FRANKLIN, CAROL MARLEY. Second Order Predictor-Corrector Pairs. (1976)
Directed by: Dr. Michael Willett. Pp. 53.

The purpose of this study is to introduce multistep methods for approximating the solutions of ordinary differential equations and to study the general convergent, second order, predictor-corrector pair. Given the values of a solution function, either exact or approximate, at certain points on the x-axis, a multistep method approximates the values of that function at other points on the x-axis. Necessary and sufficient conditions for the convergence of a multistep method are the conditions of consistency and stability. A multistep method which approximates the value of the solution function y at x_N with no knowledge of the value of the derivative y' at x_N is called a predictor. A corrector is a multistep method which uses the value of the derivative of y at x_N to approximate y at x_N. The predictor is used to generate a rough estimate of the solution function at a point x_N then the corrector is used to correct the predicted value at x_N. Used together they constitute a predictor-corrector pair. If the differential equation to be solved is of the form $y' = \lambda y$, the predictor can be substituted in the corrector to yield a linear recursion. Constraints on the roots of the recursion which would ensure convergence to the true solution are discussed along with their applicability to equations of the form $y' = f(x,y)$. Experiments provided guidelines for choosing coefficients to give fast convergence.

SECOND ORDER PREDICTOR-CORRECTOR PAIRS

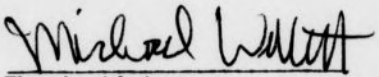
by

Carol Marley Franklin

A Thesis Submitted to
the Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Greensboro
1976

Approved by



Thesis Adviser

APPROVAL SHEET

This thesis has been approved by the following committee of the Faculty of the Graduate School at the University of North Carolina at Greensboro.

Thesis Adviser Richard Willett

Committee Members David G. Kerr

Hughes B. Hoyle, III

14 May 76

Date of Acceptance by Committee

ACKNOWLEDGEMENTS

The writer wishes to express her deep appreciation to Dr. Michael Willett for his unfailing patience, assistance, and understanding throughout this study.

Appreciation is extended to Dr. David Herr and Dr. Hughes Hoyle for their helpful suggestions and cooperation as members of the examining committee.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter I

Introduction

Differential equations are functional equations that involve derivatives. For example,

$$\frac{dx}{dt} - 2x = 0 \tag{1.1}$$

is a differential equation. Equation 1.1 states that the derivative of a function x is equal to twice the function. The equation therefore provides a partial description of the quantity x. To solve the differential equation means to determine a function $x(t)$ defined on some interval which satisfies the equation.

Differential equations arise in almost every branch of science and engineering. They are frequently encountered in physics, chemistry, and electrical engineering. A differential equation describes the flow of current through a wire; another differential equation describes the rate of a chemical reaction. Differential equations also describe the motions of heavenly bodies.

A differential equation may have many solutions, one solution, or no solution. Chapter II sets forth several conditions sufficient for the existence of a unique solution. In Chapter II we shall also illustrate various methods for solving differential equations.

Frequently a solution to a differential equation is known to exist, but it cannot be found by known analytical techniques. In these cases the solution may need to be approximated by numerical techniques. Chapter III introduces a class of approximation techniques called multistep methods.

Chapter IV defines a particular type of multistep method known as a predictor-corrector pair. In Chapter V we study in detail through both theoretical and empirical means the second order predictor-corrector method.

Chapter VI is a summary of the major ideas presented in this study.

Chapter II

The Initial Value Problem

Frequently, the solution of mathematically formulated problems in science and engineering is a function which satisfies an equation containing derivatives of that function. Such an equation is called a differential equation. In this chapter we shall introduce the concept of a differential equation and illustrate several analytical solution techniques.

Because there are so many types of differential equations, it is helpful to divide them into two broad categories. If the function satisfying the differential equation depends on one independent variable, the derivatives in the equation will be ordinary derivatives and the equation will be called an ordinary differential equation. If, on the other hand, the unknown function depends on several independent variables, the derivatives appearing in the differential equation may be partial derivatives. In this case the equation is called a partial differential equation.

Consider the differential equation

$$y' = y + 1$$

where y' denotes the first derivative of y with respect to the independent variable x. A solution to the equation is the function $y = e^x - 1$ because $y' = e^x$ and $y' = (e^x - 1) + 1 = y + 1$. Another solution to the equation is the function $y = 2e^x - 1$. In fact any

function of the form $y = \lambda e^x - 1$ where λ is a constant is a solution to the above equation. Consider the differential equation $(y')^2 + y^2 = 0$. The only solution to this differential equation is $y(x) = 0$. Finally, consider the differential equation $|\sqrt{y'}| + 2 = 0$. Clearly this equation has no solution. Thus, ordinary differential equations may have an infinite number of solutions, a unique solution, or no solution.

By the order of a differential equation is meant the largest integer n such that an n^{th} derivative appears in the equation. Because the highest order derivative in the sample equation is a first derivative, the equation is a first order differential equation.

The graphs of the solutions $y = \lambda e^x - 1$ to the sample differential equation represent a one-parameter family of curves. No two curves in the family intersect. To select a particular solution from the infinite set of solutions is equivalent to specifying one of the curves from the one-parameter family of curves. To do this one must merely specify a point through which the desired curve passes. That is, of all the solution functions, one may seek the function $y = f(x)$ which satisfies the following condition.

$$y(x_o) = y_o$$

Such a condition is called an initial condition. The initial condition and an ordinary differential equation constitute an initial value problem. Hereafter we shall call a first order initial value problem an IVP.

Consider the following IVP.

$$y' = y + 1 \qquad y(o) = 1 \qquad (') = (\frac{d}{dx}) \qquad (2.1)$$

The solution we would be seeking is a curve in the xy-plane passing through $(0,1)$ and satisfying the equation in (2.1).

More generally, for an n^{th} order ordinary differential equation, n initial conditions are usually needed to specify a particular solution, if any solution exists.

Because an initial value problem may not have a solution, it would be useful to be able to determine beforehand if a solution to the IVP exists and if such a solution is unique.

The following theorem gives fairly general and easily checked sufficient conditions for the existence of a unique solution to a first order differential equation.

Theorem 2.1 (Existence and Uniqueness Theorem): Consider the IVP

(i) $y' = f(x,y) \quad ['] \quad = \quad [\frac{d}{dx}]$

(ii) $y(a) = y_0$.

Let D be an open connected subset of the xy-plane containing the point (a,y_o). If the function $f : R^2 \to R$ satisfies

(a) f continuous on D

(b) there exists an L such that

$$|f(x,y_1) - f(x,y_2)| \leq L|y_1 - y_2| \quad \text{for all}$$

$$(x,y_1), (x,y_2) \in D,$$

then the IVP (i) (ii) has a unique solution on a closed interval $I = [a-h, a+h]$, for some $h > o$. Furthermore this solution is a continuous, differentiable function on I.

Proof: We outline the basic steps of the proof. Since (a, y_o) is a point of D, there is a closed rectangle R with center (a,y_o), length 2c, and width 2d such that R is wholly contained in D (see Figure 1).

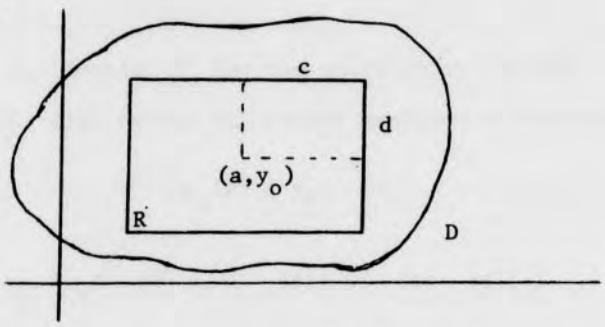


Figure 1. Closed rectangle R in D.

$$R = \{ (x,y) \mid |x-a| \leq c, |y-y_o| \leq d\} \subseteq D.$$

Since $f(x,y)$ is continuous on R and R is closed, then f is bounded on R; that is, there exists a number $M > 0$ such that $|f(x,y)| < M$ on R.

Consider the two lines through the point (a,y_o).

$$y - y_o = M(x - a),$$

$$y - y_o = -M(x - a).$$

Let

$h = \min \{\frac{d}{M}, c\}$. In outline form we shall show that the interval $[a-h, a+h]$ (or $|x-a| \leq h$) and the shaded region

of Figure 2 (which we label T) provide a domain for the solution to the IVP.

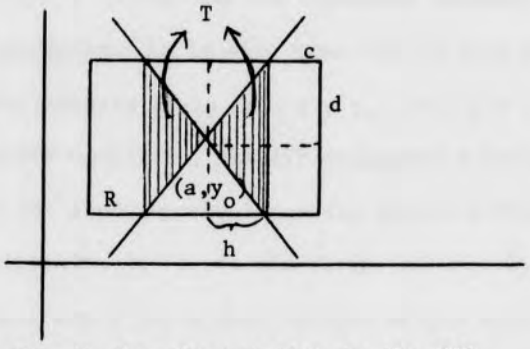


Figure 2. Domain T for the solution to the IVP.

For $x\in[a-h, a+h]$ define the Picard sequence of functions.

$$\phi_o(x) \equiv y_o$$

$$\phi_k(x) \equiv y_o + \int_a^x f\{ s, \phi_{k-1}(s)\}\, ds \quad \text{for} \quad k=1, 2, \ldots$$

(1) From induction and the fact that f is continuous on T we have that the function $\phi_n(x)$ exists and is continuous for all n. Since $|\phi_n(x) - y_o| \le d$ on $|x-a| \le h$, $\phi_n(x)$ has a continuous derivative.

(2) $$\left|\phi_n(x) - \phi_{n-1}(x)\right| \le \frac{M(Kh)^n}{K\, n!} \quad \text{on} \quad |x-a| \le h.$$

(3) Applying the Weierstrass M-test, $\{\phi_n\}$ converges uniformly on $|x-a| \le h$, say to a function $y(x)$.

(4) $y(x)$ satisfies $\frac{dy(x)}{dx} = f(x,y(x))$ on $|x-a| \le h$ and $y(a) = y_o$.

(5) $y(x)$ is the only function satisfying $\frac{dy(x)}{dx} = f(x, y(x))$, $y(a) = y_o$ on $|x-a| \le h$.

Q.E.D.

The constant L is known as a Lipschitz constant and the condition on the function f involving the Lipschitz constant is called the Lipschitz condition. It is also true that if the function f is continuous on the infinite strip $a \leq x \leq b$, $-\infty < y < \infty$ and if it satisfies the Lipschitz condition, the IVP possesses a unique solution on $[a,b]$. In our study we shall require that the function f be continuous and Lipschitz in y on the strip $a \leq x \leq b$, $-\infty < y < \infty$. Hence, we shall apply the weaker version of the existence and uniqueness theorem to ensure a unique solution for the IVP on the interval $[a,b]$.

Many ordinary differential equations can be solved explicitly. Most linear differential equations fall into this class. Also in this class are some nonlinear equations which are separable, exact, or homogeneous. Sometimes the method of integrating factors yields explicit solutions in other nonlinear differential equations. Each of these methods for solving ordinary differential equations is discussed below. Boyce [1] gives a detailed treatment of each of the following methods.

The general first order linear ordinary differential equation can be written

$$y' + p(x)y = g(x)$$

where p and g are continuous in x. Consider the IVP consisting of the above differential equation and the initial condition $y(a) = y_0$. Also consider the function μ defined by $\mu(x) = \exp[\int_a^x p(t)dt]$. The function μ is called an integrating factor because when the differential equation is multiplied by μ, the left side becomes the

derivative of $\mu(x)y$. A method for finding the solution proceeds as follows.

$$y' + p(x)y = g(x)$$

$$\mu(x)[y' + p(x)y] = \mu(x)g(x)$$

$$[\mu(x)y]' = \mu(x)g(x)$$

$$\mu(x)y = \int_a^x \mu(s)g(s)ds + y_o$$

or

$$y(x) = \frac{1}{\mu(x)} \left[\int_a^x \mu(s)g(s)ds + y_o \right] .$$

If the nonlinear equation $y' = f(x,y)$ can be put in the form $M(x) + N(y)\, y' = 0$ where M is a function of x only and N is a function of y only, the equation is said to be <u>separable</u>. An implicit solution to the equation is easily formed by rewriting the equation in the form

$$M(x) = -N(y)y'$$

and integrating both sides with respect to x.

$$\int_a^x M(s)ds = \int_a^x -N(y(s))y'(s)ds$$

Making the change of variables $z = y(s)$ we have

$$\int_a^x M(s)ds = \int_{y(a)}^{y(x)} -N(z)dz$$

If the resultant equation can be solved for y, an explicit solution is obtained. Consider the IVP

$$\frac{dy}{dx} = \frac{3x^2 + 4x + 2}{2(y-1)} , \qquad y(o) = -1. \tag{2.2}$$

Equation 2.2 can be rewritten as

$$2(y-1)y' = (3x^2 + 4x + 2).$$

Therefore, the implicit solution is given by

$$y^2-2y = x^3 + 2x^2 + 2x + 3.$$

Solving explicitly for y we get

$$y = 1- \sqrt{x^3 + 2x^2 + 2x + 4} \qquad \text{(checks in 2.2).}$$

Assume we are able to write the nonlinear equation $y' = f(x,y)$ in the form

$$M(x,y) + N(x,y)y' = 0, \ \frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} \tag{2.3}$$

where the operations on M and N denote partial differentiation. Equation 2.3 is referred to as an <u>exact</u> differential equation for the following reason. Assume that we have found a function $F(x,y)$ so that

$$\frac{\partial F}{\partial x} = M(x,y) \qquad \frac{\partial F}{\partial y} = N(x,y) \ .$$

Then the left side of the differential equation in (2.3) is the total derivative of F with respect to x so that the solution $y(x)$ satisfies the implicit relationship

$$F(x,y) = c$$

for some constant c. If there is no differentiable solution $y(x)$ to this implicit relationship then the original differential equation has

no solution. The second condition in (2.3) guarantees that

$$\frac{\partial}{\partial y}\left[\frac{\partial F}{\partial x}\right] = \frac{\partial}{\partial x}\left[\frac{\partial F}{\partial y}\right]$$

so that the order of differentiation (and conversely, of integration) is unimportant.

Suppose the given IVP is

$$(y \cos x + 2xe^y) + (\sin x + x^2e^y + 2)y' = 0,$$
$$y(o) = 1.$$

$M_y(x,y) = N_x(x,y)$ where the subscripts denote partial differentiation. Because $M(x,y)$ represents $F_x(x,y)$, we integrate $M(x,y)$ with respect to x to obtain an expression for F:

$$F(x,y) = y \sin x + x^2e^y + h(y)$$

where h is a function of y. The derivative of this expression for $F(x,y)$ with respect to y must equal $N(x,y)$.

$$F_y(x,y) = \sin x + x^2e^y + h'(y)$$
$$= \sin x + x^2e^y + 2$$

This implies that $h'(y) = 2$ or $h(y) = 2y + c$. Therefore, the desired function is $F(x,y) = y \sin x + x^2e^y + 2y + c$. Substituting the initial condition in the equation $F(x,y) = o$, we find that $c = 2$. The solution $y(x)$ therefore satisfies $y \sin x + x^2e^y + 2y = 2$ and the solution of the differential equation reduces to solving this implicit relationship.

$$y' = \frac{dy}{dx} = x\frac{dv}{dx} + v = xv' + v$$

Substitute this expression for $\frac{dy}{dx}$ into the above equation.

$$x\frac{dv}{dx} + v = F(v) \qquad\qquad xv' + v = F(v)$$

Thus the homogeneous equation can always be transformed into a separable equation and solved by the method described above for separable equations. The solution to the original equation is then obtained by replacing v with y/x.

The techniques above are aimed at finding an expression for the solution $y(x)$ if the solution exists. Because some of the algebraic steps may not be reversible it is possible to apply the techniques and to to find an apparent solution when in fact no solution exists. Therefore, the derived solution should be substituted in $y' = f(x,y)$ to see if it is in fact a solution.

The above techniques illustrate a few of the analytical methods designed to produce an explicit expression for $y(x)$. But many differential equations which occur in practice cannot be handled by known analytical techniques. In these cases it is necessary to develop methods designed to approximate the solution. One type of numerical approximation is a multistep method which we shall discuss in the next chapter.

Chapter III

Multistep Methods

In this chapter we shall be concerned with finding a solution to the initial value problem

$$y'(x) = f(x,y),\ y(x_o) = y_o \tag{3.1}$$

where $x\epsilon[x_o,b]$, $-\infty<x_o<b<\infty$ and where $f(x,y)$ is a member of the class of functions Lip $[x_o,b]$. For any interval $[x_o,b]$ the class of functions Lip$[x_o,b]$ is defined by $f\epsilon$ Lip $[x_o,b]$ if and only if

(i) $f:\ R^2 \to R$

(ii) f is continuous on $[x_o,b] \times R$

(iii) there exists a constant L so that

$$|f(x,y_1) - f(x,y_2)| \leq L\,|y_1 - y_2|$$

uniformly in $x\epsilon[x_o,b]$.

Some IVP's can be solved in closed form; that is, as finite combinations of elementary functions, exponential functions, logarithms, or indefinite integrals. More often, however, such a closed form solution cannot be found by known techniques. Although in this case a closed form solution $y(x)$ over a continuous range of the independent variable is unavailable, there are numerical approximation techniques which can approximate the solution at a set of discrete points $x_o, x_1, \cdots, x_n$. Such techniques, called discrete variable methods, are generally

algorithms which yield for each point x_n a number y_n which is an approximation to the exact solution $y(x_n)$ at x_n. Because discrete variable methods can be applied to almost any differential equation, they are widely used when analytical solutions cannot be found.

Discrete variable methods which are used for solving initial value problems and which require k preceding values $y_{n+k-1}, y_{n+k-2}, \ldots, y_n$ to calculate y_{n+k} are called multistep methods. In this chapter we shall discuss multistep methods and the criteria for "good" multistep methods. We shall also introduce the general linear recursion.

To approximate the solution to the initial value problem in equation 3.1 at a point b by using a multistep method, the method is applied to the interval $[x_o,b]$. The interval is divided into N subdivisions each of width h. The width h is called the stepsize of the multistep method and it is calculated by $h = \frac{b-x_o}{N}$. The endpoints of the subdivisions, or nodes, form a set of discrete points, namely $\{x_i \mid x_i = x_o + ih,\ i=0, 1, \ldots, N\}$ on which the multistep method acts.

Suppose a multistep method requiring two preceding values is chosen. Then two values are needed to start using the multistep method. Suppose that we somehow obtain a good approximation y_1 to $y(x_1)$. Then using the true value $y_o = y(x_o)$ we could use the multistep method to generate y_2, an approximation to $y(x_2)$. The method continues to use the two previous values until the interval is traversed and an approximation at $b = x_N$ is obtained. We now turn our attention to precisely how the successive approximations are generated.

We define the general linear k^{th} order multistep method by the difference equation

$$y_{n+k} = \alpha_1 y_{n+k-1} + \alpha_2 y_{n+k-2} + \ldots + \alpha_k y_n \tag{3.2}$$

$$+ \; h \left[\beta_o y'_{n+k} + \beta_1 y'_{n+k-1} + \ldots + \beta_k y'_n \right] \quad n = 0, 1, 2, \ldots$$

where k is a fixed integer, $y'_m = f(x_m, y_m)$ $(m = 0, 1, 2, \ldots)$, α_i and $\beta_i (i = 0, 1, \ldots, k)$ denote real constants which do not depend on n, and h is the stepsize. We shall assume that $|\alpha_k| + |\beta_k| > 0$; otherwise, the k^{th} order method could also be written as a method of order $k-1$. The term linear here means that the expression for y_{n+k} is a linear combination of the y_i and y'_i.

As is evident from the definition, the <u>order</u> of a multistep method refers to the number of preceding values of y needed to calculate y_{n+k}. Thus a second order method, written as

$$y_{n+2} = \alpha_1 y_{n+1} + \alpha_2 y_n + h \left[\beta_o y'_{n+2} + \beta_1 y'_{n+1} + \beta_2 y'_n \right] ,$$

uses the two values y_{n+1} and y_n to calculate y_{n+2}. Obtaining values to begin the multistep method will be discussed below.

It is natural to require that the initial k values of y_i used to start the iteration converge to y_o as the stepsize gets small; that is, $y_i = y_i(h)$ satisfies $\lim_{h \to o} y_i(h) = y_o$, $i = 0, 1, \ldots, k-1$. Such a choice is termed a <u>consistent choice of initial values</u>. This condition requires that as the initial x_i get close to x_o, the initial choices for the y_i must get close to y_o. This is consistent with the fact that the unknown solution $y(x)$ is continuous on

$[x_o,b]$. The exact starting values are certainly a consistent choice; however, in practice exact values are unattainable since this would require knowing a closed form solution for the differential equation. Although there are several means of obtaining a consistent set of initial values, the simplest method involves using the truncated Taylor series.

In its simplest form, convergence of a multistep method means that if successive approximations to the true solution at b are made with h being decreased for each approximation, the sequence of numbers formed from these approximations at b will not only converge, but they will converge to the true solution $y(b)$ if a consistent choice of initial values is used. In the definition of convergence we would like to incorporate more than the behavior of the method at the endpoint b. Choose $f\epsilon Lip[x_o,b]$. In this case there exists a unique solution to the initial value problem $y' = f(x,y)$, $y(x_o) = y_o$ on $[x_o,b]$. For each $x\epsilon[x_o,b]$ we can construct a sequence of approximations to $y(x)$. Divide the interval $[x_o,x]$ into N equal subdivisions. Use the method to get an approximation y_N at x. Since y_N depends on x, denote it by $y_N(x)$. A multistep method is said to converge to the solution of the IVP at $x\epsilon[x_o,b]$ if $\lim_{N\to\infty} y_N(x) = y(x)$. The method is said to converge to the solution of the IVP on $[x_o,b]$ if it converges to the solution of the IVP for every $x\epsilon[x_o,b]$. A formal definition of convergence follows.

Definition 3.1. A multistep method M is called convergent on $[x_o,b]$ if M converges to the solution of any IVP

(i) $y' = f(x,y)$ (ii) $y(x_o) = y_o$

for $f \in Lip[x_o,b]$ and $y_o \in R$ on $[x_o,b]$ and if a consistent choice of initial values is used.

At this point we shall consider some restrictions on the constant coefficients of equation 3.2 which will ensure convergence of the multi-step method. For convenience we associate with equation 3.2 the polynomials $F(x)$ and $g(x)$ defined by

$$F(x) = x^k - \alpha_1 x^{k-1} - \alpha_2 x^{k-2} - \ldots - \alpha_k \tag{3.3}$$

$$g(x) = \beta_o x^k + \beta_1 x^{k-1} + \ldots + \beta_k.$$

We shall present a simple necessary condition for convergence involving $F(x)$, but first it will be necessary to discuss the solution of a general linear recursion.

Define the k^{th} order homogeneous recursion by

$$v_{n+k} = \alpha_1 v_{n+k-1} + \ldots + \alpha_k v_n, \quad n = 0, 1, 2, \ldots \tag{3.4}$$

or

$$v_{n+k} - \alpha_1 v_{n+k-1} - \ldots - \alpha_k v_n = 0.$$

The k^{th} order recursion requires k starting values to generate the $(k+1)^{th}$ value. The recursion then continues to use the previous k values to generate the next value. Consider the second order recursion given below.

$$v_{n+2} = v_{n+1} + 6v_n \quad \text{or} \quad v_{n+2} - v_{n+1} - 6v_n = 0$$

If $v_0 = -1$ and $v_1 = 5$ are chosen as starting values, the sequence generated by the recursion is $(-1,5,-1,29,23,\ldots)$. Define the shift operator E so that

$$Ev_n \equiv v_{n+1} \quad \text{and} \quad (\alpha E^s)v_n \equiv \alpha v_{n+s} \quad \text{for any complex } \alpha.$$

Then the recursion in equation 3.4 can be written

$$(E^k - \alpha_1 E^{k-1} - \ldots - \alpha_k E^0)v_n = 0.$$

Finally, we define the characteristic polynomial for the recursion in equation 3.4 as the function $F(x) = x^k - \alpha_1 x^{k-1} - \ldots - \alpha_k$. With the above definitions we can express the k^{th} order homogeneous recursion as $F(E)v_n \equiv 0$. The set of all solutions to a homogeneous linear recursion forms a vector space over the complex numbers.

As might be expected, the roots of the characteristic polynomial F are closely related to the solution of the recursion. The proofs of the following lemmas may be found in any textbook containing a discussion of linear recursions.

Lemma 3.1. $u_n \equiv r^n$ satisfies the recursion $F(E)u_n \equiv 0$ if and only if r is a root of the characteristic polynomial F.

Lemma 3.2. If $F(x) = (x-r_1)^{m_1} \ldots (x-r_s)^{m_s}$ $r_i \neq r_j$, then the sequence U_{ij} whose n^{th} term is given by

$$n^i r_j^n \qquad 1 \leq j \leq s$$

$$0 \leq i \leq m_j - 1$$

is a solution to the recursion $F(E)u_n \equiv 0$ and the set $\{U_{ij}\}$ is a basis for the solution space.

Now return to the general multistep method given by the two polynomials in equation 3.3. We say that the method satisfies the stability condition if the magnitude of each root of $F(x)$ is bounded by 1 and the roots of magnitude 1 are simple.

Assume that we are dealing with a method satisfying the stability condition. We now proceed to illustrate heuristically why this condition is necessary for convergence. One can show that for several classes of functions $f(x,y)$, the error term $e_n = y_n - y(x_n)$ satisfies a linear recursion whose homogeneous part has $F(x)$ as its characteristic polynomial. If the roots $r_1, r_2, \ldots, r_k$ of $F(x)$ are distinct, then the principal term in e_n is

$$Ar_1^n + Br_2^n + \ldots + Cr_k^n.$$

The roots must be less or equal to one in magnitude. The terms of the recursion which contain roots less than one in magnitude become negligible as n becomes infinite. The remaining terms which contain roots equal to one in magnitude are bounded by their constant coefficients as n becomes infinite. Thus a stable method is one for which the error will remain bounded as the interval is divided into finer and finer subdivisions. For example, the error in the choice of the initial values will not be magnified as a stable multistep method proceeds across the interval to b. We now formalize the above discussion in the following theorem.

Theorem 3.1. Necessary conditions for convergence of the linear multistep method in equation 3.2 are:

(1) $F(r) = 0$ implies that $|r| \leq 1$ and

(2) $F(r) = 0$ and $|r| = 1$ implies that r is a simple root of $F(x)$.

Proof: The conditions (1) and (2) are merely a restatement of the notion of stability. Since the method is convergent, it is convergent for the initial value problem

$$f(x,y) = 0, \qquad y(0) = 0$$

whose exact solution is $y(x) = 0$. Because $f(x,y) = 0$, the multistep method in equation 3.2 reduces to the following difference equation with constant coefficients.

$$y_{n+k} = \alpha_1 y_{n+k-1} + \alpha_2 y_{n+k-2} + \ldots + \alpha_k y_n \tag{3.5}$$

First we shall show that any root of F is bounded by 1. Let r_1 be a root of $F(x) = x - \alpha_1 x^{k-1} - \alpha_2 x^{k-2} - \ldots - \alpha_k$. Assume that $|r_1| > 1$ and let

$$u_n \equiv \begin{cases} h(r_1^n + \bar{r}_1^n) & \text{if } r_1 \text{ is complex} \\ h r_1^n & \text{if } r_1 \text{ is real.} \end{cases} \tag{3.6}$$

(Since F has real coefficients, its complex roots occur in conjugate pairs.) By lemma 3.2 $\{u_n\}$ satisfies the recursion in equation 3.5. Consider the k initial values determined by equation 3.6. For a fixed i,

$$\lim_{h\to 0} u_i = 0 = y_o \qquad 0 \le i \le k-1$$

Therefore, the choice of initial values is consistent. To approximate the value of the solution at a point x, we apply the multistep method to the interval $[0,x]$. Then the stepsize is given by $h = \frac{x}{N}$ where N is the number of subdivisions. If r_1 is real then this implies that the approximation at x is given by

$$u_N = hr_1^N = \frac{x}{N} r_1^N$$

But $\lim_{h\to 0} |u_N| = \lim_{h\to 0} \frac{x}{N} |r_1|^N = \lim_{N\to\infty} \frac{x}{N} |r_1|^N =$

$$x \lim_{N\to\infty} \frac{|r_1|^N}{N} = x \lim_{N\to\infty} \frac{e^{N \log|r_1|}}{N} = x \lim_{N\to\infty} \frac{\left(e^{\log|r_1|}\right)^N}{N}$$

$$= x \lim_{N\to\infty} \frac{|r_1|^N \log|r_1|}{1} \quad \text{which is infinite.}$$

This contradicts the hypothesis that the method is convergent. Therefore $|r_1| \le 1$. A similar result holds when r_1 is complex. Therefore, condition (1) is true.

Next we show by an argument similar to the above that any root of magnitude 1 is a simple root. Let r_1 be a root of $F(x)$ such that

$|r_1| = 1$ and assume that r_1 is a multiple root.

Let

$$u_n \equiv \begin{cases} hn(r_1^n + \bar{r}_1^n) & \text{if } r_1 \text{ is complex} \\ hnr_1^n & \text{if } r_1 \text{ is real} \end{cases}$$

Then $\{u_n\}$ satisfies the recursion. The k initial values are consistent. If r_1 is real and if the multistep method is applied to the interval $[o,x]$, the approximation at x is given by

$$u_N = hNr_1^N.$$

$$\text{But } \lim_{h\to o} |u_N| = \lim_{h\to o} \frac{x}{N} N|r_1|^N = x \lim_{N\to\infty} |r_1|^N = x$$

which is not zero unless x is zero. This contradicts the hypothesis that the method is convergent. Therefore, if $|r_1| = 1$, then r_1 is a simple root and the theorem is proved. Q.E.D.

While the condition of stability prevents a small error near the beginning of the interval from growing so fast that convergence is jeopardized, stability alone is not sufficient for convergence. A condition which will guarantee that the difference equation in equation 3.2 is a good approximation to the given differential equation is needed in addition to stability. Such a condition is the consistency condition. Until now we have emphasized that a multistep method yields an approximation to the IVP at $x_N = b$. The method, however, also yields an approximation y_i to the solution at every node x_i as the method

steps across the interval to b. For each $x\epsilon[x_o,b]$ let $y_N(x)$ be the approximation given by the multistep method acting on the interval $[x_o,x]$. Then $y_N(x)$ is a function on $[x_o,b]$. Assume that the sequence of functions $\{y_N(x)\}$ converges uniformly on $[x_o,b]$ to some function $y(x)$. Then the method is called consistent if $y(x)$ is a solution to the IVP.

Lemma 3.3. The multistep method is consistent if and only if $F(1) = 0$ and $F'(1) = g(1)$.

Proof: (Necessity). The proof of the necessity of the conditions entails examining two particular differential equations and will not be presented here.

(Sufficiency). Assume that $F(1) = 0$, $F'(1) = g(1)$, and that $y_N(t)$ converges uniformly to $y(t)$ on $[x_o,x]$. We want to show that $y(t)$ is a solution to the IVP. Define the identity operator I by $(\alpha I^s)v_n = \alpha v_n$. Set $F(t) = (t-1)f_1(t)$ since $F(1) = 0$. By definition of the polynomials F and g, the multistep method is written

$$F(E)y_n = hg(E)y_n'.$$

But $F(E)y_n = f_1(E)(E-1)y_n = f_1(E)(y_{n+1}-y_n)$.

Thus $f_1(E)(y_{n+1}-y_n) = hg(E)y_n'$. So $\sum_{k=o}^{n} f_1(E)(y_{k+1}-y_k)$

$= \sum_{k=o}^{n} hg(E)y_k'$. This is equivalent to

$$f_1(E)(y_{n+1}-y_o) = \sum_{k=o}^{n} hg(E)y_n' . \qquad (3.7)$$

Set $F_n \equiv \sum_{k=o}^{n} hy'_k = \sum_{k=o}^{n} hf(t_k,y_n(t_k))$. We are assuming that the approximation at t_k is given by $y_n(t_k)$. Then equation 3.7 becomes

$$f_1(E)(y_{n+1}-y_o) = g(E)F_n. \tag{3.8}$$

Since by assumption $y_n(t)$ converges uniformly to $y(t)$ on $[x_o,x]$, $f(t,y_n(t))$ converges uniformly to $f(t,y(t))$ on $[x_o,x]$. This follows from the fact that $f(t,y)$ is Lipschitz in y and therefore uniformly continuous in y.

To show that F_n converges to $\int_{x_o}^{x} f(t,y(t))dt$ on the interval $[x_o,b]$ we examine the amount by which F_n differs from $\int_{x_o}^{x} f(t,y(t))dt$; that is, $|F_n - \int_{x_o}^{x} f(t,y(t))dt|$. Let $\epsilon>0$. Writing F_n in terms of the derivative at $n+1$ points and adding and subtracting $\sum_{k=o}^{n} hf(t_k,y(t_k))$ we have the following.

$$|F_n - \int_{x_o}^{x} f(t,y(t))dt| \leq \sum_{k=o}^{n} h|f(t_k,y_n(t_k)) - f(t_k,y(t_k))|$$

$$+ |\sum_{k=o}^{n} hf(t_k,y(t_k)) - \int_{x_o}^{x} f(t,y(t))dt|.$$

Since $y_n(t)\to y(t)$ uniformly in t as $n\to\infty$ and since f is uniformly continuous in y, we have $|f(t_k,y_n(t_k)) - f(t_k,y(t_k))| < \frac{\epsilon}{3(x-x_o)}$ for all $n\geq N_1$ for some N_1. Then for $n\geq N_1$,

$$\sum_{k=o}^{n} h\,|f(t_k,y_n(t_k)) - f(t_k,y(t_k))| < \left(\frac{x-x_o}{n}\right) \frac{\epsilon(n+1)}{3(x-x_o)} < \frac{\epsilon}{2}.$$

Since $\sum_{k=o}^{n} hf(t_k,y(t_k))$ is a Riemann sum for $\int_{x_o}^{x} f(t,y(t))dt$ and $\int_{x_o}^{x} f(t,y(t))dt$ exists, $\left|\sum_{k=o}^{n} hf(t_k,y(t_k)) - \int_{x_o}^{x} f(t,y(t))dt\right| < \frac{\epsilon}{2}$ for all $n \geq N_2$ for some N_2. Therefore for all $n \geq \max\{N_1,N_2\}$ we have $|F_n - \int_{x_o}^{x} f(t,y(t))dt| < \epsilon$. Thus F_n converges to $\int_{x_o}^{x} f(t,y(t))dt$ for each $x \in [x_o,b]$. We apply limits to both sides of equation (3.8). For the right side of (3.8) we have

$$\lim_{n\to\infty} g(E)F_n = \lim_{n\to\infty} (\beta_o F_{n+k} + \beta_1 F_{n+k-1} + \ldots + \beta_k F_n)$$

$$= \beta_o \int_{x_o}^{x} f(t,y(t))dt + \beta_1 \int_{x_o}^{x} f(t,y(t))dt + \ldots + \beta_k \int_{x_o}^{x} f(t,y(t))dt$$

$$= (\beta_o + \beta_1 + \ldots + \beta_k) \int_{x_o}^{x} f(t,y(t))dt = g(1) \int_{x_o}^{x} f(t,y(t))dt.$$

Since $f_1(x) = x^{k-1} - a_1' x^{k-2} - \ldots - a_{k-1}'$ we have the following for the left side of (3.8).

$$\lim_{n\to\infty} f_1(E)(y_{n+1} - y_o) =$$

$$\lim_{n\to\infty} [(y_{n+k} - a_1' y_{n+k-1} - \ldots - a_{k-1}' y_{n+1}) - (y_{k-1} - a_1' y_{k-2} - \ldots - a_{k-1}' y_o)].$$

As $n \to \infty$ the subdivisions of $[x_o,x]$ become finer and the values y_{n+k}, y_{n+k-1}, $\ldots$, y_{n+1} approach $y(x)$ and the values y_{k-1}, y_{k-2}, $\ldots$, y_o approach $y(x_o)$. Therefore $\lim_{n\to\infty} f_1(E)(y_{n+1} - y_o) = y(x) - y_o$.

Therefore, $f_1(1)(y(x) - y_o) = g(1) \int_{x_o}^{x} f(t,y(t))dt$ for each $x \in [x_o,b]$.

Since $f_1(1) = F'(1) \neq 0$ because 1 is a simple root and $g(1) = f_1(1)$, the equation above becomes

$$y(x) = y_o + \int_{x_o}^{x} f(t,y(t))dt \quad \text{or}$$

$y'(x) = \frac{dy}{dx} = f(x,y(x))$. Therefore, $y(x)$ satisfies the differential equation and the multistep method is consistent. Q.E.D.

The conditions of stability and consistency are not only necessary for convergence, but they are also sufficient for convergence.

Theorem 3.2. A stable and consistent linear multistep method is convergent.

The proof of theorem 3.2 is quite involved [4], but it is based on showing that stability and consistency place a bound on the error which goes to 0 uniformly with h.

The following is a derivation of the general, convergent, second order, linear multistep method. The second order method is written in the form

$$y_{n+2} = a_1 y_{n+1} + a_2 y_n + h[b_o y'_{n+2} + b_1 y'_{n+1} + b_2 y'_n] \tag{3.9}$$

which is associated with the two polynomials

$$F(x) = x^2 - a_1 x - a_2$$

$$g(x) = b_o x^2 + b_1 x + b_2.$$

The first step is to impose the conditions of consistency. One condition is that 1 be a root of F. This yields the restriction that $a_2 = 1 - a_1$ as follows:

$$F(1) = (1)^2 - a_1(1) - a_2 = 0$$

$$1 \quad -a_1 \quad -a_2 = 0$$

$$1 \quad -a_1 = a_2$$

Another restriction on the coefficients is obtained by imposing the second condition of consistency; that is, $F'(1) = g(1)$:

$$F'(x) = 2x - a_1$$

$$F'(1) = 2 - a_1$$

$$g(1) = b_o + b_1 + b_2 = 2 - a_1$$

$$b_o + b_1 + b_2 + a_1 = 2.$$

The polynomial $F(x)$ has two roots since it is a second degree polynomial. Consistency requires that one root be 1. For the method to be stable, the second root must be less than or equal to 1 in magnitude and not equal to 1. We now obtain an expression for the second root r of $F(x)$.

$$F(x) = (x-1)(x-r)$$

$$= x^2 - (1+r)x + r$$

It is readily seen that since

$$F(x) = x^2 - a_1 x - a_2$$

$$= x^2 - (1-a_2)x - a_2$$

$-a_2 = r$. Since $-a_2$ cannot equal 1, but the magnitude of a_2 must

be less than or equal to 1, then a_2 lies in the interval $(-1,1]$. Thus, by choosing the coefficients in equation (3.9) so that $a_2 \epsilon (-1,1], a_2 = 1-a_1$, and $b_o + b_1 + b_2 + a_1 = 2$, the linear multistep method will be convergent.

Once restrictions are placed on the coefficients of equation (3.9) to make the multistep method convergent, additional conditions may be imposed to further improve the method. Since a continuous function on a closed interval can be uniformly approximated by a polynomial, it seems reasonable that if a multistep method yields the exact solution for a large class of functions $f(x,y)$ for which the solution to the IVP is a polynomial in x then one would have more confidence in the approximations for other differential equations. The degree of precision is the largest number M such that all polynomials $y(x)$ of degree less than or equal to M satisfy the multistep method exactly. By satisfying the multistep method exactly we mean that when $y(x_n)$ is substituted for y_n, the multistep method will result in an equality. The degree of precision might be maximized as a means of further restricting the coefficients and improving the multistep method. Ironically, a high degree of precision conflicts with the stability condition. Dahlquist [2], who first proved the equivalence of convergence with stability and consistency also proved that for a k-step method where k is even, the maximum degree of precision cannot exceed $k+2$ without rendering the iteration unstable ($k+1$ if k is odd). We maximize the degree of precision as follows. We shall assume that $y(x)$ is a successively higher degree polynomial, substitute $y(x_n)$ for y_n

in the multistep method, assume the method is still an equality, and thereby determine additional restrictions on the coefficients. Let $y(x) \equiv 1$. Then $y'(x) = 0$. Choose the a's and b's in (3.9) so that $y(x) \equiv 1$ satisfies (3.9); that is,

$$1 = a_1 + a_2 + h\,[b_o \cdot 0 + b_1 \cdot 0 + b_2 \cdot 0]$$

$$1 = a_1 + a_2.$$

Then let $y(x) \equiv x$, $y'(x) = 1$. Further restrict the a's and b's in (3.9) so that $y(x) \equiv x$ also satisfies equation (3.9).

$$x_{n+2} = a_1 x_{n+1} + a_2 x_n + h\,[b_o \cdot 1 + b_1 \cdot 1 + b_2 \cdot 1]$$

$$x_o + (n+2)h = a_1(x_o + (n+1)h) + a_2\,(x_o + nh) +$$

$$h[b_o + b_1 + b_2].$$

$$(1-a_1-a_2)(x_o+nh) = h(a_1 + a_2 + b_o + b_1 + b_2 - 2)$$

Since $1-a_1-a_2 = 0$ we have

$$2 = a_1 + a_2 + b_o + b_1 + b_2$$

Continuing in this fashion we have the following. If $y(x) \equiv x^k$, $y'(x) = kx^{k-1}$ then equation (3.9) becomes

$$2^k = a_1 + k[2^{k-1}b_o + b_1].$$

The equations obtained from assuming that $y(x)$ is a polynomial are listed below:

$$1 = a_1 + a_2 \qquad \text{from } y(x) = 1$$

$$1 = b_o + b_1 + b_2 - a_2 \qquad \text{from } y(x) = x$$

$$3 = -a_2 + 4b_o - 2b_1 \qquad \text{from } y(x) = x^2$$

$$7 = 12b_o + 3b_1 - a_2 \qquad \text{from } y(x) = x^3$$

$$15 = 32b_o + 4b_1 - a_2 \qquad \text{from } y(x) = x^4 \qquad (3.10)$$

Note that the first two equations reiterate the conditions of consistency.

The restrictions on the coefficients can now be used to locate a second order convergent method with the highest degree of precision. Solving the equations above yields a set of coefficients in terms of a_2:

$$b_o = \frac{5-a_2}{12} \qquad b_2 = \frac{1-5a_2}{12}$$

$$b_1 = \frac{8+8a_2}{12} \qquad a_1 = 1-a_2$$

$$a_2 = (-1,1].$$

We prefer to leave a_2 arbitrary for now and use other techniques to determine its best value in Chapter V. Only the first four equations were used, so that the degree of precision is 3. The second order method can now be written.

$$y_{n+2} = (1-c)y_{n+1} + cy_n + \frac{h}{12}\,[(5-c)y'_{n+2} + (8+8c)y'_{n+1} + (5c-1)y'_n] \qquad (3.11)$$

where $c=a_2$ is used to simplify the notation. If we had used the

fourth equation we would have $c = a_2 = 1$ so that all of the coefficients would have been completely determined.

Chapter IV

Predictor - Corrector Methods

Consider the pair of multistep methods given below.

$$y_{n+k} = \sum_{i=1}^{k} a_i^* y_{n+k-i} + h \sum_{i=1}^{k} b_i^* y'_{n+k-i} \tag{4.1}$$

$$y_{n+k} = \sum_{i=1}^{k} a_i y_{n+k-i} + h \sum_{i=o}^{k} b_i y'_{n+k-i} \tag{4.2}$$

The reader should note that y_{n+k} appears only on the left side of equation (4.1) so that (4.1) gives an <u>explicit</u> determination of y_{n+k}. Equation (4.1) is called a <u>predictor</u>. On the other hand, y_{n+k} appears on both sides of equation (4.2) (within the derivative on the right) so that y_{n+k} is <u>implicitly</u> determined by (4.2). Equation (4.2) is called a <u>corrector</u>.

Assume for the moment that $y_{n+k-1} \cdots y_n$ in equation (4.2) are fixed. Then (4.2) has the form

$$y_{n+k} = G(y_{n+k}) \tag{4.3}$$

that is, the corrector is a function of the single variable y_{n+k}. Equation (4.3) suggests the iteration

$$y_{n+k}^{(i+1)} = G\,(y_{n+k}^{(i)}) \quad i = 0, 1, 2, \cdots \tag{4.4}$$

whose purpose is to improve or correct the value of y_{n+k}.

The two equations 4.1 and 4.2 can be used together as follows. The k values $y_{n+k-1} \cdots y_n$ are substituted into the predictor which generates an initial y_{n+k}, say $y^{(o)}_{n+k}$. With this starting value and with $y_{n+k-1}, \ldots, y_n$ fixed, the corrector is iterated m times in the form (4.4) to successively generate $y^{(1)}_{n+k}, \ldots, y^{(m)}_{n+k}$. We then accept $y^{(m)}_{n+k} = y_{n+k}$. The calculations then proceed as above to produce an approximation at the next node. When equations 4.1 and 4.2 are used together as described above, they constitute a predictor-corrector method.

The convergent second order multistep method which we derived in Chapter III is given below.

$$y_{n+2} = (1-c)y_{n+1} + cy_n - \frac{h}{12}\left[(5-c)y'_{n+2} + (8+8c)y'_{n+1} + (5c-1)y'_n\right] \qquad (4.5)$$

Because y'_{n+2} appears on the right, equation 4.5 is a corrector. To derive a convergent second order multistep method in the predictor form, we go back to the equations 3.10 used in maximizing the degree of precision and set $b_o = o$.

$$1 = a_1 + a_2 \qquad \text{from } y(x) = 1$$

$$1 = b_1 + b_2 - a_2 \qquad \text{from } y(x) = x$$

$$3 = -a_2 + 2b_1 \qquad \text{from } y(x) = x^2$$

$$7 = 3b_1 - a_2 \qquad \text{from } y(x) = x^3$$

$$15 = 4b_1 - a_2 \qquad \text{from } y(x) = x^4$$

Solving the above equations yields the coefficients

$$b_o = 0 \qquad b_2 = \frac{a_2-1}{2} \qquad a_2 \epsilon (-1,1]$$

$$b_1 = \frac{3+a_2}{2} \qquad a_1 = 1-a_2$$

In order to leave a_2 free, only the first three equations were used, so that the degree of precision is 2. The predictor method thus derived is

$$y_{n+2} = (1-p)y_{n+1} + py_n + \frac{h}{2}\left[(3+p)y'_{n+1} + (p-1)y'_n\right] \tag{4.6}$$

where $p = a_2$.

It is important to note that even though the predictor and the corrector are separately convergent, at this point we can say very little about the convergence of their combined operations because the net result of predicting and correcting is not a multistep method. In the next chapters we will restrict coefficients so that the net result of predicting and correcting is convergent and we will discuss ways of finding the best values of p and c so that the predictor and one iteration of the corrector will accurately approximate the solutions to a large class of differential equations.

Chapter V

Two Step Predictor-Corrector

In this chapter we shall examine in detail the general two-step predictor-corrector pair with only one iteration of the corrector. The performance of a multistep method when $f(x,y) = p(x)$ is a polynomial in x is optimized by raising the degree of precision of the method.

We shall then consider another class of functions; namely, $f(x,y) = \lambda y$ for real λ. We consider this class of differential equations for several reasons. First, the exact solution $y(x) = y_0 e^{\lambda x}$ is known so that the accuracy of the approximation given by the predictor-corrector pair can be readily determined. Secondly, differential equations whose solutions resemble the exponential occur frequently in practice. Also, the analytical techniques introduced for this class of differential equations are the same techniques which are applied to other classes of differential equations.

Given below are the predictor-corrector equations which were derived in Chapter IV. The substitution $f(x,y) = \lambda y$ has been made.

$$(P) \quad y_{n+2} = (1-p)y_{n+1} + py_n + \frac{h}{2}\left[(3+p)\lambda y_{n+1} + (p-1)\lambda y_n\right]$$

$$(C) \quad y_{n+2} = (1-c)y_{n+1} + cy_n + \frac{h}{12}\left[(5-c)\lambda y_{n+2} + (8+8c)\lambda y_{n+1} + (5c-1)\lambda y_n\right]$$

Note that y_{n+2} now appears explicitly on the right side of the corrector. Because the equations are simple and of low order, the predictor can be substituted for y_{n+2} in the corrector.

The resultant equation represents the predictor-corrector method solved explicitly for y_{n+2}.

$$y_{n+2} = Ay_{n+1} + By_n \tag{5.1}$$

where $A = \frac{1}{24}\left[24-24c + 26H + 14cH - 10pH + 2cpH + 15H^2 - 3cH^2 + 5pH^2 - cpH^2\right]$

and $B = \frac{1}{24}\left[24c + 10pH - 2cpH + 5pH^2 - cpH^2 - 5H^2 + cH^2 + 10cH - 2H\right]$, $H = \lambda h$.

Equation 5.1 is a homogeneous second order linear recursion.

Both the predictor and the corrector were convergent second order multistep methods. Substituting the predictor into the corrector does not, however, ensure that the resultant procedure is convergent. An examination of how well the approximation generated by (5.1) agrees with the true solution for $f(x,y) = \lambda y$ will lead to a restriction on the allowable values of H as a function of p and c.

The characteristic polynomial of the recursion in (5.1) is $F(x) = x^2 - Ax - B = (x-R_1)(x-R_2)$. Using the quadratic formula we find that

$$R_1 = \frac{A + \sqrt{A^2+4B}}{2}, \quad R_2 = \frac{A - \sqrt{A^2+4B}}{2} \tag{5.2}$$

The following theorem will be used to show that the approximation to $y(x)$ provided by (5.1) converges to the true solution for arbitrary λ.

Theorem 5.1 If $h = \frac{x}{N}$ then $\lim_{N\to\infty} R_1^N = e^{\lambda x}$.

Proof: Consider $R_1 = \frac{A + \sqrt{A^2+4B}}{2}$. Recall that R_1 is a function of p, c, λ, and h. To simplify R_1 we first check to see if $A^2 + 4B$ is a perfect square. If this were so then equation 5.3 would be satisfied for some D, E, and F with $H = h\lambda$.

$$A^2+4B = \left[DH^2 + EH + F\right]^2 \tag{5.3}$$

Expanding $A^2 + 4B$ and equating the coefficients D^2, F, and 2EF with like terms on the left yields

$$D = \frac{-15+3c-5p+cp}{24}, \quad E = \frac{11-7c+5p-cp}{12}, \quad F = c+1 \tag{5.4}$$

Evaluating $[DH^2+EH+F]^2$ with the values of D, E, and F given in (5.4) results in an expression which is not equal to A^2+4B. Hence, A^2+4B is not a perfect square. However, we can write A^2+4B as the sum of the square in (5.3) and the X given below.

$$A^2+4B = \left[DH^2+EH+F\right]^2 + X \tag{5.5}$$

where $X = H^2\left[\frac{H}{144}(360-72c+120p-44cp) + (2+2c)\right]$

Let $G = \frac{H}{144}\left[(360-72c+120p-44cp) + (2+2c)\right]$.

Recall the binomial expansion formula for finding the square root of a sum $u + v$.

$$(u+v)^{1/2} = u^{1/2} + \frac{1}{2}u^{-1/2}v - \frac{1}{8}u^{-3/2}v^2 + \frac{1}{16}u^{-5/2}v^3 - \frac{5}{128}u^{-7/2}v^4 + \ldots$$

Let $\sqrt{u} = DH^2 + EH + F = C$ and let $v = X = H^2G$. Note that $\lim_{H\to 0} u = F^2 \neq 0$. $F^2 = (c+1)^2 \neq 0$ because $c \neq -1$. Then $u \neq 0$ and we may apply the binomial expansion formula to the right side of (5.5) to find the square root of A^2+4B.

$$(u+v)^{1/2} = (C^2+H^2G)^{1/2} = C + \frac{1}{2}\frac{H^2G}{C} - \frac{1}{8}\frac{H^4G^2}{C^3} + \frac{1}{16}\frac{H^6G^3}{C^5} - \frac{5}{128}\frac{H^8G^4}{C^7} + \ldots$$

Factoring out H^2 gives

$$(u+v)^{1/2} = C + H^2\left[\frac{1}{2}\frac{G}{C} - \frac{1}{8}\frac{H^2G^2}{C^3} + \frac{1}{16}\frac{H^4G^3}{C^5} - \frac{5}{128}\frac{H^6G^4}{C^7} + \ldots\right] \qquad (5.6)$$

Let the alternating series within the brackets of equation 5.6 be denoted $\sum_{n=1}^{\infty} a_n$. Consider the geometric series $\sum_{n=1}^{\infty} b_n$ where $b_n = \frac{G}{C}\left[\frac{H^2G}{C^2}\right]^n$. Since the constant coefficients in Σa_n are less than one in magnitude, $|a_n| \leq |b_n|$ for all n. This implies that if Σb_n is convergent, then Σa_n will be convergent by the comparison test for series. The geometric series Σb_n will converge if and only if $|\frac{H^2G}{C^2}|<1$ or $|H^2G|<|C^2|$. But for some N_o we see that $|H^2G|<|C^2|$ for all $N \geq N_o$ with $H = h\lambda = \frac{x\lambda}{N}$. Hence Σb_n is convergent and therefore Σa_n is convergent. Equation (5.5) can rewritten $(u+v)^{1/2} = C + H^2(\Sigma a_n)$ or

$$(A^2+4B)^{1/2} = DH^2 + EH + F + O(H^2) \text{ terms.} \qquad (5.7)$$

Here $O(H^2)$ denotes terms such that $\lim_{H\to 0} \frac{O(H^2)}{H^2}$ exists. Now we may simplify R_1. Substituting the results given in (5.7) into the expression for R_1 given in (5.2) yields

$$R_1 = \frac{A + DH^2 + EH + F}{2} + O(H^2).$$

Upon simplification we have

$$R_1 = 1 + H + O(H^2).$$

A problem in Henrici [4] demonstrates that

$$\lim_{N\to\infty} \left[1 + \frac{\lambda x}{N} + O(H^2)\right]^N = e^{\lambda x}.$$

Therefore, $\lim_{N\to\infty} R_1^N = e^{\lambda x}$. Q.E.D.

We have thus far restricted ourselves to differential equations of the form $f(x,y) = \lambda y$. Further assume that these equations have initial values given at zero so that $y_o = y(0)$. To approximate the values of y at some point x, we divide the interval $[0,x]$ into N subdivisions. Then by the theory of linear recursions the approximation to $y(x)$ is given by y_N in (5.1) where

$$y_N = aR_1^N + bR_2^N \tag{5.8}$$

for some coefficients a and b which depend on y_o and h.

To determine the coefficients a and b we solve the pair of simultaneous equations which result from setting N in equation (5.8) equal to 0 and then equal to 1.

$$y_o = a + b$$

$$y_1 = aR_1 + bR_2$$

$$a = \frac{1}{R_2 - R_1} \quad (R_2 y_o - y_1)$$

$$b = \frac{1}{R_2 - R_1} \quad (-R_1 y_o + y_1)$$

Recall the expressions for R_1 and R_2 in terms of A and B in equations (5.1) and (5.2). Since

$$\lim_{h \to o} A = 1-c \qquad \text{and} \qquad \lim_{h \to o} B = c$$

we have that

$$\lim_{h \to o} R_1 = 1 \qquad \text{and} \qquad \lim_{h \to o} R_2 = -c.$$

Also, y_1 approaches y_o as h approaches 0, if y_1 is a consistent choice. Thus

$$\lim_{h \to o} a = y_o \,, \qquad \lim_{h \to o} b = o.$$

Therefore, if y_1 is a consistent choice then the fact that a approaches y_o as $h \to o$ together with theorem (5.1) shows that aR_1^N converges to $y_o e^{\lambda x}$, the true solution.

Since aR_1^N approximates the true solution, R_1 is called the <u>principal root</u>. The other root R_2 is called an <u>extraneous root</u>. The term bR_2^N represents the error e_N inherent in using a difference

equation to approximate a differential equation. Its value bears no relation to the true solution. If $|R_2|>|R_1|$ then the component of the numerical solution which corresponds to bR_2^N will dominate and as a result the numerical solution will not approximate the true solution. This domination of the numerical component which converges to the true solution by a component which is not related to the true solution is known as numerical instability.

At first glance it might seem that in order to prevent numerical instability and thereby ensure that our method converges to the true solution we must only require that $|R_2|\leq|R_1|$. However, Lapidus [5] demonstrates that for the case $f(x,y) = \lambda y$ the error e_N satisfies a linear recursion having the same roots as the recursion in (5.8). Therefore

$$e_N = cR_1^N + dR_2^N$$

where c and d are constants determined by the starting values y_o, y_1. Then for the approximation y_N provided by the recursion to be accurate, e_N must not grow as N gets larger. Requiring e_N to be nonincreasing as N increases is equivalent to requiring $|R_1|\leq 1$ and $|R_2|\leq 1$. A method for which the roots R_1, R_2 are less than or equal to one in magnitude is termed absolutely stable.

Think of $H = \lambda h$ as a complex number for a moment. The region in the complex H-plane in which the magnitudes of the roots in (5.8) are less than or equal to one is called the stability region. For a fixed value of H in the stability region the error e_N will not grow but in fact will diminish to zero as N becomes larger. That is, larger

values of N correspond to x_N values further from x_o. When the roots are less than one in magnitude, e_N will not grow with larger N. A stable choice of H therefore keeps the method from propagating errors over large distances. We certainly want to operate our method in the stability region. We shall discuss further the concept of stability after finding the stability regions for our method in (5.8).

In analyzing the stability region we shall calculate only its intersection with the real axis. To do this it is helpful to adopt the notation of Hall [3]. We rewrite the predictor-corrector equations for the case $f(x,y) = \lambda y$ as

$$(P) \quad F(E)y_n = Hg(E)y_n$$

$$(C) \quad F^*(E)y_n = Hg^*(E)y_n$$

where $F(s) = s^2 - (1-p)s - p$

$$g(s) = \left[\frac{3+p}{2}\right] s + \frac{p-1}{2}$$

$$F^*(s) = s^2 - (1-c)s - c$$

$$g^*(s) = \left[\frac{5-c}{12}\right] s^2 + \left[\frac{8+8c}{12}\right] s + \frac{5c-1}{12}$$

$$H = h\lambda .$$

The following polynomials are called the stability polynomials for the predictor and the corrector respectively.

$$L(s) = F(s) - Hg(s) \qquad (P)$$

$$L^*(s) = F^*(s) - Hg^*(s). \qquad (C)$$

The stability polynomial for the recurrence relation which results from substituting the predictor into the corrector is then found to be

$$L^*(s) + \theta L(s), \quad \theta = H\left[\frac{5-c}{12}\right] \tag{5.9}$$

To ensure that the recursive method is absolutely stable, we must determine the region in the complex H-plane where the roots of the stability polynomial in (5.9) are less than one in magnitude. Toward this end, the equation $L^*(s) + \theta L(s) = 0$ is solved for H.

$$H = \frac{bF(s) - g^*(s) \pm \sqrt{[g^*(s) - bF(s)]^2 + 4bg(s)F^*(s)}}{2bg(s)}$$

$$b = \frac{5-c}{12} .$$

Experience [3] has shown that the intersections of the boundary of the stability region with the real axis can be found by setting $s = \pm 1$. For $s = \pm 1$ we find the following values of H.

$$(s=1) \quad H_1 = \frac{-12(c+1)}{(5-c)\ (p+1)} \quad \text{and} \quad 0 \tag{5.10}$$

$$(s = -1) \quad H_{-1} = \frac{(7+c-5p+cp) \pm \sqrt{(-c-7+5p-cp)^2 - 48(5-c)(1-c)}}{2(c-5)} \tag{5.11}$$

When the expression under the radical in (5.11) is nonnegative, there are four real values of H on the boundary of the stability region, giving two disjoint stability intervals along the real H-axis; otherwise, there are two values which are the endpoints of a single stability

interval. Note that zero is always in the stability region. Since the method is operated by decreasing h to zero, it is reassuring that h can be decreased to zero from within the stability region. The stability regions in each of these cases appear schematically in Figure 3 below.

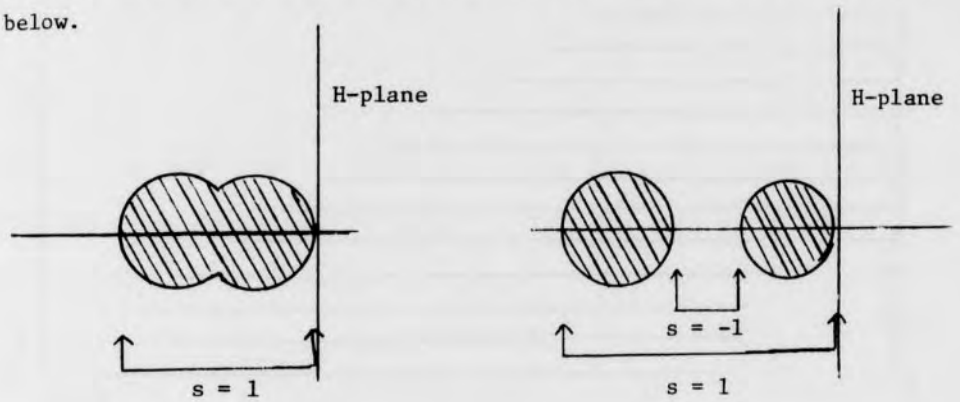


Figure 3. Stability regions.

Note that the stability intervals are located on the negative real axis. This is equivalent to $h\lambda<0$ or $\lambda<0$ since $h>0$. Notice that when $\lambda<0$ the true solution to the differential equation $f(x,y) = \lambda y$ is decreasing as x increases. If we choose p, c, h, and λ so that the method is absolutely stable, that is, $|R_1|\leq 1$ and $|R_2|\leq 1$, then we have that the error term e_N is also nonincreasing with increasing N. Most authors automatically restrict their considerations to $\lambda < 0$ because absolute stability would require this anyway.

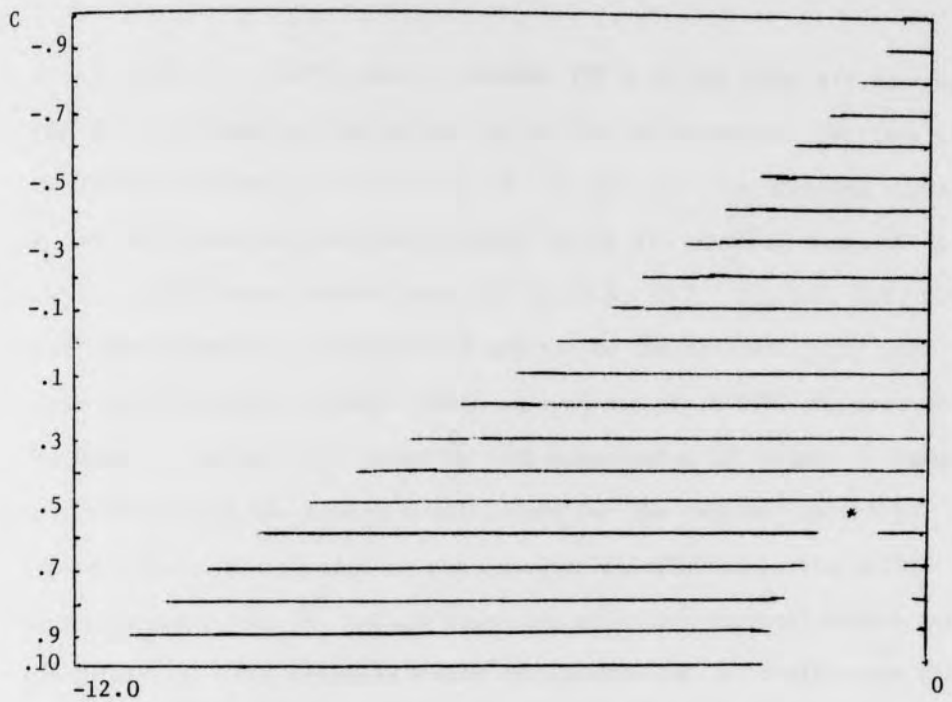


Figure 4. Splitting of stability intervals

Figure 4 shows how a single stability interval splits into two disjoint stability intervals as c ranges over its possible values for a fixed p. For p values ranging from -.9 to 1.0 in increments of 0.1, Appendix A gives the critical c values where the stability interval splits into two intervals. These values were found by setting the expression under the radical sign in (5.11) equal to zero and solving for c in terms of p.

$$c = \frac{5p^2-2p-151 + \sqrt{-2304p + 13824}}{p^2 + 2p - 47} \qquad (5.12)$$

An experiment was performed to see if the critical c values in (5.12) could provide any guidelines for choosing a c coefficient to go

with a given p coefficient. Consider IVP's of the form $f(x,y) = \lambda y$, $y(o) = 1$. We applied the method to the five differential equations of the above form where $\lambda = -1, -3, -5, -7,$ and -9. The starting value y_1 in each case was the exact value. Using all possible pairs of p and c coefficients where $p,c \in \{-0.9, -0.8, -0.7, \ldots, 0.7, 0.8, 0.9, 1.0\}$ the method was successively applied to the interval $[0,2]$ with finer and finer subdivisions until an accuracy of .000005 was reached. The best c value for a given p value was taken to be the c value that resulted in the fewest subdivisions for the required accuracy. Figure 5 shows the results of the computer calculations. The solid line represents the c values where the stability interval splits into two intervals. The triangle symbol represents the c coefficient which provided the fastest convergence for the required accuracy when $\lambda = -1$; the diamond symbol corresponds to $\lambda = -3$; the square corresponds to $\lambda = -5$; the cross corresponds to $\lambda = -7$; and the circle corresponds to $\lambda = -9$. Notice that for smaller values of λ, that is, values of -5 and less, the c value providing the fastest convergence to the true solution is less than the critical c value where the stability interval splits. As the λ values decrease, so does the value of the true solution at the endpoint 2. It would seem then that as λ decreases, the effects of error propogation would become more evident in the approximation and, hence, stability considerations would become more important. It may be that the best c coefficient is always less than the critical c value when stability considerations are very important. Otherwise, when the value of the solution at the endpoint is large enough so that

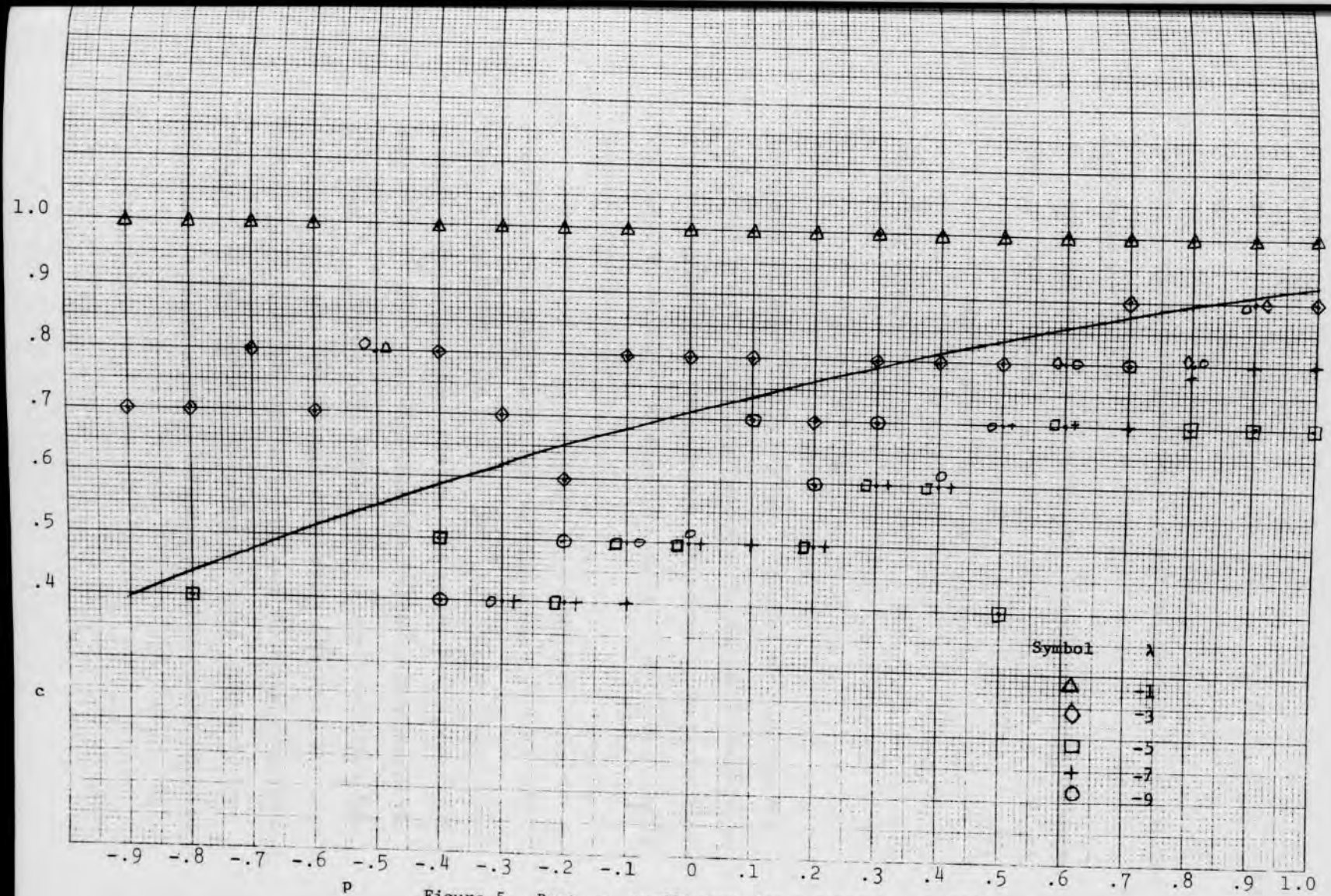


Figure 5. Best c coefficient for a given p coefficient

stability considerations are not as important, the best c coefficient is between 0.8 and 1.0.

The predictor-corrector method was also applied to the following IVP's.

$$f(x,y) = 5x^6 - 5xy - 3x^2, \; y(o) = 0$$

whose true solution is $y(x) = x^5 - x^3$ and

$$f(x,y) = 6x^7 - 4x^3 - 6xy, \; y(o) = 0$$

whose true solution is $y(x) = x^6 - x^4$. Both of the solutions are polynomials which have a degree exceeding the degree of precision of both the predictor and the corrector. Therefore, we would not expect the method to be exact. In every case the c coefficient providing fastest convergence was $c = 1.0$.

Several other IVP's were tested. In almost every case, the c coefficient providing the fastest convergence was $c = 1.0$. This could be explained by the fact that when $c = 1.0$, the degree of precision of the corrector is raised from degree 3 to degree 4. In fact, $c = 1$ gives the unique convergent corrector with degree of precision 4. Apparently the higher degree of precision overshadows stability considerations for many differential equations. But it is interesting to note that this is not the case for certain differential equations with exponentially decreasing solutions. Therefore, in choosing a c coefficient the following guidelines may be employed.

For a given p value try $c = 1$. If that choice of c is unsatisfactory then use the c value just below the critical c value.

Chapter VI

Summary

An initial value problem consists of an ordinary differential equation and an initial condition, $f(x,y) = y'$, $y(x_o) = y_o$. The initial value problem has a unique solution on the interval $[x_o, b]$ if the derivative is continuous on the strip $x_o \leq x \leq b$, $-\infty < y < \infty$ and if it satisfies a Lipschitz condition.

Discrete variable methods are numerical techniques which approximate values of a function at discrete points. A multistep method is a type of discrete variable method which uses exact or approximate values of a function at equally spaced points on the x-axis to approximate values of that function at other points on the x-axis. If a linear multistep method is stable and consistent and if a consistent choice of initial values is used, the method will be convergent.

A predictor is a multistep method which approximates the value of a function at x_N with no knowledge of the value of the derivative y' at x_N; a corrector is a multistep method which uses $y'(x_N)$ to approximate $y_N(x_N)$. The predictor is used to generate a rough estimate of the given function at a point x_N, then the corrector is applied to correct or improve the predicted value at x_N. Used together in this way they constitute a predictor-corrector pair.

If the initial value problem is of the form $y' = \lambda y$, $y(x_o) = y_o$ for real lambda, the predictor can be inserted in the corrector to yield

a recursive method having terms containing the coefficients p, c, and $H = h\lambda$. When the predictor and the corrector are second order methods with degree of precision two and three, respectively, the recursive method is convergent if the values of p and c lie in the interval $(-1, 1]$.

BIBLIOGRAPHY

1. Boyce, William E. and Richard C. DiPrima. Introduction to Differential Equations. New York, 1970.

2. Dahlquist, G. "Convergence and Stability in the Numerical Integration of Ordinary Differential Equations," Math. Scand., vol. 4, 1956. pp. 33-53.

3. Hall, George. "Stability Analysis of Predictor-Corrector Algorithms of Adams Type," SIAM J. NUMER. ANAL., vol. 11, 1974. pp. 494-505.

4. Henrici, Peter. Discrete Variable Methods in Ordinary Differential Equations. New York, 1962.

5. Lapidus, Leon and John H. Seinfield. Numerical Solution of Ordinary Differential Equations. New York, 1971.

6. Manougian, Manoug N. and Robert A. Northcutt. Ordinary Differential Equations. Columbus, Ohio, 1973.

p = -0.950	c = 0.3760
p = -0.900	c = 0.3973
p = -0.850	c = 0.4181
p = -0.800	c = 0.4385
p = -0.750	c = 0.4585
p = -0.700	c = 0.4781
p = -0.650	c = 0.4972
p = -0.600	c = 0.5160
p = -0.550	c = 0.5344
p = -0.500	c = 0.5523
p = -0.450	c = 0.5699
p = -0.400	c = 0.5871
p = -0.350	c = 0.6039
p = -0.300	c = 0.6203
p = -0.250	c = 0.6324
p = -0.200	c = 0.6520
p = -0.150	c = 0.6674
p = -0.100	c = 0.6823
p = 0.050	c = 0.6969
p = 0.000	c = 0.7112
p = 0.050	c = 0.7250
p = 0.100	c = 0.7386
p = 0.150	c = 0.7518
p = 0.200	c = 0.7646
p = 0.250	c = 0.7771
p = 0.300	c = 0.7893
p = 0.350	c = 0.8011
p = 0.400	c = 0.8126
p = 0.450	c = 0.8237
p = 0.500	c = 0.8345
p = 0.550	c = 0.8450
p = 0.600	c = 0.8552
p = 0.650	c = 0.8650
p = 0.700	c = 0.8744
p = 0.750	c = 0.8836
p = 0.800	c = 0.8924
p = 0.850	c = 0.9009
p = 0.900	c = 0.9090
p = 0.950	c = 0.9168
p = 1.000	c = 0.9243

APPENDIX A

Critical c values