3.1.4 Asymptotic Analysis

With insight from previous examples we define some notions regarding convergence and uniformity. We observed that substitution of a perturbation series into a differential equation does not always lead to a valid approximate solution. Ideally we would like to say that a few terms in a truncated perturbation series provides, for a given \( \varepsilon \), an approximate solution for the entire range of the independent variable \( t \). Unfortunately, as we have seen, this is not always the case. Failure of this regular perturbation method is the rule rather than the exception.

To aid the analysis of approximate solutions we introduce some basic notation and terminology that permits the comparison of two functions as their common argument approaches some fixed value. These comparisons are called order relations.

Let \( f(\varepsilon) \) and \( g(\varepsilon) \) be defined in some neighborhood (or punctured neighborhood) of \( \varepsilon = 0 \). We write

\[
f(\varepsilon) = o(g(\varepsilon)) \quad \text{as} \quad \varepsilon \to 0,
\]

if

\[
\lim_{\varepsilon \to 0} \frac{|f(\varepsilon)|}{|g(\varepsilon)|} = 0,
\]

and we write

\[
f(\varepsilon) = O(g(\varepsilon)) \quad \text{as} \quad \varepsilon \to 0,
\]

if there exists a positive constant \( M \) such that

\[
|f(\varepsilon)| \leq M|g(\varepsilon)|
\]

for all \( \varepsilon \) in some neighborhood (punctured neighborhood) of zero. The comparison function \( g \) is called a gauge function.

In this definition, \( \varepsilon \to 0 \) may be replaced by a one-sided limit or by \( \varepsilon \to \varepsilon_0 \), where \( \varepsilon_0 \) is any finite or infinite number, with the domain of \( f \) and \( g \) defined appropriately. If (1.23) holds, we say \( f \) is little oh of \( g \) as \( \varepsilon \to 0 \), and if (1.24) holds we say \( f \) is big oh of \( g \) as \( \varepsilon \to 0 \). Common gauge functions are \( g(\varepsilon) = \varepsilon^n \) for some exponent \( n \), and \( g(\varepsilon) = e^{\alpha(\ln \varepsilon)^m} \) for exponents \( \alpha \) and \( m \). The statement \( f(\varepsilon) = O(1) \) means \( f \) is bounded in a neighborhood of \( \varepsilon = 0 \), and \( f(\varepsilon) = o(1) \) means \( f(\varepsilon) \to 0 \) as \( \varepsilon \to 0 \). If \( f = o(g) \), then \( f \) goes to zero faster than \( g \) goes to zero as \( \varepsilon \to 0 \), and we write \( f(\varepsilon) \ll g(\varepsilon) \). The following examples illustrate a few methods to prove order relations.

Example 3.2

Verify \( \varepsilon^2 \ln \varepsilon = o(\varepsilon) \) as \( \varepsilon \to 0^+ \). By L’Hôpital’s rule,

\[
\lim_{\varepsilon \to 0^+} \frac{\varepsilon^2 \ln \varepsilon}{\varepsilon} = \lim_{\varepsilon \to 0^+} \frac{\ln \varepsilon}{(1/\varepsilon)} = \lim_{\varepsilon \to 0^+} \frac{1/\varepsilon^2}{-1/\varepsilon^2} = 0.
\]

Example 3.3

Verify \( \sin \varepsilon = O(\varepsilon) \) as \( \varepsilon \to 0^+ \). By the mean value theorem there is a number \( c \) between 0 and \( \varepsilon \) such that

\[
\sin \varepsilon = \sin 0 = 0.
\]

Hence \( |\sin \varepsilon| = |\varepsilon \cos \varepsilon| \leq |\varepsilon| \), because \( |\cos \varepsilon| \leq 1 \). An alternate argument is to note that \( \sin \varepsilon/\varepsilon \to 1 \) as \( \varepsilon \to 0^+ \). Because the limit exists, the function \( \sin \varepsilon/\varepsilon \) must be bounded for \( 0 < \varepsilon < \varepsilon_0 \), for some \( \varepsilon_0 \). Therefore \( |\sin \varepsilon|/\varepsilon \leq M \), for some constant \( M \) and \( \sin \varepsilon = O(\varepsilon) \).

The order definitions may be extended to functions of \( \varepsilon \) and another variable \( t \) lying in an interval \( I \). First we review the notion of uniform convergence. Let \( h(t, \varepsilon) \) be a function defined for \( \varepsilon \) in a neighborhood of \( \varepsilon = 0 \), possibly not including the value \( \varepsilon = 0 \) itself, and for \( t \) in some interval \( I \), either finite or infinite. We say

\[
\lim_{\varepsilon \to 0} h(t, \varepsilon) = 0 \quad \text{uniformly on} \quad I,
\]

if the convergence to zero is at the same rate for each \( t \in I \); that is, if, for any positive number \( \eta \) there can be chosen a positive number \( \varepsilon_0 \), independent of \( t \), such that \( |h(t, \varepsilon)| < \eta \) for all \( t \in I \), whenever \( |\varepsilon| < \varepsilon_0 \). In other words, if \( h(t, \varepsilon) \) can be made arbitrarily small over the entire interval \( I \) by choosing \( \varepsilon \) small enough, then the convergence is uniform. If merely \( \lim_{\varepsilon \to 0} h(t_0, \varepsilon) = 0 \) for each fixed \( t_0 \in I \), then we say that the convergence is pointwise on \( I \).

To prove \( \lim_{\varepsilon \to 0} h(t, \varepsilon) = 0 \) uniformly on \( I \) it is sufficient to find a function \( H(t) \) such that \( |h(t, \varepsilon)| \leq H(t) \) holds for all \( t \in I \), with \( H(t) \to 0 \) as \( \varepsilon \to 0 \). To prove that convergence is not uniform on \( I \), it is sufficient to produce a \( t \in I \) such that \( |h(t, \varepsilon)| \geq \eta \) for some positive \( \eta \), regardless of how small \( \varepsilon \) is chosen.

Let \( f(t, \varepsilon) \) and \( g(t, \varepsilon) \) be defined for all \( t \in I \) and all \( \varepsilon \) in a (punctured) neighborhood of \( \varepsilon = 0 \). We write

\[
f(t, \varepsilon) = o(g(t, \varepsilon)) \quad \text{as} \quad \varepsilon \to 0,
\]

if

\[
\lim_{\varepsilon \to 0} \left| \frac{f(t, \varepsilon)}{g(t, \varepsilon)} \right| = 0.
\]
pointwise on \( I \). If the limit is uniform on \( I \), we write \( f(t, \varepsilon) = o(g(t, \varepsilon)) \) as \( \varepsilon \to 0 \), uniformly on \( I \). If there exists a positive function \( M(t) \) on \( I \) such that

\[
|f(t, \varepsilon)| \leq M(t)g(t, \varepsilon)
\]

for all \( t \in I \) and \( \varepsilon \) in some neighborhood of zero, then we write

\[
f(t, \varepsilon) = O(g(t, \varepsilon)) \quad \text{as} \quad \varepsilon \to 0, \quad t \in I.
\]

If \( M(t) \) is a bounded function on \( I \), we write

\[
f(t, \varepsilon) = O(g(t, \varepsilon)) \quad \text{as} \quad \varepsilon \to 0, \quad \text{uniformly on} \quad I.
\]

The big oh and little oh notations permit us to make quantitative statements about the error in a given approximation. We can make the following definition. A function \( y(t, \varepsilon) \) is a uniformly valid asymptotic approximation to a function \( y(t, \varepsilon) \) on an interval \( I \) as \( \varepsilon \to 0 \) if the error \( E(t, \varepsilon) \equiv y(t, \varepsilon) - y_0(t, \varepsilon) \) converges to zero as \( \varepsilon \to 0 \) uniformly for \( t \in I \). We often express the fact that \( E(t, \varepsilon) \) is little oh or big oh of \( \varepsilon^n \) (for some \( n \)) as \( \varepsilon \to 0 \) to make an explicit statement regarding the rate at which the error goes to zero, and whether or not the convergence is uniform.

**Example 3.4**

Let

\[
y(t, \varepsilon) = e^{-\varepsilon t}, \quad t > 0, \quad \varepsilon \ll 1.
\]

The first three terms of the Taylor expansion in powers of \( \varepsilon \) provide an approximation

\[
y_0(t, \varepsilon) = 1 - t\varepsilon + \frac{1}{2} t^2 \varepsilon^2.
\]

The error is

\[
E(t, \varepsilon) = e^{-\varepsilon t} - 1 + t\varepsilon - \frac{1}{2} t^2 \varepsilon^2 = -\frac{1}{3!} t^3 \varepsilon^3 + \cdots.
\]

For a fixed \( t \) the error can be made as small as desired by choosing \( \varepsilon \) sufficiently small. Thus \( E(t, \varepsilon) = O(\varepsilon^2) \) as \( \varepsilon \to 0 \). If \( \varepsilon \) is fixed, however, regardless of how small, \( t \) may be chosen large enough so that the approximation is totally invalid. Thus, the approximation is not uniform on \( I = [0, \infty) \). Clearly, by choosing \( t = 1/\varepsilon \) we have \( E(1/\varepsilon, \varepsilon) = e^{-t} - \frac{1}{2} t^2 \), which is not small. We may not write \( E(t, \varepsilon) = O(\varepsilon^2) \) as \( \varepsilon \to 0 \) uniformly on \( [0, \infty) \). 

**Example 3.5**

(Transcendally small) Some functions cannot be gauged against polynomials. Consider \( f(\varepsilon) = e^{-1/\varepsilon} \) for small, positive \( \varepsilon \). From calculus, this function does not have a Taylor expansion about \( \varepsilon \); further,

\[
\lim_{\varepsilon \to 0^+} \frac{e^{-1/\varepsilon}}{\varepsilon^n} = \lim_{n \to \infty} \varepsilon^n e^{-n} = 0,
\]

because exponentials decay faster than polynomials grow. This means

\[
e^{-1/\varepsilon} = o(\varepsilon^n) \quad \text{as} \quad \varepsilon \to 0^+
\]

for any \( n = 0, 1, 2, \ldots \). Thus, \( e^{-1/\varepsilon} \) goes to zero faster than any positive power of \( \varepsilon \). In this case we say the function \( f(\varepsilon) \) is transcendently small. 

The difficulty of these definitions with regard to differential equations is that the exact solution to the equation is seldom known, and therefore a direct error estimate cannot be made. Consequently, we require some notion of how well an approximate solution satisfies the differential equation and the auxiliary conditions. For definiteness consider the differential equation in (1.1). We say that an approximate solution \( y_\varepsilon(t, \varepsilon) \) satisfies the differential equation (1.1) uniformly for \( t \in I \) as \( \varepsilon \to 0 \) if

\[
r(t, \varepsilon) = F(t, y_\varepsilon(t, \varepsilon), y_\varepsilon(t, \varepsilon), y_\varepsilon(t, \varepsilon), \varepsilon) \to 0
\]

uniformly on \( I \) as \( \varepsilon \to 0 \). The quantity \( r(t, \varepsilon) \) is the residual error, which measures how well the approximate solution \( y_\varepsilon(t, \varepsilon) \) satisfies the equation.

**Example 3.6**

Consider the initial value problem

\[
\dot{y} + y^2 + \varepsilon y = 0, \quad t > 0, \quad 0 < \varepsilon \ll 1,
\]

\[
y(0) = 0, \quad \dot{y}(0) = 1.
\]

Substituting the perturbation series \( y = y_0 + \varepsilon y_1 + \cdots \) gives the initial value problem

\[
y_0 + \varepsilon y_1^0 = 0, \quad t > 0,
\]

\[
y_0(0) = 0, \quad \dot{y}_0(0) = 1,
\]

for the leading-order term \( y_0 \). It is easily found that \( y_0(t) = \ln(t+1) \), and hence

\[
r(t, \varepsilon) = e - \ln(t+1)
\]
Thus \( r(t, \varepsilon) = O(\varepsilon) \) as \( \varepsilon \to 0 \), but not uniformly on \([0, \infty)\). On any finite interval \([0, T]\), however, we have \( |c \ln(t+1)| \leq \varepsilon \ln(T+1) \), and so \( r(t, \varepsilon) = O(\varepsilon) \) as \( \varepsilon \to 0 \) uniformly on \([0, T]\).

Specifically, the regular perturbation method produces an asymptotic expansion

\[
\varrho_0(t) + \varrho_1(t)\varepsilon + \varrho_2(t)\varepsilon^2 + \cdots,
\]

for which an approximate solution can be obtained by taking the first few terms. Such an expansion in the integral powers of \( \varepsilon \), that is, \( 1, \varepsilon, \varepsilon^2, \ldots \), is called an asymptotic power series. In some problems the expansion may take the form

\[
\varrho_0(t) + \varrho_1(t)\sqrt{\varepsilon} + \varrho_2(t)\varepsilon + \varrho_3(t)\varepsilon^{3/2} + \cdots,
\]

in terms of the sequence \( 1, \sqrt{\varepsilon}, \varepsilon, \sqrt{\varepsilon^2}, \ldots \). In yet other problems the required expansion must have the form

\[
\varrho_0(t) + \varrho_1(t)\ln \varepsilon + \varrho_2(t)\varepsilon + \varrho_3(t)\varepsilon^2 + \varrho_4(t)\varepsilon^3 + \cdots.
\]

The type of expansion depends on the problem. In general, we say a sequence of gauge functions \( \{g_n(t, \varepsilon)\} \) is an asymptotic sequence as \( \varepsilon \to 0 \), if

\[
g_{n+1}(t, \varepsilon) = o(g_n(t, \varepsilon)) \quad \text{as} \quad \varepsilon \to 0,
\]

for \( n = 0, 1, 2, \ldots \). That is, each term in the sequence tends to zero faster than its predecessor, as \( \varepsilon \to 0 \). Given a function \( g(t, \varepsilon) \) and an asymptotic sequence \( \{g_n(t, \varepsilon)\} \) as \( \varepsilon \to 0 \), the formal series

\[
\sum_{n=0}^{\infty} g_n(t, \varepsilon)
\]

is said to be an asymptotic expansion of \( g(t, \varepsilon) \) as \( \varepsilon \to 0 \), if

\[
g(t, \varepsilon) = \sum_{n=0}^{\infty} a_n g_n(t, \varepsilon) = o(g(t, \varepsilon)), \quad \text{as} \quad \varepsilon \to 0,
\]

for every \( N \). In other words, for any partial sum the remainder is little oh of the last term. If the limits just cited are uniform for \( t \in I \), then we speak of a uniform asymptotic sequence and uniform asymptotic expansion. In most cases the sequence \( \{g_n(t, \varepsilon)\} \) is the form of a product, \( g_n(t, \varepsilon) = \varrho_n(t)\delta_n(\varepsilon) \), as in the previous examples. The notation

\[
y = \sum_{n=0}^{\infty} a_n g_n(t, \varepsilon)
\]

often denotes an asymptotic expansion.

The formal series (1.26) need not converge to be valuable. The value of such expansions, although perhaps divergent, is that often only a few terms are required to obtain an accurate approximation, whereas a convergent Taylor series may yield an accurate approximation only if many terms are calculated. A rather obvious question arises. If a given approximate solution \( \varrho_n(t, \varepsilon) \) satisfies the differential equation uniformly on \( t \in I \), is it in fact a uniformly valid approximation to the exact solution \( y(t, \varepsilon) \)? A complete discussion of this question is beyond our scope, but a few remarks are appropriate to caution the reader regarding the nature of this problem. Probably more familiar is the situation in linear algebra where we consider a linear system of equations \( Ax = b \). Let \( x_0 \) be an approximate solution. A measure of how well it satisfies the system is the magnitude of the residual vector \( r \) defined by \( r = Ax - b \). If \( r = 0 \), then \( x_0 \) must be the exact solution \( x \). But if \( |r| \) is small, it does not necessarily follow that the magnitude \( |x - x_0| \) is small, where \( e = x - x_0 \) is the error. In ill-conditioned systems, where \( \det A = 0 \) it is not true that even a small residual may not imply a small error. A similar state of affairs exists for differential equations. Therefore one must proceed cautiously in interpreting the validity of a perturbation solution. Numerical calculations or the computation of additional correction terms may aid in the interpretation. Often a favorable comparison with experiment leads one to conclude that an approximation is valid.

**EXERCISES**

1. In a spring-mass problem assume that the restoring force is \(-ky\) and that there is a force \( k \) which is \( \alpha \) times the normal force \( k \). With initial conditions \( y(0) = A, y'(0) = 0 \), determine the correct form and magnitude scales for small damping and show that the problem can be written in dimensionless form as

\[
y'' + c(y')^2 + y = 0,
\]

\[
y(0) = 1, \quad y'(0) = 0,
\]

with \( \varepsilon = \alpha A/m \) is a dimensionless parameter and primes denote the derivative with respect to the scaled time \( \tilde{t} \).

2. Consider the initial value problem

\[
u'' = u', \quad t > 0, \quad u(0) = 1, \quad u'(0) = -1.
\]

Find a two-term perturbation approximation for \( 0 < \varepsilon \ll 1 \) and compare it graphically to a six-term Taylor series approximation (centered at \( t = 0 \)).