
PART THREE

PERTURBATION METHODS

When I hear you give your reasons, the thing always appears to me to be so ridiculously simple that I could easily do it myself, though at each successive instance of your reasoning I am baffled until you explain your process.

—Dr. Watson, *A Scandal in Bohemia*
Sir Arthur Conan Doyle

The local analysis methods of Part II are powerful tools, but they cannot provide global information on the behavior of solutions at two distantly separated points. They cannot predict how a change in initial conditions at $x = 0$ will affect the asymptotic behavior as $x \rightarrow +\infty$. To answer such questions we must apply the methods of global analysis which will be developed in Part IV. Since global methods are perturbative in character, in this part we will first introduce the requisite mathematical concepts: perturbation theory in Chap. 7 and summation theory in Chap. 8.

Perturbation theory is a collection of methods for the systematic analysis of the global behavior of solutions to differential and difference equations. The general procedure of perturbation theory is to identify a small parameter, usually denoted by ε , such that when $\varepsilon = 0$ the problem becomes soluble. The global solution to the given problem can then be studied by a local analysis about $\varepsilon = 0$. For example, the differential equation $y'' = [1 + \varepsilon/(1 + x^2)]y$ can only be solved in terms of elementary functions when $\varepsilon = 0$. A perturbative solution is constructed by local analysis about $\varepsilon = 0$ as a series of powers of ε :

$$y(x) = y_0(x) + \varepsilon y_1(x) + \varepsilon^2 y_2(x) + \cdots.$$

This series is called a perturbation series. It has the attractive feature that $y_n(x)$ can be computed in terms of $y_0(x), \dots, y_{n-1}(x)$ as long as the problem obtained by setting $\varepsilon = 0$, $y'' = y$, is soluble, which it is in this case. Notice that the perturbation series for $y(x)$ is *local* in ε but that it is *global* in x . If ε is very small, we expect that $y(x)$ will be well approximated by only a few terms of the perturbation series.

The local analysis methods of Part II are other examples of perturbation theory. There the expansion parameter is $\varepsilon = x - x_0$ or $\varepsilon = 1/x$ if $x_0 = \infty$.

Perturbation series, like asymptotic expansions, often diverge for all $\varepsilon \neq 0$. However, since ε is not necessarily a small parameter, the optimal asymptotic approximation may give very poor numerical results. Thus, to extract maximal information from perturbation theory, it is necessary to develop sophisticated techniques to “sum” divergent series and to accelerate the convergence of slowly converging series. Methods to achieve these goals are presented in Chap. 8. Summation methods also apply to the local series expansions derived in Part II.

In perturbation theory it is convenient to have an asymptotic order relation that expresses the relative magnitudes of two functions more precisely than \ll but less precisely than \sim . We define

$$f(x) = O[g(x)], \quad x \rightarrow x_0,$$

and say “ $f(x)$ is at most of order $g(x)$ as $x \rightarrow x_0$ ” or “ $f(x)$ is ‘O’ of $g(x)$ as $x \rightarrow x_0$ ” if $f(x)/g(x)$ is bounded for x near x_0 ; that is, $|f(x)/g(x)| < M$, for some constant M if x is sufficiently close to x_0 . Observe that if $f(x) \sim g(x)$ or if $f(x) \ll g(x)$ as $x \rightarrow x_0$, then $f(x) = O[g(x)]$ as $x \rightarrow x_0$. If $f \ll g$ as $x \rightarrow x_0$, then any $M > 0$ satisfies the definition, while if $f \sim g$ ($x \rightarrow x_0$), only $M > 1$ can work.

In perturbation theory one may calculate just a few terms in a perturbation series. Whether or not this series is convergent, the notation “O” is very useful for expressing the order of magnitude of the first neglected term when that term has not been calculated explicitly.

Examples

1. $x \sin x = O(x)$ ($x \rightarrow 0$ or $x \rightarrow \infty$);
2. $e^{-1/x} = O(x^n)$ ($x \rightarrow 0+$) for all n ;
3. $x^5 = O(x^2)$ ($x \rightarrow 0+$);
4. $e^x = 1 + x + (x^2/2) + O(x^3)$ ($x \rightarrow 0$);
5. $\text{Ai}(x) = \frac{1}{2}\pi^{-1/2}x^{-1/4}e^{-2x^{3/2}/3}[1 - \frac{5}{48}x^{-3/2} + O(x^{-3})]$ ($x \rightarrow +\infty$).

PERTURBATION SERIES

You have erred perhaps in attempting to put colour and life into each of your statements instead of confining yourself to the task of placing upon record that severe reasoning from cause to effect which is really the only notable feature about the thing. You have degraded what should have been a course of lectures into a series of tales.

—Sherlock Holmes, *The Adventure of the Copper Beeches*
Sir Arthur Conan Doyle

(E) 7.1 PERTURBATION THEORY

Perturbation theory is a large collection of iterative methods for obtaining approximate solutions to problems involving a small parameter ε . These methods are so powerful that sometimes it is actually advisable to introduce a parameter ε temporarily into a difficult problem having no small parameter, and then finally to set $\varepsilon = 1$ to recover the original problem. This apparently artificial conversion to a perturbation problem may be the only way to make progress.

The thematic approach of perturbation theory is to decompose a tough problem into an infinite number of relatively easy ones. Hence, perturbation theory is most useful when the first few steps reveal the important features of the solution and the remaining ones give small corrections.

Here is an elementary example to introduce the ideas of perturbation theory.

Example 1 *Roots of a cubic polynomial.* Let us find approximations to the roots of

$$x^3 - 4.001x + 0.002 = 0. \quad (7.1.1)$$

As it stands, this problem is not a perturbation problem because there is no small parameter ε . It may not be easy to convert a particular problem into a tractable perturbation problem, but in the present case the necessary trick is almost obvious. Instead of the single equation (7.1.1) we consider the *one-parameter family* of polynomial equations

$$x^3 - (4 + \varepsilon)x + 2\varepsilon = 0. \quad (7.1.2)$$

When $\varepsilon = 0.001$, the original equation (7.1.1) is reproduced.

It may seem a bit surprising at first, but it is easier to compute the approximate roots of the family of polynomials (7.1.2) than it is to solve just the one equation with $\varepsilon = 0.001$. The reason

for this is that if we consider the roots to be functions of ε , then we may further assume a perturbation series in powers of ε :

$$x(\varepsilon) = \sum_{n=0}^{\infty} a_n \varepsilon^n. \quad (7.1.3)$$

To obtain the first term in this series, we set $\varepsilon = 0$ in (7.1.2) and solve

$$x^3 - 4x = 0. \quad (7.1.4)$$

This expression is easy to factor and we obtain in *zeroth-order* perturbation theory $x(0) = a_0 = -2, 0, 2$.

A *second-order* perturbation approximation to the first of these roots consists of writing (7.1.3) as $x_1 = -2 + a_1 \varepsilon + a_2 \varepsilon^2 + O(\varepsilon^3)$ ($\varepsilon \rightarrow 0$), substituting this expression into (7.1.2), and neglecting powers of ε beyond ε^2 . The result is

$$(-8 + 8) + (12a_1 - 4a_1 + 2 + 2)\varepsilon + (12a_2 - a_1 - 6a_1^2 - 4a_2)\varepsilon^2 = O(\varepsilon^3), \quad \varepsilon \rightarrow 0. \quad (7.1.5)$$

It is at this step that we realize the power of generalizing the original problem to a family of problems (7.1.2) with variable ε . It is because ε is *variable* that we can conclude that the coefficient of each power of ε in (7.1.5) is *separately* equal to zero. This gives a sequence of equations for the expansion coefficients a_1, a_2, \dots :

$$\varepsilon^1: \quad 8a_1 + 4 = 0; \quad \varepsilon^2: \quad 8a_2 - a_1 - 6a_1^2 = 0;$$

and so on. The solutions to the equations are $a_1 = -\frac{1}{2}$, $a_2 = \frac{1}{8}$, \dots . Therefore, the perturbation expansion for the root x_1 is

$$x_1 = -2 - \frac{1}{2}\varepsilon + \frac{1}{8}\varepsilon^2 + \dots \quad (7.1.6)$$

If we now set $\varepsilon = 0.001$, we obtain x_1 from (7.1.6) accurate to better than one part in 10^9 . The same procedure gives

$$x_2 = 0 + \frac{1}{2}\varepsilon - \frac{1}{8}\varepsilon^2 + O(\varepsilon^3), \quad x_3 = 2 + 0 \cdot \varepsilon + 0 \cdot \varepsilon^2 + O(\varepsilon^3), \quad \varepsilon \rightarrow 0.$$

(Successive coefficients in the perturbation series for x_3 all vanish because $x_3 = 2$ is the exact solution for all ε .) All three perturbation series for the roots converge for $\varepsilon = 0.001$. Can you prove that they converge for $|\varepsilon| < 1$? (See Prob. 7.6.)

This example illustrates the three steps of perturbative analysis:

1. Convert the original problem into a perturbation problem by introducing the small parameter ε .
2. Assume an expression for the answer in the form of a perturbation series and compute the coefficients of that series.
3. Recover the answer to the original problem by summing the perturbation series for the appropriate value of ε .

Step (1) is sometimes ambiguous because there may be many ways to introduce an ε . However, it is preferable to introduce ε in such a way that the *zeroth-order* solution (the leading term in the perturbation series) is obtainable as a closed-form analytic expression. Perturbation problems generally take the form of a soluble equation [such as (7.1.4)] whose solution is altered slightly by a perturbing term [such as $(2 - x)\varepsilon$]. Of course, step (1) may be omitted when the original problem already has a small parameter if a perturbation series can be developed in powers of that parameter.

Step (2) is frequently a routine iterative procedure for determining successive coefficients in the perturbation series. A *zeroth-order* solution consists of finding the leading term in the perturbation series. In Example 1 this involves solving the *unperturbed problem*, the problem obtained by setting $\varepsilon = 0$ in the perturbation problem. A *first-order* solution consists of finding the first two terms in the perturbation series, and so on. In Example 1 each of the coefficients in the perturbation series is determined in terms of the previous coefficients by a simple linear equation, even though the original problem was a nonlinear (cubic) equation.

Generally it is the existence of a closed-form zeroth-order solution which ensures that the higher-order terms may also be determined as closed-form analytical expressions.

Step (3) may or may not be easy. If the perturbation series converges, its sum is the desired answer. If there are several ways to reduce a problem to a perturbation problem, one chooses the way that is the best compromise between difficulty of calculation of the perturbation series coefficients and rapidity of convergence of the series itself. However, many series converge so slowly that their utility is impaired. Also, we will shortly see that perturbation series are frequently divergent. This is not necessarily bad because many of these divergent perturbation series are asymptotic. In such cases, one obtains a good approximation to the answer when ε is very small by summing the first few terms according to the optimal truncation rule (see Sec. 3.5). When ε is not small, it may still be possible to obtain a good approximation to the answer from a slowly converging or divergent series using the summation methods discussed in Chap. 8.

Let us now apply these three rules of perturbation theory to a slightly more sophisticated example.

Example 2 *Approximate solution of an initial-value problem.* Consider the initial-value problem

$$y'' = f(x)y, \quad y(0) = 1, y'(0) = 1, \tag{7.1.7}$$

where $f(x)$ is continuous. This problem has no closed-form solution except for very special choices for $f(x)$. Nevertheless, it can be solved perturbatively.

First, we introduce an ε in such a way that the unperturbed problem is solvable:

$$y'' = \varepsilon f(x)y, \quad y(0) = 1, y'(0) = 1. \tag{7.1.8}$$

Second, we assume a perturbation expansion for $y(x)$ of the form

$$y(x) = \sum_{n=0}^{\infty} \varepsilon^n y_n(x), \tag{7.1.9}$$

where $y_0(0) = 1, y_0'(0) = 1,$ and $y_n(0) = 0, y_n'(0) = 0$ ($n \geq 1$).

The zeroth-order problem $y'' = 0$ is obtained by setting $\varepsilon = 0$, and the solution which satisfies the initial conditions is $y_0 = 1 + x$. The n th-order problem ($n \geq 1$) is obtained by substituting (7.1.9) into (7.1.8) and setting the coefficient of ε^n ($n \geq 1$) equal to 0. The result is

$$y_n'' = y_{n-1} f(x), \quad y_n(0) = y_n'(0) = 0. \tag{7.1.10}$$

Observe that perturbation theory has replaced the intractable differential equation (7.1.7) with a sequence of inhomogeneous equations (7.1.10). In general, any inhomogeneous equation may be solved routinely by the method of variation of parameters whenever the solution of the associated homogeneous equation is known (Sec. 1.5). Here the homogeneous equation is

precisely the unperturbed equation. Thus, it is clear why it is so crucial that the unperturbed equation be soluble.

The solution to (7.1.10) is

$$y_n = \int_0^x dt \int_0^t ds f(s) y_{n-1}(s), \quad n \geq 1. \quad (7.1.11)$$

Equation (7.1.11) gives a simple iterative procedure for calculating successive terms in the perturbation series (7.1.9):

$$\begin{aligned} y(x) = & 1 + x + \varepsilon \int_0^x dt \int_0^t ds (1+s) f(s) \\ & + \varepsilon^2 \int_0^x dt \int_0^t ds f(s) \int_0^s dv \int_0^v du (1+u) f(u) + \cdots \end{aligned} \quad (7.1.12)$$

Third, we must sum this series. It is easy to show that when N is large, the N th term in this series is bounded in absolute value by $\varepsilon^N x^{2N} K^N (1 + |x|)/(2N)!$, where K is an upper bound for $|f(t)|$ in the interval $0 \leq |t| \leq |x|$. Thus, the series (7.1.12) is convergent for all x . We also conclude that if $x^2 K$ is small, then the perturbation series is rapidly convergent for $\varepsilon = 1$ and an accurate solution to the original problem may be achieved by taking only a few terms.

How do these perturbation methods for differential equations compare with the series methods that were introduced in Chap. 3? Suppose $f(x)$ in (7.1.7) has a convergent Taylor expansion about $x = 0$ of the form

$$f(x) = \sum_{n=0}^{\infty} f_n x^n. \quad (7.1.13)$$

Then another way to solve for $y(x)$ is to perform a local analysis of the differential equation near $x = 0$ by substituting the series solution

$$y(x) = \sum_{n=0}^{\infty} a_n x^n, \quad a_0 = a_1 = 1, \quad (7.1.14)$$

and computing the coefficients a_n . As shown in Chap. 3, the series in (7.1.14) is guaranteed to have a radius of convergence at least as large as that in (7.1.13).

By contrast, the perturbation series (7.1.9) converges for all finite values of x , and not just those inside the radius of convergence of $f(x)$. Moreover, the perturbation series converges even if $f(x)$ has no Taylor series expansion at all.

Example 3 *Comparison of Taylor and perturbation series.* The differential equation

$$y'' = -e^{-x} y, \quad y(0) = 1, \quad y'(0) = 1, \quad (7.1.15)$$

may be solved in terms of Bessel functions as

$$y(x) = \frac{[Y_0(2) + Y_0(2)]J_0(2e^{-x/2}) - [J_0(2) + J_0(2)]Y_0(2e^{-x/2})}{J_0(2)Y_0(2) - J_0'(2)Y_0(2)}.$$

The local expansion (7.1.14) converges everywhere because e^{-x} has no finite singularities. Nevertheless, a fixed number of terms of the perturbation series (7.1.9) (see Prob. 7.11) gives a much

better approximation than the same number of terms of the Taylor series (7.1.14) if x is large and positive (see Fig. 7.1).

In addition, the perturbation methods of Example 2 are immediately applicable to problems where local analysis cannot be used. For example, an approximate solution of the formidable-looking nonlinear two-point boundary-value problem

$$y'' + y = \frac{\cos x}{3 + y^2}, \quad y(0) = y\left(\frac{\pi}{2}\right) = 2, \quad (7.1.16)$$

may be readily obtained using perturbation theory (see Prob. 7.14).

Thus, the ideas of perturbation theory apply equally well to problems requiring local or global analysis.

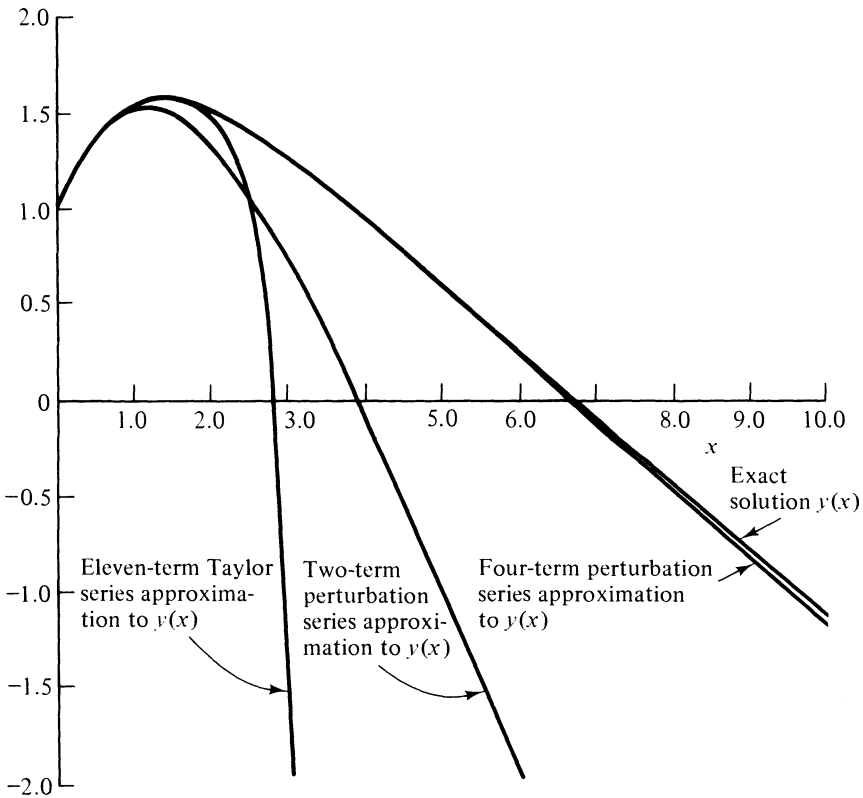


Figure 7.1 A comparison of Taylor series and perturbation series approximations to the solution of the initial-value problem $y'' = -e^{-x}y$ [$y(0) = 1, y'(0) = 1$] in (7.1.15). The exact solution to the problem is plotted. Also plotted are an 11-term Taylor series approximation of the form in (7.1.14) and 2- and 4-term perturbation series approximations of the form in (7.1.3) with $\epsilon = 1$. The global perturbative approximation is clearly far superior to the local Taylor series.

(E) 7.2 REGULAR AND SINGULAR PERTURBATION THEORY

The formal techniques of perturbation theory are a natural generalization of the ideas of local analysis of differential equations in Chap. 3. Local analysis involves approximating the solution to a differential equation near the point $x = a$ by developing a series solution about a in powers of a small parameter, either $x - a$ for finite a or $1/x$ for $a = \infty$. Once the leading behavior of the solution near $x = a$ (which we would now refer to as the zeroth-order solution!) is known, the remaining coefficients in the series can be computed recursively.

The strong analogy between local analysis of differential equations and formal perturbation theory may be used to classify perturbation problems. Recall that there are two different types of series solutions to differential equations. A series solution about an *ordinary* point of a differential equation is always a Taylor series having a nonvanishing radius of convergence. A series solution about a *singular* point does not have this form (except in rare cases). Instead, it may either be a convergent series not in Taylor series form (such as a Frobenius series) or it may be a divergent series. Series solutions about singular points often have the remarkable property of being meaningful near a singular point yet not existing at the singular point. [The Frobenius series for $K_0(x)$ does not exist at $x = 0$ and the asymptotic series for $\text{Bi}(x)$ does not exist at $x = \infty$.]

Perturbation series also occur in two varieties. We define a *regular* perturbation problem as one whose perturbation series is a power series in ε having a nonvanishing radius of convergence. A basic feature of all regular perturbation problems (which we will use to identify such problems) is that the exact solution for small but nonzero $|\varepsilon|$ smoothly approaches the unperturbed or zeroth-order solution as $\varepsilon \rightarrow 0$.

We define a *singular* perturbation problem as one whose perturbation series either does not take the form of a power series or, if it does, the power series has a vanishing radius of convergence. In singular perturbation theory there is sometimes no solution to the unperturbed problem (the exact solution as a function of ε may cease to exist when $\varepsilon = 0$); when a solution to the unperturbed problem does exist, its qualitative features are distinctly different from those of the exact solution for arbitrarily small but nonzero ε . In either case, the exact solution for $\varepsilon = 0$ is *fundamentally different in character* from the “neighboring” solutions obtained in the limit $\varepsilon \rightarrow 0$. If there is no such abrupt change in character, then we would have to classify the problem as a regular perturbation problem.

When dealing with a singular perturbation problem, one must take care to distinguish between the *zeroth-order* solution (the leading term in the perturbation series) and the solution of the unperturbed problem, since the latter may not even exist. There is no difference between these two in a regular perturbation theory, but in a singular perturbation theory the zeroth-order solution may depend on ε and may exist only for nonzero ε .

The examples of the previous section are all regular perturbation problems. Here are some examples of singular perturbation problems:

Example 1 *Roots of a polynomial.* How does one determine the approximate roots of

$$\varepsilon^2 x^6 - \varepsilon x^4 - x^3 + 8 = 0? \tag{7.2.1}$$

We may begin by setting $\varepsilon = 0$ to obtain the unperturbed problem $-x^3 + 8 = 0$, which is easily solved:

$$x = 2, 2\omega, 2\omega^2, \tag{7.2.2}$$

where $\omega = e^{2\pi i/3}$ is a complex root of unity. Note that the unperturbed equation has only three roots while the original equation has six roots. This abrupt change in the character of the solution, namely the disappearance of three roots when $\varepsilon = 0$, implies that (7.2.1) is a singular perturbation problem. Part of the exact solution ceases to exist when $\varepsilon = 0$.

The explanation for this behavior is that the three missing roots tend to ∞ as $\varepsilon \rightarrow 0$. Thus, for those roots it is no longer valid to neglect $\varepsilon^2 x^6 - \varepsilon x^4$ compared with $-x^3 + 8$ in the limit $\varepsilon \rightarrow 0$. Of course, for the three roots near 2, 2ω , and $2\omega^2$, the terms $\varepsilon^2 x^6$ and εx^4 are indeed small as $\varepsilon \rightarrow 0$ and we may assume a regular perturbation expansion for these roots of the form

$$x_k(\varepsilon) = 2e^{2\pi i k/3} + \sum_{n=1}^{\infty} a_{n,k} \varepsilon^n, \quad k = 1, 2, 3. \tag{7.2.3}$$

Substituting (7.2.3) into (7.2.1) and comparing powers of ε , as in Example 1 of Sec. 7.1, gives a sequence of equations which determine the coefficients $a_{n,k}$.

To track down the three missing roots we first estimate their orders of magnitude as $\varepsilon \rightarrow 0$. We do this by considering all possible dominant balances between pairs of terms in (7.2.1). There are four terms in (7.2.1) so there are six pairs to consider:

- (a) Suppose $\varepsilon^2 x^6 \sim \varepsilon x^4$ ($\varepsilon \rightarrow 0$) is the dominant balance. Then $x = O(\varepsilon^{-1/2})$ ($\varepsilon \rightarrow 0$). It follows that the terms $\varepsilon^2 x^6$ and εx^4 are both $O(\varepsilon^{-1})$. But $\varepsilon x^4 \ll x^3 = O(\varepsilon^{-3/2})$ as $\varepsilon \rightarrow 0$, so x^3 is the biggest term in the equation and is not balanced by any other term. Thus, the assumption that $\varepsilon^2 x^6$ and εx^4 are the dominant terms as $\varepsilon \rightarrow 0$ is inconsistent.
- (b) Suppose $\varepsilon x^4 \sim x^3$ as $\varepsilon \rightarrow 0$. Then $x = O(\varepsilon^{-1})$. It follows that $\varepsilon x^4 \sim x^3 = O(\varepsilon^{-3})$. But $x^3 \ll \varepsilon^2 x^6 = O(\varepsilon^{-4})$ as $\varepsilon \rightarrow 0$. Thus, $\varepsilon^2 x^6$ is the largest term in the equation. Hence, the original assumption is again inconsistent.
- (c) Suppose $\varepsilon^2 x^6 \sim 8$ so that $x = O(\varepsilon^{-1/3})$ ($\varepsilon \rightarrow 0$). Hence, $x^3 = O(\varepsilon^{-1})$ is the largest term, which is again inconsistent.
- (d) Suppose $\varepsilon x^4 \sim 8$ so that $x = O(\varepsilon^{-1/4})$ ($\varepsilon \rightarrow 0$). Then $x^3 = O(\varepsilon^{-3/4})$ is the biggest term, which is also inconsistent.
- (e) Suppose $x^3 \sim 8$. Then $x = O(1)$. This is a consistent assumption because the other two terms in the equation, $\varepsilon^2 x^6$ and εx^4 , are negligible compared with x^3 and 8, and we recover the three roots of the unperturbed equation $x = 2, 2\omega$, and $2\omega^2$.
- (f) Suppose $\varepsilon^2 x^6 \sim x^3$ ($\varepsilon \rightarrow 0$). Then $x = O(\varepsilon^{-2/3})$. This is consistent because $\varepsilon^2 x^6 \sim x^3 = O(\varepsilon^{-2})$ is bigger than $\varepsilon x^4 = O(\varepsilon^{-5/3})$ and $8 = O(1)$ as $\varepsilon \rightarrow 0$.

Thus, the magnitudes of the three missing roots are $O(\varepsilon^{-2/3})$ as $\varepsilon \rightarrow 0$. This result is a clue to the structure of the perturbation series for the missing roots. In particular, it suggests a *scale transformation* for the variable x :

$$x = \varepsilon^{-2/3} y. \tag{7.2.4}$$

Substituting (7.2.4) into (7.2.1) gives

$$y^6 - y^3 + 8\varepsilon^2 - \varepsilon^{1/3} y^4 = 0. \tag{7.2.5}$$

This is now a *regular* perturbation problem for y in the parameter $\varepsilon^{1/3}$ because the unperturbed problem $y^6 - y^3 = 0$ has six roots $y = 1, \omega, \omega^2, 0, 0, 0$. Now, no roots disappear in the limit $\varepsilon^{1/3} \rightarrow 0$.

The perturbative corrections to these roots may be found by assuming a regular perturbation expansion in powers of $\varepsilon^{1/3}$ (it would not be possible to match powers in an expansion having only integral powers of ε):

$$y = \sum_{n=0}^{\infty} y_n(\varepsilon^{1/3})^n. \quad (7.2.6)$$

Having established that we are dealing with a singular perturbation problem, it is no surprise that the perturbation series for the roots x is not a series in integral powers of ε .

Nevertheless, when $y_0 = 0$ we find that $y_1 = 0$ and $y_2 = 2, 2\omega$, and $2\omega^2$. Thus, since the first two terms in this series vanish, $x = \varepsilon^{-2/3}y$ is not really $O(\varepsilon^{-2/3})$ but rather $O(1)$ and we have reproduced the three finite roots near $x = 2, 2\omega, 2\omega^2$. Moreover, only every third coefficient in (7.2.6), y_2, y_5, y_8, \dots , is nonvanishing, so we have also reproduced the regular perturbation series in (7.2.3)!

Example 2 *Appearance of a boundary layer.* The boundary-value problem

$$\varepsilon y'' - y' = 0, \quad y(0) = 0, y(1) = 1, \quad (7.2.7)$$

is a singular perturbation problem because the associated unperturbed problem

$$-y' = 0, \quad y(0) = 0, y(1) = 1, \quad (7.2.8)$$

has no solution. (The solution to this first-order differential equation, $y = \text{constant}$, cannot satisfy both boundary conditions.) The solution to (7.2.7) cannot have a regular perturbation expansion of the form $y = \sum_{n=0}^{\infty} y_n(x)\varepsilon^n$ because y_0 does not exist.

There is a close parallel between this example and the previous one. Here, the highest derivative is multiplied by ε and in the limit $\varepsilon \rightarrow 0$ the unperturbed solution loses its ability to satisfy the boundary conditions because a solution is lost. In the previous example the highest power of x is multiplied by ε and in the limit $\varepsilon \rightarrow 0$ some roots are lost.

The exact solution to (7.2.7) is easy to find:

$$y(x) = \frac{e^{x/\varepsilon} - 1}{e^{1/\varepsilon} - 1}. \quad (7.2.9)$$

This function is plotted in Fig. 7.2 for several small positive values of ε . For very small but nonzero ε it is clear from Fig. 7.2 that y is almost constant except in a very narrow interval of thickness $O(\varepsilon)$ at $x = 1$, which is called a *boundary layer*. Thus, outside the boundary layer the exact solution satisfies the left boundary condition $y(0) = 0$ and almost but not quite satisfies the unperturbed equation $y' = 0$.

It is not obvious how to construct a perturbative approximation to a differential equation whose highest derivative is multiplied by ε until it is known how to construct an analytical expression for the zeroth-order approximation. A new technique called asymptotic matching must be introduced (see Sec. 7.4 and Chap. 9) to solve this problem.

Example 3 *Appearance of rapid variation on a global scale.* In the previous example we saw that the exact solution varies rapidly in the neighborhood of $x = 1$ for small ε and develops a discontinuity there in the limit $\varepsilon \rightarrow 0+$. A solution to a boundary-value problem may also develop discontinuities throughout a large region as well as in the neighborhood of a point.

The boundary-value problem $\varepsilon y'' + y = 0$ [$y(0) = 0, y(1) = 1$] is a singular perturbation problem because when $\varepsilon = 0$, the solution to the unperturbed problem, $y = 0$, does not satisfy the boundary condition $y(1) = 1$. The exact solution, when ε is not of the form $(n\pi)^{-2}$ ($n = 0, 1, 2, \dots$), is $y(x) = \sin(x/\sqrt{\varepsilon})/\sin(1/\sqrt{\varepsilon})$. Observe that $y(x)$ becomes discontinuous throughout the inter-

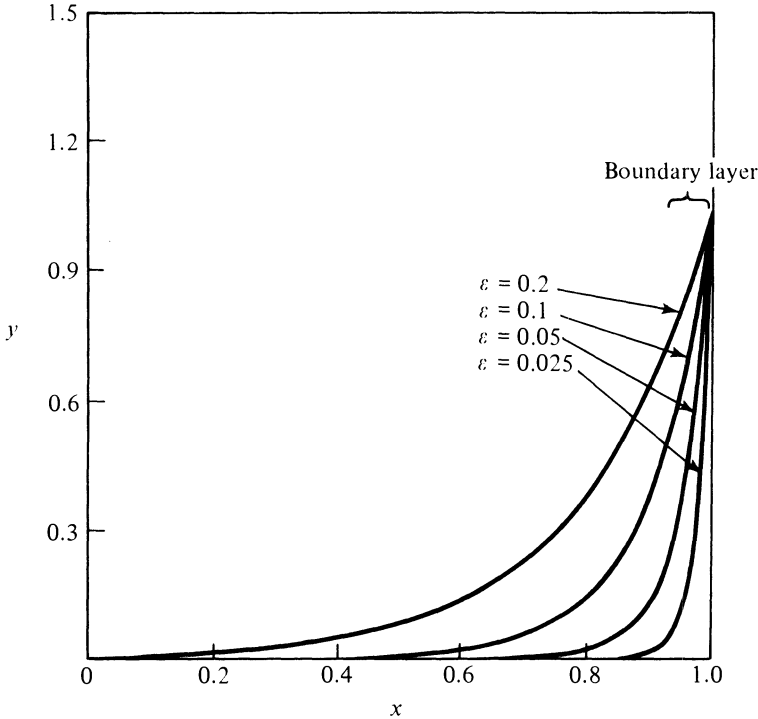


Figure 7.2 A plot of $y(x) = (e^{x/\epsilon} - 1)/(e^{1/\epsilon} - 1)$ ($0 \leq x \leq 1$) for $\epsilon = 0.2, 0.1, 0.05, 0.025$. When ϵ is small $y(x)$ varies rapidly near $x = 1$; this localized region of rapid variation is called a boundary layer. When ϵ is negative the boundary layer is at $x = 0$ instead of $x = 1$. This abrupt jump in the location of the boundary layer as ϵ changes sign reflects the singular nature of the perturbation problem.

val $0 \leq x \leq 1$ in the limit $\epsilon \rightarrow 0+$ (see Fig. 7.3). When $\epsilon = (\pi n)^{-2}$, there is no solution to the boundary-value problem.

When the solution to a differential-equation perturbation problem varies rapidly on a global scale for small ϵ , it is not obvious how to construct a leading-order perturbative approximation to the exact solution. The best procedure that has evolved is called WKB theory (see Chap. 10).

Example 4 *Perturbation theory on an infinite interval.* The initial-value problem

$$y'' + (1 - \epsilon x)y = 0, \quad y(0) = 1, \quad y'(0) = 0, \tag{7.2.10}$$

is a regular perturbation problem in ϵ over the finite interval $0 \leq x \leq L$. In fact, the perturbation solution is just

$$y(x) = \cos x + \epsilon\left(\frac{1}{4}x^2 \sin x + \frac{1}{4}x \cos x - \frac{1}{4} \sin x\right) + \epsilon^2\left(-\frac{1}{32}x^4 \cos x + \frac{5}{48}x^3 \sin x + \frac{7}{16}x^2 \cos x - \frac{7}{16}x \sin x\right) + \dots, \tag{7.2.11}$$

which converges for all x and ϵ , with increasing rapidity as $\epsilon \rightarrow 0+$ for fixed x .

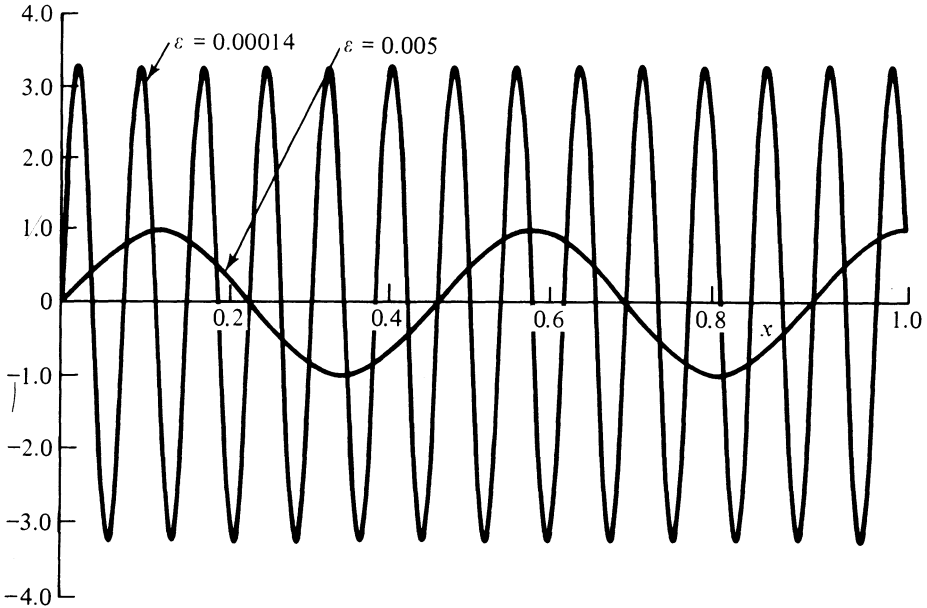


Figure 7.3 A plot of $y(x) = [\sin(x\epsilon^{-1/2})]/[\sin(\epsilon^{-1/2})]$ ($0 \leq x \leq 1$) for $\epsilon = 0.005$ and 0.00014 . As ϵ gets smaller the oscillations become more violent; as $\epsilon \rightarrow 0+$, $y(x)$ becomes discontinuous over the entire interval. The WKB approximation is a perturbative method commonly used to describe functions like $y(x)$ which exhibit rapid variation on a global scale.

However, this same initial-value problem must be reclassified as a singular perturbation problem over the semi-infinite interval $0 \leq x < \infty$. While the exact solution does approach the solution to the unperturbed problem as $\epsilon \rightarrow 0+$ for fixed x , it does not do so uniformly for all x (see Fig. 7.4). The zeroth-order solution is bounded and oscillatory for all x . But when $\epsilon > 0$, local analysis of the exact solution for large x shows that it is a linear combination of exponentially increasing and decreasing functions (Prob. 7.20). This change in character of the solution occurs because it is certainly wrong to neglect ϵx compared with 1 when x is bigger than $1/\epsilon$. In fact, a more careful argument shows that the term ϵx is not a small perturbation unless $x \ll \epsilon^{-1/2}$ (Prob. 7.20).

Example 4 shows that the interval itself can determine whether a perturbation problem is regular or singular. We examine more examples having this property in the next section on eigenvalue problems. The feature that is common to all such examples is that an n th-order perturbative approximation bears less and less resemblance to the exact solution as x increases.

For these sorts of problems Chap. 11 introduces new perturbative procedures called multiple-scale methods which substantially improve the rather poor predictions of ordinary perturbation theory. The particular problem in Example 4 is reconsidered in Prob. 11.13.

Example 5 *Roots of a high-degree polynomial.* When a perturbation problem is regular, the perturbation series is convergent and the exact solution is a smooth analytic function of ϵ for

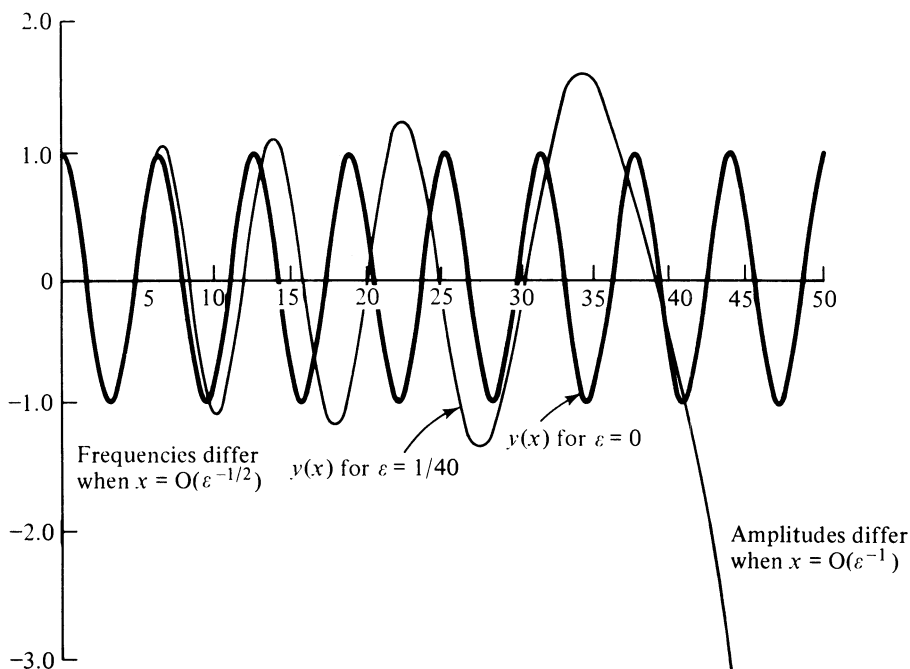


Figure 7.4 Exact solutions to the initial-value problem $y'' + (1 - \varepsilon x)y = 0$ [$y(0) = 1$, $y'(0) = 0$] in (7.2.10) for $\varepsilon = 0$ and $\varepsilon = \frac{1}{40}$. Although this is a regular perturbation problem on the finite interval $0 \leq x \leq L$, it is a singular perturbation problem on the infinite interval $0 \leq x \leq \infty$ because the perturbed solution ($\varepsilon > 0$) is not close to the unperturbed solution ($\varepsilon = 0$), no matter how small ε is. When $x = O(\varepsilon^{-1/2})$ the frequencies begin to differ (the curves become phase shifted) and when $x = O(\varepsilon^{-1})$ the amplitudes differ (one curve remains finite while the other grows exponentially).

sufficiently small ε . However, just what is “sufficiently small” may vary enormously from problem to problem. A striking example by Wilkinson concerns the roots of the polynomial

$$\prod_{k=1}^{20} (x - k) + \varepsilon x^{19} = x^{20} - (210 - \varepsilon)x^{19} + \cdots + 20! \quad (7.2.12)$$

The perturbation εx^{19} is regular, since no roots are lost in the limit $\varepsilon \rightarrow 0$; the roots of the unperturbed polynomial lie at 1, 2, 3, ..., 20.

Let us now take $\varepsilon = 10^{-9}$ so that the perturbation in the coefficient of x^{19} is of relative magnitude $10^{-9}/210$, or roughly 10^{-11} . For such a small regular perturbation one might expect the 20 roots to be only very slightly displaced from their $\varepsilon = 0$ values. The actual displaced roots are given in Table 7.1. One is surprised to find that while some roots are relatively unchanged by the perturbation, others have paired into complex conjugates. The qualitative effect on the roots of varying ε is shown in Figs. 7.5 and 7.6. In these plots the paths of the roots are traced as a function of ε . As $|\varepsilon|$ increases, the roots coalesce into pairs of complex conjugate roots. Evidently, a “small” perturbation is one for which $|\varepsilon| < 10^{-11}$, while $|\varepsilon| \gtrsim 10^{-10}$ is a “large” perturbation for at least some of the roots. Low-order regular perturbation theory may be used to understand this behavior (Probs. 7.22 and 7.23).

Table 7.1 Roots of the Wilkinson polynomial (7.2.12) with $\varepsilon = 10^{-9}$

The first column lists the unperturbed ($\varepsilon = 0$) roots 1, 2, ..., 20; the second column gives the results of first-order perturbation theory (see Prob. 7.22); the third column gives the exact roots. The unperturbed roots at 13 and 14, 15 and 16, and 17 and 18 are perturbed into complex-conjugate pairs. Observe that while first-order perturbation theory is moderately accurate for the real perturbed roots near 1, 2, ..., 12, 19, 20, it cannot predict the locations of the complex roots (but see Prob. 7.23)

Unperturbed root	First-order perturbation theory	Exact root
1	1.000 000 000 0	1.000 000 000 0
2	2.000 000 000 0	2.000 000 000 0
3	3.000 000 000 0	3.000 000 000 0
4	4.000 000 000 0	4.000 000 000 0
5	5.000 000 000 0	5.000 000 000 0
6	5.999 999 941 8	5.999 999 941 8
7	7.000 002 542 4	7.000 002 542 4
8	7.999 994 030 4	7.999 994 031 5
9	9.000 839 327 5	9.000 841 033 5
10	9.992 405 941 6	9.992 518 124 0
11	11.046 444 571	11.050 622 592
12	11.801 496 835	11.832 935 987
13	13.605 558 629	13.349 018 036 ± 0.532 765 750 <i>i</i>
14	12.667 031 557	
15	17.119 065 220	15.457 790 724 ± 0.899 341 526 2 <i>i</i>
16	13.592 486 027	
17	18.904 402 150	17.662 434 477 ± 0.704 285 236 9 <i>i</i>
18	17.004 413 300	
19	19.309 013 459	19.233 703 334
20	19.956 900 195	19.950 949 654

This example shows that the roots of high-degree polynomials may be extraordinarily sensitive to changes in the coefficients of the polynomial, even though the perturbation problem so obtained is regular. It should serve as ample warning to a “number cruncher” not to trust computer output without sufficient understanding of the nature of the problem being solved.

(I) 7.3 PERTURBATION METHODS FOR LINEAR EIGENVALUE PROBLEMS

In this section we show how perturbation theory can be used to approximate the eigenvalues and eigenfunctions of the Schrödinger equation

$$\left[-\frac{d^2}{dx^2} + V(x) + W(x) - E \right] y(x) = 0, \tag{7.3.1}$$

subject to the boundary condition

$$\lim_{|x| \rightarrow \infty} y(x) = 0. \tag{7.3.2}$$

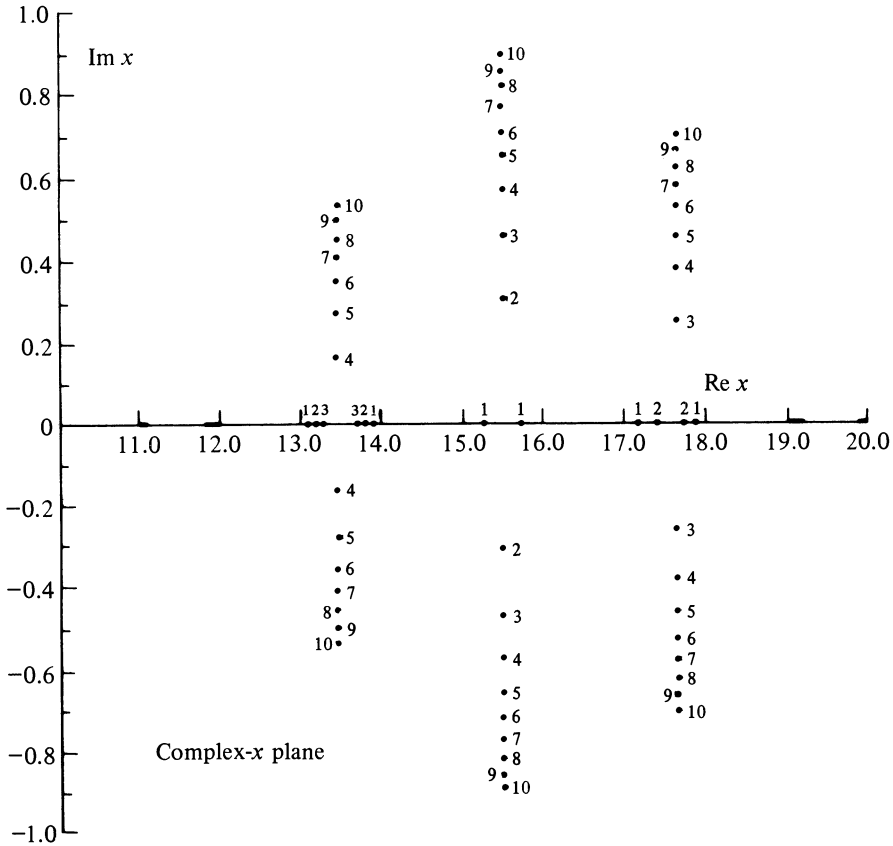


Figure 7.5 Roots of the Wilkinson polynomial $(x - 1)(x - 2)(x - 3) \cdots (x - 20) + \epsilon x^{19}$ in (7.2.12) for 11 values of ϵ . When $\epsilon = 0$ the roots shown are 10, 11, ..., 20. As ϵ is allowed to increase very slowly, the roots move toward each other in pairs along the real- x axis and then veer off in opposite directions into the complex- x plane. We have plotted the roots for $\epsilon = 0, 10^{-10}, 2 \times 10^{-10}, 3 \times 10^{-10}, \dots, 10^{-9}$. Some of the roots are numbered to indicate the value of ϵ to which they correspond; that is, 6 means $\epsilon = 6 \times 10^{-10}$, 3 means $\epsilon = 3 \times 10^{-10}$, and so on. The roots starting at 11, 12, 19, and 20 move too slowly to be seen as individual dots. We conclude from this plot that very slight changes in the coefficients of a polynomial can cause drastic changes in the values of some of the roots; one must be cautious when performing numerical calculations.

In (7.3.1) E is called the energy eigenvalue and $V + W$ is called the potential. We assume that $V(x)$ and $W(x)$ are continuous functions and that both $V(x)$ and $V(x) + W(x)$ approach ∞ as $|x| \rightarrow \infty$.

We suppose that the function $V(x) + W(x)$ is so complicated that (7.3.1) is not soluble in closed form. One can still prove from the above assumptions that nontrivial solutions $[y(x) \neq 0]$ satisfying (7.3.1) and (7.3.2) exist for special discrete values of E , the allowed eigenvalues of the equation (see Sec. 1.8). On the other hand, we assume that removing the term $W(x)$ from (7.3.1) makes the equation an

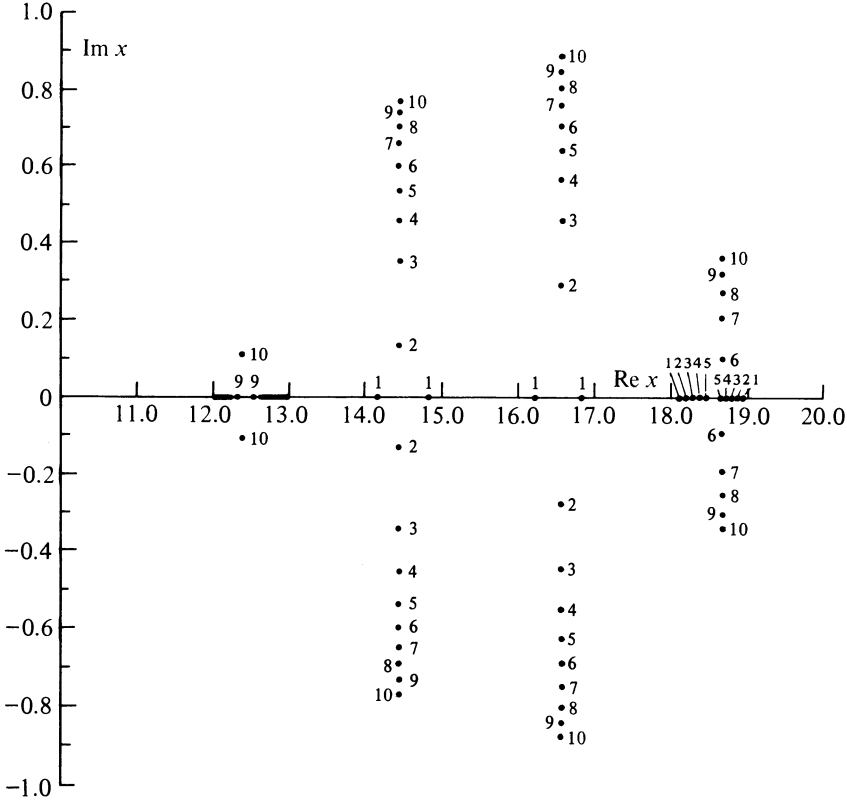


Figure 7.6 Same as in Fig. 7.5 except that the values of ϵ are $0, -10^{-10}, -2 \times 10^{-10}, -3 \times 10^{-10}, \dots, -10^{-9}$. The roots pair up and veer off into the complex- x plane, but the pairs are not the same as in Fig. 7.5.

exactly soluble eigenvalue problem. This suggests using perturbation theory to solve the family of eigenvalue problems in which $W(x)$ is replaced by $\epsilon W(x)$:

$$\left[-\frac{d^2}{dx^2} + V(x) + \epsilon W(x) - E \right] y(x) = 0. \tag{7.3.3}$$

Our assumptions on the nature of $V(x)$ and $W(x)$ leave no choice about where to introduce the parameter ϵ if the unperturbed problem is to be exactly soluble.

Example 1 *An exactly soluble eigenvalue problem.* Several exactly soluble eigenvalue problems are given in Sec. 1.8. One such example, which is used extensively in this section, is obtained if we take $V(x) = x^2/4$. The unperturbed problem is the Schrödinger equation for the quantum-mechanical harmonic oscillator, which is just the parabolic cylinder equation

$$-y'' + \frac{x^2}{4}y - Ey = 0. \tag{7.3.4}$$

We have already shown that solutions to this equation behave like $e^{\pm x^2/4}$ as $|x| \rightarrow \infty$.

There is a discrete set of values of E for which a solution that behaves like $e^{-x^2/4}$ as $x \rightarrow \infty$ also behaves like $e^{-x^2/4}$ as $x \rightarrow -\infty$ (see Example 4 of Sec. 3.5 and Example 9 of Sec. 3.8). These values of E are

$$E = n + \frac{1}{2}, \quad n = 0, 1, 2, \dots, \tag{7.3.5}$$

and the associated eigenfunctions are parabolic cylinder functions

$$y_n(x) = D_n(x) = e^{-x^2/4} \text{He}_n(x), \tag{7.3.6}$$

where $\text{He}_n(x)$ is the Hermite polynomial of degree n : $\text{He}_0(x) = 1$, $\text{He}_1(x) = x$, $\text{He}_2(x) = x^2 - 1, \dots$

In general, once an eigenvalue E_0 and an eigenfunction $y_0(x)$ of the unperturbed problem

$$\left[-\frac{d^2}{dx^2} + V(x) - E_0 \right] y_0(x) = 0 \tag{7.3.7}$$

have been found, we may seek a perturbative solution to (7.3.3) of the form

$$E = \sum_{n=0}^{\infty} E_n \varepsilon^n, \tag{7.3.8}$$

$$y(x) = \sum_{n=0}^{\infty} y_n(x) \varepsilon^n. \tag{7.3.9}$$

Substituting (7.3.8) and (7.3.9) into (7.3.3) and comparing powers of ε gives the following sequence of equations:

$$\left[-\frac{d^2}{dx^2} + V(x) - E_0 \right] y_n(x) = -W y_{n-1}(x) + \sum_{j=1}^n E_j y_{n-j}(x), \tag{7.3.10}$$

$n = 1, 2, 3, \dots,$

whose solutions must satisfy the boundary conditions

$$\lim_{|x| \rightarrow \infty} y_n(x) = 0, \quad n = 1, 2, 3, \dots \tag{7.3.11}$$

Equation (7.3.10) is linear and inhomogeneous. The associated homogeneous equation is just the unperturbed problem and thus is soluble by assumption. However, technically speaking, only *one* of the two linearly independent solutions of the unperturbed problem (the one that satisfies the boundary conditions) is assumed known. Therefore, we proceed by the method of reduction of order (see Sec. 1.4); to wit, we substitute

$$y_n(x) = y_0(x) F_n(x), \tag{7.3.12}$$

where $F_0(x) = 1$, into (7.3.10). Simplifying the result using (7.3.7) and multiplying by the integrating factor $y_0(x)$ gives

$$\frac{d}{dx} [y_0^2(x) F_n'(x)] = y_0^2(x) \left[W(x) F_{n-1}(x) - \sum_{j=1}^n E_j F_{n-j}(x) \right]. \tag{7.3.13}$$

If we integrate this equation from $-\infty$ to ∞ and use $y_0^2(x)F_n'(x) = y_0(x)y_n'(x) - y_0'(x)y_n(x) \rightarrow 0$ as $|x| \rightarrow \infty$, we obtain the formula for the coefficient E_n :

$$E_n = \frac{\int_{-\infty}^{\infty} y_0(x) \left[W(x)y_{n-1}(x) - \sum_{j=1}^{n-1} E_j y_{n-j}(x) \right] dx}{\int_{-\infty}^{\infty} y_0^2(x) dx}, \quad n = 1, 2, 3, \dots, \quad (7.3.14)$$

from which we have eliminated all reference to $F_n(x)$. [The sum on the right side of (7.3.14) is defined to be 0 when $n = 1$.]

Integrating (7.3.13) twice gives the formula for $y_n(x)$:

$$y_n(x) = y_0(x) \int_a^x \frac{dt}{y_0^2(t)} \int_{-\infty}^t ds y_0(s) \left[W(s)y_{n-1}(s) - \sum_{j=1}^n E_j y_{n-j}(s) \right], \quad n = 1, 2, 3, \dots \quad (7.3.15)$$

Observe that in (7.3.15) a is an arbitrary number at which we choose to impose $y_n(a) = 0$. This means we have fixed the overall normalization of $y(x)$ so that $y(a) = y_0(a)$ [assuming that $y_0(a) \neq 0$]. If $y_0(t)$ vanishes between a and x , the integral in (7.3.15) seems formally divergent; however, $y_n(x)$ satisfies a differential equation (7.3.10) which has no finite singular points. Thus, it is possible to define $y_n(x)$ everywhere as a finite expression (see Prob. 7.24).

Equations (7.3.14) and (7.3.15) together constitute an iterative procedure for calculating the coefficients in the perturbation series for E and $y(x)$. Once the coefficients $E_0, E_1, \dots, E_{n-1}, y_0, y_1, \dots, y_{n-1}$ are known, (7.3.14) gives E_n , and once E_n has been calculated (7.3.15) gives y_n . The remaining question is whether or not these perturbation series are convergent.

Example 2 *A regular perturbative eigenvalue problem.* Let $V(x) = x^2/4$ and $W(x) = x$. It may be shown (Prob. 7.25) that the perturbation series for $y(x)$ is convergent for all ϵ and that the series for E has vanishing terms of order ϵ^n for $n \geq 3$. This is a regular perturbation problem.

Example 3 *A singular perturbative eigenvalue problem.* It may be shown (Prob. 7.26) that if $V(x) = x^2/4$ and $W(x) = x^4/4$, then the perturbation series for the smallest eigenvalue for positive ϵ is

$$E(\epsilon) \sim \frac{1}{2} + \frac{3}{4}\epsilon - \frac{21}{8}\epsilon^2 + \frac{333}{16}\epsilon^3 + \dots, \quad \epsilon \rightarrow 0+. \quad (7.3.16)$$

The terms in this series appear to be getting larger and suggest that this series may be divergent for all $\epsilon \neq 0$. Indeed, (7.3.16) diverges for all ϵ because the n th term satisfies $E_n \sim -(-3)^n \Gamma(n + \frac{1}{2}) \sqrt{6/\pi} 3^{n/2} (n \rightarrow \infty)$. (This is a nontrivial result that we do not explain here.)

The divergence of the perturbation series in Example 3 indicates that the perturbation problem is singular. A simple way to observe the singular behavior is to compare $e^{-x^2/4}$, the controlling factor of the large- x behavior of the unperturbed ($\epsilon = 0$) solution, with $e^{-x^3\sqrt{\epsilon}/6}$, the controlling factor of the large- x behavior for $\epsilon \neq 0$. There is an abrupt change in the nature of the solution when we pass to the limit ($\epsilon \rightarrow 0+$). This phenomenon occurs because the perturbing term $\epsilon x^4/4$ is not small compared with $x^2/4$ when x is large.

If the functions $V(x)$ and $W(x)$ in Example 3 were interchanged, then the resulting eigenvalue problem would be a regular perturbation problem because εx^2 is a small perturbation of x^4 for all $|x| < \infty$. However, the unperturbed problem, $(-d^2/dx^2 + x^4/4 - E_0)y_0(x) = 0$, is not soluble in closed form. Thus, it would not be possible to use (7.3.14) and (7.3.15) to compute the coefficients in the perturbation series analytically.

Also note that if the boundary conditions in Example 3 were given at $x = \pm A$, $A < \infty$, then the perturbation theory would be regular. This is because here εx^4 is a small perturbation of x^2 . However, it is much more difficult to solve the unperturbed problem on a finite interval.

Thus, one is forced to accept a solution to Example 3 in the form of a divergent series. Fortunately, this series is one of many that may be summed by Padé theory to give a finite and unique result (see Sec. 8.3).

Example 4 *Another regular perturbation problem.* When $V = x^2/4$ and $W = |x|$ the perturbation problem is regular. But unlike the problem in Example 2, this perturbation series is not convergent for all ε ; the series in (7.3.8) and (7.3.9) have finite radii of convergence. The significance of the finite radius of convergence is discussed in Sec. 7.5.

(D) 7.4 ASYMPTOTIC MATCHING

The purpose of this section is to introduce the notion of matched asymptotic expansions. Asymptotic matching is an important perturbative method which is used often in both boundary-layer theory (Chap. 9) and WKB theory (Chap. 10) to determine analytically the approximate global properties of the solution to a differential equation. Asymptotic matching is usually used to determine a uniform approximation to the solution of a differential equation and to find other global properties of differential equations such as eigenvalues. Asymptotic matching may also be used to develop approximations to integrals.

The principle of asymptotic matching is simple. The interval on which a boundary-value problem is posed is broken into a sequence of two or more *overlapping* subintervals. Then, on each subinterval perturbation theory is used to obtain an asymptotic approximation to the solution of the differential equation valid on that interval. Finally, the matching is done by requiring that the asymptotic approximations have the same functional form on the overlap of every pair of intervals. This gives a sequence of asymptotic approximations to the solution of the differential equation; by construction, each approximation satisfies all the boundary conditions given at various points on the interval. Thus, the end result is an approximate solution to a boundary-value problem valid over the entire interval.

Asymptotic matching bears a slight resemblance to an elementary technique for solving boundary-value problems called *patching*. Patching is helpful when the differential equation can be solved in closed form. Here is a simple example:

Example 1 *Patching.* The method of patching may be used to solve the boundary-value problem $y'' - y = e^{-|x|}$ [$y(\pm\infty) = 0$]. There are two regions to consider. When $x \leq 0$, the most general solution which satisfies the boundary condition $y(-\infty) = 0$ is

$$y(x) = ae^x + \frac{1}{2}xe^x, \quad (7.4.1)$$