(2m^2-1)^2}, \quad c_2 > 0, \tag{138}$$

$$\phi(\xi) = \sqrt{-\frac{c_2 m^2}{c_4(m^2+1)}} sn\left(\sqrt{-\frac{c_2}{m^2+1}}\xi\right), \quad c_0 = \frac{d_2^2 m^2}{c_4(m^2+1)^2}, \quad c_2 < 0, \tag{139}$$

and

$$\phi(\xi) = \sqrt{-\frac{c_2}{c_4(2-m^2)}} dn\left(\sqrt{\frac{c_2}{2-m^2}}\xi\right), \quad c_0 = \frac{d_2^2(1-m^2)}{c_4(2-m^2)^2}, \quad c_2 > 0. \tag{140}$$

Case 4 If $c_0 = c_1 = c_4 = 0$, Eq. (131) admits two bell shaped solitary wave solutions

$$\phi(\xi) = -\frac{c_2}{c_3} \operatorname{sech}^2\left(\frac{\sqrt{c_2}}{2}\xi\right), \quad c_2 > 0, \tag{141}$$

and

$$\phi(\xi) = -\frac{c_2}{c_3} \operatorname{sech}^2\left(\frac{\sqrt{-c_2}}{2}\xi\right), \quad c_2 < 0, \tag{142}$$

a rational polynomial type solution

$$\phi(\xi) = \frac{1}{c_3\xi^2}, \quad c_2 = 0, \tag{143}$$

Case 5 If $c_3 > 0$, $c_4 = 0$, Eq. (131) admits a Weiertrass elliptic functions type solution

$$\phi(\xi) = \wp\left(\frac{\sqrt{c_3}}{2}\xi, -\frac{4c_1}{c_3}, -\frac{4c_0}{c_3}\right). \tag{144}$$

Let's specifically see how the algebraic method works. For a given nonlinear differential equation, say, in two variables x, t

$$F(u, u_t, u_x, u_{xx}, u_{xt}, u_{tt}, \ldots) = 0 \tag{145}$$

where F is a polynomial function with respect to the indicated variables or some function which can be reduced to a polynomial function by using some transformations.

By using the traveling wave transformation

$$u = u(\xi), \xi = x - \lambda t, \tag{146}$$

Eq. (145) is reduced to an ordinary differential equation with constant coefficients

$$G(U, U', U'', U''', \ldots) = 0 \tag{147}$$

A transformation was presented by Fan [5] in the form

$$u(x) = A_0 + \sum_{i=1}^{n} A_i \phi^{i-1}(\xi), \tag{148}$$

with the new variable $\phi(\xi)$ satisfying Eq. (131), where A_0, A_i, d_j are constants.

Substituting (148) into (147) along with (131), we can determine the parameter n in (148). And then by substituting (148) with the concrete n into (147) and equating the coefficients of these terms $\omega^i \omega'^j (i = 0, 1, 2, \ldots; j = 0, 1)$, we obtain a system of algebraic equations with respect to other parameters A_0, A_i, d_j, λ. By solving the system, if available, we may determine these parameters. Therefore we establish a transformation (147) between (146) and (148). If we know the solutions of (148), then we can obtain the solutions of (147) (or (145)) by using (146).

A New Transformation and a New Theorem

In Yu-Jie Ren's Ph.D Dissertation of Dalian University of Technology [40], in order to develop the above algebraic method, we first presented the following new transformation (150) and new Theorem 4. Then we proved the theorem by means of the Maple computer algebraic system.

Theorem 4 ([40]) *Suppose n and m are any integers, a first-order nonlinear ordinary differential equation with any degree in the form of*

$$\frac{d}{d\xi}\phi(\xi) = \varepsilon\sqrt{c_0 + \sum_{i=1}^{r} c_i \cdot \phi^i(\xi)}\,, \quad c_i = consts\,, \quad i = 1, 2, \ldots, r\,. \tag{149}$$

If we substitute the following new transform into Eq. (149)

$$\phi(\xi) = (g(\xi))^{\frac{n}{m}}\,, \tag{150}$$

then Eq. (149) can be transformed into an ordinary differential equation

$$\left(\frac{d}{d\xi}g(\xi)\right)^2 = m^2 n^{-2} \sum_{i=0}^{r} c_i (g(\xi))^{2+\frac{n(i-2)}{m}}\,. \tag{151}$$

We give the proof of Theorem 4 [40] by using Maple as follows:

Proof

Step 1 Importing the following Maple program at the Maple Command Window

$$eq := \mathrm{diff}(phi(xi), xi)^2 - \mathrm{sum}(c[i] * \mathrm{phi}(xi)^i, i = 0..r);$$

Eq. (149) is displayed at a computer screen (after implementing the above program) as follows:

$$eq := \left(\frac{\mathrm{d}}{\mathrm{d}\xi}\phi(\xi)\right)^2 - \sum_{i=0}^{r} c_i(\phi(\xi))^i\,.$$

Step 2 Importing the following Maple program at the Maple Command Window

$$eq := \mathrm{subs}(phi(xi) = (g(xi))^(n/m), eq)\,;$$

the following result is displayed at the screen (after running the above program):

$$eq := \left((g(\xi))^{\frac{n}{m}}\right)^2 n^2 \left(\frac{\mathrm{d}}{\mathrm{d}\xi}g(\xi)\right)^2 m^{-2}(g(\xi))^{-2} - \sum_{i=0}^{r} c_i \left((g(\xi))^{\frac{n}{m}}\right)^i\,.$$

Step 3 Importing the following Maple program at the Maple Command Window

$$eq := \mathrm{simplify}(eq)\,; \quad eq := \mathrm{expand}(eq)\,; \quad eq := \mathrm{numer}(eq)\,;$$

the following result is displayed at the screen (after running the above program):

$$eq := \left((g(\xi))^{\frac{n}{m}}\right)^2 n^2 \left(\frac{\mathrm{d}}{\mathrm{d}\xi}g(\xi)\right)^2 - \sum_{i=0}^{r} c_i \left((g(\xi))^{\frac{n}{m}}\right)^i m^2 (g(\xi))^2\,.$$

Step 4 Importing the following Maple program at the Maple Command Window

$$eq := \mathrm{subs}(\mathrm{diff}(g(xi), xi) = G(xi), g(xi) = g, eq)\,;$$
$$eq := eq * (g^(n/m))^(-2)\,; \quad eq := \mathrm{simplify}(eq)\,;$$
$$eq := n^2 * G(xi)^2 - (\mathrm{sum}(c[i] * g^(-2 * (n - m)/m) + i * n/m), i = 0..r)) * m^2\,;$$
$$eq := \mathrm{simplify}(eq);$$

the following result is displayed at the screen (after running the above program):

$$eq := n^2(G(\xi))^2 - g^{-2\frac{n-m}{m}} \sum_{i=0}^{r} c_i \left(g^{\frac{n}{m}}\right)^i m^2$$

$$eq := n^2(G(\xi))^2 - \sum_{i=0}^{r} c_i g^{\frac{-2n+2m+ni}{m}} m^2\,.$$

Step 5 Importing the following Maple program at the Maple Command Window

$$eq := \mathrm{subs}(g = g(xi), G(xi) = \mathrm{diff}(g(xi), xi), eq)\,;$$

the following result is displayed at the screen (after running the above program):

$$eq := n^2\left(\frac{\mathrm{d}}{\mathrm{d}\xi}g(\xi)\right)^2 - \sum_{i=0}^{r} c_i (g(\xi))^{\frac{-2n+2m+ni}{m}} m^2\,. \tag{152}$$

We can reduce Eq. (152) to (151).

□

Remark 6 The above transformation (150), Theorem 4, and its proof by means of Maple were first presented by us in Yu-Jie Ren's Ph.D Dissertation of Dalian University of Technology [40].

A New Mechanization Method to Find the Exact Solutions of a First-Order Nonlinear Ordinary Differential Equation with any Degree Using Maple and Its Application

According to Theorem 4 in Subsect. "The the Exp-Bäcklund Transformation Method and Its Application in (1 + 1)-Dimensional KdV Equation", in [40] we first presented the following mechanization method to find the exact solutions of a first-order nonlinear ordinary differential equation with any degree. The validity and reliability of our method are tested by its application to a first-order nonlinear ordinary differential equation with six degrees [40]. Now, we simply describe our mechanization method as follows:

Step 1 Import the following Maple program at the Maple Command Window

$$eq := n^2 * \mathrm{diff}(g(xi), xi)^2 - \mathrm{sum}(c[i] \\ *g(xi)^((-2*n+2*m+n*i)/m), i = 0..r)*m^2;$$

Eq. (151) is displayed at a computer screen (after implementing the above program) as follows

$$eq := n^2\left(\frac{\mathrm{d}}{\mathrm{d}\xi}g(\xi)\right)^2 - \sum_{i=0}^{r} c_i(g(\xi))^{\frac{-2n+2m+ni}{m}} m^2$$

Step 2 Import the following Maple programs with some values of the degree r and parameter (n, m) in new transform (150) at the Maple Command Window. For example,

$$eq621 := \mathrm{subs}(r = 6, m = 2, n = 1, eq);$$
$$eq621 := \mathrm{simplify}(eq621);$$

the following result is displayed at the screen (after running the above program):

$$eq621 := \left(\frac{\mathrm{d}}{\mathrm{d}\xi}g(\xi)\right)^2 - 4c_0 g(\xi) - 4c_1(g(\xi))^{3/2} - 4c_2(g(\xi))^2 - 4c_3(g(\xi))^{5/2} - 4c_4(g(\xi))^3 - 4c_5(g(\xi))^{7/2} - 4c_6(g(\xi))^4 . \tag{153}$$

Step 3 According to the output results in Step 2, we choose the coefficients of ci, $i = 1, 2, \ldots, m(< n)$ to be zero. Then we import their Maple program. For example, we import the Maple program as follows:

$$eq246 := \mathrm{subs}(c[0] = 0, c[1] = 0, c[3] = 0, c[5] = 0, eq621);$$

the following result is displayed at the screen (after running the above program):

$$eq246 := \left(\frac{\mathrm{d}}{\mathrm{d}\xi}g(\xi)\right)^2 - 4c_2(g(\xi))^2 - 4c_4(g(\xi))^3 - 4c_6(g(\xi))^4 . \tag{154}$$

Step 4 Import the Maple program for solving the output equation in Step 3. For example, we import the Maple program for solving the output Eq. (154) as follows:

$$\mathrm{dsolve}(eq246, g(xi));$$

the following formal solutions of (154) are displayed at the screen (after running the above program):

$$g_1(\xi) = 1/2\frac{-c_4 - \sqrt{c_4^2 - 4c_6c_2}}{c_6}, \tag{155}$$

$$g_2(\xi) = 1/2\frac{-c_4 + \sqrt{c_4^2 - 4c_6c_2}}{c_6}, \tag{156}$$

$$g_3(\xi) = 4\left(e^{C_1\sqrt{c_2}}\right)^2 c_2\left(e^{\xi\sqrt{c_2}}\right)^{-2} \left(-4c_6c_2 + \frac{\left(e^{C_1\sqrt{c_2}}\right)^4}{\left(e^{\xi\sqrt{c_2}}\right)^4} - 2\frac{\left(e^{C_1\sqrt{c_2}}\right)^2 c_4}{\left(e^{\xi\sqrt{c_2}}\right)^2} + c_4^2\right)^{-1}, \tag{157}$$

$$g_4(\xi) = -4\left(e^{\xi\sqrt{c_2}}\right)^2 c_2\left(e^{C_1\sqrt{c_2}}\right)^{-2} \left(4c_6c_2 - \frac{\left(e^{\xi\sqrt{c_2}}\right)^4}{\left(e^{C_1\sqrt{c_2}}\right)^4} + 2\frac{\left(e^{\xi\sqrt{c_2}}\right)^2 c_4}{\left(e^{C_1\sqrt{c_2}}\right)^2} - c_4^2\right)^{-1}. \tag{158}$$

Import the Maple program for reducing the above solutions of (154). For example, we import the Maple program for reducing the (157) and (158) as follows:

$$g3 := \mathrm{simplify}(g3); g4 := \mathrm{simplify}(g4);$$

the following results are displayed at a computer screen (after running the above program):

$$g_3(\xi) = 4\frac{e^{2\sqrt{c_2}(C_1+\xi)}c_2}{-4c_6c_2e^{4\xi\sqrt{c_2}} + e^{4C_1\sqrt{c_2}} - 2e^{2\sqrt{c_2}(C_1+\xi)}c_4 + c_4^2e^{4\xi\sqrt{c_2}}}, \tag{159}$$

$$g_4(\xi) = 4 \frac{e^{2\sqrt{c_2}(C_1+\xi)} c_2}{-4c_6 c_2 e^{4C_1\sqrt{c_2}} + e^{4\xi\sqrt{c_2}} - 2e^{2\sqrt{c_2}(C_1+\xi)} c_4 + {c_4}^2 e^{4C_1\sqrt{c_2}}} . \quad (160)$$

Step 5 We discuss the above solutions under different conditions in Step 3.

For example, we discuss the solutions under different conditions $c_4^2 - 4c_6c_2 < 0$ or $c_4^2 - 4c_6c_2 > 0$ or $c_4^2 - 4c_6c_2 = 0$ or $c_2 < 0$ or $c_2 > 0$ or $c_2 = 0$. When

$$c_2 > 0, 4c_6c_2 - c_4^2 < 0 , \quad (161)$$

we may import the following Maple program:

$$\mathrm{assume}(c[2] > 0, (4 * c[2] * c[6] - c[4]^2) < 0);$$
$$eq246jie := \mathrm{dsolve}(eq246, g(xi));$$

six solutions of a first-order nonlinear ordinary differential equation with six degrees under condition (161), which includes two new solutions, are displayed at a computer screen (after running the above program). Here we omit them due to the length of our article. Importing the following Maple program for reducing the two new solutions above,

$$eq246jie1 := \mathrm{subs}(2 * xi * \mathrm{sqrt}(c[2]) - 2 *_C 1 * \mathrm{sqrt}(c[2]) = eta, eq246jie1);$$
$$eq246jie2 := \mathrm{subs}(2 * xi * \mathrm{sqrt}(c[2]) - 2 *_C 1 * \mathrm{sqrt}(c[2]) = eta, eq246jie2);$$

the following results are displayed at a computer screen (after running the above program):

$$(\phi_{11}(\xi))^2 = -2 \frac{c_2\left(-{c_4}^2 + (\tanh(\eta))^2{c_4}^2 - \tanh(\eta)\sqrt{{c_4}^2 - 4c_6c_2}\sqrt{{c_4}^2((\tanh(\eta))^2 - 1)}\right)}{(4c_6(\tanh(\eta))^2c_2 - {c_4}^2)\,c_4} , \quad (162)$$

$$(\phi_{12}(\xi))^2 = -2 \frac{c_2\left(-{c_4}^2 + (\tanh(\eta))^2{c_4}^2 + \tanh(\eta)\sqrt{{c_4}^2 - 4c_6c_2}\sqrt{{c_4}^2((\tanh(\eta))^2 - 1)}\right)}{(4c_6(\tanh(\eta))^2c_2 - {c_4}^2)c_4} . \quad (163)$$

Importing the following Maple program for reducing (162) and (163):

$$eq246jie3 := \mathrm{subs}(\mathrm{sqrt}(c[4]^2 * (\tanh(eta)^2 - 1)) = I * \mathrm{abs}(c[4] * \mathrm{sech}(eta)), -c[4] + \tanh(eta)^2 * c[4] = -\mathrm{sech}(eta)^2 * c[4], eq246jie3);$$
$$eq246jie4 := \mathrm{subs}(\mathrm{sqrt}(c[4]^2 * (\tanh(eta)^2 - 1)) = I * \mathrm{abs}(c[4] * \mathrm{sech}(eta)), -c[4] + \tanh(eta)^2 * c[4] = -\mathrm{sech}(eta)^2 * c[4], eq246jie4);$$
$$phi[1](xi) := epsilon * (eq246jie3)^(1/2);$$
$$phi[2](xi) := epsilon * (eq246jie4)^(1/2);$$

the following results are displayed at a computer screen (after running the above program):

$$\phi_1(\xi) = \varepsilon \sqrt{\frac{2c_2\left(c_4^2\,\mathrm{sech}^2(\eta) + i|c_4\,\mathrm{sech}(\eta)|\tanh(\eta)\sqrt{{c_4}^2 - 4c_6c_2}\right)}{c_4\left(4c_6\tanh^2(\eta)c_2 - {c_4}^2\right)}} , \quad (164)$$

$$\phi_2(\xi) = \varepsilon \sqrt{\frac{2c_2\left({c_4}^2\,\mathrm{sech}^2(\eta) - i|c_4\,\mathrm{sech}(\eta)|\tanh(\eta)\sqrt{{c_4}^2 - 4c_6c_2}\right)}{c_4\left(4c_6\tanh^2(\eta)c_2 - {c_4}^2\right)}} . \quad (165)$$

where $\eta = 2\sqrt{c_2}(\xi - C_1)$.

Step 6 We need to further discuss the results found by sometimes adding new conditions so that the results are simpler in form. For example, we can import the following Maple program for getting rid of the absolute value sign in the above results:

$$\mathrm{assume}(c[4] * \mathrm{sech}(eta) < 0);$$
$$eq246jie3 := 2 * c[2] * (c[4]^2 * \mathrm{sech}(eta)^2 + I * \tanh(eta) * sqrt(c[4]^2 - 4 * c[6] * c[2]) * abs(c[4] * sech(eta)))/(4 * c[6] * \tanh(eta)^2 * *c[2] - c[4]^2)/c[4];$$
$$eq246jie4 := 2 * c[2] * (c[4]^2 * \mathrm{sech}(eta)^2 - I * \tanh(eta)sqrt(c[4]^2 - 4 * c[6] * c[2]) * abs(c[4]* \mathrm{sech}(eta)))/(4 * c[6] * \tanh(eta)^2 * c[2] - c[4]^2)/c[4]; phi[1,1](xi) := epsilon * (eq246jie3)^(1/2);$$
$$phi[1,2](xi) := epsilon * (eq246jie4)^(1/2);$$

the following results are displayed at a computer screen (after running the above program):

$$\phi_{1,1}(\xi) = \varepsilon\sqrt{\frac{2c_2\operatorname{sech}(\eta)\left(c_4\operatorname{sech}(\eta) - i\tanh(\eta)\sqrt{c_4^2 - 4c_6c_2}\right)}{4c_6(\tanh(\eta))^2c_2 - c_4^2}}, \tag{166}$$

$$\phi_{1,2}(\xi) = \varepsilon\sqrt{\frac{2c_2\operatorname{sech}(\eta)\left(c_4\operatorname{sech}(\eta) + i\tanh(\eta)\sqrt{c_4^2 - 4c_6c_2}\right)}{4c_6(\tanh(\eta))^2c_2 - c_4^2}}. \tag{167}$$

Importing the following Maple program for reducing and discussing above results:

```
assume(c[4] * sech(eta) > 0);
eq246jie3 := 2 * c[2] * (c[4]^2 * sech(eta)^2 + I
  * tanh(eta) * sqrt(c[4]^2 - 4 * c[6] * c[2])*
  abs(c[4] * sech(eta)))/(4 * c[6] * tanh(eta)^2
  * c[2] - c[4]^2)/c[4];
eq246jie4 := 2 * c[2] * (c[4]^2 * sech(eta)^2
  - I * tanh(eta) * sqrt(c[4]^2 - 4 * c[6] * c[2])*
  abs(c[4] * sech(eta)))/(4 * c[6] * tanh(eta)^2
  * c[2] - c[4]^2)/c[4];
phi[2, 1](xi) := epsilon * (eq246jie3)^(1/2);
phi[2, 2](xi) := epsilon * (eq246jie4)^(1/2);
```

the following results are shown at a computer screen (after running the above program):

$$\phi_{2,1}(\xi) = \varepsilon\sqrt{\frac{2c_2\operatorname{sech}(\eta)\left(c_4\operatorname{sech}(\eta) + i\tanh(\eta)\sqrt{c_4^2 - 4c_6c_2}\right)}{4c_6(\tanh(\eta))^2c_2 - c_4^2}}, \tag{168}$$

$$\phi_{2,2}(\xi) = \varepsilon\sqrt{\frac{2c_2\operatorname{sech}(\eta)\left(c_4\operatorname{sech}(\eta) - i\tanh(\eta)\sqrt{c_4^2 - 4c_6c_2}\right)}{4c_6(\tanh(\eta))^2c_2 - c_4^2}}. \tag{169}$$

By using this method, we obtained some new types of general solution of a first-order nonlinear ordinary differential equation with six degrees and presented the following theorem in [40,43,44].

Theorem 5 *The following nonlinear ordinary differential equation with six degrees*

$$\frac{d\phi(\xi)}{d\xi} = \varepsilon\sqrt{\begin{array}{c} c_0 + c_1\phi(\xi) + c_2\phi^2(\xi) + c_3\phi^3(\xi) \\ +c_4\phi^4(\xi) + c_5\phi^5(\xi) + c_6\phi^6(\xi)\end{array}}, \quad c_i = constant\,, \quad i = 0, 1, 2, 3, 4, 5, 6 \tag{170}$$

admits many kinds of fundamental solutions which depend on the values and constraints between c_i, $i = 0, 1, 2, 3, 4, 5, 6$.

Some of the solutions are listed in the following sections.

Case 1. If $c_0 = c_1 = c_3 = c_5 = 0$, Eq. (170) admits the following constant solutions

$$\phi_{1,2}(\xi) = \varepsilon\sqrt{\frac{-c_4 + \sqrt{c_4^2 - 4c_6c_2}}{2c_6}}, \tag{171}$$

$$\phi_{3,4}(\xi) = \varepsilon\sqrt{\frac{-c_4 - \sqrt{c_4^2 - 4c_6c_2}}{2c_6}}, \tag{172}$$

and the following exponential type solutions

$$\phi_{5,6}(\xi) = 2\varepsilon\sqrt{\frac{e^{2\sqrt{c_2}(\xi+C_1)}c_2}{\begin{array}{c}-4c_6c_2e^{4C_1\sqrt{c_2}} + e^{4\xi\sqrt{c_2}} \\ -2e^{2\sqrt{c_2}(\xi+C_1)}c_4 + c_4^2e^{4C_1\sqrt{c_2}}\end{array}}}, \tag{173}$$

$$\phi_{7,8}(\xi) = 2\varepsilon\sqrt{\frac{e^{2\sqrt{c_2}(\xi+C_1)}c_2}{\begin{array}{c}-4c_6c_2e^{4\xi\sqrt{c_2}} + e^{4C_1\sqrt{c_2}} \\ -2e^{2\sqrt{c_2}(\xi+C_1)}c_4 + c_4^2e^{4\xi\sqrt{c_2}}\end{array}}}. \tag{174}$$

When we take different values and constraints of $4c_2c_6 - c_4^2$, the solutions (5.43) and (5.44) can be written in different formats as follows:

Case 1.1. If $4c_2c_6 - c_4^2 < 0$, Eq. (170) admits the following tanh-sech hyperbolic type solutions

$$\phi_{1,2}(\xi) = \varepsilon\sqrt{\frac{\begin{array}{c}2c_2\operatorname{sech}(\eta)\Big[c_4\operatorname{sech}(\eta) \\ +\varepsilon i\sqrt{c_4^2 - 4c_6c_2}\tanh(\eta)\Big]\end{array}}{4c_6c_2\tanh^2(\eta) - c_4^2}}, \quad c_2 > 0\,, \tag{175}$$

$$\phi_{3,4}(\xi) = \varepsilon\sqrt{\frac{2c_2\operatorname{sech}(\eta)\Big[c_4\operatorname{sech}(\eta) - \varepsilon i\sqrt{c_4{}^2 - 4c_6c_2}\tanh(\eta)\Big]}{4c_6c_2\tanh^2(\eta) - c_4{}^2}}\,, \quad c_2 > 0\,, \tag{176}$$

and the following tan-sec triangular type solutions

$$\phi_{5,6}(\xi) = \varepsilon\sqrt{\frac{2c_2\sec(\zeta)\Big[-c_4\sec(\zeta) + \varepsilon\sqrt{c_4{}^2 - 4c_6c_2}\tan(\zeta)\Big]}{4c_6c_2\tan^2(\zeta) + c_4{}^2}}\,, \quad c_2 < 0\,, \tag{177}$$

$$\phi_{7,8}(\xi) = \varepsilon\sqrt{\frac{2c_2\sec(\zeta)\Big[-c_4\sec(\zeta) - \varepsilon\sqrt{c_4{}^2 - 4c_6c_2}\tan(\zeta)\Big]}{4c_6c_2\tan^2(\zeta) + c_4{}^2}}\,, \quad c_2 < 0\,, \tag{178}$$

where $\eta = 2\sqrt{c_2}(\xi - C_1)$ and C_1 is any constant, $\zeta = 2\sqrt{-c_2}(\xi - C_1)$ and C_1 is any constant.

Case 1.2. If $4c_2c_6 - c_4{}^2 > 0$, Eq. (170) admits the following sinh-cosh hyperbolic type solutions

$$\phi_{9,10}(\xi) = \varepsilon\sqrt{\frac{2c_2\Big[c_4 + \varepsilon\sqrt{4c_2c_6 - c_4{}^2}\sinh(\eta)\Big]}{4c_2c_6\sinh^2(\eta) - c_4{}^2\cosh^2(\eta)}}\,, \quad c_2 > 0\,, \tag{179}$$

$$\phi_{11,12}(\xi) = \varepsilon\sqrt{\frac{2c_2\Big[c_4 - \varepsilon\sqrt{4c_2c_6 - c_4{}^2}\sinh(\eta)\Big]}{4c_2c_6\sinh^2(\eta) - c_4{}^2\cosh^2(\eta)}}\,, \quad c_2 > 0\,, \tag{180}$$

and the following sin-cos triangular type solutions

$$\phi_{13,14}(\xi) = \varepsilon\sqrt{\frac{2c_2\Big[-c_4 + \varepsilon i\sqrt{4c_2c_6 - c_4{}^2}\sin(\zeta)\Big]}{4c_6c_2\sin^2(\zeta) + c_4{}^2\cos^2(\zeta)}}\,, \quad c_2 < 0 \tag{181}$$

$$\phi_{15,16}(\xi) = \varepsilon\sqrt{\frac{2c_2\Big[-c_4 - \varepsilon i\sqrt{4c_6c_2 - c_4{}^2}\sin(\zeta)\Big]}{4c_6c_2\sin^2(\zeta) + {}_4{}^2\cos^2(\zeta)}}\,, \quad c_2 < 0\,; \tag{182}$$

where $\eta = 2\sqrt{c_2}\,(\xi - C_1)$, $\zeta = 2\sqrt{-c_2}(\xi - C_1)$ and C_1 is any constant.

Case 1.3. If $4c_2c_6 - c_4{}^2 = 0$, Eq. (170) admits the following tanh hyperbolic type solutions

$$\phi_{17,18,19,20}(\xi) = \varepsilon\sqrt{\frac{2c_2}{e^{-\varepsilon 2\sqrt{c_2}(\xi - C_1)} - c_4}} = \varepsilon\sqrt{\frac{2c_2\left(1 + \varepsilon\tanh\left(\sqrt{c_2}(\xi - C_1)\right)\right)}{1 - c_4 - \varepsilon(1 + c_4)\tanh\left(\sqrt{c_2}(\xi - C_1)\right)}}\,, \quad c_2 > 0\,, \tag{183}$$

and the following tan triangular type solutions

$$\phi_{21,22,23,24}(\xi) = \varepsilon\sqrt{\frac{2c_2}{e^{-\varepsilon 2i\sqrt{-c_2}(\xi - C_1)} - c_4}} = \varepsilon\sqrt{\frac{2c_2\left(1 + \varepsilon\tan\left(i\sqrt{-c_2}(\xi - C_1)\right)\right)}{1 - c_4 - \varepsilon(1 + c_4)\tan\left(i\sqrt{-c_2}(\xi - C_1)\right)}}\,, \quad c_2 < 0 \tag{184}$$

where C_1 is any constant.

Case 2. If $c_5 = c_6 = 0$, Eq. (170) admits the solutions in Theorem 6.1.

Remark 7 By using our mechanization method using Maple to find the exact solutions of a first-order nonlinear ordinary differential equation with any degree, we can obtain some new types of general solution of a first-order nonlinear ordinary differential equation with $r(= 7, 8, 9, 10, 11, 12, \ldots)$ degree [40]. We do not list the solutions here in order to avoid unnecessary repetition.

Summary of the Generalized Algebra Method

In this section, based on a first-order nonlinear ordinary differential equation with any degree (149) and its the exact solutions obtained by using our mechanization method via Maple, we will develop the algebraic methods [50,51] for constructing the traveling wave solutions and present a new generalized algebraic method and its algorithms [40, 43,44]. The KdV equation is chosen to illustrate our algorithm so that more families of new exact solutions are obtained which contain both non-traveling solutions and traveling wave solutions. We outline the main steps of our generalized algebraic method as follows:

Step 1. For given nonlinear differential equations, with some physical fields $u_i(t, x_1, x_2, \ldots, x_m)$, $(i = 1, 2, \ldots, n)$

in $m+1$ independent variables $t, x_1, x_2, \dots, x_m$,

$$F_j(u_1,\dots,u_n,u_{1,t},\dots,u_{n,t},u_{1,x_1},\dots,u_{n,x_m},\\ u_{1,tt},\dots,u_{n,tt},u_{1,tx_1},\dots u_{n,tx_m},\dots)=0\,,\\ j=1,2,\dots,n\,, \quad (185)$$

where $j=1,2,\dots,n$.

By using the following more general transformation, which we first present here,

$$u_i(t,x_1,x_2,\dots,x_m)=U_i(\xi)\,,\\ \xi=u_i(t,x_1,x_2,\dots,x_m)=U_i(\xi)\,,\\ \xi=\alpha_0(t)+\sum_{i=1}^{m-1}\alpha_i(x_i,x_{i+1},\dots,x_m,t)\cdot\beta_i(x_i)\,, \quad (186)$$

where $\alpha_0(t), \alpha_i(x_i, x_{i+1}, \dots, x_m, t)$ and $\beta_i(x_i)$, $i=1,2,\dots,m-1$ are functions to be determined later. For example, when $n=1$, we may take

$$\xi=\alpha_0(t)+\alpha_1(x_1t)\cdot\beta_1(x_1)\,,$$

where $\alpha_0(t), \alpha_1(x_1t)$ and $\beta_1(x_1)$ are undetermined functions.

Then Eq. (185) is reduced to nonlinear differential equations

$$G_j(U_1,\dots,U_n,U_1',\dots,U_n',U_1'',\dots,U_n'',\dots)=0\,,\\ j=1,2,\dots,n\,, \quad (187)$$

where $G_j(j=1,2,\dots,n)$ are all polynomials of $U_i(i=1,2,\dots,n)$, $\alpha_0(t), \alpha_i(x_i,x_{i+1},\dots,x_m,t)$, $\beta_i(x_i)$, $i=1,2,\dots,m-1$ and their derivatives. If G_k of them is not a polynomial of $U_i(i=1,2,\dots,n)$, $\alpha_i(x_i,x_{i+1},\dots,x_m,t)$, $\beta_i(x_i)$, $i=1,2,\dots,m-1$, $\alpha_0(t)$ and their derivatives, then we may use new variable $v_i(\xi)(i=1,2,\dots,n)$ which makes G_k become a polynomial of $v_i(\xi)(i=1,2,\dots,n)$, $\alpha_0(t), \alpha_i(x_i,x_{i+1},\dots,x_m,t)$ and $\beta_i(x_i)$, $i=1,2,\dots,m-1$ and their derivatives. Otherwise the following transformation will fail to seek solutions of Eq. (185).

Step 2. We introduce a new variable $\phi(\xi)$ which is a solution of the following ODE

$$\frac{d\phi(\xi)}{d\xi}=\varepsilon\sqrt{\begin{array}{c}c_0+c_1\phi(\xi)+c_2\phi^2(\xi)+c_3\phi^3(\xi)\\+c_4\phi^4(\xi)+\cdots+c_r\phi^r(\xi)\end{array}}\,,\\ r=0,1,2,3,\dots\,. \quad (188)$$

Then the derivatives with respect to the variable ξ become the derivatives with respect to the variable ϕ.

Step 3. By using the new variable ϕ, we expand the solution of Eq. (185) in the form:

$$U_i=a_{i,0}(X)+\sum_{k=1}^{n_i}(a_{i,k}(X)\phi^k(\xi(X))\\ +b_{i,k}(X)\phi^{-k}(\xi(X)))\,. \quad (189)$$

where $X=(x_1,x_2,\dots,x_m,t)$, $\xi=\xi(X)$, $a_{i,0}(X)$, $a_{i,k}(X), b_{i,k}(X)(i=1,2,\dots,n;k=1,2,\dots,n_i)$ are all differentiable functions of X to be determined later.

Step 4. In order to determine $n_i(i=1,2,\dots,n)$ and r, we may substitute (188) into (187) and balance the highest derivative term with the nonlinear terms in Eq. (187). By using the derivatives with respect to the variable ϕ, we can obtain a relation for n_i and r, from which the different possible values of n_i and r can be obtained. These values lead to the series expansions of the solutions for Eq. (185).

Step 5. Substituting (189) into the given Eq. (185) and collecting coefficients of polynomials of ϕ^k, ϕ^{-k}, and $\phi^i\phi^{-k}\sqrt{\sum_{j=1}^{r}c_j\phi^j(\xi)}$, with the aid of Maple, then setting each coefficient to zero, we will get a system of over-determined partial differential equations with respect to $a_{i,0}(X), a_{i,k}(X), b_{i,k}(X)$, $i=1,2,\dots, n; k=1,2, \dots, n_i c_j$, $j=0,1,\dots,r\alpha_0(t), \alpha_i(x_i,x_{i+1},\dots,x_m,t)$ and $\beta_i(x_i)$, $i=1,2,\dots,m-1$.

Step 6. Solving the over-determined partial differential equations with Maple,then we can determine $a_{i,0}(X)$, $a_{i,k}(X), b_{i,k}(X)$, $i=1,2,\dots,n$; $k=1,2,\dots,n_i)c_j$, $j=0,1,\dots,r\alpha_0(t), \alpha_i(x_i,x_{i+1},\dots,x_m,t)$ and $\beta_i(x_i)$, $i=1,2,\dots,m-1$.

Step 7. From the constants $a_{i,0}(X), a_{i,k}(X), b_{i,k}(X)(i=1,2,\dots,n;k=1,2,\dots,n_i)c_j(j=0,1,\dots,r)$, $\alpha_0(t)$, $\alpha_i(x_i,x_{i+1},\dots,x_m,t)$ and $\beta_i(x_i)$, $i=1,2,\dots,m-1$ obtained in Step 6 to Eq. (188), we can then obtain all the possible solutions.

Remark 8 When $c_5=c_6=0$ and $b_{i,k}=0$, Eq. (170) and transformation (189) just become the ones used in our previous method [50,51]. However, if $c_5\neq 0$ or $c_6\neq 0$, we may obtain solutions that cannot be found by using the methods [50,51]. It should be pointed out that there is no method to find all solutions of nonlinear PDEs. But our method can be used to find more solutions of nonlinear PDEs, and with the exact solutions obtained by using our mechanization method via Maple, we will develop the algebraic methods [50,51] for constructing the traveling wave solutions and present a new generalized algebraic method and its algorithms [40,43,44].

Remark 9 By the above description, we find that our method is more general than the method in [50,51]. We have improved the method [50,51] in five aspects: First, we extend the ODE with four degrees (131) into the ODE with any degree (188) and get its new general solutions by using our mechanization method via Maple [40,43,44]. Second, we change the solution of Eq. (185) into a more general solution (189) and get more types of new rational solutions and irrational solutions. Third, we replace the traveling wave transformation (146) in [50,51] by a more general transformation (186). Fourth, we suppose the coefficients of the transformation (186) and (189) are undetermined functions, but the coefficients of the transformation (146) in [50,51] are all constants. Fifth, we present a more general algebra method than the method given in [50,51], which is called the generalized algebra method, to find more types of exact solutions of nonlinear differential equations based upon the solutions of the ODE (188). This can obtain more general solutions of the NPDEs than the number obtained by the method in [50,51].

The Generalized Algebra Method to Find New Non-traveling Waves Solutions of the (1 + 1)-Dimensional Generalized Variable-Coefficient KdV Equation

In this section, we will make use of our generalized algebra method and symbolic computation to find new non-traveling waves solutions and traveling waves solutions of the following (1 + 1) – dimensional Generalized Variable – Coefficient KdV equation [16].

Propagation of weakly nonlinear long waves in an inhomogeneous waveguide is governed by a variable – coefficient KdV equation of the form [15]

$$u_t(x,t) + 6u(x,t)u_x(x,t) + B(t)u_{xxx}(x,t) = 0\,. \quad (190)$$

where $u(x,t)$ is the wave amplitude, t the propagation coordinate, x the temporal variable and $B(t)$ is the local dispersion coefficient.

The applicability of the variable-coefficient KdV equation (190) arises in many areas of physics as, for example, for the description of the propagation of gravity-capillary and interfacial-capillary waves, internal waves and Rossby waves [15]. In order to study the propagation of weakly nonlinear, weakly dispersive waves in inhomogeneous media, Eq. (190) is rewritten as follows [16]

$$u_t(x,t) + 6A(t)u(x,t)u_x(x,t) + B(t)u_{xxx}(x,t) = 0. \quad (191)$$

which now has a variable nonlinearity coefficient $A(t)$. Here, Eq. (191) is called a (1 + 1)-dimensional generalized variable-coefficient KdV (gvcKdV) equation.

In order to find new non-traveling waves solutions and traveling waves solutions of the following (1 + 1)-dimensional gvcKdV equation (191) by using our generalized algebra method and symbolic computation, we first take the following new general transformation, which we first present here

$$u(x,t) = u(\xi)\,, \quad \xi = \alpha(x,t)\beta(t) - r(t)\,, \quad (192)$$

where $\alpha(x,t)$, $\beta(t)$ and $r(t)$ are functions to be determined later.

By using the new variable $\phi = \phi(\xi)$ which is a solution of the following ODE

$$\frac{d\phi(\xi)}{d\xi} = \varepsilon\sqrt{\begin{matrix} c_0 + c_1\phi(\xi) + c_2\phi^2(\xi) + c_3\phi^3(\xi) \\ +c_4\phi^4(\xi) + c_5\phi^5(\xi) + c_6\phi^6(\xi) \end{matrix}}\,. \quad (193)$$

we expand the solution of Eq. (191) in the form [40,43]:

$$u = a_0(X) + \sum_{i=1}^{n}(a_i(X)\phi^i(\xi(X)) + b_i(X)\phi^{-i}(\xi(X))) \quad (194)$$

where $X = (x,t)$, $a_0(X), a_i(X), b_i(X)$ $(i = 1, 2, \ldots, n)$ are all differentiable functions of X to be determined later.

Balancing the highest derivative term with the nonlinear terms in Eq. (191) by using the derivatives with respect to the variable ϕ, we can determine the parameter $n = 2$ in (194). In addition, we take $a_0(X) = a_0(t)$, $a_i(X) = a_i(t), b_i(X) = b_i(t), i = 1, 2$ in (194) for simplicity, then substituting them and (192) into (194) along with $n = 2$, leads to:

$$u(x,t) = a_0(t) + a_1(t)\phi(\xi) + a_2(t)(\phi(\xi))^2 + \frac{b_1(t)}{\phi(\xi)} + \frac{b_2(t)}{(\phi(\xi))^2}\,. \quad (195)$$

where $\xi = \alpha(x,t)\beta(t) - r(t)$, and $\alpha(x,t)$, $\beta(t)$, $r(t)$, $a_0(t)$, $a_i(t)$ and $b_i(t)$, $i = 1, 2$ are all differentiable functions of x or t to be determined later.

By substituting (195) into the given Eq. (191) along with (193) and the derivatives of ϕ, and collecting coefficients of polynomials of ϕ^k and $\phi^\tau\sqrt{\sum_{j=1}^{6} c_j\phi^j(\xi)}$, with the aid of Maple, then setting each coefficient to zero, we will get a system of over-determined partial differential equations with respect to $A(t)$, $B(t)$, $\alpha(x,t)$, $\beta(t)$, $r(t)$,

$a_0(t)$, $a_i(t)$ and $b_i(t)$, $i = 1, 2$ as follows:

$$
\begin{aligned}
&21B(t)(\beta(t))^3 - \varepsilon a_1(t)\left(\frac{\partial}{\partial x}\alpha(x,t)\right)^3 c_4 c_6 + 18A(t)\\
&- \varepsilon\left(\frac{\partial}{\partial x}\alpha(x,t)\right)\beta(t)a_1(t)a_2(t)c_6 - 3B(t)(\beta(t))^3\\
&- \varepsilon b_1(t)\left(\frac{\partial}{\partial x}\alpha(x,t)\right)^3 {c_6}^2 = 0\,,\\
&12b_2(t) - \varepsilon\left(\frac{\mathrm{d}}{\mathrm{d}t}r(t)\right)c_2 - 8B(t)(\beta(t))^3\\
&- \varepsilon b_2(t)\left(\frac{\partial}{\partial x}\alpha(x,t)\right)^3 {c_2}^2 - 2b_2(t)\\
&- \varepsilon\left(\frac{\partial}{\partial t}\alpha(x,t)\right)\beta(t)c_2 - 2B(t)\beta(t)\\
&- \varepsilon b_2(t)\left(\frac{\partial^3}{\partial x^3}\alpha(x,t)\right)c_2 - 2b_2(t)\\
&- \varepsilon\alpha(x,t)\left(\frac{\mathrm{d}}{\mathrm{d}t}\beta(t)\right)c_2 - 6A(t)\\
&- \varepsilon\left(\frac{\partial}{\partial x}\alpha(x,t)\right)\beta(t)(b_1(t))^2 c_2\\
&- 12A(t) - \varepsilon\left(\frac{\partial}{\partial x}\alpha(x,t)\right)\beta(t)a_0(t)b_2(t)c_2 - 12A(t)\\
&- \varepsilon\left(\frac{\partial}{\partial x}\alpha(x,t)\right)\beta(t)(b_2(t))^2 c_4 = 0\,,\\
&b_1(t) - \varepsilon\left(\frac{\mathrm{d}}{\mathrm{d}t}r(t)\right)c_6 - a_1(t) - \varepsilon\left(\frac{\mathrm{d}}{\mathrm{d}t}r(t)\right)c_4\\
&+ 6A(t) - \varepsilon\left(\frac{\partial}{\partial x}\alpha(x,t)\right)\beta(t)a_2(t)b_1(t)c_4\\
&+ a_1(t) - \varepsilon\alpha(x,t)\left(\frac{\mathrm{d}}{\mathrm{d}t}\beta(t)\right)c_4 - b_1(t)\\
&- \varepsilon\alpha(x,t)\left(\frac{\mathrm{d}}{\mathrm{d}t}\beta(t)\right)c_6 + 6A(t)\\
&- \varepsilon\left(\frac{\partial}{\partial x}\alpha(x,t)\right)\beta(t)a_0(t)a_1(t)c_4 - 6A(t)\\
&- \varepsilon\left(\frac{\partial}{\partial x}\alpha(x,t)\right)\beta(t)a_0(t)b_1(t)c_6\\
&- 4B(t)(\beta(t))^3 - \varepsilon b_1(t)\left(\frac{\partial}{\partial x}\alpha(x,t)\right)^3 c_2 c_6\\
&+ B(t)\beta(t) - \varepsilon a_1(t)\left(\frac{\partial^3}{\partial x^3}\alpha(x,t)\right)c_4\\
&+ a_1(t) - \varepsilon\left(\frac{\partial}{\partial t}\alpha(x,t)\right)\beta(t)c_4 - b_1(t)\\
&- \varepsilon\left(\frac{\partial}{\partial t}\alpha(x,t)\right)\beta(t)c_6 - 6A(t)\\
&- \varepsilon\left(\frac{\partial}{\partial x}\alpha(x,t)\right)\beta(t)a_1(t)b_2(t)c_6 + 7B(t)(\beta(t))^3\\
&- \varepsilon a_1(t)\left(\frac{\partial}{\partial x}\alpha(x,t)\right)^3 c_2 c_4 + 18A(t)\\
&- \varepsilon\left(\frac{\partial}{\partial x}\alpha(x,t)\right)\beta(t)a_1(t)a_2(t)c_2 - B(t)\beta(t)\\
&- \varepsilon b_1(t)\left(\frac{\partial^3}{\partial x^3}\alpha(x,t)\right)c_6 = 0\,,
\end{aligned}
\tag{196}
$$

$$\ldots\ldots\ldots\ldots$$

because there are so many over-determined partial differential equations, only a few of them are shown here for convenience. Solving the over-determined partial differential equations with Maple, we have the following solutions.

Case 1

$$
\begin{aligned}
&A(t) = A(t)\,, \quad B(t) = B(t)\,, \quad \alpha(x,t) = F_1(t)\,,\\
&\beta(t) = \beta(t)\,, \quad a_2(t) = C_1\,, \quad a_1(t) = C_2\,,\\
&b_2(t) = C_3\,, a_0(t) = C_5\,,\\
&r(t) = \int F_1(t)\frac{\mathrm{d}}{\mathrm{d}t}\beta(t) + \left(\frac{\mathrm{d}}{\mathrm{d}t}F_1(t)\right)\beta(t)\mathrm{d}t + C_6\,,\\
&b_1(t) = C_4\,,
\end{aligned}
\tag{197}
$$

where $A(t)$, $B(t)$, $\beta(t)$, $F_1(t)$ are arbitrary functions of t, and C_i, $i = 1, 2, 3, 4, 5$ are arbitrary constants.

Case 2

$$
\begin{aligned}
&b_1(t) = 0\,, \quad A(t) = A(t)\,, \quad B(t) = B(t)\,,\\
&\alpha(x,t) = F_1(t)\,, \quad \beta(t) = \beta(t)\,, \quad a_2(t) = C_1\,,\\
&a_0(t) = C_4\,, \quad a_1(t) = C_2\,,\\
&r(t) = \int F_1(t)\frac{\mathrm{d}}{\mathrm{d}t}\beta(t) + \left(\frac{\mathrm{d}}{\mathrm{d}t}F_1(t)\right)\beta(t)\mathrm{d}t + C_5\,,\\
&b_2(t) = C_3\,,
\end{aligned}
\tag{198}
$$

where $A(t)$, $B(t)$, $\beta(t)$, $F_1(t)$ are arbitrary functions of t, and C_i, $i = 1, 2, 3, 4$ are arbitrary constants.

Case 3

$$
\begin{aligned}
&\beta(t) = 0\,, \quad A(t) = A(t)\,, \quad B(t) = B(t)\,,\\
&a_2(t) = C_1\,, \quad a_1(t) = C_2\,, \quad b_2(t) = C_3\,,\\
&a_0(t) = C_5\,, \quad b_1(t) = C_4\,, \quad r(t) = C_6\,,\\
&\alpha(x,t) = \alpha(x,t)\,,
\end{aligned}
\tag{199}
$$

where $\alpha(x,t)$ are arbitrary functions of x and t, $A(t)$, $B(t)$ are arbitrary functions of t, and C_i, $i = 1, 2, 3, 4, 5$ are arbitrary constants.

Case 4

$$\begin{aligned}
&b_2(t)=0\,,\quad A(t)=A(t)\,,\quad B(t)=B(t)\,,\\
&\alpha(x,t)=F_1(t)\,,\quad \beta(t)=\beta(t)\,,\quad a_2(t)=C_1\,,\\
&a_0(t)=C_4\,, a_1(t)=C_2\,,\\
&r(t)=\int F_1(t)\frac{\mathrm{d}}{\mathrm{d}t}\beta(t)+\left(\frac{\mathrm{d}}{\mathrm{d}t}F_1(t)\right)\beta(t)\mathrm{d}t+C_5\,,\\
&b_1(t)=C_3\,,
\end{aligned}\tag{200}$$

where $A(t)$, $B(t)$, $F_1(t)$, $\beta(t)$ are arbitrary functions of t, and C_i, $i=1,2,3,4,5$ are arbitrary constants.

Case 5

$$\begin{aligned}
&b_2(t)=0\,,\quad \beta(t)=0\,,\quad A(t)=A(t)\,,\\
&B(t)=B(t)\,,\quad a_2(t)=C_1\,,\quad a_0(t)=C_4\,,\\
&a_1(t)=C_2\,,\quad r(t)=C_5\,,\quad b_1(t)=C_3\,,\\
&\alpha(x,t)=\alpha(x,t)\,,
\end{aligned}\tag{201}$$

where $\alpha(x,t)$ are arbitrary functions of x and t, $A(t)$, $B(t)$ are arbitrary functions of t, and C_i, $i=1,2,3,4,5$ are arbitrary constants.

Case 6

$$\begin{aligned}
&b_1(t)=\frac{C_2c_4}{c_6}\,,\quad \beta(t)=0\,,\quad b_2(t)=0\,,\\
&A(t)=A(t)\,,\quad B(t)=B(t)\,,\quad a_2(t)=C_1\,,\\
&a_1(t)=C_2\,,\quad r(t)=C_4\,,\quad a_0(t)=C_3\,,\\
&\alpha(x,t)=\alpha(x,t)\,,
\end{aligned}\tag{202}$$

where $\alpha(x,t)$ are arbitrary functions of x and t, $A(t)$, $B(t)$ are arbitrary functions of t, and C_i, $i=1,2,3,4$ are arbitrary constants.

Case 7

$$\begin{aligned}
&B(t)=B(t)\,,\quad b_1(t)=0\,,\quad a_0(t)=C_3\,,\\
&A(t)=0\,,\quad b_2(t)=C_2\,,\quad a_1(t)=0\,,\quad a_2(t)=0\,,\\
&\alpha(x,t)=F_1(t)x+F_2(t)\,,\quad \beta(t)=\frac{C_1}{F_1(t)}\,,\\
&r(t)=\int\frac{C_1,\left(F_1(t)\frac{\mathrm{d}}{\mathrm{d}t}F_2(t)-\left(\frac{\mathrm{d}}{\mathrm{d}t}F_1(t)\right)F_2(t)+4c_2B(t){C_1}^2(F_1(t))^2\right)}{(F_1(t))^2}\mathrm{d}t+C_4\,,
\end{aligned}\tag{203}$$

where $F_1(t)$, $B(t)$ are arbitrary functions of t, and C_i, $i=1,2,3,4$ are arbitrary constants.

Case 8

$$\begin{aligned}
&B(t)=B(t)\,,\quad b_1(t)=0\,,\quad \alpha(x,t)=F_1(t)\,,\\
&\beta(t)=\beta(t)\,,\quad A(t)=0\,,\quad a_1(t)=0\,,\\
&a_2(t)=0\,,\\
&r(t)=\int F_1(t)\frac{\mathrm{d}}{\mathrm{d}t}\beta(t)+\left(\frac{\mathrm{d}}{\mathrm{d}t}F_1(t)\right)\beta(t)\mathrm{d}t+C_3\,,\\
&a_0(t)=C_2\,,\quad b_2(t)=C_1\,,
\end{aligned}\tag{204}$$

where $A(t)$, $B(t)$, $\beta(t)$, $F_1(t)$ are arbitrary functions of t, and C_i, $i=1,2,3$ are arbitrary constants.

Case 9

$$\begin{aligned}
&A(t)=A(t)\,,\quad B(t)=B(t)\,,\quad b_1(t)=0\,,\\
&\alpha(x,t)=F_1(t)\,,\quad \beta(t)=\beta(t)\,,\\
&a_2(t)=C_1\,,\quad a_0(t)=C_3\,,\quad b_2(t)=C_2\,,\\
&a_1(t)=0\,,\\
&r(t)=\int F_1(t)\frac{\mathrm{d}}{\mathrm{d}t}\beta(t)+\left(\frac{\mathrm{d}}{\mathrm{d}t}F_1(t)\right)\beta(t)\mathrm{d}t+C_4\,,
\end{aligned}\tag{205}$$

where $A(t)$, $B(t)$, $\beta(t)$, $F_1(t)$ are arbitrary functions of t, and C_i, $i=1,2,3,4$ are arbitrary constants.

Case 10

$$\begin{aligned}
&A(t)=A(t)\,,\quad B(t)=B(t)\,,\quad b_1(t)=0\,,\\
&\alpha(x,t)=F_1(t)\,,\quad \beta(t)=\beta(t)\,,\\
&a_2(t)=C_1\,,\quad b_2(t)=0\,,\quad a_1(t)=0\,,\\
&r(t)=\int F_1(t)\frac{\mathrm{d}}{\mathrm{d}t}\beta(t)+\left(\frac{\mathrm{d}}{\mathrm{d}t}F_1(t)\right)\beta(t)\mathrm{d}t+C_3\,,\\
&a_0(t)=C_2\,,
\end{aligned}\tag{206}$$

where $A(t)$, $B(t)$, $\beta(t)$, $F_1(t)$ are arbitrary functions of t, and C_i, $i=1,2,3,4$ are arbitrary constants.

Case 11

$$\begin{aligned}
&A(t)=A(t)\,,\quad B(t)=B(t)\,,\quad b_1(t)=0\,,\\
&a_2(t)=C_1\,,\quad b_2(t)=0\,,\quad a_1(t)=0\,,\\
&\beta(t)=0\,,\quad r(t)=C_3\,,\\
&\alpha(x,t)=\alpha(x,t)\,,\quad a_0(t)=C_2\,,
\end{aligned}\tag{207}$$

where $\alpha(x,t)$ are arbitrary functions of x and t, $A(t)$, $B(t)$ are arbitrary functions of t, and C_i, $i=1,2$ are arbitrary constants. Because there are so many solutions, only a few of them are shown here for convenience.

So we get new general forms of solutions of equations (191):

$$u(x,t) = a_0(t) + a_1(t)\phi(\alpha(x,t)\beta(t) - r(t)) + a_2(t)(\phi(\alpha(x,t)\beta(t) - r(t)))^2 + \frac{b_1(t)}{\phi(\alpha(x,t)\beta(t) - r(t))} + \frac{b_2(t)}{(\phi(\alpha(x,t)\beta(t) - r(t)))^2}. \tag{208}$$

where $\alpha(x,t)$, $\beta(t)$, $r(t)$, $a_0(t)$, $a_i(t)$, $b_i(t)$, $i = 1, 2$ satisfy (197)–(207) respectively, and the variable $\phi(\alpha(x,t)\beta(t) - r(t))$ takes the solutions of the Eq. (193). For example, we may take the variable $\phi(\alpha(x,t)\beta(t) - r(t))$ as follows:

Type 1. If $4c_2c_6 - c_4{}^2 < 0$, corresponding to Eq. (193), $\phi(\alpha(x,t)\beta(t) - r(t))$ is taken, we get the following four tanh-sech hyperbolic type solutions

$$\phi_{1,2}((\alpha(x,t)\beta(t) - r(t))) = \varepsilon\sqrt{\frac{2c_2\,\mathrm{sech}(\eta)[c_4\,\mathrm{sech}(\eta) + \varepsilon i\sqrt{c_4{}^2 - 4c_6c_2}\tanh(\eta)]}{4c_6c_2\tanh^2(\eta) - c_4{}^2}}, \quad c_2 > 0\,, \tag{209}$$

$$\phi_{3,4}((\alpha(x,t)\beta(t) - r(t))) = \varepsilon\sqrt{\frac{2c_2\,\mathrm{sech}(\eta)[c_4\,\mathrm{sech}(\eta) - \varepsilon i\sqrt{c_4{}^2 - 4c_6c_2}\tanh(\eta)]}{4c_6c_2\tanh^2(\eta) - c_4{}^2}}, \quad c_2 > 0\,, \tag{210}$$

and the following tan-sec triangular type solutions

$$\phi_{5,6}((\alpha(x,t)\beta(t) - r(t))) = \varepsilon\sqrt{\frac{2c_2\sec(\zeta)[-c_4\sec(\zeta) + \varepsilon\sqrt{c_4{}^2 - 4c_6c_2}\tan(\zeta)]}{4c_6c_2\tan^2(\zeta) + c_4{}^2}}, \quad c_2 < 0\,, \tag{211}$$

$$\phi_{7,8}((\alpha(x,t)\beta(t) - r(t))) = \varepsilon\sqrt{\frac{2c_2\sec(\zeta)[-c_4\sec(\zeta) - \varepsilon\sqrt{c_4{}^2 - 4c_6c_2}\tan(\zeta)]}{4c_6c_2\tan^2(\zeta) + c_4{}^2}}, \quad c_2 < 0\,, \tag{212}$$

where $\eta = 2\sqrt{c_2}((\alpha(x,t)\beta(t) - r(t)) - C_1)$ and C_1 is any constant, $\zeta = 2\sqrt{-c_2}((\alpha(x,t)\beta(t) - r(t)) - C_1)$ and C_1 is any constant.

Type 2. If $4c_2c_6 - c_4{}^2 > 0$, corresponding to Eq. (193), $\phi(\alpha(x,t)\beta(t) - r(t))$ is taken, we get the following four sinh-cosh hyperbolic type solutions

$$\phi_{9,10}((\alpha(x,t)\beta(t) - r(t))) = \varepsilon\sqrt{\frac{2c_2[c_4 + \varepsilon\sqrt{4c_2c_6 - c_4{}^2}\sinh(\eta)]}{4c_2c_6\sinh^2(\eta) - c_4{}^2\cosh^2(\eta)}}, \quad c_2 > 0\,, \tag{213}$$

$$\phi_{11,12}((\alpha(x,t)\beta(t) - r(t))) = \varepsilon\sqrt{\frac{2c_2[c_4 - \varepsilon\sqrt{4c_2c_6 - c_4{}^2}\sinh(\eta)]}{4c_2c_6\sinh^2(\eta) - c_4{}^2\cosh^2(\eta)}}, \quad c_2 > 0\,, \tag{214}$$

and the following sin-cos triangular type solutions

$$\phi_{13,14}((\alpha(x,t)\beta(t) - r(t))) = \varepsilon\sqrt{\frac{2c_2[-c_4 + \varepsilon i\sqrt{4c_2c_6 - c_4{}^2}\sin(\zeta)]}{4c_6c_2\sin^2(\zeta) + c_4{}^2\cos^2(\zeta)}}, \quad c_2 < 0 \tag{215}$$

$$\phi_{15,16}((\alpha(x,t)\beta(t) - r(t))) = \varepsilon\sqrt{\frac{2c_2[-c_4 - \varepsilon i\sqrt{4c_6c_2 - c_4{}^2}\sin(\zeta)]}{4c_6c_2\sin^2(\zeta) + {}_4{}^2\cos^2(\zeta)}}, \quad c_2 < 0, \tag{216}$$

where $\eta = 2\sqrt{c_2}((\alpha(x,t)\beta(t) - r(t)) - C_1)$, $\zeta = 2\sqrt{-c_2}((\alpha(x,t)\beta(t) - r(t)) - C_1)$,and C_1 is any constant.

Type 3. If $4c_2c_6 - c_4{}^2 = 0$, corresponding to Eq. (193), $\phi(\alpha(x,t)\beta(t) - r(t))$ is taken, we get the following two sech-tanh hyperbolic type solutions $4c_2c_6 - c_4{}^2 = 0$, Eq. (170) admits the following tanh hyperbolic type solutions

$$\phi_{17,18,19,20}(\alpha(x,t)\beta(t) - r(t)) = \varepsilon\sqrt{\frac{2c_2\{1 + \varepsilon\tanh[\sqrt{c_2}(\alpha(x,t)\beta(t) - r(t) - C_1)]\}}{1 - c_4\varepsilon(1 + c_4)\tanh[\sqrt{c_2}(\alpha(x,t)\beta(t) - r(t) - C_1)]}}, \quad c_2 > 0\,, \tag{217}$$

and the following tan triangular type solutions

$$\phi_{21,22,23,24}(\alpha(x,t)\beta(t) - r(t)) = \varepsilon\sqrt{\frac{2c_2\{1 + \varepsilon\tan[i\sqrt{-c_2}(\alpha(x,t)\beta(t) - r(t) - C_1)]\}}{1 - c_4 - \varepsilon(1 + c_4)\tan[i\sqrt{-c_2}(\alpha(x,t)\beta(t) - r(t) - C_1)]}}, \quad c_2 < 0\,, \tag{218}$$

where C_1 is any constant.

Substituting (209)–(218) and (197)–(207) into (208), respectively, we get many new irrational solutions and rational solutions of combined hyperbolic type or triangular type solutions of Eq. (191).

For example, when we select $A(t), B(t), \alpha(x,t), \beta(t), r(t), a_0(t), a_i(t)$ and $b_i(t), i = 1, 2$ to satisfy Case 1, we can easily get the following irrational solutions of combined hyperbolic type or triangular type solutions of Eq. (191):

$$\begin{aligned} u(x,t) &= C_5 + C_2\phi(F_1(t)\beta(t) - r(t)) \\ &+ C_1(\phi(F_1(t)\beta(t) - r(t)))^2 \\ &+ \frac{C_4}{\phi(F_1(t)\beta(t) - r(t))} + \frac{C_3}{(\phi(F_1(t)\beta(t) - r(t)))^2}, \end{aligned} \tag{219}$$

where $A(t), B(t), \beta(t), F_1(t)$ are arbitrary functions of t, C_i, $i = 1, 2, 3, 4, 5$ are arbitrary constants, and

$$r(t) = \int F_1(t)\frac{\mathrm{d}}{\mathrm{d}t}\beta(t) + \left(\frac{\mathrm{d}}{\mathrm{d}t}F_1(t)\right)\beta(t)\mathrm{d}t + C_6 .$$

Substituting (209)–(218) into (219), respectively, we get the following new irrational solutions and rational solutions of combined hyperbolic type or triangular type solutions of Eq. (191).

Case 1. If $4c_2c_6 - c_4{}^2 < 0, c_2 > 0$, corresponding to (209), (210) and (197), Eq. (191) admits the following four tanh-sech hyperbolic type solutions

$$\begin{aligned} u_{1,2}(x,t) &= C_5 \\ &+ \frac{C_4}{\varepsilon\sqrt{\frac{2c_2\operatorname{sech}(\eta)[c_4\operatorname{sech}(\eta)\ +\varepsilon i\sqrt{c_4^2 - 4c_6c_2}\tanh(\eta)]}{4c_6c_2\tanh^2(\eta) - c_4{}^2}}} \\ &+ \frac{C_3(4c_6c_2\tanh^2(\eta) - c_4{}^2)}{2c_2\operatorname{sech}(\eta)[c_4\operatorname{sech}(\eta) + \varepsilon i\sqrt{c_4{}^2 - 4c_6c_2}\tanh(\eta)]} \\ &+ C_2\varepsilon\sqrt{\frac{2c_2\operatorname{sech}(\eta)[c_4\operatorname{sech}(\eta) + \varepsilon i\sqrt{c_4{}^2 - 4c_6c_2}\tanh(\eta)]}{4c_6c_2\tanh^2(\eta) - c_4{}^2}} \\ &+ \frac{2c_2C_1\operatorname{sech}(\eta)[c_4\operatorname{sech}(\eta) + \varepsilon i\sqrt{c_4{}^2 - 4c_6c_2}\tanh(\eta)]}{4c_6c_2\tanh^2(\eta) - c_4{}^2}, \end{aligned} \tag{220}$$

and

$$\begin{aligned} u_{3,4}(x,t) &= C_5 \\ &+ \frac{C_3(4c_6c_2\tanh^2(\eta) - c_4{}^2)}{2c_2\operatorname{sech}(\eta)[c_4\operatorname{sech}(\eta) - \varepsilon i\sqrt{c_4{}^2 - 4c_6c_2}\tanh(\eta)]} \\ &+ \frac{C_4}{\varepsilon\sqrt{\frac{2c_2\operatorname{sech}(\eta)[c_4\operatorname{sech}(\eta) - \varepsilon i\sqrt{c_4{}^2 - 4c_6c_2}\tanh(\eta)]}{4c_6c_2\tanh^2(\eta) - c_4{}^2}}} \\ &+ C_2\varepsilon\sqrt{\frac{2c_2\operatorname{sech}(\eta)[c_4\operatorname{sech}(\eta) - \varepsilon i\sqrt{c_4{}^2 - 4c_6c_2}\tanh(\eta)]}{4c_6c_2\tanh^2(\eta) - c_4{}^2}} \\ &+ \frac{2c_2C_1\operatorname{sech}(\eta)[c_4\operatorname{sech}(\eta) - \varepsilon i\sqrt{c_4{}^2 - 4c_6c_2}\tanh(\eta)]}{4c_6c_2\tanh^2(\eta) - c_4{}^2}, \end{aligned} \tag{221}$$

where $\eta = 2\sqrt{c_2}((\alpha(x,t)\beta(t) - r(t)) - C_1)$, $A(t), B(t), \beta(t), F_1(t)$ are arbitrary functions of t, C_i, $i = 1, 2, 3, 4, 5$ are arbitrary constants, and

$$r(t) = \int F_1(t)\frac{\mathrm{d}}{\mathrm{d}t}\beta(t) + \left(\frac{\mathrm{d}}{\mathrm{d}t}F_1(t)\right)\beta(t)\mathrm{d}t + C_6 .$$

Case 2. If $4c_2c_6 - c_4{}^2 < 0, c_2 < 0$, corresponding to (211), (212) and (197), Eq. (191) admits the following four tan-sec triangular type solutions

$$\begin{aligned} u_{5,6}(x,t) &= C_5 \\ &+ \frac{C_4}{\varepsilon\sqrt{\frac{2c_2\sec(\zeta)[-c_4\sec(\zeta) + \varepsilon\sqrt{c_4{}^2 - 4c_6c_2}\tan(\zeta)]}{4c_6c_2\tan^2(\zeta) + c_4{}^2}}} \\ &+ \frac{C_3(4c_6c_2\tan^2(\zeta) + c_4{}^2)}{2c_2\sec(\zeta)[-c_4\sec(\zeta) + \varepsilon\sqrt{c_4{}^2 - 4c_6c_2}\tan(\zeta)]} \\ &+ C_2\varepsilon\sqrt{\frac{2c_2\sec(\zeta)[-c_4\sec(\zeta) + \varepsilon\sqrt{c_4{}^2 - 4c_6c_2}\tan(\zeta)]}{4c_6c_2\tan^2(\zeta) + c_4{}^2}} \\ &+ \frac{2c_2C_1\sec(\zeta)[-c_4\sec(\zeta) + \varepsilon\sqrt{c_4{}^2 - 4c_6c_2}\tan(\zeta)]}{4c_6c_2\tan^2(\zeta) + c_4{}^2}, \end{aligned} \tag{222}$$

and

$$\begin{aligned} u_{7,8}(x,t) &= C_5 \\ &+ \frac{C_4}{\varepsilon\sqrt{\frac{2c_2\sec(\zeta)[-c_4\sec(\zeta) - \varepsilon\sqrt{c_4{}^2 - 4c_6c_2}\tan(\zeta)]}{4c_6c_2\tan^2(\zeta) + c_4{}^2}}} \end{aligned}$$

$$+\frac{C_3(4c_6c_2\tan^2(\zeta)+{c_4}^2)}{2c_2\sec(\zeta)[-c_4\sec(\zeta)-\varepsilon\sqrt{{c_4}^2-4c_6c_2}\tan(\zeta)]}+C_2\varepsilon\sqrt{\frac{2c_2\sec(\zeta)[-c_4\sec(\zeta)-\varepsilon\sqrt{{c_4}^2-4c_6c_2}\tan(\zeta)]}{4c_6c_2\tan^2(\zeta)+{c_4}^2}}+\frac{2c_2C_1\sec(\zeta)[-c_4\sec(\zeta)-\varepsilon\sqrt{{c_4}^2-4c_6c_2}\tan(\zeta)]}{4c_6c_2\tan^2(\zeta)+{c_4}^2}, \tag{223}$$

where $\zeta=2\sqrt{-c_2}((\alpha(x,t)\beta(t)-r(t))-C_1)$ and the rest of the parameters are the same as with Case 1.

Case 3. If $4c_2c_6-{c_4}^2>0, c_2>0$, corresponding to (213), (214) and (197), Eq. (191) admits the following four sinh-cosh hyperbolic type solutions

$$u_{9,10}(x,t)=C_5+C_2\varepsilon\sqrt{\frac{2c_2[c_4+\varepsilon\sqrt{4c_2c_6-{c_4}^2}\sinh(\eta)]}{4c_2c_6\sinh^2(\eta)-{c_4}^2\cosh^2(\eta)}}+\frac{2c_2C_1[c_4+\varepsilon\sqrt{4c_2c_6-{c_4}^2}\sinh(\eta)]}{4c_2c_6\sinh^2(\eta)-{c_4}^2\cosh^2(\eta)}+\frac{C_4}{\varepsilon\sqrt{\frac{2c_2[c_4+\varepsilon\sqrt{4c_2c_6-{c_4}^2}\sinh(\eta)]}{4c_2c_6\sinh^2(\eta)-{c_4}^2\cosh^2(\eta)}}}+\frac{2c_2C_3[c_4+\varepsilon\sqrt{4c_2c_6-{c_4}^2}\sinh(\eta)]}{4c_2c_6\sinh^2(\eta)-{c_4}^2\cosh^2(\eta)}, \tag{224}$$

and

$$u_{11,12}(x,t)=C_5+C_2\varepsilon\sqrt{\frac{2c_2[c_4+\varepsilon\sqrt{4c_2c_6-{c_4}^2}\sinh(\eta)]}{4c_2c_6\sinh^2(\eta)-{c_4}^2\cosh^2(\eta)}}+\frac{2c_2C_1[c_4+\varepsilon\sqrt{4c_2c_6-{c_4}^2}\sinh(\eta)]}{4c_2c_6\sinh^2(\eta)-{c_4}^2\cosh^2(\eta)}+\frac{C_4}{\varepsilon\sqrt{\frac{2c_2[c_4+\varepsilon\sqrt{4c_2c_6-{c_4}^2}\sinh(\eta)]}{4c_2c_6\sinh^2(\eta)-{c_4}^2\cosh^2(\eta)}}}+\frac{2c_2C_3[c_4+\varepsilon\sqrt{4c_2c_6-{c_4}^2}\sinh(\eta)]}{4c_2c_6\sinh^2(\eta)-{c_4}^2\cosh^2(\eta)}, \tag{225}$$

where $\eta=2\sqrt{c_2}((\alpha(x,t)\beta(t)-r(t))-C_1)$ and the rest of the parameters are the same as with Case 1.

Case 4. If $4c_2c_6-{c_4}^2>0, c_2<0$, corresponding to (215), (216) and (197), Eq. (191) admits the following four sin-cos triangular type solutions

$$u_{13,14}(x,t)=C_5+\frac{C_4}{\varepsilon\sqrt{\frac{2c_2[-c_4+\varepsilon i\sqrt{4c_2c_6-{c_4}^2}\sin(\zeta)]}{4c_6c_2\sin^2(\zeta)+{c_4}^2\cos^2(\zeta)}}}+\frac{C_3(4c_6c_2\sin^2(\zeta)+{c_4}^2\cos^2(\zeta))}{2c_2[-c_4+\varepsilon i\sqrt{4c_2c_6-{c_4}^2}\sin(\zeta)]}+C_2\varepsilon\sqrt{\frac{2c_2[-c_4+\varepsilon i\sqrt{4c_2c_6-{c_4}^2}\sin(\zeta)]}{4c_6c_2\sin^2(\zeta)+{c_4}^2\cos^2(\zeta)}}+\frac{2c_2C_1\left[-c_4+\varepsilon i\sqrt{4c_2c_6-{c_4}^2}\sin(\zeta)\right]}{4c_6c_2\sin^2(\zeta)+{c_4}^2\cos^2(\zeta)}, \tag{226}$$

and

$$u_{15,16}(x,t)=C_5+\frac{C_4}{\varepsilon\sqrt{\frac{2c_2\left[-c_4-\varepsilon i\sqrt{4c_2c_6-{c_4}^2}\sin(\zeta)\right]}{4c_6c_2\sin^2(\zeta)+{c_4}^2\cos^2(\zeta)}}}+\frac{C_3(4c_6c_2\sin^2(\zeta)+{c_4}^2\cos^2(\zeta))}{2c_2\left[-c_4-\varepsilon i\sqrt{4c_2c_6-{c_4}^2}\sin(\zeta)\right]}+C_2\varepsilon\sqrt{\frac{2c_2\left[-c_4-\varepsilon i\sqrt{4c_2c_6-{c_4}^2}\sin(\zeta)\right]}{4c_6c_2\sin^2(\zeta)+{c_4}^2\cos^2(\zeta)}}+\frac{2c_2C_1\left[-c_4-\varepsilon i\sqrt{4c_2c_6-{c_4}^2}\sin(\zeta)\right]}{4c_6c_2\sin^2(\zeta)+{c_4}^2\cos^2(\zeta)}, \tag{227}$$

where $\zeta=2\sqrt{-c_2}((\alpha(x,t)\beta(t)-r(t))-C_1)$ and the rest of the parameters are the same as with Case 1.

Case 5. If $4c_2c_6-{c_4}^2=0, c_2>0$, corresponding to (217) and (197), Eq. (191) admits the following four tanh hyperbolic type solutions

$$u_{17,18,19,20}(x,t)=C_5+C_2-\varepsilon\sqrt{\frac{2c_2\left[1+\varepsilon\tanh\left(\frac{\eta}{2}\right)\right]}{1-c_4-\varepsilon(1+c_4)\tanh\left(\frac{\eta}{2}\right)}}+\frac{2c_2C_1\left[1+\varepsilon\tanh\left(\frac{\eta}{2}\right)\right]}{1-c_4-\varepsilon(1+c_4)\tanh\left(\frac{\eta}{2}\right)}+\frac{C_4}{-\varepsilon\sqrt{\frac{2c_2\left[1+\varepsilon\tanh\left(\frac{\eta}{2}\right)\right]}{1-c_4-\varepsilon(1+c_4)\tanh\left(\frac{\eta}{2}\right)}}}+\frac{C_3\left[1-c_4-\varepsilon(1+c_4)\tanh\left(\frac{\eta}{2}\right)\right]}{2c_2\left[1+\varepsilon\tanh\left(\frac{\eta}{2}\right)\right]}, \tag{228}$$

where $\eta=2\sqrt{c_2}((\alpha(x,t)\beta(t)-r(t))-C_1)$ and the rest of the parameters are the same as with Case 1.

Case 6. If $4c_2c_6 - {c_4}^2 = 0, c_2 < 0$, corresponding to (218) and (197), Eq. (191) admits the following two sectan triangular type solutions

$$\begin{aligned} &u_{21,22,23,24}(x,t) \\ &= C_5 + C_2 - \varepsilon\sqrt{\frac{2c_2\left[1+\varepsilon\tan\left(\frac{i\zeta}{2}\right)\right]}{1-c_4-\varepsilon(1+c_4)\tan\left(\frac{i\zeta}{2}\right)}} \\ &\quad - \frac{2c_2C_1\left[1+\varepsilon\tan\left(\frac{i\zeta}{2}\right)\right]}{1-c_4-\varepsilon(1+c_4)\tan\left(\frac{i\zeta}{2}\right)} \\ &\quad + \frac{C_4}{-\varepsilon\sqrt{\frac{2c_2\left[1+\varepsilon\tan\left(\frac{i\zeta}{2}\right)\right]}{1-c_4-\varepsilon(1+c_4)\tan\left(\frac{i\zeta}{2}\right)}}} \\ &\quad - \frac{C_3\left(1-c_4-\varepsilon(1+c_4)\tan\left(\frac{i\zeta}{2}\right)\right)}{2c_2\left[1+\varepsilon\tan\left(\frac{i\zeta}{2}\right)\right]}, \end{aligned} \tag{229}$$

where $\zeta = 2\sqrt{-c_2}((\alpha(x,t)\beta(t) - r(t)) - C_1)$ and the rest of the parameters are the same as with Case 1.

Remark 10 We may further generalize (194) as follows:

$$\begin{aligned} u &= a_0(X) \\ &+ \sum_{i=0}^{m} \frac{\sum_{r_{i1}+\cdots+r_{in}=i} a_{r_{i1},\ldots,r_{ni}}(X) G_{1i}^{r_{i1}}(\xi_{1i}(X))\ldots G_{ni}^{r_{ni}}(\xi_{ni}(X))}{\left[\sum_{l_{i1}+\cdots+l_{in}=i} b_{l_{i1},\ldots,l_{ni}}(X) G_{1i}^{l_{i1}}(\xi_{1i}(X))\ldots G_{ni}^{l_{ni}}(\xi_{ni}(X)) + c_i(X)\right]^{\tau}}, \end{aligned} \tag{230}$$

where $\sum_{l_{i1}+\cdots+l_{in}=i} b^2_{l_{i1},\ldots,l_{ni}}(X) + c_i^2(X) \neq 0$, $a_0(X)$, $\xi_{1i}(X), \ldots, \xi_{ni}(X)$, $a_{r_{i1},\ldots,r_{ni}}(X)$, $b_{l_{i1},\ldots,l_{ni}}(X)$ and $c_i(X)$, $i = 0,1,2,\ldots,m$ are all differentiable functions to be determined later, τ is a constant, $G_{1i}(\xi_{1i}(X)),\ldots$, $G_{ni}(\xi_{ni}(X))$ are all $\phi(\xi_{ki}(X))$ or $\phi^{-1}(\xi_{ki}(X))$ or some derivatives $\phi^{(j)}(\xi_{ki}(X))$, $k = 1,2,\ldots,n; i = 0,1,2,\ldots,m; j = \pm1,\pm2,\ldots$. We can get many new explicit solutions of Eq. (191).

A New Exp-N Solitary-Like Method and Its Application in the (1 + 1)-Dimensional Generalized KdV Equation

In this section, in order to develop the Exp-function method [31], we present two new generic transformations, a new Exp-N solitary-like method, and its algorithm [40, 47]. In addition, we apply our method to construct new exact solutions of the $(1+1)$-dimensional classical generalized KdV(gKdV) equation.

Summary of the Exp-N Solitary-Like Method

In the following we would like to outline the main steps of our Exp-N solitary-like method [40,47]:

Step 1. For a given NLEE system with some physical fields $u_m(t,x_1,x_2,\ldots,x_{n-1})$, $(m = 1,2,\ldots,n)$ in n independent variables $t, x_1, x_2, \ldots, x_{n-1}$,

$$\begin{aligned} &F_m(u_1,u_2,\ldots,u_n,u_{1,t},u_{2,t},\ldots,u_{n,t}, \\ &u_{1,x_1},u_{2,x_1},\ldots,u_{n,x_1},u_{1,tt},u_{2,tt},\ldots,u_{n,tt},\ldots) = 0, \\ &\qquad (m = 1,2,\ldots,n)\,. \end{aligned} \tag{231}$$

We introduce a new generic transformation

$$\begin{cases} u_m(t,x_1,\ldots,x_{n-1}) = u_m(\xi_1,\xi_2) & (m = 1,2,\ldots,n)\,, \\ \xi_j = p_{0j}(t) + \\ \sum_{i=1}^{n-1} p_{ij}(t,x_1,x_2,\ldots,x_{i-1})\cdot q_{ij}(x_i)\,, & j = 1,2\,, \end{cases} \tag{232}$$

where we let $p_{1j}(t,x_0) = p_{1j}(t)$, here $p_{ij}(t,x_1,x_2,\ldots,x_{i-1})$, $q_{ij}(x_i)$, $j = 1,2$; $i = 1,2,\ldots,n-1$ and $p_{0j}(t)$ are functions to be determined later. Then the nonlinear partial differentials of Eq. (231) are reduced to a nonlinear partial differential equation with respect to $\xi_j, j = 1,2, p_{ij}(t,x_1,x_2,\ldots,x_{i-1}), q_{ij}(x_i), j = 1,2$; $i = 1,2,\ldots,n-1$ and $p_{0j}(t)$.

Step 2. Let $X = (t,x_1,\ldots,x_{n-1})$. Then we introduce a new transformation in terms of Exp-function rational formal expansion in the following form based on the idea of the Exp-function method [41] by the new generic transformation (232).

$$u_m(\xi_1,\xi_2) = g_m(X) + \frac{\sum_{j=-c}^{d} a_{mj}(X)exp(j\xi_1)}{\sum_{k=-p}^{q} b_{mk}(X)exp(k\xi_2)}\,, \quad (m = 1,2,\ldots,n)\,, \tag{233}$$

where c, d, p and q are any positive integers, $\xi_1, \xi_2, g_m(X)$, $a_{mj}(X), j = -c,\ldots,d$, and $b_{mk}(X), k = -p,\ldots,q; m = 1,2,\ldots,n$ are functions to be determined later.

Step 3. We take a numerical value of c, d, p and q, substitute (233) into the partial differential equations obtained in Step 1, and then set all coefficients of $\exp(k\eta), k = 0,1,2,\ldots$ to zero to get an over-determined partial differential equation with respect to $g_m(X)$, $a_{mj}(X), j = -c,\ldots,d, b_{mk}(X), k = -p,\ldots,q; m = 1, 2,\ldots,n, p_{ij}(t,x_1,x_2,\ldots,x_{i-1}), q_{ij}(x_i), j = 1,2; i = 1, 2,\ldots,n-1$ and $p_{0j}(t)$.

Step 4. Solving the over-determined partial differential equations in Step 3 by use of Maple, we would end up with explicit expressions for $g_m(X), a_{mj}(X), j = -c, \dots, d$, $b_{mk}(X), k = -p, \dots, q$; $m = 1, 2, \dots, n$, $p_{ij}(t, x_1, x_2, \dots, x_{i-1}), q_{ij}(x_i), j = 1, 2; i = 1, 2, \dots, n-1$ and $p_{0j}(t)$.

Step 5. Substituting the results obtained in Step 4 into (232) and (233), we can then obtain new rational solutions of Exp-functions as follows

$$u_m(X) = g_m(X) + \frac{\sum_{j=-c}^{d} a_{mj}(X) \exp\left(j\left(p_{01}(t) + \sum_{i=1}^{n-1} p_{i1}(t, x_1, x_2, \dots, x_{i-1}) \cdot q_{i1}(x_i)\right)\right)}{\sum_{k=-p}^{q} b_{mk}(X) \exp\left(k\left(p_{02}(t) + \sum_{i=1}^{n-1} p_{i2}(t, x_1, x_2, \dots, x_{i-1}) \cdot q_{i2}(x_i)\right)\right)}, \quad (m = 1, 2, \dots, n) \qquad (234)$$

Step 6. If we again take a numerical value of c, d, p and q and repeat Step 3, Step 4, and Step 5, then we can obtain other rational solutions of Exp-functions Eq. (231). Now, repeat the process. We can obtain a family of new N solitary-like solutions of Eq. (231).

Remark 11 The new transformations (232) and (233) proposed here are more general than the transformation in the Exp-function method [31]. We have obtained many families of non-traveling waves solutions and traveling waves solutions of the $(2+1)$-dimensional generalized KdV–Burgers equation by using the two transformations in [40,47].

Remark 12 Our methods are more powerful, simpler, and more convenient than the Exp-function method [31]. The Exp-function method [31] is used only to obtain a traveling waves solution of lower dimensional NLPDE because it makes use of the homogeneous balance method. In this paper we improve the method to be more powerful such that it can be used to obtain many families of non-traveling waves solutions and traveling waves solutions of higher dimensional NLPDEs because we don't use the homogeneous balance method and make use of our transformations (232) and (233).

The Application of the Exp-N Solitary-Like Method in the $(1+1)$-Dimensional Generalized KdV Equation

In this section, we will present a new transformation, and then make use of the transformation and our method via symbolic computation to find the solutions of the $(1+1)$-dimensional classical generalized Korteweg–de Vries (gKdV) equation.

The $(1+1)$-dimensional gKdV equation has the following form [7,8,9]:

$$u_t(x,t) + \alpha u^p(x,t) u_x(x,t) + \beta u_{xxx}(x,t) = 0\,, \qquad (235)$$

where the coefficient of the nonlinear term α and the coefficient of the dispersive term β are independent of x and t.

According to the above steps, to seek traveling wave solutions of Eq. (235), we present the following new transformation here. When $p \geq 2$

$$u(x,t) = f^{\frac{2}{p}}(x,t)\,. \qquad (236)$$

By substituting (236) into Eq. (235), we get the following equation:

$$\begin{aligned}
&p^2 f^2(x,t) f_t(x,t) + \alpha p^2 f^4(x,t) f_x(x,t) \\
&+ 4\beta f_x^3(x,t) + 6p\beta f(x,t) f_x(x,t) f_{xx}(x,t) \\
&- 6p\beta f_x^3(x,t) + \beta p^2 f_{xxx}(x,t) f^2(x,t) \\
&- 3\beta p^2 f(x,t) f_x(x,t) f_{xx}(x,t) + 2\beta p^2 f_x^3(x,t) = 0\,.
\end{aligned} \qquad (237)$$

We introduce a new generic transformation [40,47]

$$f(x,t) = g + \frac{\sum_{i=-m_1}^{m_2} k_i e^{i(\alpha_1 x + \beta_1 t + \gamma_1)}}{\sum_{j=-n_1}^{n_2} h_j e^{j(\alpha_2 x + \beta_2 t + \gamma_2)}}\,, \qquad (238)$$

where m_1, m_2, n_1, n_2 are any positive integers, $g, \alpha_1, \beta_1, \alpha_2, \beta_2, \gamma_1, \gamma_2, k_i, i = -m_1, \dots, m_2$, and $h_j, j = -n_1, \dots, n_2$ are all constants to be determined later.

With the aid of Maple, substituting (238) into (237) and setting $\xi = \alpha_1 x + \beta_1 t + \gamma_1, \eta = \alpha_2 x + \beta_2 t + \gamma_2$, yields the following equation (because the equation is so long, just part of the equation is shown here for simplification)

$$\begin{aligned}
&\alpha p^2 \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^4 \alpha_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} \sum_{j=-n_1}^{n_2} h_j e^{j\eta} \\
&- p^2 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^4 \sum_{i=-m_1}^{m_2} k_i e^{i\xi} \beta_2 \\
&\quad \sum_{j=-n_1}^{n_2} h_j j e^{j\eta} \beta\, p^2 g^2 \\
&+ \alpha_1{}^3 \sum_{i=-m_1}^{m_2} k_i i^3 e^{i\xi} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^5 g^2 \\
&- \beta\, p^2 \sum_{i=-m_1}^{m_2} k_i e^{i\xi} \alpha_2{}^3 \\
&\quad \sum_{j=-n_1}^{n_2} h_j j^3 e^{j\eta} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^4 g^2 \\
&+ \alpha\, p^2 g^4 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^5 \alpha_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} \\
&+ 6\,\beta\, p \alpha_1{}^3 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^4 \\
&\quad \sum_{i=-m_1}^{m_2} k_i i^2 e^{i\xi} g
\end{aligned}$$

$$+ 6\,\beta\,p \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^2 \alpha_2{}^3 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta} \sum_{j=-n_1}^{n_2} h_j j^2 e^{j\eta} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^2 g$$

$$- 12\,\beta\,p \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^2 \alpha_2{}^3 \left(\sum_{j=-n_1}^{n_2} h_j j e^{j\eta}\right)^3 g \sum_{j=-n_1}^{n_2} h_j e^{j\eta}$$

$$+ 12\,\beta\,\alpha_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} \sum_{j=-n_1}^{n_2} h_j e^{j\eta} \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^2 \alpha_2{}^2 \left(\sum_{j=-n_1}^{n_2} h_j j e^{j\eta}\right)^2$$

$$- 12\,\beta\,\alpha_1{}^2 \left(\sum_{i=-m_1}^{m_2} k_i i e^{i\xi}\right)^2 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^2 \sum_{i=-m_1}^{m_2} k_i e^{i\xi} \alpha_2 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta}$$

$$+ 4\,\alpha\,p^2 g \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^2 \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^3 \alpha_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi}$$

$$- 4\,\alpha\,p^2 g^3 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^2 \alpha_2 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta}$$

$$+ 6\,\alpha\,p^2 g^2 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^2 \alpha_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi}$$

$$- 2\,p^2 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^2 \beta_2 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta} g$$

$$- 3\,\beta\,p^2 \alpha_1{}^3 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^4 \sum_{i=-m_1}^{m_2} k_i i^2 e^{i\xi} g$$

$$- 4\,\alpha\,p^2 g \sum_{j=-n_1}^{n_2} h_j e^{j\eta} \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^4 \alpha_2 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta}$$

$$+ 2\,\beta\,p^2 \alpha_1{}^3 \sum_{i=-m_1}^{m_2} k_i i^3 e^{i\xi} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^4 g \sum_{i=-m_1}^{m_2} k_i e^{i\xi}$$

$$- 6\,\beta\,p^2 \sum_{i=-m_1}^{m_2} k_i e^{i\xi} \alpha_2{}^3 \left(\sum_{j=-n_1}^{n_2} h_j j e^{j\eta}\right)^3 g^2 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^2$$

$$+ 3\,\beta\,p^2 \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^3 \alpha_2{}^3 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta} \sum_{j=-n_1}^{n_2} h_j j^2 e^{j\eta} \sum_{j=-n_1}^{n_2} h_j e^{j\eta}$$

$$+ 6\,\beta\,p \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^3 \alpha_2{}^3 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta} \sum_{j=-n_1}^{n_2} h_j j^2 e^{j\eta} \sum_{j=-n_1}^{n_2} h_j e^{j\eta}$$

$$+ 6\,\beta\,p^2 \alpha_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} \alpha_2{}^2 \left(\sum_{j=-n_1}^{n_2} h_j j e^{j\eta}\right)^2 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 g^2$$

$$- 3\,\beta\,p^2 \alpha_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} \alpha_2{}^2 \sum_{j=-n_1}^{n_2} h_j j^2 e^{j\eta} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^4 g^2$$

$$+ 6\,\beta\,p\alpha_1{}^3 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 \sum_{i=-m_1}^{m_2} k_i i^2 e^{i\xi} \sum_{i=-m_1}^{m_2} k_i e^{i\xi}$$

$$- 6\,\beta\,p\alpha_1{}^3 \left(\sum_{i=-m_1}^{m_2} k_i i e^{i\xi}\right)^3 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3$$

$$- 2\,\beta\,p^2 \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^3 \alpha_2{}^3 \left(\sum_{j=-n_1}^{n_2} h_j j e^{j\eta}\right)^3$$

$$- 6\,\beta\,p \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^3 \alpha_2{}^3 \left(\sum_{j=-n_1}^{n_2} h_j j e^{j\eta}\right)^3$$

$$+ 2\,\beta\,p^2 \alpha_1{}^3 \left(\sum_{i=-m_1}^{m_2} k_i i e^{i\xi}\right)^3 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3$$

$$- 6\,\alpha\,p^2 g^2 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^2 \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^3 \alpha_2 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta} +$$

$$6\,\beta\,p\alpha_1{}^2 \left(\sum_{i=-m_1}^{m_2} k_i i e^{i\xi}\right)^2 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^2 \sum_{i=-m_1}^{m_2} k_i e^{i\xi} \alpha_2 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta}$$

$$+ 9\,\beta\,p^2 \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^2 \alpha_2{}^3 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta} \sum_{j=-n_1}^{n_2} h_j j^2 e^{j\eta} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^2 g$$

$$- 6\,\beta\,p \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^2 \alpha_2 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta} \alpha_1{}^2 \sum_{i=-m_1}^{m_2} k_i i^2 e^{i\xi} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^2$$

$$+ 6\,\beta\,p^2 \alpha_1{}^2 \left(\sum_{i=-m_1}^{m_2} k_i i e^{i\xi}\right)^2 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 \alpha_2 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta} g$$

$$- 6\,\beta\,p\alpha_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^2 \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^2 \alpha_2{}^2 \sum_{j=-n_1}^{n_2} h_j j^2 e^{j\eta}$$

$$+ 6\,\beta\,p\alpha_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} \sum_{j=-n_1}^{n_2} h_j e^{j\eta} \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^2 \alpha_2{}^2 \left(\sum_{j=-n_1}^{n_2} h_j j e^{j\eta}\right)^2$$

$$- 2\,\beta\,p^2 \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^2 \alpha_2{}^3 \sum_{j=-n_1}^{n_2} h_j j^3 e^{j\eta} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 g$$

$$+ 4\,\alpha\,p^2 g^3 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^4 \sum_{i=-m_1}^{m_2} k_i e^{i\xi} \alpha_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi}$$

$$+ 2\,p^2 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^4 \beta_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} g \sum_{i=-m_1}^{m_2} k_i e^{i\xi}$$

$$- 6\,\beta\,p\alpha_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 \sum_{i=-m_1}^{m_2} k_i e^{i\xi} \alpha_2{}^2 \sum_{j=-n_1}^{n_2} h_j j^2 e^{j\eta} g$$

$$- 6\,\beta\,p^2 \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^2 \alpha_2{}^3 \left(\sum_{j=-n_1}^{n_2} h_j j e^{j\eta}\right)^3 g \sum_{j=-n_1}^{n_2} h_j e^{j\eta}$$

$$- 3\,\beta\,p^2 \alpha_1{}^3 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 \sum_{i=-m_1}^{m_2} k_i i^2 e^{i\xi} \sum_{i=-m_1}^{m_2} k_i e^{i\xi}$$

$$- 3\,\beta\,p^2 \alpha_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 \sum_{i=-m_1}^{m_2} k_i e^{i\xi} \alpha_2{}^2 \sum_{j=-n_1}^{n_2} h_j j^2 e^{j\eta} g$$

$$+ 24\,\beta\,p\alpha_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^2 \sum_{i=-m_1}^{m_2} k_i e^{i\xi} \alpha_2{}^2 \left(\sum_{j=-n_1}^{n_2} h_j j e^{j\eta}\right)^2 g$$

$$- 6\,\beta\,p \sum_{i=-m_1}^{m_2} k_i e^{i\xi} \alpha_2 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta} \alpha_1{}^2$$

$$
\begin{aligned}
&\quad \textstyle\sum_{i=-m_1}^{m_2} k_i i^2 e^{i\xi} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 g \\
&- 3\,\beta\, p^2 \textstyle\sum_{i=-m_1}^{m_2} k_i e^{i\xi} \alpha_2 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta} {\alpha_1}^2 \\
&\quad \textstyle\sum_{i=-m_1}^{m_2} k_i i^2 e^{i\xi} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 g \\
&- 3\,\beta\, p^2 {\alpha_1}^2 \textstyle\sum_{i=-m_1}^{m_2} k_i i^2 e^{i\xi} \alpha_2 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta} \\
&\quad \left(\textstyle\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^4 g^2 \\
&- 12\,\beta\, p {\alpha_1}^2 \left(\textstyle\sum_{i=-m_1}^{m_2} k_i i e^{i\xi}\right)^2 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 \\
&\quad \alpha_2 \textstyle\sum_{j=-n_1}^{n_2} h_j j e^{j\eta} g \\
&+ 6\,\beta\, p^2 \textstyle\sum_{i=-m_1}^{m_2} k_i e^{i\xi} {\alpha_2}^3 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta} \\
&\quad \textstyle\sum_{j=-n_1}^{n_2} h_j j^2 e^{j\eta} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 g^2 \\
&+ p^2 \left(\textstyle\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 \beta_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} \\
&\quad \left(\textstyle\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^2 \\
&+ \beta\, p^2 {\alpha_1}^3 \textstyle\sum_{i=-m_1}^{m_2} k_i i^3 e^{i\xi} \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 \\
&\quad \left(\textstyle\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^2 \\
&- \beta\, p^2 \left(\textstyle\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^3 {\alpha_2}^3 \sum_{j=-n_1}^{n_2} h_j j^3 e^{j\eta} \\
&\quad \left(\textstyle\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^2 \\
&- 4\,\beta \left(\textstyle\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^3 {\alpha_2}^3 \left(\sum_{j=-n_1}^{n_2} h_j j e^{j\eta}\right)^3 \\
&+ 4\,\beta\, {\alpha_1}^3 \left(\textstyle\sum_{i=-m_1}^{m_2} k_i i e^{i\xi}\right)^3 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^3 \\
&- p^2 \left(\textstyle\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^2 \left(\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^3 \\
&\quad \beta_2 \textstyle\sum_{j=-n_1}^{n_2} h_j j e^{j\eta} \\
&- \alpha\, p^2 \left(\textstyle\sum_{i=-m_1}^{m_2} k_i e^{i\xi}\right)^5 \alpha_2 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta} + p^2
\end{aligned}
$$

$$
\begin{aligned}
&\left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^5 \beta_1 \sum_{i=-m_1}^{m_2} k_i i e^{i\xi} g^2 \\
&- \alpha p^2 g^4 \left(\sum_{j=-n_1}^{n_2} h_j e^{j\eta}\right)^4 \sum_{i=-m_1}^{m_2} k_i e^{i\xi} \alpha_2 \sum_{j=-n_1}^{n_2} h_j j e^{j\eta} \\
&= 0\,. \qquad (239)
\end{aligned}
$$

where m_1, m_2, n_1, n_2 are all any positive integers which we may take arbitrarily.

For example, we can take $n_2 = 2, m_2 = 2, n_1 = 1$, $m_1 = 1$ and substitute them into (239), and then set all coefficients of $exp(k\eta)$ and $exp(j\xi)$, $j, k = 0, 1, 2, \ldots$ to zero to get an over-determined equation with respect to $h_{-1}, h_0, h_1, h_2, k_0, k_1, k_2, k_{-1}, \alpha_1, \beta_1, \alpha_2, \beta_2$ and g. Solving the over-determined equations by use of Maple, we obtain the following solutions of h_{-1}, $h_0, h_1, h_2, k_0, k_1, k_2, k_{-1}, \alpha_1, \beta_1, \alpha_2, \beta_2$ and g:

Case 1.

$$
\begin{aligned}
&\beta_2 = -\frac{4\beta {\alpha_2}^3}{p^2}\,, \quad h_1 = \frac{p^2 \alpha {k_0}^2}{8{\alpha_2}^2 \beta h_{-1}(3p + 2 + p^2)}\,, \\
&k_{-1} = 0\,, \quad g = 0\,, \gamma_2 = \gamma_2\,, \quad \alpha_2 = \alpha_2\,, \\
&k_0 = k_0\,, \quad \alpha_1 = \alpha_1\,, \quad \beta_1 = \beta_1\,, \quad \gamma_1 = \gamma_1\,, \\
&h_{-1} = h_{-1}\,, \quad h_0 = 0\,, \quad k_2 = 0\,, \quad k_1 = 0\,, \\
&h_2 = 0\,, \qquad (240)
\end{aligned}
$$

where $\alpha, \beta, \alpha_2, \gamma_2, \alpha_2, k_0, \alpha_1, \beta_1, \gamma_1$ and k_0 are any constants.

Case 2.

$$
\begin{aligned}
&\gamma_2 = \gamma_2\,, \quad k_0 = k_0\,, \quad \alpha_1 = \alpha_1\,, \quad \beta_1 = \beta_1\,, \\
&\gamma_1 = \gamma_1\,, \quad k_2 = k_2\,, \quad k_1 = k_1\,, \quad h_{-1} = 0\,, \\
&\beta_2 = \beta_2\,, \quad \alpha_2 = \alpha_2\,, \quad k_{-1} = k_{-1}\,, \quad h_2 = 0\,, \\
&h_0 = 0\,, \quad h_1 = 0\,, \quad g = g\,, \qquad (241)
\end{aligned}
$$

where $\alpha, \beta, \gamma_2, k_0, \alpha_1, \beta_1, \gamma_1, k_2, k_1, \beta_2, \alpha_2, k_{-1}$ and g are any constants.

Case 3.

$$
\begin{aligned}
&\alpha_2 = -2\alpha_1\,, \quad \beta_2 = -2\beta_1\,, \quad \gamma_2 = \gamma_2\,, \quad \alpha_1 = \alpha_1\,, \\
&\beta_1 = \beta_1\,, \quad \gamma_1 = \gamma_1\,, \quad h_{-1} = h_{-1}\,, \quad h_0 = 0\,, \\
&k_1 = 0\,, \quad h_2 = 0\,, \quad k_{-1} = 0\,, \quad k_0 = 0\,, \\
&h_1 = 0\,, \quad k_2 = k_2, g = g\,, \qquad (242)
\end{aligned}
$$

where $\alpha, \beta, \alpha_1, \beta_1, \gamma_2, \beta_1, \gamma_1, h_{-1}, k_2$ and g are any constants.

Case 4.

$$
\begin{aligned}
&\gamma_2 = \gamma_2\,, \quad \alpha_2 = \alpha_2\,, \quad \alpha_1 = \alpha_1\,, \quad \beta_1 = \beta_1\,, \\
&\gamma_1 = \gamma_1\,, \quad h_{-1} = h_{-1}\,, \quad k_2 = 0\,, \quad k_1 = 0\,, \\
&k_{-1} = 0\,, \quad \beta_2 = \beta_2\,, \quad h_1 = h_1\,, \quad h_2 = h_2\,, \\
&h_0 = h_0\,, \quad k_0 = 0\,, \quad g = g\,, \qquad (243)
\end{aligned}
$$

where $\gamma_2, \alpha_2, \alpha_1, \beta_1, \gamma_1, h_{-1}, \beta_2, h_1, h_2$ and g are any constants.

Case 5.

$$
\begin{aligned}
&\gamma_2 = \gamma_2, k_0 = k_0, \alpha_1 = \alpha_1\,, \quad \beta_1 = \beta_1\,, \\
&\gamma_1 = \gamma_1\,, \quad h_{-1} = h_{-1}\,, \quad k_2 = 0\,, \quad k_1 = 0\,, \\
&k_{-1} = 0\,, \quad h_1 = h_1\,, \quad h_2 = h_2\,, \quad h_0 = h_0\,, \\
&\alpha_2 = 0, \beta_2 = 0\,, \quad g = g\,, \qquad (244)
\end{aligned}
$$

where $\alpha, \beta, \gamma_2, k_0, \alpha_1, \beta_1, \gamma_1, h_{-1}, h_1, h_2, h_0$ and g are any constants.

Case 6.

$$\begin{aligned}&\beta_2=-\beta_1\,,\quad \alpha_2=-\alpha_1\,,\quad \gamma_2=\gamma_2\,,\quad \alpha_1=\alpha_1\,,\\&\beta_1=\beta_1\,,\quad \gamma_1=\gamma_1\,,\quad h_{-1}=h_{-1}\,,\quad h_0=0\,,\\&k_2=0\,,\quad h_2=0\,,\quad k_{-1}=0\,,\quad k_0=0\,,\\&\qquad h_1=0\,,\quad g=g\,,\quad k_1=k_1\,,\end{aligned}\tag{245}$$

where $\alpha, \beta, \beta_1, \alpha_1, \gamma_2, \gamma_1, h_{-1}, g$ and k_1 are any constants.

Case 7.

$$\begin{aligned}&\gamma_2=\gamma_2\,,\quad \gamma_1=\gamma_1\,,\quad h_{-1}=h_{-1}\,,\quad k_2=0\,,\\&k_1=0\,,\quad \beta_2=\beta_2\,,\quad \alpha_2=\alpha_2\,,\quad k_{-1}=k_{-1}\,,\\&\beta_1=\beta_2\,,\quad h_2=0\,,\quad h_0=0\,,\quad h_1=0\,,\\&\qquad \alpha_1=\alpha_2\,,\quad k_0=0\,,\quad g=g\,,\end{aligned}\tag{246}$$

where $\alpha, \beta, \gamma_2, \gamma_1, h_{-1}, \beta_2, \alpha_2, k_{-1}, \alpha_1$ and g are any constants.

Case 8.

$$\begin{aligned}&\gamma_2=\gamma_2\,,\quad k_0=k_0\,,\quad \gamma_1=\gamma_1\,,\quad h_0=h_0\,,\\&\beta_1=0\,,\quad \alpha_1=0\,,\quad k_2=k_2\,,\quad k_1=k_1\,,\\&h_{-1}=0\,,\quad \beta_2=\beta_2\,,\quad \alpha_2=\alpha_2\,,\quad k_{-1}=k_{-1}\,,\\&\qquad h_2=0\,,\quad h_1=0\,,\quad g=g\,,\end{aligned}\tag{247}$$

where $\alpha, \beta, \gamma_2, k_0, \gamma_1, h_0, k_2, k_1, \beta_2, \alpha_2, k_{-1}$ and g are any constants.

Case 9.

$$\begin{aligned}&\gamma_2=\gamma_2\,,\quad \gamma_1=\gamma_1\,,\quad k_2=0\,,\quad k_1=0\,,\\&\alpha_1=-\alpha_2\,,\quad \beta_1=-\beta_2\,,\quad g=0\,,\quad h_1=h_1\,,\\&h_{-1}=0\,,\quad \beta_2=\beta_2\,,\quad \alpha_2=\alpha_2\,,\quad k_{-1}=k_{-1}\,,\\&\qquad h_2=0\,,\quad h_0=0\,,\quad k_0=0\,,\end{aligned}\tag{248}$$

where $\alpha, \beta, \gamma_2, \gamma_1, \alpha_2, \beta_2, h_1, \beta_2, \alpha_2$ and k_{-1} are any constants.

Case 10.

$$\begin{aligned}&\gamma_2=\gamma_2, \beta_1=\beta_1, \gamma_1=\gamma_1\,,\quad k_2=0\,,\\&k_1=0\,,\quad \alpha_1=-\alpha_2, h_1=h_1, \beta_2=-\beta_1\,,\\&h_{-1}=0\,, \alpha_2=\alpha_2\,,\quad k_{-1}=k_{-1}, h_2=0\,,\\&\qquad h_0=0, k_0=0, g=g\,,\end{aligned}\tag{249}$$

where $\alpha, \beta, \gamma_2, \beta_1, \gamma_1, \alpha_2, h_1, k_{-1}$ and g are any constants.

Case 11.

$$\begin{aligned}&\beta_1=\frac{1}{2}\beta_2\,,\quad \gamma_2=\gamma_2\,,\quad \gamma_1=\gamma_1\,,\quad k_1=0\,,\\&k_{-1}=0\,,\quad h_1=h_1\,,\quad \alpha_2=0\,,\quad \alpha_1=0\,,\\&k_2=k_2\,,\quad h_{-1}=0\,,\quad \beta_2=\beta_2\,,\quad h_2=0\,,\\&\qquad h_0=0\,,\quad k_0=0\,,\quad g=g\,,\end{aligned}\tag{250}$$

where $\beta_1=\frac{1}{2}\beta_2, \alpha, \beta, \gamma_2, \gamma_1, h_1, k_2, \beta_2$ and g are any constants.

Case 12.

$$\begin{aligned}&\beta_2=2\beta_1\,,\quad \alpha_1=\frac{1}{2}\alpha_2\,,\quad \gamma_2=\gamma_2\,,\quad \gamma_1=\gamma_1\,,\\&k_1=0\,,\quad k_{-1}=0\,,\quad h_1=h_1\,,\quad \alpha_2=\alpha_2\,,\\&\beta_1=\beta_1\,,\quad k_2=k_2\,,\quad h_{-1}=0\,,\quad h_2=0\,,\\&\qquad h_0=0\,,\quad k_0=0\,,\quad g=g\,,\end{aligned}\tag{251}$$

where $\alpha, \beta, \gamma_2, \gamma_1, h_1, \alpha_2, \beta_1, k_2$ and g are any constants.

Case 13.

$$\begin{aligned}&\beta_1=-2\beta_2\,,\quad h_1=0\,,\quad k_2=0\,,\quad \alpha_2=-1/2\alpha_1\,,\\&g=0\,,\quad h_2=h_2\,,\quad k_{-1}=k_{-1}\,,\quad \alpha_1=\alpha_1\,,\\&\gamma_2=\gamma_2\,,\quad \gamma_1=\gamma_1\,,\quad k_1=0\,,\quad h_{-1}=0\,,\\&\qquad \beta_2=\beta_2\,,\quad h_0=0\,,\quad k_0=0\,,\end{aligned}\tag{252}$$

where $\alpha, \beta, \beta_2, \alpha_1, h_2, \gamma_2, \gamma_1, \beta_2$ and k_{-1} are any constants.

Case 14.

$$\begin{aligned}&\alpha_1=2\alpha_2\,,\quad \beta_1=2\beta_2\,,\quad \alpha_2=\alpha_2\,,\quad h_1=0\,,\\&k_2=0, h_2=h_2\,,\quad \gamma_2=\gamma_2\,,\quad \gamma_1=\gamma_1\,,\\&k_{-1}=0\,,\quad k_1=k_1\,,\quad h_{-1}=0\,,\quad \beta_2=\beta_2\,,\\&\qquad h_0=0\,,\quad k_0=0\,,\quad g=g\,,\end{aligned}\tag{253}$$

where $\alpha, \beta, \alpha_2, \beta_2, h_2, \gamma_2, \gamma_1, k_1, \beta_2$ and g are any constants.

Case 15.

$$\begin{aligned}&\beta_1=-2\beta_2\,,\quad h_1=0\,,\quad k_2=0\,,\quad \alpha_2=-1/2\alpha_1\,,\\&g=0\,,\quad h_2=h_2\,,\quad k_{-1}=k_{-1}\,,\quad \alpha_1=\alpha_1\,,\\&\gamma_2=\gamma_2\,,\quad \gamma_1=\gamma_1\,,\quad k_1=0\,,\quad h_{-1}=0\,,\\&\qquad \beta_2=\beta_2\,,\quad h_0=0\,,\quad k_0=0\,,\end{aligned}\tag{254}$$

where $\alpha, \beta, \beta_2, \alpha_1, h_2, \alpha_1, \gamma_2, \gamma_1, \beta_2$ and k_{-1} are any constants.

Case 16.

$$\begin{aligned}&\beta_1 = 2\beta_2\,, \quad h_1 = 0\,, \quad k_2 = 0\,, \quad h_2 = h_2\,,\\ &\gamma_2 = \gamma_2\,, \quad \gamma_1 = \gamma_1\,, \quad k_{-1} = 0\,, \quad k_1 = k_1\,,\\ &\alpha_1 = 2\alpha_2\,, \quad \alpha_2 = \alpha_2\,, \quad h_{-1} = 0\,, \quad \beta_2 = \beta_2\,,\\ &\qquad h_0 = 0\,, \quad k_0 = 0\,, \quad g = g\,, \end{aligned} \tag{255}$$

where $\alpha, \beta, \beta_2, \gamma_2, \gamma_1, k_1, \alpha_2, \beta_2$ and g are any constants.

Case 17.

$$\begin{aligned}&\alpha_1 = -2\alpha_2\,, \quad h_1 = 0\,, \quad k_2 = 0\,, \quad h_2 = h_2\,,\\ &\gamma_2 = \gamma_2\,, \quad \gamma_1 = \gamma_1\,, \quad \alpha_2 = \alpha_2\,, \quad \beta_1 = -2\beta_2\,,\\ &k_{-1} = k_{-1}\,, \quad k_1 = 0\,, \quad h_{-1} = 0\,, \quad \beta_2 = \beta_2\,,\\ &\qquad h_0 = 0\,, \quad k_0 = 0\,, \quad g = g\,, \end{aligned} \tag{256}$$

where $\alpha, \beta, \alpha_2, h_2, \gamma_2, \gamma_1, \beta_2, k_{-1}, \beta_2$ and g are any constants.

Case 18.

$$\begin{aligned}&h_2 = h_2\,, \quad \gamma_2 = \gamma_2\,, \quad \gamma_1 = \gamma_1\,, \quad h_1 = 0\,,\\ &k_{-1} = 0\,, \quad \alpha_2 = \alpha_2\,, \quad \alpha_1 = \alpha_2\,, \quad k_2 = k_2\,,\\ &\beta_1 = \beta_2\,, \quad k_1 = 0\,, \quad h_{-1} = 0\,, \quad \beta_2 = \beta_2\,,\\ &\qquad h_0 = 0\,, \quad k_0 = 0\,, \quad g = g\,, \end{aligned} \tag{257}$$

where $\alpha, \beta, h_2, \gamma_2, \gamma_1, \alpha_2, k_2, \beta_2$ and g are any constants.

Case 19.

$$\begin{aligned}&h_1 = 0\,, \quad k_0 = k_0\,, \quad h_0 = h_0\,, \quad h_2 = 0\,,\\ &\beta_1 = 0\,, \quad \alpha_1 = 0\,, \quad \gamma_2 = \gamma_2\,, \quad \gamma_1 = \gamma_1\,,\\ &k_1 = k_1\,, \quad \alpha_2 = \alpha_2\,, \quad k_2 = k_2\,, \quad k_{-1} = k_{-1}\,,\\ &\qquad h_{-1} = 0\,, \quad \beta_2 = \beta_2\,, \quad g = g\,, \end{aligned} \tag{258}$$

where $\alpha, \beta, k_0, h_0, \gamma_2, \gamma_1, k_1, \alpha_2, k_2, k_{-1}, \beta_2$ and g are any constants.

Substituting (240)–(258) into (236) and (238), respectively, we can obtain the exact solutions of Eq. (235).

$$u(x,t) = \sqrt[2/p]{g + \frac{k_0 + 2\cosh_{k_1,k_{-1},1}(\alpha_1 x + \beta_1 t + \gamma_1) + k_2 e^{2(\alpha_1 x + \beta_1 t + \gamma_1)}}{h_0 + 2\cosh_{h_1,h_{-1},1}(\alpha_2 x + \beta_2 t + \gamma_2) + h_2 e^{2(\alpha_2 x + \beta_2 t + \gamma_2)}}}\,, \tag{259}$$

where $\cosh_{k_1,k_{-1},1}(\alpha_1 x + \beta_1 t + \gamma_1)$ and $\cosh_{h_1,h_{-1},1}(\alpha_2 x + \beta_2 t + \gamma_2)$ are all the generalized hyperbolic cosine functions, here $k_i, h_i, i = 0, \pm 1, 2; \alpha_j, \beta_j, \gamma_j, j = 1, 2$ and g satisfy (240)–(258), respectively.

For example, if we substitute (240) into (236) and (238), we obtain the following solutions of Eq. (235).

$$u_1(x,t) = \sqrt[2/p]{\frac{k_0 + 2\cosh_{0,0,1}(\alpha_1 x + \beta_1 t + \gamma_1)}{2\cosh_{h_1,h_{-1},1}(\alpha_2 x - 4\frac{\beta\alpha_2{}^3 t}{p^2} + \gamma_2)}}\,, \tag{260}$$

where $\cosh_{0,0,1}(\alpha_1 x + \beta_1 t + \gamma_1)$ and $\cosh_{h_1,h_{-1},1}(\alpha_2 x - 4(\beta\alpha_2{}^3 t)/p^2 + \gamma_2)$ are all the generalized hyperbolic cosine function, here $h_1 = (p^2\alpha k_0{}^2)/(8\alpha_2{}^2\beta h_{-1}(3p + 2 + p^2))$, and $\alpha, \beta, \alpha_2, \gamma_2, \alpha_2, k_0, \alpha_1, \beta_1, \gamma_1, k_0$ are any constants.

Remark 13 When $m_1, m_2, n_1, n_2 = 1, 2, \ldots$, we can obtain a family of N solitary-like solutions of Eq. (235).

In summary, we suggest two new generic transformations, a new Exp–N solitary-like method and its algorithm, to find new non-traveling waves solutions and traveling waves solutions of higher dimensional NLPDEs by our transformations. The suggested method is more powerful than the method proposed by Dr. He [31] to seek more exact solutions of higher dimensional NLPDEs. The $(1+1)$-dimensional generalized KdV equation is chosen to illustrate our algorithm such that new exact solutions are found.

The Exp-Bäcklund Transformation Method and Its Application in (1 + 1)-Dimensional KdV Equation

In this section, a new method and its algorithm, which is called Exp-Bäcklund transformation method, is presented to find more exact solutions of NLPDEs based on the idea of the Exp-function method [31]. We choose the $(1+1)$-dimensional KdV equations to illustrate the effectiveness and convenience of our algorithm. As a result, many new solutions are obtained.

Summary of the Exp-Bäcklund Transformation Method

In the following we would like to outline the main steps of our method:

Step 1. For a given NLEEs, with $u_i(t, x_1, x_2, \ldots, x_{n-1})$ $(i = 1, 2, \ldots, n)$ in n independent variables $t, x_1, x_2, \ldots, x_{n-1}$,

$$\begin{aligned}&F_j(u_1, \ldots, u_n, u_{1,t}, \ldots, u_{n,t}, u_{1,x_1}, \ldots, u_{n,x_{n-1}},\\ &\quad u_{1,tt}, \ldots, u_{n,tt}, u_{1,tx_1}, \ldots u_{n,tx_{n-1}}, \ldots) = 0\,, \end{aligned} \tag{261}$$

where $j = 1, 2, \ldots, n$.

We obtain an auto-Bäcklund transformation of Eqs. (261) using our method of constructing auto-Bäcklund transformation (if it exist)

$$u_i(t,x_1,\dots,x_{n-1}) = \sum_{j=0}^{m_i} u_{ij}(t,x_1,\dots,x_{n-1}) f^{j-m_i}(t,x_1,\dots,x_{n-1}), \quad i=1,2,\dots,n. \tag{262}$$

where $u_{ij}(t,x_1,\dots,x_{n-1}), i = 1,2,\dots,n; j = 0,1,2,\dots,m_i-1$ are differentiable functions which have already been obtained. But $f(t,x_1,\dots,x_{n-1})$ are any differentiable functions, $u_{i,m_i}, i=1,2,\dots,n$ are the seed solutions of Eq. (261), and $m_i, i=1,2,\dots,n$ are positive integers which have already been obtained.

Step 2. Substituting (262) into Eq. (261), and setting the coefficients of these terms $1/(f^k(t,x,y)), k=0,1,2,\dots$ to zero yields a set of over-determined partial differential equations with respect to $f(t,x,y), u_{i,m_i}, i=1,2,\dots,n$.

Step 3. We consider the variable in the form

$$\eta_i = p_{i0}(t) + \sum_{k=1}^{n-1} p_{ik}(t,x_1,\dots,x_{k-1})\cdot x_k, \quad i=1,2 \tag{263}$$

and make the transformation

$$f(t,x_1,\dots,x_{n-1}) = f(\eta_1,\eta_2). \tag{264}$$

Then we express the $f(t,x_1,\dots,x_{n-1})$ in (262) and (264) to be the following forms

$$f(\eta_1,\eta_2) = \frac{\sum_{n=-c}^{d} a_n exp(n\eta_1)}{\sum_{m=-p}^{q} b_m exp(m\eta_2)}, \tag{265}$$

where $\sum_{k=0}^{n-1} p_k \neq 0$, c,d,p and q are any positive integers, $a_n, n=-c,\dots,d$, $b_m, m=-p,\dots,q$, and $p_{ik}, k=0,1,2,\dots,n-1$are functions or constants to be determined later.

Step 4. If we substitute (262)–(265) into the over-determined partial differential equations in Step 2 and yield a set of over-determined partial differential equations with respect to $exp(k\eta_i), i=1,2; k=0,\pm1,\pm2,\dots$, $a_n, n=-c,\dots,d$, $b_m, m=-p,\dots,q$, $u_{i,m_i}, i = 1,2,\dots,n$, and $p_{ik}, k=0,1,2,\dots,n-1$. Then we collect the coefficients of the polynomials with respect to $exp(k\eta_i), i=1,2; k=0,\pm1,\pm2,\dots$ and set each coefficient to zero, we will get a system of over-determined partial differential of $a_n, n=-c,\dots,d$, $b_m, m=-p,\dots,q$, and $p_{ik}, k=0,1,2,\dots,n-1$, $u_{i,m_i}, i=1,2,\dots,n$.

Step 5. We solve the over-determined partial differential in Step 4 and obtain $a_n, n=-c,\dots,d$, $b_m, m=-p,\dots,q$, $u_{i,m_i}, i=1,2,\dots,n$, and $p_{ik}, k = 0,1,2,\dots,n-1$.

Step 6. Substituting the results obtained in Step 5 into (262)–(265), then we can obtain a family of rational solutions of Exp-functions as follows

$$u_i(t,x_1,\dots,x_{n-1}) = \sum_{j=0}^{m_i} u_{ij}(t,x_1,\dots,x_{n-1}) \left(\frac{\begin{array}{l}\sum_{n=-c}^{d} a_n exp(np_{10}(t) \\ +n\sum_{k=1}^{n-1} p_{1k}(t,x_1,\dots,x_{k-1})\cdot x_k)\end{array}}{\begin{array}{l}\sum_{m=-p}^{q} b_m exp(mp_{20}(t) \\ +m\sum_{k=1}^{n-1} p_{2k}(t,x_1,\dots,x_{k-1})\cdot x_k)\end{array}} \right)^{j-m_i}, \quad i=1,2,\dots,n, \tag{266}$$

where c, d, p and q are any positive integers.

The Application of the Exp-Bäcklund Transformation Method in (1 + 1)-Dimensional KdV Equation

In this section, we will make use of our Exp-Bäcklund transformation method [40,46] and symbolic computation to find the exact solutions of the following $(1+1)$-dimensional KdV equation [1,2,3,4]

$$u_t(x,t) + \alpha u(x,t)u_x(x,t) + \beta u_{xxx}(x,t) = 0, \tag{267}$$

where the coefficient of the nonlinear term α and the coefficient of the dispersive term β is independent of x and t.

Set the auto-Bäcklund transformation of Eq. (267) to have the following form:

$$u(x,t) = \sum_{j_1=0}^{m} u_j(x,t) f^{j-m}(x,t), \tag{268}$$

where $f(x,t)$ and $u_j(x,t), j=0,1,2,\dots,m-1$ are all differential functions to be determined later, and $u_m(x,t)$ is the trivial seed solution of Eq. (267).

By balancing the highest order linear term and nonlinear terms in Eq. (267), we get $m=2$ and (268) has the following formal

$$u(x,t) = u_0(x,t)f^{-2}(x,t) + u_1(x,t)f^{-1}(x,t) + u_2(x,t), \tag{269}$$

We take the trivial seed solution as

$$u_2 = u_2(x,t)\,. \tag{270}$$

With the aid of Maple symbolic computation software, substituting (269) and (270) into (267), and collecting all terms with $f^{-i}(x,t)$, $i = 0, 1, 2, \ldots$, we obtain

$$\begin{aligned}
&[-2\alpha u_0^2(x,t)f_x(x,t) - 24\beta u_0(x,t)f^3(x,t)]\frac{1}{f^5}\\
&+ \{-\beta[-18u_0(x,t)f_x(x,t)f_{xx}(x,t) - 18u_{0x}(x,t)f_x^2(x,t)\\
&\qquad + 6u_1(x,t)f_x^3(x,t)]\\
&\quad + \alpha u_0(x,t)[u_{0x}(x,t) - u_1(x,t)f_x(x,t)]\\
&\quad - 2\alpha u_1(x,t)u_0(x,t)f_x(x,t)\}\frac{1}{f^4}\\
&+ \{-\beta[-6u_1(x,t)f_x(x,t)f_{xx}(x,t)\\
&\qquad + 6u_{0xx}(x,t)f_x(x,t) + 2u_0(x,t)f_{xxx}(x,t)\\
&\qquad + 6u_{0x}(x,t)f_{xx}(x,t) - 6u_{1x}(x,t)f_x^2(x,t)]\\
&\quad - 2u_0(x,t)f_t(x,t) + \alpha u_0(x,t)u_{1x}(x,t)\\
&\quad + \alpha u_1(x,t)[u_{0x}(x,t) - u_1(x,t)f_x(x,t)]\\
&\quad - 2\alpha u_2(x,t)u_0(x,t)f_x(x,t)\}\frac{1}{f^3}\\
&+ \{-\beta[-u_{0xxx}(x,t) + 3u_{1x}(x,t)f_{xx}(x,t)\\
&\qquad + 3u_{1xx}(x,t)f_x(x,t) + u_1(x,t)f_{xxx}(x,t)]\\
&\quad + \alpha u_0(x,t)u_{2x}(x,t) + \alpha u_1(x,t)u_{1x}(x,t)\\
&\quad + \alpha u_2(x,t)[u_{0x}(x,t) - u_1(x,t)f_x(x,t)]\\
&\quad + u_{0t}(x,t) - u_1(x,t)f_t(x,t)\}\frac{1}{f^2}\\
&+ (u_{1t}(x,t) + \alpha u_1(x,t)u_{2x}(x,t) + \alpha u_2(x,t)u_{1x}(x,t)\\
&+ \beta u_{1xxx}(x,t))\frac{1}{f} + u_{2t}(x,t) + \alpha u_2(x,t)u_{2x}(x,t)\\
&+ \beta u_{2xxx}(x,t) = 0\,,
\end{aligned} \tag{271}$$

Setting the coefficient of f^{-5} in (271) to be zero, we obtain a differential equation

$$-2\alpha u_0^2(x,t)f_x(x,t) - 24\beta u_0(x,t)f^3(x,t) = 0\,, \tag{272}$$

which has solution

$$u_0(x,t) = -\frac{12\beta}{\alpha}f_x^2(x,t)\,. \tag{273}$$

Setting the coefficient of f^{-4}in (271) to be zero, we obtain a differential equation

$$\begin{aligned}
&-\beta[-18u_0(x,t)f_x(x,t)f_{xx}(x,t)\\
&\qquad - 18u_{0x}(x,t)f_x^2(x,t) + 6u_1(x,t)f_x^3(x,t)]\\
&+ \alpha u_0(x,t)[u_{0x}(x,t) - u_1(x,t)f_x(x,t)]\\
&- 2\alpha u_1(x,t)u_0(x,t)f_x(x,t) = 0\,,
\end{aligned} \tag{274}$$

we can get the following expressions:

$$u_1(x,t) = \frac{12\beta}{\alpha}f_{xx}(x,t)\,. \tag{275}$$

By (273), (275), (270) and (269), we obtain an auto-Bäcklund transformation of Eq. (267) as follows

$$u(x,t) = -\frac{12\beta}{\alpha}\left(\frac{f_{xx}(x,t)}{f(x,t)} - \frac{f_x^2(x,t)}{f^2(x,t)}\right) + u_2(x,t) \tag{276}$$

where $u_2(x,t)$ is a seed solution of Eq. (267), $f(x,t)$ is any function that satisfies Eq. (272), Eq. (274), and the following equations

$$\begin{aligned}
&-\beta[-6u_1(x,t)f_x(x,t)f_{xx}(x,t) + 6u_{0xx}(x,t)f_x(x,t)\\
&\qquad + 2u_0(x,t)f_{xxx}(x,t) + 6u_{0x}(x,t)f_{xx}(x,t)\\
&\qquad - 6u_{1x}(x,t)f_x^2(x,t)]\\
&- 2u_0(x,t)f_t(x,t) + \alpha u_0(x,t)u_{1x}(x,t)\\
&+ \alpha u_1(x,t)[u_{0x}(x,t) - u_1(x,t)f_x(x,t)]\\
&- 2\alpha u_2(x,t)u_0(x,t)f_x(x,t) = 0\,,\\[1ex]
&-\beta[-u_{0xxx}(x,t) + 3u_{1x}(x,t)f_{xx}(x,t)\\
&\qquad + 3u_{1xx}(x,t)f_x(x,t) + u_1(x,t)f_{xxx}(x,t)]\\
&+ \alpha u_0(x,t)u_{2x}(x,t) + \alpha u_1(x,t)u_{1x}(x,t)\\
&+ \alpha u_2(x,t)[u_{0x}(x,t) - u_1(x,t)f_x(x,t)] + u_{0t}(x,t)\\
&- u_1(x,t)f_t(x,t) = 0\,,\\[1ex]
&u_{1t}(x,t) + \alpha u_1(x,t)u_{2x}(x,t)\\
&+ \alpha u_2(x,t)u_{1x}(x,t) + \beta u_{1xxx}(x,t) = 0\,.
\end{aligned} \tag{277}$$

We take $f(x,t)$ to be of the following form

$$f(x,t) = \frac{\sum_{i=-m_1}^{m_2} k_i(t)e^{i(p_1(t)x+q_1(t))}}{\sum_{j=-n_1}^{n_2} h_j(t)e^{j(p_2(t)x+q_2(t))}}\,, \tag{278}$$

where n_1, n_2, m_1 and m_2 are any positive integers, $p_l(t), q_l(t)$, $l = 1, 2$; $k_i(t)$, $i = -m_1, \ldots, m_2$ and $h_j(t)$, $j = -n_1, \ldots, n_2$ are the coefficients to be determined later.

By substituting (278) into the given Eq. (277) with (273) and (275), and then collecting the coefficients of the polynomials of $e^{i(p_1(t)x+q_1(t))}$ and $e^{j(p_2(t)x+q_2(t))}$, $i, j = 0, 1, 2, \ldots$, then setting each coefficient to zero, we will get a system of over-determined partial differential equations with respect to $k_i(t)$, $i = -m_1, \ldots, m_2$; $h_j(t)$, $j = -n_1, \ldots, n_2$; $p_k(t), q_k(t)$, $k = 1, 2$ and $u_2(x, t)$. We solve the over-determined partial differential equations and obtain many families of solutions of $k_i(t), i = -m_1, \ldots, m_2; h_j(t)$, $j = -n_1, \ldots, n_2; p_k(t), q_k(t), k = 1, 2$ and $u_2(x, t)$. Finally, substituting them into (276) with (278), we can obtain the following solutions of Eq. (267)

$$
\begin{aligned}
u = & -\frac{12\beta\left(-\sum_{i=-m_1}^{m_2} k_i(t) i p_1(t) e^{i\xi_1} \sum_{j=-n_1}^{n_2} h_j(t) e^{j\xi_2} + \sum_{i=-m_1}^{m_2} k_i(t) e^{i\xi_1} \sum_{j=-n_1}^{n_2} h_j(t) j p_2(t) e^{j\xi_2}\right)^2}{\alpha\left(\sum_{j=-n_1}^{n_2} h_j(t) e^{j\xi_2}\right)^2 \left(\sum_{i=-m_1}^{m_2} k_i(t) e^{i\xi_1}\right)^2} \\
& - \frac{12\beta {p_1}^2(t)\left(\sum_{i=-m_1}^{m_2} k_i(t) i e^{i\xi_1}\right)^2}{\alpha\left(\sum_{i=-m_1}^{m_2} k_i(t) e^{i\xi_1}\right)^2} + \frac{12\beta {p_1}^2(t) \sum_{i=-m_1}^{m_2} k_i(t) i^2 e^{i\xi_1}}{\alpha \sum_{i=-m_1}^{m_2} k_i(t) e^{i\xi_1}} \\
& + \frac{12\beta {p_2}^2(t)\left(\sum_{j=-n_1}^{n_2} h_j(t) j e^{j\xi_2}\right)^2}{\left(\sum_{j=-n_1}^{n_2} h_j(t) e^{j\xi_2}\right)^2 \alpha} - \frac{12\beta {p_2}^2(t) \sum_{j=-n_1}^{n_2} h_j(t) j^2 e^{j\xi_2}}{\sum_{j=-n_1}^{n_2} h_j(t) e^{j\xi_2} \alpha} + u_2(t) ,
\end{aligned}
\tag{279}
$$

where n_1, n_2, m_1 and m_2 are any positive integers, and $\xi_1 = p_1(t)x + q_1(t)$ and $\xi_2 = p_2(t)x + q_2(t)$.

For example, when we take $n_1 = 1, n_2 = 1, m_1 = 1$ and $m_2 = 1$, substituting (278) into the given Eq. (277) with (273) and (275), and then collecting the coefficients of the polynomials of $e^{i(p_1(t)x+q_1(t))}$ and $e^{j(p_2(t)x+q_2(t))}$, i, $j = 0, 1, 2, \ldots$, then setting each coefficient to zero, we will get a system of over-determined partial differential equations with respect to $p_l(t), q_l(t), l = 1, 2; k_i(t), i = -1$, $0, 1$ and $h_j(t), j = -1, 0, 1$ as follows

$$
\begin{aligned}
& - p_2(t)k_0(t)h_1^4(t)k_{-1t}(t) \\
& - p_2(t)k_{-1}(t)h_1^4(t)k_{0t}(t) \\
& + 4\beta p_2(t)k_0(t)h_1^4(t)p_1^3(t)k_{-1}(t) \\
& + 2p_2(t)k_{-1}(t)h_1^3(t)k_0(t)h_{1t}(t) \\
& + p_2(t)k_0(t)h_1^4(t)k_{-1}(t)q_{1t}(t) \\
& + p_1(t)k_{-1}(t)h_1^4(t)k_0(t)q_{2t}(t) \\
& + 2u_2(x,t)\alpha p_1(t)k_{-1}(t)h_1^4(t)p_2(t)k_0(t) \\
& + 2u_2(x,t)\alpha p_2^2(t)k_{-1}(t)h_1^4(t)k_0(t) \\
& + 2p_2(t)k_{-1}(t)h_1^4(t)k_0(t)q_{2t}(t) \\
& + p_1(t)k_{-1}(t)h_1^3(t)k_0(t)h_{1t}(t) \\
& - p_1(t)k_{-1}(t)h_1^4(t)k_{0t}(t) \\
& + 4\beta p_1(t)k_{-1}(t)h_1^4(t){p_2}^3(t)k_0(t) \\
& + 2\beta p_2^4(t)k_{-1}(t)h_1^4(t)k_0(t) \\
& + 6\beta p_2^2(t)k_0(t)h_1^4(t)p_1^2(t)k_{-1}(t) = 0 ,
\end{aligned}
$$

$$
\begin{aligned}
& p_1(t)k_1(t)h_1^4(t)k_{0t}(t) - 4\beta p_2(t)k_0(t)h_1^4(t)p_1^3(t)k_1(t) \\
& - 4\beta p_1(t)k_1(t)h_1^4(t)p_2^3(t)k_0(t) \\
& + 2p_2(t)k_0(t)h_1^4(t)k_1(t)q_{2t}(t) + 2\beta p_2^4(t)k_1(t)h_1^4(t)k_0(t) \\
& + 2u_2(x,t)\alpha p_2^2(t)k_0(t)h_1^4(t)k_1(t) \\
& - 2u_2(x,t)\alpha p_1(t)k_1(t)h_1^4(t)p_2(t)k_0(t) \\
& - p_2(t)k_1(t)h_1^4(t)k_{0t}(t) + 6\beta p^{2_2}(t)k_0(t)h_1^4(t)p_1^2k_1(t) \\
& + 2p_2(t)k_0(t)h_1^3(t)k_1(t)h_{1t}(t) \\
& - p_1(t)k_1(t)h_1^3(t)k_0(t)h_{1t}(t) - p_2(t)k_0(t)h_1^4(t)k_1(t)q_{1t}(t) \\
& - p_2(t)k_0(t)h_1^4(t)k_{1t}(t) \\
& - p_1(t)k_1(t)h_1^4(t)k_0(t)q_{2t}(t)4\beta p_1(t)k_1(t)h_{-1}^4(t)p_2^3(t)k_0(t) \\
& + p_2(t)k_0(t)h_{-1}^4(t)k_{1t}(t) + p_2(t)k_0(t)h_{-1}^4(t)k_1(t)q_{1t}(t) \\
& + 2p_2(t)k_0(t)h_{-1}^4(t)k_1(t)q_{2t}(t) \\
& + p_1(t)k_1(t)h_{-1}^4(t)k_0(t)q_{2t}(t) \\
& + 6\beta p_2^2(t)k_0(t)h_{-1}^4(t)p_1^2(t)k_1(t) \\
& + 2u_2(x,t)\alpha p_1(t)k_1(t)h_{-1}^4(t)p_2(t)k_0(t) \\
& - p_1(t)k_1(t)h_{-1}^3(t)k_0(t)h_{-1t}(t) \\
& + 2u_2(x,t)\alpha p_2^2(t)k_0(t)h_{-1}^4(t)k_1(t) \\
& - 2p_2(t)k_0(t)h_{-1}^3(t)k_1(t)h_{-1t}(t) \\
& + 2\beta p_2^4(t)k_1(t)h_{-1}^4(t)k_0(t) \\
& + 4\beta p_2(t)k_0(t)h_{-1}^4(t)p_1^3k_1(t) + p_1(t)k_1(t)h_{-1}^4(t)k_{0t}(t) \\
& + p_2(t)k_1(t)h_{-1}^4(t)k_{0t}(t) = 0 ,
\end{aligned}
\tag{280}
$$

$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

We solve the over-determined partial differential equations and obtain the following many families of solutions:

Case 1

$$u_1(x,t) = -\frac{12\beta(k_{-1}(t))^2(p_1(t))^2(\mathrm{e}^{-p_1(t)x-q_1(t)})^2}{\alpha(k_{-1}(t)\mathrm{e}^{-p_1(t)x-q_1(t)}+C_1)^2} + \frac{12\beta k_{-1}(t)(p_1(t))^2\mathrm{e}^{-p_1(t)x-q_1(t)}}{\alpha(k_{-1}(t)\mathrm{e}^{-p_1(t)x-q_1(t)}+C_1)} + u_2(t)\,, \quad (281)$$

where $q_1(t) = \int \frac{\frac{\mathrm{d}}{\mathrm{d}t}k_{-1}(t)-(p_1(t))^3k_{-1}(t)\beta-p_1(t)u_2(t)\alpha k_{-1}(t)}{k_{-1}(t)}\mathrm{d}t + C_3$, $k_{-1}(t), p_1(t)$ are any functions of t, and α, β, C_i, $i = 1, 3$ are any constants.

Case 2

$$u_2(x,t) = -\frac{12\beta(C_1p_2(t)+k_{-1}(t)(p_2(t)-p_1(t))\mathrm{e}^{-\xi_1} + C_2(p_2(t)+p_1(t))\mathrm{e}^{\xi_1})^2}{\alpha(k_{-1}(t)\mathrm{e}^{-\xi_1}+C_1+C_2\mathrm{e}^{\xi_1})^2} + \frac{12\beta(C_1(p_2(t))^2+C_2(p_2(t)+p_1(t))^2\mathrm{e}^{\xi_1} + (p_1(t)-p_2(t))^2k_{-1}(t)\mathrm{e}^{-\xi_1})}{\alpha(k_{-1}(t)\mathrm{e}^{-\xi_1}+C_1+C_2\mathrm{e}^{\xi_1})}, \quad (282)$$

where $\xi_1 = p_1(t)x + q_1(t)$, $k_{-1}(t), q_1(t), p_1(t)$ are any functions of t, and α, β, C_i, $i = 1, 2$ are any constants.

Case 3

$$u_3(x,t) = -\frac{12\beta(k_{-1}(t)(p_2(t)-p_1(t))\mathrm{e}^{-\xi_1} + k_1(t)(p_2(t)+p_1(t))\mathrm{e}^{\xi_1}+C_1p_2(t))^2}{\alpha(k_{-1}(t)\mathrm{e}^{-\xi_1}+C_1+k_1(t)\mathrm{e}^{\xi_1})^2} + \frac{12\beta(C_1(p_2(t))^2+k_1(t)(p_2(t)+p_1(t))^2\mathrm{e}^{\xi_1} + (p_1(t)-p_2(t))^2k_{-1}(t)\mathrm{e}^{-\xi_1})}{\alpha(k_{-1}(t)\mathrm{e}^{-\xi_1}+C_1+k_1(t)\mathrm{e}^{\xi_1})}, \quad (283)$$

where $\xi_1 = p_1(t)x + q_1(t)$, $k_{-1}(t), k_1(t), q_1(t), p_1(t)$ are any functions of t, and α, β, C_1 are any constants.

Case 4

$$u_4(x,t) = -\frac{12\beta(p_1(t))^2(k_1(t))^2(\mathrm{e}^{p_1(t)x+q_1(t)})^2}{\alpha(C_1+k_1(t)\mathrm{e}^{p_1(t)x+q_1(t)})^2} + \frac{12\beta(p_1(t))^2k_1(t)\mathrm{e}^{p_1(t)x+q_1(t)}}{\alpha(C_1+k_1(t)\mathrm{e}^{p_1(t)x+q_1(t)})} + u_2(t)\,, \quad (284)$$

where $q_1(t) = -\int \frac{\beta(p_1(t))^3k_1(t)+u_2(t)\alpha p_1(t)k_1(t)+\frac{\mathrm{d}}{\mathrm{d}t}k_1(t)}{k_1(t)}\mathrm{d}t + C_3$, $k_1(t), p_1(t)$ are the functions of t, and α, β, C_i, $i = 1, 3$ are any constants.

Case 5

$$u_5(x,t) = \frac{2C_1h_0(t)(u_2(t)\alpha-6\beta(p_2(t))^2)\mathrm{e}^{-\xi_2} + u_2(t)\alpha C_1^2\mathrm{e}^{-2\xi_2}+u_2(t)\alpha(h_0(t))^2}{(C_1\mathrm{e}^{-\xi_2}+h_0(t))^2\alpha}, \quad (285)$$

where $\xi_2 = p_2(t)x + q_2(t)$, $h_0(t), p_2(t), q_2(t)$ are the functions of t, and α, β, C_1 are any constants.

Case 6

$$u_6(x,t) = -\frac{2C_1h_0(t)(6\beta(p_2(t))^2-u_2(t)\alpha)\mathrm{e}^{-\xi_2} + 2C_1h_1(t)(24\beta(p_2(t))^2-u_2(t)\alpha)}{(C_1\mathrm{e}^{-\xi_2}+h_0(t)+h_1(t)\mathrm{e}^{\xi_2})^2\alpha} - \frac{2h_0(t)h_1(t)(6\beta(p_2(t))^2-u_2(t)\alpha)\mathrm{e}^{\xi_2} - u_2(t)\alpha((h_1(t))^2+C_1^2)\mathrm{e}^{-2\xi_2}-u_2(t)\alpha(h_0(t))^2}{(C_1\mathrm{e}^{-\xi_2}+h_0(t)+h_1(t)\mathrm{e}^{\xi_2})^2\alpha}, \quad (286)$$

where $\xi_2 = p_2(t)x + q_2(t)$, $h_0(t), h_1(t), q_1(t), q_2(t)$ are the functions of t, and α, β, C_1 are any constants.

Case 7

$$u_7(x,t) = \frac{2h_0(t)h_1(t)(u_2(t)\alpha-6\beta(p_2(t))^2)\mathrm{e}^{\xi_2} + 2h_0(t)h_{-1}(t)(u_2(t)\alpha-6\beta(p_2(t))^2)\mathrm{e}^{-\xi_2}}{(h_{-1}(t)\mathrm{e}^{-\xi_2}+h_0(t)+h_1(t)\mathrm{e}^{\xi_2})^2\alpha} + \frac{u_2(t)\alpha((h_{-1}(t))^2\mathrm{e}^{-2\xi_2}+(h_1(t))^2\mathrm{e}^{2\xi_2}) - 48\beta h_{-1}(t)(p_2(t))^2h_1(t) + u_2(t)\alpha(2h_{-1}(t)h_1(t)+(h_0(t))^2)}{(h_{-1}(t)\mathrm{e}^{-\xi_2}+h_0(t)+h_1(t)\mathrm{e}^{\xi_2})^2\alpha}, \quad (287)$$

where $\xi_2 = p_2(t)x + q_2(t)$, $h_{-1}(t), h_0(t), h_1(t), q_1(t)$, $q_2(t)$ are the functions of t, and α, β are any constants.

Case 8

$$u_8(x,t) = \frac{2C_1k_0(t)(6\beta(p_1(t))^2+u_2(t)\alpha)\mathrm{e}^{-\xi_1} + u_2(t)\alpha C_1^2\mathrm{e}^{-2\xi_1}+u_2(t)\alpha(k_0(t))^2}{\alpha(C_1\mathrm{e}^{-\xi_1}+k_0(t))^2}, \quad (288)$$

where $\xi_1 = p_1(t)x + q_1(t)$, $k_0(t), q_1(t), q_2(t)$ are the functions of t, and α, β, C_1 are any constants.

Case 9

$$u_9(x,t) = \frac{u_2(t)\alpha((k_0(t))^2+2C_1k_1(t)) + 48\beta C_1(p_1(t))^2k_1(t)+u_2(t)\alpha(k_1(t))^2\mathrm{e}^{2\xi_1}}{\alpha(C_1\mathrm{e}^{-\xi_1}+k_0(t)+k_1(t)\mathrm{e}^{\xi_1})^2} + \frac{2k_0(t)k_1(t)(6\beta(p_1(t))^2+u_2(t)\alpha)\mathrm{e}^{\xi_1} + 2C_1k_0(t)(6\beta(p_1(t))^2+u_2(t)\alpha)\mathrm{e}^{-\xi_1} + u_2(t)\alpha C_1^2\mathrm{e}^{-2\xi_1}}{\alpha(C_1\mathrm{e}^{-\xi_1}+k_0(t)+k_1(t)\mathrm{e}^{\xi_1})^2}, \quad (289)$$

where $\xi_1 = p_1(t)x + q_1(t)$, $k_0(t), k_1(t), q_1(t), q_2(t)$ are the functions of t, and α, β, C_1 are any constants.

Case 10

$$u_{10}(x,t) = -\frac{12\beta(p_1(t))^2(k_1(t))^2(\mathrm{e}^{p_1(t)x+q_1(t)})^2}{\alpha(k_0(t)+k_1(t)\mathrm{e}^{p_1(t)x+q_1(t)})^2} + \frac{12\beta(p_1(t))^2k_1(t)\mathrm{e}^{p_1(t)x+q_1(t)}}{\alpha(k_0(t)+k_1(t)\mathrm{e}^{p_1(t)x+q_1(t)})} + u_2(t)\,, \quad (290)$$

where $q_1(t) = \int \frac{k_0(t)\frac{\mathrm{d}}{\mathrm{d}t}k_1(t) - k_1(t)\frac{\mathrm{d}}{\mathrm{d}t}k_0(t) + k_0(t)k_1(t)\cdot(p_1(t))^3\beta + u_2(t)\alpha p_1(t)k_1(t)k_0(t)}{k_0(t)k_1(t)}\mathrm{d}t + C_2$, $k_0(t), k_1(t), q_2(t)$ are the functions of t, and α, β, C_2 are any constants.

Case 11

$$\begin{aligned} u_{11}(x,t) = \\ &- \frac{\begin{array}{l}12\beta(k_{-1}(t)(p_2(t)-p_1(t))\mathrm{e}^{-\xi_1} \\ +k_1(t)(p_2(t)+p_1(t))\mathrm{e}^{\xi_1} + k_0(t)p_2(t))^2\end{array}}{\alpha(k_{-1}(t)\mathrm{e}^{-\xi_1}+k_0(t)+k_1(t)\mathrm{e}^{\xi_1})^2} \\ &+ \frac{\begin{array}{l}12\beta(k_0(t)(p_2(t))^2 + k_1(t)(p_2(t)+p_1(t))^2\mathrm{e}^{\xi_1} \\ +(p_1(t)-p_2(t))^2k_{-1}(t)\mathrm{e}^{-\xi_1})\end{array}}{\alpha(k_{-1}(t)\mathrm{e}^{-\xi_1}+k_0(t)+k_1(t)\mathrm{e}^{\xi_1})} \\ &+ u_2(t)\,, \quad (291)\end{aligned}$$

where $\xi_1 = p_1(t)x + q_1(t), k_{-1}(t), k_0(t), k_1(t), q_1(t), q_2(t)$ are the functions of t, and α, β are any constants.

Case 12

$$u_{12}(x,t) = -\frac{12\beta(k_{-1}(t))^2(p_1(t))^2(\mathrm{e}^{-p_1(t)x-q_1(t)})^2}{\alpha(k_{-1}(t)\mathrm{e}^{-p_1(t)x-q_1(t)}+k_0(t))^2} + \frac{12k_{-1}(t)(p_1(t))^2\mathrm{e}^{-p_1(t)x-q_1(t)}\beta}{\alpha(k_{-1}(t)\mathrm{e}^{-p_1(t)x-q_1(t)}+k_0(t))} + u_2(t)\,, \quad (292)$$

where $q_1(t) = \int \frac{k_0(t)\frac{\mathrm{d}}{\mathrm{d}t}k_{-1}(t) - u_2(t)\alpha p_1(t)k_{-1}(t)k_0(t) - k_0(t)(p_1(t))^3k_{-1}(t)\beta - k_{-1}(t)\frac{\mathrm{d}}{\mathrm{d}t}k_0(t)}{k_0(t)k_{-1}(t)}\mathrm{d}t + C_2$, $k_{-1}(t), k_0(t), q_2(t)$ are the functions of t, and α, β, C_2 are any constants.

Remark 14 From the above application, it is easily seen that our method is more powerful, more convenient, and more general than the method presented by He [31]. Firstly, our method is convenient to obtain the exact solutions of higher order and higher dimensional NLEEs because we do not use the homogeneous balance method to get the value of c, d, p, q in (265). Secondly, we can obtain more general exact solutions including non-traveling wave solutions and traveling wave solutions by using our method. For example, the solution (282) of Eq. (267) is a more general exact solution including the non-traveling wave solutions with constant coefficients, the non-traveling wave solutions with variable coefficients, the traveling wave solutions with constant coefficients, and the traveling wave solutions with variable coefficients of Eq. (267).

Example 1 The solution (282) of Eq. (267) includes abundant non-traveling wave solutions with variable coefficients. For example, when we take $C_1 = 1/2, C_2 = 1$, $p_2(t) = \cos(t^2), u_2(t) = 1/2, p_1(t) = \tanh^2(t), q_1(t) = \sin(t^4), k_{-1}(t) = \cosh(t), k_0(t) = \cos(t^2), \beta = 2, \alpha = 3$, we can obtain a non-traveling wave solution $u_{2,1}(x,t)$ with variable coefficients as follows:

$$\begin{aligned} u_{2,1}(x,t) = \\ &\frac{\begin{array}{l}8\cos^2(t^2)(2\cosh(t)\mathrm{e}^{-\xi_1}+1+2\mathrm{e}^{\xi_1}) \\ +32\tanh^2(t)\cos(t^2)(\mathrm{e}^{\xi_1}-\cosh(t)\mathrm{e}^{-\xi_1}) \\ +16\tanh^4(t)(\cosh(t)\mathrm{e}^{-\xi_1}+\mathrm{e}^{\xi_1})\end{array}}{2\cosh(t)\mathrm{e}^{-\xi_1}+1+2\mathrm{e}^{\xi_1}} \\ &-8\frac{\begin{array}{l}(2\tanh^2(t)(\mathrm{e}^{\xi_1}-\cosh(t)\mathrm{e}^{-\xi_1}) \\ +\cos(t^2)(2\cosh(t)\mathrm{e}^{-\xi_1}+1+2\mathrm{e}^{\xi_1}))^2\end{array}}{(2\cosh(t)\mathrm{e}^{-\xi_1}+1+2\mathrm{e}^{\xi_1})^2} + 1/2\,, \quad (293)\end{aligned}$$

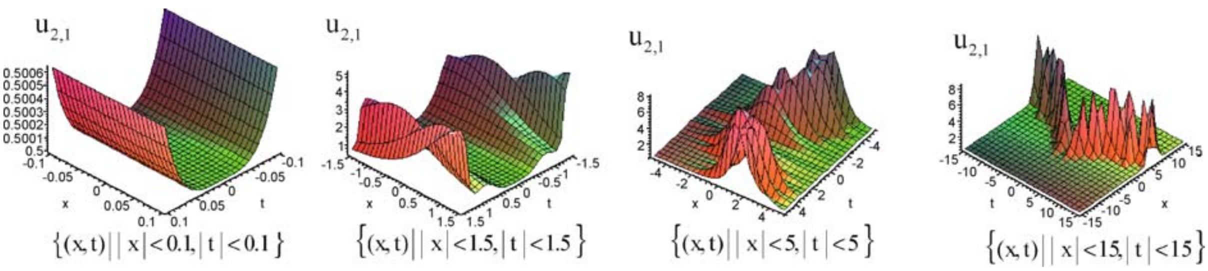
where $\xi_1 = \tanh^2(t)x + \sin(t^4)$.

Some figures of the long-term behavior of the $u_{2,1}(x,t)$ from $\{(x,t)||x| < 0.1, |t| < 0.1, x \in R, t \in R\}$ to $\{(x,t)||x| < 10^9, |t| < 10^9, x \in R, t \in R\}$ are shown by following Fig. 1 and Fig. 2. From Fig. 1 and Fig. 2, it is easy to find the global existence, asymptotic behaviors as $t \to +\infty$ and $|x| \to +\infty$ and the scattering properties of the solution $u_{2,1}(x,t)$ of Eq. (267). It is worth noting that the shape of $u_{2,1}(x,t)$ degenerates gradually from a set of waves to a set of solitons on a trigonal curve as the $|x|$ and $|t|$ in the $u_{2,1}(x,t)$ continue to increase.

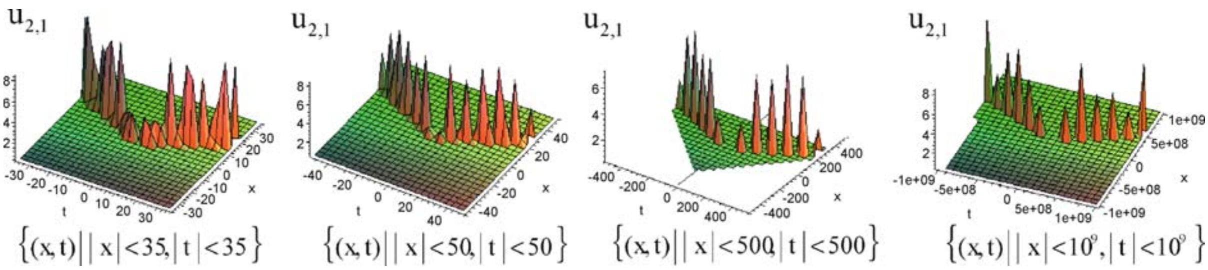
Example 2 The solution (282) of Eq. (267) includes abundant non-traveling wave solutions with constant coefficients. For example, when we take $k_{-1}(t) = 1/2, \beta = 2, \alpha = 3, k_0(t) = 2, C_2 = 1, C_1 = 1/2, p_1(t) = 1/2, p_2(t) = 1/3, u_2(t) = 1/2$, and $q_1(t) = \sin(t^4)$, we can obtain a non-traveling wave solution $u_{2,2}(x,t)$ with constant coefficients as follows:

$$u_{2,2}(x,t) = -\frac{2(\mathrm{e}^{-x/2-\sin(t^4)} - 10\mathrm{e}^{x/2+\sin(t^4)} - 2)^2}{9(\mathrm{e}^{-x/2-\sin(t^4)} + 1 + 2\mathrm{e}^{x/2+\sin(t^4)})^2} + \frac{8 + 100\mathrm{e}^{x/2+\sin(t^4)} + 2\mathrm{e}^{-x/2-\sin(t^4)}}{9\mathrm{e}^{-x/2-\sin(t^4)} + 9 + 18\mathrm{e}^{x/2+\sin(t^4)}} + 1/2\,. \quad (294)$$

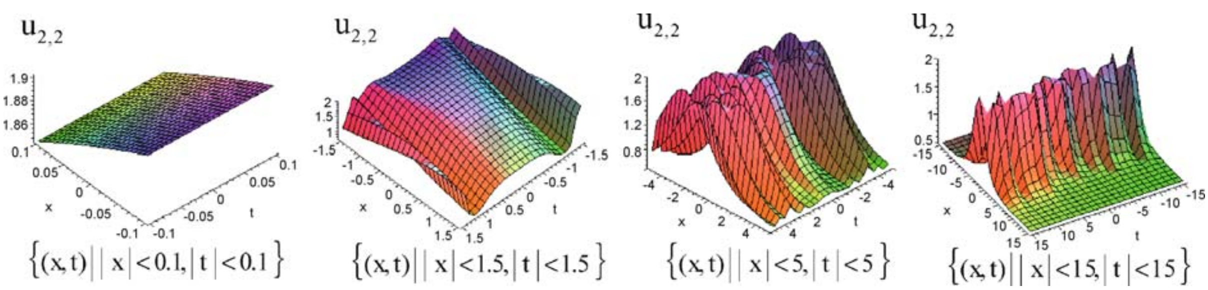
Some figures of the long-term behavior of the $u_{2,2}(x,t)$ from $\{(x,t)||x| < 0.1, |t| < 0.1, x \in R, t \in R\}$



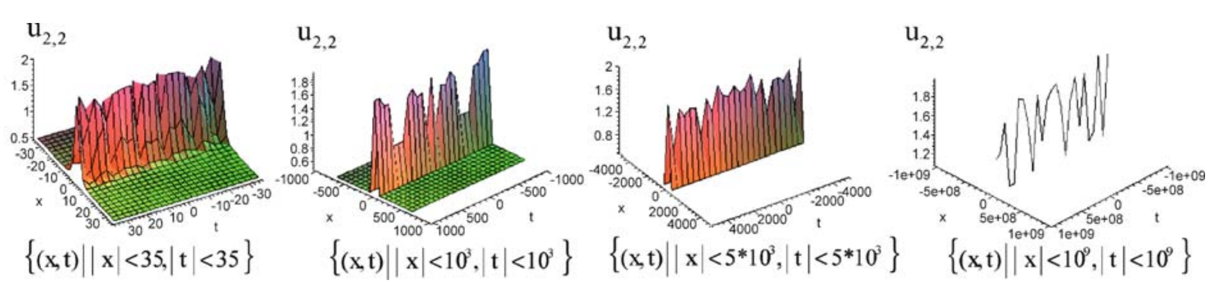
Korteweg–de Vries Equation (KdV), Different Analytical Methods for Solving the, Figure 1
The evolution of a non-traveling wave solution $u_{2,1}(x,t)$ with variable coefficients of Eq. (267) from $\{(x,t)||x| < 0.1, |t| < 0.1, x \in R, t \in R\}$ to $\{(x,t)||x| < 15, |t| < 15, x \in R, t \in R\}$



Korteweg–de Vries Equation (KdV), Different Analytical Methods for Solving the, Figure 2
The evolution of a non-traveling wave solution $u_{2,1}(x,t)$ with variable coefficients of Eq. (267) from $\{(x,t)||x| < 35, |t| < 35, x \in R, t \in R\}$ to $\{(x,t)||x| < 10^9, |t| < 10^9, x \in R, t \in R\}$



Korteweg–de Vries Equation (KdV), Different Analytical Methods for Solving the, Figure 3
The evolution of a non-traveling wave solution $u_{2,2}(x,t)$ with constant coefficients of Eq. (267) from $\{(x,t)||x| < 0.1, |t| < 0.1, x \in R, t \in R\}$ to $\{(x,t)||x| < 15, |t| < 15, x \in R, t \in R\}$



Korteweg–de Vries Equation (KdV), Different Analytical Methods for Solving the, Figure 4
The evolution of a non-traveling wave solution $u_{2,2}(x,t)$ with constant coefficients of Eq. (267) from $\{(x,t)||x| < 35, |t| < 35, x \in R, t \in R\}$ to $\{(x,t)||x| < 10^9, |t| < 10^9, x \in R, t \in R\}$

to $\{(x,t)||x|<10^9, |t|<10^9, x\in R, t\in R\}$ are shown by following Fig. 3 and Fig. 4.

Example 3 The solution (282) of Eq. (267) includes abundant traveling wave solutions with constant coefficients. For example, when $q_1(t)=\sin(t^4)$ in (294) is replaced by $q_1(t)=t/3$, we can obtain a traveling wave solution $u_{2,3}(x,t)$ with constant coefficients of Eq. (267) as follows:

$$u_{2,3}(x,t) = -\frac{2(\mathrm{e}^{-x/2-t/3}-10\mathrm{e}^{x/2+t/3}-2)^2}{9(\mathrm{e}^{-x/2-t/3}+1+2\mathrm{e}^{x/2+t/3})^2} + \frac{8+100\mathrm{e}^{x/2+t/3}+2\mathrm{e}^{-x/2-t/3}}{9\mathrm{e}^{-x/2-t/3}+9+18\mathrm{e}^{x/2+t/3}}+1/2\,. \quad (295)$$

Some figures of the long-term behavior of the $u_{2,3}(x,t)$ from $\{(x,t)||x|<0.1, |t|<0.1, x\in R, t\in R\}$ to $\{(x,t)||x|<423, |t|<423, x\in R, t\in R\}$ are shown by following Fig. 5 and Fig. 6. It is worth noting that the shape of the $u_{2,3}(x,t)$ degenerates gradually from a wave to a row of solitons on a straight line as the $|x|$ and $|t|$ in $u_{2,3}(x,t)$ continue to increase.

Example 4 The solution (282) of Eq. (267) includes abundant traveling wave solutions with variable coefficients. For example, when $p_1(t)=\tanh^2(t), q_1(t)=\sin(t^4)$ in (293) is replaced by $p_1(t)=1/2, q_1(t)=t/3$, we can obtain a traveling wave solution $u_{2,4}(x,t)$ with variable coefficients of Eq. (267) as follows:

$$u_{2,4}(x,t) = \frac{\begin{array}{l}1+12\cosh(t)\mathrm{e}^{-x/2-t/3}+12\mathrm{e}^{x/2+t/3}\\ +4\mathrm{e}^{x+2t/3}+72\cosh(t)+4\cosh^2(t)\mathrm{e}^{-x-2t/3}\end{array}}{2(2\cosh(t)\mathrm{e}^{-x/2-t/3}+1+2\mathrm{e}^{x/2+t/3})^2}\,. \quad (296)$$

Some figures of the long-term behavior of the $u_{2,4}(x,t)$ from $\{(x,t)||x|<0.1, |t|<0.1, x\in R, t\in R\}$ to $\{(x,t)||x|<10^9, |t|<10^9, x\in R, t\in R\}$ are shown by following Fig. 7 and Fig. 8. From Fig. 7 and Fig. 8, it is easy to find the global existence, asymptotic behaviors as $t\to+\infty$ and $|x|\to+\infty$ and the scattering properties of the solution $u_{2,4}(x,t)$ of Eq. (267). It is worth noting that the shape of $u_{2,4}(x,t)$ degenerates gradually from a trigonal wave to a set of solitons on a trigonal curve as $|x|$ and $|t|$ in $u_{2,4}(x,t)$ continue to increase. But the plane has become a scraggly surface over $\{(x,t)||x|<10^9, |t|<10^9, x\in R, t\in R\}$, namely, the $u_{2,4}(x,t)$ appears a mild disorder.

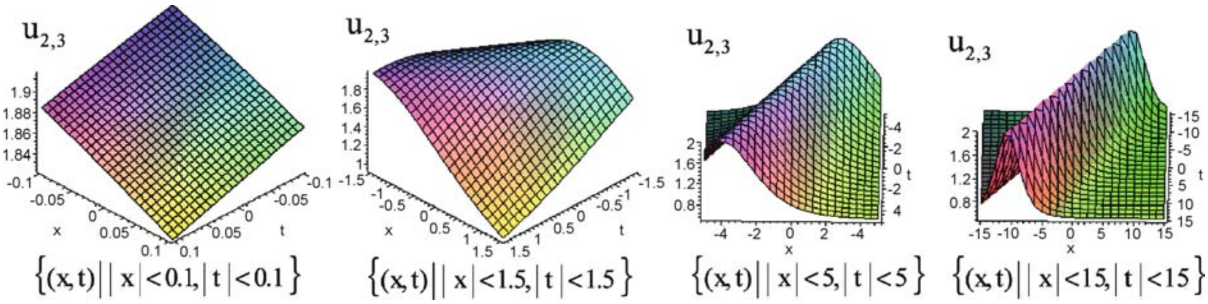
Remark 15 From the above long-term behavior of the $u_{2,1}(x,t), u_{2,2}(x,t), u_{2,3}(x,t), u_{2,4}(x,t)$, we find easily that the stabilization of the traveling wave solutions $u_{2,3}(x,t)$ with constant coefficients is best as $t\to+\infty$ and $|x|\to+\infty$. In addition, we still discover that the longterm behavior of a solution, which includes the global existence, asymptotic behaviors as $t\to+\infty$ and $|x|\to+\infty$, and the scattering properties of the solution, depends mainly on the functions that constitute the solution and the selection of all arbitrary constants and the parameters in the solution. We choose other nonlinear evolution equations [39,40,45,47] to test and obtain similar results. Therefore, we devise a new guess as follows.

R-guess The long-term behavior of a solution for a given NLEEs (261), which includes the global existence, asymptotic behaviors as $t\to+\infty$ and $|x|\to+\infty$, and the scattering properties of the solution, depends mainly on the functions that constitute the solution and the selection of all arbitrary constants and parameters in the solution.

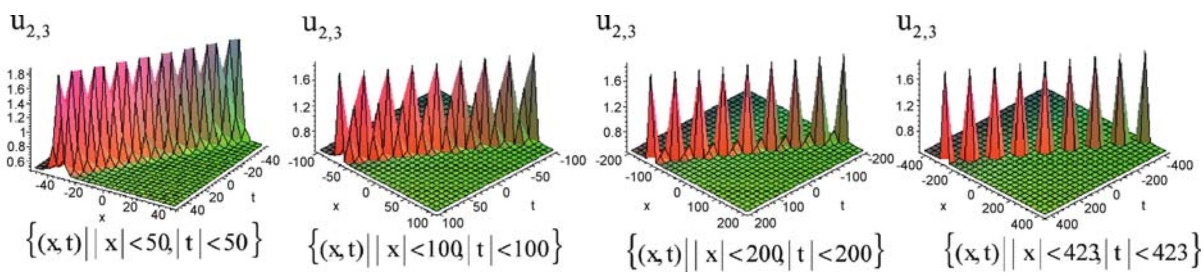
Future Directions

In recent years, directly searching for exact solutions of NLPDEs has become more and more attractive partly due to the availability of computer symbolic systems like Maple or mathematica which allow us to perform complicated and tedious algebraic calculation on a computer, as well as helping us to find new exact solutions of NLPDEs in mathematical physics. Although many powerful analytical methods for solving NLPDEs have been presented, the methods fail to satisfy the developmental needs in physics, mechanics, chemistry, biology, computer science, etc. There is much new work to be done on analytical methods for solving NLPDEs. The main future directions are as follows:

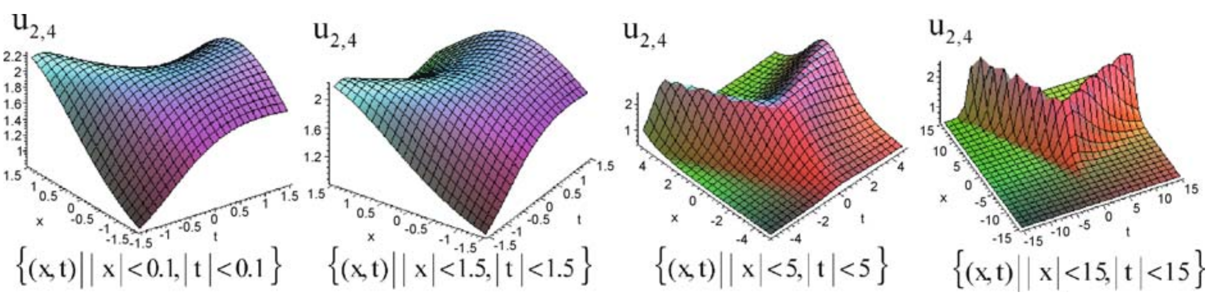
1. Researching new and uniform analytical methods based on the ideas of unification methods, algorithm realization and mechanization for solving NLPDEs (we have done some new work in Sect. "The Generalized Hyperbolic Function–Bäcklund Transformation Method and Its Application in the $(2+1)$-Dimensional KdV Equation" of this article).
2. Researching mechanization methods and developing their computer software for showing the long-playing traveling state of the exact solutions of NLPDEs (we have done some new work in [40]).
3. Developing and improving existing analytical methods, computer algebraic systems, and software so that analytical methods can be actualized with complete mechanization by their computer software at a familiar computer (we have done some new work in [39,40,41,42,43,44,45,46,47,48]).



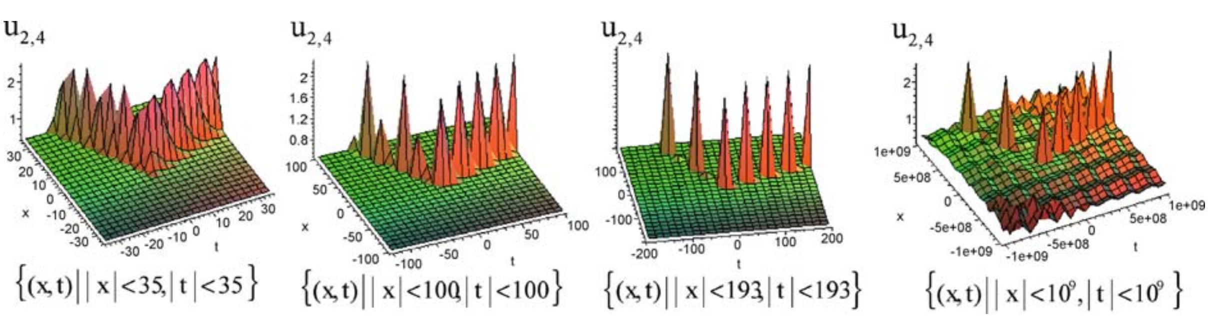
Korteweg–de Vries Equation (KdV), Different Analytical Methods for Solving the, Figure 5
The evolution of a traveling wave solution $u_{2,3}(x,t)$ with constant coefficients of Eq. (267) from $\{(x,t)||x| < 0.1, |t| < 0.1, x \in R, t \in R\}$ to $\{(x,t)||x| < 15, |t| < 15, x \in R, t \in R\}$



Korteweg–de Vries Equation (KdV), Different Analytical Methods for Solving the, Figure 6
The evolution of a traveling wave solution $u_{2,3}(x,t)$ with constant coefficients of Eq. (267) from $\{(x,t)||x| < 50, |t| < 50, x \in R, t \in R\}$ to $\{(x,t)||x| < 423, |t| < 423, x \in R, t \in R\}$



Korteweg–de Vries Equation (KdV), Different Analytical Methods for Solving the, Figure 7
The evolution of a traveling wave solution $u_{2,4}(x,t)$ with variable coefficients of Eq. (267) from $\{(x,t)||x| < 0.1, |t| < 0.1, x \in R, t \in R\}$ to $\{(x,t)||x| < 15, |t| < 15, x \in R, t \in R\}$



Korteweg–de Vries Equation (KdV), Different Analytical Methods for Solving the, Figure 8
The evolution of a traveling wave solution $u_{2,4}(x,t)$ with variable coefficients of Eq. (267) from $\{(x,t)||x| < 35, |t| < 35, x \in R, t \in R\}$ to $\{(x,t)||x| < 10^9, |t| < 10^9, x \in R, t \in R\}$

4. Researching new analytical and numerical complex methods and their computer software for solving those NLPDEs that can't make use of analytical methods. The complex methods can be actualized with complete mechanization by their computer software at a familiar computer (we have done some new work in [40,53]).
5. One can calculate the solutions of NLPDEs as easily as counting or addition by using a small calculator with the development of mathematics and computer science (we have done some new work in [39,40,41,42,43,44,45,46,47,48,52,53,54] and Ren 2055 and 2007, and Ren und Sun 2004).

Acknowledgment

The author would like to express his gratitude to all the people whose efforts contributed to the various editions of this paper. One reviewer who made useful recommendations for this paper is Professor M. Atef Helal at Mathematics Department, Cairo University, Egypt. The author is grateful to Professor M. Atef Helal for his time and valuable suggestions. The author thanks Editor-in-Chief Robert Meyers of the Encyclopedia of Complexity and System Science, Professor M.S. El Naschie at Frankfurt Institute for the advancement of Theoretical Physics, Frankfurt, Germany, Professor Hong-Qing Zhang at Department of Applied Mathematics, Dalian University of Technology, Dalian, PR China, and Professor Abdul-Majid Wazwaz at Department of Mathematics and Computer Science, Saint Xavier University, Chicago, USA, for their help. This work has been supported partially by a National Key Basic Research Project of China (2004 CB 318000).

Bibliography

Primary Literature

1. 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
2. Khaters AH, El-Kalaawy H, Helal MA (1997) Two new classes of exact solutions for the KdV equation via Bäcklund transformations. Chaos Solitons Fractals 8(12):1901–1909
3. Helal MA (2001) A Chebyshev spectral method for solving Korteweg–de Vries equation with hydrodynamical application. Chaos Solitons Fractals 12:943–950
4. Brugarino T, Sciacca M (2008) A direct method to find solutions of some type of coupled Korteweg–de Vries equations using hyperelliptic functions of genus two. Phys Lett A 372(11):1836–1840
5. Leibovitch S (1970) Weakly nonlinear waves in rotating fluids. J Fluid Mech 42:803–825
6. RM Miura, CS Gardner, MD Kruskal (1968) Korteweg–de Vries equation and generalizations. II Existence of conservation laws and constants of motion. J Math Phys 9:1204–1209
7. MA Helal, MS Mehanna (2007) A comparative study between two different methods for solving the general Korteweg–de Vries equation (GKdV). Chaos Solitons Fractals 33:725–739
8. Hamdia S, Enrighta WH, Schiesserb WE, Gottliebc JJ (2005) Exact solutions and conservation laws for coupled generalized Korteweg–de Vries and quintic regularized long wave equations. Nonlinear Anal 63:1425–1434
9. Sarma J (2008) Solitary wave solution of higher-order Korteweg–de Vries equation. Chaos Solitons Fractals, in press
10. Karpman VI (1998) Evolution of radiating solitons described by the fifth-order Korteweg–de Vries type equations. Phys Lett A 244:394–396
11. Hassan MM (2004) Exact solitary wave solutions for a generalized KdV–Burgers equation. Chaos Solitons Fractals 19:1201–1206
12. Khare A, Cooper F (1993) One-parameter family of soliton solutions with compact support in a class of generalized Korteweg–de Vries equation. Phys Rev E 48(6):4843–4
13. Dod'an K (2005) An application for the higher order modified KdV equation by decomposition method. Commun Nonlinear Sci Num Simul 10:693–702
14. Shen J, Xu W (2007) Travelling wave solutions in a class of generalized Korteweg–de Vries equation. Chaos Solitons Fractals 34:1299–1306
15. Clarke S, Malomed BA, Grimshaw R (2002) Dispersion management for solitons in a Korteweg–de Vries system. Chaos 12(1):8–15
16. Houria T (2007) Energy transfer in a dispersion-managed Korteweg–de Vries system. Math Comput Simul 76:283–292
17. Kaya D, Aassila M (2002) An application for a generalized KdV equation by the decomposition method. Physic Lett A 299:201–206
18. Shek ECM, Chow KW (2008) The discrete modified Korteweg–de Vries equation with non-vanishing boundary conditions: Interactions of solitons. Chaos Solitons Fractals 36:296–302
19. Zhang S (2007) A generalized auxiliary equation method and its application to the $(2+1)$-dimensional KdV equations. Appl Math Comput 188:1–6
20. Song L, Zhang H (2007) A new Korteweg–de Vries equation-based sub-equation method and its application to the $(2+1)$-dimensional Korteweg–de Vries equation. Appl Math Comput 187:1368–1372
21. Hirota R (1971) Exact solution of the Korteweg–de Vries equation for multiple collisions of solitons. Phys Rev Lett 27:1192–1194
22. Wadati M, Sanuki H, Konno K (1975) Relationships among inverse method, Bäcklund transformation and an infinite number of conservation laws. Prog Theor Phys 53:419–436
23. Weiss J, Tabor M, Carnevale G (1983) The Painlevel property for partial differential equations. J Math Phys 24:522–526
24. Fan EG, Zhang HQ (1998) A note on the homogeneous balance method. Phys Lett A 246:403–406
25. Wazwaz AM (2005) The tanh and the sine-cosine methods for the complex modified KdV and the generalized KdV equations. Comput Math Appl 45:1101–1112
26. Wang ML (1996) Exact solution for a compound KdV–Burgers equations. Phys Lett A 213:279–287
27. He JH (2005) Homotopy perturbation method for bifurcation of nonlinear problems. Int J Nonlinear Sci Numer Simul 6:207–208

28. He JH (2004) Variational principles for some nonlinear partial differential equations with variable coefficients. Chaos Solitons Fractals 19:847–851
29. He JH (2006) Some asymptotic methods for strongly nonlinear equations. Int J Modern Phys B 20:1141–1199
30. He JH (2006) Non-perturbative Methods for Strongly Nonlinear Problems, Dissertation. Verlag im Internet, Berlin
31. He JH, Wu XH (2006) Exp-function method for nonlinear wave equations. Chaos Solitons Fractals 30:700–708
32. Abassy TA, El-Tawil MA, Saleh HK (2004) The solution of KdV and mKdV equations using adomian pade approximation. Int J Nonlinear Sci Numer Simul 5:327–340
33. Liu SK, Fu ZT, Liu SD, Zhao Q (2001) Jacobi elliptic function expansion method and periodic wave solutions of nonlinear wave equations. Phys Lett A 289:69–74
34. Zhou YB, Wang ML, Wang YM (2003) Periodic wave solutions to a coupled KdV equations with variable coefficients. Phys Lett A 308:31–36
35. Chen Y, Yan ZY (2006) Weierstrass semi-rational expansion method and new doubly periodic solutions of the generalized Hirota–Satsuma coupled KdV system. Appl Math Comput 177:85–91
36. Chen Y, Wang Q (2006) A unified rational expansion method to construct a series of explicit exact solutions to nonlinear evolution equations. Appl Math Comput 177:396–409
37. Zhang S, Xia TC (2006) A further improved extended Fan sub-equation method and its application to the $(3+1)$-dimensional Kadomstev–Petviashvili equation. Phys Lett A 356:119–123
38. Xu GQ, Li ZB (2005) Exact traveling wave solutions of the Whitham–Broer–Kaup and Broer–Kaup–Kupershmidt equations. Chaos Solitons Fractals 24:549–556
39. Ren YJ, Zhang HQ (2006) New generalized hyperbolic function and Auto-Bäcklund transformation to find new exact solutions of the $(2+1)$-dimensional NNV equation. Phys Lett A 357:438–448
40. Ren YJ (2007) Some investigation on the methods for finding solutions of nonlinear evolution equations and realization of mathematics mechanization, Yujie Ren' Ph.D Dissertation. Dalian University of Technology, Dalian, p 39–50
41. Ren YJ, Zhang HQ (2006) A generalized F-expansion method to find abundant families of Jacobi Elliptic Function solutions of the $(2+1)$-dimensional Nizhnik–Novikov–Veselov equation. Chaos Solitons Fractals 27(4):959–979
42. Ren YJ, Liu ST, Zhang HQ (2006) On a Generalized Improved F-expansion method. Commun Theor Phys 45(1):15–28
43. Ren YJ, Liu ST, Zhang HQ (2007) A new generalized algebra method and its application in the $(2+1)$ dimensional Boiti–Leon–Pempinelli equation. Chaos Solitons Fractals 32(5):1655–1665
44. Ren YJ, Zhang DH, Cheng F, Zhang HQ (2008) New generalized method to construct new non-traveling waves solutions and traveling waves solutions of K-D equations. Int J Comput Math 1:1–16
45. Ren YJ, He JH, Zhang HQ (2008) A new Exp-Bäcklund transformation method and its application. Chaos Solitons Fractals (accepted)
46. Ren YJ (2007) A new Exp-Bäcklund method. Ph D Dissertation. Dalian University of Technology, p 158–171
47. Ren YJ, He JH, Zhang HQ (2008) A new Exp-N solitary-like method and its application in the $(2+1)$-dimensional generalized KdV–Burgers equation. Chaos Solitons Fractals (accepted)
48. Ren YJ (2007) A new the Exp-N soliton-like method. Ph D Dissertation. Dalian University of Technology, p 172–178
49. Wu YT, Geng XG, Hu XB, Zhu SM (1999) A generalized Hirota–Satsuma coupled Korteweg–de Vries equation and Miura transformations. Phys Lett A 255:259–264
50. Fan EG, Dai HH (2003) A direct approach with computerized symbolic computation for finding a series of traveling waves to nonlinear equations. Computer Physics Communications 153:17–21
51. Yan ZY (2004) An improved algebra method and its applications in nonlinear wave equations. Chaos Solitons Fractals 21:1013–1021
52. Wang DS, Ren YJ, Zhang HQ (2005) Further extended sih-cosh and sin-cos methods and new non-travelng wave solutions of the $(2+1)$-dimensional dispersive long wave equation. Appl Math E-notes 5:157–163
53. Ren YJ, Li YJ, Song FG, Zhang HQ (2006) A new compound operation method of the numerical solutions and analytic solution of partial differential equation. International Conf. on Multimedia, Information Technology and its Applications, Dalian, China, p 328–330
54. Ren YJ, Song FG, Li YJ, Zhang HQ (2006) A mechanization method on the mixed operation of the numerical solutions, analytic solution, and showing their figures simultaneously for differential equation. International Conf. on Multimedia, Information Technology and its Applications, Dalian, China, p 276–278

Books and Reviews

Ablowitz MJ, Clarkson PA (1991) Soliton, Nonlinear Evolution Equations and Inverse Scattering. Cambridge University Press, New York

Dodd RK, Eilbeck JC, Gibbon JD, Morris HC (1984) Solitons and nonlinear wave equations. Academic Press, London

Drazin PG, Johnson RS (1989) Solitons: an introduction. Cambridge University Press, Cambridge

Fan EG (2004) Integrable System and Computer Algebra. Science Press, Beijing

Guo BL (1996) Nonlinear Evolution Equation [M]. Shanghai Science and Technology Press, Shanghai

Liu SK, Liu SD (2000) Nonlinear Equations in Physics. Peking University Press, Beijing

Miura RM (ed) (1976) Backlaund Transformation C. Springer Lecture Notes in Mathematica, vol 515. Springer, New York

Miurs MR (1978) Bäcklund Transformation. Springer, Berlin

Ren YJ (2005) Advanced Mathematics – Multivariable Calculus and Experiments (MATLAB edition). China Machine Press, Beijing

Ren YJ (2007) Numerical Analysis and its MATLAB Realization. Higher Education Press, Beijing

Ren YJ, Sun WH (2004) Advanced Mathematics – Unitary Calculus and Experiments (MATLAB edition). China Machine Press, Beijing

Rogerscand Shaduick WR (1982) Backland Transformation and their Applications. Academic Press, New York

Whitham GB (1974) Linear and nonlinear waves. Wiley, New York

Zhuo LY et al (2000) Nonlinear Physics Threoy and Application [M]. Science Press, Peking

Korteweg–de Vries Equation (KdV), History, Exact *N*-Soliton Solutions and Further Properties of the

YI ZANG
Department of Mathematics,
Zhejiang Normal University, Jinhua, China

Article Outline

Glossary
Definition of the Subject
Introduction
Inverse Scattering Transform for the KdV Equation
Exact N-soliton Solutions of the KdV Equation
Further Properties of the KdV Equation
Future Directions
Bibliography

Glossary

Soliton A soliton is a solitary wave which asymptotically preserves its shape and velocity upon nonlinear interaction with other solitary waves, or more generally, with another (arbitrary) localized disturbance.

Korteweg–de Vries equation In mathematics, the Korteweg-de Vries (KdV) equation is a mathematical model of waves on shallow water surfaces. It is particularly famous as the prototypical example of an exactly solvable model, that is, a nonlinear partial differential equation whose solutions can be exactly and precisely specified.

Inverse scattering transform The identification of the KdV equation as an isospectral flow of the Schrödinger operator enabled Gardner, Greene, Kruskal and Miura (GGKM) to devise a method of solving the KdV equation (with 'rapidly decreasing' boundary conditions), called the inverse scattering or inverse spectral transform (IST). This is a direct generalization of the Fourier transform used to solve linear equations and can be represented by essentially the same scheme.

Hirota bilinear method In 1971, Hirota developed a direct method for finding N-soliton solutions of nonlinear evolution equations, also named Hirota bilinear method. The key step is to transform the equation into a bilinear form, from which we can get soliton solutions successively by means of a kind of perturbational technique.

Definition of the Subject

Among integrable equations is the celebrated KdV equation, which serves as a model equation governing weakly nonlinear long waves whose phase speed attains a simple maximum for waves of infinite length. It motivates us to explore beauty hidden in nonlinear differential (and difference) equations. The equation is named for Diederik Korteweg and Gustav de Vries. The remarkable and exceptional discovery of the inverse scattering transform is one of important developments in the field of applied mathematics, which comes from the study of the KdV equation. There are various algebraic and geometric characteristics that the KdV equation possesses, for example, infinitely many symmetries and infinitely many conserved densities, the Lax representation, bi-Hamiltonian structure, loop group, and the Darboux-Bäcklund transformation. More significantly, many physically important solutions to the KdV equation can be presented explicitly through a simple, specific form, called the Hirota bilinear form.

Introduction

A nice story about the history and the underlying physical properties of KdV equation can be found at an Internet page of the Herriot-Watt University in Edinburgh (Scotland). The following text is taken from that page:

Over one hundred and fifty years ago, which conducting experiments to determine the most efficient design for canal boats, a young Scottish engineer named John Scott Russell (1808–1882) made a remarkable scientific discovery. Here his original text as he described it in Russell [1]:

I believe I shall best introduce this phenomenon by describing the circumstances of my own first acquaintance with it. 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, a name which it now very generally bears.

He followed this observation with a number of experiments during which he determined the shape of a solitary wave to be that of a $sech^2()$ function. He also determined the relationship of the speed of the wave to its amplitude. At that time there was no equation describing such water waves and having such a solution. Thus John Scott Russell discovered the solution to an as yet unknown equation!

Further investigations were undertaken by Airy, Stokes, Boussinesq and Rayleigh in an attempt to understand this phenomenon. Boussinesq and Rayleigh independently obtained approximate descriptions of the solitary wave; Boussinesq derived a one-dimensional nonlinear evolution equation, which now bears his name, in order to obtain his results.

These investigations provoked much lively discussion and controversy as to whether the inviscid equations of water waves would possess such solitary wave solutions. The issue was finally resolved by Korteweg and de Vries [2]. After performing a Galilean and a variety of scaling transformations, the KdV equation can be written in simplified form:

$$u_t - 6uu_x - u_{xxx} = 0\,, \tag{1}$$

where subscripts denote partial differentiations. In particular, if we now assume a solution in the form of a traveling wave $u(x,t) = f(x+ct)$, then Eq. (1) can be integrated: imposing the boundary conditions at large distances that $u(x,t)$ tends to 0 sufficiently fast as $x \longrightarrow \pm\infty$, we find the exact solution

$$u(x,t) = \frac{c}{2}\operatorname{sech}^2 \frac{1}{2}\sqrt{c}(x+ct+\delta)\,, \tag{2}$$

where δ is the phase. This clearly represents the solitary wave observed by John Scott Russell and shows that the peak amplitude is exactly half the speed. Thus larger solitary waves have greater speeds. This suggests a numerical experiment: start with two solitary wave solutions, with centers well separated and the larger to the right. Initially, with negligible overlap, they will evolve *independently* as solitary wave solutions. However, the larger, faster one will start to overtake the smaller and the non linearity will play a significant role. For most dispersive evolution equations these solitary waves would scatter in elastically and lose 'energy' to radiation. Not so for the KdV equation: after a fully nonlinear interaction, the solitary waves reemerge, retaining their identities (same speed and form), suffering nothing more than a phase shift (modified δ's, representing a displacement of their centers). It was after a similar numerical experiment that Kruskal and Zabusky [3] coined the name 'soliton', to reflect the particle-like behavior of the solitary waves under interaction.

It was not until the mid 1960's when applied scientists began to use modern digital computers to study nonlinear wave propagation that the soundness of Russell's early ideas began to be appreciated. He viewed the solitary wave as a self-sufficient dynamic entity, a "thing" displaying many properties of a particle. From the modern perspective it 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 tornado's to the Great Red Spot of Jupiter, from the elementary particles of matter to the elementary particles of thought. For a more detailed and technical account of the solitary wave, see [4,5].

Inverse Scattering Transform for the KdV Equation

We now briefly discuss the inverse scattering transform for the KdV Eq. (1). First consider the Miura map $u = -v_x - v^2 + \lambda$, which can be viewed as a Riccati equation for v and thus linearized by the substitution $v = \psi_x/\psi$, giving

$$L\psi \equiv \psi_{xx} + u\psi = \lambda\psi\,. \tag{3}$$

This is the time-independent Schrödinger equation with $u(x,t)$ playing the role of potential and λ the energy. It is important to realize that *t is not the time of the time-independent Schrödinger equation*. We think of x as the spatial variable and t as a parameter. Considered as a Sturm–Liouville eigenvalue problem it is natural to ask how λ and ψ change with t as $u(x,t)$ evolves from some initial state according to the KdV equation. GGKM [6,7] discovered the remarkable fact that the (discrete part of the) spectrum necessarily remains constant in 'time' while the corresponding wave functions ψ evolve according to a very simple *linear differential equation*.

Today (following Lax [8]) we usually take the opposite route. We postulate that ψ evolves through a linear differential equation:

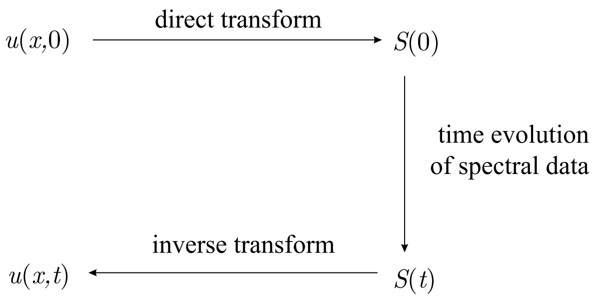
$$\psi_t = P\psi\,. \tag{4}$$

Eq. (3) and (4) form an overdetermined system, whose integrability conditions can be written:

$$L_t = [P,L] = PL - LP\,. \tag{5}$$

A consequence of (5) is that all eigenvalues corresponding to the function $u(x,t)$ remain constant. When P is given by:

$$P = -\left(4\partial^3 + 6u\partial + 3u_x\right)\,, \qquad \text{(where } \partial \text{ is a differential operator)} \tag{6}$$



Korteweg–de Vries Equation (KdV), History, Exact *N*-Soliton Solutions and Further Properties of the, Figure 1
Inverse Scattering Transform

(5) reduces to the KdV Eq. (1). Since, under these conditions, the spectrum of L remains constant, the KdV equation is referred to as an *isospectral flow*. Eq. (5) is called the *Lax representation* of the KdV equation.

The identification of the KdV equation as an isospectral flow of the Schrödinger operator enabled GGKM to devise a method of solving the KdV equation (with 'rapidly decreasing' boundary conditions), called the inverse scattering or inverse spectral transform (IST). This is a direct generalization of the Fourier transform used to solve linear equations and can be represented by essentially the same scheme (Fig. 1).

For these boundary conditions, the solutions of (3) are characterized by their asymptotic properties as $x \to \pm\infty$. For a given potential function, the continuous and discrete spectrum are treated separately. Corresponding to the continuous spectrum ($\lambda = -k^2$) the solutions are asymptotically oscillatory, characterized by two coefficients:

$$\psi \sim \begin{cases} T(k)\mathrm{e}^{-ikx} & x \to -\infty \\ \mathrm{e}^{-ikx} + R(k)\mathrm{e}^{ikx} & x \to +\infty \end{cases} \tag{7}$$

subject to the condition $|R|^2 + |T|^2 = 1$. The constants $R(k)$ and $T(k)$ are respectively called the *reflection and transmission* coefficients, from their quantum mechanical interpretation. Under mild conditions on the potential function, the Schrödinger operator L has only a finite number of discrete eigenvalues $\{\kappa_n^2\}$. The corresponding eigenfunctions are square integrable and are the 'bound states' of quantum mechanics with asymptotic properties (for $\kappa_n > 0$):

$$\psi_n \sim \begin{cases} \widetilde{c}_n \mathrm{e}^{\kappa_n x} & x \to -\infty \\ c_n \mathrm{e}^{-\kappa_n x} & x \to +\infty \end{cases} . \tag{8}$$

The *direct scattering transform* constructs the quantities $\{T(k), R(k), \kappa_n, c_n\}$ from a given potential function. The important inversion formula were derived by Gel'fand and Levitan in 1955 [9]. These enable the potential u to be constructed out of the spectral or scattering data $S = \{R(k), \kappa_n, c_n\}$. This is considerably more complicated than the Inverse Fourier Transform, involving the solution of a nontrivial integral equation, whose kernel is built out of the scattering data (see [10,11,12,13] for descriptions of this).

To solve the KdV equation we first construct the scatting data $S(0)$ from the initial condition $u(x,0)$. As a consequence of (4) (with an additional constant) with the given boundary conditions, the scatting data evolves in a very simple way. Indeed, we can give explicit formula:

$$R(k,t) = R(k,0)\mathrm{e}^{8ik^3 t}, \quad c_n(t) = c_n(0)\mathrm{e}^{-4\kappa_n^3 t} . \tag{9}$$

Using the inverse scattering transform on the scattering data $S(t)$, we obtain the potential $u(x,t)$ and thus the solution to the initial value problem for the KdV equation.

This process cannot be carried out *explicitly* for arbitrary initial data, although, in this case, it gives a great deal of information about the solution $u(x,t)$. However, whenever the reflection coefficient is zero, the kernel of Gel'fand-Levitan integral equation becomes separable and explicit solutions can be found. It is in this way that the N-soliton solution is constructed by IST from the initial condition:

$$u(x,0) = N(N+1)\mathrm{sech}^2 x . \tag{10}$$

The general formula for the multi-soliton solution is given by:

$$u(x,t) = 2(\ln \det M)_{xx} , \tag{11}$$

where M is a matrix built out of the discrete scattering data.

Exact *N*-soliton Solutions of the KdV Equation

Besides the IST, there are several analytical methods for obtaining solutions of the KdV equation, such as Hirota bilinear method [14,15,16], Bäcklund transformation [17], Darboux transformation [18], and so on. The existence of such analytical methods reflects a rich algebraic structure of the KdV equation. In Hirota's method, we transform the equation into a bilinear form, from which we can get soliton solutions successively by means of a kind of perturbational technique. The Bäcklund transformation is also employed to obtain solutions from a known solution of the concerned equation. In what follows, we will mainly discuss the Hirota bilinear method to derive the N-soliton solutions of the KdV equation.

It is well known that Hirota developed a direct method for finding N-soliton solutions of nonlinear evolution equations. In particular, we shall discuss the KdV bilinear form

$$D_x\left(D_t + D_x^3\right)f \cdot f = 0\,, \tag{12}$$

by the dependent variable transformation

$$u(x,t) = 2(\ln f)_{xx}\,. \tag{13}$$

Here the Hirota bilinear operators are defined by

$$D_x^m D_t^n a{\cdot}b = (\partial_x - \partial_{x\prime})^m(\partial_t - \partial_{t\prime})^n a(x,t)b(x\prime,t\prime)|_{x\prime=x,t\prime=t}\,. \tag{14}$$

We expand f as power series in a parameter ε

$$f(x,t) = 1 + f^{(1)}\varepsilon + f^{(2)}\varepsilon^2 + \cdots + f^{(j)}\varepsilon^j + \cdots\,. \tag{15}$$

Substituting (15) into (12) and equating coefficients of powers of ε gives the following recursion relations for the $f^{(n)}$

$$\varepsilon\colon f^{(1)}_{xxxx} + f^{(1)}_{xt} = 0\,, \tag{16}$$

$$\varepsilon^2\colon f^{(2)}_{xxxx} + f^{(2)}_{xt} = -\frac{1}{2}(D_xD_t + D_x^4)f^{(1)} \cdot f^{(1)}\,, \tag{17}$$

$$\varepsilon^3\colon f^{(3)}_{xxxx} + f^{(3)}_{xt} = -\left(D_xD_t + D_x^4\right)f^{(1)} \cdot f^{(2)}\,, \tag{18}$$

and so on. N-soliton solutions of the KdV equation are found by assuming that $f^{(1)}$ has the form

$$f^{(1)} = \sum_{j=1}^{N}\exp(\eta_j), \quad \eta_j = k_jx - \omega_jt + x_{j0}\,, \tag{19}$$

and k_j, ω_j and x_{j0} are constants, provided that the series (15) truncates.

For $N = 1$, we take

$$f^{(1)} = \exp(\eta_1),$$

and by solving (16) we find that

$$f^{(n)} = 0, \quad \text{for } n = 2, 3, \ldots\,.$$

Therefore we have

$$f_1 = 1 + \exp(\eta_1), \quad \omega_1 = -k_1^3\,,$$

and

$$u(x,t) = \frac{k_1^2}{2}\operatorname{sech}^2\frac{1}{2}\left(k_1x - k_1^3t + x_{10}\right)\,. \tag{20}$$

For $N = 2$, the two-soliton solution for the KdV equation is similarly obtained from

$$u(x,t) = 2(\ln f_2)_{xx}\,,$$

where

$$f_2 = 1 + \exp(\eta_1) + \exp(\eta_2) + \exp(\eta_1 + \eta_2 + A_{12})\,. \tag{21}$$

Frequently the N-soliton solutions are obtained as follows

$$f_N = \sum_{\mu=0,1}\exp\left(\sum_{j=1}^{N}\mu_j\eta_j + \sum_{1\le j<l}^{N}\mu_j\mu_lA_{jl}\right)\,, \tag{22}$$

where the first $\sum_{\mu=0,1}$ means a summation over all possible combinations of $\mu_j = 0, 1$ and $\sum_{1\le j<l}^{N}$ means a summation over all possible pairs (j, l) chosen from the set $\{1, 2, \ldots, N\}$, with the condition that $j < l$.

Further Properties of the KdV Equation

Conservation Laws

It was this numerical evidence which prompted Kruskal and his co-workers in Princeton to analytically investigate the KdV equation. Initially, Miura investigated local conservation laws:

$$\partial_tT + \partial_xF = 0\,, \tag{23}$$

where T and F are polynomials in $u(x,t)$ and its x-derivatives and where ∂_t and ∂_x denote total derivatives. With appropriate boundary conditions this leads to a conserved quantity (constant of the motion). Integrating (23) with respect to x we get:

$$\partial_t\int_A^B T + [F]_A^B = 0\,. \tag{24}$$

Under periodic (in x) boundary conditions with $A - B$ an integer multiple of the period or with $u(x,t)$ rapidly decreasing as $x \to \pm\infty$ and $(A, B) = (-\infty, +\infty)$, the square bracket in (24) vanishes and we have the constant of motion $\partial_t\int_A^B Tdx$. The quantities T and F are respectively called conserved density and flux. Each T is, in fact, only determined up to an exact x-derivative, so defines an equivalence class of conserved densities:

$$T \sim T + \partial_xS\,, \tag{25}$$

since this only adds $\partial_t S$ to F and leaves the value of $\int T dx$ unchanged. For the KdV equation the first three are:

$$\begin{aligned} T_0 &= u\,, & F_0 &= -u_{xx} - 3u^2,\\ T_1 &= \frac{1}{2}u^2\,, & F_1 &= -uu_{xx} + \frac{1}{2}u_x^2 - 2u^3\,,\\ T_2 &= u^3 - \frac{1}{2}u_x^2\,, & F_2 &= u_x u_{xxx} - \frac{1}{2}u_{xx}^2\\ & & &\quad - 3u^2u_{xx} + 6uu_x^2 - \frac{9}{2}u^4\,. \end{aligned} \tag{26}$$

The first of these is just the equation itself. These three conservation laws have same physical interpretation, so it was no surprise that they exist. However, Miura discovered several more by direct calculation and was led to the conjecture that there should be *infinitely many*.

In order to ascertain whether the KdV equation was the only such equation with so many conservation laws Miura investigated equations of the form:

$$u_t = u_{xxx} + 6u^n u_x\,, \tag{27}$$

and found that for $n = 1$ and $n = 2$ (and *only* these values) there existed many conservation laws. With a slight change in notation the second of these, called the modified KdV (MKdV) equation, can be written:

$$v_t = v_{xxx} - 6v^2 v_x = \left(v_{xx} - 2v^3\right)_x\,. \tag{28}$$

The first few conservation laws correspond to conserved densities and fluxes:

$$\begin{aligned} \tilde{T}_{-1} &= v\,, & \tilde{F}_{-1} &= -v_{xxx} + 2v^3,\\ \tilde{T}_0 &= v^2\,, & \tilde{F}_0 &= -2vv_{xx} + v_x^2 + 3v^4,\\ \tilde{T}_1 &= \frac{1}{2}\left(v_x^2 + v^4\right)\,, & \tilde{F}_1 &= -v_x v_{xxx} + \frac{1}{2}v_{xx}^2\\ & & &\quad - 2v^3 v_{xx} + 6v^2v_x^2 + 2v^6\,. \end{aligned} \tag{29}$$

The Lax Hierarchy

In [8] Lax reformulated GGKM's discovery [6,7] of the isospectral nature of the KdV equation in algebraic form:

$$\left.\begin{aligned} L\psi &\equiv (\partial^2 + u)\psi = \lambda\psi\\ \psi_t = P\psi &\equiv -\left(4\partial^3 + 6u\partial + 3u_x\right)\psi\\ \lambda_t &= 0 \end{aligned}\right\}$$
$$\Rightarrow L_t = [P, L] \equiv PL - LP\,. \tag{30}$$

This led Lax to an interesting generalization:

$$\left.\begin{aligned} L\psi &\equiv (\partial^2 + u)\psi = \lambda\psi\\ \psi_{t_m} = P_{(m)}\psi &\equiv \partial^{2m+1} + \sum_{i=0}^{2m-1} b_i\partial^i\\ \lambda_t &= 0 \end{aligned}\right\}$$
$$\Rightarrow L_{t_m} = [P_{(m)}, L]\,. \tag{31}$$

The integrability condition $L_{t_m} = [P_{(m)}, L]$ is explicitly written as:

$$u_{t_m} = (2m+1)u_x - 2b_{2m-1x}\partial^{2m} + \cdots + (P_{(m)}u - Lb_0)\,. \tag{32}$$

Equating coefficients of $\partial^i, i = 0, \ldots, 2m$, gives us $2m + 1$ equations, from which we deduce the $2m$ coefficients $b_0, \ldots, b_{2m-1}$:

$$\begin{aligned} b_{2m-1} &= \frac{1}{2}(2m+1)u,\\ b_{2m-2} &= \frac{1}{4}(2m+1)(2m-1)u_x, \ldots \end{aligned} \tag{33}$$

together with the isospectral flow:

$$u_{t_m} = P(m)u - Lb_0\,. \tag{34}$$

Nontrivial flows only exist for odd-order equations, since the adjoint of the integrability condition implies $P^\dagger = -P$. There exists an infinite hierarchy of such isospectral flows, the first three of which are:

$$u_{t_0} = u_x\,, \tag{35}$$

$$u_{t_1} = \frac{1}{4}(u_{xxx} + 6uu_x)\,, \tag{36}$$

$$u_{t_2} = \frac{1}{16}(u_{xxxxx} + 10uu_{xxx} + 20u_x u_{xx} + 30u^2u_x)\,, \tag{37}$$

corresponding respectively to operators:

$$P_{(0)} = \partial\,, \tag{38}$$

$$P_{(1)n} = \partial^3 + \frac{3}{4}(u\partial + \partial u)\,, \tag{39}$$

$$\begin{aligned} P_{(2)} &= \partial^5 + \frac{5}{4}(u\partial^3 + \partial^3 u)\\ &\quad + \frac{5}{16}((3u^2 - uxx)\partial + \partial(3u^2 - u_{xx}))\,. \end{aligned} \tag{40}$$

Future Directions

In mathematics, the KdV equation is a mathematical model of waves on shallow water surfaces. It is particularly famous as the prototypical example of an exactly solvable model, that is, a nonlinear partial differential equation whose solutions can be exactly and precisely specified. The solutions in turn are the prototypical examples of solitons; these may be found by means of the inverse scattering transform. The mathematical theory behind the KdV equation is rich and interesting, and, in the broad sense, is a topic of active mathematical research.

The KdV equation has several connections to physical problems. In addition to being the governing equation of the string in the Fermi–Pasta–Ulam problem in the continuum limit, it approximately describes the evolution of long, one-dimensional waves in many physical settings, including shallow-water waves with weakly nonlinear restoring forces, long internal waves in a density-stratified ocean, ion-acoustic waves in a plasma, acoustic waves on a crystal lattice, and more.

The scientists' interest for analytical solutions of the KdV equation stems from the fact that in applying numerical methods to nonlinear partial differential equations, the KdV equation is well suited as a test object, since having an analytical solution statements can be made on the quality of the numerical solution in comparing the numerical result to the exact result. More significantly, the existence of several analytical methods reflects a rich algebraic structure of the KdV equation. It is hoped that the study of the KdV equation could further assist in understanding, identifying and classifying nonlinear integrable differential equations and their exact solutions.

Bibliography

Primary Literature

1. Russell JS (1845) Report on Waves. Proc Roy Soc, Edinburgh, pp 421–445
2. 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 Serl 39(5):422–443
3. Zabusky NJ, Kruskal MD (1965) Interaction of Solitons in a Collisionless Plasma and the recurrence of initial States. Phys Rev Lett 15:240–243
4. 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. Springer Series in Nonlinear Dynamics. Springer, New York, Berlin, Heidelberg etc. pp 150–281
5. Scott AC, Chu FY, Mclaughlin DW (1973) The Soliton: A New Concept in Applied Science. Proc IEEE 61(10):1443–1483
6. Gardner CS, Greene JM, Kruskal MD, Miura RM (1967) Method for solving the Korteweg–de Vries equation. Phys Rev Lett 19:1095–1097
7. Gardner CS, Greene JM, Kruskal MD, Miura RM (1974) The Korteweg–de Vries equation and generalizations VI. Methods for exact solution. Commun Pure Appl Math 27:97–133
8. Lax PD (1968) Integrals of nonlinear equations of evolution and solitary waves. Commun. Pure Appl. Math 21:467–490
9. Gel'fand IM, Levitan BM (1951) On the determination of a differential equation from its spectral function. Izv Akad Nauk SSR Ser Math 15:309–360
10. Ablowitz MJ, Clarkson PA (1991) Solitons, Nonlinear Evolution Equations and Inverse Scattering. London Math Soc, vol 149, CUP, Cambridge
11. Fordy AP (1990) Soliton Theory: A Survey of Results. MUP, Manchester
12. Newell AC (1985) Solitons in Mathematics and Physics. SIAM, Philadelphia
13. Novikov SP, Manakov SV, Pitaevskii LP, Zakharov VE (1984) Theory of Solitons. Plenum, NY
14. Hirota R (1971) Exact soliton of the KdV equation for multiple collisions of solitons. Phys Rev Lett 27:1192–1194
15. Hirota R, Satsuma J (1977) Nonlinear evolution equation generated from the Bäcklund transformation for the Boussinesq equation. Prog Theor Phys Suppl 57:797–807
16. Hirota R (1980) Direct methods in soliton theory. In: Bullough RK, Caudrey PJ (ed) Topics of Modern Physics, vol 17. Springer, New York, pp 157–176
17. Wahlquist HD, Estabrook FB (1973) Bäcklund transformations for solitons of the Korteweg–de Vries equation. Phys Rev Lett 31:1386–1390
18. Matveev VB, Salle MA (1991) Darboux Transformations and Solitons. Springer, Berlin

Books and Reviews

Ablowitz MJ, Clarkson PA (1991) Solitons, Nonlinear Evolution Equations and Inverse Scattering. Cambridge University Press, New York

Ablowitz MJ, Segur H (1981) Solitons and the Inverse Scattering Transform. SIAM, USA

Drazin PG, Johnson RS (1989) Solitons: an Introduction. Cambridge University Press, New York

Konopelchenko BG (1993) Solitons in Multidimensions: Inverse Spectral Transform Method. World Scientific, Singapore

Rogers C, Shadwick WR (1982) Bäcklund Transformations and Their Application. Academic Press, New York

Whitham GB (1974) Linear and Nonlinear Waves. Wiley, New York

Korteweg–de Vries Equation (KdV) and Modified Korteweg–de Vries Equations (mKdV), Semi-analytical Methods for Solving the

DOĞAN KAYA[1,2]

[1] Department of Mathematics, Firat University, Elazig, Turkey

[2] Dean of Engineering Faculty, Ardahan University, Ardahan, Turkey

Article Outline

Glossary
Definition of the Subject
Introduction
An Analysis of the Semi-analytical Methods and Their Applications
Future Directions
Bibliography

Glossary

Korteweg–de Vries equation The classical nonlinear equations of interest usually admit for the existence of a special type of the traveling wave solutions, which are either solitary waves or solitons.

Modified Korteweg–de Vries This equation is a modified form of the classical KdV equation in the nonlinear term.

Soliton This concept can be regarded as solutions of nonlinear partial differential equations.

Exact solution A solution to a problem that contains the entire physics and mathematics of a problem, as opposed to one that is approximate, perturbative, closed, etc.

Adomian decomposition method, Homotopy analysis method, Homotopy perturbation method and Variational iteration method These are some of the semi-analytic/numerical methods for solving ODE or PDE in literature.

Definition of the Subject

In this study, some semi-analytical/numerical methods are applied to solve the Korteweg–de Vries (KdV) equation and the modified Korteweg–de Vries (mKdV) equation, which are characterized by the solitary wave solutions of the classical nonlinear equations that lead to solitons. Here, the classical nonlinear equations of interest usually admit for the existence of a special type of the traveling wave solutions which are either solitary waves or solitons. These approaches are based on the choice of a suitable differential operator which may be ordinary or partial, linear or nonlinear, deterministic or stochastic. It does not require discretization, and consequently massive computation.

In this scheme the solution is performed in the form of a convergent power series with easily computable components. This section is particularly concerned with the Adomian decomposition method (ADM) and the results obtained are compared to those obtained by the variational iteration method (VIM), homotopy analysis method (HAM), and homotopy perturbation method (HPM). Some numerical results of these particular equations are also obtained for the purpose of numerical comparisons of those considered approximate methods. The numerical results demonstrate that the ADM is relatively accurate and easily implemented.

Introduction

In this part, a certain nonlinear partial differential equation is introduced which is characterized by the solitary wave solutions of the classical nonlinear equations that lead to solitons [8,9,11,64,70]. The classical nonlinear equations of interest usually admit for the existence of a special type of the traveling wave solutions which are either solitary waves or solitons. In this study, a few solutions arising from the semi-analytical work on the Korteweg–de Vries (KdV) equation and modified Korteweg–de Vries (mKdV) equation will be reviewed.

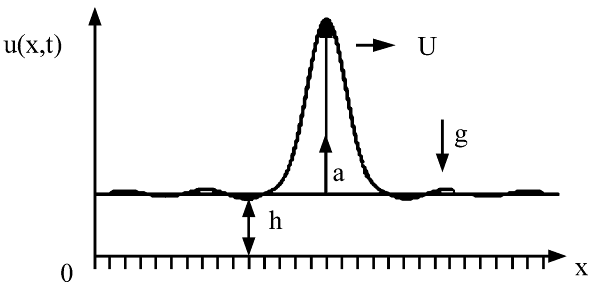
In this work, a brief history of the above mentioned nonlinear KdV equation is given and how this type of equation leads to the soliton solutions is then presented as an introduction to the theory of solitons. This theory is an important branch of applied mathematics and mathematical physics. In the last decade this topic has become an active and productive area of research, and applications of the soliton equations in physical cases have been considered. These have important applications in fluid mechanics, nonlinear optics, ion plasma, classical and quantum fields theories, etc.

The best introduction for this study may be the Scottish naval engineer J. Scott Russell's seminal 1844 report to the Royal Society. In the time of the eighteen century, Scott Russell's report was called "Report on Waves" [57]. Around forty years later, starting from Scott Russell's experimental observations in 1934, the theoretical scientific work of Lord Rayleigh and Joseph Boussinesq around 1870 [56] independently confirmed Russell's prediction and derived formula (1) from the equation of motion for an inviscid incompressible liquid. They also gave the solitary wave profile $z = u(x, t)$ as

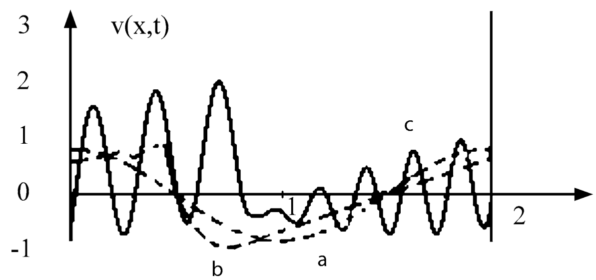
$$u(x, t) = a\ \mathrm{sech}^2\left[\beta(x - Ut)\right] \tag{1}$$

where $\beta^2 = 3a \div \left\{4h^2(h + a)\right\}$ any $a > 0$. They drew the solitary wave profile as in the following

Finally, two Dutch scientists, Korteweg and de Vries, developed a nonlinear partial differential equation to model the propagation of shallow water waves applicable to situation in 1895 [43]. This work was really what Scott Russell fortuitously witnessed. This famous classical equation is known simply as the KdV equation. Korteweg and



Korteweg–de Vries Equation (KdV) and Modified KdV (mKdV), Semi-analytical Methods for Solving the, Figure 1
A solitary wave



Korteweg–de Vries Equation (KdV) and Modified KdV (mKdV), Semi-analytical Methods for Solving the, Figure 2
Development of solitary waves: a initial profile at $t = 0$, b profile at $t = \pi^{-1}$, and c wave profile at $t = (3.6)\pi^{-1}$ [13]

de Vries published a theory of shallow water waves which reduced Russell's observations to its essential features. The nonlinear classical dispersive equation was formulated by Korteweg and de Vries in the form

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} + \varepsilon\frac{\partial^3 u}{\partial x^3} + \gamma\, u\frac{\partial u}{\partial x} = 0\,, \tag{2}$$

where $c,\ \varepsilon,\ \gamma$ are physical parameters. This equation now plays a key role in soliton theory.

After setting the theory and applications of the KdV solitary waves, scientists should know a numerical model for the such nonlinear equations. This model was constructed and published as a Los Alamos Scientific Laboratory Report in 1955 by Fermi, Pasta, and Ulam (FPU). This was a numerical model of a discrete nonlinear mass–spring system [13]. The natural question arises of how the energy would be distributed among all modes in this nonlinear discrete system. The energy would be distributed uniformly among all modes in accordance with the principle of the equipartition of energy stated by Fermi, Pasta and Ulam. This model follows the form of a coupled nonlinear ordinary differential equation system, where w_i is a function of t, and this system is written as

$$\begin{aligned}\frac{m}{k}\frac{\mathrm{d}^2 w_i}{\mathrm{d}t^2} &= (w_{i+1} + w_{i-1} - 2w_i)\\ &\quad + \alpha\left[(w_{i+1} - w_i)^2 - (w_i - w_{i-1})^2\right]\end{aligned} \tag{3}$$

where w_i is the displacement of the ith mass from the equilibrium position, $i = 1, 2, \ldots, n$, $w_0 = w_n = 0$, k is the linear spring constant and α (> 0) measures the strength of nonlinearity.

In the late 1960s, Zabusky and Kruskal numerically studied, and then analytically solved, the KdV equation [71] from the result of the FPU experiment inspiration. They came to the result that the stable pulse-like waves could exist in a problem described by the KdV equation from their numerical study. A remarkable aspect of the discovered solitary waves is that they retain their shapes and speeds after a collision. After the observation of Zabusky and Kruskal, they called to these waves "solitons", because the character of the waves had a particle-like nature. In fact, they considered the initial value problem for the KdV equation in the form

$$v_t + v v_x + \delta\, v_{\mathrm{xxx}} = 0\,, \tag{4}$$

where $\delta = (h/\ell)^2$, ℓ is a typical horizontal length scale, with the initial condition

$$v(x,0) = \cos\,\pi\,x, \qquad 0 \le x \le 2 \tag{5}$$

and the periodic boundary conditions with period 2, so that $v(x,t) = v(x+2,t)$ for all t. Their numerical study with $\sqrt{\delta} = 0.022$ produced a lot of new interesting results, which are shown as follows.

In 1967, this numerical development was placed on a settled mathematical basis with C.S. Gardner, J.M. Greene, M.D. Kruskal and R.M. Miura's discovery of the inverse-scattering-transform method [14,15]. This way of calculation was an ingenious method for finding the exact solution of the KdV equation. Almost simultaneously, Miura [51] and Miura et al. [52] formulated another ingenious method to derive an infinite set of conservation laws for the KdV equations by introducing the so-called Miura transformation. Subsequently, Hirota [28,29,30] constructed analytical solutions of the KdV equation which provide the description of the interaction among N solitons for any positive integral N.

The soliton concept can be regarded with solutions of nonlinear partial differential equations. The soliton solution of a nonlinear equation is usually used as single wave. If there are several soliton solutions, these solutions are called "solitons." On the other hand, if a soliton separates infinitely from other soliton, this soliton is a single wave. Besides, a single wave solution cannot be a sech^2 function for equations different from nonlinear equations, such as the KdV equation. But this solution can be sech or $\tan^{-1}(e^{\alpha x})$.

At this stage, one could ask what the definition of a soliton solution is. It is not easy to define the soliton concept. Wazwaz [64] describes solitons as solutions of nonlinear differential equations such that

1. A long and shallow water wave should not lose its permanent forms;
2. A long and shallow water wave of the solution is localized, which means that either the solutions decay exponentially to zero such as the solitons admitted by the KdV equation, or approach a constant at infinity such as the solitons provided by the SG equation;
3. A long and shallow water wave of the solution can interact with other solitons and still preserve its character.

There is also a more formal definition of this concept, but these definitions require substantial mathematics [12]. On the other hand, the phenomena of the solitons use not quite the same as have stated above three properties in all expression of the solutions. For example, the concept referred to as "light bullets" in nonlinear optics are often called solitons despite losing energy during interaction. This idea can be found at an internet web page of the Simon Fraser University British Columbia, Canada [50].

Nonlinear phenomena play a crucial role in applied mathematics and physics. Calculating exact and numerical solutions, particularly traveling wave solutions, of nonlinear equations in mathematical physics play an important role in soliton theory [8,70]. It has recently become more interesting to obtain exact solutions of nonlinear partial differential equations utilizing symbolical computer programs (such as Maple, Matlab, and Mathematica) which facilitate complex and tedious algebraical computations. It is also important to find exact solutions of nonlinear partial differential equations. These equations exist because of the mathematical models of complex physical phenomenon that arise in engineering, chemistry, biology, mechanics and physics. Various effective methods have been developed to understand the mechanisms of these physical models, to help physicians and engineers and to ensure knowledge for physical problems and its applications.

Many explicit exact methods have been introduced in literature [8,9,11,31,32,33,42,62,64,65,66,67,69,70]. Some of them are: Bäcklund transformation, Cole–Hopf transformation, Generalized Miura Transformation, Inverse Scattering method, Darboux transformation, Painleve method, similarity reduction method, tanh (and its variations) method, homogeneous balance method, Exp-function method, sine–cosine method and so on. There are also many numerical methods implemented for nonlinear equations [16,17,27,34,35,37,38,39,40,54,61]. Some of them are: finite difference methods, finite elements method, Sinc–Galerkin method and some approximate or semi-analytic/numerical methods such as Adomian decomposition method, variational analysis method, homotopy analysis method, homotopy perturbation method and so on.

An Analysis of the Semi-analytical Methods and Their Applications

The recent years have seen a significant development in the use of various methods to find the numerical and semi-analytical/numerical solution of a linear or nonlinear, deterministic or stochastic ODE or PDE. In this work, we will only discuss the nonlinear PDE case of the problem.

Before giving the semi-analytical and numerical implementations of the considered methods, may draw some conclusions from Liao's book [46] and the recent new papers [1,7,10] for comparisons of the ADM, HAM, HPM and VIM. ADM can be applied to solve linear or nonlinear ordinary and partial differential equations, no matter whether they contain small/large parameters, and thus is rather general. In some cases, the Adomian decomposition series converge rapidly. However, this method has some restrictions, e. g. if you are solving nonlinear equation you have to sort out Adomian's polynomials. In the same case, this needs to take more calculations. Liao [46] pointed out that "in general, convergence regions of power series are small, thus acceleration techniques are often needed to enlarge convergence regions. This is mainly due to the fact that a power series is often not an efficient set of base functions to approximate a nonlinear problem, but unfortunately ADM does not provide us with freedom to use different base functions. Like the artificial small parameter method and the d-expansion method, ADM itself also does not provide us with a convenient way to adjust the convergence region and the rate of approximation solutions".

Many authors [1,7,10] made their own studies in which they gave implementations of equations by using HAM, HPM, and then found numerical solutions. Their results show that: (i) Liao's HAM can produce much better approximations than the previous solutions for nonlinear differential equations. (ii) They compared the approximations of these two methods and they observe that although

HAM is faster than HPM (see Figs. 1–6 in [10]), both of the methods converge to the exact solution quite fast. HAM and HPM are in some cases similar to each other; e. g., the obtained solution function of the equation is shorter if HPM is used instead of HAM. The corresponding results converge more rapidly if HAM is used. (iii) The other advantage of the HAM is that it gives the flexibility to choose an auxiliary parameter h to adjust and control the convergence and its rate for the solutions series and the defined different functions which originate from the nature of the considered problems [1]. (iv) If one compares the approximated numerical solution with the corresponding exact solution by using HAM and HPM, one can see that when the small parameter h is increased the error of the first method is less than the second one.

Chowdhury and co worker [7], show that their obtained numerical results by the 5-term HAM are exactly the same as the ADM solutions and HPM solutions for the special case of the auxiliary parameter $h = -1$ and auxiliary function $H(x) = 1$. Because of this conclusion, they admitted that the HPM and the ADM is a special case of HAM.

Adomian Decomposition Method

The aim of the present section is to give an outline and implementation of the Adomian decomposition method (ADM) for nonlinear wave equations, and to obtain analytic and approximate solutions which are obtained in a rapidly convergent series with elegantly computable components by this method. The approach is based on the choice of a suitable differential operator which may be ordinary or partial, linear or nonlinear, deterministic or stochastic [4,5,6,63,64]. It allows one to obtain a decomposition series solution of the equation which is calculated in the form of a convergent power series with easily computable components.

The inhomogeneous problem is quickly solved by observing the self-canceling "noise" terms, where the sum of the components vanishes in the limit. Many tests which model problems from mathematical physics, linear and nonlinear, are discussed to illustrate the effectiveness and the performance of the ADM. Adomian and Rach [5] and Wazwaz [63] have investigated the phenomena of the self-canceling "noise" terms where the sum of the components vanishes in the limit. An important observation was made that the "noise" terms appear for nonhomogenous cases only. Further, it was formally justified that if terms in u_0 are canceled by terms in u_1, even though u_1 includes further terms, then the remaining non-canceled terms in u_1 constitute the exact solution of the equation.

ADM is valid for ordinary and partial differential equations, whether or not they contain small/large parameters, and thus is rather general. Moreover, the Adomian approximation series converge quickly. However, this method has some restrictions. Approximates solutions given by ADM often contain polynomials. In general, convergence regions of power series are small, thus acceleration techniques are often needed to enlarge convergence regions. This is mainly due to the fact that power series is often not an efficient set of base functions to approximate a nonlinear problem, but unfortunately ADM does not provide us with freedom to use different base functions.

An outline of the method will be given here in order to obtain analytic and approximate solutions by using the ADM. Considering a generalized KdV equation [41]

$$u_t + \alpha\, u^m u_x + u_{xxx} = 0\,, \quad u(x,0) = g(x)\,, \tag{6}$$

where α and $m > 0$ are constants. An operator form of this equation can be written as

$$L_t\,(u) + \alpha N u + L_{xxx}\,(u) = 0\,, \tag{7}$$

where $L_t \equiv \partial/(\partial t)$, $Nu = u^m u_x$ and $L_{xxx} \equiv \partial^3/(\partial x^3)$. It is assumed that L_t^{-1} is an integral operator given by $L_t^{-1} \equiv \int_0^t (.)\mathrm{d}t$. Operating with the integral operator L_t^{-1} on both sides of (7) with

$$L_t^{-1} L_t\,(u) = -\alpha L_t^{-1}\,(Nu) - L_t^{-1} L_{xxx}\,(u)\ . \tag{8}$$

Therefore, it follows that

$$u(x,t) = u(x,0) - \alpha L_t^{-1}\,(Nu) - L_t^{-1} L_{xxx}\,(u)\ .$$

We find that the zeroth component is obtained by

$$u_0 = u(x,0)\,, \tag{9}$$

which is defined by all terms that arise from the initial condition and from integrating the source term and decomposing the unknown function $u(x,t)$ as a sum of components defined by the decomposition series

$$u(x,t) = \sum_{n=0}^{\infty} u_n(x,t)\,. \tag{10}$$

The nonlinear term $u^m u_x$ can be decomposed into the infinite series of polynomials given by

$$Nu = u^m u_x = \sum_{n=0}^{\infty} A_n\,,$$

where the components $u_i(x,t)$ will be determined recurrently, and the A_n polynomials are the so-called Adomian polynomials [64] of $u_0, u_1, u_2, \cdots, u_i$ defined by

$$A_n = \frac{1}{n!}\left[\frac{d^n}{d\lambda^n}\Phi\left(\sum_{k=1}^{\infty}\lambda^k u_k\right)\right]\Bigg|_{\lambda=0}, \quad n \geq 0\,. \tag{11}$$

Substituting (11) and (9) into (8) gives rise to

$$u_{n+1} = -L_t^{-1}(A_n) - L_t^{-1}L_{xxx}(u_n), \quad n \geq 0 \tag{12}$$

where L_t^{-1} is the previously given integration operator. The solution $u(x,t)$ must satisfy the requirements imposed by the initial conditions. The decomposition method provides a reliable technique that requires less work if compared with the traditional techniques. To give a clear overview of the methodology, the following examples will be discussed.

Example 1 Let's consider the mKdV equation (6) for the value of $m = 1$, which is called the classical KdV equation, and take this equation with the following initial value condition [36],

$$u(x,0) = \frac{3\lambda}{\alpha} - \frac{3\lambda}{\alpha}\tanh^2\left(\frac{\sqrt{\lambda}}{2}x\right), \tag{13}$$

and then basically Eq. (6) is taken in an operator form exactly in the same manner as the form of the Eq. (6) and using (9) to find the zeroth component of $u_0 = u(x,0)$. Taking into consideration (12) with (11) and one can calculate the rest of the terms of the series with the aid of Mathematica and then write these terms to the Eq. (10), yielding

$$\begin{aligned}
&u(x,t) = \\
&\frac{1}{4\alpha}\left\{3t^2\lambda^4\left(-2+\cosh\left(x\sqrt{\lambda}\right)\right)\operatorname{sech}^4\left(\frac{x\sqrt{\lambda}}{2}\right)\right\} \\
&+\frac{1}{64\alpha}\left\{t^4\lambda^7\Bigg(-33-26\cosh\left(x\sqrt{\lambda}\right)\right. \\
&\left.+\cosh\left(2x\sqrt{\lambda}\right)\Bigg)\operatorname{sech}^6\left(\frac{x\sqrt{\lambda}}{2}\right)\right\} \\
&+\frac{1}{8\alpha}\left\{t^3\lambda^{\frac{11}{2}}\operatorname{sech}^5\left(\frac{x\sqrt{\lambda}}{2}\right)\Bigg(-11\sinh\left(\frac{x\sqrt{\lambda}}{2}\right)\right. \\
&\left.+\sinh\left(\frac{3x\sqrt{\lambda}}{2}\right)\Bigg)\right\} \\
&+\frac{1}{640\alpha}\left\{t^5\lambda^{\frac{7}{2}}\operatorname{sech}^7\left(\frac{x\sqrt{\lambda}}{2}\right)\Bigg(302\sinh\left(\frac{x\sqrt{\lambda}}{2}\right)\right. \\
&\left.-57\sinh\left(\frac{3x\sqrt{\lambda}}{2}\right)+\sinh\left(\frac{5x\sqrt{\lambda}}{2}\right)\Bigg)\right\} \\
&+\frac{1}{\alpha}\left\{9t\lambda^{\frac{5}{2}}\operatorname{sech}^2\left(\frac{x\sqrt{\lambda}}{2}\right)\tanh\left(\frac{x\sqrt{\lambda}}{2}\right)\right\} \\
&-\frac{1}{\alpha}\left\{6t\lambda^{\frac{5}{2}}\operatorname{sech}^4\left(\frac{x\sqrt{\lambda}}{2}\right)\tanh\left(\frac{x\sqrt{\lambda}}{2}\right)\right\} \\
&-\frac{1}{\alpha}\left\{6t\lambda^{\frac{5}{2}}\operatorname{sech}^2\left(\frac{x\sqrt{\lambda}}{2}\right)\tanh^3\left(\frac{x\sqrt{\lambda}}{2}\right)\right\} \\
&+\frac{1}{\alpha}\left\{3\lambda\left(1-\tanh^2\left(\frac{x\sqrt{\lambda}}{2}\right)\right)\right\}+\cdots
\end{aligned}$$

Example 2 Consider the mKdV equation (6) with the values of $m = 2$ and $\alpha = 6$, it has the traveling wave solution which can be obtained subject to the initial condition [41]

$$u(x,0) = \sqrt{c}\operatorname{sech}\left[k+\sqrt{c}x\right], \tag{14}$$

for all $c \geq 0$, where k is an arbitrary constant.

Again, to find the solution of this equation, simply take the equation in a operator form exactly in the same manner as the form of Eq. (6) and use (12) to find the zeroth component of $u_0 = u(x,0)$ and obtain sequential terms by using (12) with (11) to determine the other individual terms of the decomposition series with the aid of Mathematica, to get

$$\begin{aligned}
u(x,t) &= \sqrt{c}\operatorname{sech}\left(k+\sqrt{c}x\right) \\
&+\frac{1}{4}c^{\frac{7}{2}}t^2\left(-3+\cosh\left(2k+2\sqrt{c}x\right)\right)\operatorname{sech}^3\left(k+\sqrt{c}x\right) \\
&+\frac{1}{192}c^{\frac{13}{2}}t^4\Big(115-76\cosh\left(2k+2\sqrt{c}x\right) \\
&+\cosh\left(4k+4\sqrt{c}x\right)\Big)\operatorname{sech}^5\left(k+\sqrt{c}x\right) \\
&+\frac{1}{23040}\Big\{c^{\frac{19}{2}}t^6\Big(-11774+10543\cosh\left(2k+2\sqrt{c}x\right) \\
&-722\cosh\left(4k+4\sqrt{c}x\right)+\cosh\left(6k+6\sqrt{c}x\right)\Big) \\
&\times\operatorname{sech}^7\left(k+\sqrt{c}x\right)\Big\} \\
&+\frac{1}{4}c^5t^3\operatorname{sech}^4\left(k+\sqrt{c}x\right)\Big(-23\sinh\left(k+\sqrt{c}x\right) \\
&+\sinh\left(3k+3\sqrt{c}x\right)\Big)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{1920}\left\{c^{8}t^{5}\operatorname{sech}^{6}\left(k+\sqrt{c}x\right)\left(1682\sinh\left(k+\sqrt{c}x\right)\right.\right.\\
&\left.\left.-237\sinh\left(3k+3\sqrt{c}x\right)+\sinh\left(5k+5\sqrt{c}x\right)\right)\right\}\\
&+\frac{1}{322560}\left\{c^{11}t^{7}\operatorname{sech}^{8}\left(k+\sqrt{c}x\right)\right.\\
&\left(-259723\sinh\left(k+\sqrt{c}x\right)+60657\sinh\left(3k+3\sqrt{c}x\right)\right.\\
&\left.\left.-2179\sinh\left(5k+5\sqrt{c}x\right)+\sinh\left(7k+7\sqrt{c}x\right)\right)\right\}\\
&+c^{2}t\operatorname{sech}^{3}\left(k+\sqrt{c}x\right)\tanh\left(k+\sqrt{c}x\right)\\
&\qquad+c^{2}t\operatorname{sech}\left(k+\sqrt{c}x\right)\tanh^{3}\left(k+\sqrt{c}x\right)+\cdots
\end{aligned}
$$

Homotopy Analysis Method In this section, the homotopy analysis method (HAM) [44,46] is considered. HAM has been constructed and successfully implemented as an approximate and numerical solution for many types of nonlinear problems [18,19,45,46,47,48,49,58,59,60] and the references cited therein. A very nice explanation of the basic ideas of the HAM, its relationships with other analytic techniques, and some of its applications in science and engineering are given in Liao's book [46]. There are many groups of methods similar to HAM in the literature which are implemented nonlinear problems. Most of these groups of methods are in principle based on a Taylor series in an embedding parameter. If one could guess the initial function and the auxiliary linear operator well, then one can get very good approximations in a few terms, especially for a small value of the variable of the series.

For the purpose of illustration of the HAM [44], the mKdV equation is written in the operator form as

$$L(u)+\alpha\, u^{m}u_{x}+u_{xxx}=0\,, \tag{15}$$

where L is a linear operator: $L\equiv\partial/(\partial t)$. Equation (15) can be written in a nonlinear operator form as

$$
\begin{aligned}
N\left[\varphi\left(x,t;q\right)\right]&=\frac{\partial\varphi\left(x,t;q\right)}{\partial t}\\
&+\alpha\varphi^{m}\left(x,t;q\right)\frac{\partial\varphi\left(x,t;q\right)}{\partial x}+\frac{\partial^{3}\varphi\left(x,t;q\right)}{\partial x^{3}}\,,
\end{aligned}\tag{16}
$$

where $q\in[0,1]$ is an embedding parameter and $\varphi\left(x,t;q\right)$ is a function.

From $u(x,0)=U_0(x)$, $-\infty<x<\infty$, it is straightforward to express the solution u by a set of base functions

$$\{e_n(x)\,t^n,\ \ n\geq 0\}\,,$$

where $e_n(x)$ as a coefficient is a function with respect to x. This provides us with the so-called Rule of Solution Expression.

Following Liao's method [44,46], let $u(x,0)=U_0(x)$ indicate an initial guess of the exact solution u, $h\neq 0$, an auxiliary parameter, $H(x,t)\neq 0$ an auxiliary function. A zero-order deformation equation is constructed as

$$
\begin{aligned}
&(1-q)\,L\left[\varphi\left(x,t;q\right)-u_0(x,t)\right]\\
&\qquad=qhH(x,t)\,N\left[\varphi\left(x,t;q\right)\right]\,,
\end{aligned}\tag{17}
$$

with the initial condition

$$\varphi\left(x,0;q\right)=U_0(x)\,. \tag{18}$$

When $q=0$ and 1, the above equation has the solution

$$\varphi(x,t;0)=u_0(x,t) \tag{19}$$

and

$$\varphi(x,t;1)=u(x,t) \tag{20}$$

respectively.

Assume the auxiliary function $H(x,t)$ and the auxiliary parameter h are properly chosen so that $\varphi\left(x,t;q\right)$ can be expressed by the Taylor series

$$\varphi\left(x,t;q\right)=u_0(x,t)+\sum_{n=1}^{\infty}u_n(x,t)\,q^n\,, \tag{21}$$

where

$$u_n(x,t)=\left.\frac{1}{n!}\frac{\partial^n\varphi\left(x,t;q\right)}{\partial q^n}\right|_{q=0} \tag{22}$$

and that the above series is convergent at $q=1$. Equations (19) and (20) then yield

$$u(x,t)=u_0(x,t)+\sum_{n=1}^{\infty}u_n(x,t)\,. \tag{23}$$

For the sake of simplicity, define the vectors

$$\vec{u}_n(x,t)=\{u_0(x,t),u_1(x,t),\ldots,u_n(x,t)\} \tag{24}$$

differentiating the zero-order deformation Eq. (17) n times with respect to the embedding parameter q, then setting $q=0$, and finally dividing by $n!$, the nth-order deformation equation is written as

$$
\begin{aligned}
&L\left[u_n(x,t)-\chi_n u_{n-1}(x,t)\right]\\
&\qquad=hH(x,t)\,R_n\left[\vec{u}_{n-1}(x,t)\right]\,,
\end{aligned}\tag{25}
$$

where

$$R_n\left[\vec{u}_{n-1}(x,t)\right] = \frac{1}{(n-1)!}\left\{\frac{\partial^{n-1}}{\partial q^{n-1}} N\left[\sum_{m=0}^{\infty} u_m(x,t)\, q^m\right]\right\}\Bigg|_{q=0} \quad (26)$$

and

$$\chi_n = \begin{cases} 0\,, & n \le 1 \\ 1\,, & n > 1 \end{cases} \quad (27)$$

with the initial condition

$$u_n(x,0) = 0\,, \qquad n \ge 1\,. \quad (28)$$

Therefore the nth order approximation of $u(x,t)$ is given by

$$u(x,t) \approx u_0(x,t) + \sum_{m=1}^{N} u_m(x,t)\,. \quad (29)$$

Example 1 Let's consider the mKdV equation (6) for the value of $m = 1$, which is called the classical KdV equation, and take this equation with the following initial value condition

$$u(x,0) = u_0 = \frac{3\lambda}{\alpha}\left(1 - \tanh^2\left(\frac{\sqrt{\lambda}}{2}x\right)\right)\,. \quad (30)$$

All related formulae are the same as those given from (25),

$$R_n\left[u_{n-1}\right] = \frac{\partial u_{n-1}}{\partial t} + \alpha\sum_{i=0}^{n-1} u_i \frac{\partial u_{n-1-i}}{\partial x} + \frac{\partial^3 u_{n-1}}{\partial x^3}\,, \quad (31)$$

using (24) and (25) with (31) and the initial function (30), three terms of the series (29) can be calculated as

$$\begin{aligned} u(x,t) &= \frac{3\lambda}{\alpha}\left(1 - \tanh^2\left(\frac{\sqrt{\lambda}}{2}x\right)\right) \\ &- \frac{3ht\lambda^{\frac{5}{2}}\operatorname{sech}^2\left(\frac{x\sqrt{\lambda}}{2}\right)\tanh\left(\frac{x\sqrt{\lambda}}{2}\right)}{\alpha} \\ &+ \frac{1}{4\alpha}\Bigg\{3ht\lambda^{\frac{5}{2}}\operatorname{sech}^4\left(\frac{x\sqrt{\lambda}}{2}\right) \\ &\times\left(-2+\cosh\left(x\sqrt{\lambda}\right)\right) - 2(1+h)\sinh\left(x\sqrt{\lambda}\right)\Bigg\} \\ &+ \frac{1}{8\alpha}\Bigg\{ht\lambda^{\frac{5}{2}}\operatorname{sech}^2\left(\frac{x\sqrt{\lambda}}{2}\right)\Bigg\{12h(1+h)t\lambda^{\frac{3}{2}} \\ &\times\left(-2+\cosh\left(x\sqrt{\lambda}\right)\right)\operatorname{sech}^2\left(\frac{x\sqrt{\lambda}}{2}\right) \\ &- h^2t^2\lambda^3\operatorname{sech}^3\left(\frac{x\sqrt{\lambda}}{2}\right) \\ &\times\left(-11\sinh\left(\frac{x\sqrt{\lambda}}{2}\right) + \sinh\left(\frac{3x\sqrt{\lambda}}{2}\right)\right) \\ &-24(1+h)^2\tanh\left(\frac{x\sqrt{\lambda}}{2}\right)\Bigg\}\Bigg\} \\ &+ \frac{1}{64\alpha}\Bigg\{ht\lambda^{\frac{5}{2}}\operatorname{sech}^2\left(\frac{x\sqrt{\lambda}}{2}\right) \\ &\times\Bigg\{144h(1+h)^2t\lambda^{\frac{3}{2}}\left(-2+\cosh\left(x\sqrt{\lambda}\right)\right) \\ &\times\operatorname{sech}^2\left(\frac{x\sqrt{\lambda}}{2}\right) \\ &+ h^3t^3\lambda^{\frac{9}{2}}\left(33 - 26\cosh\left(x\sqrt{\lambda}\right) + \cosh\left(2x\sqrt{\lambda}\right)\right) \\ &\times\operatorname{sech}^4\left(\frac{x\sqrt{\lambda}}{2}\right) - 24h^2(1+h)t^2\lambda^3 \\ &\times\operatorname{sech}^3\left(\frac{x\sqrt{\lambda}}{2}\right) \\ &\times\left(-11\sinh\left(\frac{x\sqrt{\lambda}}{2}\right) + \sinh\left(\frac{3x\sqrt{\lambda}}{2}\right)\right) \\ &\qquad - 192(1+h)^3\tanh\left(\frac{x\sqrt{\lambda}}{2}\right)\Bigg\}\Bigg\} + \cdots \end{aligned}$$

Example 2 In this example, the mKdV equation (6) is considered for the values of $m = 2$ and $\alpha = 6$. It has the traveling wave solution which can be obtained subject to the initial condition

$$u(x,0) = \sqrt{c}\operatorname{sech}\left[k + \sqrt{c}x\right]\,. \quad (32)$$

For implementation of the HAM for this particular equation with the initial function (32), all related formulae are the same as those given from (15) and (25) with the following $R_n\left[u_{n-1}\right]$:

$$\begin{aligned} R_n\left[u_{n-1}\right] &= \frac{\partial u_{n-1}}{\partial t} \\ &+ \alpha\sum_{i=0}^{n-1}\sum_{j=0}^{i} u_j u_{i-j}\frac{\partial u_{n-1-i}}{\partial x} + \frac{\partial^3 u_{n-1}}{\partial x^3}\,. \end{aligned} \quad (33)$$

Using (24) and (25) with (33) by with the initial function (32), we have

$$
\begin{aligned}
u(x,t) &= \sqrt{c}\,\mathrm{sech}\left[k+\sqrt{c}x\right] \\
&- c^2 ht\,\mathrm{sech}\left[k+\sqrt{c}x\right]\tanh\left[k+\sqrt{c}x\right] \\
&+ \frac{1}{4}c^2 ht\,\mathrm{sech}\left[k+\sqrt{c}x\right] \\
&\times \left(2c^{\frac{3}{2}}ht\left(1-2\,\mathrm{sech}^2\left(k+\sqrt{c}x\right)\right)\right. \\
&\left. -4(1+h)\tanh\left(k+\sqrt{c}x\right)\right) \\
&+ \frac{1}{24}c^2 ht\,\mathrm{sech}\left(k+\sqrt{c}x\right) \\
&\times \left(12c^{\frac{3}{2}}h(1+h)t\left(-3+\cosh\left(2\left(k+\sqrt{c}x\right)\right)\right)\right) \\
&\times \mathrm{sech}^2\left(k+\sqrt{c}x\right) + c^3h^2t^2\,\mathrm{sech}^3\left(k+\sqrt{c}x\right) \\
&\times \left(23\sinh\left(k+\sqrt{c}x\right)-\sinh\left(3\left(k+\sqrt{c}x\right)\right)\right) \\
&\qquad - 24(1+h)^2\tanh\left(k+\sqrt{c}x\right)+\cdots
\end{aligned}
$$

which are the three terms of the approximate series solution (29).

Homotopy Perturbation Method

In 1998, Liao's early idea to construct the one-parameter family of equations was introduced by He's so-called "homotopy perturbation method" [20,21,23,24], which is only a special case of the homotopy analysis method [45]. But this homotopy perturbation method is almost the same as Liao's early one-parameter family equation which can be found in [44]. There are some minor differences, and detailed discussions and proofs can be seen in literature [20,21,23,24]. The authors in [20,21,23,24], have proved that these two methods lead to the same solutions of the considered equation for high-order approximations. This is also mainly because of the Taylor series properties. In conclusion of this, nothing is new in He's idea, except the new name "homotopy perturbation method" [58].

To illustrate HPM, consider the following nonlinear differential equation

$$A(u) - f(r) = 0\,, \quad r \in \Omega \tag{34}$$

with the boundary condition

$$B\left(u, \frac{\partial u}{\partial n}\right) = 0\,, \quad r \in \Gamma\,, \tag{35}$$

where $A(u)$ is written as follows:

$$A(u) = L(u) + N(u)\,. \tag{36}$$

A is a general differential operator, B is a boundary operator, $f(r)$ is a given analytic right hand side function, and Γ is the boundary of the domain Ω. The operator A can be generally divided into two parts L and N, where L is linear and N is the nonlinear term. So, Eq. (34) can be rewritten as follows:

$$L(u) + N(u) - f(r) = 0\,. \tag{37}$$

By the homotopy technique [45], a homotopy $v(r,p) : \Omega \times [0,1] \to \Re$ is obtained which satisfies

$$
\begin{aligned}
&\mathrm{H}(v,p) \\
&\quad = (1-p)\left[L(v) - L(u_0)\right] + p\left[A(v) - f(r)\right] \\
&\qquad = 0\,, \quad p \in [0,1]\,, \quad r \in \Omega\,,
\end{aligned} \tag{38}
$$

where $p \in [0,1]$ is an embedding parameter, and u_0 is an initial approximation of Eq. (34) which satisfies the boundary conditions. Obviously, from (38) the result will be

$$
\begin{aligned}
H(v,0) &= L(v) - L(u_0) &= 0\,, \\
H(v,1) &= A(v) - f(r) &= 0\,,
\end{aligned}
$$

changing process of p from zero to unity is just that of $v(r,p)$ from $u_0(r)$ to $u(r)$. In topology, this is called deformation, and $L(v) - L(u_0)$ and $A(v) - f(r)$ are called homotopy.

We consider v as the following:

$$v = v_0 + pv_1 + p^2v_2 + p^3v_3 + \cdots\,. \tag{39}$$

According to HPM, the best approximation solution of Eq. (37) can be explained as a series of the power of p,

$$u = \lim_{p\to 1} v = v_0 + v_1 + v_2 + v_3 + \cdots\,. \tag{40}$$

The above convergence is given in [21]. Some results have been discussed in [20,23,24].

Example 1 Here, the mKdV equation (6) is again considered for the value of $m = 1$, which is called classical KdV equation; first we construct a homotopy as follows:

$$(1-p)\left[\dot{Y} - \dot{u_0}\right] + p\left[\dot{Y} + \alpha YY' + Y'''\right] = 0\,, \tag{41}$$

where

$$\dot{Y} = \frac{\partial Y}{\partial t}\,,\ Y' = \frac{\partial Y}{\partial x}\,,\ Y''' = \frac{\partial^3 Y}{\partial x^3} \quad \text{and} \quad p \in [0,1]\,.$$

With the initial approximation

$$Y_0 = u_0 = \frac{3\lambda}{\alpha} - \frac{3\lambda}{\alpha}\tanh^2\left(\frac{\sqrt{\lambda}}{2}x\right)\,,$$

suppose the solution of Eq. (41) has the form:

$$Y = Y_0 + pY_1 + p^2 Y_2 + p^3 Y_3 + \ldots = \sum_{n=0}^{\infty} p^n Y_n(x,t)\,. \quad (42)$$

Substituting Eq. (42) into Eq. (41) and arranging the coefficients of "p" powers, we have

$$p^0: \dot{Y}_0 - \dot{u}_0 = 0\,, \quad (43)$$

$$p^1: \dot{Y}_1 + \dot{u}_0 + \alpha Y_0 Y_0' + Y_0''' = 0\,, \quad (44)$$

$$p^2: \dot{Y}_2 + \alpha Y_0 Y_1' + \alpha Y_1 Y_0' + Y_1''' = 0\,, \quad (45)$$

$$p^3: \dot{Y}_3 + \alpha Y_0 Y_2' + \alpha Y_1 Y_1' + \alpha Y_2 Y_0' + Y_2''' = 0\,, \quad (46)$$

$$\vdots$$

and finally using Mathematica, the solutions of the equation can be obtained as follows:

$$Y_0 = \frac{3\lambda}{\alpha} - \frac{3\lambda}{\alpha}\tanh^2\left(\frac{\sqrt{\lambda}}{2}x\right), \quad (47)$$

$$Y_1 = \frac{3t\lambda^{\frac{5}{2}}}{\alpha}\operatorname{sech}^2\left(\frac{x\sqrt{\lambda}}{2}\right)\tanh\left(\frac{x\sqrt{\lambda}}{2}\right) \quad (48)$$

$$Y_2 = \frac{3t^2\lambda^4}{4\alpha}\left(-2+\cosh\left(x\sqrt{\lambda}\right)\right)\operatorname{sech}^4\left(\frac{x\sqrt{\lambda}}{2}\right), \quad (49)$$

$$Y_3 = \frac{3t^3\lambda^{\frac{11}{2}}}{8\alpha}\operatorname{sech}^5\left(\frac{x\sqrt{\lambda}}{2}\right) \times \left(-11\sinh\left(\frac{x\sqrt{\lambda}}{2}\right) + \sinh\left(\frac{3x\sqrt{\lambda}}{2}\right)\right) \quad (50)$$

$$\vdots$$

The above terms of the series (42) could be calculated. When the series (42) is considered with the terms (43)–(46) and supposing $p = 1$, an approximate series solution of the considered KdV equation is obtained as follows:

$$\begin{aligned} u(x,t) &= \frac{3\lambda}{\alpha} - \frac{3\lambda}{\alpha}\tanh^2\left(\frac{\sqrt{\lambda}}{2}x\right) \\ &+ \frac{3t\lambda^{\frac{5}{2}}}{\alpha}\operatorname{sech}^2\left(\frac{x\sqrt{\lambda}}{2}\right)\tanh\left(\frac{x\sqrt{\lambda}}{2}\right) \\ &+ \frac{3t^2\lambda^4}{4\alpha}\left(-2+\cosh\left(x\sqrt{\lambda}\right)\right)\operatorname{sech}^4\left(\frac{x\sqrt{\lambda}}{2}\right) \\ &+ \frac{3t^3\lambda^{\frac{11}{2}}}{8\alpha}\operatorname{sech}^5\left(\frac{x\sqrt{\lambda}}{2}\right) \\ &\quad \left(-11\sinh\left(\frac{x\sqrt{\lambda}}{2}\right) + \sinh\left(\frac{3x\sqrt{\lambda}}{2}\right)\right) + \cdots \end{aligned}$$

Example 2 Let's consider the mKdV equation (6) with the values of $m = 2$ and $\alpha = 6$. For this method's implementation, a homotopy is constructed as follows:

$$(1-p)\left[\dot{Y} - \dot{u}_0\right] + p\left[\dot{Y} + 6Y^2Y' + Y'''\right] = 0\,. \quad (51)$$

With the initial approximation

$$Y_0 = u_0 = \sqrt{c}\operatorname{sech}\left[k + \sqrt{c}x\right],$$

suppose the solution of Eq. (41) has the form:

$$Y = Y_0 + pY_1 + p^2 Y_2 + p^3 Y_3 + \ldots = \sum_{n=0}^{\infty} p^n Y_n(x,t)\,. \quad (52)$$

Substituting Eq. (52) into Eq. (51) and arranging the coefficients of "p" powers, we have

$$p^0: \dot{Y}_0 - \dot{u}_0 = 0\,, \quad (53)$$

$$p^1: \dot{Y}_1 + \dot{u}_0 + 6Y_0^2 Y_0' + Y_0''' = 0\,, \quad (54)$$

$$p^2: \dot{Y}_2 + 12 Y_0 Y_1 Y_0' + 6Y_0^2 Y_1' + Y_1''' = 0\,, \quad (55)$$

$$p^3: \dot{Y}_3 + 12 Y_0 Y_2 Y_0' + 6 Y_1^2 Y_0' + 12 Y_0 Y_1 Y_1' + Y_2''' = 0 \quad (56)$$

$$\vdots$$

Finally, using Mathematica, the few terms of the series solution of the Eq. (51) can be obtained as follows:

$$\begin{aligned} u(x,t) &= \sqrt{c}\operatorname{sech}\left(k+\sqrt{c}x\right) \\ &+ \frac{1}{4}c^{\frac{7}{2}}t^2\left(-3+\cosh\left(2\left(k+\sqrt{c}x\right)\right)\right) \\ &\times \operatorname{sech}^3\left(k+\sqrt{c}x\right) \\ &+ \frac{1}{24}c^5t^3\operatorname{sech}^4\left(k+\sqrt{c}x\right) \\ &\times \left(-23\sinh\left(k+\sqrt{c}x\right) + \sinh\left(3\left(k+\sqrt{c}x\right)\right)\right) \\ &\quad + c^2t\operatorname{sech}\left(k+\sqrt{c}x\right)\tanh\left(k+\sqrt{c}x\right) + \cdots \end{aligned}$$

Variational Iteration Method

In this section, a kind of semi-analytical technique is considered for a non-linear problem called the variational iteration method (VIM) [20,22,25]. It is proposed by He [20], and the method is based on the use of restricted variations and correction functionals. In this technique, a correction functional is defined by a general Lagrange multiplier using the ingenuity of variational theory. This method is implemented to get approximate solutions for many linear and non-linear problems in physics and mathematics [2,3,26,53,55,68]. When the method is implemented on the differential equation, it does not require the presence of small parameters and the solution of the equation is obtained as a sequence of iterates. The method does not require that the nonlinearities be differentiable with respect to the dependent variable and its derivatives [2,3,26,53,55,68].

To illustrate the VIM, let's consider the nonlinear KdV equation (6):

$$Lu + Nu = 0\,, \tag{57}$$

where L and N are linear and nonlinear operators, respectively. In [2,3,26,53,55,68], He proposed the VIM where a correction functional for Eq. (57) can be written as:

$$\begin{aligned} &u_{n+1}(x,t) \\ &= u_n(x,t) + \int_0^t \lambda\left\{Lu_n(x,\tau) + \tilde{N}u_n(x,\tau)\right\} d\tau\,, \\ &\qquad n \geq 0\,, \end{aligned} \tag{58}$$

where λ is a general Lagrange multiplier [20], which can be identified optimally via the variational theory, and $\tilde{u}_n$ is a restricted variation, which means $\delta\tilde{u}_n = 0$. It is required first to determine the Lagrangian multiplier λ that will be identified optimally via integration by parts. The successive approximations $u_{n+1}(x,t)$, $n \geq 0$, of the solution $u(x,t)$ will be readily obtained upon using the Lagrangian multiplier obtained and by using any selective function u_0. The initial values $u(x,0)$ and $u_t(x,0)$ are usually used for the selective zeroth approximation u_0. Having λ determined, then several approximations $u_j(x,t)$, $j \geq 0$, can be determined. Consequently, the solution is given by

$$u = \lim_{n\to\infty} u_n\,. \tag{59}$$

In what follows, the VIM will apply to two nonlinear KdV equations with similar initial functions to illustrate the strength of the method and to compare the given above methods.

Example 1 Let's consider the mKdV equation (6) for the value of $m = 1$ with the following initial function

$$u(x,0) = \frac{3\lambda}{\alpha} - \frac{3\lambda}{\alpha}\tanh^2\left(\frac{\sqrt{\lambda}}{2}x\right)\,, \tag{60}$$

where λ, α are arbitrary constant. The correction functional for this equation is

$$\begin{aligned} &u_{n+1}(x,t) \\ &= u_n(x,t) + \int_0^t \lambda\Big\{\big(u_n(x,\tau)\big)_\tau + \frac{\alpha}{2}\big(\tilde{u}_n^2(x,\tau)\big)_x \\ &\qquad + \big(u_n(x,\tau)\big)_{xxx}\Big\} d\tau\,, \quad n \geq 0\,, \end{aligned} \tag{61}$$

where λ is the general Lagrange multiplier [20] whose optimal value is found using variational theory. u_0 is an initial solution with or without unknown parameters. In the case of no unknown parameters, u_0 should satisfy initial-boundary conditions and $\tilde{u}_n^2$ is the restricted variation [22], i. e., $\tilde{u}_n^2 = 0$. To find the optimal value of λ, we have

$$\begin{aligned} &\delta u_{n+1}(x,t) \\ &= \delta u_n(x,t) + \delta\int_0^t \lambda\Big\{\big(u_n(x,\tau)\big)_\tau + \frac{\alpha}{2}\big(\tilde{u}_n^2(x,\tau)\big)_x \\ &\qquad + \big(u_n(x,\tau)\big)_{xxx}\Big\} d\tau\,, \end{aligned} \tag{62}$$

and this yields the stationary conditions

$$\lambda'(\tau) = 0\,, \quad 1 + \lambda(\tau)|_{\tau=t} = 0\,.$$

This in turn gives $\lambda = -1$ therefore, Eq. (51) can be written as following

$$\begin{aligned} &u_{n+1}(x,t) \\ &= u_n(x,t) - \int_0^t \Big\{\big(u_n(x,\tau)\big)_\tau + \frac{\alpha}{2}\big(\tilde{u}_n^2(x,\tau)\big)_x \\ &\qquad + \big(u_n(x,\tau)\big)_{xxx}\Big\} d\tau\,. \end{aligned} \tag{63}$$

Substituting the value of $m = 1$ into the mKdV equation (6) with an initial value (60), and then substituting this into Eq. (63) and using Mathematica, the solutions of the Eq. (6) with initial value (60) can be obtained as follows:

$$u(x,t) = \frac{1}{560\alpha\left(1 + \cosh\left(x\sqrt{\lambda}\right)\right)}$$

$$
\begin{aligned}
&\times \left\{105\lambda\left\{32 + t^4\lambda^6 \operatorname{sech}^2\left(\frac{x\sqrt{\lambda}}{2}\right)\right.\right.\\
&\times \left\{88 + 4\left(-149 + 6t^2\lambda^3\right)\operatorname{sech}^2\left(\frac{x\sqrt{\lambda}}{2}\right)\right.\\
&+ \left(1095 - 124t^2\lambda^3\right)\operatorname{sech}^4\left(\frac{x\sqrt{\lambda}}{2}\right)\\
&+ 198\left(-3 + t^2\lambda^3\right)\operatorname{sech}^6\left(\frac{x\sqrt{\lambda}}{2}\right)\\
&\left.\left.-99t^2\lambda^3\operatorname{sech}^8\left(\frac{x\sqrt{\lambda}}{2}\right)\right\}\right\} - 4t\lambda^{\frac{5}{2}}\left\{-140\left(6 + t^2\lambda^3\right)\right.\\
&+ 9t^2\lambda^3\operatorname{sech}^2\left(\frac{x\sqrt{\lambda}}{2}\right)\left\{140 - 141t^2\lambda^3\right.\\
&- 35\left(4 + 3t^2\lambda^3\right)\operatorname{sech}^2\left(\frac{x\sqrt{\lambda}}{2}\right)\\
&+ \left(497t^2\lambda^3 - 80t^4\lambda^6\right)\operatorname{sech}^4\left(\frac{x\sqrt{\lambda}}{2}\right)\\
&+ 20t^2\lambda^3\left(-21 + 20t^2\lambda^3\right)\operatorname{sech}^6\left(\frac{x\sqrt{\lambda}}{2}\right)\\
&- 630t^4\lambda^6\operatorname{sech}^8\left(\frac{x\sqrt{\lambda}}{2}\right)\\
&\left.\left.+ 315t^4\lambda^6\operatorname{sech}^{10}\left(\frac{x\sqrt{\lambda}}{2}\right)\right\}\right\}\\
&\left.\times \tanh\left(\frac{x\sqrt{\lambda}}{2}\right)\right\} + \cdots
\end{aligned}
$$

Example 2 Again the mKdV equation (6) is considered with the values of $m = 2$ and $\alpha = 6$ and the following initial function

$$u(x, 0) = \sqrt{c}\operatorname{sech}\left[k + \sqrt{c}x\right]. \tag{64}$$

The correction functional for this equation is

$$
\begin{aligned}
&u_{n+1}(x, t)\\
&= u_n(x, t) - \int_0^t \left\{\left(u_n(x, \tau)\right)_\tau + 2\left(\tilde{u}_n^3(x, \tau)\right)_x\right.\\
&\qquad \left. + \left(u_n(x, \tau)\right)_{xxx}\right\}\mathrm{d}\tau, \quad n \geq 0.
\end{aligned}\tag{65}
$$

Substituting the value of $m = 2$ into the mKdV equation (6) with initial value (64), and then substituting this into Eq. (65) and using Mathematica, the solutions of the considered equation with initial value (64) can be obtained as follows:

$$
\begin{aligned}
u(x, t) &= \frac{1}{4}\operatorname{sech}\left(k + \sqrt{c}x\right)\\
&\times \left\{2c^{\frac{7}{2}}t^2\left(1 - 2\operatorname{sech}^2\left(k + \sqrt{c}x\right)\right)\right.\\
&+ 2c^5t^3\operatorname{sech}^5\left(k + \sqrt{c}x\right)\\
&\times \left(-17\sinh\left(k + \sqrt{c}x\right) + 3\sinh\left(3\left(k + \sqrt{c}x\right)\right)\right)\\
&+ 3c^{\frac{13}{2}}t^4\left(-3 + \cosh\left(2\left(k + \sqrt{c}x\right)\right)\right)\\
&\times \operatorname{sech}^4\left(k + \sqrt{c}x\right)\tanh^2\left(k + \sqrt{c}x\right)\\
&\left. + 4\left(\sqrt{c} + c^2 t\tanh\left(k + \sqrt{c}x\right)\right)\right\} + \cdots
\end{aligned}
$$

We have just written the series solution for the mKdV as three terms rather than four terms.

Numerical Experiments For numerical comparisons purposes, two KdV equations are considered. The first one is the classical KdV and the second one is the modified KdV equation. The formula of numerical results for ADM, HAM and HPM are given as follows

$$\lim_{n\to\infty}\phi_n = u(x, t)$$

$$\text{where}\quad \phi_n(x, t) = \sum_{k=0}^{n} u_k(x, t), \quad n \geq 0. \tag{66}$$

The formula of numerical results for VIM is as (59). The recurrence relations of the methods are given as in (12), (25), (43)–(46) and (58), respectively.

Moreover, the decomposition series of the KdV equation's solutions generally converge very rapidly in real physical problems [5,6,64]. Results were obtained about the speed of convergence of this method, providing us methods to solve linear and nonlinear functional equations. The convergence of the HAM and VIM are numerically shown in some works [3,55] and references there in. Here, how all these methods are converged to their corresponding exact solutions has been proved.

Numerical approximations show a high degree of accuracy and in most cases of ϕ_n, the n-term approximation is accurate for quite low values of n. The solutions are very rapidly convergent by utilizing the ADM, VIM, HPM and HAM. The obtained numerical results justify the advantage of these methodologies, and even with few terms the approximation is accurate. Furthermore, as the all these methods do not require discretization of the variables, i. e.

Korteweg–de Vries Equation (KdV) and Modified KdV (mKdV), Semi-analytical Methods for Solving the, Table 1
Comparison between the absolute error of the solution of KdV equation (6) ($m = 1$) at various values of t and $x = 0.5$ for $\alpha = 1$, $\lambda = 0.01$ in ADM, VIM and HAM ($h = -1$)

t	$x = 0$			$x = 5$		
	ADM	HAM	VIM	ADM	HAM	VIM
0.1	2.9925×10^{-8}	2.9925×10^{-8}	$3. \times 10^{-8}$	0.0000130999	0.0000145274	0.0000145274
0.2	1.197×10^{-7}	1.197×10^{-7}	1.2×10^{-7}	0.0000261535	0.0000291008	0.000029101
0.3	2.69323×10^{-7}	2.69323×10^{-7}	2.69998×10^{-7}	0.0000391608	0.0000437201	0.0000437206
0.4	4.78795×10^{-7}	4.78795×10^{-7}	4.79995×10^{-7}	0.0000521216	0.0000583853	0.0000583862
0.5	7.48113×10^{-7}	7.48113×10^{-7}	7.49987×10^{-7}	0.000065036	0.0000730962	0.0000730976
0.6	1.07727×10^{-6}	1.07727×10^{-6}	1.07997×10^{-6}	0.0000779036	0.0000878527	0.0000878548
0.7	1.46628×10^{-6}	1.46628×10^{-6}	1.46995×10^{-6}	0.0000907246	0.000102655	0.000102658
0.8	1.91512×10^{-6}	1.91512×10^{-6}	1.91992×10^{-6}	0.000103499	0.000117502	0.000117506
0.9	2.42379×10^{-6}	2.42379×10^{-6}	2.42987×10^{-6}	0.000116226	0.000132395	0.0001324
1.	2.9923×10^{-6}	2.9923×10^{-6}	2.9998×10^{-6}	0.000128906	0.000147333	0.000147339

t	$x = 10$			$x = 15$		
	ADM	HAM	VIM	ADM	HAM	VIM
0.1	0.0000207071	0.0000229046	0.0000229046	0.0000216022	0.0000238682	0.0000238682
0.2	0.0000413971	0.000045826	0.0000458261	0.0000432119	0.0000477289	0.0000477289
0.3	0.0000620701	0.000068764	0.0000687642	0.0000648289	0.0000715819	0.0000715819
0.4	0.0000827257	0.0000917187	0.000091719	0.0000864532	0.0000954272	0.0000954271
0.5	0.000103364	0.00011469	0.00011469	0.000108085	0.000119265	0.000119265
0.6	0.000123985	0.000137677	0.000137678	0.000129723	0.000143094	0.000143094
0.7	0.000144588	0.000160681	0.000160682	0.000151369	0.000166916	0.000166916
0.8	0.000165173	0.0001837	0.000183702	0.000173022	0.00019073	0.000190729
0.9	0.000185741	0.000206736	0.000206738	0.000194682	0.000214535	0.000214535
1.	0.00020629	0.000229787	0.00022979	0.000216348	0.000238333	0.000238332

Korteweg–de Vries Equation (KdV) and Modified KdV (mKdV), Semi-analytical Methods for Solving the, Table 2
Comparison between the absolute error of the solution for Eq. (6) ($m = 1$) by HAM and HPM at various values of t and $x = 0.5$ with different values of h for HAM, in fact HPM is the value of $h = -1$ for HAM

t	HPM	HAM ($h = -1$)	HAM ($h = -1.4$)	HAM ($h = -0.9$)	HAM ($h = -0.6$)
0.1	1.60354×10^{-6}	1.60354×10^{-6}	1.60838×10^{-6}	1.60346×10^{-6}	1.59877×10^{-6}
0.2	3.26676×10^{-6}	3.26676×10^{-6}	3.27654×10^{-6}	3.26662×10^{-6}	3.25727×10^{-6}
0.3	4.98966×10^{-6}	4.98966×10^{-6}	5.00446×10^{-6}	4.98946×10^{-6}	4.97551×10^{-6}
0.4	6.77222×10^{-6}	6.77222×10^{-6}	6.79214×10^{-6}	6.77196×10^{-6}	6.75346×10^{-6}
0.5	8.61444×10^{-6}	8.61444×10^{-6}	8.63955×10^{-6}	8.61411×10^{-6}	8.59111×10^{-6}
0.6	0.0000105163	0.0000105163	0.0000105467	0.0000105159	0.0000104885
0.7	0.0000124777	0.0000124777	0.0000125135	0.0000124773	0.0000124455
0.8	0.0000144988	0.0000144988	0.0000145401	0.0000144984	0.0000144621
0.9	0.0000165795	0.0000165795	0.0000166263	0.000016579	0.0000165384
1.	0.0000187197	0.0000187197	0.0000187722	0.0000187192	0.0000186744

time and space, it is not effected by computation round off errors and one is not faced with the necessity of large computer memory and time.

In order to verify numerically whether the proposed methodologies lead to higher accuracy, the numerical solutions can be evaluated using the n-term approximation (66) and (59). The numerical results of the Eq. (6) (for KdV $m = 1$ and for mKdV $m = 2$) for the various values of x and t are illustrated in Tables 1 and 3 KdV and mKdV, respectively. These tabulated results show the differences

Korteweg–de Vries Equation (KdV) and Modified KdV (mKdV), Semi-analytical Methods for Solving the, Table 3
Comparison between the absolute error of the solution of Eq. (6) ($m = 2$) for $c = 1,\ k = -7$

$\lvert u_{\text{Exact}} - \varphi_3\rvert$ for ADM					
$t\lvert x$	0.1	0.2	0.3	0.4	0.5
0.1	8.23236×10^{-9}	1.29161×10^{-7}	6.41349×10^{-7}	1.98865×10^{-6}	4.76447×10^{-6}
0.2	9.098×10^{-9}	1.42742×10^{-7}	7.08789×10^{-7}	2.19776×10^{-6}	5.26547×10^{-6}
0.3	1.00546×10^{-8}	1.57752×10^{-7}	7.83317×10^{-7}	2.42885×10^{-6}	5.81914×10^{-6}
0.4	1.11118×10^{-8}	1.74338×10^{-7}	8.65678×10^{-7}	2.68424×10^{-6}	6.43099×10^{-6}
0.5	1.228×10^{-8}	1.92667×10^{-7}	9.56694×10^{-7}	2.96645×10^{-6}	7.10715×10^{-6}

$\lvert u_{\text{Exact}} - u_3\rvert$ for VIM					
$t\lvert x$	0.1	0.2	0.3	0.4	0.5
0.1	8.19516×10^{-9}	1.2858×10^{-7}	6.38476×10^{-7}	1.97978×10^{-6}	4.74334×10^{-6}
0.2	9.04779×10^{-9}	1.41958×10^{-7}	7.0491×10^{-7}	2.18579×10^{-6}	5.23696×10^{-6}
0.3	9.98686×10^{-9}	1.56692×10^{-7}	7.78082×10^{-7}	2.4127×10^{-6}	5.78065×10^{-6}
0.4	1.10203×10^{-8}	1.72908×10^{-7}	8.58611×10^{-7}	2.66243×10^{-6}	6.37904×10^{-6}
0.5	1.21566×10^{-8}	1.90738×10^{-7}	9.47155×10^{-7}	2.93702×10^{-6}	7.03703×10^{-6}

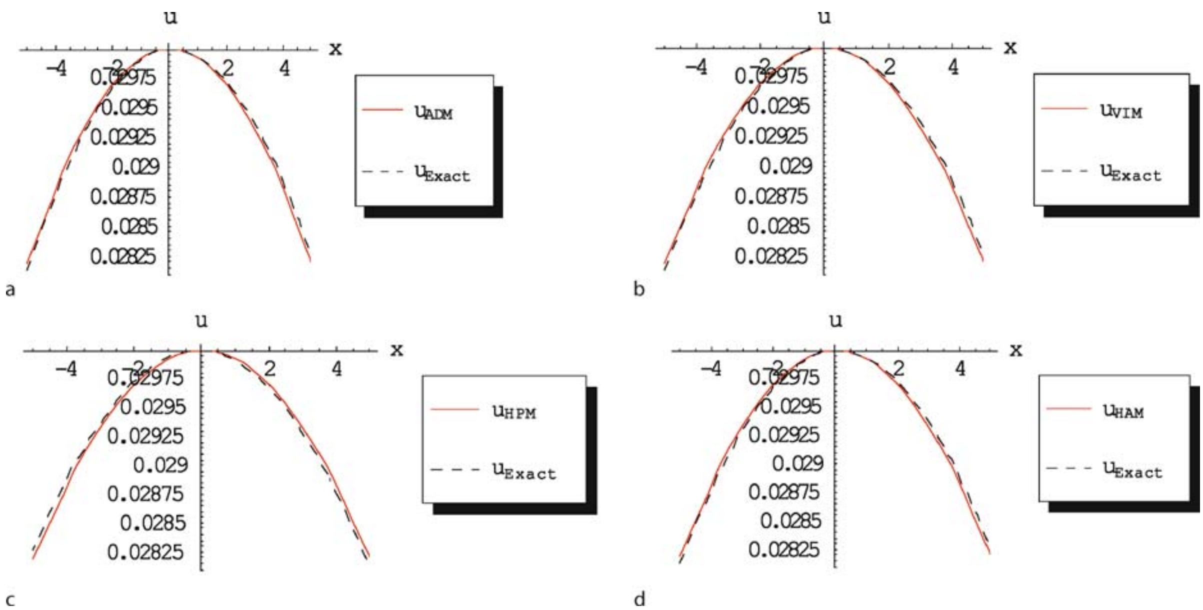
$\lvert u_{\text{Exact}} - \varphi_3\rvert$ for HPM $\equiv$ HAM with value of $h = -1$					
$t\lvert x$	0.1	0.2	0.3	0.4	0.5
0.1	8.19516×10^{-9}	1.2858×10^{-7}	6.38476×10^{-7}	1.97978×10^{-6}	4.74334×10^{-6}
0.2	9.04779×10^{-9}	1.41958×10^{-7}	7.0491×10^{-7}	2.18579×10^{-6}	5.23696×10^{-6}
0.3	9.98686×10^{-9}	1.56692×10^{-7}	7.78082×10^{-7}	2.4127×10^{-6}	5.78065×10^{-6}
0.4	1.10203×10^{-8}	1.72908×10^{-7}	8.58611×10^{-7}	2.66243×10^{-6}	6.37904×10^{-6}
0.5	1.21566×10^{-8}	1.90738×10^{-7}	9.47155×10^{-7}	2.93702×10^{-6}	7.03703×10^{-6}

$\lvert u_{\text{Exact}} - \varphi_3\rvert$ for HAM with value of $h = -0.9$					
$t\lvert x$	0.1	0.2	0.3	0.4	0.5
0.1	2.1809×10^{-9}	1.26484×10^{-7}	1.18355×10^{-7}	1.28924×10^{-7}	5.68147×10^{-7}
0.2	2.41029×10^{-9}	1.39787×10^{-7}	1.30802×10^{-7}	1.42489×10^{-7}	6.27919×10^{-7}
0.3	2.66382×10^{-9}	1.54489×10^{-7}	1.44558×10^{-7}	1.57481×10^{-7}	6.93984×10^{-7}
0.4	2.94403×10^{-9}	1.70737×10^{-7}	1.59761×10^{-7}	1.74053×10^{-7}	7.67005×10^{-7}
0.5	3.25373×10^{-9}	1.88694×10^{-7}	1.76562×10^{-7}	1.9237×10^{-7}	8.47719×10^{-7}

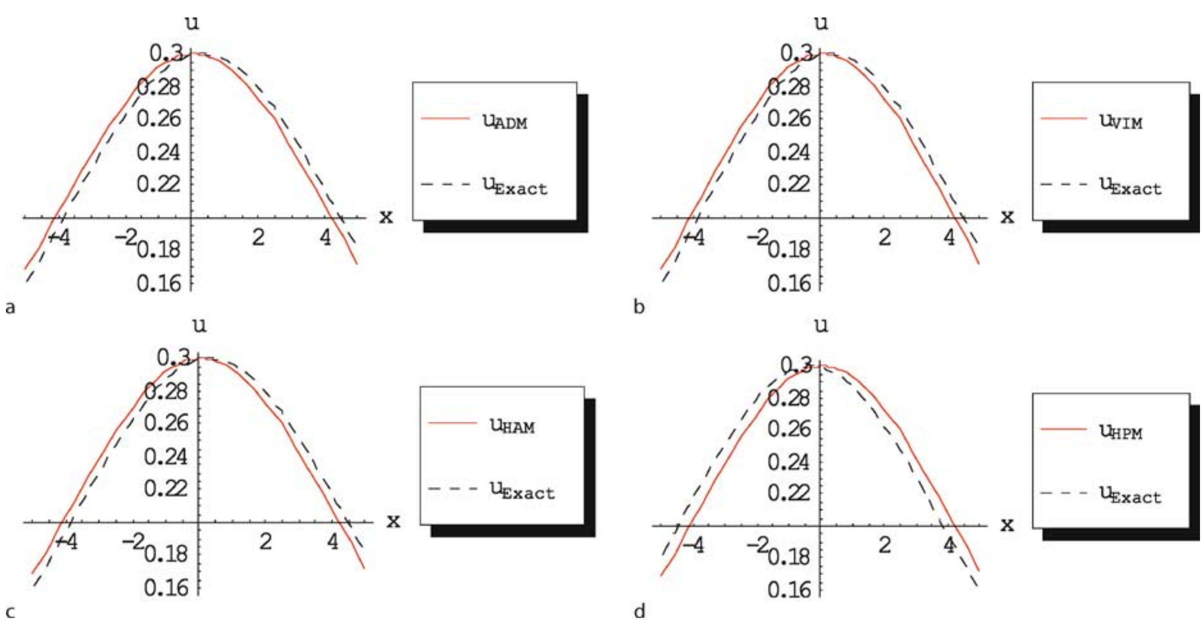
of the absolute errors between the exact solution and the numerical solution for the ADM, VIM and HAM. One can conclude that the values on the interval $0 \le x \le 15$ and for some small values of t all the considered methods give very close numerical solutions to a corresponding KdV equation. If one closely looks at Tables 1 and 3 then it is evident that the numerical results are almost the same for a few terms of the series solutions, such as the first three terms. The calculation for the later terms of the series has stopped because the scheme VIM does not work very well in the sense of computer time. This problem needs to be approached in a completely different way. Moreover, it has also been numerically proved that HPM is a special case of the HAM which is tabulated in Table 2 for the various values of h for interval $-1.4 \le h \le 0.6$ and similar constant values as in Table 1. It is numerically shown that when we take the value of $h = -1$ in the formula HAM, then (numerically) HAM $\equiv$ HPM in Table 2. It is also our experience that the choice of the h value depends on the chosen equation and the solution of the equation, too. This can be seen in Tables 2 and 3.

Overall, one can see that closeness of the approximate solutions and exact solutions of the all methods give almost equal absolute values of error, except for the approximate solutions from HAM. The numerical solutions with a special value of $h = -0.9$ for the HAM perform very well, as one can see in Table 3. It is also a valuable opportunity for us to use HAM among the considered methods because it gives the users flexibility to choose different values of h, depending on which value is necessary for the problem.

The graphs of the numerical values of the exact and approximate solutions of the classical KdV equation are

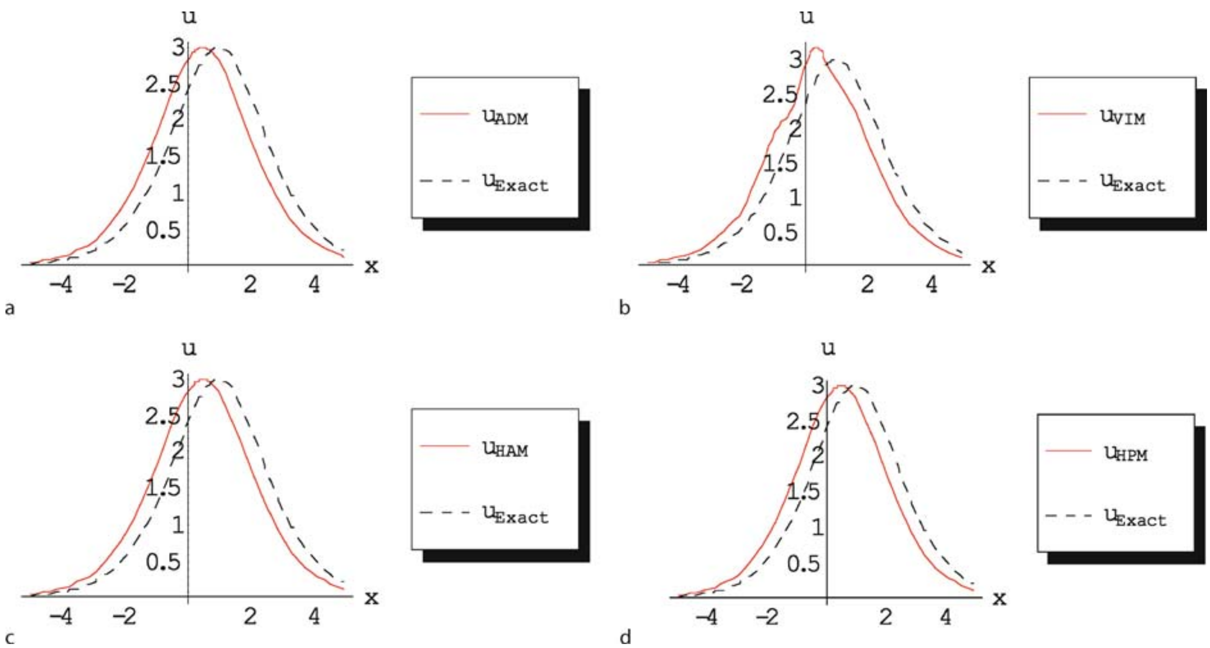


Korteweg–de Vries Equation (KdV) and Modified KdV (mKdV), Semi-analytical Methods for Solving the, Figure 3
A comparison by the ADM, VIM, HPM and HAM with exact solution of Eq. (6) with ($m = 1$) at $t = 0.5$ for $\alpha = 1,\ \lambda = 0.01$ in a–c and $\alpha = 1,\ \lambda = 0.01,\ h = -1$ in d, respectively

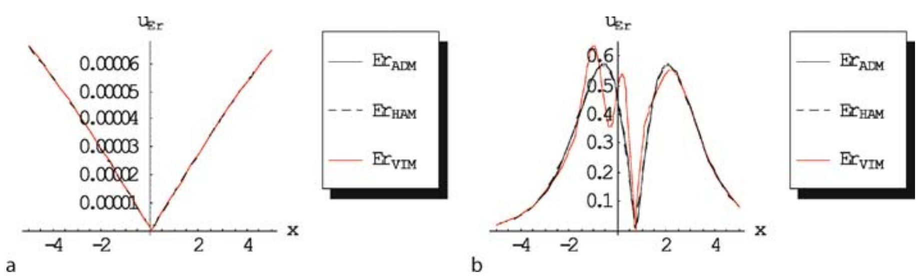


Korteweg–de Vries Equation (KdV) and Modified KdV (mKdV), Semi-analytical Methods for Solving the, Figure 4
A comparison by the ADM, VIM, HAM and HPM with exact solution of Eq. (6) with ($m = 1$) at $t = 0.5$ for $\alpha = 1,\ \lambda = 0.1$ in a–c and $\alpha = 1,\ \lambda = 0.1,\ h = -1$ in d, respectively

depicted in Figs. 3–6. In Figs. 3–5, and numerical results are given for the KdV equation and corresponding exact solution by using the considered methods. It is to be noted that only four terms were used in evaluating the approximate solutions with different values of the constant $\lambda = 0.01,\ 0.1,\ 1$ in order to get how close the approximate solutions were to the exact solution. A very good approximation to the actual solution of the equation was



Korteweg–de Vries Equation (KdV) and Modified KdV (mKdV), Semi-analytical Methods for Solving the, Figure 5
A comparison by the ADM, VIM, HAM and HPM with exact solution of Eq. (6) with ($m = 1$) at $t = 0.5$ for $\alpha = 1,\ \lambda = 1$ in a–c and $\alpha = 1,\ \lambda = 1,\ h = -1$ in d, respectively



Korteweg–de Vries Equation (KdV) and Modified KdV (mKdV), Semi-analytical Methods for Solving the, Figure 6
These two graphs is a comparison of $u_{\mathrm{Er}} = |u_{\mathrm{Exact}} - u_{\mathrm{Approx}}|$ for ADM, HAM and VIM of the solutions of Eq. (6) with ($m = 1$) at $t = 0.5$ for $\alpha = 1,\ \lambda = 0.01,\ h = -1$ in a and $\alpha = 1,\ \lambda = 1,\ h = -1$ in b, respectively

achieved by only using the four terms of the series derived above (59) and (66) for all considered methods with the value of $\lambda = 0.01$ which are depicted in Fig. 3. This is because of the nature of the series methods. The closeness of the numerical results for approximate solutions and the exact solution diverges for the other value of $\lambda = 0.1$ in all considered methods which are depicted in Fig. 4. This divergence is very clear in Fig. 5 for the results of the all the methods. But a sharp divergence can be seen in the numerical results of VIM especially at the value of $x = 0.5$, in Fig. 5b. In fact, these are illustrated in Tables 1–2 and Figs. 3–4. It is evident that the overall errors can be made smaller by adding new terms of the series (59) and (66).

One can see from Fig. 6a, the absolute values of the numerical results for approximate solutions and exact solutions are very small at the value of $0 \le x \le 1$ but other than this values of the x neither positive nor negative values of the x gives small errors for all considered methods with the value of $\lambda = 0.01$. However, in all considered methods with the value of $\lambda = 1$ a relatively small absolute error is given at the value of $0 \le x \le 1$, but all the methods are given biggest absolute error at a value of x around ± 2. The VIM is poorly performed at the value of x around -2, but besides this all the other methods are performed relatively well at the other values of the x, which can be seen in Fig. 4b.

Lastly, the clear conclusion can be draw from the numerical results that the ADM and HAM algorithms provide highly accurate numerical solutions without spatial discretizations for nonlinear partial differential equations. It is also worth noting that the advantage of the approximation of the series methodologies displays a fast convergence of the solutions. The illustrations show the dependence of the rapid convergence depends on the character and behavior of the solutions just as in a closed form solutions. Finally, it can be pointed out that, for given equations with initial values $u(x,0)$, the corresponding analytical and numerical solutions are obtained according to the recurrence relations (12), (25), (43)–(46) and (58) using Mathematica package version of Mathematica 4 in PC computer.

Future Directions

Nonlinear phenomena play a crucial role in applied mathematics and physics. Furthermore, when an original nonlinear equation is directly calculated, the solution will preserve the actual physical characters of solutions. Explicit solutions to the nonlinear equations are of fundamental importance. Various effective methods have been developed to understand the mechanisms of these physical models, to help physicists and engineers, and to ensure knowledge for physical problems and their applications.

In the future, the scientist will be very busy doing many works via applications of the nonlinear evolution equations in order to obtain exact and numerical solutions. One has to find out more efficient and effective methods to solve the nonlinear problems in applied mathematical areas of constructing either an exact solution or a numerical solution. Scientists have long way to go in this nonlinear study.

Bibliography

Primary Literature

1. Abbasbandy S (2006) The application of homotopy analysis method to nonlinear equations arising in heat transfer. Phys Lett A 360:109–13
2. Abdou MA, Soliman AA (2005) New applications of variational iteration method. Phys D 211:1–8
3. Abdou MA, Soliman AA (2005) Variational iteration method for solving Burger's and coupled Burger's equations. J Comput Appl Math 181:245–251
4. Adomian G (1988) A Review of the Decomposition Method in Applied Mathematics. J Math Anal Appl 135:501–544
5. Adomian G, Rach R (1992) Noise Terms in Decomposition Solution Series. Comput Math Appl 24:61–64
6. Adomian G (1994) Solving Frontier Problems of Physics: The decomposition method. Kluwer, Boston
7. Chowdhury MSH, Hashim I, Abdulaziz O (2009) Comparison of homotopy analysis method and homotopy-perturbation method for purely nonlinear fin-type problems. Commun Nonlinear Sci Numer Simul 14:371–378
8. Debtnath L (1997) Nonlinear Partial Differential Equations for Scientist and Engineers. Birkhauser, Boston
9. Debtnath L (2007) A brief historical introduction to solitons and the inverse scattering transform – a vision of Scott Russell. Int J Math Edu Sci Techn 38:1003–1028
10. Domairry G, Nadim N (2008) Assessment of homotopy analysis method and homotopy perturbation method in non-linear heat transfer equation. Intern Commun Heat Mass Trans 35:93–102
11. Drazin PG, Johnson RS (1989) Solutions: An Introduction. Cambrige University Press, Cambridge
12. Edmundson DE, Enns RH (1992) Bistable light bullets. Opt Lett 17:586
13. Fermi E, Pasta J, Ulam S (1974) Studies of nonlinear problems. American Math Soci, Providence, pp 143–156
14. Gardner CS, Greene JM, Kruskal MD, Miura RM (1967) Method for solving the Korteweg–de Vries equation. Phys Rev Lett 19:1095–1097
15. Gardner CS, Greene JM, Kruskal MD, Miura RM (1974) Korteweg–de Vries equation and generalizations, VI, Methods for exact solution. Commun Pure Appl Math 27:97–133
16. Geyikli T, Kaya D (2005) An application for a Modified KdV equation by the decomposition method and finite element method. Appl Math Comp 169:971–981
17. Geyikli T, Kaya D (2005) Comparison of the solutions obtained by B-spline FEM and ADM of KdV equation. Appl Math Comp 169:146–156
18. Hayat T, Abbas Z, Sajid M (2006) Series solution for the upper-convected Maxwell fluid over a porous stretching plate. Phys Lett A 358:396–403
19. Hayat T, Sajid M (2007) On analytic solution of thin film flow of a fourth grade fluid down a vertical cylinder. Phys Lett A 361:316–322
20. He JH (1997) A new approach to nonlinear partial differential equations. Commun Nonlinear Sci Numer Simul 2:203–205
21. He JH (1999) Homotopy perturbation technique. Comput Math Appl Mech Eng 178:257–262
22. He JH (1999) Variation iteration method – a kind of non-linear analytical technique: some examples. Int J Nonlinear Mech 34:699–708
23. He JH (2000) A coupling method of homotopy technique and a perturbation technique for nonlinear problems. Int J Nonlinear Mech 35:37–43
24. He JH (2003) Homotopy perturbation method a new nonlinear analytical technique. Appl Math Comput 135:73–79
25. He JH (2006) Some asymptotic methods for strongly nonlinear equations. Int J Mod Phys B 20:1141–1199
26. He JH, Wu XH (2006) Construction of solitary solution and compacton-like solution by variational iteration method. Chaos Solitons Fractals 29:108–113
27. Helal MA, Mehanna MS (2007) A comparative study between two different methods for solving the general Korteweg–de Vries equation. Chaos Solitons Fractals 33:725–739

28. Hirota R (1971) Exact solution of the Korteweg–de Vries equation for multiple collisions of solitons. Phys Rev Lett 27:1192–1194
29. Hirota R (1973) Exact N-solutions of the wave equation of long waves in shallow water and in nonlinear lattices. J Math Phys 14:810–814
30. Hirota R (1976) Direct methods of finding exact solutions of nonlinear evolution equations. In: Miura RM (ed) Bäcklund Transformations. Lecture Notes in Mathematics, vol 515. Springer, Berlin, pp 40–86
31. Inan IE, Kaya D (2006) Some Exact Solutions to the Potential Kadomtsev–Petviashvili Equation. Phys Lett A 355:314–318
32. Inan IE, Kaya D (2006) Some Exact Solutions to the Potential Kadomtsev–Petviashvili Equation. Phys Lett A 355:314–318
33. Inan IE, Kaya D (2007) Exact Solutions of the Some Nonlinear Partial Differential Equations. Phys A 381:104–115
34. Kaya D (2003) A Numerical Solution of the Sine–Gordon Equation Using the Modified Decomposition Method. Appl Math Comp 143:309–317
35. Kaya D (2006) The Exact and Numerical Solitary-wave Solutions for Generalized Modified Boussinesq Equation. Phys Lett A 348:244–250
36. Kaya D Some Methods for the Exact and Numerical Solutions of Nonlinear Evolution Equations. Paper presented at the 2nd International Conference on Mathematics: Trends and Developments (ICMTD), Cairo, Egypt, Dec 2007, pp 27–30 (in press)
37. Kaya D, Al-Khaled K (2007) A Numerical Comparison of a Kawahara Equation. Phys Lett A 363:433–439
38. Kaya D, El-Sayed SM (2003) An Application of the Decomposition Method for the Generalized KdV and RLW Equations. Chaos Solitons Fractals 17:869–877
39. Kaya D, El-Sayed SM (2003) Numerical Soliton-like Solutions of the Potential Kadomstev–Petviashvili Equation by the Decomposition Method. Phys Lett A 320:192–199
40. Kaya D, El-Sayed SM (2003) On a Generalized Fifth Order KdV Equations. Phys Lett A 310:44–51
41. Kaya D, Inan IE (2005) A Convergence Analysis of the ADM and an Application. Appl Math Comp 161:1015–1025
42. Khater AH, El-Kalaawy OH, Helal MA (1997) Two new classes of exact solutions for the KdV equation via Bäcklund transformations. Chaos Solitons Fractals 8:1901–1909
43. Korteweg DJ, de Vries H (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 39:422–443
44. Liao SJ (1992) The proposed homotopy analysis technique for the solution of nonlinear problems. Ph D Thesis, Shanghai Jiao Tong University
45. Liao SJ (1995) An approximate solution technique not depending on small parameters: A special example. Int J Nonlinear Mech 30:371
46. Liao SJ (2003) Beyond Perturbation: Introduction to Homotopy Analysis Method. Chapman & Hall/CRC Press, Boca Raton
47. Liao SJ (2004) On the homotopy analysis method for nonlinear problems. Appl Math Comp 147:499
48. Liao SJ (2004) On the homotopy analysis method for nonlinear problems. Appl Math Comput 147:499–513
49. Liao SJ (2005) Comparison between the homotopy analysis method and homotopy perturbation method. Appl Math Comp 169:1186–1194
50. Light Bullet Home Page, http://www.sfu.ca/~renns/lbullets.html
51. Miura RM (1968) Korteweg–de Vries equations and generalizations, I; A remarkable explicit nonlinear transformations. J Math Phys 9:1202–1204
52. Miura RM, Gardner CS, Kruskal MD (1968) Korteweg–de Vries equations and generalizations, II; Existence of conservation laws and constants of motion. J Math Phys 9:1204–1209
53. Momani S, Abuasad S (2006) Application of He's variational iteration method to Helmholtz equation. Chaos Solitons Fractals 27:1119–1123
54. Polat N, Kaya D, Tutalar HI (2006) A Analytic and Numerical Solution to a Modified Kawahara Equation and a Convergence Analysis of the Method. Appl Math Comp 179:466–472
55. Ramos JI (2008) On the variational iteration method and other iterative techniques for nonlinear differential equations. Appl Math Comput 199:39–69
56. Rayleigh L (1876) On Waves. Lond Edinb Dublin Philos Mag 5:257
57. Russell JS (1844) Report on Waves. In: Proc 14th meeting of the British Association for the Advancement of Science, BAAS, London
58. Sajid M, Hayat T (2008) Comparison of HAM and HPM methods in nonlinear heat conduction and convection equations, Nonlinear Analysis. Real World Appl 9:2296–2301
59. Sajid M, Hayat T (2008) The application of homotopy analysis method for thin film flow of a third order fluid. Chaos Solitons Fractal 38:506–515
60. Sajid M, Hayat T, Asghar S (2007) Comparison between HAM and HPM solutions of thin film flows of non-Newtonian fluids on a moving belt. Nonlinear Dyn 50:27–35
61. Shawagfeh N, Kaya D (2004) Series Solution to the Pochhammer–Chree Equation and Comparison with Exact Solutions. Comp Math Appl 47:1915–1920
62. Ugurlu Y, Kaya D Exact and Numerical Solutions for the Cahn–Hilliard Equation. Comput Math Appl
63. Wazwaz AM (1997) Necessary Conditions for the Appearance of Noise Terms in Decomposition Solution Series J Math Anal Appl 5:265–274
64. Wazwaz AM (2002) Partial Differential Equations: Methods and Applications. Balkema, Rotterdam
65. Wazwaz AM (2007) Analytic study for fifth-order KdV-type equations with arbitrary power nonlinearities. Comm Nonlinear Sci Num Sim 12:904–909
66. Wazwaz AM (2007) A variable separated ODE method for solving the triple sine-Gordon and the triple sinh-Gordon equations. Chaos Solitons Fractals 33:703–710
67. Wazwaz AM (2007) The extended tanh method for abundant solitary wave solutions of nonlinear wave equations. App Math Comp 187:1131–1142
68. Wazwaz AM (2007) The variational iteration method for solving linear and nonlinear systems of PDEs. Comput Math Appl 54:895–902
69. Wazwaz AM, Helal MA (2004) Variants of the generalized fifth-order KdV equation with compact and noncompact structures. Chaos Solitons Fractals 21:579–589
70. Whitham GB (1974) Linear and Nonlinear Waves. Wiley, New York
71. Zabusky NJ, Kruskal MD (1965) Interactions of solitons in collisionless plasma and the recurrence of initial states. Phys Rev Lett 15:240–243

Books and Reviews

The following, referenced by the end of the paper, is intended to give some useful for further reading.

For another obtaining of the KdV equation for water waves, see Kevorkian and Cole (1981); one can see the work of the Johnson (1972) for a different water-wave application with variable depth, for waves on arbitrary shears in the work of Freeman and Johnson (1970) and Johnson (1980) for a review of one and two-dimensional KdV equations. In addition to these; one can see the book of Drazin and Johnson (1989) for some numerical solutions of nonlinear evolution equations. In the work of the Zabusky, Kruskal and Deam (F1965) and Eilbeck (F1981), one can see the motion pictures of soliton interactions. See a comparison of the KdV equation with water wave experiments in Hammack and Segur (1974).

For further reading of the classical exact solutions of the nonlinear equations can be seen in the works: the Lax approach is described in Lax (1968); Calogero and Degasperis (1982, A.20), the Hirota's bilinear approach is developed in Matsuno (1984), the Bäckland transformations are described in Rogers and Shadwick (1982); Lamb (1980, Chap. 8), the Painleve properties is discussed by Ablowitz and Segur (1981, Sect. 3.8), In the book of Dodd, Eilbeck, Gibbon and Morris (1982, Chap. 10) can found review of the many numerical methods to solve nonlinear evolution equations and shown many of their solutions.

Korteweg–de Vries Equation (KdV), Some Numerical Methods for Solving the

MUSTAFA INC
Department of Mathematics, Fırat University,
Elazığ, Turkey

Article Outline

Glossary
Definition of the Subject
Introduction
Some Numerical Methods for Solving the Korteweg–de Vries (KdV) Equation
The Adomian Decomposition Method (ADM)
The Homotopy Analysis Method (HAM)
The Variational Iteration Method (VIM)
The Homotopy Perturbation Method (HPM)
Numerical Applications and Comparisons
Conclusions and Discussions
Future Directions
Bibliography

Glossary

Absolute error When a real number x is approximated by another number x^*, the error is $x - x^*$. The absolute error is $|x - x^*|$.

Adomian decomposition method The method was introduced and developed by George Adomian. The method proved to be powerful, effective, and can easily handle a wide class of linear or nonlinear differential equations.

Adomian polynomials It is well known now that the ADM suggests that the unknown linear function u may be presented by the decomposition series $u = \sum_{n=0}^{\infty} u_n$, the nonlinear term $F(u)$, such as u^2, u^3, $\sin u$, e^u, u_x^2, etc. can be expressed by an infinite series of the so-called Adomian polynomials A_n.

Finite difference method The method handles the differential equation by replacing the derivatives in the equation with difference quotients.

Homotopy analysis method The HAM was introduced by Shi-Jun Liao in 1992. For details see Sect. "The Variational Iteration Method (VIM)".

Homotopy perturbation method The HPM was introduced by Ji-Huan He in 1999. For details see Sect. "Numerical Applications and Comparisons".

Korteweg–de Vries equation This is one of the simplest and most useful nonlinear model equations for solitary waves.

Lagrange multiplier The multiplier in the functional should be chosen such that its correction solution is superior to its initial approximation (trial function) and is the best within the flexibility of trial function, accordingly we can identify the multiplier by variational theory.

Initial conditions The PDEs mostly arise to govern physics phenomena such as heat distribution, wave propagation phenomena. Most of the PDEs depend on the time t. The initial values of the dependent variable u at the starting time $t = 0$ should be prescribed.

Solitons It is interesting to point out that there is no precise definition of a soliton. However, a soliton can be defined as a solution of a nonlinear partial differential equation.

Variational iteration method The VIM was first proposed by Ji-Huan He. This method has been employed to solve a large variety of linear and nonlinear problems with approximations converging rapidly to accurate solutions.

Definition of the Subject

In this work, the Adomian decomposition method (ADM), the homotopy analysis method (HAM), the variational iteration method (VIM), the homotopy perturbation method (HPM) and the explicit finite difference method (EFDM) are implemented to investigate numerical solutions of the Korteweg–de Vries (KdV) equation with initial condition. The five methods are compared and it is shown that the HAM and EFDM are more efficient and effective than the ADM, the VIM and the HPM, and also converges to its exact solution more rapidly. The HAM contains the auxiliary parameter $\hbar$ which provides us with a simple way to adjust and control the convergence region of solution series.

Introduction

John Scott Russell first experimentally observed the solitary wave, a long water wave without change in shape, on the Edinburgh-Glasgow Canal in 1834. He called it the great wave of translation and, then, reported his observations at the British Association in his 1844 paper "Report on waves" [1].

After 60 years from this discovery, the two scientists D.J. Korteweg and G. de Vries formulated a mathematical model equation to provide an explanation of the phenomenon observed by Russell [2]. They derived the famous equation, the Korteweg–de Vries (KdV) equation, for the propagation of waves in one dimension on the surface of water. This is one of the simplest and most useful nonlinear model equations for solitary waves, and it represents the longtime evolution of wave phenomena in which the steepening effect of the nonlinear term is counterbalanced by dispersion. The KdV equation is a nonlinear partial differential equation of third order.

Modern developments in the theory and applications of the KdV equation began with the seminal work published as a Los Alamos Scientific Laboratory Report in 1955 by Fermi, Pasta and Ulam (FPU) on a numerical model of a discrete nonlinear mass-spring system [3]. This curious result of the FPU experiment inspired Zabusky and Kruskal [4] marked the birth of the new concept of the soliton, a name intended to signify particle like quantities. They found that stable pulse like waves could exist in a system described by the KdV equation. A remarkable quality of these solitary waves was that they could collide with each other and yet preserve their shapes and speeds after the collision. This means that a collision between KdV solitary waves is elastic. Subsequently, Zabusky [5] confirmed, numerically, the actual physical interaction of two solitons, and Lax [6] gave a rigorous analytical proof that the identities of two distinct solitons are preserved through the nonlinear interaction governed by the KdV equation. Subsequently, a paper by Gardner et al. [7] demonstrated that it was possible to write many exact solutions to the equation by using ideas from scattering theory. They discovered the first integrable nonlinear partial differential equation. Gardner et al. [8] and Hirota [9,10] constructed analytical solutions of the KdV equation that provide the description of the interaction among N solitons for any positive integer N. Experimental confirmation of solitons and their interactions has been provided successfully by Zabusky and Galvin [11], Hammack and Segur [12], and Weidman and Maxworthy [13]. During the past fifteen years a rather complete mathematical description of solitons has been developed.

The nondispersive nature of the soliton solutions to the KdV equation arises not because the effects of dispersion are absent, but because they are balanced by nonlinearities in the system [14,15,16,17]. The KdV equation [2,7,18] arises in several areas of nonlinear physics such as hydrodynamics, plasma physics, etc.

During the last five decades or so an exciting and extremely active area of research has been devoted to the construction of exact and numerical solutions for a wide class of nonlinear equations. This includes the most famous nonlinear one of Korteweg and de-Vries.

Exact solutions of the KdV equation have been used for many powerful methods such as the application of the Jacobian elliptic function [19,20,21], the powerful inverse scattering transform [22], the Painleve analysis [23], the Backlund transformations method [24], the Lie group theoretical methods [25], the direct algebraic method [26], the tanh-method [27], the sine-cosine method [28], the homogeneous balance method [29], the Riccati expansion method with constant coefficient [30] and the mapping method [31]. In addition, Sawada and Kotera [32], Rosales [33], Whitham [34], Wadati and Sawada [35,36] have all employed perturbation techniques. Recently, Helal and El-Eissa [37], Khater et al. [38,39], Helal [40] have studied analytically and numerically some physical problems that lead to the KdV equation and the soliton solutions have been obtained [17].

The KdV equation is a generic equation for the study of weakly nonlinear long waves. The KdV type of equation has been an important class of nonlinear evolution equations with numerous applications in physical sciences and engineering fields. For example, in plasma physics, these equations give rise to the ion acoustic solitons [41]; in geophysical fluid dynamics, they describe a long wave in shallow seas and deep oceans [42,43]. Their strong presence is exhibited in cluster physics, super deformed nuclei, fission, thin film, radar and rheology [44,45], optical-fiber communications [46] and superconductors [47].

Thus many analytical solutions of the KdV equation are found, and their existence and uniqueness have been studied for a certain class of initial functions [14]. The numerical solutions of the KdV equation are essential because of solutions which are not analytically available. Many methods have been proposed for numerical treatment of the KdV equation for the various boundary and initial conditions. For appropriate initial conditions, Gradner et al. [7] have shown the existence and uniqueness of solutions of the KdV equation. For the KdV equation, several quite successful numerical methods have been available such as spectral/pseudo spectral methods [48,49], finite-difference methods and Fourier spectral methods developed by many authors from both the theoretical and computational points of view [50,51]. The exponential finite-difference method is used to solve the KdV equation with the initial and boundary conditions by Bahadır [52]. Jain et al. [53] developed a numerical method for solving the KdV equation by using splitting method and quintic spline approximation technique. Soliman [54] worked numerical solutions for the KdV equation based on the collocation method using septic splines as element shape functions were set up. Frauendiener and Klein [55] presented the hyperelliptic theta-functions with spectral methods for numerical solutions of the KP and KdV equations. Bhatta and Bhatti [56] studied numerical solution of the KdV equation by using modified Bernstein polynomials. Helal and Mehanna [57] presented a comparative study between the Adomian decomposition method and the finite-difference method for solving the general KdV equation. Kutluay et al. [58] used the heat balance integral method to the KdV equation prescribed by appropriate homogenous boundary conditions and a set of initial conditions to obtain its approximate analytical solutions at small times. An analytical-numerical method was applied to the KdV equation with a variant of boundary and initial conditions to obtain its numerical solutions at small times by Özer and Kutluay [59]. Recently, Dehghan and Shokri [60] proposed a numerical scheme to solve the third-order nonlinear KdV equation using collocation points and approximating the solution using multiquadric radial basis function. Dag and Dereli [61] presented numerical solution of the KdV equation by using the meshless method based on the collocation with radial basis functions.

Some Numerical Methods for Solving the Korteweg–de Vries (KdV) Equation

The Korteweg–de Vries (KdV) equation has the following form:

$$u_t - 6uu_x + u_{xxx} = 0\,, \tag{1}$$

subject to initial condition

$$u(x,0) = f(x)\,, \tag{2}$$

where $u(x,t)$ is a differentiable function and $f(x)$ is bounded. We shall assume that the solution $u(x,t)$, along with its derivatives, tends to zero as $|x| \to \infty$. The second and third terms uu_x and u_{xxx} represent the nonlinear convection and dispersion effects, respectively. The solitons of the KdV equation arise as a balance between nonlinear convection and dispersion terms. The nonlinear effect causes the steepening of the waveform [61], while the dispersion effect makes the waveform spread. Due to the competition of these two effects, a stationary waveform (solitary wave) exists. The solution is obtained among the nonlinear and dispersion by stability. The soliton solutions of nonlinear wave equations shown by Wadati [62,63] have the following properties:

- Some certain waves propagate which does not change its special behavior.
- Reginal waves do not lose their properties and also they are stable towards the collisions.

The basic purpose of this work is to approach the KdV Eq. (1) with initial condition (2) differently, by using four semi inverse methods and the EFDM. In this work, we will use the Adomian decomposition method (ADM), the homotopy analysis method (HAM), the variational iteration method (VIM), the homotopy perturbation method (HPM) and the explicit finite difference method (EFDM). The results are compared with those obtained using the five methods. More details for the five methods can be found in next sections.

The Adomian Decomposition Method (ADM)

The ADM was first introduced by Adomian in the beginning of [64,65,66]. The method is useful for obtaining both a closed form and the explicit solution and numerical approximations of linear or nonlinear differential equations and it is also quite straight forward for writing computer codes. This method has been applied to obtain a formal solution to a wide class of stochastic and deterministic problems in science and engineering involving algebraic, differential, integro-differential, differential delay, integral and partial differential equations [67,68,69,70,71,72,73,74,75,76]. The convergence of ADM for partial differential equations was presented by Cherruault [77]. Application and convergence of this method for nonlinear partial differential equations are found in [78,79,80,81].

In general, it is necessary to construct the solution of the problems in the form of a decomposition series solution. In the simplest case, the solution can be developed as a Taylor series expansion about the function, not the point at which the initial condition and integration of the right-hand side function of the problem are determined (the first term u_0 of the decomposition series for $n \geq 0$). The sum of the $u_0, u_1, u_2, \ldots$ terms are simply the decomposition series

$$u(x,t) = \sum_{n=0}^{\infty} u_n(x,t) \ . \tag{3}$$

Suppose that the differential equation operator, including both linear and nonlinear terms, can be formed as

$$Lu + Ru + Nu = F(x,t) \ , \tag{4}$$

with initial condition

$$u(x,0) = g(x) \ , \tag{5}$$

where L is the higher-order derivative which is assumed to be invertible, R is a linear differential operator of order less than L, N is the nonlinear term and $F(x,t)$ is a source term. We next apply the inverse operator L^{-1} to both sides of Eq. (4) and using the given condition (5) to obtain

$$u(x,t) = g(x) + f(x,t) - L^{-1}(Ru) - L^{-1}(Nu) \ , \tag{6}$$

where the function $f(x,t)$ represents the terms arising from integrating the source term $F(x,t)$ and from using the given conditions, all are assumed to be prescribed. The nonlinear term can be written as

$$Nu = \sum_{n=0}^{\infty} A_n \ , \tag{7}$$

where A_n is the Adomian polynomials. These polynomials are defined as

$$A_n = \frac{1}{n!} \frac{\mathrm{d}^n}{\mathrm{d}\lambda^n} \left[N\left(\sum_{n=0}^{\infty} \lambda^k u_k(x,t) \right) \right]_{k=0} \ , \quad n = 0,1,2,\ldots \tag{8}$$

for example

$$\begin{aligned} A_0 &= N(u_0) \ , \\ A_1 &= u_1 N'(u_0) \ , \\ A_2 &= u_2 N'(u_0) + \frac{1}{2} u_1 N''(u_0) \ , \\ A_3 &= u_3 N'(u_0) + u_1 u_2 N''(u_0) + \frac{1}{6} u_1^3 N'''(u_0) \ , \end{aligned} \tag{9}$$

and so on, the other polynomials can be constructed in a similar way [66]. As indicated before, Adomian method defines the solution u by an infinite series of components given by Eq. (4) and the components $u_0, u_1, u_2, \ldots$ are usually recurrently determined. Thus, the formal recursive relation is defined by

$$\begin{cases} u_0(x,t) &= g(x) + f(x,t) \ , \\ u_{n+1}(x,t) &= -L^{-1}(Ru_n) - L^{-1}(Nu_n) \ , \quad n \geq 0 \ , \end{cases} \tag{10}$$

which are obtained for all components of u. As a result, the terms of the series $u_0, u_1, u_2, \ldots$ are identified and the exact solution may be entirely determined by using the approximation

$$u(x,t) = \lim_{n\to\infty} \varphi_n(x,t) \ , \tag{11}$$

where

$$\varphi_n(x,t) = \sum_{k=0}^{n-1} u_k(x,t) \ , \tag{12}$$

or

$$\begin{cases} \varphi_0 = u_0 \ , \\ \varphi_1 = u_0 + u_1 \ , \\ \varphi_2 = u_0 + u_1 + u_2 \ , \\ \vdots \\ \varphi_n = u_0 + u_1 + u_2 + \ldots + u_{n-1}, \quad n \geq 0 \ . \end{cases} \tag{13}$$

The Homotopy Analysis Method (HAM)

The HAM was developed in 1992 by Liao in [82,83,84,85]. This method has been successfully applied by many authors [86,87,88,89,90,91]. The HAM contains the auxiliary parameter $\hbar$ which provides us with a simple way to adjust and control the convergence region of solution series for large or small values of x and t.

We consider the following differential equation

$$N[u(x,t)] = 0 \ , \tag{14}$$

where N is a nonlinear differential operator, x and t denote independent variables, $u(x,t)$ is an unknown function. By means of the HAM, one first constructs a zero-order deformation equation

$$(1-p)\,\pounds\left[\phi(x,t;p) - u_0(x,t)\right] = p\,\hbar\,N\left[\phi(x,t;p)\right] \ , \tag{15}$$

where $p \in [0,1]$ is the embedding parameter, $\hbar \neq 0$ is a non-zero auxiliary parameter, is an auxiliary linear operator, $u_0(x,t)$ is an initial guess of $u(x,t)$ and $\phi(x,t;p)$

is a unknown function. It is important that one has great freedom to choose auxiliary things in the HAM. Obviously, when $p = 0$ and $p = 1$, it holds

$$\phi(x, t; 0) = u_0(x, t) , \quad \phi(x, t; 1) = u(x, t) , \tag{16}$$

respectively. The solution $\phi(x, t; p)$ varies from the initial guess $u_0(x, t)$ to the solution $u(x, t)$. Expanding $\phi(x, t; p)$ in Taylor series about the embedding parameter

$$\phi(x, t; p) = u_0(x, t) + \sum_{m=1}^{+\infty} u_m(x, t)\, p^m , \tag{17}$$

where

$$u_m(x, t) = \frac{1}{m!} \frac{\partial^m \phi(x, t; p)}{\partial p^m}\bigg|_{p=0} . \tag{18}$$

The convergence of the series (17) depends upon the auxiliary parameter $\hbar$. If it is convergent at $p = 1$, one has

$$u(x, t) = u_0(x, t) + \sum_{m=1}^{+\infty} u_m(x, t) . \tag{19}$$

According to the definition (19), the governing equation can be deduced from the zero-order deformation Eq. (15). Define the vector

$$\vec{u}_n = \{u_0(x, t), u_1(x, t), \dots, u_n(x, t)\} .$$

Differentiating Eq. (15) m-times with respect to the embedding parameter p and then setting $p = 0$ and finally dividing them by $m!$, we have the so-called mth-order deformation equation

$$£\left[u_m(x, t) - \chi_m u_{m-1}(x, t)\right] = \hbar\, \Re_m\left(\vec{u}_{m-1}\right) , \tag{20}$$

where

$$\Re_m\left(\vec{u}_{m-1}\right) = \frac{1}{(m-1)!} \frac{\partial^{m-1} N\left[\phi(x, t; p)\right]}{\partial p^{m-1}}\bigg|_{p=0} , \tag{21}$$

and

$$\chi_m = \begin{cases} 0, & m \le 1 , \\ 1, & m > 1 . \end{cases} \tag{22}$$

It should be emphasized that $u_m(x, t)$ for $m \ge 1$ is governed by the nonlinear Eq. (20) with the linear boundary conditions that come from the original problem, which can be easily solved by symbolic computation software such as Maple and Mathematica.

The Variational Iteration Method (VIM)

The VIM was first proposed by He [92,93,94,95,96]. This method has been employed to solve a large variety of linear and nonlinear problems with approximations converging rapidly to accurate solutions. The idea of VIM is to construct a correction functional by a general Lagrange multiplier. The multiplier in the functional should be chosen such that its correction solution is superior to its initial approximation (trial function) and is the best within the flexibility of the trial function, accordingly we can identify the multiplier by variational theory. The initial approximation can be freely chosen with possible unknowns, which can be determined by imposing the boundary/initial conditions.

We consider the following general differential equation:

$$Lu + Nu = g , \tag{23}$$

where L and N are linear and nonlinear operators respectively, and $g(t)$ is the source inhomogeneous term.

According to the VIM, the terms of a sequence $\{u_n\}$ are constructed such that this sequence converges to the exact solution. u_ns are calculated by a correction functional as follows:

$$u_{n+1}(x, t) = u_n(x, t) + \int_0^t \lambda \left\{Lu_n(\xi) + N\tilde{u}_n(\xi) - g(\xi)\right\} d\xi , \quad n \ge 0 , \tag{24}$$

where λ is the general Lagrange multiplier, which can be identified optimally by the variational theory, the subscript n denotes the nth approximation and $\tilde{u}_n$ considered as a restricted variation, i. e. $\delta\tilde{u}_n = 0$. For linear problems, the exact solution can be obtained by only one iteration step due to the fact that the Lagrange multiplier can be exactly identified. In nonlinear problems, in order to determine the Lagrange multiplier in a simple manner, the nonlinear terms have to be considered as restricted variations. Consequently, the exact solution may be obtained by using

$$u(x, t) = \lim_{n\to\infty} u_n(x, t) . \tag{25}$$

This method has been employed to solve a large variety of linear and nonlinear problems with approximations converging to accurate solutions. This technique was used to find the solutions of partial differential equations, ordinary differential equations, boundary-value problems and integral equations by many authors [97,98,99,100, 101,102,103,104,105].

The Homotopy Perturbation Method (HPM)

The HPM was developed by He [106,107,108,109]. In recent years, the applications of HPM in nonlinear problems has been devoted to scientists and engineers [110,111,112,113].

We consider the following general nonlinear differential equation:

$$A\left(y\right) - f\left(r\right) = 0\,, \quad r \in \Omega\,, \tag{26}$$

with boundary conditions

$$B\left(y, \partial y/\partial n\right) = 0\,, \quad r \in \Gamma\,, \tag{27}$$

where A is a general differential operator, B is a boundary operator, $f(r)$ is a known analytic function and Γ is the boundary of the domain Ω. The operator A can be generally divided into two parts L and N, where L is linear and N is nonlinear. Therefore, Eq. (26) can be written as follows:

$$L\left(y\right) + N\left(y\right) - f\left(r\right) = 0\,. \tag{28}$$

We construct a homotopy of Eq. (26) $y\left(r, p\right) : \Omega \times [0, 1] \to R$ which satisfies

$$\begin{aligned} H\left(y, p\right) &= \left(1-p\right)\left[L\left(y\right) - L\left(y_0\right)\right] + p\left[A\left(y\right) - f\left(r\right)\right] \\ &= 0\,, \quad r \in R\,, \end{aligned} \tag{29}$$

which is equivalent to

$$\begin{aligned} H\left(y, p\right) &= L\left(y\right) - L\left(y_0\right) + pL\left(y_0\right) + p\left[A\left(y\right) - f\left(r\right)\right] \\ &= 0\,, \end{aligned} \tag{30}$$

where $p \in [0, 1]$ is an embedding parameter and y_0 is an initial approximation which satisfies the boundary conditions. It follows from Eqs. (29) and (30) that

$$\begin{aligned} &H\left(y, 0\right) = L\left(y\right) - L\left(y_0\right) = 0 \quad \text{and} \\ &H\left(y, 1\right) = A\left(y\right) - f\left(r\right) = 0\,. \end{aligned} \tag{31}$$

Thus, the changing process of p from 0 to 1 is just that of $y\left(r, p\right)$ from $y_0\left(r\right)$ to $y\left(r\right)$. In topology this is called deformation and $L\left(y\right) - L\left(y_0\right)$ and $A\left(y\right) - f\left(r\right)$ are called homotopic. Here the embedding parameter is introduced much more naturally, unaffected by artificial factors; further it can be considered as a small parameter for $0 \le p \le 1$. So it is very natural to assume that the solution of (29) and (30) can be expressed as

$$y\left(x\right) = u_0\left(x\right) + pu_1\left(x\right) + p^2 u_2\left(x\right) + \cdots\,. \tag{32}$$

According to HPM, the approximate solution of (26) can be expressed as a series of the power of p, i. e.,

$$y = \lim_{p \to 1} y\left(x\right) = u_0\left(x\right) + u_1\left(x\right) + u_2\left(x\right) + \cdots\,. \tag{33}$$

The convergence of series (33) has been proved by He in [106].

Numerical Applications and Comparisons

We now consider the initial value problem associated with the KdV equation:

$$\begin{cases} u_t - 6uu_x + u_{xxx} = 0\,, \\ u\left(x, 0\right) = -\frac{c}{2}\sec h^2\left(\frac{\sqrt{c}}{2}x\right)\,, \end{cases} \tag{34}$$

with $u(x, t)$ is a sufficiently smooth function and c is the velocity of the wavefront.

In this section, we apply the above described four methods and the explicit finite difference method on the KdV Eq. (34) so that the comparisons are made numerically.

The ADM for Eq. (14)

According to the ADM [64,65,66], Eq. (34) can be written in an operator form

$$\begin{cases} Lu = 6uu_x - u_{xxx}\,, \\ u\left(x, 0\right) = f\left(x\right) = -\frac{c}{2}\sec h^2\left(\frac{\sqrt{c}}{2}x\right)\,, \end{cases} \tag{35}$$

where the differential operator L is

$$L \equiv \frac{\partial}{\partial t}\,, \tag{36}$$

and the inverse operator L^{-1} is an integral operator given by

$$L^{-1}\left(\circ\right) = \int_0^t \left(\circ\right) \mathrm{d}t\,. \tag{37}$$

The ADM [64,65,66] assumes a series solution for the unknown function $u(x, t)$ given by Eq. (3) and the nonlinear term $Nu = uu_x$ can be decomposed into a infinite series of polynomials given by Eq. (5). Operating with the integral operator (37) on both sides of (35) and using the initial condition, we get

$$u\left(x, t\right) = f\left(x\right) + L^{-1}\left(6uu_x - u_{xxx}\right)\,. \tag{38}$$

Substituting (3) and (5) into the functional Eq. (38) yields

$$\sum_{n=0}^{\infty} u_n(x,t) = f(x) + L^{-1}\left[6\left(\sum_{n=0}^{\infty} A_n\right) - \left(\sum_{n=0}^{\infty} u_n\right)_{xxx}\right]. \tag{39}$$

The ADM admits that the zeroth component $u_0(x,t)$ be identified by all terms that arise from the initial conditions and from the source terms if exist, and as a result, the remaining components $u_n(x,t)$, $n \geq 1$ can be determined by using the recurrence relation:

$$\begin{cases} u_0(x,t) = f(x) = -\frac{c}{2}\sec h^2\left(\frac{\sqrt{c}}{2}x\right), \\ u_{k+1}(x,t) = L^{-1}\left(6A_k - (u_k)_{xxx}\right), \quad k \geq 0, \end{cases} \tag{40}$$

where A_k, $k \geq 0$ are Adomian polynomials that represent the nonlinear term (uu_x) and given by

$$\begin{cases} A_0 = u_0 u_{0x}, \\ A_1 = u_{0x}u_1 + u_0 u_{1x}, \\ A_2 = u_1 u_{1x} + u_{0x}u_2 + u_0 u_{2x}, \\ A_3 = u_1 u_{2x} + u_{1x}u_2 + u_{0x}u_3 + u_0 u_{3x}. \end{cases} \tag{41}$$

Thus, some of the symbolically computed components are found as:

$$\begin{aligned} u_0(x,t) &= -\frac{c}{2}\sec h^2\left(\frac{\sqrt{c}}{2}x\right), \\ u_1(x,t) &= -\frac{c^{5/2}}{2}\sec h^2\left(\frac{\sqrt{c}}{2}x\right)\tanh\left(\frac{\sqrt{c}}{2}x\right)t, \\ u_2(x,t) &= -\frac{c^4}{8}\left(\cosh\left[\sqrt{c}x\right]-2\right)\sec h^4\left(\frac{\sqrt{c}}{2}x\right)t^2, \\ u_3(x,t) &= -\frac{c^{11/2}}{48}\sec h^5\left(\frac{\sqrt{c}}{2}x\right) \\ &\quad \cdot\left(-11\sinh\left(\frac{\sqrt{c}}{2}x\right)+\sinh\left(\frac{3\sqrt{c}}{2}x\right)\right)t^3, \\ u_4(x,t) &= -\frac{c^7}{384}\sec h^6\left(\frac{\sqrt{c}}{2}x\right) \\ &\quad \cdot\left[33-26\cosh\left[\sqrt{c}x\right]+\cosh\left[2\sqrt{c}x\right]\right]t^4, \end{aligned}$$

and so on. In this manner the other components of the decomposition series can be easily obtained using any symbolic program. The solution in a series for is given by

$$\begin{aligned} u(x,t) &= \sum_{n=0}^{\infty} u_n(x,t) \\ &= u_0(x,t) + u_1(x,t) + u_2(x,t) + u_3(x,t) + \cdots, \\ &= -\frac{c}{2}\sec h^2\left(\frac{\sqrt{c}}{2}x\right) \\ &\quad -\frac{c^{5/2}}{2}\sec h^2\left(\frac{\sqrt{c}}{2}x\right)\tanh\left(\frac{\sqrt{c}}{2}x\right)t \\ &\quad -\frac{c^4}{8}\left(\cosh\left[\sqrt{c}x\right]-2\right)\sec h^4\left(\frac{\sqrt{c}}{2}x\right)t^2 \\ &\quad -\frac{c^{11/2}}{48}\sec h^5\left(\frac{\sqrt{c}}{2}x\right) \\ &\quad \cdot\left(-11\sinh\left(\frac{\sqrt{c}}{2}x\right)+\sinh\left(\frac{3\sqrt{c}}{2}x\right)\right)t^3 \\ &\quad +\ldots, \end{aligned} \tag{42}$$

so that the exact solution

$$u(x,t) = -\frac{c}{2}\sec h^2\left(\frac{\sqrt{c}}{2}(x-ct)\right), \tag{43}$$

is readily obtained.

The convergence analysis of the ADM applied to the KdV Eq. (35) without the initial condition has been conducted by Helal and Mehanna [55].

The HAM for Eq. (34)

We will apply the HAM to the Eq. (34) to illustrate the strength of the method and to establish exact solution for this Eq. (34). We choose the linear operator

$$£\left[\phi(x,t;p)\right] = \frac{\partial\phi(x,t;p)}{\partial t}, \tag{44}$$

with property

$$£[k] = 0, \tag{45}$$

where k is a constant. We now define a nonlinear operator

$$\begin{aligned} N\left[\phi(x,t;p)\right] &= \frac{\partial\phi(x,t;p)}{\partial t} \\ &\quad - 6\phi(x,t;p)\frac{\partial\phi(x,t;p)}{\partial x} + \frac{\partial^3\phi(x,t;p)}{\partial x^3}. \end{aligned} \tag{46}$$

We construct the zeroth-order deformation equation

$$(1-p)\,\pounds\left[\phi\,(x,t;p)-u_0\,(x,t)\right]=p\,\hbar\,N\left[\phi\,(x,t;p)\right]. \tag{47}$$

For $p=0$ and $p=1$, we can write

$$\begin{cases}\phi\,(x,t;0)=u\,(x,0)=u_0\,(x,t)\,,\\ \phi\,(x,t;1)=u\,(x,t)\,.\end{cases} \tag{48}$$

So we obtain mth-order deformation equation

$$\pounds\left[u_m\,(x,t)-\chi_m\,u_{m-1}\,(x,t)\right]=\hbar\,\Re_m\left(\vec{u}_{m-1}\right), \tag{49}$$

where

$$\Re_m\left(\vec{u}_{m-1}\right)=\frac{\partial\phi_{m-1}\,(x,t;p)}{\partial t}-6\sum_{n=0}^{m-1}\phi_n\,(x,t;p)\,\frac{\partial\phi_{m-1-n}\,(x,t;p)}{\partial x}+\frac{\partial^3\phi_{m-1}\,(x,t;p)}{\partial x^3}\,. \tag{50}$$

Now the solution of the mth-order deformation Eq. (49) for $m\geq 1$ become

$$u_m\,(x,t)=\chi_m\,u_{m-1}\,(x,t)+\hbar\,\pounds^{-1}\left[\Re_m\left(\vec{u}_{m-1}\right)\right]. \tag{51}$$

Thus we get the following terms by using (51) for Eq. (34):

$$u_0\,(x,t)=-\frac{c}{2}\sec h^2\left(\frac{\sqrt{c}}{2}x\right),$$

$$u_1\,(x,t)=\frac{c^{5/2}}{2}\hbar\sec h^2\left(\frac{\sqrt{c}}{2}x\right)\tanh\left(\frac{\sqrt{c}}{2}x\right)t\,,$$

$$u_2\,(x,t)=-\frac{c^{5/2}}{2}\hbar t\sec h^4\left(\frac{\sqrt{c}}{2}x\right)\left[c^{3/2}\hbar t\cosh\left(\sqrt{c}x\right)-2\left(c^{3/2}\hbar t+(1+\hbar)\sinh\left(\sqrt{c}x\right)\right)\right],$$

$$\begin{aligned}u_3(x,t)=\frac{c^{5/2}}{48}\hbar t\sec h^2\left(\frac{\sqrt{c}}{2}x\right)\Bigg[&c^3\hbar^2t^2\sec h^3\left(\frac{\sqrt{c}}{2}x\right)\\&\cdot\sinh\left(\frac{3\sqrt{c}}{2}x\right)+24\hbar(1+\hbar)\tanh\left(\frac{\sqrt{c}}{2}x\right)\\&+\sec h^2\left(\frac{\sqrt{c}}{2}x\right)\Bigg(-12c^{3/2}\hbar(1+\hbar)t\cosh(\sqrt{c}x)\\&+12(1+\hbar)\sinh(\sqrt{c}x)\\&+c^{3/2}\hbar t\left(24+24\hbar-11c^{3/2}\hbar t\tanh\left(\frac{\sqrt{c}}{2}x\right)\right)\Bigg)\Bigg],\end{aligned}$$

and so on. In this manner the other components of the decomposition series can be easily obtained using any symbolic program.

Liao [82,83,84,85] showed that whatever a solution series converges to it will be one of the solutions of considered problem. Liao [82,83,84,85] showed the approximate solutions obtained by the HAM to be controlled by the auxiliary parameter and the rate of convergence of $\hbar$.

It is noted that the approximate solutions converge at $-2.1\leq\hbar\leq-1.1$ for 5th-order approximation and at $-4\leq\hbar\leq 2$ for 10th-order approximation for the KdV Eq. (34) when $x=10,\ t=0.5,\ c=1$ and $c=5$ (see Figs. 2 and 3).

The VIM for Eq. (34)

To solve the KdV Eq. (34) by means of the VIM, we construct a correction functional which reads as

$$u_{n+1}\,(x,t)=u_n\,(x,t)+\int_0^t\lambda\left(\frac{\partial u_n\,(x,\tau)}{\partial\tau}-N\tilde{u}_n\,(x,\tau)+\frac{\partial^3\tilde{u}_n\,(x,\tau)}{\partial x^3}\right)\mathrm{d}\tau\,, \tag{52}$$

where λ is the general Lagrange multiplier [109] whose optimal value is found using variational theory, $u_0\,(x,t)$ is an initial approximation which must be chosen suitable and $N\tilde{u}_n$ is the restricted variation, i. e., $\delta\tilde{u}_n=0$ [107]. To find the optimal values of λ we have

$$\delta u_{n+1}\,(x,t)=\delta u_n\,(x,t)+\delta\int_0^t\lambda\left(\frac{\partial u_n\,(x,\tau)}{\partial\tau}\right)\mathrm{d}\tau, \tag{53}$$

that results

$$\delta u_{n+1}\,(x,t)=\delta u_n\,(x,t)\,(1+\lambda)-\int_0^t\delta u_n\,(x,t)\,\lambda'\,\mathrm{d}\tau\,, \tag{54}$$

which yields

$$\lambda'\,(\tau)=0\,|_{\tau=t}\,,\,1+\lambda\,(\tau)=0\,|_{\tau=t}\,. \tag{55}$$

Thus we have

$$\lambda\,(t)=-1\,, \tag{56}$$

and we obtain the following iteration formula

$$u_{n+1}(x,t) = u_n(x,t) - \int_0^t \left(\frac{\partial u_n(x,\tau)}{\partial \tau} - 6u_n(x,\tau)\frac{\partial u_n(x,\tau)}{\partial x} + \frac{\partial^3 u_n(x,\tau)}{\partial x^3} \right) \mathrm{d}\tau \,. \quad (57)$$

Now using (57) we can find the solution of the KdV Eq. (34) as a convergent sequence. We get the following components:

$$u_0(x,t) = -\frac{c}{2}\sec h^2\left(\frac{\sqrt{c}}{2}x\right),$$

$$u_1(x,t) = -\frac{c}{2}\sec h^3\left(\frac{\sqrt{c}}{2}x\right)\left[\cosh\left(\frac{\sqrt{c}}{2}x\right) + c^{3/2}t\sinh\left(\frac{\sqrt{c}}{2}x\right)\right],$$

$$\begin{aligned} u_2(x,t) = -\frac{c}{32}\sec h^7\left(\frac{\sqrt{c}}{2}x\right)\Bigg[& -2\left(4c^3t^2-5\right) \cdot \cosh\left(\frac{\sqrt{c}}{2}x\right) + \left(5-c^3t^2\right)\cosh\left(\frac{3\sqrt{c}}{2}x\right) \\ & + \cosh\left(\frac{5\sqrt{c}}{2}x\right) + c^3t^2\cosh\left(\frac{5\sqrt{c}}{2}x\right) \\ & + 4c^{3/2}t\sinh\left(\frac{\sqrt{c}}{2}x\right) - 60c^{9/2}t^3\sinh\left(\frac{\sqrt{c}}{2}x\right) \\ & + 6c^{3/2}t\sinh\left(\frac{3\sqrt{c}}{2}x\right) + 12c^{9/2}t^3\sinh\left(\frac{3\sqrt{c}}{2}x\right) \\ & + 2c^{3/2}t\sinh\left(\frac{5\sqrt{c}}{2}x\right)\Bigg], \end{aligned}$$

and so on. In the same manner the rest of components of the iteration formulae (57) can be obtained using any symbolic packages.

So we obtain

$$u(x,t) = \lim_{n\to\infty} u_n(x,t) = -\frac{c}{2}\sec h^2\left(\frac{\sqrt{c}}{2}(x-ct)\right). \quad (58)$$

The HPM for Eq. (34)

To investigate the approximate solution of Eq. (34), we first construct a homotopy as follows:

$$(1-p)\left(\dot{Y}-\dot{u}_0\right) + p\left(\dot{Y}-6YY'+Y'''\right) = 0\,, \quad (59)$$

where "primes" denote differentiation with respect to x, and "dot" denotes differentiation with respect to t. In order to obtain the unknowns of $Y_i(x,t)$, $i=1,2,3,4$, we must construct and solve the following system which includes four equations with three unknowns, considering the initial condition of $Y(x,0) = u(x,0)$ and having the initial approximation of Eq. (34):

$$\begin{aligned} p^0 &: \quad \dot{Y}_0 - \dot{u}_0 = 0\,, \\ p^1 &: \quad \dot{Y}_1 + \dot{u}_0 - 6Y_0Y_0' + Y_0''' = 0\,, \\ p^2 &: \quad \dot{Y}_2 - 6Y_0Y_1' - 6Y_1Y_0^0 + Y_1''' = 0\,, \\ p^3 &: \quad \dot{Y}_3 - 6Y_0Y_2' - 6Y_1Y_1^{2\prime} - 6Y_2Y_0' + Y_2''' = 0\,. \end{aligned} \quad (60)$$

Thus, we will obtain:

$$u(x,t) = \lim_{p\to 1} Y_k(x,t)\,, \quad k = 0,1,2,3,\cdots\,. \quad (61)$$

To calculate the terms of the homotopy series (61) for $u(x,t)$, we substitute the initial condition into the systems (60) and finally using Mathematica, the solutions of the equation can be obtained as follows:

$$\begin{aligned} u_0(x,t) &= -\frac{c}{2}\sec h^2\left(\frac{\sqrt{c}}{2}x\right), \\ u_1(x,t) &= -\frac{c^{5/2}}{2}\sec h^2\left(\frac{\sqrt{c}}{2}x\right)\tanh\left(\frac{\sqrt{c}}{2}x\right)t\,, \\ u_2(x,t) &= \frac{c^{5/2}}{2}t^2\sec h^4\left(\frac{\sqrt{c}}{2}x\right)\left[-c^{3/2}\cosh\left(\sqrt{c}x\right)+2c^{3/2}\right], \end{aligned} \quad (62)$$

and so on. In this manner the other components can be easily obtained. Substituting Eq. (62) into Eq. (61):

$$\begin{aligned} u(x,t) &= Y_0(x,t) + Y_1(x,t) + Y_2(x,t) + \cdots \\ &= -\frac{c}{2}\sec h^2\left(\frac{\sqrt{c}}{2}x\right) \\ &\quad - \frac{c^{5/2}}{2}\sec h^2\left(\frac{\sqrt{c}}{2}x\right)\tanh\left(\frac{\sqrt{c}}{2}x\right)t \\ &\quad - \frac{c^4}{2}t^2\sec h^4\left(\frac{\sqrt{c}}{2}x\right)\left[\cosh\left(\sqrt{c}x\right)-2\right] + \cdots\,. \end{aligned} \quad (63)$$

The EFDM for Eq. (34)

The finite difference methods are the most frequently used and universally applicable. These methods are approximate in the sense that the derivatives at a point are approximated by difference quotients over a small interval [114]. In order to obtain a finite difference replacement of the KdV Eq. (34) the region to be examined is divided into equal rectangular meshes with sides Δx and Δt parallel to the x- and t-axes respectively. The function $u(x, t)$ is approximated by $u_i^n = u(i\Delta x, n\Delta t)$ where i and n are integers and $i = n = 0$ is the origin. Let us define

$$U_t\mid_{(i,n)} = \frac{U_i^{n+1} - U_i^{n-1}}{2\Delta t}, \tag{64}$$

$$U_x\mid_{(i,n)} = \frac{U_{i+1}^n - U_{i-1}^n}{2\Delta x}, \tag{65}$$

$$U_{xxx}\mid_{(i,n)} = \frac{1}{2(\Delta x)^3}\left[U_{i+2}^n - 2U_{i+1}^n + 2U_{i-1}^n - U_{i-2}^n\right]. \tag{66}$$

The explicit scheme computes the value of the numerical solution at the forward time step in terms of known values at the previous time step. An explicit scheme for solving the KdV equation produced originally by Zabusky and Galvin [11] is centered in time and space. Substituting (64)–(66) into the KdV Eq. (34) with

$$U\mid_{(i,n)} = \frac{1}{3}\left(U_{i-1}^n + U_i^n + U_{i+1}^n\right), \tag{67}$$

leads to

$$\begin{aligned} U_i^{n+1} &= U_i^{n-1} + 2\Delta t\left(U_{i-1}^n + U_i^n + U_{i+1}^n\right)\left(U_{i+1}^n - U_{i-1}^n\right) \\ &\quad - \frac{\Delta t}{(\Delta x)^3}\left(U_{i+2}^n - 2U_{i+1}^n + 2U_{i-1}^n - U_{i-2}^n\right). \end{aligned} \tag{68}$$

For the initial step, we use a scheme which is forward in time and centered in space

$$\begin{aligned} U_i^1 &= U_i^0 + \Delta t\left(U_{i-1}^0 + U_i^0 + U_{i+1}^0\right)\left(U_{i+1}^0 - U_{i-1}^0\right) \\ &\quad - \frac{\Delta t}{(\Delta x)^3}\left(U_{i+2}^0 - 2U_{i+1}^0 + 2U_{i-1}^0 - U_{i-2}^0\right). \end{aligned} \tag{69}$$

It is clear that Eq. (68) is a three-level scheme of time, i. e. in order to obtain U_i at the time level $n+1$, we need the following values of U_{i-2}, U_{i-1}, U_{i+1} and U_{i+2} at the previous time level n in addition to the value of U_i at the time level $n-1$. The explicit difference scheme (68) has second order accuracy in Δt and Δx as the truncation error is $O(\Delta t)^2 + O(\Delta x)^2$ and is also consistent with KdV equation. A stability analysis of the nonlinear numerical scheme (68) using Fourier mode method is not easy to handle unless it is assumed that U, in the nonlinear term, is locally constant. This is equivalent to replacing the term (67) in Eq. (68) by U^*. This linearized scheme for the KdV Eq. (34) has stability condition [115]:

$$\frac{\Delta t}{\Delta x}\left[-6\left|U^*\right| + \frac{4}{(\Delta x)^2}\right] \le 1. \tag{70}$$

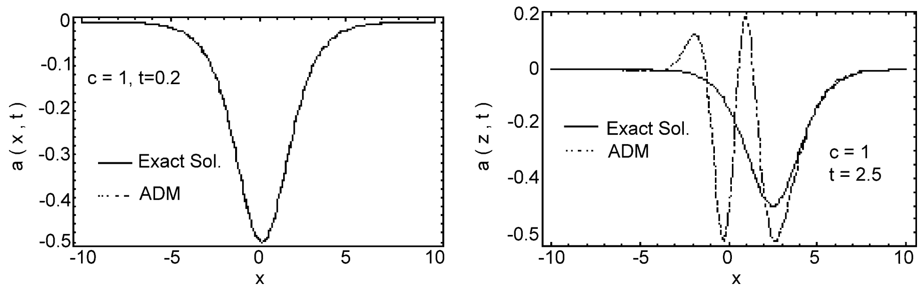
In Table 7, h is x-direction mesh size and k is t-direction time step.

Conclusions and Discussions

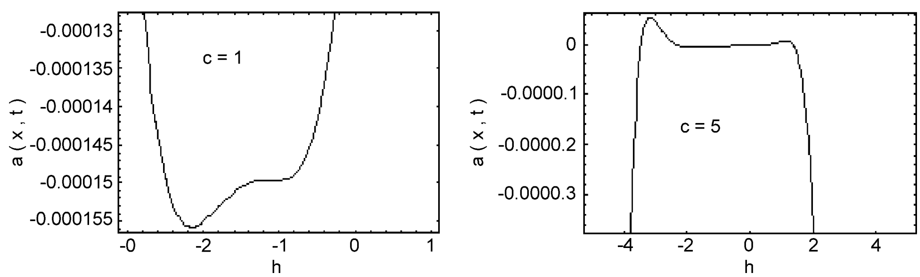
In this paper, the ADM, HAM, VIM, HPM and EFDM were used for the KdV Eq. (34) with initial conditions. We made comparisons of the numerical results obtained by using ADM, HAM, VIM, HPM and EFDM. The approximate solutions to the equation has been also calculated by using the ADM, HAM, VIM, HPM and EFDM without any need to use transformation techniques and linearization of the equation. The ADM, HAM and HPM avoids the difficulties and massive computational work by determining analytic solutions of the nonlinear equations. The main advantage of HAM, VIM and HPM are to overcome the difficulty arising in calculating Adomian's polynomials in the ADM. But the main disadvantage of VIM and EFDM, in general, very big terms and the consuming time to compute it is big, so we need a large computer memory and time. In addition, it is difficult and takes a lot of time to calculate the terms after the fifth term in the VIM for the KdV Eq. (14). It is convenient to examine Tables 1, 2, 3, 4, 5, 6, 7 and Figs. 1, 3, 4, 5 for comparing exact solution and numerical solutions which one gets by using these five methods. Tables and figures are calculated for these methods by using five terms and $c = 1$, except Tables 2 and 4. The results which are obtained by the EFDM are better than the results which are obtained by the other four (semi-inverse) methods for five terms (see Tables 1, 3, 5, 6, 7. If you increase the numbers of terms in the semi-inverse methods, we can see that the results are better (see Tables 2 and 4. Obtaining serial solutions with the semi-inverse methods are easier and more efficient than pure numerical methods (the finite difference methods, the finite element methods, B-spline methods, etc.). Finding numerical results by the EFDM are difficult and take time. The present methods give nearly the same results for $c = 1$ and large values of x but the HAM is the more sensitive method than the others. Because it is possible to obtain a rapidly convergent solution for large and small values of x and t by choosing convenient helper the auxiliary parameter $\hbar$. For this it is enough to consider Figs. 1, 3, 4, 5. Finally, we point out that, for given equations with initial value, the corresponding analytical and numerical solutions are obtained according to the recurrence equations by using Mathematica.

Future Directions

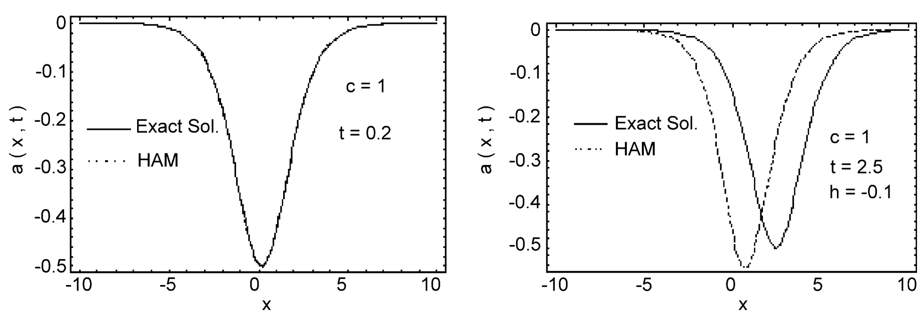
The Korteweg–de Vries equation, which is the most important equation of mathematical physics, has solitons and



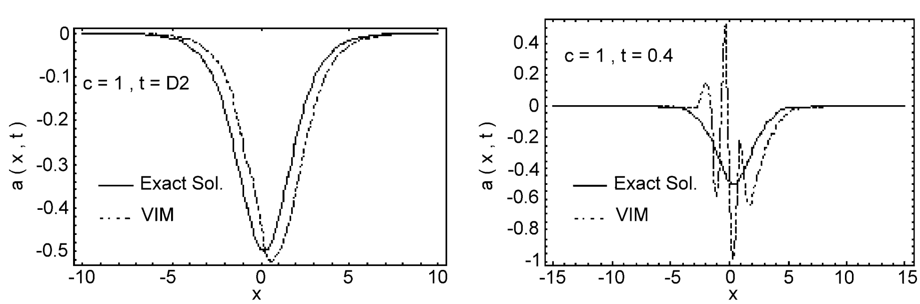
Korteweg–de Vries Equation (KdV), Some Numerical Methods for Solving the, Figure 1
The comparison of the ADM (φ_5) and the exact solutions for Eq. (34) at $c = 1$



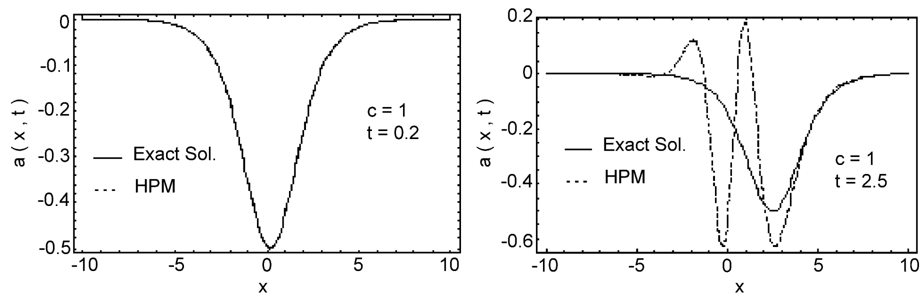
Korteweg–de Vries Equation (KdV), Some Numerical Methods for Solving the, Figure 2
The $\hbar$-curves for $x = 10$, $t = 0.5$, $c = 1$ for 5th-order approximation and $x = 10$, $t = 0.5$, $c = 5$ for 10th-order approximation of $u(x, t)$, respectively



Korteweg–de Vries Equation (KdV), Some Numerical Methods for Solving the, Figure 3
The comparison of the HAM (5th-order approximation) and the exact solutions for Eq. (34) at $c = 1$



Korteweg–de Vries Equation (KdV), Some Numerical Methods for Solving the, Figure 4
The comparison of the VIM (five iterations) and the exact solutions for Eq. (34) at $c = 1$



Korteweg–de Vries Equation (KdV), Some Numerical Methods for Solving the, Figure 5
The comparison of the HPM (five iterations) and the exact solutions for Eq. (34) at $c = 1$

Korteweg–de Vries Equation (KdV), Some Numerical Methods for Solving the, Table 1
Error between the ADM (φ_5) and exact solutions for the KdV Eq. (14) at $c = 1$

t/x	10	15	20	25	30
0.1	2.00679×10^{-4}	1.35233×10^{-6}	9.11171×10^{-9}	6.13942×10^{-11}	4.13671×10^{-13}
0.2	2.22178×10^{-4}	1.49452×10^{-6}	1.00712×10^{-8}	6.78512×10^{-11}	4.57177×10^{-13}
0.3	2.45102×10^{-4}	1.65169×10^{-6}	1.11292×10^{-8}	7.49865×10^{-11}	5.05255×10^{-13}
0.4	2.70872×10^{-4}	1.82535×10^{-6}	1.22991×10^{-8}	8.28713×10^{-11}	5.58381×10^{-13}
0.5	2.99336×10^{-4}	2.01722×10^{-6}	1.35919×10^{-8}	9.15815×10^{-11}	6.17071×10^{-13}

Korteweg–de Vries Equation (KdV), Some Numerical Methods for Solving the, Table 2
Error between the ADM (φ_{10}) and exact solutions for the KdV Eq. (34) at $c = 5$

t/x	10	15	20	25	30
0.1	2.00679×10^{-8}	1.35233×10^{-13}	9.11171×10^{-18}	6.13942×10^{-23}	4.1367×10^{-28}
0.2	2.22178×10^{-8}	1.49452×10^{-12}	1.00712×10^{-18}	6.78512×10^{-23}	4.5717×10^{-27}
0.3	2.45102×10^{-7}	1.65169×10^{-12}	1.11292×10^{-17}	7.49865×10^{-22}	5.0525×10^{-27}
0.4	2.70872×10^{-7}	1.82535×10^{-12}	1.22991×10^{-17}	8.28713×10^{-22}	5.5838×10^{-26}
0.5	2.99336×10^{-6}	2.01722×10^{-11}	1.35919×10^{-16}	9.15815×10^{-21}	6.1707×10^{-26}

Korteweg–de Vries Equation (KdV), Some Numerical Methods for Solving the, Table 3
Error between the HAM (5th-order approximation) and exact solutions for the KdV Eq. (34) at $c = 1$ and $\hbar = -0.5$

t/x	10	15	20	25	30
0.1	1.94277×10^{-4}	1.30915×10^{-6}	8.82101×10^{-9}	5.94355×10^{-11}	4.00473×10^{-13}
0.2	2.08022×10^{-4}	1.40179×10^{-6}	9.44517×10^{-9}	6.36412×10^{-11}	4.28811×10^{-13}
0.3	2.22925×10^{-4}	1.50222×10^{-6}	1.01219×10^{-8}	6.82009×10^{-11}	4.59534×10^{-13}
0.4	2.39102×10^{-4}	1.61125×10^{-6}	1.08565×10^{-8}	7.31508×10^{-11}	4.92886×10^{-13}
0.5	2.56683×10^{-4}	1.72975×10^{-6}	1.16549×10^{-8}	7.85304×10^{-11}	5.29134×10^{-13}

Korteweg–de Vries Equation (KdV), Some Numerical Methods for Solving the, Table 4
Error between the HAM (10th-order approximation) and exact solutions for the KdV Eq. (34) at $c = 5$ and $\hbar = 0.5$

t/x	10	15	20	25	30
0.1	8.12684×10^{-9}	1.13334×10^{-13}	1.58053×10^{-18}	2.20415×10^{-23}	3.0738×10^{-28}
0.2	3.62126×10^{-8}	5.05121×10^{-13}	7.04272×10^{-18}	9.82156×10^{-23}	1.3696×10^{-27}
0.3	8.86781×10^{-7}	1.23668×10^{-12}	1.72463×10^{-17}	2.40512×10^{-22}	3.3541×10^{-27}
0.4	1.87306×10^{-7}	2.61211×10^{-12}	3.64277×10^{-17}	5.08009×10^{-22}	7.0845×10^{-27}
0.5	5.02718×10^{-7}	7.01076×10^{-12}	9.77698×10^{-17}	1.36347×10^{-21}	1.9014×10^{-26}

Korteweg–de Vries Equation (KdV), Some Numerical Methods for Solving the, Table 5
Error between the VIM (5 terms) and exact solutions for the KdV Eq. (34) at $c = 1$

t/x	10	15	20	25	30
0.1	2.33263×10^{-4}	1.57189×10^{-6}	1.05913×10^{-8}	7.13638×10^{-11}	4.24233×10^{-13}
0.2	2.99151×10^{-4}	2.01592×10^{-6}	1.35833×10^{-8}	9.15216×10^{-11}	6.16666×10^{-13}
0.3	3.81873×10^{-4}	2.57322×10^{-6}	1.73383×10^{-8}	1.16824×10^{-10}	7.87156×10^{-13}
0.4	4.84308×10^{-4}	3.26302×10^{-6}	2.19864×10^{-8}	1.48141×10^{-10}	9.98164×10^{-13}
0.5	6.09593×10^{-4}	4.10597×10^{-6}	2.76657×10^{-8}	1.86417×10^{-10}	1.25602×10^{-12}

Korteweg–de Vries Equation (KdV), Some Numerical Methods for Solving the, Table 6
Error between the HPM (five iterations) and exact solutions for the KdV Eq. (34) at $c = 1$

t/x	10	15	20	25	30
0.1	2.00679×10^{-4}	1.35232×10^{-6}	9.11171×10^{-9}	6.13942×10^{-11}	4.13671×10^{-13}
0.2	2.21782×10^{-4}	1.49452×10^{-6}	1.00734×10^{-8}	6.78513×10^{-11}	4.57177×10^{-13}
0.3	2.45102×10^{-4}	1.65169×10^{-6}	1.11296×10^{-8}	7.49865×10^{-11}	5.05255×10^{-13}
0.4	2.70871×10^{-4}	1.82535×10^{-6}	1.22991×10^{-8}	8.28716×10^{-11}	5.58381×10^{-13}
0.5	2.99337×10^{-4}	2.01722×10^{-6}	1.35919×10^{-8}	9.15815×10^{-11}	6.17071×10^{-13}

Korteweg–de Vries Equation (KdV), Some Numerical Methods for Solving the, Table 7
Error between the EFDM and exact solutions for the KdV Eq. (34) at $c = 1$, $h = 0.05$ and $k = 0.025$

t/x	10	15	20	25	30
0.1	6.15156×10^{-6}	4.14842×10^{-8}	2.79520×10^{-10}	1.88339×10^{-12}	1.26902×10^{-14}
0.2	7.54177×10^{-6}	5.09160×10^{-8}	3.43074×10^{-10}	2.31161×10^{-12}	1.55755×10^{-14}
0.3	9.14755×10^{-6}	6.19126×10^{-8}	4.17173×10^{-10}	1.70489×10^{-12}	1.89396×10^{-14}
0.4	1.10086×10^{-5}	7.46977×10^{-8}	5.03358×10^{-10}	3.39162×10^{-12}	1.38607×10^{-14}
0.5	1.35102×10^{-5}	8.94343×10^{-8}	6.03380×10^{-10}	4.06554×10^{-12}	2.73934×10^{-14}

solitary wave solutions. It was solved by using analytical and numerical methods by many authors. We have mentioned these analytical and numerical methods in the introduction. The theoretical studies on the KdV equation have been nearly completed, but we may find new soliton solutions by producing new analytical methods or expanding the old methods. It is also possible to perform same operations on numerical methods. We may show the interaction of KdV solitons with each other by numerical.

Bibliography

Primary Literature

1. Russell JS (1844) Reports on waves. Report of the 14th Meeting of the British Association for the Advancement of Science, London, pp 311–390
2. 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. Phil Mag 39:422–443
3. Fermi E, Pasta J, Ulam S (1955) Studies of nonlinear problems I. Los Alamos Report LA-1940, Los Alamos, New Mexico
4. Zabusky NJ, Kruskal MD (1965) Interaction of solitons in a collisionless plasma and the recurrence on initial states. Phys Rev Lett 15:240–243
5. Zabusky NJ (1967) A synergetic approach to problems of nonlinear dispersive wave propagation and interaction. In: Ames WF (ed) Proceedings of the Symposium on Nonlinear Partial Differential Equations. Academic Press, Boston, pp 223–258
6. Lax PD (1968) Integrals of nonlinear equations of evolution and solitary waves. Commun Pure Appl Math 21:467–490
7. Gardner CS, Greene JM, Kruskal MD, Miura RM (1967) Method for solving the KdV equation. Phys Rev Lett 19:1095–1097
8. Gardner CS, Greene JM, Kruskal MD, Miura RM (1974) Korteweg–de Vries equation and generalizations VI methods for exact solutions. Commun Pure Appl Math 27:97–133
9. Hirota R (1971) Exact solution of the Korteweg–de Vries equation for multiple collisions of solitons. Phys Rev Lett 27:1192–1194
10. Hirota R (1973) Exact envelope-soliton solutions of a nonlinear wave equation. J Math Phys 14:805–809
11. Zabusky NJ, Galvin CJ (1971) Shallow water waves, the Korteweg–de Vries equation and solitons. J Fluid Mech 47:811–824
12. Hammack JL, Segur H (1974) The Korteweg–de Vries equation and water waves, Part 2, comparison with experiments. J Fluid Mech 65:289–314
13. Weidman PD, Maxworthy T (1978) Experiments on strong interactions, between solitary waves. J Fluid Mech 85:417–431
14. Dodd RK, Eilbeck JC, Gibbon JD, Morris HC (1982) Solitons and Nonlinear Wave Equations. Academic Press, New York

15. Lombdahl PS (1984) What is a soliton? Los Alamos Sci 10:27–31
16. Debnath L (1997) Nonlinear Partial Differential Equations for Scientists and Engineers. Birkhauser, Boston
17. Helal MA (2002) Soliton solution of some nonlinear partial differential equations and its applications in fluid mechanics. Chaos Solitons Fractals 13:1917–1929
18. Davidson RC (1972) Methods in Nonlinear Plasma Theory. Academic Press, New York
19. Lamb GL (1980) Elements of Soliton Theory. Wiley, New York
20. Drazin PG, Johnson RS (1989) Solitons: An Introduction. Cambridge University Press, New York
21. Drazin PG (1992) Nonlinear Systems. Cambridge University Press, New York
22. Khater AH, Helal MA, El-Kalaawy OH (1998) Backlund transformations: exact solutions for the KdV and Calogero-Degasperis-Fokas mKdV equations. Math Meth Appl Sci 21:719–731
23. Khater AH, El-Sabbagh MF (1989) The pain level property and coordinates transformations. Nuovo Cimento B 104:123–129
24. Wahlquist HD, Estabrook FB (1973) Backlund transformation for solutions of the KdV equation. Phys Rev Lett 31:1386–1390
25. Olver PJ (1986) Applications of Lie Groups to Differential equations Graduate Texts in Mathematics. Springer, Berlin
26. Hereman W, Korpel A, Banerjee PP (1985) A general physical approach to solitary wave construction from linear solutions. Wave Motion 7:283–289
27. Malfliet W (1992) Solitary wave solutions of nonlinear wave equations. Am J Phys 60:650–654
28. Yan CT (1996) New explicit solitary wave solutions and periodic wave solutions for Whitham-Broer-Kaup equation in shallow water. Phys Lett 224:77–84
29. Wang ML (1996) Exact solutions for a compound KdV-Burgers equation. Phys Lett 213(5–6):279–287
30. Yan ZY, Zhang HQ (2001) A simple transformation forn nonlinear waves. Phys Lett 285:355–362
31. Peng YZ (2003) Exact solutions for some nonlinear partial differential equations. Phys Lett 314(5–6):401–408
32. Sawada K, Kotera T (1974) A method for finding N-solitons of the KdV equation and KdV-like equation. Prog Theor Phys 51:1355–1367
33. Rosales RR (1978) Exact solutions of some nonlinear evolution equations. Stud Appl Math 59:117–157
34. Whitham GB (1927) Linear and Nonlinear Waves. Pure and Applied Mathematics. Wiley/Interscience, New York
35. Wadati M, Sawada K (1980) New representations of the soliton solution for the KdV equation. J Phys Soc Jpn 48:312–318
36. Wadati M, Sawada K (1980) Application of the trace method to the modified KdV equation. J Phys Soc Jpn 48:319–326
37. Helal MA, El-Eissa HN (1996) Shallow water waves and KdV equation. PUMA 7:263–282
38. Khater AH, El-Kalaawy, Helal MA (1997) Two new classes of exact solutions for the KdV equation via Backlund transformations. Chaos Solitons Fractals 8:1901–1909
39. Khater AH, Helal MA, Seadawy AR (2000) General soliton solutions of an n-dimensional nonlinear Schrödinger equations. Nuovo Cimento B 115:1303–1311
40. Helal MA (2001) Chebyshev spectral method for solving KdV equation with hydro-dynamical application. Chaos Solitons Fractals 12:943–950
41. Das G, Sarma J (1999) A new mathematical approach for finding the solitary waves in dusty plasma. Phys Plasmas 6:4394–4397
42. Osborne A (1995) The inverse scattering transform: Tools for the nonlinear fourier analysis and filtering of ocean surfaces waves. Chaos Solitons Fractals 5:2623–2637
43. Ostrovsky L, Stepanyants YA (1989) Do internal solitons exist in the ocean? Rev Geophys 27:293–310
44. Ludu A, Draayer JP (1998) Nonlinear modes of liquid drops as solitary waves. Phy Rev Lett 80:2125–2128
45. Reatto L, Galli D (1999) What is a ROTON? Int J Modern Phys B 13:607–616
46. Turitsyn S, Aceves A, Jones C, Zharnitsky V (1998) Average dynamics of the optical soliton in communication lines with dispersion management: Analytical results. Phys Rev E 58:48–51
47. Coffey MW (1996) Nonlinear dynamics of vortices in ultraclean type-II superconductors: Integrable wave equations in cylindrical geometry. Phys Rev B 54:1279–1285
48. Gou BY, Shen J (2001) On spectral approximations using modified Legendre rational functions: Application to the KdV equation on the half line. Indiana Univ Math J 50:181–204
49. Huang WZ, Sloan DM (1992) The pseudo spectral method for third-order differential equations. SIAM J Numer Anal 29:1626–1647
50. Man HP, Sun WW (2000) A Legendre-Petrov-Galerkin and Chebyshev collocation methods for third-order differential equations. SIAM J Numer Anal 38:1425–1438
51. Man HP, Sun WW (2001) Optimal error estimates of the Legendre-Petrov-Galerkin method for the KdV equation. SIAM J Numer Anal 39:1380–1394
52. Bahadır AR (2005) Exponential finite-difference method applied to the KdV equation for small times. Appl Math Comput 160:675–682
53. Jain PC, Shankar R, Bhardwaj D (1997) Numerical solution of the KdV equation. Chaos Solitons Fractals 8:943–951
54. Soliman AA (2004) Collocation solution of the KdV equation using septic splines. Int J Comput Math 81:325–331
55. Frauendiener J, Klein C (2006) Hyperelliptic theta-functions and spectral methods: KdV and KP solutions. Lett Math Phys 76:249–267
56. Bhatta DD, Bhatti MI (2006) Numerical solution of KdV equation using modified Bernstein polynomials. Appl Math Comput 174:1255–1268
57. Helaln MA, Mehanna MS (2007) A comparative study between two different methods for solving the GKdV equation. Chaos Solitons Fractals 33:729–739
58. Kutluay S, Bahadır AR, Özdeş A (2000) A small time solutions for the KdV equation. Appl Math Comput 107:203–210
59. Özer S, Kutluay S (2005) An analytical numerical method for solving the KdV equation. Appl Math Comput 164:789–797
60. Dehghan M, Shokri A (2007) A numerical method for KdV equation using collocation and radial basis functions. Nonlinear Dyn 50:111–120
61. Dağ I, Dereli Y (2008) Numerical solutions of KdV equation using radial basis functions. Appl Math Model 32:535–546
62. Wadati M (1973) The modified Kortweg–de Vries equation. J Phys Soc Jpn 34:1289–1296
63. Wadati M (2001) Introduction to solitons. Pramana J Phys 57(5–6):841–847
64. Adomian G (1989) Nonlinear Stochastic Systems and Applications to Physics. Kluwer, Boston

65. Adomian G (1994) Solving Frontier Problems of Physics: The Decomposition Method. Kluwer, Boston
66. Adomian G (1998) A review of the decomposition method in applied mathematics. J Math Anal Appl 135:501–544
67. Lesnic D, Elliott L (1999) The decomposition approach to inverse heat conduction. J Math Anal Appl 232:82–98
68. Wazwaz AM (2002) Partial Differential Equations: Methods and Applications. Balkema Publication, Lisse
69. Wazwaz AM (2003) An analytical study on the third-order dispersive partial differential equation. Appl Math Comput 142:511–520
70. Cherruault Y (1998) Modéles et Méthodes Math ématiques pour les Sciences du Vivant. Presses Universitaires de France, Paris
71. Dehghan M (2004) Application of ADM for two-dimensional parabolic equation subject to nonstandard boundary specification. Appl Math Comput 157:549–560
72. Dehghan M (2004) The solution of a nonclassic problem for one-dimensional hyberbolic equation using the decomposition procedure. Int J Comput Math 81:979–989
73. Babolian E, Biazar J, Vahidi AR (2004) The decomposition method applied to systems of Fredholm integral equations of the second kind. Appl Math Comput 148:443–452
74. El-Sayed MS (2002) The modified decomposition method for solving nonlinear algebraic equations. Appl Math Comput 132:589–597
75. Inc M (2006) On numerical Jacobi elliptic function solutions of (1+1)-dimensional dispersive long wave equation by the decomposition method. Appl Math Comput 173:372–382
76. Inc M (2007) Exact solutions with solitary patterns for the Zakharov-Kuznetsov equations with fully nonlinear dispersion. Chaos Solitons Fractals 33:1783–1790
77. Cherruault Y (1990) Convergence of Adomian's decomposition method. Math Comput Modelling 14:83–86
78. Ngarhasta N, Some B, Abboui K, Cherruault Y (2002) New numerical study of Adomian method applied to a diffusion model. Kybernetes 31:61–75
79. Mavoungou T, Cherruault Y (1992) Convergence of Adomian's method and applications to nonlinear partial differential equations. Kybernetes 21:13–25
80. Inc M (2006) On numerical soliton solution of the KK equation and convergence analysis of the decomposition method. Appl Math Comput 172:72–85
81. Hashim I, Noorani MS, Al-Hadidi MRS (2006) Solving the generalized Burgers-Huxley equation using the Adomian decomposition method. Math Comput Model 43:1404–1411
82. Liao SJ (2002) An analytic approximation of the drag coefficient for the viscous flow past a sphere. Int J Nonlinear Mech 37:1–18
83. Liao SJ (2003) Beyond Perturbation: Introduction to the Homotopy Analysis Method. Champan, Boca Raton
84. Liao SJ (2003) On the analytic solution of magnetohydrodynamic flows non-Newtonian fluids over a stretching sheet. J Fluid Mech 488:189–212
85. Liao SJ (2005) A new branch of solutions of boundary-layer flows over an impermeable stretched plate. Int J Heat Mass Transfer 48:2529–2539
86. Sajid M, Hayat T, Asghar S (2006) On the analytic solution of the steady flow of a fourth grade fluid. Phys Lett 355:18–26
87. Hayat T, Abbas Z, Sajid M (2006) Series solution for the upper-convected Maxwell fluid over a porous stretching plate. Phys Lett 358:396–403
88. Hayat T, Khan M (2005) Homotopy solutions for a generalized second-grade fluid past a porous plate. Nonlinear Dynamics 42:395–405
89. Abbasbandy S (2006) The application of homotopy analysis method to nonlinear equations arising in heat transfer. Phys Lett 360:109–113
90. Abbasbandy S (2007) The application of homotopy analysis method to solve a generalized Hirota-Satsuma coupled KdV equation. Phys Lett 361:478–483
91. Inc M (2007) On exact solution of Laplace equation with Dirichlet and Neumann boundary conditions by the homotopy analysis method. Phys Lett 365:412–415
92. He JH (1997) A new approach to nonlinear partial differential equations. Commun Nonlinear Sci Numer Simul 2:230–235
93. He JH (1997) Variational iteration method for delay differential equations. Commun Nonlinear Sci Numer Simul 2:235–236
94. He JH (1999) Variational iteration method a kind of nonlinear analytical technique: Some examples. Int J Nonlinear Mech 34:699–708
95. He JH (2000) Variational iteration method for autonomous ordinary differential systems. Appl Math Comput 114:115–123
96. He JH (2006) Some asymptotic methods for strongly nonlinear equations. Int J Modern Phys B 20:1141–1199
97. Abdou MA, Soliman AA (2005) Variational iteration method for solving Burger's and coupled Burger's equations. J Comput Appl Math 181:245–251
98. Abdou MA, Soliman AA (2005) New applications of variational iteration method. Phys D 211:1–8
99. Bildik N, Konuralp A (2006) The use of variational iteration method, differential transform method and Adomian decomposition method for solving different types of nonlinear partial differential equations. Int J Nonlinear Sci Numer Simul 7:65–70
100. Draganescu G (2006) Application of a variational iteration method to linear and nonlinear viscoelastic models with fractional derivatives. J Math Phys 47:082902
101. Tatari M, Dehghan M (2007) He's variational iteration method for computing a control parameter in a semi-linear inverse parabolic equation. Chaos Solitons Fractals 33:671–677
102. Wazwaz AM (2007) The variational iteration method for exact solutions of Laplace equation. Phys Lett 363:260–262
103. Tatari M, Dehghan M (2007) Solution of problems in calculus of variations via He's variational iteration method. Phys Lett 362:401–406
104. Inc M (2007) An approximate solitary wave solution with compact support for the modified KdV equation. Appl Math Comput 184:631–637
105. Inc M (2007) Exact and numerical solutions with compact support for nonlinear dispersive K(m,p) equations by the variational iteration method. Physica 375:447–456
106. He JH (1999) Homotopy perturbation technique. Comput Meth Appl Mech Eng 178:257–262
107. He JH (2005) Application of homotopy perturbation method to nonlinear wave equations. Chaos Solitons Fractals 26:695–700
108. He JH (2005) Homotopy perturbation method for bifurcation of nonlinear problems. Int J Nonlinear Sci Numer Simul 6:207–208

109. He JH (2006) Non-perturbative Methods for Strongly Nonlinear Problems. Dissertation. de-Verlag im Internet GmbH, Berlin
110. Ganji DD, Rajabi A (2006) Assesment of homotopy perturbation and perturbation methods in heat radiation equations. Int Commun Heat Mass Transf 33:391–400
111. Siddiqui AM, Mahmood R, Ghori QK (2006) Homotopy perturbation method for thin film flow of a fourth grade fluid down a vertical cylinder. Phys Lett 352:404–410
112. Ganji DD, Rafei M (2006) Solitary wave solutions for a generalized Hirota-Satsuma coupled KdV equation by homotopy perturbation method. Phys Lett 356:131–137
113. Chowdhury MSH, Hashim I (2007) Solutions of a class of singular second-order IVPs by homotopy-perturbation method. Phys Lett 356(5–6):439–447
114. Smith GD (1987) Numerical Solution of Partial Differential Equation: Finite Difference Methods. Oxford University Press, New York
115. Vilegenthart AC (1971) On finite difference methods for the Korteweg–de Vries equations. J Eng Math 5(2):137–155

Books and Reviews

Abbasbandy S (2007) An approximation solution of a nonlinear equation with Riemann–Liouville's fractional derivatives by He's variational iteration method. J Comp Appl Math 207: 53–58

Abbasbandy S (2007) A new application of He's variational iteration method for quadratic Riccati differential equation by using Adomian's polynomials. J Comp Appl Math 207:59–63

Deeba E, Khuri SA, Xie S (2000) An algorithm for solving boundary value problems. J Comp Phys 159:125–138

El-Sayed SM (2002) The modified decomposition method for solving nonlinear algebraic equations. Appl Math Comput 132:589–597

Ganji DD, Sadighi A (2007) Application of homotopy-perturbation and variational iteration methods to nonlinear heat transfer and porous media equations. J Comp Appl Math 207:24–34

Hayat T, Sajid M (2007) Analytic solution for axisymmetric flow and heat transfer of a second grade fluid past a stretching sheet. Int J Heat Mass Transf 50:75–84

Kaya D (2005) An application for the higher order modified KdV equation by decomposition method. Commun Nonlinear Sci Numer Simul 10(6):693–702

Kaya D, Aassila M (2002) An application for a generalized KdV equation by the decomposition method. Phys Lett 299(2):201–206

Khelifa S, Cherruault Y (2000) New results for the Adomian method. Kybernetes 29(3):332–354

Öziş T, Yıldırım A (2007) A note on He's homotopy perturbation method for van der Pol oscillator with very strong nonlinearity. Chaos Solitons Fractals 34(3):989–991

Wang C, Wu Y, Wu W (2005) Solving the nonlinear periodic wave problems with the Homotopy Analysis Method. Wave Motion 41(4):329–337

Wazwaz AM (1997) A First Course in Integral Equations. World Scientific, London

Wu Y, Wang C, Liao S (2005) Solving the one-loop soliton solution of the Vakhnenko equation by means of the Homotopy analysis method. Chaos Solitons Fractals 23(5):1733–1740