PERTURBATION METHODS SPRING 2016

LESSON 1: Preliminaries

The aims and terminology of Perturbation Theory are introduced.

1 Introduction

Mathematical models of scientific problems seldom yield exact solutions, especially when the models are nonlinear. One must therefore resort to numerics, or to approximate analytical techniques. Perturbation theory is a collection of methods that allow the construction of approximate series solutions to problems that involve a small parameter ϵ . The basic requirement is that the reduced problem, corresponding to $\epsilon = 0$, be tractable, thus providing the *leading-order solution*. Corrections to this solution, typically leading to an asymptotic series, can then be constructed by means of an iterative procedure.

Put another way, suppose that one has the solution to a given mathematical problem. Perturbation theory is then concerned with how to solve a neighboring problem, *i.e.*, one that is a slightly altered version of the original. The measure of the slight alteration is the small parameter ϵ . For a PDE model, say, the alteration may be in the shape of the domain, in the auxiliary conditions, or in the PDE itself.

This course will focus largely on ODE and PDE models where the solution $u = f(s; \epsilon)$ is a function of the parameter ϵ and the independent variable s. Both u and s may be scalars or vectors. The leading approximation corresponding to $\epsilon = 0$ is $u_0 = f(s; 0)$. Perturbation theory develops corrections u_1, u_2, \ldots to the leading approximation when ϵ is small but nonzero. The analysis has both global and local aspects, i.e., s may range globally over a broad domain \mathcal{D}_s while ϵ is confined to a local neighborhood of zero.

Although we shall discuss a number of established procedures, perturbation theory remains an art. When faced with a new problem, a significant amount of trial and error may ensue. The intent is to develop the skills, gain experience and most importantly, acquire a determination to exploit the small parameter. The approach will be careful but formal rather than rigorous; we shall aim for a certain consistency but not adopt a theorem-proof approach.

Roughly speaking, when small effects (*i.e.*, small changes in formulation) lead to small consequences (*i.e.*, small changes in the solution), we speak of a *regular-perturbation problem*; otherwise the problem is said to be a *singular-perturbation problem*. The latter are the more frequent and the more interesting, and will receive a bulk of the attention in this course.

2 Formal computation of perturbation series

Unless indicated otherwise, ϵ will be treated as a small quantity approaching zero through positive values. The following simple examples present some key ideas.

2.1 Functions defined by algebraic equations

In each of the following examples, we set aside the fact that the quadratic can be solved exactly. Instead we look for a solution in the form of a perturbation series for small ϵ .

Example 2.1. Roots of $x^2 + (1 + \epsilon)x - 2 = 0$.

We write the equation as $x^2 + x - 2 = -\epsilon x$. Note that for $\epsilon = 0$ the problem reduces to

$$x^{2} + x - 2 = (x - 1)(x + 2) = 0.$$

yielding $x_0^{(1)}=1$ and $x_0^{(2)}=-2$ as the leading approximations to the two roots. In order to find corrections to the root $x^{(1)}$ for ϵ small but nonzero, we write the quadratic as

$$x = 1 - \epsilon \frac{x}{x+2}.$$

A correction to the root can be found by substituting $x = x_0^{(1)}$ on the RHS, leading to

$$x_1^{(1)} = 1 - \epsilon \frac{x_0^{(1)}}{x_0^{(1)} + 2} = 1 - \frac{1}{3}\epsilon.$$

This procedure can be continued in an obvious way to compute improved corrections, and a similar procedure can be employed for the second root. Alternatively, we can be guided by the above result to seek a series expansion for either root of the form

$$x = x_0 + \epsilon x_1 + \epsilon^2 x_2 + \cdots.$$

Substitution into the quadratic leads to

$$(x_0 + \epsilon x_1 + \epsilon^2 x_2 + \cdots)^2 + (1 + \epsilon)(x_0 + \epsilon x_1 + \epsilon^2 x_2 + \cdots) - 2 = 0,$$

or, upon expanding the square root and collecting like powers of ϵ , to

$$(x_0^2 + x_0 - 2) + \epsilon(2x_0x_1 + x_1 + x_0) + \epsilon^2(2x_0x_2 + x_1^2 + x_2 + x_1) + \dots = 0.$$

If the above equation is to hold for ϵ small but otherwise arbitrary, then the coefficient of each power of ϵ must separately vanish. Thus we are led to the hierarchy of equations

$$x_0^2 + x_0 - 2 = 0,$$

$$(2x_0 + 1)x_1 + x_0 = 0,$$

$$(2x_0 + 1)x_2 + x_1^2 + x_1 = 0,$$

$$\dots \dots \dots \dots \dots \dots$$

which can be solved to get

$$(x_0 - 1)(x_0 + 2) = 0,$$

$$x_1 = -\frac{x_0}{2x_0 + 1},$$

$$x_2 = -\frac{x_1^2 + x_1}{2x_0 + 1},$$

$$\dots \dots \dots$$

The first of the above equations yields the leading-order approximations to the two roots,, $x_0^{(1)} = 1$ and $x_0^{(2)} = -2$, while the subsequent equations yield the higher-order corrections for each. Straightforward computations yield the two roots as

$$x^{(1)} = 1 - \frac{1}{3}\epsilon + \frac{2}{27}\epsilon^2 + \cdots, ,$$

$$x^{(2)} = -2 - \frac{2}{3}\epsilon - \frac{2}{27}\epsilon^2 + \cdots.$$

Example 2.2. Roots of $x^2 + \epsilon x - \epsilon = 0$.

A power-series expansion in ϵ of the type used in the example above does not work. To see this, we substitute

$$x = x_0 + \epsilon x_1 + \epsilon^2 x_2 + \cdots$$

into the quadratic and collect terms as before to get

$$x_0^2 + \epsilon(2x_0x_1 + x_0 - 1) + \epsilon^2(2x_0x_2 + x_1^2 + x_1) + \dots = 0,$$

which leads to the hierarchy of equations

$$x_0^2 = 0,$$

$$2x_0x_1 = 1 - x_0,$$

$$2x_0x_2 = -x_1^2 - x_1,$$

The first of the above equations finds $x_0 = 0$ as a double root, and then the second offers the contradiction 0 = 1. To find our way out of the difficulty we return to the quadratic and write it as

$$x^2 = \epsilon - \epsilon x. \tag{2.1}$$

The leading-order approximation to each root, obtained by setting $\epsilon = 0$, is $x_0 = 0$, as already noted above. When this approximation is substituted into the RHS of the above equation to find the corrected approximation x_1 , the result is

$$x_1^2 = \epsilon - \epsilon x_0 = \epsilon,$$

yielding $x_1^{(1)} = \sqrt{\epsilon}$ and $x_1^{(2)} = -\sqrt{\epsilon}$.

[Alternatively the argument could be developed as follows. Since each root is zero for $\epsilon = 0$, it is reasonable to assume that each root will be small for ϵ nonzero but small. In equation (2.1) then, the dominant term on the RHS is ϵ , since the term ϵx , being the product of ϵ with a small quantity, must be smaller than ϵ . The dominant balance in equation (2.1) must therefore be between x^2 on the LHS and ϵ on the RHS, yielding the approximations $x = \pm \sqrt{\epsilon}$.

Proceeding in a similar fashion the next-stage correction to, say, the first root, is given by

$$(x_2^{(1)})^2 = \epsilon - \epsilon x_1^{(1)} = \epsilon - \epsilon^{3/2}.$$

Therefore, taking the positive square root again,

$$x_2^{(1)} = \left(\epsilon - \epsilon^{3/2}\right)^{1/2} = \sqrt{\epsilon}(1 - \sqrt{\epsilon})^{1/2} = \sqrt{\epsilon} - \frac{1}{2}\epsilon + \cdots,$$

where we have applied the binomial expansion of the square root in the last step. The upshot is that the expansion proceeds in a series of powers of $\sqrt{\epsilon}$. A similar fate awaits the second root.

Once the form of the expansion has been ascertained, we can begin with the ansatz that each root has a series expansion proceeding in half powers of ϵ , *i.e.*,

$$x = x_0 + \epsilon^{1/2} x_1 + \epsilon x_2 + \epsilon^{3/2} x_3 + \cdots$$

Substitution into the quadratic and proceeding as in Example 2.1 above leads to the results

$$x^{(1)} = \sqrt{\epsilon} - \frac{1}{2}\epsilon + \frac{1}{8}\epsilon^{3/2} + \cdots,$$

$$x^{(2)} = -\sqrt{\epsilon} - \frac{1}{2}\epsilon - \frac{1}{8}\epsilon^{3/2} + \cdots.$$

Example 2.3. Roots of $\epsilon x^2 + x + 1 = 0$.

We write the equation as

$$x + 1 = -\epsilon x^2$$
.

and note that for $\epsilon = 0$ the leading approximation determines only one root, $x_0^{(1)} = -1$. A power series expansion carried out along the lines of those in the two examples above yields

$$x^{(1)} = -1 - \epsilon - 2\epsilon^2 + \cdots. {(2.2)}$$

If the second root is to be found, then it is clear that in the limit $\epsilon \to 0$ the quadratic should remain a quadratic, *i.e.*, the ϵx^2 term must survive. This implies that x^2 must become unbounded as $\epsilon \to 0$. To ascertain how large x must be, we use the notion of dominant balance introduced in Example 1.2 above. Let

$$x = \frac{X}{\delta},$$

where $\delta(\epsilon) \to 0$ as $\epsilon \to 0$ and X is bounded. Then the quadratic becomes

$$\frac{\epsilon}{\delta^2}X^2 + \frac{1}{\delta}X + 1 = 0.$$

As $\epsilon \to 0$, the second term becomes unbounded like $1/\delta$, and therefore much larger than the third. We also recognize that the first term has to be reinstated. Therefore, dominant balance must occur between the first two terms, requiring in turn that

$$\frac{\epsilon}{\delta^2} = \frac{1}{\delta},$$

which leads to $\delta = \epsilon$. We can now proceed in one of two ways. For the second root we can assume an expansion of the form

$$x \sim \frac{1}{\epsilon} x_{-1} + x_0 + \epsilon x_1 + \cdots,$$

which, after the by-now-routine computations, yields

$$x^{(2)} = -\frac{1}{\epsilon} + 1 + \epsilon + \cdots$$
 (2.3)

Otherwise, we can rescale the original quadratic by setting $x = X/\epsilon$, which produces the equation

$$X^2 + X + \epsilon = 0.$$

The expansion

$$X = X_0 + \epsilon X_1 + \epsilon^2 X_2 + \cdots$$

is easily found to produce the following expansions for the two roots:

$$X^{(1)} = -\epsilon - \epsilon^2 - 2\epsilon^3 + \cdots,$$

$$X^{(2)} = -1 + \epsilon + \epsilon^2 + \cdots.$$

On setting $X = \epsilon x$, the original expansions (2.2) and (2.3) are recovered.

2.2 Function defined by an ODE

In the following example we seek the perturbation expansion in ϵ , of the solution $u(t, \epsilon)$ of an ODE, which is a function of the independent variable t and the small parameter ϵ .

Example 2.4. We consider a model of damped vertical motion. Let a particle of mass M be launched vertically upwards with initial speed U_0 . In the absence of any air resistance the mathematical model for the motion is

$$M\frac{dU}{dT} = -Mg, \ U(0) = U_0,$$

where U is the displacement and T the time. The solution, obtained through simple quadrature, is $U = U_0 - gT$. A convenient reference velocity for the model is U_0 , and reference time U_0/g . With

$$U = U_0 u$$
, $T = (U_0/q)t$,

the problem assumes the dimensionless form

$$\frac{du}{dt} = -1, \ u(0) = 1,$$

leading to the solution u = 1 - t.

Air resistance may be included in the model by assuming a drag force linearly dependent on velocity, a reasonable assumption if velocities are not too high. Then, the motion is governed by

$$M\frac{dU}{dT} = -Mg - KU, \ U(0) = U_0,$$

where the drag coefficient K has units of force/speed, or mass/time. The reference quantities introduced above lead to the dimensionless form

$$\frac{du}{dt} = -1 - (KU_0/Mg)u,$$

where the combination KU_0/Mg , a dimensionless drag constant, will be denoted by ϵ . The problem then becomes

$$\frac{du}{dt} = -1 - \epsilon u, \quad u(0) = 1. \tag{2.4}$$

The transformation to dimensionless form is important. We shall be concerned with developing a solution for 'small' drag. It is insufficient to simply state that small drag means a small value of K, because the numerical value of K depends on the choice of units. The important consideration is the size of K compared to another quantity having the same dimensions; here that quantity is Mg/U_0 . Put another way, ϵ measures the ratio of the damping force KU_0 to the gravitational pull Mg at the moment of launch.

It is possible to solve the damped model exactly, but let us leave the exact solution aside for a moment and attempt to develop an approximate solution in the form of a perturbation expansion by an iterative process. The basic idea is that the small term ϵu may be expected to have a small effect on the solution which, to a zeroth-order approximation, is simply the undamped version $u^{(0)} = 1 - t$. Higher-order approximations u^j may be constructed by solving the equation

$$\frac{du^{(j)}}{dt} = -1 - \epsilon u^{(j-1)}, \ u^{(j)}(0) = 1.$$

The first two are easily constructed to be

$$u^{(1)} = 1 - t + \epsilon(t^2/2 - t),$$

and

$$u^{(2)} = 1 - t + \epsilon(t^2/2 - t) + \epsilon^2(t^2/2 - t^3/6),$$

and the procedure may be continued indefinitely.

An alternative to the iteration is to simply assume a power-series expansion for the solution at the outset, i.e.,

$$u(t;\epsilon) = u_0(t) + \epsilon u_1(t) + \epsilon^2 u_2(t) + \cdots$$

Substitution of the expansion into the damped equation yields

$$\frac{du_0}{dt} + \epsilon \frac{du_1}{dt} + \epsilon^2 \frac{du_2}{dt} + \dots = -1 - \epsilon u_0 - \epsilon^2 u_1 + \dots$$

while the initial condition reads

$$u_0(0) + \epsilon u_1(0) + \epsilon^2 u_2(0) + \dots = 1.$$

Assuming that the above pair of equations holds for all value of ϵ in some interval containing zero, we may compare coefficients of powers of ϵ , thereby leading to a hierarchy of problems:

$$\epsilon^0: \frac{du_0}{dt} = -1, \ u_0(0) = 1,$$

$$\epsilon^1: \quad \frac{du_1}{dt} = -u_0, \ u_1(0) = 0,$$

$$\epsilon^2: \quad \frac{du_2}{dt} = -u_1, \ u_2(0) = 0,$$

and so on. The solutions are

$$u_0 = 1 - t$$
, $u_1 = t^2/2 - t$, $u_2 = t^2/2 - t^3/6$,

resulting in the same series expansion,

$$u = 1 - t + \epsilon \left(-t + \frac{t^2}{2} \right) + \epsilon^2 \left(\frac{t^2}{2} - \frac{t^3}{6} \right) + \cdots,$$
 (2.5)

as generated earlier by iteration.

2.3 Function defined by an integral

Our last example considers the expansion of a function $f(\epsilon)$ that is defined by an integral.

Example 2.5. Consider

$$f(\epsilon) = \int_0^\infty \frac{e^{-t}}{1 + \epsilon t} dt. \tag{2.6}$$

We expand $1/(1+\epsilon t)$ in a binomial series and integrate termwise, to get

$$f(\epsilon) = \int_0^\infty e^{-t} [1 - \epsilon t + \epsilon^2 t^2 - \epsilon^3 t^3 + \epsilon^4 t^4 + \cdots] dt$$

= 1 - \epsilon + 2! \epsilon^2 - 3! \epsilon^3 + 4! \epsilon^4 + \cdots . (2.7)

Here we have used the result

$$\int_0^\infty e^{-t}t^n dt = n! \quad \text{for positive integers } n.$$

Alternatively, the series (2.7) can be obtained by expanding $f(\epsilon)$ in a Taylor series about $\epsilon = 0$.

2.4 Quality of the series

Up to this point the focus in each of the examples above has been on the construction of the series. It is reasonable to ask how well the series approximates the function that is being expanded, and how the quality of the approximation depends upon the number of terms computed and upon ϵ if the function is $f(\epsilon)$, and upon both ϵ and t if the function is $f(t;\epsilon)$. These questions are difficult to answer in general, which we shall find as the course proceeds. However, the above examples are simple enough to supply the answers.

• For the quadratic equations considered in Examples 2.1 - 2.3, the exact solutions for the roots are easily written down. In Example 2.3, for instance, the exact roots are

$$x^{(1)} = \frac{-1 + (1 - 4\epsilon)^{1/2}}{2\epsilon}, \quad x^{(2)} = \frac{-1 - (1 - 4\epsilon)^{1/2}}{2\epsilon}.$$

Series expansions for the roots can be constructed by using the binomial expansion,

$$(1-4\epsilon)^{1/2} = 1 - 2\epsilon - 2\epsilon^2 - 4\epsilon^3 + \cdots$$

This a convergent series for $|\epsilon| < 1/4$. When substituted into the exact solutions above, it leads to the series expansions

$$x^{(1)} = -1 - \epsilon - 2\epsilon^2 + \cdots,$$

$$x^{(2)} = -\frac{1}{\epsilon} + 1 + \epsilon + \cdots,$$

which agree with those found in Example 2.3 above. As the series are convergent for ϵ sufficiently small, we know that the error of approximation can be made arbitrarily small by retaining a large enough number of terms in the series, and that the error declines for a fixed number of terms as $\epsilon \to 0$. Similar conclusions hold for Examples 2.1 and 2.2.

• In Example 2.4 the exact solution of the linear ODE problem (2.4) is

$$u(t;\epsilon) = \frac{(1+\epsilon)e^{-\epsilon t} - 1}{\epsilon}.$$

To write this expression as a series we replace $e^{-\epsilon t}$ by its Taylor polynomial

$$e^{-\epsilon t} = \sum_{k=0}^{n+1} \frac{(-1)^k \epsilon^k t^k}{k!} + R_{n+1},$$

where the remainder is given by

$$R_{n+1} = \frac{(-1)^{n+2} \epsilon^{n+2} t^{n+2}}{(n+2)!} e^{-\epsilon \tau}, \quad \text{where} \quad 0 \le \tau \le t.$$
 (2.8)

Substitution of the above expansion into the exact solution leads to the series

$$u = \left(\frac{1}{\epsilon} + 1\right) \sum_{k=0}^{n+1} \frac{(-1)^k \epsilon^k t^k}{k!} + \left(\frac{1}{\epsilon} + 1\right) R_{n+1} - \frac{1}{\epsilon}$$

$$= \sum_{k=0}^{n+1} \frac{(-1)^k \epsilon^{k-1} t^k}{k!} + \sum_{k=0}^{n+1} \frac{(-1)^k \epsilon^k t^k}{k!} + \left(\frac{1}{\epsilon} + 1\right) R_{n+1} - \frac{1}{\epsilon}$$

$$= \sum_{k=1}^{n+1} \frac{(-1)^k \epsilon^{k-1} t^k}{k!} + \sum_{k=0}^{n+1} \frac{(-1)^k \epsilon^k t^k}{k!} + \left(\frac{1}{\epsilon} + 1\right) R_{n+1}$$

$$= \sum_{k=0}^{n} \frac{(-1)^{k+1} \epsilon^k t^{k+1}}{(k+1)!} + \sum_{k=0}^{n} \frac{(-1)^k \epsilon^k t^k}{k!} + \frac{(-1)^{n+1} \epsilon^{n+1} t^{n+1}}{(n+1)!} + \left(\frac{1}{\epsilon} + 1\right) R_{n+1}$$

$$= \sum_{k=0}^{n} \epsilon^k \left(\frac{(-1)^k t^k}{k!} - \frac{(-1)^k t^{k+1}}{(k+1)!}\right) + \hat{R}_{n+1}, \tag{2.9}$$

where

$$\hat{R}_{n+1} = \frac{(-1)^{n+1} \epsilon^{n+1} t^{n+1}}{(n+1)!} + \left(\frac{1}{\epsilon} + 1\right) R_{n+1}.$$

Upon substitution for R_{n+1} from (2.8), the above expression becomes

$$\hat{R}_{n+1} = \frac{(-1)^{n+1} \epsilon^{n+1} t^{n+1}}{(n+1)!} + \frac{(-1)^{n+2} \epsilon^{n+1} t^{n+2}}{(n+2)!} e^{-\epsilon \tau} + \frac{(-1)^{n+2} \epsilon^{n+2} t^{n+2}}{(n+2)!} e^{-\epsilon \tau}, \tag{2.10}$$

and leads to the bound

$$|\hat{R}_{n+1}| \le \epsilon^{n+1} \left(\frac{t^{n+1}}{(n+1)!} + \frac{t^{n+2}}{(n+2)!} \right) + \epsilon^{n+2} \frac{t^{n+2}}{(n+2)!}. \tag{2.11}$$

In (2.9) we have a series expansion of $u(t;\epsilon)$ in powers of ϵ , and a remainder. The coefficients of the series are functions of t. The first few terms of the series are

$$u = 1 - t + \epsilon \left(-t + \frac{t^2}{2} \right) + \epsilon^2 \left(\frac{t^2}{2} - \frac{t^3}{6} \right) + \cdots,$$

in agreement with the series (2.5) found above directly from the ODE.

Let the time interval of interest be $t \in [0, T]$, where T is arbitrary but fixed. Then the bound (2.11) on the remainder $\hat{R}_{n+1}(t, n; \epsilon)$ can be further expressed as

$$|\hat{R}_{n+1}(t,n;\epsilon)| \le |\hat{R}_{n+1}(T,n;\epsilon)| = \epsilon^{n+1} \left(\frac{T^{n+1}}{(n+1)!} + \frac{T^{n+2}}{(n+2)!} \right) + \epsilon^{n+2} \frac{T^{n+2}}{(n+2)!}.$$
 (2.12)

The RHS is independent of t and approaches zero as $n \to \infty$ for any fixed ϵ . Thus the series (2.9) is uniformly convergent for $t \in [0,T]$, and hence the error in the entire interval can be made arbitrarily small by retaining a large enough number of terms. However, if $t \in [0,\infty)$, then it is not possible to obtain a t-independent bound on the remainder, and hence the series is only pointwise convergent. The larger t is, the greater is the number of terms needed to maintain a given accuracy.

• The series (2.7) generated for $f(\epsilon)$ defined by the integral (2.6) in Example 2.5 is divergent for all ϵ except $\epsilon = 0$, as can be easily seen by observing that the absolute value of the general term of the series, $n!\epsilon^n$, is unbounded for any $\epsilon > 0$ as $n \to \infty$. Thus this series would appear to offer little by way of an approximation to f. However, as we shall see shortly, the divergent series can indeed yield useful approximations as well, but in a sense that is quite different from that provided by a convergent series.

3 Series in perturbation theory

Perturbation problems generally lead to what are known as asymptotic series. Before introducing such series and examining how they differ from convergent series, we introduce some needed terminology.

3.1 Gauge functions and order relations

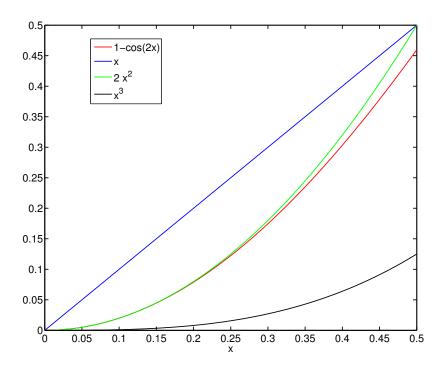


Figure 1: Comparison of $1 - \cos 2x$ with power functions x, x^2 and x^3 as $x \to 0$.

An intuitive sense of how a given function $f(\epsilon)$ behaves, or looks like, in the vicinity of a point, say $\epsilon = 0$, can be gathered by comparing $f(\epsilon)$ to a gauge function $g(\epsilon)$, with which one is more familiar and whose behavior one understands. For example, consider $f(\epsilon) = 1 - \cos 2\epsilon$. We know that f is a periodic function with period π , but our interest is in the behavior of f for small values of ϵ , i.e., near $\epsilon = 0$. It is a simple matter to check that $f(\epsilon) = 0$, but one wishes to know more. What does f look like when ϵ is small? How rapidly does it approach zero as $\epsilon \to 0$? We know several functions that vanish at the origin, such as the powers ϵ , ϵ^2 , ϵ^3 , etc. Let us compare the behavior of f with that of powers of increasing degree by considering the following limits.

$$\lim_{\epsilon \to 0} \frac{1 - \cos 2\epsilon}{\epsilon} = 0,$$

$$\lim_{\epsilon \to 0} \frac{1 - \cos 2\epsilon}{\epsilon^2} = 2,$$

$$\lim_{\epsilon \to 0} \frac{1 - \cos 2\epsilon}{\epsilon^3} = \infty.$$

The limiting values show that $f(\epsilon)$ approaches zero faster than the linear function, as fast as a multiple of the quadratic, and not as fast as the cubic. The graphs in Figure 1 reinforce these conclusions. Here we have used the powers as known functions with which to gauge the behavior of a more complicated function.

We introduce a new notation by rewriting the above results as follows.

$$1 - \cos 2\epsilon = o(\epsilon) \text{ as } \epsilon \to 0,$$

$$1 - \cos 2\epsilon = O(\epsilon^2) \text{ as } \epsilon \to 0,$$

$$\epsilon^3 = o(1 - \cos 2\epsilon) \text{ as } \epsilon \to 0.$$

Here, O and o are $order\ symbols$, defined below.

Def 1. f = O(g) as $\epsilon \to 0$ if there exist positive constants k and δ such that $|f| \le k|g|$ whenever $0 < \epsilon < \delta$.

Def 2. f = o(g) as $\epsilon \to 0$ if for any positive k however small there exists a positive δ such that $|f| \le k|g|$ whenever $0 < \epsilon < \delta$.

More simply, if $g(\epsilon)$ does not vanish in a neighborhood of $\epsilon = 0$ except possibly at $\epsilon = 0$, then it is a simple matter to show that if

$$\lim_{\epsilon \to 0} \left| \frac{f}{g} \right| = L,$$

where L is nonzero and bounded, then f = O(g), and if

$$\lim_{\epsilon \to 0} \frac{f}{q} = 0,$$

then f = o(g). The notations $f \ll g$ and $g \gg f$ are equivalent to f = o(g).

The mathematical statement $f(\epsilon) = O(g(\epsilon))$ is read as 'f is of order g' or 'f is big "Oh" of g.' The mathematical statement $f(\epsilon) = o(g(\epsilon))$ is read as 'f is little "Oh" of g' or 'f is of higher order than g.'

For example, $\sin \epsilon = O(\epsilon)$, $1 - \cos \epsilon = O(\epsilon^2)$, $\epsilon^{\alpha} = o(\epsilon^{\beta})$ for $\alpha > \beta$, $\epsilon^{\alpha} \ln(1/\epsilon) = o(1)$ for any $\alpha > 0$, $\sin(1/\epsilon) = O(1)$, $\ln(1 + \sin \epsilon) = O(\epsilon)$.

Remarks

- If the limit L is precisely 1 in **Def 1** above, then we write $f(\epsilon) \sim g(\epsilon)$ and say that f is asymptotically equal to, or behaves like, g as $\epsilon \to 0$.
- One prefers the sharpest estimate, e.g., $\sin \epsilon = O(\epsilon)$ rather than the formally correct but less informative choices $\sin \epsilon = O(1)$, o(1), $o(\sqrt{\epsilon})$, $o(\sqrt{\epsilon})$, etc. If the function has a Taylor expansion about $\epsilon = 0$ then the leading term in the expansion provides the sharpest order estimate.
- The mathematical order is distinct from the physical order of magnitude since the constant of proportionality is not taken into account. Thus, $k\epsilon = O(\epsilon)$ whether k = 1 or k = 10,000.
- In general, when one examines the order of $f(s;\epsilon)$, then the constants k, δ and L depend upon s. If they can be chosen independently of s for $s \in \mathcal{D}_s$, then the ordering is said to be *uniform* in \mathcal{D}_s . Thus, for $\epsilon \to 0$ and $f = \epsilon/s$, $f = O(\epsilon)$ uniformly in $1 \le s \le 2$ but not in $0 < s \le 2$, the nonuniformity occurring near s = 0. Similarly, $e^{-s/\epsilon} = o(\epsilon^{\alpha})$ for any $\alpha \ge 0$ for s > 0 but the ordering is not uniform at s = 0.

Def 3. The family of gauge functions $\{g_n(\epsilon)\}$ is said to define an asymptotic sequence or gauge sequence as $\epsilon \to 0$ if

$$g_{n+1}(\epsilon) = o(g_n(\epsilon))$$
 for $n = 0, 1, 2, 3, \cdots$ as $\epsilon \to 0$.

For example, $\{\epsilon^n\}$ is an asymptotic sequence as $\epsilon \to 0$, as is $\{1/\{\ln(1/\epsilon)\}^n\}$.

3.2 Asymptotic series

We are now ready to define an asymptotic series, also known as asymptotic expansion.

Def 4. Let $\{g_n(\epsilon)\}$ be a gauge sequence as $\epsilon \to 0$. Then, the series

$$a_0g_0(\epsilon) + a_1g_1(\epsilon) + a_2g_2(\epsilon) + a_3g_3(\epsilon) + \cdots$$

is said to be an asymptotic expansion, as $\epsilon \to 0$, to the function $G(\epsilon)$ if

$$G(\epsilon) - \sum_{k=0}^{n} a_k g_k(\epsilon) = o(g_n(\epsilon))$$
 as $\epsilon \to 0$. (3.1)

In other words, the error incurred in approximating G by the partial sum $\sum_{k=0}^{n} a_k g_k(\epsilon)$ is of higher order than the last gauge function in the partial sum. We write

$$G(\epsilon) \sim a_0 g_0(\epsilon) + a_1 g_1(\epsilon) + a_2 g_2(\epsilon) + a_3 g_3(\epsilon) + \cdots$$
 as $\epsilon \to 0$. (3.2)

The definition (3.1) allows us to relate the coefficients of the asymptotic series to the function $G(\epsilon)$. Thus for n = 0, (3.1) reads

$$G(\epsilon) = a_0 g_0(\epsilon) + o(g_0(\epsilon)),$$

which leads to

$$\lim_{\epsilon \to 0} \frac{G(\epsilon)}{g_0(\epsilon)} = a_0.$$

Similarly, for n = 1, (3.1) reads

$$G(\epsilon) - a_0 g_0(\epsilon) = a_1 g_1(\epsilon) + o(g(\epsilon)),$$

whence

$$\lim_{\epsilon \to 0} \frac{G(\epsilon) - a_0 g_0(\epsilon)}{g_1(\epsilon)} = a_1.$$

Proceeding analogously one is led to the general result

$$\lim_{\epsilon \to 0} \frac{G(\epsilon) - \sum_{k=0}^{n-1} a_k g_k(\epsilon)}{q_n(\epsilon)} = a_n.$$

This procedure to evaluate the coefficients, although a direct consequence of the definition of an asymptotic expansion, is usually too cumbersome to use. Asymptotic series are often computed more directly, with the aid of Taylor and binomial expansions, as will be shown shortly.

Remarks.

• It is important to reiterate the distinction between a convergent series and an asymptotic series. In both cases the series approximates a function. The convergent series approximates the function within the interval of convergence, while the asymptotic series approximates the function in the vicinity of a point. At any point within the interval of convergence, the error of approximation by the convergent series can be made arbitrarily small by retaining a sufficiently large number of terms n in the nth partial sum. In an asymptotic series, the error associated with the nth partial sum is of higher order than, i.e., approaches zero faster than, the nth term as $\epsilon \to 0$. Consider the series

$$\sum_{0}^{\infty} a_m \epsilon^m,$$

and the partial sum

$$S_n = \sum_{0}^{n} a_m \epsilon^m.$$

The series converges to $f(\epsilon)$ for a given ϵ if

$$\lim_{n \to \infty} [f(\epsilon) - S_n(\epsilon)] = 0,$$

while it is asymptotic to $f(\epsilon)$ if

$$f(\epsilon) - S_n(\epsilon) = o(\epsilon^n)$$
, i.e., $\lim_{\epsilon \to 0} \frac{f(\epsilon) - S_n(\epsilon)}{\epsilon^n} = 0$.

Convergent: $f - S_n \to 0$ as $n \to \infty$, ϵ fixed.

Asymptotic: $f - S_n = o(g_n)$ as $\epsilon \to 0$, n fixed.

• A convergent series is always asymptotic. As an illustration consider $f(\epsilon) = e^{-\epsilon}$. A series representation of f is the Taylor expansion about $\epsilon = 0$,

$$e^{-\epsilon} = \sum_{k=0}^{n} \frac{(-1)^k \epsilon^k}{k!} + R_n,$$
 (3.3)

where the remainder R_n is given by

$$R_n = \frac{\epsilon^{n+1}}{(n+1)!} (-1)^{n+1} e^{-\hat{\epsilon}},$$

and $\hat{\epsilon}$ lies between 0 and ϵ . For ϵ positive the remainder has the bound

$$|R_n| \le \frac{\epsilon^{n+1}}{(n+1)!}.\tag{3.4}$$

Suppose that we wish to approximate the exponential by the n-th degree Taylor polynomial obtained by dropping the remainder R_n in (3.3). Suppose further that the approximation is needed for a moderate value of ϵ , say $\epsilon = 1$ or 10. Now (3.4) shows that for any finite ϵ , R_n approaches zero as $n \to \infty$, confirming that the Taylor series for the exponential is convergent, so that the approximation of the exponential by the Taylor polynomial can be made arbitrarily accurate by taking n large enough. The relative error is given by

$$\left| \frac{R_n}{e^{-\epsilon}} \right| = |e^{\epsilon} R_n| \le \frac{\epsilon^{n+1}}{(n+1)!} e^{\epsilon}.$$

Figure 2 shows the graph of the relative error bound as a function of n for $\epsilon = 1$ and $\epsilon = 10$. Note that for the larger value of ϵ the series converges rather slowly as 40 terms are needed to obtain a relative accuracy of 10^{-4} .

If one is interested only in small values of ϵ then an alternate measure of accuracy can be employed. Instead of fixing ϵ and asking how the error depends upon n, we fix n and ask how the error depends upon ϵ . The error bound (3.4) applies again and we now assess the smallness of the remainder by comparing it with the last term (*i.e.*, the highest-degree term) of the Taylor polynomial. The ratio of the error bound to the last term of the Taylor polynomial in (3.3) is

$$|R_n|/|(-1)^n\epsilon^n/n! \le \frac{\epsilon^{n+1}}{(n+1)!} / \frac{\epsilon^n}{n!} = \frac{\epsilon}{n+1}.$$

This ratio approaches zero as $\epsilon \to 0$, for any fixed n. Thus the remainder is of higher order than the last term retained, and this property is the hallmark of an asymptotic series.

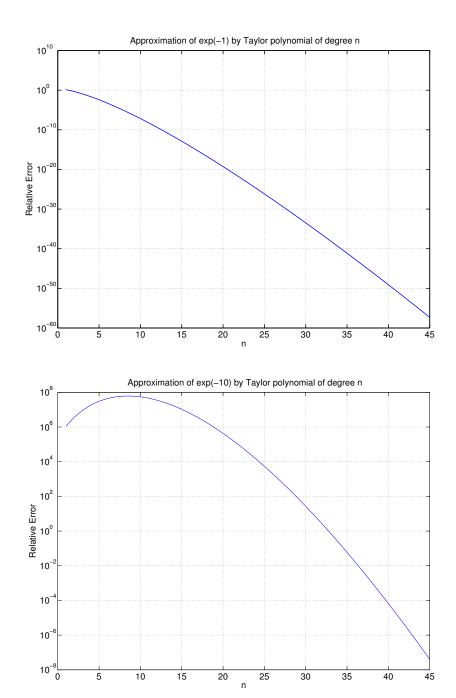


Figure 2: Approximation of $\exp(-1)$ and $\exp(-10)$ by Taylor polynomials.

• Computation of asymptotic expansions of functions is the main object of the course. Determination of the asymptotic sequence is a part of the exercise. The sequence is not entirely arbitrary, because it must be sufficiently complete to capture the behavior of the function. Within this restriction, the number of alternatives is unlimited. For example,

$$\sin 2\epsilon \sim 2\epsilon - \frac{4}{3}\epsilon^3 + \frac{4}{15}\epsilon^5 \cdots$$

$$\sim 2\tan \epsilon - 2\tan^3 \epsilon - 2\tan^5 \epsilon + \cdots$$

$$\sim 2\ln(1+\epsilon) + \ln(1+\epsilon^2) - 2\ln(1+\epsilon^3) + \cdots$$

It is clear, however, that the first choice above is the simplest, and in some sense, natural.

Exercise: Show that the above expansions are equivalent, i.e., the last two reduce to the first.

• Several functions may have the same expansion with respect to a given asymptotic sequence. For example, the functions

$$f_1(\epsilon) = \frac{1}{1 - \epsilon}, \quad f_2(\epsilon) = \frac{1}{1 - \epsilon} + e^{-1/\epsilon} \quad \text{and} \quad \frac{1}{1 - \epsilon} + \frac{2e^{-1/\epsilon^2}}{\epsilon}$$

all have the asymptotic expansion $\sum_{n=0}^{\infty} \epsilon^n$. The difference of any two of these functions has zero asymptotic expansion with respect to the given gauge sequence and is said to be *subdominant* or *transcendentally small* with respect to the sequence. Thus $e^{-1/\epsilon}$ is transcendentally small with respect to the sequence $\{\epsilon^n\}$ and ϵ is transcendentally small with respect to the sequence $\{1/[\ln(1/\epsilon)]^n\}$.

Since asymptotic expansions with respect to a given gauge sequence are unique, the manner of their
determination is not important. Rather than using the limit process to compute the coefficients, Taylor
and binomial expansions are frequently useful; they also yield the natural gauge sequence automatically.
Here are some examples.

Examples 3.1.

(i) A 2-term asymptotic expansion of $\sin(\sin x)$ as $x \to 0$. First we expand the outer sine in a Taylor expansion, to get

$$f(x) = \sin x - \frac{1}{3!} \sin^3 x + O(\sin^5 x).$$

Then we use the expansion of $\sin x$,

$$\sin x = x - \frac{x^3}{3!} + O(x^5).$$

Upon substituting the second expansion into the first, and retaining terms up to $O(x^3)$, we obtain

$$\sin(\sin x) = [x - \frac{x^3}{3!} + O(x^5)] - \frac{1}{3!}[x - \frac{x^3}{3!} + O(x^5)]^3 + O(x^5)$$

$$= x - \frac{x^3}{3!} - \frac{x^3}{3!} + O(x^5)$$

$$= x - 2\frac{x^3}{3!} + O(x^5)$$

(ii) A 3-term asymptotic expansion of $\ln(1+\frac{x}{1+x^2})$ as $x\to 0$. First we use the Taylor expansion of $\ln(1+u)$ to write

$$\ln(1+u) = u - \frac{u^2}{2} + \frac{u^3}{3} + O(u^4)$$
 as $u \to 0$.

Then we expand $u = x/(1+x^2)$ in a binomial expansion,

$$u = \frac{x}{1+x^2} = x(1+x^2)^{-1} = x(1-x^2+O(x^4)) = x-x^3+O(x^5).$$

Upon substituting the second expansion into the first, and retaining terms up to $O(x^3)$, we obtain

$$\ln\left(1 + \frac{x}{1+x^2}\right) = x - x^3 + O(x^5) - \frac{1}{2}[x - x^3 + O(x^5)]^2 + \frac{1}{3}[x - x^3 + O(x^5)]^3 + O(x^4)$$

$$= x - x^3 - \frac{1}{2}[x^2 + O(x^4)] + \frac{1}{3}[x^3 + O(x^5)]$$

$$= x - \frac{1}{2}x^2 - \frac{2}{3}x^3 + O(x^4).$$

• Thus far we have considered the behavior of functions $f(\epsilon)$ in the vicinity of $\epsilon = 0$. Now let us consider the expansion, as $\epsilon \to 0$, of $f(x;\epsilon)$, where $x \in \mathcal{D}$ is the independent variable and ϵ is a parameter. Then the definition **Def 4** of the asymptotic expansion given earlier in (3.1), is modified as follows.

Def 4a. With respect to the gauge sequence $\{g_n(\epsilon)\}$ we write the asymptotic expansion of $f(x,\epsilon)$ as

$$f(x;\epsilon) \sim \sum_{k=0}^{N} a_k(x)g_k(\epsilon)$$
 (3.5)

for x = O(1) and every integer $N \ge 0$. The requirement that x = O(1) is equivalently that the limit $\epsilon \to 0$ is taken with x fixed.

If the above expansion holds for all $x \in \mathcal{D}$ then the expansion is said to be uniform or uniformly valid. Otherwise the expansion is non-uniform. A symptom of non-uniformity is a breakdown or disordering of the expansion for certain x. These ideas are elaborated in the following example.

Example 3.2. Consider

$$f(x;\epsilon) = \frac{x\sqrt{1+\epsilon x}}{x+\epsilon}, \quad x \ge 0.$$
 (3.6)

For x = O(1) we write

$$f(x;\epsilon) = \left(1 + \frac{\epsilon}{x}\right)^{-1} (1 + \epsilon x)^{1/2},$$

and use binomial expansions to get

$$f(x;\epsilon) \sim 1 + \epsilon \left(\frac{x}{2} - \frac{1}{x}\right)$$
 as $\epsilon \to 0$. (3.7)

For x finite and bounded away from zero this is a valid expansion, as the second term is o(1) with respect to the first.

For x near zero the expansion breaks down. Breakdown occurs when $x = O(\epsilon)$ because then ϵ/x in the second term is of the same order as the leading term of the expansion. A similar breakdown occurs for large x because when $x = O(1/\epsilon)$, ϵx in the second term becomes of order unity.

In the regions of non-uniformity we develop new expansions after rescaling the independent variable. For x near zero we let $x = \epsilon X$, where X is treated as O(1). Then we write

$$f(\epsilon X; \epsilon) = \frac{X\sqrt{1+\epsilon^2 X}}{1+X}$$

$$\sim \frac{X}{1+X} \left(1 + \frac{1}{2}\epsilon^2 X\right). \tag{3.8}$$

We note that this expansion recovers the correct value f = 0 at X = 0.

For x large we let $x = \xi/\epsilon$ where $\xi = O(1)$. Then we have

$$f(\xi/\epsilon;\epsilon) = \frac{\xi\sqrt{1+\xi}}{\xi+\epsilon^2}$$

$$\sim \sqrt{1+\xi}\left(1-\frac{\epsilon^2}{\xi}\right). \tag{3.9}$$

Thus three different series are needed to correctly expand $f(x;\epsilon)$ in the domain $x \ge 0$ for $\epsilon \to 0$.

• In order to construct asymptotic expansions directly from the governing problems (such as ODEs or PDEs), it will be necessary to carry out certain operations on the expansions, such as addition, multiplication, integration with respect to x, and differentiation with respect to x, etc. Some of these operations are not always permissible, especially differentiation (allowed only when it is known that the derivative has an asymptotic expansion as well). For example, the following is a perfectly reasonable asymptotic expression,

$$f(x,\epsilon) \sim \sin(x/\epsilon) + \epsilon \sin(x/\epsilon^4) + \epsilon^2 \sin(e^{1/\epsilon}x) + \cdots$$

since each term is of higher order than its predecessor, but its x-derivative,

$$f'(x,\epsilon) \sim \frac{1}{\epsilon} \cos(x/\epsilon) + \frac{1}{\epsilon^3} \cos(x/\epsilon^4) + e^{1/\epsilon} \cos(e^{1/\epsilon}x) + \cdots,$$

is not. We shall proceed under the assumption that the above operations, including differentiation, are valid, recognizing that violation of this assumption will lead to non-uniformities.

3.2.1 Divergent series

We return to Example 2.5 for which we had constructed a divergent series in powers of ϵ . In order to examine whether this series has any utility, consider again the integral

$$f(\epsilon) = \int_0^\infty \frac{e^{-t}}{1 + \epsilon t} dt.$$

Upon using

$$\frac{1}{1+x} = \sum_{n=0}^{N} (-1)^n x^n + (-1)^{N+1} \frac{x^{N+1}}{1+x},$$

we see that

$$f(\epsilon) = \int_0^\infty \frac{e^{-t}}{1 + \epsilon t} dt = \int_0^\infty e^{-t} \left[\sum_{n=0}^N (-1)^n \epsilon^n t^n + (-1)^{N+1} \epsilon^{N+1} \frac{t^{N+1}}{1 + \epsilon t} \right] dt$$
$$= \sum_{n=0}^N (-1)^n \epsilon^n n! + R_N,$$

where

$$R_N = (-1)^{N+1} \epsilon^{N+1} \int_0^\infty \frac{t^{N+1} e^{-t}}{1 + \epsilon t} dt.$$

We can bound the integral in the remainder as

$$I_N = \int_0^\infty \frac{t^{N+1}e^{-t}}{1+\epsilon t} dt < \int_0^\infty t^{N+1}e^{-t} dt = (N+1)!,$$

so that the remainder can be written as

$$R_N = (-1)^{N+1} \epsilon^{N+1} I_N$$
, where $I_N \le (N+1)!$. (3.10)

Now consider the formal series

$$\sum_{0}^{\infty} (-1)^n \epsilon^n n!.$$

Since the nth term is unbounded as $n \to \infty$, the series diverges for any $\epsilon > 0$. However, for any fixed N,

$$f(\epsilon) - \sum_{n=0}^{N} (-1)^n \epsilon^n n! = R_N = O(\epsilon^{N+1}) = o(\epsilon^N),$$

so that the series is asymptotic. Thus we are justified in writing

$$f(\epsilon) \sim \sum_{0}^{\infty} (-1)^n \epsilon^n n!.$$

Now let us explore the behavior of the series for small but finite values of ϵ . (Recall that the definition of the asymptotic series is silent on how small the error is when ϵ is small but finite.)

Criterion 1: If all we know about the remainder R_N is that it alternates in sign with N (as it does here), then the exact value $f(\epsilon)$ is always bracketed by two successive partial sums,

$$S_N = \sum_{0}^{N} (-1)^n \epsilon^n n!$$
 and $S_{N+1} = \sum_{0}^{N+1} (-1)^n \epsilon^n n!$.

Therefore either partial sum, S_N or S_{N+1} , produces the optimal approximation when $T_{N+1} = S_{N+1} - S_N$ has the least value, where T_{N+1} is the (N+1) th term of the series. (The average $(S_N + S_{N+1})/2$ is an even better approximation.)

Criterion 2: If we have a bound on the remainder (as we do here), then the optimal N corresponds to minimizing R_N . Equation (3.10) shows that

$$|R_N| \le \epsilon^{N+1} (N+1)!.$$

For given ϵ the above bound is minimized by selecting N+1 to be the closest integer to $1/\epsilon$.

For $\epsilon = 0.2$ the following table lists the first 12 terms of the series, as well as the corresponding partial sums. Figure 3 displays the same information graphically.

ep = 2.0e-01 exact = 8.5211e-01

N	T(N)	S(N)
1	1.0000e+00	1.0000e+00
2	-2.0000e-01	8.0000e-01
3	8.0000e-02	8.8000e-01
4	-4.8000e-02	8.3200e-01
5	3.8400e-02	8.7040e-01
6	-3.8400e-02	8.3200e-01
7	4.6080e-02	8.7808e-01
8	-6.4512e-02	8.1357e-01
9	1.0322e-01	9.1679e-01
10	-1.8579e-01	7.3099e-01
11	3.7159e-01	1.1026e+00
12	-8.1750e-01	2.8509e-01

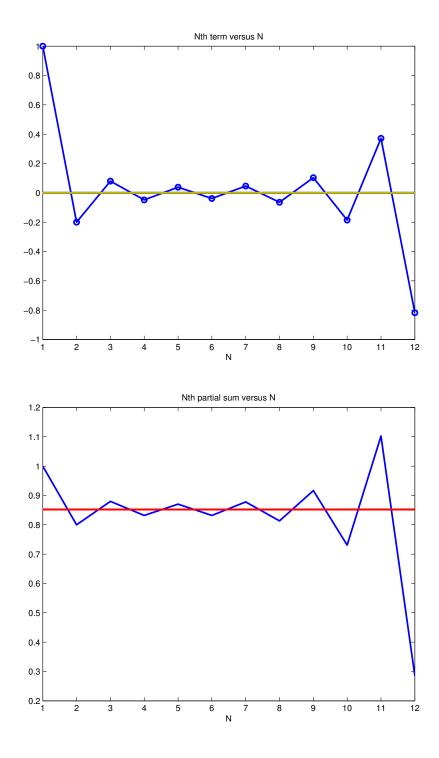


Figure 3: Plots of the N th term and the N th partial sum of the series expansion of the integral $f(\epsilon)$ in Example 2.5 for $\epsilon=0.2$. In the lower plot the exact value of f(0.2)=0.85211 is represented by the red line.

We note that the terms in the series first decrease and then increase. The smallest terms are the 5th and the 6th, identical in size but opposite in sign. Therefore, according to Criterion 1 above the partial sums $S_4 = 0.83200$ and $S_5 = 0.87040$ yield results that are roughly of similar accuracy, given that the exact answer is 0.85211.

According to Criterion 2 above, S_4 is optimal as the closest integer to $1/\epsilon$ is N+1=5. Thus both criteria yield similar results.

The graph clearly shows that keeping additional terms will reduce the accuracy, in contrast to what happens for convergent series; this is the essential difference between asymptotic and convergent series.

The corresponding results for $\epsilon = 0.15$ are shown in the table below and in Figure 4.

```
ep = 1.5e-01
exact = 8.8193e-01
                                S(N)
   N
             T(N)
       1.0000e+00
                      1.0000e+00
   1
   2
      -1.5000e-01
                      8.5000e-01
   3
       4.5000e-02
                     8.9500e-01
   4
      -2.0250e-02
                     8.7475e-01
       1.2150e-02
                     8.8690e-01
   5
   6
      -9.1125e-03
                     8.7779e-01
       8.2012e-03
   7
                     8.8599e-01
      -8.6113e-03
                     8.7738e-01
       1.0334e-02
                     8.8771e-01
   9
       -1.3950e-02
                      8.7376e-01
   10
        2.0925e-02
                      8.9469e-01
   11
       -3.4527e-02
                      8.6016e-01
   12
   13
        6.2149e-02
                      9.2231e-01
   14
       -1.2119e-01
                      8.0112e-01
        2.5450e-01
                       1.0556e+00
   15
   16
       -5.7262e-01
                      4.8299e-01
        1.3743e+00
   17
                      1.8573e+00
   18
       -3.5045e+00
                     -1.6472e+00
```

Now the 7th term is the smallest and the sums $S_6 = 0.87779$ or $S_7 = 0.88599$ give the optimal approximations to the exact result 0.88193.

Remarks.

- Truncating the series at the smallest term is often used as a rule of thumb for optimal truncation. However, it is justified only when the remainder is known to be alternating in sign.
- In most practical problems one is able to generate only a handful of terms in the series (often only one or two), as computation of higher-order terms may involve a prohibitive amount of effort. Additionally, it is often impossible to estimate the remainder. Experience, experiment, and intuition play important roles in deciding when to trust the results. Having said that, the scientific literature is replete with problems where asymptotic analysis has led to a remarkable increase in the understanding of a variety of phenomena in fields as diverse as areodynamics and biology. In such situations, the value of the asymptotic result lies not as much in the numerical accuracy, but in the insight it provides into the structure of the solution.

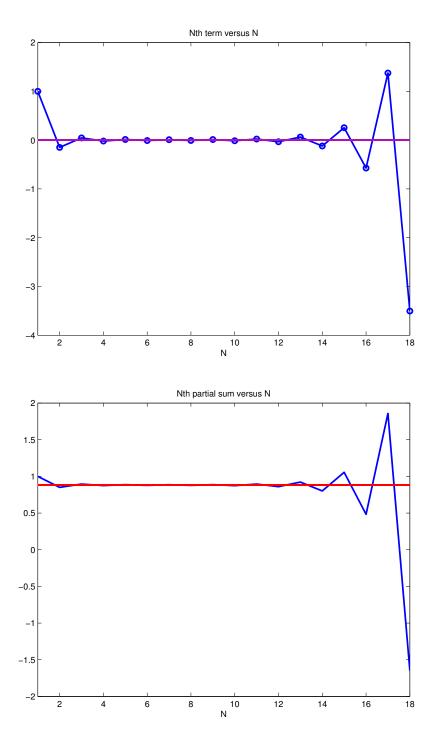


Figure 4: Plots of the N th term and the N th partial sum of the series expansion of the integral $f(\epsilon)$ in Example 2.5 for $\epsilon = 0.15$. In the lower plot the exact value of f(0.15) = 0.88193 is represented by the red line.