5

Advantages, limitations, and open questions

As all things have their good and bad sides, so too does the homotopy analysis method. Here, we make some discussions about the advantages and limitations of this method and point out some open questions.

5.1 Advantages

Compared with perturbation techniques and nonperturbation methods such as Lyapunov's artificial small parameter method, the δ-expansion method, and Adomian's decomposition method, the homotopy analysis method has some or all of the following advantages.

Firstly, unlike all previous analytic techniques, the homotopy analysis method provides us with great freedom to express solutions of a given nonlinear problem by means of different base functions. Therefore, we can approximate a nonlinear problem more efficiently by choosing a proper set of base functions. This is because the convergence region and rate of a series are chiefly determined by the base functions used to express the solution.

Secondly, unlike all previous analytic techniques, the homotopy analysis method always provides us with a family of solution expressions in the auxiliary parameter $\hbar$, even if a nonlinear problem has a unique solution. The convergence region and rate of each solution expression among the family might be determined by the auxiliary parameter $\hbar$. So, the auxiliary parameter $\hbar$ provides us with an additional way to conveniently adjust and control the convergence region and rate of solution series. By means of the so-called $\hbar$-curves it is easy to find out the so-called valid regions of $\hbar$ to gain a convergent solution series. In addition, the so-called homotopy-Padé technique is often more efficient than the traditional Padé technique and is in some cases even independent of the auxiliary parameter $\hbar$.

Thirdly, unlike perturbation techniques, the homotopy analysis method is independent of any small or large quantities. So, the homotopy analysis method can be applied no matter if governing equations and boundary/initial conditions of a given nonlinear problem contain small or large quantities or not.

Finally, the homotopy analysis method logically contains Lyapunov's artificial small parameter method, the δ-expansion method, and Adomian's decomposition method, and therefore unifies these nonperturbation methods and is more general than them.

It should be pointed out that the homotopy analysis method is based on the following assumptions:

(A) There exists the solution of the zero-order deformation equation in the whole region of the embedding parameter $q \in [0, 1]$.

(B) All of the high-order deformation equations have solutions.

(C) All Taylor series expanded in the embedding parameter q converge at $q = 1$.

Fortunately, the homotopy analysis method provides us with great freedom to choose initial approximation, the auxiliary linear operator, the auxiliary function, and the auxiliary parameter $\hbar$. This kind of freedom provides us with the great possibility to ensure that all of these assumptions may be satisfied. So, the above assumptions do little damage to the homotopy analysis method. In fact, nearly all of the above-listed advantages of the homotopy analysis method come from such kinds of freedom. According to Theorem 3.1 and Theorem 3.3, as long as a solution series given by the homotopy analysis method converges, it must be one of the solutions of a given nonlinear problem. Thus, we need only focus on choosing proper initial approximations, auxiliary linear operators, auxiliary functions, and proper values of $\hbar$ to ensure that solution series converge. Therefore, it is this kind of freedom that establishes a cornerstone of the validity and the flexibility of the homotopy analysis method.

5.2 Limitations

However, such kinds of freedom seem too great for us. Up to now, there are no rigorous theories to direct us to choose the initial approximations, auxiliary linear operators, auxiliary functions, and auxiliary parameter $\hbar$. From the practical viewpoints, we propose some fundamental rules such as the rule of solution expression, the rule of coefficient ergodicity, and the rule of solution existence, which play important roles within the homotopy analysis method. The rule of solution expression provides us with a starting point. It is under the rule of solution expression that initial approximations, auxiliary linear operators, and the auxiliary functions are determined. The rule of coefficient ergodicity and the rule of solution existence play important roles in determining the auxiliary function and ensuring that the high-order deformation equations are closed and have solutions.

The rule of coefficient ergodicity is based on the completeness, and the rule of solution existence is straightforward. So, the rule of coefficient ergodicity and the rule of solution existence are reasonable. Unfortunately, the rule of solution expression implies such an assumption that we should have, more or less, some knowledge about a given nonlinear problem *a prior*. How can we get such kind of prior knowledge before we solve a problem that is completely new for us? How can we know that a set of base functions is better than others and is more efficient to approximate a nonlinear problem which we know nothing? So, theoretically, this assumption impairs the homotopy analysis method, although we can always attempt some base functions even if a given nonlinear problem is completely new for us. Fortunately, it seems that solutions of a nonlinear problem could be expressed by many different kinds of base functions, as illustrated in Chapter 2.

As mentioned in Chapter 2, the idea of avoiding the so-called secular term was proposed by a lot of researchers such as Lindstedt [52], Bohlin [53], Poincaré [54], Gyldén [55], and so on, and the rule of solution expression can be regarded as its generalization. However, for a completely new problem, how can we know that a term belongs to the so-called secular term or not? So, in fact, many previous analytic techniques also imply the assumption that some prior knowledge should be known. And this assumption also impairs these methods, although such kinds of damage seem tiny compared to other serious restrictions of these previous methods and thus are often neglected.

5.3 Open questions

To overcome the above-mentioned limitation of the homotopy analysis method, it is necessary to propose some pure mathematical theorems to direct us to choose the initial approximation, the auxiliary linear operator, and the auxiliary function. These mathematical theorems should be valid in rather general cases without any prior knowledge so that we can apply them without any physical backgrounds. Up to now, it is even an open question if such kinds of pure mathematical theorems exist or not.

Although the homotopy analysis method has been successfully applied to many nonlinear problems such as those illustrated in this book and published in some journals, it is unclear if this method is valid for nonlinear problems with discontinue or chaotic solutions. To the best of the author's knowledge, chaos is generally investigated by numerical techniques and hardly expressed analytically. Up to now, it seems not very clear what kind of base functions is efficient to analytically express a chaotic solution. Recently, Norden E. Huang et al. [80] developed the empirical mode decomposition method and showed that nonlinear and nonstationary time series can be expressed by the so-called “instrinic mode functions”. However, up to now, we do not know how to use the “instrinic mode functions” to gain *analytic* expressions of chaotic solutions of a given nonlinear problem without numerically solving it.