Chapter 7

Numerical Solution of Ordinary Differential Equations

Introduction

We assume in this chapter that the reader is familiar with the elementary theory of ordinary differential equations. Hence, although we state a few of the theorems governing existence and uniqueness of solutions we neither prove them nor claim that the list is exhaustive. Many excellent textbooks exist on this subject, some we list at the end of the chapter. The reader should note as she proceeds through this chapter that the flavor when numerically solving initial values problems versus boundary value problems is very different.

When solving initial value problems, we may, without loss of generality, restrict ourselves to first order nonlinear systems of the form

$$y' = f(x, y) \ , \ \ y(x_0) = y_0 \ , \tag{7.0.1}$$

where y, y_0, x, and x_0 are vectors if appropriate. This is possible since nth order equations can always be transformed into a system of such equations. Many methods are known for solving (7.0.1). We discuss and illustrate by example only the simpler among these, such as Euler's method, Heun's method, Runge-Kutta methods, and predictor-corrector methods. These are still typically applied when numerically solving partial differential equations that evolve in time. We also carry out an error analysis, discussing the computational difficulties that may arise due to instability and roundoff error.

Finally, we consider the numerical solution of stiff equations. These commonly occur for systems that model a physical process with widely varying time constants. For example, a chemical system may be dominated for a short time by one very fast reaction followed by a second reaction that is extremely slow. Care must be taken when selecting a method for the system otherwise the fast reaction dictates an extremely small step size to maintain numerical stability. Quite likely, the small step size in incompatible with monitoring the behavior of the slower second phase.

For the case of the numerical solution of boundary value problems, we restrict ourselves to the second order equation

$$y'' + f(x, y, y') = 0 \ , \quad y(0) = y(1) = 0 \, .$$

Here we again describe a selected group of the well-known simple methods that lead to quickly obtainable approximations. Included among these are Numerov's method, the shooting method combined with initial value methods, and methods that lead to the solution of a system of algebraic equations. This latter class includes finite difference methods and techniques that use "elements" such as cubic splines. We also discuss the numerical solution of eigenvalue problems and the Bubnov-Galerkin method.

The topic of this chapter presents unusual difficulties to students. It makes extensive use of numerical integration and, in addition, calls on concepts from two fields: differential equations and difference equations. There is a vast literature of differential equations and an almost equally vast literature of numerical solutions of these equations. If this chapter serves to orient the reader in his subsequent study of the literature, it has served its purpose.

7.1 Illustrative Examples of Ordinary Differential Equations

In this section we discuss examples in order to elucidate essential points of the theory of differential equations, and to illustrate the kind of preliminary analysis that often precedes numerical solution. Throughout this chapter a prime denotes differentiation with respect to x.

The simplest first order linear differential equation,

$$y' = f(x) \, , \tag{7.1.1}$$

immediately admits a formal integral solution

$$y(x) = y_0 + \int_{x_0}^{x} f(t)dt \tag{7.1.2}$$

where $y(x_0) = y_0$. Once the point $y(x_0) = y_0$ is prescribed, the solution $y(x)$ of (7.1.1) is unique. Similarly, n auxiliary conditions must be prescribed for an equation of order n (e.g. the equation $y^{(n)}(x) = f(x)$) to have a unique solution. Usually the integral in (7.1.2) cannot be explicitly expressed. We must therefore resort to a method of numerical integration as described in Chapter 5 in order to approximate $y(x)$.

The general form of a first order linear differential equation is

$$y' = a(x)y + f(x)\,. \tag{7.1.3}$$

The circumstances gauranteeing existence and uniqueness are specified in the following theorem.

Theorem 7.1.1 *If a and f are continuous in the interval $\alpha < x\beta$, and then there exists a unique differentiable solution $y = y(x)$ of (7.1.3) defined on (α, β) and satisfying $y(x_0) = y_0$.*

The solution of this equation is

$$y(x) = \exp\{A(x)\}\left[y_0 + \int_{x_0}^{x} \exp\{-A(t)\}f(t)dt\right] \tag{7.1.4}$$

where

$$A(x) = \int_{x_0}^{x} a(t)dt\ ,\ \ y_0 = y(x_0)\,. \tag{7.1.5}$$

From a standpoint of analysis, the solution is complete. However, it may be inconvenient for purposes of numerical evaluation. For example, it may not be possible to explicitly express either $A(x)$, or the integral

$$\int_{x_0}^{x} \exp\{-A(t)\}f(t)dt\,. \tag{7.1.6}$$

In addition, the numerical evaluation of these integrals may be difficult if either $a(t)$ or $f(t)$ has a singularity on the interval $[x_0, x]$ (see Problem 7.1.6 and Chapter 5) since the standard quadrature rules of Chapter 5 require more than just continuity of the integrand. If neither $A(t)$ nor the integral (7.1.6) can be expressed explicitly, it may be much simpler to use a direct method from this chapter as an approximation technique.

Example 7.1.2 Consider the differential equation

$$xy' + 2y = 4x^2 \ , \ \ y(1) = 1 + \varepsilon \, . \tag{7.1.7}$$

Rewriting the equation to match the form of (7.1.3) we see that $a(x) = -2/x$ from which simple integrations yield $A(x) = -2\log x$ and

$$y(x) = x^2 + \frac{\varepsilon}{x^2} \, .$$

If $\varepsilon = 0$ then this solution is valid for all x otherwise we must take care to exclude $x = 0$ from consideration. However, even in the former case the numerical student should expect difficulties due to round off error.

Next we turn our attention to nonlinear first order equations of the form

$$y' = f(x, y) \ , \ \ y(x_0) = y_0 \, . \tag{7.1.8}$$

These problems differ significantly from the linear case. Perhaps most important for our discussion is that we can not write down a general analytic solution such as (7.1.4). Hence approximation techniques take an even a more important role.

Theorem 7.1.3 *If $f(x, y)$ and $\partial f/\partial y$ are continuous functions of x and y in a rectangle R given by $\alpha < x < \beta$ and $\gamma < y < \delta$ where $(x_0, y_0) \in R$ then a unique solution of (7.1.8) exists in an interval $(x_0 - h, x_0 + h) \subseteq (\alpha, \beta)$.*

Example 7.1.4

$$y' = x^2 + y^2 \tag{7.1.9}$$

By Theorem 7.1.3 we know that a unique solution of the differential equation exists. Indeed, the solution can be expressed explicitly (but inconveniently) in terms of Bessel functions. The isoclines

$$y' = \text{const} \tag{7.1.10}$$

aid in graphing solution curves. Here they are the circles

$$\{(x, y) : x^2 + y^2 = c^2\} \tag{7.1.11}$$

that are graphed in Figure 7.1.1 for $c_1 < 1 < c_2$. The solution curves passing through the x-axis at x_0 have the slopes

$$y'(x_0) = x_0^2 \, . \tag{7.1.12}$$

Numerical difficulties may be expected in the region

$$\{(x, y) : x^2 + y^2 > 1\} \tag{7.1.13}$$

where the slopes of the solution curves of (7.1.9) are very steep.

Figure 7.1.1: Isoclines of $x^2 + y^2 = c^2$.

Example 7.1.5 For an example not addressed by Theorem 7.1.3 consider the nonlinear differential equation

$$(y')^2 = \frac{1 - y^2}{1 - x^2}\, a(x)\ ,\quad x_0 \le x \le x_1 \tag{7.1.14}$$

where a is a continuous function of x on the interval $[x_0, x_1]$. Except for particular functions a, the solutions of (7.1.14) cannot be explicitly expressed.

A nonlinear differential equation that is linear in the highest derivative is called quasi-linear. If we take $a(x) = 1$ and then take the square root of each side of (7.1.14) we obtain two quasi-linear equations, namely

$$y' = \left(\frac{1 - y^2}{1 - x^2}\right)^{1/2} \quad \text{and} \quad y' = -\left(\frac{1 - y^2}{1 - x^2}\right)^{1/2} . \tag{7.1.15}$$

These equations are defined only in the regions

$$\begin{aligned} R_1 &= \{(x, y) : x^2 \le 1\ ,\ y^2 \le 1\} \\ R_2 &= \{(x, y) : x^2 \ge 1\ ,\ y^2 \ge 1\}\,. \end{aligned}$$

In R_1, the solutions of (7.1.15) are a one-parameter family of hyperbolas. For example, with reference to Figure 7.1.2, the solid curves are solutions to the first equation in (7.1.15) while the dotted curves are solutions to the second equation in (7.1.15) (solutions given by $x^2 - 2cxy + y^2 = 1 - c^2$ where $y(0) = 1 - c^2$). The solution curves are tangent to the boundary of R_1 at A, B, C, and D, so that starting from D and proceeding to C along the curve, y' increases from 0 to ∞. This property makes it difficult to solve (7.1.15) numerically.

Figure 7.1.2:

We next turn our attention to the general second order linear differential equation

$$y'' + a(x)y' + b(x)y = c(x)\ , \quad x_0 < x < x_1 \tag{7.1.16}$$

$$y(x_0) = y_0 \quad , \quad y'(x_0) = y_0' \tag{7.1.17}$$

where a, b and c are continuous functions defined on (x_0, x_1).

Theorem 7.1.6 *If $a(x)$, $b(x)$, and $c(x)$ are continuous functions of x on (x_0, x_1) then the solution of (7.1.16) exists on (x_0, x_1) and is unique.*

Despite its apparent simplicity, it is seldom possible to express the general solution of (7.1.16) in a form such as (7.1.4), involving only a finite number of integrations. If $c(x) = 0$ on $[x_0, x_1]$, then (7.1.16) is said to be homogeneous. Otherwise it is said to be nonhomogeneous. Once a nontrivial solution ($y(x) \neq 0$ on $[x_0, x_1]$) of the homogeneous equation is known, we can reduce (7.1.16) to a first order linear equation, i.e., we can explicitly express a solution in closed form (see Problem 7.1.3). The homogeneous form of (7.1.16) is

$$y'' + a(x)y' + b(x)y = 0\,.$$

The presence of $a(x)y'$ in (7.1.16) may be inconvenient in numerical work and it may be advantageous to transform (7.1.16) into another second order linear equation where there is no first derivative term present. This may be achieved in the homogeneous case via

$$y = v\exp\left[-\frac{1}{2}\,A\right]\ ,\ A(x) = \int_{x_0}^{x} a(t)dt \tag{7.1.18}$$

so that

$$v'' + g(x)v = 0\ ;\ g(x) = b(x) - \frac{1}{2}\,a'(x) - \frac{1}{4}\,a^2(x)\,. \tag{7.1.19}$$

If (7.1.16) is to be solved subject to the initial condition

$$y(x_0) = y_0\ ,\ y'(x_0) = y_0'$$

then the corresponding initial conditions for v become

$$v(x_0) = y_0\ ,\ v'(x_0) = y_0' + \frac{1}{2}\,a(x_0)y_0\,.$$

The solution of the problem (7.1.19) can be expressed in series form

$$v(x) = \sum_{n=0}^{\infty} v_n(x) \tag{7.1.20}$$

where

$$\begin{aligned} v_0(x) &= v(x_0) + v'(x_0)(x - x_0) \\ v_n(x) &= \int_{x_0}^{x} (t - x)g(t)v_{n-1}(t)dt\ ,\ \ n = 1, 2, 3\ldots \end{aligned} \tag{7.1.21}$$

Notice the linear dependence of the solution on the initial values. The procedure (7.1.20)–(7.1.21) may be a worthwhile procedure in applications if $|g(t)|$ and the length of interval $[x_0, x_1]$ are not too large. For if

$$|g(t)| \le M\ ,\ \ |v_0(t)| \le \mu \tag{7.1.22}$$

on the interval $[x_0, x_1]$, then it may be shown (Problem 7.1.2) that

$$|v_n(x)| \leq \frac{\mu M^n |x - x_0|^{2n}}{(2n)\,!} \tag{7.1.23}$$

Example 7.1.7 The equation

$$y'' + c(x)y = 0\;, \quad x \leq x \leq x_1\,, \tag{7.1.24}$$

where c is a large, positive and slowly varying function of x, has solutions that oscillate rapidly. It is difficult to obtain an accurate solution of this equation by standard numerical methods. This difficulty can be eliminated by the Madelong transformation

$$y = u(x)\cos[v(x)]\,. \tag{7.1.25}$$

In this transformation, y is expressed in terms of its modulus $u(x)$ and phase $v(x)$. Differentiating (7.1.25), we get

$$\begin{aligned} y' &= u'\cos(v) - uv'\sin(v) \\ y'' &= [u'' - u(v')^2]\cos v - [uv'' + 2u'v']\sin v \end{aligned}$$

If we substitute these expressions into the original expression and require that the coefficients of $\cos v$ and $\sin v$ vanish separately, we obtain the two equations

$$\begin{aligned} uv'' + 2u'v' &= 0 \\ u'' + [c - (v')^2]v &= 0 \end{aligned} \tag{7.1.26}$$

The first equation in (7.1.26) can be integrated at once, to yield

$$v'(x) = \frac{k}{u^2}$$

where k is a constant, so that

$$v(x) = k\int_{x_0}^{x} \frac{dt}{u^2(t)}\,. \tag{7.1.27}$$

The second equation in (7.1.26) then becomes

$$u'' + \left[c - \frac{k^2}{u^4}\right] u = 0\,. \tag{7.1.28}$$

If (7.1.24) is to be solved subject to the initial conditions

$$y(x_0) = y_0\;, \quad y'(x_0) = y_0'$$

then the corresponding initial conditions for (7.1.28) become

$$u(x_0) = y_0\;, \quad u'(x_0) = y_0'\,.$$

We may now choose the constant k^2 in (7.1.28) to make this equation simpler to solve numerically. One such choice is

$$k^2 = c(x_0)y_0^4\,.$$

The WKB method described in Chapter 6 may also be used to obtain an approximate solution of (7.1.24).

Example 7.1.8 An n'th order equation in one dependent variable can always be rewritten as a system of n first order equations. For example, in the Van der Pol equation

$$y'' + \mu(y^2 - 1)y' + y = 0 \tag{7.1.29}$$

where μ is a constant, we rename y as y_1 and y' as y_2. Then (7.1.29) becomes

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= \mu(y_1^2 - 1)y_2 - y_1 \end{aligned} \tag{7.1.30}$$

It is not always possible to proceed in the opposite direction, i.e., to replace a system of n first order equations in n dependent variables by a single n'th order equation in one independent variable.

The transformation from (7.1.29) to (7.1.30) makes it possible to use a numerical method designed for solving a first order equation, to solve an n'th order equation.

For first order equations a number of theorems are known to have practical consequences. The Lipschitz condition plays a fundamental role in existence and uniqueness theorems.

Definition 7.1.9 *Let $f(x, \vec{y})$, where $\vec{y} = (y_1, \ldots, y_n)$, be a function of the $n + 1$ variables $(x, \vec{y})$ in a convex domain D in $\Re^{n+1}$. We say that the function $f(x)$ satisfies Lipschitz conditions with respect to $\vec{y}$ in D if there exists a constant M (the Lipschitz constant) such that for all $(x, \vec{y})$ and $(x, \vec{z})$ in D*

$$|f(x, \vec{y}) - f(x, \vec{z})| \leq M \sum_{i=1}^{n} |y_i - z_i| \, , \tag{7.1.31}$$

where $\vec{z} = (z_1, \ldots, z_n)$.

In particular, the Lipschitz condition is satisfied if the partial derivatives of f with respect to y_i are all bounded in D; for then the moduli of the derivatives have a common upper bound M, that can serve as the Lipschitz constant.

Theorem 7.1.10 *Let the system of differential equations take the form*

$$y_i' = f_i(x, \vec{y}, \alpha, \beta, \ldots) \;, \quad i = 1, 2 \ldots, n$$

where $\alpha, \beta, \ldots$ are parameters. Then there exists a unique solution $y(x)$ that is a continuous vector function of the initial value $\vec{y}(x_0)$ and of the parameter $\alpha, \beta \ldots$ in every domain where the functions f_i are continuous functions of $(x, \vec{y}, \alpha, \beta, \ldots)$ and satisfy the Lipschitz condition with respect to y.

In this chapter our concern is mainly with a single first order equation, with a system of two first order equations, or with second order equations. Example 7.1.8 above illustrates that the analysis of a single first order equation can usually be extended to a system of n first order equations.

Problems

7.1.1 Find a power series solution of the problem

$$y'' = xy \ , \ \ y(0) = 1 \ , \ \ y'(0) = 0 \, .$$

7.1.2 Verify the inequality (7.1.23).

7.1.3 Let $Y(x)$ be a known solution of the differential equation

$$y'' + f(x)y' + g(x)y = 0 \ , \quad 0 \leq x \leq 1$$

where $Y(x) \neq 0$ in $[0,1]$. Assume that $u = vY$ and find a solution of

$$u'' + f(x)u' + g(x)u = h(x) \ , \quad 0 \leq x \leq 1 \ , \ \ u(0) = 1 \ , \ \ u'(0) = b \, .$$

7.1.4 Show that if $y = y(x)$ solves the problem

$$y'' + f(x,y) = 0 \ , \ \ 0 \leq x \leq 1 \ ; \quad y(0) = a \ , \ \ y(1) = b \, ,$$

then $y(x)$ also satisfies the equation

$$y(x) = a(1-x) + bx + \int_0^1 G(x,t) f(t, y(t)) dt$$

where

$$G(x,t) = \begin{cases} x(1-t) & \text{if } 0 \leq t \leq x \\ t(1-x) & \text{if } x \leq t \leq 1 \end{cases} .$$

Conversely, show that any twice differentiable function that satisfies (7.1.10) also satisfies the first equation given here.

7.1.5 Let M and N be two given functions of x and y such that $\frac{M}{y} = \frac{N}{x}$. Show that the differential equation

$$N(x,y)y' + M(x,y) = 0$$

has the “solution”

$$\int_{x_0}^{x} M(x, y_0) dx + \int_{y_0}^{y} N(x,t) dt = \text{const.}$$

7.1.6 Let m and k be the smallest positive integers such that $k = \frac{qm}{p+q}$, where p and q are integers, such that $q > 0$, $p + q > 0$. Show that the transformation $x = y^k$ reduces the integral $\int_0^1 x^{p/q} F(x)dx$ to the integral $\int_0^1 \frac{mq}{p+q}\, y^{m-1} F(y^k)dy$, i.e., it removes the singularity in the first integral if F has no singularities.

7.2 Methods for First Order Equations

Our goal in this section is two fold. With respect to the initial value problem

$$y' = f(x, y) \ , \ \ y(x_0) = y_0 \tag{7.2.1}$$

defined on the interval $[x_0, a]$, we outline the methods that are detailed in subsequent sections and second we explain the rationale behind these methods. We make an overriding assumption that a solution of the initial value problem on the interval $[x_0, a]$ exists and is unique. We denote this solution by $Y(x)$ where upper case notation distinguishes a true solution from an approximate.

We begin with the outline. In Section 7.3 we consider the Taylor series method applied to (7.2.1). As a practical tool, the Taylor series method is a reasonable choice only if (7.2.1) admits repeated differentiation. To make our point, the student of calculus need only think of computing the fifth derivative of the product of two functions. Clearly the circumstances are rare when Taylor series are useful other than as a theoretical tool.

We present difference methods in Section 7.4. These methods lead to approximations at the points

$$x_n = x_0 + n \ , \ \ n = 0, 1, 2, \ldots .$$

and, hence, are called discrete approximations. More specifically, y_n denotes the numerical estimate of the solution $Y(x_n)$. For notational purposes it is convenient to represent $Y(x_n)$ as Y_n and f_n as $f(x_n, y_n)$.

General k-step methods are the subject of Section 7.5. These methods take the form the form

$$y_{n+1} = \alpha_1 y_n + \alpha_2 y_{n-1} + \cdots + \alpha_k y_{n+1-k} + h(\beta_0 f_{n+1} + \beta_1 f_n + \cdots + \beta_k f_{n+1-k}) \,. \tag{7.2.2}$$

where $\alpha_1, \ldots, \alpha_k$ are coefficients whose specification distinguishes one k-step method from another. If the coefficients $\alpha_1, \ldots, \alpha_k$ and $\beta_0, \beta_1, \ldots, \beta_k$ in (7.2.2) are chosen such that

$$\gamma_0 \equiv \sum_{j=1}^{k} \alpha_j - 1 = 0 \ , \quad \gamma_1 \equiv \sum_{j=1}^{k} j\alpha_j - \sum_{j=0}^{k} \beta_j = 0 \tag{7.2.3}$$

and

$$\begin{aligned} \gamma_r &\equiv \sum_{j=0}^{k} \Big[(1-j)^r \alpha_j + r(1-j)^{q-1} \beta_j \Big] = 0 \ , \quad r = 2, \ldots, p \\ \alpha_0 &\equiv 1 \end{aligned} \tag{7.2.4}$$

then the method is said to have order p. The equations in (7.2.3) and (7.2.4) are obtained if we demand (7.2.2) to be satisfied exactly for $y(x) = x^r$, $r = 0, 1, \ldots, p$. Problem (7.2.1) shows that if the f in (7.2.1) is of class C^p in x and y and if $y_j = Y_j$ for $j = n+1-k, \ldots, n$ then

$$y_{n+1} - Y_{n+1} = \frac{\gamma_{p+1}}{(p+1)\,!}\, Y^{(p+1)}(\zeta) h^{p+1} \ , \quad x_{n+1-k} < \zeta < x_{n+1} \tag{7.2.5}$$

Explicit procedures for obtaining the constants α_j and β_j are explained in the following sections.

If $\beta_0 = 0$, then (7.2.2) is said to be explicit, since we can then explicitly solve for y_{n+1}, assuming that $y_n, y_{n-1}, \ldots, y_{n+1-k}$ are known. If $\beta_0 \neq 0$, then (7.2.2) is said to be implicit, since in that case y_{n+1} cannot in general be explicitly expressed. Specific forms of (7.2.2) are then used as correctors, for correcting the error of predictors (a predictor is almost always an explicit formula—see the examples in the next section) or else y_{n+1} is found via one of the methods of Chapter 3 for solving nonlinear equations. Except for the case of stiff problems (see Section 7.8), y_{n+1} is usually obtained via the following method of successive approximations:

$$\begin{aligned} y_{n+1}^{[0]} &= y_n \\ y_{n+1}^{[\nu+1]} &= h\beta_0 f\left(x_{n+1}, y_{n+1}^{[\nu]}\right) - u_n \ , \quad \nu = 0, 1, 2, \ldots \end{aligned} \tag{7.2.6}$$

where

$$u_n = \sum_{j=1}^{k} \left[\alpha_j y_{n+1-j} - h\beta_j f_{n+1-j}\right] . \tag{7.2.7}$$

If f is of class C^p in x and y and if h is sufficiently small then $y_{n+1}^{[\nu+1]} \to y_{n+1}$ as $\nu \to \infty$ and $y_{n+1}^{[\nu+1]}$ converges to the unique solution of (7.2.2) (see Problem 7.2.2). In practice, h is usually chosen sufficiently small so that no more than two or three iterations are required to compute y_{n+1} to the desired accuracy.

The integer k in (7.2.2) depends on the order p of the method. For larger orders p, it is necessary to take k larger in the construction of (7.2.2).

We also remark that the error expressed in (7.2.4) is proportional to $Y^{(p+1)}(\zeta)$ (of course, this is in general unknown) and h^{p+1}. In practice it is desirable that h be small so that this error may be small. However, if h is too small, the effect of roundoff error (see Section 7.13), the cost of computation as well as the effect of extraneous solutions (see Section 5.5) of (7.2.2) may destroy the accuracy of the computation. These extraneous solutions may grow more rapidly than the solution of (7.2.1) and mask it completely. Thus in practice we usually depend on crude criteria for obtaining a solution within a tolerable error, a criteria that is usually based on using (7.2.2) to compute an approximate solution for different values of h. Automatic programs exist that compute the step-size required to achieve a certain accuracy.

Equation (7.2.2) is simply a nonlinear finite difference equation of order k. Normally we do not attempt to obtain an explicit solution of this equation; instead we simply compute $y_k, y_{k+1}, \ldots$ and so on, repeating the process with a different value of h if necessary, until the error is acceptable. Even if we can explicitly solve the difference equation, we usually prefer a step-by-step solution for practical reasons.

Runge-Kutta methods are discussed in Section 7.6. The general Runge-Kutta method for solving (7.2.1) takes the form

$$y_1 = y_0 + h \sum_{j=0}^{n} a_j k_j \tag{7.2.8}$$

where

$$\begin{aligned} k_0 &= f(x_0 + \alpha_0 h, y_0) \\ k_1 &= f(x_0 + \alpha_1 h, y_0 + h\beta_{10} k_0) \\ k_2 &= f(x_0 + \alpha_2 h, y_0 + h\beta_{20} k_0 + h\beta_{21} k_1) \\ &\vdots \\ k_n &= f\left(x_0 + \alpha_n h, y_0 + \sum_{i=0}^{n-1} \beta_{ni} k_i\right) . \end{aligned} \tag{7.2.9}$$

The constants α_j, β_{ij} and a_j are chosen such that if y_0 is exact. As a result the sum on the right-hand side of (7.2.8) approximates the correct value of Y_1 to within an error of $0(h^p)$ as $h \to 0$, where p is as large as possible.

We also mention algorithms for computing an approximate solution of (7.2.1) using the notion of "extrapolation to the limit."

Euler's method

$$y_1 = y_0 + hf_0 \,,$$

is just the special case of (7.2.2) where we take $k = 1$, $\alpha_1 = 1$, $\beta_0 = 0$, is probably the simplest of all the methods for solving (7.2.1). It is also not very accurate, and its primary use therefore is to compute a rough approximation to the solution of (7.2.1).

Euler's method, the Taylor series method, Runge-Kutta methods, and methods based on extrapolation to the limit may be used to compute initial values for k-step methods, where $k > 1$.

Among methods in use, predictor-corrector methods based on Adams-type formulas and Runge-Kutta methods are the most popular. The predictor-corrector methods are self-checking, and they enable the user to obtain the solution within an acceptable tolerance. The popularity Runge-Kutta methods enjoy is because they are self-starting. Hence in contrast to predictor-corrector schemes, Runge-Kutta methods do not require an extra procedure to obtain the necessary starting values. We expect that due to their relative efficiency, Fehlberg's versions of Runge-Kutta methods will become popular during the next few years.

The usual method of estimating the error of Runge-Kutta methods recomputes the solution using the smaller step-size $h/2$. By making such a comparison at each step, we can determine the h required to achieve a certain tolerance of error.

Tests carried out by numerical analysts have shown that solution methods based on extrapolation to the limit can be very effective. The most popular of these methods is the method of Bulirsch and Stoer (see chapter references). Unfortunately this method is not yet fully understood. The main difficulty is that in contrast to the case of polynomial approximations, it is not always possible to tell a priori when a certain rational approximation scheme such as the Thiele algorithm (Problem 4.2.5) or the ϵ-algorithm produces a reliable approximation to a given function. Similar statements may be made regarding other methods of extrapolation to the limit.

It is hoped that the methods described in this chapter are some help to users faced with solving a problem of the type (7.2.1). In case of difficulty, we direct the user to more extensive literature(see chapter references).

Problems

7.2.1 Use the Taylor theorem expression

$$w(t) = \sum_{r=0}^{p} \frac{w^{(r)}(0)}{r\,!}\, t^r + \frac{1}{p\,!} \int_0^k (t-u)_+^p w^{(p+1)}(u)du$$

where $(t-u)_+ = \max(0, t-u)$ to prove that if (7.2.3) and (7.2.4) are satisfied, then

$$\sum_{s=0}^{k} [\alpha_s w(k-1) - \beta_s w'(k-s)] = \frac{\gamma_{p+1}}{(p+1)\,!}\, w^{(p+1)}(t)\,,, \tag{7.2.10}$$

for some $t \in (0, k)$. Use this result to verify (7.2.5).

7.2.2 Prove that if $|\beta_0|Mh < 1$, where M is a constant satisfying (1.28), then (7.2.2) has a unique solution y_{n+1}. Moreover, if $y_{n+1}^{[\nu]}$ is defined as in (7.2.6), then $y_{n+1}^{[\nu]} \to y_{n+1}$ as $v \to \infty$. Furthermore,

$$\left| y_{n+1}^{[\nu]} - y_{n+1} \right| \le \frac{(|\beta_0|Mh)^v}{1 - |\beta|Mh} \left| y_{n+1}^{[1]} - y_n \right| .$$

7.3 Taylor Series Method of Solution of $y' = f(x, y)$

Let us assume that the partial derivatives of f with respect to x and y are easily obtained. We set

$$y(x) = \sum_{k=0}^{\infty} \frac{y^{(k)}(x_0)}{k\,!}\, (x - x_0)^k\,. \tag{7.3.1}$$

Equation (7.2.1) then yields

$$\begin{aligned} y(x_0) &= y_0 \\ y'(x_0) &= f(x_0, y_0) \\ y''(x_0) &= f_x(x_0, y_0) + f_y(x_0, y_0)y'(x_0) \\ y^{(3)}(x_0) &= f_{xx}(x_0, y_0) + 2f_{xx}(x_0, y_0)y'(x_0) \\ &\quad + f_y(x_0, y_0)y''(x_0) + f_{yy}(x_0, y_0)y'(x_0)^2\,. \end{aligned} \tag{7.3.2}$$

Substituting these values back into the series expansion (7.3.1), we get a polynomial approximation to the solution $Y(x)$ of (7.2.1).

If $f(x,y)$ is sufficiently simple, it is sometimes possible to get an explicit recurrence relation for the coefficients in (7.3.1) by means of the identity

$$\sum_{n=0}^{\infty} a_n x^n \sum_{m=0}^{\infty} b_m x^m = \sum_{n=0}^{\infty}\left(\sum_{m=0}^{n} a_{n-m} b_m\right) x^m \tag{7.3.3}$$

Example 7.3.1 Consider a series solution of $y' = x^2 + y^2$, $y(0) = 0$. We attempt the series

$$y = \sum_{n=0}^{\infty} c_n x^n \,. \tag{7.3.4}$$

Assuming that this series converges in some neighborhood of $x = 0$ we have

$$y' = \sum_{n=0}^{\infty}(n+1)c_{n+1}x^n \;, \quad y^2 = \sum_{n=0}^{\infty}\left(\sum_{m=0}^{n} c_m c_{n-m}\right)x^n \,.$$

Substituting these results into the differential equation, we arrive at the equation

$$\sum_{n=0}^{\infty}\left[(n+1)c_{n+1} - \delta_{2n} - \sum_{m=0}^{\infty} c_m c_{n-m}\right] x^n = 0 \,,$$

where $\delta_{2n} = 1$ if $n = 2$ and 0 otherwise. Hence we get the system of equation

$$c_0 = 0 \;, \quad c_{n+1} = \frac{1}{n+1}\left[\delta_{2n} + \sum_{m=0}^{\infty} c_m c_{n-m}\right] \;, \quad n = 0, 1, 2, \ldots \,.$$

Thus the coefficients of y_k are

$$\begin{aligned} c_0 &= c_1 = c_2 = 0 \;, \quad c_3 = \frac{1}{3} \\ c_4 &= c_5 = c_6 = 0 \;, \quad c_7 = \frac{1}{7}\, c_3^2 \;, \ldots \end{aligned}$$

These relations suggest that

$$c_{4k} = c_{4k+1} = c_{4k+2} = 0 \;, \quad c_{4k+3} \neq 0 \;, \quad k = 0, 1, 2, \ldots \,.$$

Indeed, this may be verified by induction. Putting

$$\begin{aligned} n+1 &= 4k+3 \rightarrow n = 4k+2 \\ m &= 4s+3 \rightarrow n-m = 4k-4s-1 \end{aligned}$$

it follows that

$$c_3 = \frac{1}{3} \;, \quad c_{4k+3} = \frac{1}{4k+3}\sum_{s=0}^{k-1} c_{4s+3}c_{4k-4s-1} \;, \quad k = 1, 2, 3, \ldots \,.$$

We use this recurrence relation to obtain the following values of c_{4k+3} given in Table 7.3.1a. These values are used in (7.3.4) to obtain approximates of $y' = x^2 + y^2$, $y(0) = 0$ in Table 7.3.1b.

Table 7.3.1: Results for Taylor Series Applied to Example 1.

(a)

k	c_{4k+3}
0	.3333 3333
1	.0158 7302
2	.0007 6200
3	.0000 5955
4	.0000 0370
5	.0000 0023
6	.0000 0001

(b)

x	y
0.0	0.0000 000
0.1	0.0003 333
0.2	0.0026 669
0.3	0.0090 035
0.4	0.0213 594
0.5	0.0417 912
0.6	0.0724 479
0.7	0.1156 598
0.8	0.1740 883
0.9	0.2509 097
1.0	0.3502 318

Problems

7.3.1 Find a cubic polynomial approximation to the problem $y'(x) = x^2 + y^2$, $y(0) = 0$ by computing (7.3.2) at $(x_0, y_0) = (0, 0)$. To find, for example, the next nonzero coefficient in (7.3.1) for this problem via continuing with the computation in (7.3.2) is quite tedious. The method of Example 7.3.1 is far simpler.

7.3.2 (a) Show that the problem

$$\begin{bmatrix} y' \\ v' \end{bmatrix} = \begin{bmatrix} v \\ y - \frac{y^3}{x^2} \end{bmatrix} \quad ; \quad \begin{bmatrix} y(0) \\ v(0) \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \end{bmatrix}$$

has a power series solution

$$\begin{bmatrix} y \\ v \end{bmatrix} = \begin{bmatrix} \sum_{n=0}^{\infty} a_n x^{2n+1} \\ \sum_{n=0}^{\infty} (2n+1) a_n x^{2n} \end{bmatrix}$$

where $a_0 = \alpha$ and

$$a_{n+1} = \frac{1}{(2n+2)(2n+1)} \left(\left[a_n - \sum_{j=0}^{n} \sum_{i=0}^{j} a_i a_{j-i} a_{n-j} \right] \right) \quad , \quad n = 0, 1, 2, \ldots .$$

(b) Prove that the expansions in (a) for the given coefficients converge if $|x| < \sqrt{6}/\sqrt{1+\alpha^2}$.

7.4 Finite Difference Approximations to $y' = f(x,y)$

In this section we derive some simple formulas of the type (7.2.2).

If we integrate both sides of (7.2.1) from x_r to x_m, we obtain

$$y_m - y_r = \int_{x_r}^{x_m} f(x,y)dx\,, \tag{7.4.1}$$

where $x_r = x_0 + rh$. For example, if we integrate from x_0 to x_1 the identity

$$Y'(x) = Y'(0) + (x-x_0)Y''(\zeta)\ , \quad x_0 < \zeta < x$$

and replace y_0' by f_0, we obtain the classical Euler formula

$$y_1 = y_0 + hf_0 \tag{7.4.2}$$

with error (see (7.2.5))

$$y_1 - Y_1 = \frac{h^2}{2}\,Y''(\zeta)\ , \quad x_0 < \zeta < x_1\,. \tag{7.4.3}$$

The midpoint method is obtained by means of the identity

$$Y_{\pm 1} = Y(x_0 \pm h) = y_0 \pm hy_0' + \frac{h^2}{2}\,y_0'' \pm \frac{h^3}{6}\,y^{(3)}(\zeta_\pm)$$

where $\zeta_+(\zeta_-)$ is between x_0 and $x_1(x_{-1})$. By subtraction, we have

$$y_1 = y_{-1} + 2hf_0\,. \tag{7.4.4}$$

The error is given by

$$y_1 - Y_1 = \frac{h^3}{3}\,Y^{(3)}(n)\ , \quad x_{-1} < n < x_1\,. \tag{7.4.5}$$

Higher order approximations are derived by including more points. For example, if we interpolate f at x_{-2} and x_0 by a straight line and then integrate this approximation from x_{-2} to x_1, we obtain Heun's formula

$$y_1 - y_{-2} = \frac{3}{4}\,h(f_{-2} + 3f_0) \tag{7.4.6}$$

where the error is

$$y_1 - Y_1 = \frac{3}{8}\, h^4 Y^{(4)}(n)\ , \quad x_{-2} < n < x_1\,. \tag{7.4.7}$$

Similarly, if we intepolate f by a quadratic at x_{-2}, x_{-1} and x_0 and then integrate from x_{-2} to x_1, we obtain Milne's formula

$$y_1 = y_{-3} + \frac{4h}{3}\,(2f_{-2} - f_{-1} + 2f_0) \tag{7.4.8}$$

with error

$$y_1 - Y_1 = \frac{112}{360}\, h^5 Y^{(5)}(\zeta)\ , \quad x_{-3} < \zeta < x_1\,. \tag{7.4.9}$$

The rules (7.4.2), (7.4.4), (7.4.6) and (7.4.8) determine y_1 in terms of data at points to the left of x_1. In general, these relations are obtained by interpolating f to the left of x_1 with a Newton polynomial then replacing f in (7.4.1) by the Newton polynomial integrating . As may be expected by studying Chapter 4, these formulas are not very accurate. However, they do exhibit an advantage not shared by implicit rules; that is, formula (7.4.2) is self-starting hence requires no preliminary calculations. The formulas (7.4.4), (7.4.6) and (7.4.8) are not; for example, y_{-1} in (7.4.4) must be obtained prior to calculation.

Implicit formulas of the type (7.2.2) where $\beta_0 \neq 0$ are more generally more accurate than explicit formulas using the same points. To some extent this overcomes their starting difficulties. For instance, the formula

$$y_1 = y_0 + \frac{h}{2}\,(f_0 + f_1) \tag{7.4.10}$$

is more accurate than (7.4.2), since its error

$$y_1 - Y_1 = -\frac{h^3}{12}\, Y^{(3)}(\xi)\ , \quad x_0 < \xi < x_1 \tag{7.4.11}$$

exhibits a higher power of h. Of course, we caution that this new implicit procedure requires greater continuity. We refer to (7.4.10) as the trapezoidal rule. It is obtained by integrating from x_0 to x_1 the straight line that interpolates f at x_0 and at x_1. If f is a nonlinear function of y then we may not be able to solve (7.4.10) for y_1 directly, and it may be necessary to use the iterative methods given by (7.2.6) and (7.2.7). The equations (7.4.2) and

(7.4.10) may be used together as a predictor and corrector:

$$\begin{aligned} y_1^p &= y_0 + hf_0 \\ y_1^c &= y_0 + \frac{h}{2}(f_0 + f_1^p), \end{aligned} \tag{7.4.12}$$

where $f_1^p = f(x_1, y_1^p)$. This predictor-corrector pair is our first illustration of a *self-correcting formula*. If the difference $y_1^c - y_1^p$ is not within the tolerance of an acceptable error, the interval h must be decreased. The midpoint method (7.4.4) can be combined with the trapezoidal scheme (7.4.10) from x_{-1} to x_1 to yield the predictor-corrector scheme

$$\begin{aligned} y_1^p &= y_{-1} + 2hf_0 \\ y_1^c &= y_{-1} + h[f_{-1} + f_1^p] \end{aligned} \tag{7.4.13}$$

Equations (7.4.5) and (7.4.11) show that

$$\begin{aligned} y_1^p - Y_1 &= \frac{h^3}{3}\, Y^{(3)}(\zeta) \\ y_1^c - Y_1 &= -\frac{h^3}{3}\, Y^{(3)}(\eta) \end{aligned} \tag{7.4.14}$$

where ζ and η are between x_{-1} and x_1. The opposite sign of the error in these two formulas almost always makes the difference $y_1^p - y_1^c$ a reliable check on the error of computation.

Finally, the implicit formula

$$y_1 = y_{-1} + \frac{h}{3}(f_{-1} + 4f_0 + f_1) \tag{7.4.15}$$

with error

$$y_1 - Y_1 = -\frac{h^5}{90}\, Y^{(5)}(n) \tag{7.4.16}$$

may be used as a corrector for any of the procedures (7.4.2), (7.4.4), (7.4.6), or (7.4.8). The formula (7.4.15) is obtained by interpolating f at x_{-1}, x_0 and x_1 by a parabola and then integrating this parabola from x_{-1} to x_1.

The formula (7.4.15) is effective when used directly along with the iteration scheme (7.2.6)–(7.2.7). The result takes the form

$$\begin{aligned} y_1^{[0]} &= y_0 \\ y_1^{[v+1]} &= \frac{h}{3} f(x_1, y_1^{[v]}) + y_{-1} + \frac{h}{3}(4f_0 + f_{-1}), \quad v = 0, 1, 2, \ldots . \end{aligned} \tag{7.4.17}$$

A convenient property of implicit formulas such as (7.4.15) is that they are to some extent self-checking. For instance, if we suppose that (7.4.15) proceedes to a given point y_1 then we use (7.4.15) in the backward direction, i.e.,

$$y_{-1} = y_1 - \frac{h}{3}[f_{-1} + 4f_0 + f_1]$$

and verify the values already calculated. This practice is particularly useful in detecting unwanted growth of the solution (see Section 7.5).

Example 7.4.1 Solution of $y' = x^2 + y^2$, $y(0) = 0$, by Euler's method, $y_1 = y_0 + hf_0 = y_0 + h(x_0^2 + y_0^2)$. The results obtained are given in Table 7.4.1. Notice, upon comparing with the results of Table 7.3.1(b), that the error of the Euler method is quite large.

Table 7.4.1: Solution of $y' = x^2 + y^2$, $y(0) = 0$ by Euler's Method

x	y
0.0	0.0
0.1	0.0
0.2	0.001
0.3	0.0050 001
0.4	0.0140 026
0.5	0.0300 222
0.6	0.0551 123
0.7	0.0914 156
0.8	0.1412 518
0.9	0.2072 470
1.0	0.2925 421

Example 7.4.2 Solution of $y' = x^2 + y^2$, $y(0) = 0$, using the trapezoidal formula

$$y_1 = y_0 + \frac{h}{2}\left(x_0^2 + y_0^2 + x_1^2 + y_1^2\right).$$

Here y_1 can be conveniently solved for using the successive approximation technique of (7.2.6). The results obtained are given in Table 7.4.2. Notice that the results in Table 7.4.2 are considerably more accurate than those of the Euler method in Table 7.4.1.

Table 7.4.2: Solution of $y' = x^2 + y^2$, $y(0) = 0$ by the Trapezoidal Formula

x	y
0.0	0.0
0.1	0.0005 000
0.2	0.0030 005
0.3	0.0095 055
0.4	0.0220 343
0.5	0.0426 495
0.6	0.0735 106
0.7	0.1169 148
0.8	0.1756 922
0.9	0.2529 344
1.0	0.3528 586

Example 7.4.3 Solution of $y' = x^2 + y^2$, $y(0) = 0$, by the Milne-Simpson predictor-corrector method of (7.4.8) and (7.4.15):

$$\begin{aligned} y_1 &= y_{-3} + \frac{4h}{3}\left(2y'_{-2} - y'_{-1} + 2y'_0\right) \\ y_1 &= y_{-3} + \frac{2h}{3}\left(y'_{-3} + 4y'_{-1} + y'_1\right) \end{aligned}$$

The starting values at $x = 0.1, 0, .1$, and $.2$ are obtained from the series solution of Table 7.3.1.

In Table 7.4.3 the solution has been discontinued due to the large discrepancy between the results of the predictor and the corrector. In order to obtain more accuracy, a smaller mesh size should be used.

Table 7.4.3: Solution of $y' = x^2+y^2$, $y(0) = 0$ by the Milne-Simpson Method

x	y	y'
-.01	-.0003 333	.0100 001
0.0	.0000 000	.0000 000
0.1	.0003 333	.0100 001
0.2	.0026 669	.0400 071
0.3	.0090 019	Predictor
	.0090 055	Corrector
0.4	.0203 590	Predictor
	.0213 656	Corrector
0.5	.0418 046	Predictor
	.0417 794	Corrector

Problems

7.4.1 Compute a solution of the problem $y' = x^2+y^2$, $y(0) = 0$ by means of the method of (7.4.17) at $.2, .3, \ldots, 1.0$, given that $y_0 = 0$, $y_1 = 0.0003\ 33333$. Compare the results with those in Table 7.3.1.

7.4.2 Using (7.2.2) write down an arbitrary $k = 1$ step method. Compare your result to

$$\alpha_1 = 1 \quad \text{and} \quad \alpha_1 - (\beta_0 + \beta_1) = 0$$

which is obtained for the more specific case (7.2.3).

7.4.3 (a) Note that explicit methods require $\beta_0 = 0$. This yields Euler's method (7.4.2).?? I don't understand this problem.

(b) Implicit methods allow $\beta_0 \neq 0$. Define all implicit one-step methods. To what selection does (7.4.10) correspond? Justify (7.4.11) by computing γ_2 in (7.2.4).

7.5 Adams Type Methods for Solving $y' = f(x, y)$

A vast number of methods for solving (7.2.1) can be obtained by operational methods. For example, beginning with (7.4.1) ($m = 1$, $r = 0$)

$$\begin{aligned} \int_{x_0}^{x_1} f(x)dx &= h\int_0^1 f(x_0 + ph)\,dp \\ &= \frac{-h\nabla}{(1-\nabla)\ln(1-\nabla)}\, f_0\,. \end{aligned} \tag{7.5.1}$$

On expanding in powers of the backward difference ∇, we obtain the a formal series

$$\int_{x_0}^{x_1} f(x)dx = h\sum_{\ell=0}^{\infty} a_\ell \nabla^\ell f_0 \tag{7.5.2}$$

that is exact for any polynomial f. In (7.5.2)

$$\begin{aligned} a_0 &= 1 \\ a_\ell &= 1 - \sum_{k=0}^{\ell-1} \frac{a_k}{k-\ell+1}\,, \quad \ell = 1, 2, 3, \ldots\,. \end{aligned} \tag{7.5.3}$$

In particular

$$a_0 = 1\,,\ \ a_1 = \frac{1}{2}\,,\ \ a_2 = \frac{5}{12}\,,\ \ a_3 = \frac{3}{8}\,,\ \ a_4 = \frac{251}{720}\,. \tag{7.5.4}$$

Since the backward differences of f_0 employ $f_0, f_{-1}, f_{-2}, \ldots$, i.e., the values of f to the left of f_0, the formulas obtainable by taking the first k terms of the expansion on the right-hand side of (7.5.2) are all explicit, and take the form

$$y_1 = y_0 + h\sum_{\ell=0}^{k} a_\ell \nabla^\ell f_0 \tag{7.5.5}$$

where the a_k are defined in (7.5.3). Moreover, if $y_0 = Y_0, y_{-1} = Y_{-1}, \ldots, y_{-k} = Y_{-k}$, and if $f \in C^{k+1}$, then (Problem 7.5.3)

$$y_1 - Y_1 = a_{k+1}h^{k+2}Y^{(k+2)}(\eta)\,, \quad x_{-k} < \zeta < x_1\,. \tag{7.5.6}$$

Similarly, we employ the formal representation (Problem)

$$\begin{aligned} \int_{x_0}^{x_1} f(x)dx &= -\frac{h\nabla}{\log(1-\nabla)} f_1 \\ &= h\sum_{\ell=0}^{\infty} b_\ell \nabla^\ell f_1 \end{aligned} \tag{7.5.7}$$

that is exact for any polynomial f. In (7.5.7)

$$\begin{aligned} b_0 &= 1 \\ b_\ell &= -\sum_{k=0}^{\ell-1} \frac{b_k}{\ell-k+1}\ , \quad \ell = 1,2,3,\ldots . \end{aligned} \tag{7.5.8}$$

In particular

$$b_0 = 1\ ,\ b_1 = -\frac{1}{2}\ ,\ b_2 = -\frac{1}{12}\ ,\ b_3 = -\frac{1}{24}\ ,\ b_4 = -\frac{19}{720}\ . \tag{7.5.9}$$

Truncation of the expansion (7.5.7) yields the implicit formula

$$y_1 = y_0 + h\sum_{\ell=0}^{k} b_1 \nabla^\ell f_1 \tag{7.5.10}$$

where the b_ℓ are defined as in (7.5.8) and where, in the notation of (7.5.6)

$$y_1 - Y_1 = b_{k+1}h^{k+2}Y^{(k+2)}(\eta)\ , \quad x_{1-k} < \zeta < x\ . \tag{7.5.11}$$

The equations (7.5.5) and (7.5.10) are often used together as predictor and corrector respectively. It can be seen by inspection that the initial coefficients b_ℓ are smaller in magnitude than the corresponding coefficients a_ℓ. Indeed, it can be shown that $b_\ell/a_\ell \to 0$ as $\ell \to \infty$, and that a_ℓ and b_ℓ are of opposite sign for all $\ell \geq 1$. Hence if (7.5.10) is used as a corrector, then for $Y^{(k+2)}(x)$ of constant sign on $[x_0, a]$, the errors in these formulas are opposite in sign. The methods defined by (7.5.5) and (7.5.10) yield some simple predictors and correctors in Tables 7.5.1 and 7.5.2.

Table 7.5.1: Adams Predictors

$$
\begin{aligned}
y_1 - y_0 &= h \sum_{\ell=0}^{k} a_\ell \nabla^\ell f_0 \\
a_0 = 1\ , \ a_\ell &= 1 - \sum_{k=0}^{\ell-1} \frac{a_k}{\ell - k + 1}
\end{aligned}
$$

k	α_0^k	α_1^k	α_2^k	α_3^k	Error
0	1	0	0	0	$\frac{h^2}{2}\ Y''(\eta)$
1	3/2	-1/2	0	0	$\frac{5h^3}{12}\ Y'''(\eta)$
2	23/12	-16/12	5/12	0	$\frac{3h^4}{8}\ Y^{iv}(\eta)$
3	55/24	-59/24	37/24	-9/24	$\frac{251}{720}\ h^5 Y^v(\eta)$

$$y_1 - y_0 = h \sum_{\ell=0}^{k} \alpha_\ell^k f_{-\ell}$$

Problems

7.5.1 Use (a) the $k = 2$ and (b) the $k = 3$ predictor-corrector method of Table 7.5.1 to approximate $y(1)$, where $y' = y$, $y(0) = 1$, with $1/h = 4, 8, 16, 32$. For each of these cases, use the Taylor series approximation $y = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{120}$ to get the required initial values of y.

7.5.2 Verify the formula in (7.5.3). Hint: Recall the expansions

$$\frac{x}{1-x} = x^2 + x^3 + x^4 + \cdots$$

and

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

Put

$$\frac{x}{1-x} = \left(\sum_{j=0}^{\infty} a_j x^j \right) \left(\sum_{j=1}^{\infty} \frac{x^j}{j} \right)$$

and equate coefficients. (The same procedure verifies (7.5.9).)

Table 7.5.2: Adams Correctors

$$
\begin{aligned}
y_1 - y_0 &= h \sum_{\ell=0}^{k} b_\ell \nabla^\ell f_1 \\
b_0 = 1\ ,\ \ b_\ell &= -\sum_{k=0}^{\ell-1} \frac{b_k}{\ell - k + 1}
\end{aligned}
$$

k	β_0^k	β_1^k	β_2^k	β_3^k	Error
0	1	0	0	0	$\frac{-h^2}{2}\, Y''(\zeta)$
1	1/2	1/2	0	0	$\frac{-1}{12}\, h^3 Y'''(\zeta)$
2	5/12	8/12	-1/12	0	$\frac{-1}{24}\, h^4 Y^{iv}(\zeta)$
3	9/24	19/24	-5/24	1/24	$\frac{-19}{720}\, h^5 Y^{v}(\zeta)$

$$y_1 - y_0 = h \sum_{\ell=0}^{k} \beta_\ell^k f_{1-\ell}$$

7.5.3 Verify (7.5.6). Hint: Use the γ_r in (7.2.4).

7.6 Runge-Kutta Methods for Solving $y' = f(x, y)$

Let us refer to (7.2.9), and let us consider the case $n = 1$. We want to determine a_0, a_1, α_0, and α_1 so that

$$y_1 = y_0 + h\left[a_0 f(x_0 + \alpha_0 h, y_0) + a_1 f(x_0 + \alpha_1 h, y_0 + h\beta_{10} f(x_0 + \alpha_0 h, h_0))\right] \tag{7.6.1}$$

approximates $Y(x_0 + h)$ as accurately as possible in the limit as $h \to 0$. Using the notation

$$f^{(j,k)} = \frac{\partial^{j+k} f}{\partial^j x \partial^k y}$$

the function f, if sufficiently smooth, may be expanded as follows:

$$f(x_0 + u, y_0 + v) \;=\; \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{f^{(j,k)}(x_0, y_0)}{j!\, k!}\; u^j v^k$$

$$= f(x_0, y_0) + u f_x(x_0, y_0) + v f_y(x_0, y_0) + 0(h^2) \tag{7.6.2}$$

as $h \to 0$, provided that both u and v are $0(h)$ as $h \to 0$. Also

$$Y(x_0 + h) = Y(x_0) + hY'(x_0) + \frac{h^2}{2} Y''(x_0) + 0(h^3)$$

as $h \to 0$, and since $Y'(x) = f(x, Y(x))$, $Y(x_0) = y_0$,

$$Y(x_0 + h) = y_0 + hf_0 + \frac{h^2}{2} [f_x(x_0, y_0) + f_0 f_y(x_0, y_0)] + 0(h^3) \tag{7.6.3}$$

as $h \to 0$.

Using (7.6.2) we get

$$\begin{aligned} k_0 &= f(x_0, y_0) + \alpha_0 h f_x(x_0, y_0) + 0(h^2)\ , \quad h \to 0 \\ k_1 &= f(x_0, y_0) + \alpha_1 h f_x(x_0, y_0) + \beta_{10} h k_0 f_y(x_0, y_0) + 0(h^2)\ , \quad h \to 0 \\ &= f(x_0, y_0) + \alpha_1 h f_x(x_0, y_0) + \beta_{10} h f(x_0, y_0) f_y(x_0, y_0) + 0(h^2)\ , \quad h \to 0 \end{aligned} \tag{7.6.4}$$

Substituting (7.6.4) into (7.6.1) we obtain

$$\begin{aligned} y_1 &= y_0 + (a_0 + a_1) h f_0 \\ &\quad + h^2 [(a_0 \alpha_0 + a_1 \alpha_1) f_x(x_0, y_0) + a_1 \beta_{10} f_0 f_y(x_0, y_0)] \\ &\quad + 0(h^2)\ , \quad h \to 0 \end{aligned} \tag{7.6.5}$$

We now equate equal powers of h in (7.6.3) and (7.6.5) to arrive at the system of equations

$$\begin{aligned} a_0 + a_1 &= 1 \\ a_0 \alpha_0 + a_1 \alpha_1 &= \frac{1}{2} \\ a_1 \beta_{10} &= \frac{1}{2} \end{aligned} \tag{7.6.6}$$

with one solution given by

$$\begin{aligned} a_0 &= a_1 = 1/2 \\ \alpha_0 &= 0\ , \quad \alpha_1 = 1 \\ \beta_{10} &= 1\ . \end{aligned} \tag{7.6.7}$$

This solution yields the Runge-Kutta method

$$y_1 = y_0 + \frac{h}{2}(k_0 + k_1) \tag{7.6.8}$$

where by (7.6.7),

$$\begin{aligned} k_0 &= f_0 \\ k_1 &= f(x_0 + h, y_0 + hk_0) \end{aligned} \tag{7.6.9}$$

and where if $y_0 = Y_0$, then $y_1 - Y_2 = 0(h^2)$, $h \to 0$. The method (7.6.8) with k_0 and k_1 as in (7.6.9) is just the previously derived predictor-corrector method (7.4.12). However, the system (7.6.2) has other solutions that yield Runge-Kutta methods that are different from (7.4.12). All of these are second order methods.

By means of the procedure used above, we can derive more accurate Runge-Kutta methods. The most popular of these is a fourth order method that takes the form

$$y_1 = y_0 + \frac{h}{6}(k_0 + 2k_1 + 2k_2 + k_3)\,, \tag{7.6.10}$$

where

$$\begin{aligned} k_0 &= f_0 \\ k_1 &= f\left(x_0 + \frac{1}{2}\,h\;,\; y_0 + \frac{1}{2}\,hk_0\right) \\ k_2 &= f\left(x_0 + \frac{1}{2}\,h\;,\; y_0 + \frac{1}{2}\,hk_1\right) \\ k_3 &= f(x_0 + h\;,\; y_0 + hk_2) \end{aligned} \tag{7.6.11}$$

and where

$$y_1 - Y_1 = 0(h^5) \text{ as } \; h \to 0\,. \tag{7.6.12}$$

A very accurate variant of a Runge-Kutta method is due to Fehlberg. Given (7.2.1) one sets [34]

$$y = \bar{y} + y_0'(x - x_0) + \frac{1}{2}\,y_0''(x - x_0)^2 \tag{7.6.13}$$

where y_0' and $y_0''/2$ denote the Taylor coefficients of the solution y of (7.2.1), to get the problem

$$\bar{y}' = \bar{f}(x, \bar{y}) \ , \ \ \bar{y}(x_0) = \bar{y}_0 = y_0 \ , \tag{7.6.14}$$

where

$$\bar{f}(x, \bar{y}) = f(x, y) - y_0' - y_0''(x - x_0) \ .. \tag{7.6.15}$$

One next computes $\bar{y}_1$ by means of the formula

$$\bar{y}_1 = \bar{y}_0 + \frac{1}{1152}\,(500k_1 + 375k_2 + 64k_3) \tag{7.6.16}$$

where

$$\begin{aligned} k_1 &= h\bar{f}\left(x_0 + \frac{2}{5}\,h, \bar{y}_0\right) \\ k_2 &= h\bar{f}\left(x_0 + \frac{4}{5}\,h, \bar{y}_0 + \frac{4}{3}\,k_1\right) \\ k_3 &= h\bar{f}\left(x_0 + h, y_0 + \frac{5}{128}\,(27k_2 - 76k_1)\right) \ . \end{aligned} \tag{7.6.17}$$

The method is surprisingly accurate; it may be shown that

$$\bar{y}_1 - Y_1 = 0(h^7) \tag{7.6.18}$$

as $h \to 0$. The only drawback is that y_0' and y_0'' must be computed at each step. The computation of y_0' is trivial: $y_0' = f(x_0, y_0)$. Thus if y_0'' could also be easily computed at each step, the above method would be extremely efficient. Other more accurate variants of this procedure may be found in [35, 36, 37].

The actual error in (7.6.5) or (7.6.10) is difficult to estimate. Usually one applies a Runge-Kutta method at two different intervals over the same range and compares the results. A better approximation can then be obtained by extrapolation (Section 7.7).

Example: Runge-Kutta method of solution of $y' = x^2 + y^2$. $y(0) = 0$. The formula used is

$$y_1 = y_0 + \frac{h}{6}\,(k_0 + 2k_1 + 2k_2 + k_3)$$

where

$$
\begin{aligned}
k_0 &= x_0^2 + y_0^2 \\
k_1 &= \left(x_0 + \frac{1}{2}\,h\right)^2 + \left(y_0 + \frac{1}{2}\,hk_0\right)^2 \\
k_2 &= \left(x_0 + \frac{1}{2}\,h\right)^2 + \left(y_0 + \frac{1}{2}\,hk_1\right)^2 \\
k_3 &= (x_0 + h)^2 + (y_0 + hk_2)^2\,.
\end{aligned}
$$

Using an interval $h = 0.2$, we obtain the solution given in Table 7.6.1.

Table 7.6.1: Solution of $y' = x^2 + y^2$, $y(0) = 0$, by a Runge-Kutta Method.

$h = .2$		$h = 1$	
x	y	x	y
0.0	0.0000 0000	0.1	.0003 333
0.2	0.0026 6687	0.2	.0026 669
0.4	0.0213 6009	0.3	.0090 035
0.6	0.0742 5120	0.4	.0213 594
0.8	0.1759 7804	0.5	.0417 913
1.0	0.3523 4874	0.6	.0724 481
		0.7	.1156 603
		0.8	.1740 810
		0.9	.2509 078
		1.0	.3502 337

The construction of fourth and higher order Runge-Kutta methods is very complicated. Nevertheless with sixth [19] and eighth [15, 20] order Runge-Kutta methods have been constructed. For example, finding the coefficients a_j, α_j, and β_{ij} for an eighth order Runge-Kutta method (7.2.8)–(7.2.9) requires the solution of 200 nonlinear algebraic equations.

Problems

7.6.1 Verify (7.6.12).

7.6.2 Show that when $f(x, y) = f(x)$ is independent of y, the fourth order Runge-Kutta method of Equations (7.6.10)–(7.6.11) reduces to Simpson's rule.

7.7 Extrapolation to the Limit

Extrapolation to the limit techniques may be used to improve the accuracy of an approximate solution of problem

$$y' = f(x,y) \qquad\qquad y(x_0) = y_0 \tag{7.7.1}$$

computed by a finite difference method of the type discussed in the previous sections. In this section we illustrate various procedures that may be used for carrying out extrapolation to the limit. All of the schemes we discuss involve obtaining an improved approximation to $g(0)$, where $g(h)$ is a function that can be computed for $h > 0$ and is known to have an asymptotic expansion of the form

$$g(h) \sim c_0 + c_1 h^{k_1} + c_2 h^{k_2} + c_3 h^{k_3} + \cdots \tag{7.7.2}$$

as $h \to 0$. In (7.7.2) the constants $c_0, c_1, c_2, \ldots$ are known to exist but their true values are not known, and $k_1 < k_2 < k_3 < \cdots$, are positive constants.

Indeed, under the absence of roundoff error, approximate solutions of (7.7.1) computed by means of one of the methods of the previous section do take the form (7.7.2) where c_0 denotes the exact solution. For example, the following result is established in [12, page 25].

Theorem 7.7.1 *Let f be of class C^p in x and y, and let $y(x,h)$ denote the solution of (7.7.1) computed by means of Euler's method, where $x_0 < x \leq a$. Then*

$$y(x,h) = Y(x) + c_1(x)h + c_2(x)h^2 + \cdots + c_p(x)h^p + 0(h^{p+1}) \tag{7.7.3}$$

as $h \to 0$.

A similar result holds if $y(x,h)$ is computed by means of a higher order method. For example, for the case of a method of order 2, $c_1(x) = 0$, $c_2(x) \not\equiv 0$, unless $Y(x)$, the true solution of (7.7.1), is a parabola in x.

7.7.1 Richardson's Method of Extrapolation

Let us now describe the application of Richardson's method of extrapolation [17] to the function $g(h)$ that satisfies (7.7.2). The most rapid rate of convergence can be achieved in the case when the constants $k_1, k_2, k_3, \ldots,$ are known.

We define a sequence $\{g_n(h)\}_{n=1}^{\infty}$ by

$$\begin{aligned} g_1(h) &= g(h) \\ g_{n+1}(h) &= g_n(h) + \frac{g_n(h) - g_n(qh)}{q^k n - 1} \end{aligned} \tag{7.7.4}$$

where $q \neq 1$.

Theorem 7.7.2 *Let the sequence $\{g_n(h)\}_{n=1}^{\infty}$ be defined as above. Then there exist constants $c_n^{(n)}, c_{n+1}^{(n)}, c_{n+2}^{(n)}, \ldots$, such that*

$$g_n(h) \sim c_0 + c_n^{(n)} h^{k_n} + c_{n+1}^{(n)} h^{k_{n+1}} + c_{n+2}^{(n)} h^{k_{n+2}} + \cdots$$

as $h \to 0$.

Proof: The result (7.7.4) is clearly satisfied for the case $n = 1$. Assuming that the expansion (7.7.4) is valid for $n = m$, we find, upon substituting into (7.7.4), that

$$g_{m+1}(h) \sim c_0 + c_{m+1}^{(m+1)} h^{k_{m+1}} + c_{m+2}^{(m+1)} h^{k_{m+2}} + \cdots$$

as $h \to 0$, where

$$c_s^{(m+1)} = \frac{q^{k_m} - q^{k_s}}{q^{k_m} - 1}\, c_2^{(m)} \quad , \quad s = m+1, m+2, \ldots .$$

This completes the proof.

If the constants $k_1, k_2, \ldots$ are not known, the method of Equations (7.7.4) may still be applied by taking $k_n = n$, although the resulting sequence may not converge to the limit as rapidly in that case.

In practice one computes a table of entries $g_n(q^j h)$ where $0 < q < 1$ (usually $q = 1/2$) according to the layout of Table 7.7.1.

It then follows from Theorem 7.7.2 that each column of Table 7.7.1 converges to c_0, the rate of convergence of the n'th column being given by the asymptotic relation

$$g_n(q^j h) \sim c_0 + c_n^{(n)} (q^j h)^{k_n} + c_{n+1}^{(n)} (q^j h)^{k_{n+1}} + c_{n+2}^{(n)} (q^j h)^{k_{n+2}} + \cdots$$

as $q^j h \to 0$.

Table 7.7.1: Richardson's Method of Extrapolation

$g_1(h)$			
$g_1(qh)$	$g_2(h)$		
$g_1(q^2h)$	$g_2(qh)$	$g_3(h)$	
$g_1(q^3h)$	$g_2(q^2h)$	$g_3(qh)$	$g_4(h)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table 7.7.2: Richardson's Extrapolation combined with Euler's Method

h	$y_1(1,h)$	$y_2(1,h)$	$y_3(1,h)$	$y_4(1,h)$
.2	.2428567456			
		.3422274636		
.1	.2925421046		.3497656855	
		.34788113		.3502140256
.05	.0320211617		.3501557983	
		.3495887698		
.025	.3349001935			

Example 7.7.3 Euler's Method and Richardson's Method of Extrapolation. We compute $y_1(1,h)$ (see Theorem 7.7.1) for the problem $y' = x^2 + y^2$, $y(0) = 0$, by Euler's method for $h = 0.2, 0.1, .05$ and 0.025, to get the entries in column 1 of Table 7.7.2. Column 2 of this table is obtained by applying the formulas (7.7.4) with $k_1 = 1$, $q = 1/2$, while column 3 is obtained by applying (7.7.4) with $k_2 = 2$, $q = 1/2$.

7.7.2 Aitken's Method of Extrapolation

This method was previously discussed in Section 3.16. The layout again takes the form of Table 7.7.1, where, given $g_1(q^jh)$ we compute $g_2(q^jh), g_3(q^jh), \ldots$ by means of the formula

$$g_{n+1}(q^jh) = g_n(q^jh) - \frac{[g_n(q^{j+1}h) - q_n(q^jh)]^2}{g_n(q^{j+2}h) - 2g_n(g^{j+1}h) + g_n(q^jh)}\,. \tag{7.7.5}$$

Theorem 7.7.4 *Let $g_1(h)$ satisfy the relation (7.7.2) and let $0 < |q| < 1$. Then there exist constants $d_n^{(n)}, d_{n+1}^{(n)}, d_{n+3}^{(n)}, \ldots$, and $K_n^{(n)}, K_{n+1}^{(n)}, K_{n+2}^{(n)}, \ldots$, such that $K_1^{(0)} = k_1$, $K_n^{(n)} = K_n^{(n-1)}$, $n = 1, 2, 3, \ldots$, and such that*

$$g_n(q^j h) \sim c_0 + d_n^{(n)}(q^j h)K_n^{(n)} + d_{n+1}^{(n)}(q^j h)K_{n+1}^{(n)} + d_{n+2}^{(n)}(q^j h)K_{n+2}^{(n)} + \cdots \quad (7.7.6)$$

as $q^j h \to 0$.

We leave the proof of Theorem 7.7.4 to the problems.

Example 7.7.5 Euler's Method and Aitken's δ^2 Method. Starting with the values of $h_1(1, h)$ ($h = 0.2, 0.1, 0.05$) in Table 7.7.2, we compute an improved approximation $y_2^*(1, .2)$ to the value $Y(1)$ of the solution $Y(x)$ of the problem $y' = x^2 + y^2$, $y(0) = 0$, by means of (7.7.5):

$$\begin{aligned} Y(1) \simeq y_2(1, .2) &= y_1(1, .2) - \frac{[y_1(1, .2) - y_1(1, .1)]^2}{y_1(1, .2) - 2y_1(1, .1) + y_1(1, .05)} \\ &= .24285\ 67 - \frac{(0.5977\ 18)^2}{-.04218\ 87} \\ &= .32753\ 98\,. \end{aligned}$$

7.7.3 Extrapolation by Continued Fractions

The method of Bulirsch and Stoer uses a continued fraction procedure to extrapolate $g(0)$. We illustrate such a continued fraction method using the Thiele algorithm (Problem 4.2.5).

Example 7.7.6 Thiele's continued fraction method (Problem 4.2.5) yields

$$\begin{aligned} p(.2, .1) &= \frac{.2 - .1}{.24285\ 67 - .30262\ 85} = -1.6730\ 30 \\ p(.1, .05) &= \frac{.1 - .05}{.30262\ 85 - .24285\ 67} = -2.8436\ 40 \\ p(.2, .1, .05) &= \frac{.2 - .05}{-1.6730\ 30 - (-2.8436\ 40)} + .30262\ 85 \\ &= .43076\ 68 \end{aligned}$$

Hence, from Figure ???4.2.1, we get the convergent

$$r(h) = .24285\ 67 + \frac{h - .2}{-1.6730\ 30} + \frac{h - .1}{.43076\ 68 - .24285\ 67}$$

so that

$$y_1(1, 0) \simeq r(0) = .33433\ 35\,.$$

The above examples show that each of three methods yields a different improved value of the limit. The problems that follow show that if more values of $y(1,h)$ are computed by means of Euler's method, then each of the three methods of extrapolation may be used to get a very accurate approximation of $y(1,0)$.

While each of the columns in Table 7.7.1 converges to $g(0)$ for both Richardson's and Aitken's methods, it is not possible to prove, in general, that the n'th convergent of a continued fraction method converges to $g(0)$ as $n \to \infty$.

Problems

7.7.1 Continue the procedure of Example 1 of this section, by first using Euler's method to compute $y_1(1, .025)$ and $y_1(1, .0125)$ and then computing $y_n(1, .025)$ and $y_n(1, .0125)$ for $n = 2, 3, 4$, and 5 by means of the extrapolation method of (7.7.3).

7.7.2 Use the five computed values of $y_1(1, h)$ referred to in Problem 7.7.1 to construct a Thiele rational approximation, and use this approximation to extrapolate $y(1, 0)$.

7.7.3 Using the values $y_1(1, h)$ referred to in Problem 7.7.1 obtain a new sequence of approximations $y_2^*(1, h)$, $h = .2, .1, .05$, by means of Aitken's method.

7.7.4 Prove Theorem 7.7.4.

7.7.5 Devise a method of polynomial extrapolation to the limit based on:

(a) Aitken's algorithm (Example 9, Section 3.2).

(b) Neville's algorithm (Problem 3.2.4).

Apply each procedure to the values $y_n(1, h)$, $h = .2/2^{k-1}$, $k = 1, 2, 3, 4, 5$, and compare the results.

7.7.6 Let a Runge-Kutta method be of order p, i.e., if $y_0 = Y(x_0)$, then $y_1 = Y(x_1) + c_1 h^{p+1} + c_2 h^{p+2} + \cdots$. Let y_1^* and y_1^{**} be approximations of $Y(x_1)$ computed by the method using step sizes h and $h/2$ respectively.

(a) Discuss the use of the estimate

$$|y_1^{**} - Y(x_1)| \simeq \frac{|y_1^{**} - y_1^*|}{2^p - 1}$$

in the process of computation.

(b) In what sense is the approximation

$$y_1^R = y_1^{**} + \frac{y_1^{**} - y_1^*}{2^p - 1}$$

an improved approximation of $Y(x_1)$?

7.8 Direct Methods for Second Order Differential Equations

In this section we derive some numerical methods for solving the equations

$$y'' = g(x, y, z) \quad , \quad z = y' \tag{7.8.1}$$

and

$$y'' + P(x)y' + Q(x)y = R(x) \, . \tag{7.8.2}$$

We repeat a remark made in Section 7.1, namely that these equations can be solved by the methods of Sections 7.1 to 7.5, by writing them in the form of two first order equations. For example, (7.8.1) can be written in the form

$$\begin{aligned} y' &= z \\ z' &= g(x, y, z) \, . \end{aligned}$$

We can then apply any of the single equations methods to each of these equations.

The operator expansions (see 5.1.15 and 5.1.13)

$$hD = 2\sinh^{-1}(\delta/2) = 7\sum_{n=0}^{\infty} \frac{(1/2)_n(-1)^n\delta^{2n+1}}{(2n+1)2^{2n+1}n\,!} \tag{7.8.3}$$

and

$$\mu^p = \left(1 + \frac{1}{4}\,\delta^2\right)^{p/2} = \sum_{n=0}^{\infty} \frac{(-p/2)_n(-1)^n\delta^{2n}}{2^{2n}n\,!} \tag{7.8.4}$$

enable us to get expressions for higher derivatives, namely

$$\begin{aligned} h^r D^r &= \left[\frac{2}{\delta}\,\sinh^{-1}\frac{\delta}{2}\right]^r \delta^r \\ &= \left[1 - \frac{r}{24}\,\delta^2 + \frac{5r^2 + 22r}{5760}\,\delta^4 \right. \\ &\quad \left. - \frac{35r^3 + 462r^2 + 1528r}{2903\ 040}\,\delta^6 + \cdots\right]\delta^r \end{aligned} \tag{7.8.5}$$

and

$$h^r D^r = \left[\frac{1}{\mu}\left(\frac{2}{\delta}\,\sinh^{-1}\frac{\delta}{2}\right)^r\right]\mu\delta^r$$

$$= \left[1 - \frac{r+3}{24}\,\delta^2 + \frac{5r^2+52r+135}{5760}\,q^4 - \frac{35r^3+777r^2+5749r+14175}{2903\,040}\,\delta^6 + \cdots\right]\mu\delta^r\,. \tag{7.8.6}$$

In these expansions we may take $r = 0, \pm 1, \pm 2, \ldots$.

The expansions (7.8.5) and (7.8.6) may be applied to (7.8.1) to yield the following expansions:

$$\begin{aligned} \delta^2 y_0 &= \left[\frac{2}{\delta}\,\sinh^{-1}\frac{\delta}{2}\right]^{-2} h^2 D^2 y_0 = h^2\left[\frac{2}{\delta}\,\sinh^{-1}\frac{\delta}{2}\right]^{-2} g_0 \\ &= h^2\left[1 + \frac{\delta^2}{12} - \frac{\delta^4}{240} + \frac{31\delta^6}{60\,480} - \cdots\right] g \end{aligned} \tag{7.8.7}$$

$$\begin{aligned} hz_0 &= \left[\frac{1}{\mu}\left(\frac{2}{\delta}\,\sinh^{-1}\frac{\delta}{2}\right)\right]\mu\delta y_0 \\ &= \frac{1}{2}\left[\delta - \frac{\delta^3}{6} + \frac{\delta^5}{30} - \frac{\delta^7}{140} + \cdots\right]\left(y_{1/2} + y_{-1/2}\right) \end{aligned} \tag{7.8.8}$$

$$\begin{aligned} 2\mu\delta z_0 &= 2h\left[\mu\left(\frac{2}{\delta}\,\sinh^{-1}\frac{\delta}{2}\right)^{-1}\right] D^2 y_0 \\ &= 2h\left[1 + \frac{\delta^2}{6} - \frac{\delta^4}{180} + \frac{\delta^6}{1512} - \cdots\right] g_0 \end{aligned} \tag{7.8.9}$$

$$\begin{aligned} 2\mu\delta y_0 &= 2h\left[\mu\left(\frac{2}{\delta}\,\sinh^{-1}\frac{\delta}{2}\right)^{-1}\right] z_0 \\ &= 2h\left[1 + \frac{\delta^2}{6} - \frac{\delta^4}{180} + \frac{\delta^6}{1512} - \cdots\right] z_0\,. \end{aligned} \tag{7.8.10}$$

In the numerical solution of (7.8.1) these expansions are truncated, and a possible predictor-corrector computation proceeds as follows. Once y_{-1}, z_{-1}, y_0 and z_0^p are known, we predict y_1 by means of the expression

$$y_1^p = 2y_0 - y_{-1} + h^2 g(x_0, y_0, z_0^p) \tag{7.8.11}$$

which is obtained from (7.8.7). Next, we correct z_0 by means of the expression

$$hz_0 = \frac{1}{2}(y_1^p - y_{-1}) \tag{7.8.12}$$

obtainable from (7.8.8). The value of z_1 is now predicted from a truncated form of (7.8.9), namely

$$z_1^p = z_{-1} + 2hg(x_0, y_0, z_0)\,. \tag{7.8.13}$$

We can then correct y_1 by means of a truncated form of the expansion (7.8.10), namely

$$y_1 = y_{-1} + \frac{h}{3}(z_{-1} + 4z_0 + z_1^p)\,, \tag{7.8.14}$$

and we can next compute $g(x_1, y_1, z_1^p)$.

The predictor-corrector procedure of (7.8.11)–(7.8.14) eliminates nonlinear implicit schemes and successive approximations. Such an implicit method is described as follows.

Once y_{-1}, z_{-1}, y_0 and z_0 are known we can obtain more accurate values of y_1 and z_1 by using more terms in the expansions (7.8.7) and (7.8.9). For if we discard fourth and higher powers of δ in (7.8.7) and (7.8.9) we arrive at the system of two nonlinear equations

$$\begin{aligned} y_1 &= \frac{h^2}{12}\,g(x_1, y_1, z_1) + u \\ z_1 &= \frac{h}{3}\,g(x_1, y_1, z_1) + v \end{aligned} \tag{7.8.15}$$

where

$$\begin{aligned} u &= 2y_0 - y_1 + \frac{h^2}{12}(10g_0 + g_{-1}) \\ v &= z_{-1} + \frac{h}{3}(4g_0 + g_{-1}) \end{aligned} \tag{7.8.16}$$

and where $g_k = g(x_k, y_k, z_k)$, $k = -1, 0$. The system (7.8.15) is now solved by method of successive approximations:

$$\begin{bmatrix} y_1^{[0]} \\ z_1^{[0]} \end{bmatrix} = \begin{bmatrix} y_0 \\ z_0 \end{bmatrix}$$

$$\begin{bmatrix} y_1^{[v+1]} \\ z_1^{[v+1]} \end{bmatrix} = \begin{bmatrix} \frac{h^2}{12} g\left(x_1, x_1^{[v]}, z_1^{[v]}\right) + u \\ \frac{h}{3} g\left(x_1, y_1^{[v]}, z_1^{[v]}\right) + v \end{bmatrix} \quad v = 0, 1, 2, \ldots \,. \tag{7.8.17}$$

It can then be shown (Problem 7.8.4) that if $y_k = Y_k$, $z_k = Z_k$ for $k = -1, 0$, where Y is a solution of (7.8.1) and $Z = Y'$, then the solution $(y_1, z_1)^T$ of (7.8.15) satisfies the relation

$$\begin{bmatrix} y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} Y_1 + \frac{h^6}{240} Y^{(6)}(\zeta) \\ Z_1 + \frac{h^5}{90} Y^{(6)}(\eta) \end{bmatrix}, \quad \begin{matrix} x_{-1} < \zeta < x_1 \\ x_{-1} < \eta < x_1 \end{matrix} \,. \tag{7.8.18}$$

Let us next describe a procedure for solving the second order equation (7.8.2). It is possible to eliminate predictors in this case, since the equation is linear. Using the expansions (7.8.5) and (7.8.6) to replace y^2D^2 and hD respectively in Equation (7.8.2), we get the approximating equations

$$\begin{aligned} (1 \quad &+ hP_0/2)y_1 = (2 - h^2Q_0)y_0 - (1 - hP_0/2)y_{-1} + h^2R_0 \\ &+ \left[\left(\frac{\delta^4}{12} - \frac{\delta^6}{90} + \frac{\delta^8}{150} - \cdots\right) + hP_0\mu\left(\frac{\delta^3}{6} - \frac{\delta^5}{30} + \frac{\delta^7}{140} - \cdots\right)\right] y_0 \end{aligned} \tag{7.8.19}$$

where in using this equation, the $0(h^4)$ (as $h \to 0$) term in square brackets is dropped.

In the case when $P(x) \equiv 0$, in (7.8.2) we apply the operator $(1 + \delta^2/12)$ to each side of (7.8.19) to get the very accurate Numerov method

$$\begin{aligned} (1 + h^2Q_1/12)y_1 &= (2 - 5h^2Q_0/6)y_0 - (1 + h^2Q_{-1}/12)y_{-1} \\ &\quad + \frac{h^2}{12}(R_1 + 10R_0 + R_{-1}) \\ &\quad - \left[\frac{\delta^6}{240} - \frac{13\delta^8}{15120} + \cdots\right] y \,. \end{aligned} \tag{7.8.20}$$

The $0(h^6)$ (as $h \to 0$) term in square brackets is dropped when the method (7.8.20) is applied. We illustrate the accuracy of Numerov's method in Example ?? below.

Example 7.8.1 Series solution of $y'' + xy = 0$, $y(0) = 1$, $y'(0) = 0$. It is readily verified that this problem has the power series solution

$$y = \sum_{n=0}^{\infty} y_n x^{3n}$$

where

$$y_0 = 1\ ,\ y_{n+1} = -\frac{1}{(3n+2)(3n+3)}\, y_n\ ,\ n = 0, 1, 2, \dots\ .$$

The first five coefficients are $y_0 = 1$, $y_1 = -0.1666\ 6667$, $y_2 = 0.0055\ 5556$, $y_3 = -0.0000\ 7716$, $y_4 = 0.0000\ 0058$. These enable us to compute "exact" values of y given in Table 7.8.1.

Table 7.8.1: The Exact Solution of $y'' + xy = 0$, $y(0) = 1$, $y'(0) = 0$

x	y
0.0	1.00000000
0.1	0.99983334
0.2	0.99866702
0.3	0.99550405
0.4	0.98935607
0.5	0.97925332
0.6	0.96425842
0.7	0.94348383
0.8	0.91611271
0.9	0.88142272
1.0	0.83881231

Example 7.8.2 Solution of $y'' + xy = 0$, $y(0) = 1$, $y'(0) = 0$ by Numerov's method. Using Equation (7.8.20) the values y_k are computed by means of the recurrence relation

$$(1 + h^2 x_1/12)y_1 = (2 - 5h^2 x_0/6)y_0 - (1 + h^2 x_{-1}/12)y_{-1}$$

where $y(-0.1)$ is computed by means of the power series representation of Example 1. The numerical values of y_p are given in Table 7.8.2.

Problems

7.8.1 In the algorithm described in (7.8.10)–(7.8.13), let y_{-1}, z_{-1}, and x_0^p be the exact values of the solution Y and $Z = Y'$ of (7.8.1) at x_{-1} and x_0 respectively. Prove that if $g \in C^2$ then the values z_1^p and y_1 satisfy the relations $z_1^p = Z_1 + 0(h^3)$ and $y_1 = Y_1 + 0(h^3)$ as $h \to 0$.

Table 7.8.2: The Solution of $y'' + xy = 0$, $y(0) = 1$, $y'(0) = 0$ by Numerov's Method

x	y
-0.1	1.0001667
0.0	1.0000000
0.1	0.9998333
0.2	0.9986670
0.3	0.9955040
0.4	0.9853560
0.5	0.9792532
0.6	0.9642583
0.7	0.9434837
0.8	0.9161126

7.8.2 Let Y be the solution of (7.8.2) and let y_1 be determined from the (7.8.14), with the term in square brackets replaced by zero. Use the method of Problem 2.1 to prove that if P, Q, and R are of class C^2 then

$$y_1 - Y_1 = \frac{h^4}{12}\, Y^{(4)}(\zeta) - \frac{h^4}{6}\, P_0 Y^{(3)}(\eta)\ ,$$

where ζ and η are between x_{-1} and x_1.

7.8.3 In (7.8.2), let $P(x) = 0$, and let Y be a solution of this equation. Use the method of Problem 2.1 to show that if $y_{-1} = Y_{-1}$, $y_0 = Y_0$, and if y_1 is determined by means of (7.8.20) with the term in square brackets replaced by zero, then

$$y_1 - Y_1 = [1 + h^2 Q_1/12]^{-1} Y^{(6)}(\zeta) \quad , \quad \text{for some } \zeta \in (x_{-1} x_1)\,.$$

7.8.4 Prove the result of (7.8.18).

7.8.5 $*$ Nystrom developed a Runge-Kutta-type method that is especially suited to the problem

$$y'' = f(y)\ ,\ y(x_0) = y_0\ ,\ y'(x_0) = z(x_0) = y'_0\,.$$

Applied to this problem Nystrom's [21] method is

$$\begin{aligned} z_{k+1} &= z_k + \frac{h}{192}\left(23k_0 + 125k_1 - 81k_2 + 125k_3\right) \\ y_{k+1} &= y_k + hz_k + \frac{h^2}{192}\left(23k_0 + 75k_1 - 27k_2 + 25k_3\right). \end{aligned}$$

Here

$$\begin{aligned} k_0 &= f(y_k) \\ k_1 &= f\left(y_k + \tfrac{2}{5}\, hz_k + \tfrac{2}{25}\, h^2 k_0\right) \\ k_2 &= f\left(y_k + \tfrac{2}{3}\, hz_k + \tfrac{2}{9}\, h^2 k_0\right) \\ k_3 &= f\left(y_k + \tfrac{4}{5}\, hz_k + \tfrac{4}{25}\, h^2 (k_0 + k_1)\right). \end{aligned}$$

Show that

$$y_1 - Y(x_1) = 0(h^5) \quad \text{as} \quad h \to 0\,.$$

7.9 Boundary Value Problems

Let us consider the numerical solution of the boundary value problem

$$y'' = g(x, y, y')\ ,\ \ y(0) = y(1) = 0\,. \tag{7.9.1}$$

We remark that if a and b are finite, the problem

$$\frac{d^2 w}{dt^2} = g(t, w, w')\ ,\ \ w(a) = A\ ,\ \ w(b) = B \tag{7.9.2}$$

can be transformed into the problem (7.9.1) by means of the transformations

$$\begin{aligned} t &= a + (b-a)x \\ w &= y + A(1-x) + Bx\,. \end{aligned} \tag{7.9.3}$$

Not all problems of the type (7.9.1) have a solution, nor is it always possible to tell, in practice, whether or not a particular problem of the type (7.9.1) has a solution. Sometimes it is, however, possible to give sufficient conditions so that a class of problems of the type (7.9.1) have a solution. For example, if $g(x, y, y') = G(x, y)$ where $G_y(x, y) \geq 0$ for all $x \in [0, 1]$ and all real y, then the problem (7.9.1) has a unique solution. Let us illustrate the application of matrix methods to some linear boundary value problems.

Example 7.9.1 Matrix solution of the boundary problem $y'' + y + 1 = 0$, $y(0) = y(1) = 0$. We use the approximation $h^2 y'' = \delta^2 y$. The exact solution of the boundary value problem is $y = \{[1 - \cos(1)]/\sin(1)\} \sin x + \cos x - 1$. We take $h = 0.1$, and take account of symmetry about $x = 0.5$, to get the system of equations (see Section 5.5.2)

$$\begin{bmatrix} 1.99 & -1 & & & \\ -1 & 1.99 & -1 & & \\ & -1 & 1.99 & -1 & \\ & & -1 & 1.99 & -1 \\ & & & -2 & 1.99 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 0.01 \\ 0.01 \\ 0.01 \\ 0.01 \\ 0.01 \end{bmatrix}$$

which has the solution given in Table 7.9.1.

Table 7.9.1: The solution of $y'' + y + 1 = 0$, $y(0) = y(1) = 0$

Approximate	Exact
$y_1 = y_9 = 0.04958$	0.04954
$y_2 = y_8 = 0.08868$	0.08860
$y_3 = y_7 = 0.11688$	0.11678
$y_4 = y_6 = 0.13392$	0.13380
$y_5 = 0.13962$	0.13949

Example 7.9.2 Solution of $y''+y+1 = 0$, $y(0) = y(1) = 0$ by computing two solutions f_p and g_p and taking $y_p = \alpha f_p + (1 - \alpha) g_p$ (see Section 5.5.2).

We again take $h = 0.1$ and use the difference scheme

$$y_{p+1} = 1.99 y_p - y_{p-1} - 0.01 \quad , \quad p = 1, 2, 3, \ldots$$

as in the previous example. We compute f_p such that $f_0 = f_1 = 0$ and g_p such that $g_0 = 0$, $g_1 = 1$. After computing the two solutions, we compute α such that $y_{10} = \alpha f_{10} + (1 - \alpha) g_{10} = 0$. The results are contained in Table 7.9.2.

Table 7.9.2: Solution of $y'' + y + 1 = 0$, $y(0) = y(1) = 0$

$1 - \alpha = 0.04958$

x	f	g	y
0.0	0.00000	0.00000	0.00000
0.1	0.00000	1.00000	0.04958
0.2	−0.01000	1.98000	0.08868
0.3	−0.01990	2.93020	0.11688
0.4	−0.05950	3.84109	0.13392
0.5	−0.09850	4.70358	0.13962
0.6	−0.14652	5.50904	0.13392
0.7	−0.20308	6.24940	0.11688
0.8	−0.26760	6.91728	0.08868
0.9	−0.33945	7.50598	0.04959
1.0	−0.49791	8.00962	0.00000

Example 3: An eigenvalue problem, $y'' + \theta xy = 0$, $y(0) = y(1) = 0$.

(a) Exact solution by power series method. The solution of the initial value problem

$$y'' + xy = 0 \ , \quad y(0) = 0 \ , \quad y'(0) = 1$$

is

$$y(x) = x - \frac{1}{12}\, x^4 + \frac{1}{6 \cdot 12^2}\, x^7 - \cdots$$

at $x = 1$, we must have

$$y(1) = 1 - \frac{1}{12} + \frac{1}{6 \cdot 12^2} - \cdots \ .$$

The exact value of θ_1 to seven significant figures, is found using successive approximations, and is $\theta_1 = 18.95622$.

(b) Approximate solution via a system of algebraic equations is carried out by approximation $h^2 y_p''$ by $\delta^2 y_p$. This leads to the following system of homogeneous equations:

$$\begin{bmatrix} 2-\mu & -1 & & & \\ -1 & 2-2\mu & -1 & & \\ & -1 & 2-3\mu & -1 & \\ & & \cdots & \cdots & \\ & & & -1 & 2-n\mu \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{n-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

where $h = 1/n$, $\mu = \theta h^3$. Setting the determinant of this system equal to zero, we obtain the following results.

$$\begin{array}{lcl} n=2: & \mu - 2 = 0 & \theta_1 = 16 \\ n=3: & 2\mu^3 - 6\mu + 3 = 0 & \\ & \mu_1 = 0.633974 & \theta_1 = 17.11732 \\ & \mu_2 = 2.366025 & \theta_2 = 63.88268 \\ n=4: & 6\mu^3 - 22\mu^2 + 20\mu - 4 = 0 & \\ & \mu_1 = 0.27924075 & \theta_1 = 17.87141 \\ & \mu_2 = 1.0 & \theta_2 = 64.0 \\ & \mu_3 = 2.587426 & \theta_3 = 152.79526 \\ n=5: & 24\mu^4 - 100\mu^3 + 122\mu^2 - 50\mu + 5 = 0 & \\ & \mu_1 = 0.1460119 & \theta_1 = 18.2510\,49 \end{array}$$

Using Aitken's δ^2 process on the above values of θ_1, we get the improved approximation $\theta_1 = 18.92\,719$.

Problems

7.9.1 Let a and b be finite, and consider the second order differential equation problem

$$\begin{aligned} y'' + f(x)y' & +g(x)y = r(x) \ , \quad a \le x \le b \\ y(a) = A & \qquad\qquad\quad , \ y(b) = B \end{aligned}$$

(a) Replace h^2y'' by $\delta^2 y$, hy' by $\mu\delta y$, where $h = (b-a)/n$ to obtain a linear system with tridiagonal matrix U.

(b) Show that if f and g are uniformly bounded on $[a, b]$ and if $g(x) < 0$ on $[a, b]$ then the matrix U is non-singular for all h sufficiently small.

7.10 The Shooting Method of Solving Boundary Value Problems

We describe two variants of the shooting method. If we denote the solution of the initial value problem

$$y'' = g(x, y, y') \ , \ y(0) = 0 \ , \ y'(0) = \alpha \tag{7.10.1}$$

by $Y(x, \alpha)$ then we want to find α such that $Y(1) = Y(1, \alpha) = 0$. We may thus solve for $Y(1, \alpha)$ by the Runge-Kutta method and then use an iterative method of solving the nonlinear equation $Y(1, \alpha) = 0$. For Newton's method we also require $W(1) = W(1, \alpha)$ where $W(x, \alpha) = \frac{\partial y(x,\alpha)}{\partial \alpha}$, and where the function $W(x) = W(x, \alpha)$ is the solution of the initial value problem

$$\begin{aligned} W'' &= g_y(x, Y(x), Y'(x))W + g_{y'}(x, Y(x), Y'(x))W' \\ W(0) &= 0, W'(0) = 1 \, . \end{aligned} \tag{7.10.2}$$

Let us describe another variant of the shooting method. Suppose that we approximate the differential equation (7.10.1) by the difference equations

$$y_{k+1} = y_k + F(y_{k-1}) \tag{7.10.3}$$

where $y_0 = 0$, $y_1 = a$. We want to find a such that $y_n = y_n(a) = 0$. This can be done using Newton's method, provided that along with Equations (7.10.3) we also solve the equation

$$w_{k+1} = w_k + F'(y_{k-1})w_{k-1} \ ; \ w_0 = 0, w_1 = 1 \, . \tag{7.10.4}$$

Here F' denotes the derivative of F, and $w_k = dy_k/da$. Hence with $a = a_0$, we simultaneously compute y_k and w_k, $k = 2, 3, \ldots, n$, and then get an improved value of a using the Newton formula

$$a_1 = a_0 - \frac{y_n(a_0)}{w_n(a_0)} \, . \tag{7.10.5}$$

We then recompute y_n and w_n with a_0 replaced by a_1, and repeat the process until a desired accuracy is achieved.

The shooting method can also be used to compute eigenvalues. For example, we may wish to find a value of the parameter c such that the problem

$$y'' + g(x, c)y = 0 \ , \ y(0) = y(1) = 0 \tag{7.10.6}$$

has a non-trivial solution. Here we can approximate (7.10.6) by

$$y_{k+1} - 2y_k + y_{k-1} + h^2 g(x_k, c) y_k = 0 \tag{7.10.7}$$

where we take $y_0 = 0$, $y_1 = h$. Along with (7.10.7) we also solve

$$\begin{aligned} w_{k+1} & -2w_k + w_{k-1} + h^2 g_c(x_k, c) y_k + h^2 g(x_k, c) w_k = 0 \\ w_0 & = w_1 = 0 \end{aligned} \tag{7.10.8}$$

where $Y_k = dy_k/dc$. Thus starting with $c = c_0$, we get an improved value c_1 of c by means of the formula

$$c_1 = c_0 - \frac{y_n(c_0)}{w_n(c_0)}. \tag{7.10.9}$$

We then replace c_0 by c_1 and recompute y_n, w_n, and c_1, and so on, until a desired accuracy is achieved.

Example 7.10.1 The problem

$$\begin{aligned} \frac{d^2y}{dx^2} - \left(\frac{1}{U-c} + \frac{1}{U-c-c_r}\right) U' \frac{dy}{dx} - A^2 \left(\frac{U-c-c_r}{U-c}\right) y & = 0 \\ y(0) = y(1) & = 0 \end{aligned}$$

arises in weather forecasting. Here y is related to the vertical wind velocity, A and c_r are given constants and $U = U(x)$ is a given function. The problem is to determine the value of c such that a non-trivial solution exists. For the case of $U = 10x(1-x)^3$, $c_r = 2.14$, we replace y_p'' by $(y_{p+1} - 2y_p + y_{p-1})/h^2$, y_p' by $(y_{p+1} - y_{p-1})/(2h)$, take $h = 1/36$ and apply the above described Newton's method to solve for two values of c for $A = 0, 1, 2$. We thus find that

$$\begin{aligned} c & = 0.43030\,, \quad -1.56988 \quad \text{with} \quad A = 0 \\ c & = 0.36213\,, \quad -1.57459 \quad \text{with} \quad A = 1 \\ c & = 0.37474\,, \quad -1.61654 \quad \text{with} \quad A = 2\,. \end{aligned}$$

We can obtain an exact value of c in the case where $A = 0$, since the differential equation can be explicitly solved then. The exact values of c when $A = 0$ are $c = 0.42904$ and $c = -1.56904$.

Problems

7.10.1 Let the solution of Problem 7.9.1 be approximated by the solution of the difference equations (7.10.7), so that y_n is computed as a function of y_1. Thus we want to solve the nonlinear equation $(*)$ $y_n(y_1) = 0$. Construct an algorithm for solving the problem $(*)$ by combining the method of (7.10.3) with the secant method.

7.11 Solution of Linear Second Order Differential Equations using Splines

Let us subdivide the interval $[a, b]$ into n equal intervals $[x_{j-1}, x_j]$, $j = 1, 2, \ldots, n$, where $x_0 = a$, $x_j = x_0 + jh$, and where $x_n = x_0 + nh = b$. Let us interpolate a function $y(x)$ on the interval $[a, b]$ by $S(x)$ where $S(x)$ is a cubic polynomial in each of the intervals $[x_{j-1}, x_j]$, such that $S(x_j) = y_j$ and such that $S(x)$ and $S'(x)$ are continuous on $[a, b]$. $S(x)$ is then called a cubic spline. It may be readily verified that the solution is given on the interval $[x_{j-1}, x_j]$ by the equation

$$\begin{aligned} S(x) &= M_{j-1}\frac{(x_j - x)^3}{6h} + M_j\frac{(x - x_{j-1})^3}{6h} \\ &\quad + \left(y_{j-1} - \frac{h^2}{6}M_{j-1}\right)\frac{(x_j - x)}{h} \\ &\quad + \left(y_j - \frac{h^2}{6}M_j\right)\frac{(x - x_{j-1})}{h} \end{aligned} \tag{7.11.1}$$

where $M_j = S''(x_j)$ and $y_j = y(x_j)$. Hence for $j = 0, 1, \ldots, n-1$,

$$S'(x_j^+) = -\frac{h}{3}M_j - \frac{h}{6}M_{j+1} + \frac{y_{j+1} - y_j}{h} \tag{7.11.2}$$

while for $j = 1, 2, \ldots, n$,

$$S'(x_j^-) = \frac{h}{3}M_j + \frac{h}{6}M_{j-1} + \frac{y_j - y_{j-1}}{h}\,. \tag{7.11.3}$$

Continuity of $S'(x)$ on $[a, b]$ therefore implies that $S'(x_j^+) = S'(x_j^-)$, i.e. that

$$\frac{h}{6}M_{j-1} + \frac{2h}{3}M_j + \frac{h}{6}M_{j+1} = \frac{y_{j+1} - 2y_j + y_{j-1}}{h}\ , \quad j = 1, 2, \ldots, n-1\,. \tag{7.11.4}$$

Given the boundary value problem

$$\begin{aligned} y'' + P(x)y' + Q(x)y &= R(x)\ , \quad a \le x \le b \\ y(a) = A\ ,\ \ y(b) &= B\,, \end{aligned} \tag{7.11.5}$$

we can approximate the solution $y(x)$ using the above cubic spline by making the spline approximation satisy the differential equation (7.11.5) at the *knots*

x_j, $j = 0, 1, \ldots, n$. Using (7.11.2) and (7.11.3) this leads to a set of relations from which we can eliminate the unknowns $M_0, M_1, \ldots, M_n$. The result, in conjunction with the boundary conditions in (7.11.5), is a set of tridiagonal equations for the determination of $y_0, y_1, \ldots, y_n$.

For the case when $P(x) \equiv 0$ the differential equation yields

$$M_j = R_j - Q_j y_j \ , \quad j = 0, 1, \ldots, n \tag{7.11.6}$$

that when substituted into (7.11.4) yields

$$\begin{aligned} y_{j+1}\left(1 + \frac{h^2}{6}Q_{j+1}\right) & - y_j\left(2 - \tfrac{2h^2}{3}Q_j\right) + y_{j-1}\left(1 + \tfrac{h^2}{6}Q_{j-1}\right) \\ &= \tfrac{h^2}{6}\left(R_{j+1} + 4R_j + R_{j-1}\right) \ , \quad j = 1, 2, \ldots, n-1 \ . \end{aligned} \tag{7.11.7}$$

Using differences we may write this equation as

$$\delta^2 y_j + h^2\left(1 + \frac{1}{6}\delta^2\right) Q_j y_j = h^2\left(1 + \frac{1}{6}\delta^2\right) R_j \ . \tag{7.11.8}$$

Thus (7.11.7) has a local truncation error of order h^4 as $h \to 0$. The equation (7.11.7) is also similar to the Numerov formula (7.8.20) except that $\delta^2/6$ now replaces $\delta^2/12$ in (7.8.20).

Cubic splines possess a curious approximation property.

Theorem 7.11.1 *Let $f \in C^2[a,b]$, and let $S(x)$ by the cubic spline that satisfies*

$$\begin{aligned} S(x_i) &= f(x_i) \ , \quad i = 0, 1, \ldots, n \\ S'(a) &= f'(a) \ , \quad S'(b) = f'(b) \ . \end{aligned}$$

Then

$$\int_a^b [f''(x) - S''(x)]^2 dx \le \int_a^b [f''(x) - V''(x)]^2 dx \tag{7.11.9}$$

where V is any function that reduces to a polynomial of degree at most 3 on each interval $[x_{i-1}, x_i]$, $i = 1, 2, \ldots, n$.

Example: Solution of $y'' + y + 1 = 0$, $y(0) = y(1) = 0$ by use of cubic splines. For the sake of illustration, we divide the interval $[0, 1]$ into two subintervals, $[0, 0.5]$ and $[0.5, 1]$. Then using (7.11.7), we get

$$\begin{aligned} y(0.5) &= \tfrac{1}{2-\frac{2(0.5)^2}{3}} \tfrac{(0.5)^2}{6} (1 + 4 + 1) \\ &= \tfrac{3}{22} \simeq 0.1364 \,. \end{aligned}$$

Next, using (7.11.6), we get

$$M_0 = M_2 = -1 \;, \;\; M_1 = -\frac{47}{48} \,.$$

Substituting these results into (7.11.1), we get

$$S(x) = \frac{47}{48} x - \frac{1}{2} x^2 - \frac{1}{22} x^3 + \frac{u}{11} \left(x - \frac{1}{2} \right)^3 ,$$

where $u = 0$ for $x < 0.5$, $u = 1$ for $x \geq 0.5$.

We cannot expect much accuracy using such a large interval. For example, for $x = .3$, our solution yields $S(x) = S(.3) \simeq .2475$, whereas from Table 7.9.1, $y(.3) \simeq .1168$.

The case when the first derivative is present in (7.11.5), i.e. when $P(x) \not\equiv 0$, leads upon the elimination of the M_j, to a more complicated difference equation, namely

$$\begin{aligned} y_{j+1} \left(1 + \frac{h}{2} P_{j+1} + \frac{h^2}{6}\, Q_{j+1} \right) A_j - y_j \left[\left(1 + \frac{h}{2} P_{j+1} \right) A_j + \left(1 - \frac{h}{2} P_{j-1} \right) B_j - \frac{2h^2}{3} Q_j C_j \right] \\ + y_{j-1} \left(1 - \tfrac{h}{2} P_{j-1} + \tfrac{h^2}{6} Q_{j-1} \right) B_j \\ = \tfrac{h^2}{6} \left(A_j R_{j+1} + 4 C_j R_j + B_j R_{j-1} \right) \,, \;\; j = 1, 2, \ldots, n-1 \end{aligned} \tag{7.11.10}$$

where

$$\begin{aligned} A_j &= \left(1 - \tfrac{h}{3} P_{j-1} \right) \left(1 + \tfrac{h}{3} P_j \right) + \tfrac{h^2}{36} P_{j-1} P_j B_j \;\; = A_{j+1} \\ C_j &= 1 + \tfrac{7h}{24} \left(P_{j+1} - P_j \right) - \tfrac{h^2}{12} P_{j-1} P_{j+1} \,. \end{aligned}$$

Problems

7.11.1 Derive (7.11.1) by starting with the approximation

$$f''(x) \simeq S''(x) = M_{j-1}\frac{x - x_{j-1}}{h} + M_j\frac{(x_j - x)}{h}\,.$$

7.11.2 * Verify the relation (7.11.9).

7.12 The Bubnov-Galerkin Method for Solving Second Order Boundary Value Problems

Let us consider the problem

$$y'' + f(x, y, y') = 0\ ,\ \ y(0) = y(1) = 0\,. \tag{7.12.1}$$

We assume that the problem (7.12.1) has a solution $y_0(x)$, such that for some $\delta > 0$ there is no other solution $y(x)$ in the region $S_\delta(y_0) = \{y : \|y - y_0\| < \delta\}$. Here we adopt the notation

$$(u, v) = \int_0^1 u(x)v(x)dx \tag{7.12.2}$$

and $\|u\| = (u, u)^{1/2}$. Furthermore, we assume that f can be differentiated at least once with respect to each of the variables x, y, and y', for all y and y' such that $\|y - y_0\| + \|y' - y_0'\| < \delta$.

The Bubnov-Galerkin method of obtaining an approximate solution y_n of (7.12.1) is to set

$$y_n = \sum_{k=1}^{n} a_k u_k\ ,\ \ u_k(x) = 2^{1/2}\sin k\pi x\,, \tag{7.12.3}$$

and to determine the numbers a_k by solving the system of non-linear equations

$$(y_n'' + f(x, y_n, y_n'), u_k) = 0\ ,\ \ k = 1, 2, \ldots, n\,. \tag{7.12.4}$$

If we can solve this system for each n, and if $y_n \in S_\delta(y_0)$ for all n sufficiently large, it can be shown that $y_n \to y_0$ as $n \to \infty$. Moreover, it can then be shown that

$$\|y_n - y_0\| = 0\left(\frac{1}{n^2}\,\|y_n'' + f(x, y_n, y_n')\|\right) \tag{7.12.5}$$

as $n \to \infty$.

Sometimes it is convenient to obtain accurate approximate solutions using a different choice of a sequence $\{u_k\}$, such as $u_k(x) = x(1-x)P_k(x)$, where the polynomials $P_k(x)$ are orthogonal over the interval $[0, 1]$ with respect to the weight function $x^2(1-x)^2$.

Example 7.12.1 Let us use the Bubnov-Galerkin method to obtain an approximate solution of

$$y'' + y + \frac{1}{2}\, y^3 + \frac{1}{8} \sin \pi x = 0 \;,\;\; y(0) = y(1) = 0 \tag{7.12.6}$$

having the form

$$y_1 = a \sin(\pi x)\,.$$

Setting

$$\int_0^1 \left[y_1'' + y_1 + \frac{1}{2}\, y_1^3 + \frac{1}{8} \sin \pi x \right] \sin \pi x dx = 0$$

we arrive at the equation

$$\frac{1}{2}(\pi^2 - 1)a - \frac{3}{16}\, a^3 - \frac{1}{16} = 0$$

that has the approximate solution $a \simeq \frac{1}{8(\pi^2-1)}$, so that $y_1 = \frac{\sin \pi x}{8(\pi^2-1)}$. It can be shown that this approximation is accurate to five decimals. Indeed

$$\left\| y_1'' + y_1 + \frac{1}{2}\, y_1^3 + \frac{1}{8} \sin \pi x \right\| < 3 \times 10^{-6}\,.$$

Further discussions on the use of the Bubnov-Galerkin method for solving ordinary differential equations can be found in [4, 18, 30, 31]. With proper choice of the basis functions u_k, the rate of convergence of y_n to y_0 is usually very rapid. The chief practical difficulty of the method is the solution of the system nonlinear equations (7.12.3).

Problems

7.12.1 Let the Hilbert space H be defined by the class of all functions f such that

$$\|f\|^2 = \int_0^\infty |f(x)|^2 dx < \infty\,.$$

Use the Bubnov-Galerkin method to find an approximate solution of the problem

$$y'' = y - \frac{y^3}{x^2} \;,\;\; y(0) = y(\infty) = 0$$

in the form

$$y_1 = a(e^{-x} - e^{-2x})\,.$$

7.12.2 Set $u_1 = 30^{1/2} \times (1-x)$, $y_1 = au_1$ in (7.12.6) and determine the constant a such that (7.12.4) is satisfied with (u, v) defined in (7.12.2). How does the resulting solution compare with that obtained in the example of this section?

7.13 The Error of Numerical Methods for Solving Ordinary Differential Equations

Let us assume that the conditions of Theorem 7.1.10 are satisfied, namely, that $f(x, y)$ is continuous and satisfies a Lipschitz condition on a sufficiently large neighborhood of (x_0, y_0) that contains both the actual and the numerical solution of the problem

$$y' = f(x, y) \ , \ \ y(x_0) = y_0 \, . \tag{7.13.1}$$

We examine the error of Euler's method for this case [2].

Euler's method predicts an approximate solution y_p that satisfies the equation

$$y_{p+1} = y_p + hf(x_p, y_p) \quad , \quad p = 0, 1, 2, \ldots . \tag{7.13.2}$$

In practice we do not compute y_p, but rather z_p, where z_p satisfies the equation

$$z_{p+1} = z_p + hf(x_0, z_p) - R_p \tag{7.13.3}$$

where R_p is the roundoff error at the p'th step. The exact solution Y_p satisfies the equation

$$Y_{p+1} = Y_p + hf(x_p, Y_p) + T_p \tag{7.13.4}$$

where T_p is the truncation error of Euler's method, and is given by

$$T_p = \frac{h^2}{2}\, Y''(\zeta) \ , \ \ x_p < \zeta < x_{p+1} \, .$$

By taking the difference of (7.13.3) and (7.13.4), we get

$$Y_{p+1} - Y_{p+1} = Y_p - y_p + h[f(x_p, Y_p) - f(x_p, y_p)] + R_p + T_p \, . \tag{7.13.5}$$

Our assumed Lipschitz condition on f (see (7.1.31)) now tells us that

$$|f(x_p, y_p) - f(x_p, Y_p)| \le M|y_p - Y_p| \tag{7.13.6}$$

where M does not depend on x and y in the region of consideration. Hence setting $e_p = y_p - Y_p$, we get the inequality

$$|e_{p+1}| \leq |e_p|(1 + Mh) + K \;;\; e_0 = 0 \quad , \tag{7.13.7}$$

where we have assumed that $|T_p + R_p| \leq K$. Then

$$\begin{aligned}
|e_1| &\leq K \\
|e_2| &\leq K[1 + (1 + Mh)] \\
&\vdots \\
|e_p| &\leq K[1 + (1 + Mh) + (1 + Mh)^2 + \cdots + (1 + Mh)^{p-1}] \\
&\leq \frac{K[(1 + Mh)^p - 1]}{Mh} \leq \frac{e^{M(x_p - x_0)} - 1}{Mh}
\end{aligned} \tag{7.13.8}$$

since $1 + Mh \leq e^{Mh}$.

Notice that if we ignore roundoff error, then $K \leq Ch^2$, and $e_p = 0(h)$ as $h \to 0$. For moderate values of h, the roundoff error is considerably less than the truncation error. As h decreases, the total error decreases at first, and then becomes unbounded as $h \to 0$. A characteristic graph of the total error bound as a function of h is given in Figure 7.13.1.

We omit the analyses of more accurate methods, covered in the previous sections, since the analyses of these methods are similar to that of the Euler method.

Problems

7.13.1 For the case $y' = f(x, y)$, show that the local truncation error in the fourth order Runge-Kutta method is bounded by

$$MN(3.7 + 5.4M + 1.3M^2 + 0.017M^3)h^5$$

where $|f(x, y)| \leq N$, and

$$\left| \frac{f^{(i+j)}(x, y)}{\partial x^i \partial y^j} \right| \leq M/N \,.$$

7.13.2 Use the results of Problem 7.13.1 to carry out the error analysis of Section 6.13 for the case of the fourth order Runge-Kutta method of (7.6.10)–(7.6.11).

Figure 7.13.1: The Relationship Between Step Size and Error Bound

7.14 Stability of Linear Multistep Methods and Stiff Problems

The general k-step method for solving the problem

$$y' = f(x, y) \ , \ \ y(x_0) = y_0 \tag{7.14.1}$$

takes the form

$$y_{p+1} + \alpha_1 y_p + \cdots + \alpha_k y_{p-k+1} = h(\beta_0 f_{p+1} + \cdots + \beta_k f_{p-k+1}) \tag{7.14.2}$$

where $f_p = (fx_p, y_p)$, $k > 1$, the α_s and β_s are real constants, and $|\alpha_s|+|\beta_s| > 0$. In practice we do not solve the equation (7.14.2) for y_{p+1}, but rather the equation

$$Z_{p+1} + \alpha_1 Z_p + \cdots + \alpha_k Z_{p+1-k} = h(\beta_0 F_{p+1} + \cdots + \beta_k F_{p-k+1}) + R_p \tag{7.14.3}$$

where $F_p = f(x_p, Z_p)$, and where R_p denotes a roundoff error.

Roughly speaking, the equation (7.14.2) is called stable if the y_p differ only slightly from the Y_p, if all the R_p are small, where Y denotes the solution of (7.14.1). The general solution of the difference equation

$$z_{p+1} + \alpha_1 z_p + \cdots + \alpha_p z_{p+1-k} = 0 \tag{7.14.4}$$

takes the form

$$z_p = \sum_{i=1}^{q} \sum_{j=1}^{q_i-1} A_{ij} \zeta_i^p p^j \qquad (7.14.5)$$

where the A_{ij} are constants, and where the ζ_1 are the $q \leq k$ distinct roots of the polynomial equation

$$\zeta^k + \alpha_1 \zeta^{k-1} + \cdots + \alpha_k = 0\,. \qquad (7.14.6)$$

If $R_p \not\equiv 0$, each term Z_p in (7.14.3) contains unwanted extraneous terms on the right-hand side of (7.14.5). If some of the ζ_1 are larger than 1 in modulus, or else if $|\zeta_1| = 1$ and $q_1 > 1$ for some i then this unwanted solution grows and eventually destroys the accuracy of the desired solution.

The above line of reasoning suggests the following definition of stability of the k-step method (7.14.2).

Definition 7.14.1 *The multistep method (7.14.2) is called stable if and only if all roots ζ_1 of (7.14.6) are at most 1 in modulus, and all roots of modulus 1 are simple.*

The above definition can be related more precisely to one involving the growth of error of the method (7.14.2). Let $f(x, y)$ be a continuous function of x and y for all $x \in [x_0, a]$ and $y \in \left\{ y : \min_{x \in [x_0,a]} Y(x) - \alpha \leq y \leq \max_{[x_0,a]} Y(x) + \alpha \right\}$ where $Y(x)$ denotes the solution of (6.1), $\alpha > 0$, and let f satisfy the Lipschitz condition (1.28).

Definition 7.14.2 *We say that the perturbations R_p in (7.14.3) are of class $C(\delta)$ if, for some $h_0 > 0$,*

$$\sum_{p=0}^{k-1} |Z_p - y_p| + \sum_{j=k-1}^{N-1} |R_p| \leq \delta \qquad (7.14.7)$$

for all $h \leq h_0$ and all N with $x_0 + Nh < a$.

For example, perturbations such that

$$|Z_p - y_p| \leq \delta/2k \;, \; |R_p| \leq \delta h/[2(a - x_0)]$$

are of class $C(\delta)$.

Definition 7.14.3 *The multistep method (7.14.2) is called stable with respect to $f(x, y)$ if for any $\epsilon > 0$ there exists a $\delta(\epsilon) > 0$ such that the errors produced by perturbations of class $C(\delta)$ satisfy $|Z_p - y_p| \leq \epsilon (p = k, k+1, \ldots, N)$, uniformly for all $h \leq h_0$ and all N with $x_0 + Nh < a$.*

Theorem 7.14.4 *[39]: The multistep method (7.14.2) is stable in the sense of Definition 7.14.3 if and only if it is stable in the sense of Definition 7.14.1.*

We omit the proof of this theorem and refer the reader to [40, 6].

Definition 7.14.5 *The multistep method (7.14.2) is called consistent if*

$$1 + \alpha_1 + \alpha_2 + \cdots + \alpha_k = 0\,,$$

i.e., if $\zeta = 1$ is a solution of (7.14.6).

It may be shown [42] that every stable and consistent method of order greater than 1 is convergent.

We remark that by making the precision of computation sufficiently large the roundoff error of any k-step method can be made as small as we please, and hence the extraneous error can be made tolerably small. However, if the interval of integration is large, this usually demands an h so small and a precision so high that practical computation becomes unfeasible. Consider for example the following system of stiff equations:

$$\begin{bmatrix} f' \\ g' \end{bmatrix} = \begin{bmatrix} -11 & \frac{9}{12} \\ 18 & -11 \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix}. \tag{7.14.8}$$

Such a system is called stiff because the eigenvalues θ_j of the coefficient matrix have negative real parts, such that the ratio of the largest modulus of the θ_j to the smallest modulus of the θ_j is large, e.g. greater than 10.

The eigenvalues of the coefficient matrix in (7.14.8) are -2 and -20. The solution with initial values (f_0, g_0) is

$$\begin{bmatrix} f(t) \\ g(t) \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\left(f_0 + \frac{1}{2}\,g_0\right) & \frac{1}{2}\left(f_0 - \frac{1}{2}\,g_0\right) \\ f_0 + \frac{1}{2}\,g_0 & -\left(f_0 - \frac{1}{2}\,g_0\right) \end{bmatrix} \begin{bmatrix} e^{-2t} \\ e^{-20t} \end{bmatrix}. \tag{7.14.9}$$

If we supply Euler's method to (7.14.8) we get the recurrence relation

$$\begin{bmatrix} f_{p+1} \\ g_{p+1} \end{bmatrix} = (I + Ah) \begin{bmatrix} f_p \\ g_p \end{bmatrix} \tag{7.14.10}$$

where I is the unit matrix of order 2, and A denotes the coefficient matrix of the system in (7.14.8). Thus

$$\begin{bmatrix} f_p \\ g_p \end{bmatrix} = (I+hA)^p \begin{bmatrix} f_0 \\ g_0 \end{bmatrix} = B^{-1} \begin{bmatrix} 1-2h & \\ & 1-20h \end{bmatrix}^p B \begin{bmatrix} f_0 \\ g_0 \end{bmatrix} \quad (7.14.11)$$

where

$$B = \begin{bmatrix} 1 & \frac{1}{2} \\ -1 & \frac{1}{2} \end{bmatrix} B^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ 1 & 1 \end{bmatrix}. \quad (7.14.12)$$

For example if $f_0 = 3$, $g_0 = 2$, then

$$f(t) = 2e^{-2t} + e^{-20t}\ ,\ \ g(t) = 4e^{-2t} - 2e^{-20t} \quad (7.14.13)$$

and

$$f_p = 2(1-2h)^p + (1-20h)^p\ ,\ \ g_p = 4(1-2h)^p - 2(1-20h)^p\,. \quad (7.14.14)$$

The e^{-20t} component decreases very rapidly relative to the e^{-2t} component, and hence in order to obtain the same behavior in the numerical solution, the step size h must be less than 1/20. However, after a few steps $(1-20h)^p$ is so small relative to $(1-2h)^p$ that it is comparable to the round-off error. At this stage we would like to be able to increase h, in order to be able to efficiently compute a solution over a larger interval. Note, however, that we cannot take $h > 0.1$, since then the $(1-20h)^p$ component would increase in magnitude. Hence we are forced to keep h small. Clearly this is an undesirable feature of the Euler method.

We are thus led to introduce the following notion of A-stability of a k-step method:

Definition 7.14.6 *The k-step method (7.14.2) is called A-stable when applied to (7.14.1), if every solution y_p of (7.14.2) converges to zero as $p \to \infty$ for all $h > 0$. The k-step method (7.14.2) is called explicit if $\beta_0 = 0$. It can be shown [32] that an explicit k-step method cannot be A-stable.*

For example, the Euler method is an explicit two step method, and hence is not A-stable. The trapezoidal method is A-stable. When applied to the problem

$$y' = -y$$

the trapezoidal method

$$y_{p+1} = y_p + \frac{h}{2}\left[f_{p+1} + f_p\right]$$

becomes

$$y_{p+1} = \left(\frac{1 - h/2}{1 + h/2}\right) y_p \, .$$

Clearly the modulus of $(1 - h/2)/(1 + h/2)$ is less than one for all $h > 0$ and hence $y_p \to 0$ as $p \to \infty$.

It can be shown [22] that the order k of an A-stable consistent k-step method cannot be greater than 2.

Further Reading

Ordinary Differential Equations

J.F. Botha and G.F. Pinder. *Fundamental Concepts in the Numerical Solution of Differential Equations*. John Wiley and Sons, New York, 1987.

W.E. Boyce and R. C. DiPrima. *Elementary Differential Equations and Boundary Value Problems*, John Wiley and Sons, New York, 1965.

G.W. Gear. *Numerical Initial Value Problems in Ordinary Differential Equations*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1971.

H.B. Keller. *Numerical Methods for Two-Point Boundary-Value Problems*. Blaisdell Publishing Co., Waltham, Massachusetts, ?year?.

E.L. Ince. *Ordinary Differential Equations*. Dover Publications, Inc., New York, 1956.

J.D. Lambert. *Computational Methods in Ordinary Differential Equations*. John Wiley and Sons, New York, 1973.

I. Stakgold. *Boundary Value Problems of Mathematical Physics*, volume I. MacMillan Co., New York, 1967.

I. Stakgold. *Boundary Value Problems of Mathematical Physics*, volume II. MacMillan Co., New York, 1967.

Partial Differential Equations

W.F. Ames. *Numerical Methods for Partial Differential Equations.* Academic Press, New York, 2nd edition, 1977.

D.A. Anderson, J.C. Tannenhill, and R.H. Pletcher. *Computational Fluid Mechanics and Heat Transfer.* McGraw-Hill Book Company, Wasington, 1984.

G. Birkhoff and R.E. Lynch. *Numerical Solution of Elliptic Problems.* SIAM, Philadelphia, 1984.

C. Canuto, M.Y. Hussaini, A. Quarteroni, and T.A. Zang. *Spectral Methods in Fluid Dynamics.* Springer-Verlag, New York, 1987.

D. Gottlieb and S.Z. Orszag. *Numerical Analysis of Spectral Methods: Theory and Applications.* SIAM, Philadelphia, 1977.

L. Lapidus and G.F. Pinder. *Numerical Solution of Partial Differential Equations in Science and Engineering.* John Wiley and Sons, New York, 1982.

R. Peyret and T.D. Taylor. *Computational Methods for Fluid Flow.* Springer-Verlag, New York, 1983.

G. D. Smith. *Numerical Solution of Partial Differential Equations: Finite Difference Methods.* Clarendon Press, Oxford, 2nd edition, 1978.

E.C. Zachmanoglou and D.W. Thoe. *Introduction to Partial Differential Equations.* Dover Publications, Inc., New York, 1986.