3

Systematic description

In Chapter 2 the basic ideas of the homotopy analysis method are illustrated by a simple nonlinear problem. Here a systematic description is given for general nonlinear problems.

3.1 Zero-order deformation equation

In most cases a nonlinear problem can be described by a set of governing equations and initial and/or boundary conditions. For brevity, let us consider here only one nonlinear equation in a general form:

$$\mathcal{N}[u(\mathbf{r},t)] = 0, \tag{3.1}$$

where $\mathcal{N}$ is a nonlinear operator, $u(\mathbf{r},t)$ is an unknown function, and $\mathbf{r}$ and t denote spatial and temporal independent variables, respectively.

Let $u_0(\mathbf{r},t)$ denote an initial guess of the exact solution $u(\mathbf{r},t)$, $\hbar \neq 0$ an auxiliary parameter, $H(\mathbf{r},t) \neq 0$ an auxiliary function, and $\mathcal{L}$ an auxiliary linear operator with the property

$$\mathcal{L}\left[f(\mathbf{r},t)\right] = 0 \qquad \text{when } f(\mathbf{r},t) = 0. \tag{3.2}$$

Then, using $q \in [0,1]$ as an embedding parameter, we construct such a homotopy

$$\begin{aligned} &\mathcal{H}[\Phi(\mathbf{r},t;q); u_0(\mathbf{r},t), H(\mathbf{r},t), \hbar, q] \\ &= (1-q)\left\{\mathcal{L}[\Phi(\mathbf{r},t;q) - u_0(\mathbf{r},t)]\right\} - q\,\hbar\,H(\mathbf{r},t)\,\mathcal{N}[\Phi(\mathbf{r},t;q)]. \end{aligned} \tag{3.3}$$

It should be emphasized that the above homotopy contains the so-called auxiliary parameter $\hbar$ and the auxiliary function $H(\mathbf{r},t)$. To the best of the author's knowledge, the nonzero auxiliary parameter $\hbar$ and auxiliary function $H(\mathbf{r},t)$ are introduced for the first time in this way to construct a homotopy. So, such a kind of homotopy is more general than traditional ones. The auxiliary parameter $\hbar$ and the auxiliary function $H(\mathbf{r},t)$ play important roles within the frame of the homotopy analysis method. It should be emphasized that we have great freedom to choose the initial guess $u_0(\mathbf{r},t)$, the auxiliary linear operator $\mathcal{L}$, the nonzero auxiliary parameter $\hbar$, and the auxiliary function $H(\mathbf{r}, t)$.

Let $q \in [0, 1]$ denote an embedding parameter. Enforcing the homotopy (3.3) to be zero, i.e.,

$$\mathcal{H}[\Phi(\mathbf{r}, t; q); u_0(\mathbf{r}, t), H(\mathbf{r}, t), \hbar, q] = 0,$$

we have the so-called zero-order deformation equation

$$(1 - q) \{\mathcal{L}[\Phi(\mathbf{r}, t; q) - u_0(\mathbf{r}, t)]\} = q\, \hbar\, H(\mathbf{r}, t)\, \mathcal{N}[\Phi(\mathbf{r}, t; q)], \tag{3.4}$$

where $\Phi(\mathbf{r}, t; q)$ is the solution which depends upon not only the initial guess $u_0(\mathbf{r}, t)$, the auxiliary linear operator $\mathcal{L}$, the auxiliary function $H(\mathbf{r}, t)$ and the auxiliary parameter $\hbar$ but also the embedding parameter $q \in [0, 1]$. When $q = 0$, the zero-order deformation equation (3.4) becomes

$$\mathcal{L}[\Phi(\mathbf{r}, t; 0) - u_0(\mathbf{r}, t)] = 0, \tag{3.5}$$

which gives, using the property (3.2),

$$\Phi(\mathbf{r}, t; 0) = u_0(\mathbf{r}, t). \tag{3.6}$$

When $q = 1$, since $\hbar \neq 0$ and $H(\mathbf{r}, t) \neq 0$, the zero-order deformation equation (3.4) is equivalent to

$$\mathcal{N}[\Phi(\mathbf{r}, t; 1)] = 0, \tag{3.7}$$

which is exactly the same as the original equation (3.1), provided

$$\Phi(\mathbf{r}, t; 1) = u(\mathbf{r}, t). \tag{3.8}$$

Thus, according to (3.6) and (3.8), as the embedding parameter q increases from 0 to 1, $\Phi(\mathbf{r}, t; q)$ varies (or deforms) continuously from the initial approximation $u_0(\mathbf{r}, t)$ to the exact solution $u(\mathbf{r}, t)$ of the original equation (3.1). Such a kind of continuous variation is called deformation in homotopy. This is the reason why we call (3.4) the zero-order deformation equation.

Define the so-called mth-order deformation derivatives

$$u_0^{[m]}(\mathbf{r}, t) = \left. \frac{\partial^m \Phi(\mathbf{r}, t; q)}{\partial q^m} \right|_{q=0}. \tag{3.9}$$

By Taylor's theorem, $\Phi(\mathbf{r}, t; q)$ can be expanded in a power series of q as follows:

$$\Phi(\mathbf{r}, t; q) = \Phi(\mathbf{r}, t; 0) + \sum_{m=1}^{+\infty} \frac{u_0^{[m]}(\mathbf{r}, t)}{m!}\, q^m. \tag{3.10}$$

Writing

$$u_m(\mathbf{r}, t) = \frac{u_0^{[m]}(\mathbf{r}, t)}{m!} = \frac{1}{m!} \left. \frac{\partial^m \Phi(\mathbf{r}, t; q)}{\partial q^m} \right|_{q=0} \tag{3.11}$$

and using (3.6), the power series (3.10) of $\Phi(\mathbf{r}, t; q)$ becomes

$$\Phi(\mathbf{r}, t; q) = u_0(\mathbf{r}, t) + \sum_{m=1}^{+\infty} u_m(\mathbf{r}, t)\; q^m. \tag{3.12}$$

Note that we have great freedom to choose the initial guess $u_0(\mathbf{r}, t)$, the auxiliary linear operator $\mathcal{L}$, the nonzero auxiliary parameter $\hbar$, and the auxiliary function $H(\mathbf{r}, t)$. Assume that all of them are properly chosen so that:

1. The solution $\Phi(\mathbf{r}, t; q)$ of the zero-order deformation equation (3.4) exists for all $q \in [0, 1]$.

2. The deformation derivative $u_0^{[m]}(\mathbf{r}, t)$ exists for $m = 1, 2, 3, \cdots, +\infty$.

3. The power series (3.12) of $\Phi(\mathbf{r}, t; q)$ converges at $q = 1$.

Then, from (3.8) and (3.12), we have under these assumptions the solution series

$$u(\mathbf{r}, t) = u_0(\mathbf{r}, t) + \sum_{m=1}^{+\infty} u_m(\mathbf{r}, t). \tag{3.13}$$

This expression provides us with a relationship between the exact solution $u(\mathbf{r}, t)$ and the initial approximation $u_0(\mathbf{r}, t)$ by means of the terms $u_m(\mathbf{r}, t)$ which are determined by the so-called high-order deformation equations described below.

3.2 High-order deformation equation

For brevity, define the vector

$$\vec{u}_n = \{u_0(\mathbf{r}, t), u_1(\mathbf{r}, t), u_2(\mathbf{r}, t), \cdots, u_n(\mathbf{r}, t)\}.$$

According to the definition (3.11), the governing equation of $u_m(\mathbf{r}, t)$ can be derived from the zero-order deformation equation (3.4). Differentiating the zero-order deformation equation (3.4) m times with respective to the embedding parameter q and then dividing it by $m!$ and finally setting $q = 0$, we have the so-called *mth-order deformation equation*

$$\mathcal{L}\left[u_m(\mathbf{r}, t) - \chi_m\, u_{m-1}(\mathbf{r}, t)\right] = \hbar\, H(\mathbf{r}, t)\; R_m(\vec{u}_{m-1}, \mathbf{r}, t), \tag{3.14}$$

where χ_m is defined by (2.42) and

$$R_m(\vec{u}_{m-1}, \mathbf{r}, t) = \frac{1}{(m-1)!} \left. \frac{\partial^{m-1} \mathcal{N}[\Phi(\mathbf{r}, t; q)]}{\partial q^{m-1}} \right|_{q=0}. \tag{3.15}$$

Substituting (3.12) into the above expression, we have

$$R_m(\vec{u}_{m-1}, \mathbf{r}, t) = \frac{1}{(m-1)!} \left\{ \frac{\partial^{m-1}}{\partial q^{m-1}} \mathcal{N} \left[\sum_{n=0}^{+\infty} u_n(\mathbf{r}, t)\, q^n \right] \right\} \Bigg|_{q=0}. \tag{3.16}$$

Note that the high-order deformation equation (3.14) is governed by the same linear operator $\mathcal{L}$, and the term $R_m(\vec{u}_{m-1}, \mathbf{r}, t)$ can be expressed simply by (3.15) for any given nonlinear operator $\mathcal{N}$. According to the definition (3.15), the right-hand side of Equation (3.14) is only dependent upon $\vec{u}_{m-1}$. Thus, we gain $u_1(\mathbf{r}, t), u_2(\mathbf{r}, t), \cdots$ by means of solving the linear high-order deformation equation (3.14) one after the other in order. The mth-order approximation of $u(\mathbf{r}, t)$ is given by

$$u(\mathbf{r}, t) \approx \sum_{k=0}^{m} u_k(\mathbf{r}, t). \tag{3.17}$$

We can construct the zero-order deformation equation in a form even more general than (3.4). Let $A(q), B(q)$ be complex functions analytic in the region $|q| \leq 1$, called the embedding functions, which satisfy

$$A(0) = B(0) = 0, \qquad A(1) = B(1) = 1, \tag{3.18}$$

respectively. Let

$$A(q) = \sum_{k=1}^{+\infty} \alpha_k\, q^k, \qquad B(q) = \sum_{k=1}^{+\infty} \beta_k\, q^k \tag{3.19}$$

denote the Maclaurin series of $A(q)$ and $B(q)$, respectively. Because $A(q)$ and $B(q)$ are analytic in the region $|q| \leq 1$, we have from (3.18) that

$$\sum_{k=1}^{+\infty} \alpha_k = 1, \qquad \sum_{k=1}^{+\infty} \beta_k = 1. \tag{3.20}$$

Then, we construct the zero-order deformation equation in a more general form

$$[1 - B(q)] \left\{ \mathcal{L}[\Phi(\mathbf{r}, t; q) - u_0(\mathbf{r}, t)] \right\} = A(q)\, \hbar\, H(\mathbf{r}, t)\, \mathcal{N}[\Phi(\mathbf{r}, t; q)]. \tag{3.21}$$

All other related formulae are the same, except the high-order deformation equation which is now in a more general form

$$\mathcal{L} \left[u_m(\mathbf{r}, t) - \sum_{k=1}^{m-1} \beta_k\, u_{m-k}(\mathbf{r}, t) \right] = \hbar\, H(\mathbf{r}, t)\, R_m(\vec{u}_{m-1}, \mathbf{r}, t), \tag{3.22}$$

where

$$R_m(\vec{u}_{m-1}, \mathbf{r}, t) = \sum_{k=1}^{m} \alpha_k\, \delta_{m-k}(\mathbf{r}, t) \tag{3.23}$$

under the definition

$$\delta_n(\mathbf{r}, t) = \frac{1}{n!} \left. \frac{\partial^n \mathcal{N}[\Phi(\mathbf{r}, t; q)]}{\partial q^n} \right|_{q=0}. \tag{3.24}$$

The zero-order deformation equation (3.4) and the high-order deformation equation (3.14) are clearly special cases of Equations (3.21) and (3.22) when $A(q) = B(q) = q$, respectively.

Note that, in general, a nonlinear problem might be described by a set of governing equations with related initial/boundary conditions. For the sake of brevity, only one equation (3.1) is employed here to systematically describe the basic ideas of the homotopy analysis method. However, the form of equation (3.1) is so general that it can denote either a governing equation or a boundary/initial condition. It may be a differential equation, an integral equation, an integro-differential equation, or an algebraic equation. All governing equations and boundary conditions can be treated in a similar way, although for different governing equations and intial/boundary conditions we should choose different initial approximations, different auxiliary linear operators, and different types of embedding functions $A(q), B(q)$. In addition, it is unnecessary for us to assume the existence of any small/large quantities in governing equations or initial/boundary conditions. Therefore, the analytic approach described above is very general.

3.3 Convergence theorem

The convergence of a series is important. A series is often of no use if it is convergent in a rather restricted region. In general cases, we can prove that, as long as the solution series (3.13) given by the homotopy analysis method is convergent, it must be the solution of the considered nonlinear problem.

THEOREM 3.1 Convergence theorem
As long as the series

$$u_0(\mathbf{r}, t) + \sum_{m=1}^{+\infty} u_m(\mathbf{r}, t)$$

is convergent, where $u_m(\mathbf{r}, t)$ is governed by the high-order deformation equation (3.22) under the definitions (3.23), (3.24), and (2.42), it must be a solution of Equation (3.1).

Proof: Let

$$s(\mathbf{r}, t) = u_0(\mathbf{r}, t) + \sum_{m=1}^{+\infty} u_m(\mathbf{r}, t)$$

denote the convergent series. Using (3.22) and (2.42), we have

$$
\begin{aligned}
&\hbar\, H(\mathbf{r},t) \sum_{m=1}^{+\infty} R_m(\vec{u}_{m-1},\mathbf{r},t)\\
&= \sum_{m=1}^{+\infty} \mathcal{L}\left[u_m(\mathbf{r},t) - \sum_{k=1}^{m-1} \beta_k\, u_{m-k}(\mathbf{r},t)\right]\\
&= \mathcal{L}\left[\sum_{m=1}^{+\infty} u_m(\mathbf{r},t) - \sum_{m=1}^{+\infty}\sum_{k=1}^{m-1} \beta_k\, u_{m-k}(\mathbf{r},t)\right]\\
&= \mathcal{L}\left[\sum_{m=1}^{+\infty} u_m(\mathbf{r},t) - \sum_{k=1}^{+\infty}\sum_{m=k+1}^{+\infty} \beta_k\, u_{m-k}(\mathbf{r},t)\right]\\
&= \mathcal{L}\left[\sum_{m=1}^{+\infty} u_m(\mathbf{r},t) - \sum_{k=1}^{+\infty}\beta_k\sum_{n=1}^{+\infty} u_n(\mathbf{r},t)\right]\\
&= \mathcal{L}\left[\left(1-\sum_{k=1}^{+\infty}\beta_k\right)\sum_{m=1}^{+\infty} u_m(\mathbf{r},t)\right],\\
&= \mathcal{L}\left[\left(1-\sum_{k=1}^{+\infty}\beta_k\right) s(\mathbf{r},t)\right],
\end{aligned}
$$

which gives, since $\hbar \neq 0$, $H(\mathbf{r},t) \neq 0$ and from (3.20) and (3.2),

$$
\sum_{m=1}^{+\infty} R_m(\vec{u}_{m-1},\mathbf{r},t) = 0. \tag{3.25}
$$

On the other side, we have according to the definitions (3.23) and (3.24), that

$$
\begin{aligned}
\sum_{m=1}^{+\infty} R_m(\vec{u}_{m-1},\mathbf{r},t) &= \sum_{m=1}^{+\infty}\sum_{k=1}^{m} \alpha_k\, \delta_{m-k}(\mathbf{r},t)\\
= \sum_{k=1}^{+\infty}\sum_{m=k}^{+\infty} \alpha_k\, \delta_{m-k}(\mathbf{r},t) &= \left(\sum_{k=1}^{+\infty}\alpha_k\right)\sum_{n=0}^{+\infty} \delta_n(\mathbf{r},t),
\end{aligned}
$$

which gives from (3.20), (3.24), and (3.25)

$$
\begin{aligned}
\sum_{m=1}^{+\infty} R_m(\vec{u}_{m-1},\mathbf{r},t) &= \sum_{m=0}^{+\infty} \delta_m(\mathbf{r},t)\\
&= \sum_{m=0}^{+\infty} \frac{1}{m!}\, \frac{\partial^m \mathcal{N}[\Phi(\mathbf{r},t;q)]}{\partial q^m}\bigg|_{q=0} = 0.
\end{aligned} \tag{3.26}
$$

In general, $\Phi(\mathbf{r},t;q)$ does not satisfy the original nonlinear equation (3.1). Let

$$
\mathcal{E}(\mathbf{r},t;q) = \mathcal{N}[\Phi(\mathbf{r},t;q)]
$$

denote the residual error of Equation (3.1). Clearly,

$$\mathcal{E}(\mathbf{r}, t; q) = 0$$

corresponds to the exact solution of the original equation (3.1). According to the above definition, the Maclaurin series of the residual error $\mathcal{E}(\mathbf{r}, t; q)$ about the embedding parameter q is

$$\sum_{m=0}^{+\infty} \frac{q^m}{m!} \left. \frac{\partial^m \mathcal{E}(\mathbf{r}, t; q)}{\partial q^m} \right|_{q=0} = \sum_{m=0}^{+\infty} \frac{q^m}{m!} \left. \frac{\partial^m \mathcal{N}[\Phi(\mathbf{r}, t; q)]}{\partial q^m} \right|_{q=0}.$$

When $q = 1$, the above expression gives, using (3.26),

$$\mathcal{E}(\mathbf{r}, t; 1) = \sum_{m=0}^{+\infty} \frac{1}{m!} \left. \frac{\partial^m \mathcal{E}(\mathbf{r}, t; q)}{\partial q^m} \right|_{q=0} = 0. \tag{3.27}$$

This means, according to the definition of $\mathcal{E}(\mathbf{r}, t; q)$, that we gain the exact solution of the original equation (3.1) when $q = 1$. Thus, as long as the series

$$u_0(\mathbf{r}, t) + \sum_{m=1}^{+\infty} u_m(\mathbf{r}, t)$$

is convergent, it must be one solution of the original equation (3.1). This ends the proof.

THEOREM 3.2

As long as the series

$$u_0(\mathbf{r}, t) + \sum_{m=1}^{+\infty} u_m(\mathbf{r}, t)$$

is convergent, where $u_m(\mathbf{r}, t)$ *is governed by the high-order deformation equation (3.22) under the definitions (3.23), (3.24), and (2.42), it holds that*

$$\sum_{m=1}^{+\infty} R_m(\vec{u}_{m-1}, \mathbf{r}, t) = \sum_{m=0}^{+\infty} \delta_m(\mathbf{r}, t) = 0.$$

Proof: Using (3.26), the proof of this theorem is straightforward. This ends the proof.

Note that Equation (3.14) is only a special case of Equation (3.22) when $A(q) = B(q) = q$. We therefore have the following.

THEOREM 3.3

As long as the series

$$u_0(\mathbf{r}, t) + \sum_{m=1}^{+\infty} u_m(\mathbf{r}, t)$$

is convergent, where $u_m(\mathbf{r}, t)$ *is governed by the high-order deformation equation (3.14) under the definitions (2.42) and (3.15), it must be a solution of Equation (3.1). It holds therefore that*

$$\sum_{m=1}^{+\infty} R_m(\vec{u}_{m-1}, \mathbf{r}, t) = 0.$$

According to Theorem 3.1 and Theorem 3.3, we need only focus on choosing the initial approximation $u_0(\mathbf{r}, t)$, the auxiliary linear operator $\mathcal{L}$, the embedding functions $A(q), B(q)$, the auxiliary parameter $\hbar$, and the auxiliary function $H(\mathbf{r}, t)$ to ensure that the solution series (3.13) converges. Theorem 3.2 provides us with an alternative method to estimate the convergence and accuracy of approximation series given by the homotopy analysis method.

3.4 Fundamental rules

Perturbation techniques and other nonperturbation methods for nonlinear problems are, more or less, based on some assumptions. Similarly, the homotopy analysis method is also based on the assumptions listed on page 55. Theoretically speaking, these assumptions impair the method. However, the homotopy analysis method provides us with great freedom to choose the initial approximation $u_0(\mathbf{r}, t)$, the auxiliary linear operator $\mathcal{L}$, the auxiliary parameter $\hbar$, and the auxiliary function $H(\mathbf{r}, t)$. Such freedom is so great that it is almost quite possible for us to satisfy all of these assumptions. This kind of freedom is therefore a cornerstone of the validity and flexibility of the homotopy analysis method.

However, from the view points of practical applications, the freedom seems too great. It is therefore better to have some fundamental rules to direct us to choose the initial approximation $u_0(\mathbf{r}, t)$, the auxiliary linear operator $\mathcal{L}$, and the auxiliary function $H(\mathbf{r}, t)$. We must first emphasize two facts. First, a solution of a nonlinear problem may be expressed by different sets of base functions, as illustrated in Chapter 2. Second, in many cases, from physical characteristics and boundary/initial conditions, it is often not very difficult to determine the type of base functions convenient to represent solutions of a given nonlinear problem, even without solving it. So, given a nonlinear problem, we can first choose a set of base function to present its solutions. This kind of presentation provides us with the *rule of solution expression.* For example, support that we choose a set of base functions

$$\{e_n(\mathbf{r}, t) \ | \ n = 0, 1, 2, 3, \cdots\} \tag{3.28}$$

to represent the solution $u(\mathbf{r},t)$ of Equation (3.1) by

$$u(\mathbf{r},t) = \sum_{n=0}^{+\infty} c_n \; e_n(\mathbf{r},t), \tag{3.29}$$

where c_n is a coefficient. The above expression provides the so-called *rule of solution expression* for Equation (3.1). To obey the *rule of solution expression*, the initial approximation $u_0(\mathbf{r},t)$ must be expressed by a sum of the base functions, i.e.,

$$u_0(\mathbf{r},t) = \sum_{n=0}^{M_0} a_n \; e_n(\mathbf{r},t), \tag{3.30}$$

where a_n is a coefficient and M_0 is an integer. To obey the *rule of solution expression* denoted by (3.29), the auxiliary linear operator $\mathcal{L}$ must be chosen in such a way that the solution of the equation

$$\mathcal{L}[w(\mathbf{r},t)] = 0$$

must be expressed by a sum of the base functions, say,

$$w(\mathbf{r},t) = \sum_{n=0}^{M_1} b_n \; e_n(\mathbf{r},t), \tag{3.31}$$

where b_n is a coefficient and the integer M_1 is determined by the highest order of the derivative of linear operator $\mathcal{L}$, which is generally the same as the highest order of the derivative of the original equation (3.1). This is because the solution of the high-order deformation equation (3.22) can be expressed by

$$u_m(\mathbf{r},t) = u_m^*(\mathbf{r},t) + w(\mathbf{r},t),$$

where $u_m^*(\mathbf{r},t)$ is a special solution of Equation (3.22). Furthermore, to obey the *rule of solution expression* denoted by (3.29), the auxiliary function $H(\mathbf{r},t)$ should be chosen so that the special solution $u_m^*(\mathbf{r},t)$ of the high-order deformation equation (3.22) must be expressed by a sum of the base functions, say,

$$u_m^*(\mathbf{r},t) = \mathcal{L}^{-1}[\hbar \; H(\mathbf{r},t) \; R_m(\vec{u}_{m-1},\mathbf{r},t)] = \sum_{n=0}^{M_2} d_n \; e_n(\mathbf{r},t), \tag{3.32}$$

where d_n is a coefficient and $\mathcal{L}^{-1}$ is the inverse operator of the auxiliary linear operator $\mathcal{L}$. In this way, the *rule of solution expression* directs us to choose the initial approximation $u_0(\mathbf{r},t)$, the auxiliary linear operator $\mathcal{L}$, and the auxiliary function $H(\mathbf{r},t)$. Using the so-called *rule of solution expression*, we can easily avoid the appearance of the so-called secular terms in solution expressions, as illustrated in this book. Therefore, the *rule of solution expression* practically provides us with a starting point and therefore plays a very important role within the frame of the homotopy analysis method.

It is found that, in most cases, the auxiliary function $H(\mathbf{r}, t)$ cannot be uniquely determined by above-mentioned *rule of solution expression*. Thus, more restrictions should be given to direct us to choose the auxiliary function $H(\mathbf{r}, t)$. Note that, from the view point of the completeness, each base $e_n(\mathbf{r}, t)$ of the set denoted by (3.28) should appear in the solution expression (3.29). In other words, each coefficient $c_{m,n}$ of the mth-order approximate solution

$$u(\mathbf{r}, t) \approx \sum_{n=1}^{m} u_m(\mathbf{r}, t) = \sum_{n=0}^{M_3} c_{m,n}\, e_n(\mathbf{r}, t) \tag{3.33}$$

can be modified as the order of approximation tends to infinity. This provides us with the so-called *rule of coefficient ergodicity*, i.e., *as the order of approximation tends to infinity, each base should appear in the solution expression and each coefficient can be modified.* This further restricts the choice of the auxiliary function. In many cases, using the *rule of solution expression* and the *rule of coefficient ergodicity*, we can uniquely determine the auxiliary function $H(\mathbf{r}, t)$, as illustrated in this book. Thus, the *rule of coefficient ergodicity* also plays a very important role within the frame of the homotopy analysis method.

Using (3.13), the original nonlinear problem is transformed into an infinite number of linear subproblems governed by the high-order deformation equation (3.14) or (3.22). So, if the original nonlinear problem has a solution, all of these linear subproblems should have solutions too. Thus, we have the so-called *rule of solution existence*, i.e., the initial approximation $u_0(\mathbf{r}, t)$, the auxiliary linear operator $\mathcal{L}$, and the auxiliary function $H(\mathbf{r}, t)$ should be chosen so that all of the high-order deformation equation (3.14) or (3.22) are closed and have solutions, if the original nonlinear problem has a solution. This rule further restricts the choice of the initial approximation $u_0(\mathbf{r}, t)$, the auxiliary linear operator $\mathcal{L}$, and the auxiliary function $H(\mathbf{r}, t)$.

The above-mentioned *rule of solution expression* and *rule of coefficient ergodicity*, in addition to *rule of solution existence*, direct us to choose the initial approximation $u_0(\mathbf{r}, t)$, the auxiliary linear operator $\mathcal{L}$, and the auxiliary function $H(\mathbf{r}, t)$. These rules considerably simplify the application of the homotopy analysis method.

3.5 Control of convergence region and rate

It is important to ensure that a solution series is convergent in a large enough region. In general, the convergence region and rate of solution series are mainly determined by the base functions used to represent the solution series. Unlike previous analytic techniques, the homotopy analysis method provides us with great freedom to represent solutions of a given nonlinear problem by different base functions. Therefore, by means of the homotopy analysis method, we can gain solution series convergent in a whole region having physical meanings, as illustrated in this book. The so-called *rule of solution expression* is most important and the key, which determines the choice of the initial approximation $u_0(\mathbf{r}, t)$, the auxiliary linear operator $\mathcal{L}$, and the auxiliary function $H(\mathbf{r}, t)$.

Even if the initial approximation $u_0(\mathbf{r}, t)$, the auxiliary linear operator $\mathcal{L}$, and the auxiliary function $H(\mathbf{r}, t)$ are given, we still have great freedom to choose the value of the auxiliary parameter $\hbar$. Unlike all previous analytic techniques, the homotopy analysis method always provides us with a *family* of solution expressions in the auxiliary parameter $\hbar$. It is found that the auxiliary parameter $\hbar$ often affects convergence region and rate of solution series, as shown and proved in Chapter 2. The influence of $\hbar$ on the convergence region and rate becomes obvious, especially when a "bad" set of base function is chosen, as illustrated in Chapter 2. It is found that the convergence region and rate of solution series can be easily adjusted and controlled by means of setting $\hbar$ proper values. Thus, unlike all previous analytic techniques, the homotopy analysis method provides us with a convenient way to control and adjust convergence region and rate of solution series.

3.5.1 The $\hbar$-curve and the valid region of $\hbar$

Assume that we gain a family of solution series in the auxiliary parameter $\hbar$ by means of homotopy analysis method. How does one then to choose the value of $\hbar$ to ensure that the solution series converges fast enough in a large enough region?

Many nonlinear problems contain important physical quantities such as frequency of a nonlinear oscillator, wall skin friction of viscous flow, and so on. Because we have a family of solution expressions in the auxiliary parameter $\hbar$, those physical quantities also depend upon $\hbar$. So, regarding $\hbar$ as an independent variable, it is easy to plot curves of these kinds of quantities versus $\hbar$. For example, assume that

$$\gamma = \ddot{u}(\mathbf{r}, t)|_{\mathbf{r}=0, t=0}$$

corresponds to a quantity having important physical meaning, where the dot denotes the derivative with respect to the time t. Then, γ is a function of $\hbar$ and thus can be plotted by a curve $\gamma \sim \hbar$. According to Theorem 3.1 or Theorem 3.3, all convergent series of γ given by different values of $\hbar$ converge to its exact value. So, if the solution is unique, all of them converge to the same value and therefore there exists a horizontal line segment in the figure of $\gamma \sim \hbar$ that corresponds to a region of $\hbar$ denoted by $\mathbf{R}_\hbar$. For the sake of brevity we call such a kind of curve *the $\hbar$-curve* and the corresponding region $\mathbf{R}_\hbar$ *the valid region of $\hbar$*, respectively. Thus, if we set $\hbar$ any value in the so-called valid region of $\hbar$, we are quite sure that the corresponding solution series converge. Certainly, if there exist many such kinds of quantities, we can plot corresponding $\hbar$-curves of them. And even if the term denoted by γ has no physical meanings, we can still plot the corresponding $\hbar$-curves. Obviously, the more the so-called $\hbar$-curves are plotted, the clearer it is to choose the value of $\hbar$. It is found that, for given initial approximation $u_0(\mathbf{r}, t)$, the auxiliary linear operator $\mathcal{L}$, and the auxiliary function $H(\mathbf{r}, t)$, the valid regions of $\hbar$ for different special quantities are often nearly the same for a given problem, although up to now we cannot give a mathematical proof in general. In most cases, using the same $\hbar$-curve gained by a special quantity such as γ mentioned above, we can find a proper value of $\hbar$ to ensure that the solution series of $u(\mathbf{r}, t)$ converges in the whole spacial and temporal regions having physical meanings. So, the so-called $\hbar$-curve provides us with a convenient way to show the influence of $\hbar$ on the convergence region and rate of solution series.

3.5.2 Homotopy-Padé technique

The Padé technique is widely applied to enlarge the convergence region and rate of a given series. Traditionally, the $[m, n]$ Padé approximant of $u(\mathbf{r}, t)$ is expressed by either

$$\frac{\sum_{k=0}^{m} F_k(\mathbf{r})\ t^k}{1 + \sum_{k=1}^{n} F_{m+1+k}(\mathbf{r})\ t^k}$$

or

$$\frac{\sum_{k=0}^{m} G_k(t)\ \mathbf{r}^k}{1 + \sum_{k=1}^{n} G_{m+1+k}(t)\ \mathbf{r}^k},$$

where $F_k(\mathbf{r})$ and $G_k(t)$ are functions. Note that the numerator and denominator are polynomial of either the spatial variable $\mathbf{r}$ or the temporal variable t.

The Padé technique can be employed within the frame of the homotopy analysis method. As mentioned before, the homotopy analysis method is based on such an assumption that the series (3.12) is convergent at $q = 1$ because the solution series (3.13) is obtained by setting $q = 1$ in (3.12). So, it is important to ensure that the series (3.12) is convergent at $q = 1$. We first employ the traditional Padé technique to the series (3.12) about the embedding parameter q to gain the $[m, n]$ Padé approximant

$$\frac{\sum_{k=0}^{m} W_k(\mathbf{r}, t)\ q^k}{1 + \sum_{k=1}^{n} W_{m+1+k}(\mathbf{r}, t)\ q^k},$$

where $W_k(\mathbf{r},t)$ is a function determined by the first several approximations

$$u_j(\mathbf{r},t), \qquad j = 0, 1, 2, 3, \cdots, m+n.$$

Then, using (3.8), we set $q = 1$ to get the so-called $[m, n]$ homotopy-Padé approximant

$$\frac{\sum_{k=0}^{m} W_k(\mathbf{r},t)}{1 + \sum_{k=1}^{n} W_{m+1+k}(\mathbf{r},t)}.$$

It is found that the $[m, n]$ homotopy-Padé approximant often converges faster than the corresponding traditional $[m, n]$ Padé approximant. In many cases, the $[m, m]$ homotopy-Padé approximant does not depend upon the auxiliary parameter $\hbar$. In this case, we can gain convergent solution by means of the homotopy-Padé technique even if the solution series is divergent. However, up to now, we cannot prove it in general.

It is flexible to apply the homotopy-Padé technique to accelerate related solution series. For example, we can employ it to accelerate the convergence of the series

$$\dot{u}(\mathbf{r},t) = \dot{u}_0(\mathbf{r},t) + \sum_{n=1}^{+\infty} \dot{u}_n(\mathbf{r},t).$$

First of all, from (3.12), we have the series

$$\frac{\partial \Phi(\mathbf{r},t;q)}{\partial t} = \dot{u}_0(\mathbf{r},t) + \sum_{n=1}^{+\infty} \dot{u}_n(\mathbf{r},t)\; q^n.$$

Then, applying the Padé technique to the above series about the embedding parameter q, we have the traditional $[m, n]$ Padé approximation

$$\frac{\sum_{k=0}^{m} V_k(\mathbf{r},t)\; q^k}{1 + \sum_{k=1}^{n} V_{m+1+k}(\mathbf{r},t)\; q^k},$$

where $V_k(\mathbf{r},t)$ is a function of $\mathbf{r}$ and t. Setting $q = 1$ in the above expression, we have using (3.8) the $[m, n]$ homotopy-Padé approximant

$$\dot{u}(\mathbf{r},t) \approx \frac{\sum_{k=0}^{m} V_k(\mathbf{r},t)}{1 + \sum_{k=1}^{n} V_{m+1+k}(\mathbf{r},t)}.$$

In summary, the $\hbar$-curve provides us with a convenient way to determine the valid region of $\hbar$. In addition, the so-called homotopy-Padé technique can greatly enlarge the convergence region and rate of solution series. In many cases, the homotopy-Padé technique is more efficient than the traditional Padé method and is even independent of the auxiliary parameter $\hbar$. So, by means of choosing a proper set of base functions, selecting a proper value of $\hbar$, or employing the homotopy-Padé technique, we can gain accurate approximations convergent in a large enough region within the frame of the homotopy analysis method.

3.6 Further generalization

The homotopy analysis method can be further generalized by means of the zero-order deformation equation in the form

$$\begin{aligned} &[1 - B(q)]\,\{\mathcal{L}[\Phi(\mathbf{r}, t; q) - u_0(\mathbf{r}, t)]\} \\ &= A(q)\,\hbar\,H(\mathbf{r}, t)\,\mathcal{N}[\Phi(\mathbf{r}, t; q)] + \hbar_2\,H_2(\mathbf{r}, t)\,\Pi[\Phi(\mathbf{r}, t; q); q], \end{aligned} \tag{3.34}$$

where $u_0(\mathbf{r}, t), \mathcal{L}, H(\mathbf{r}, t), \hbar, A(q)$, and $B(q)$ are defined as before, $\hbar_2$ is the second auxiliary parameter, $H_2(\mathbf{r}, t)$ is the second auxiliary function, and $\Pi[\Phi(\mathbf{r}, t); q]$ is an auxiliary operator which equals to zero when $q = 0$ and $q = 1$, i.e.,

$$\Pi[\Phi(\mathbf{r}, t; 0); 0] = \Pi[\Phi(\mathbf{r}, t; 1); 1] = 0. \tag{3.35}$$

All other related formulae are the same, except the high-order deformation equation in a more general form

$$\begin{aligned} &\mathcal{L}\left[u_m(\mathbf{r}, t) - \sum_{k=1}^{m-1} \beta_k\, u_{m-k}(\mathbf{r}, t)\right] \\ &= \hbar\,H(\mathbf{r}, t)\,R_m(\vec{u}_{m-1}, \mathbf{r}, t) + \hbar_2\,H_2(\mathbf{r}, t)\,\Delta_m(\mathbf{r}, t), \end{aligned} \tag{3.36}$$

where

$$\Delta_m(\mathbf{r}, t) = \frac{1}{m!}\,\left.\frac{\partial^m \Pi\,[\Phi(\mathbf{r}, t; q); q]}{\partial q^m}\right|_{q=0}. \tag{3.37}$$

In this way, we introduce the additional auxiliary parameter $\hbar_2$ and auxiliary function $H_2(\mathbf{r}, t)$, and more importantly, an auxiliary operator $\Pi[\Phi(\mathbf{r}, t; q); q]$. In this way the flexibility of the homotopy analysis method is further increased. Note that the solution series given by Equation (3.36) is now a family of two parameters, $\hbar$ and $\hbar_2$.

It is rather flexible to choose the auxiliary operator $\Pi[\Phi(\mathbf{r}, t; q); q]$ which satisfies the property (3.35). For example, we can choose

$$\Pi[\Phi(\mathbf{r}, t; q); q] = A(q)[1 - B(q)]F[\Phi(\mathbf{r}, t; q)], \tag{3.38}$$

where $F[\Phi(\mathbf{r}, t; q)]$ is a function, or

$$\Pi[\Phi(\mathbf{r}, t; q); q] = [1 - A(q)] \left\{ [\Phi(\mathbf{r}, t; q)]^{1+q} - \Phi(\mathbf{r}, t; q) \right\}, \qquad (3.39)$$

and so on. However, it is under investigation how to choose the additional auxiliary parameter $\hbar_2$, the additional auxiliary function $H_2(\mathbf{r}, t)$, and the auxiliary operator $\Pi[\Phi(\mathbf{r}, t; q); q]$ for a given nonlinear problem in general. For the applications of the zero-order deformation equation in the form (3.34), the reader is referred to §4.3 and §12.1.