Stable Recursions

With Applications to the Numerical Solution of Stiff Systems

J. R. CASH

Department of Mathematics
Imperial College of Science and Technology, London

1979

ACADEMIC PRESS
A subsidiary of Harcourt Brace, Jovanovich, Publishers
London New York Toronto Sydney San Francisco

ACADEMIC PRESS INC. (LONDON) LTD.
24/28 Oval Road,
London NW1

United States Edition published by
ACADEMIC PRESS INC.
111 Fifth Avenue
New York, New York 10003

British Library Cataloguing in Publication Data

Cash, J R
Stable recursions.—(Computational mathematics and applications).
1. Differential equations—Numerical solutions
2. Recursive functions
I. Title II. Series
515'.352 QA372 79-50521
ISBN 0-12-163050-1

Printed in Gt Britain by Page Bros (Norwich) Ltd., Norwich

Stable Recursions

With Applications to the Numerical Solution of Stiff Systems

COMPUTATIONAL MATHEMATICS AND APPLICATIONS

Series Editor

J. R. WHITEMAN

Institute of Computational Mathematics, Brunel University, England.

E. HINTON and D. R. J. OWEN: Finite Element Programming

M. A. JASWON and G. T. SYMM: Integral Equation Methods in Potential Theory and Elastostatics

J. R. CASH: Stable Recursions: with applications to the numerical solution of stiff systems

Preface

The demands of modern technology have produced a need for scientists to solve numerical initial and boundary value problems involving systems of ordinary differential equations. As a result a multitude of finite difference techniques have been proposed for dealing with these problems. The advent of modern computers gave many of these techniques wide application and consequently a set of *ad hoc* procedures which were known to be useful for practical computation, in some rather vague sense, were developed. For a long time, however, the theoretical aspects of these finite difference schemes remained neglected. Probably the first really successful attempt to establish finite difference schemes on a firm theoretical basis arose from the work of Dahlquist and Henrici. The latter author's book, published in 1962, has become a cornerstone for the basic theory of discretisation methods for ordinary differential equations. In a more recent book Stetter (1973) has produced a general framework which encompasses many of these methods and, as a consequence, the classical $h \to 0$ theory of discretisation methods seems to be in a relatively satisfactory state. However, both Henrici and Stetter, amongst many others, have noted that there is at least one class of equations, namely stiff systems of equations, for which this analysis is inappropriate and which requires rather special treatment.

A great deal of the early theory relating to stiff systems of equations was developed by Dahlquist. This culminated in his classic theorem that "An A-stable linear multistep method must be implicit and must have an order of accuracy less than or equal to 2". Implicit in the proof of Dahlquist's theorems, although not usually stated when his results are quoted, is the requirement that the desired solution of the linear multistep method in question has to be generated directly in the conventional way. It is usually the case that it is better to use an efficient direct method, if available, rather than an indirect one. However, as an immediate consequence of Dahlquist's and similar theorems, most reasonable finite difference methods for the numerical solution of stiff

systems of non-linear ordinary differential equations are essentially iterative. Thus it is natural to ask whether or not improvements in the stability properties of our schemes could be obtained if the solution were to be generated from the outset using an iterative technique. As we shall show in Chapter 3, if the required solution is generated using a precisely defined iteration scheme acting on the linear multistep method itself, it is possible to derive schemes which are A-stable and which have an order of accuracy greater than 2 and also to derive convergent linear k-step methods with orders of accuracy greater than $2[k/2] + 2$.

In this text we examine various classes of iteration schemes in the hope of developing finite difference procedures which can be used for the efficient numerical integration of systems of ordinary differential equations. It is pointed out that it is not meaningful merely to state whether or not a particular linear multistep method is stable. The precise way in which the required solution is to be generated must be stated and only then can the stability properties of the complete method (= discretisation scheme + precise method of generating the required solution) be discussed.

The text is intended mainly for scientific research workers and requires very few pre-requisites apart from a knowledge of some of the basic analytic properties of difference and differential equations. It is not meant as an introduction to finite difference methods but it is hoped that the material presented will help the reader to deepen his understanding of this particular class of methods. One of the key observations is that the well known class of finite difference schemes based on linear multistep methods, with which the reader should be familiar, is merely a particular case of a much wider class of integration procedures which are examined in detail in the forthcoming chapters.

As the title suggests, my main aim in writing this book has been to bring together algorithms for the efficient numerical integration of stiff systems of ordinary differential equations since this is one particular area where the existing theory is far from satisfactory. At several points in the text, however, I have departed from this basic aim and have considered the development of algorithms for the solution of various non-stiff systems in cases where it has seemed that an iterative approach may be valuable. Several of the algorithms have proved difficult to analyse with complete mathematical rigour and it has been necessary to rely almost entirely on numerical evidence to demonstrate their potential. I have, however, not worried too much if an algorithm is not backed up by the necessary mathematical theory. Rather I have proposed the algorithm in the hope that this will stimulate further research. In view of this,

the present text is not in a "cut and dried" form, but contains numerous open ended questions which seem to merit further investigation. Indeed, for this reason I have at several points deliberately suggested algorithms with virtually no theoretical backing (but with numerical results to demonstrate their potential usefulness). At many points in the text I have illustrated particular algorithms by applying them to fairly simple test problems, since it is often the case that the actual mechanics of an algorithm become more transparent if it is applied to a simple, rather than a complicated problem. I have, however, been careful to ensure that the algorithms derived also perform reasonably well on more complicated problems of practical interest.

In Chapter 1 some of the relevant theory relating to the stability properties of discretisation schemes and linear recurrence relations is introduced. Classically the theories of finite difference integration schemes and recurrence relations have been developed separately and one of the main purposes of this text is to examine and exploit as much as possible the interrelationships between these two theories. In Chapter 2 we examine in more detail the theory behind the numerical solution of linear recurrence relations, since many of our ideas for the solution of ordinary differential equations come from this source. In the first part of Chapter 2 we consider Olver's algorithm for the numerical solution of second order linear recurrence relations and examine its extension to vector equations. Although Olver's algorithm is extremely efficient for this particular class of problems it does seem to have the drawback that it cannot be extended directly to deal with non-linear equations. In order to overcome this difficulty an iterative method of solution is derived and this has the advantage that, as well as being reasonably efficient for linear problems, it may be extended directly to an important class of non-linear equations. In Section 2.5 we consider the extension of these iterative techniques to the solution of higher order recurrence relations and as a result a class of algorithms based on the Gauss-Seidel approach is constructed. In the non-linear case it is necessary to use two distinct iteration schemes, which we have referred to as the primary and the secondary schemes, and it is shown that a one-step Gauss-Seidel–modified-Newton scheme is efficient for the solution of our non-linear equations. Finally in Chapter 2 we consider a completely different class of iteration schemes which do not have as wide a range of application as those considered earlier but which do have very important practical applications in the numerical solution of ordinary differential equations.

In Chapter 3 we consider the extension of some of the recurrence relation techniques derived in the previous chapter to the numerical solution of systems of ordinary differential equations. Although, as previously mentioned, our

main interest is in developing algorithms for the numerical solution of stiff systems, for the sake of completeness we start our analysis in Chapter 3 with a discussion of the non-stiff case. A class of algorithms based on one-step Gauss-Seidel–direct-iteration schemes is developed and it is shown that these are often satisfactory for the solution of our problem. Also considered in Section 3.2 is the class of conventional Adams predictor–corrector schemes used in PECE mode and it is shown that these schemes may be extended in a natural way to yield a more general class of Adams integration schemes. In Section 3.4 we examine in more detail the stability aspects of our iterative integration procedures. It will have become clear by the time that this stage of the text is reached that the existing theory is not sufficient to cover iterative methods, since different ways of generating the required solution often give rise to methods with differing stability properties. A new concept of iterative absolute stability is derived and this helps us to recognise iterative algorithms suitable for the integration of ordinary differential equations. In Section 3.5 a class of iterative algorithms based on a Gauss–Seidel approach is developed for the integration of stiff systems. Although we are not quite able to achieve full A-stability, a class of schemes which behave like either one- or two-step methods, have high orders of accuracy, and have infinite regions of absolute stability is derived and this class is shown to be useful for the integration of stiff systems. In Section 3.6 we consider a somewhat different approach based on the class of iteration procedures developed in Section 2.7. This approach is rather unusual in that the finite difference schemes in question are not solved exactly and the primary iteration scheme is only applied a fixed number of times—usually once. Also in this section a close correspondence is derived between our algorithms and the method of deferred correction originally developed by Fox. The approach does have practical significance since it allows us to develop A-stable schemes of order greater than two. In Section 3.12 we consider the extension of some of the techniques developed earlier in Chapter 3 to deal with the numerical solution of second order equations. The algorithms derived do not have a firm theoretical basis and as a result it is necessary to rely mainly on numerical evidence to demonstrate their potential. In particular an attempt is made to extend Olver's analysis to derive "quadratic factors" of high order linear recurrence relations and this is another area where further research is called for. Finally, in Section 3.13 some of the iterative techniques developed for ordinary differential equations are extended to deal with the heat conduction equation with non-derivative boundary conditions. Although only relatively few numerical experiments have been carried out with these particular algorithms, the results obtained do seem to indicate that the approach is promising. It is once more hoped that the

algorithms will serve to stimulate further research, although there is an obvious need for additional theoretical results.

In Chapter 4 we consider a radically different approach to the solution of stiff systems. We extend some of the procedures developed in Chapter 2 to produce a class of iterative algorithms which are applied directly to the system of differential equations itself rather than to an approximating discretisation method. This class of procedures has a much more limited range of application than the algorithms derived in previous chapters, since it requires the system to have a special structure. For systems which do possess this structure, and these are quite common in practice, the algorithms are particularly powerful since they do not involve truncation error. In the second half of Chapter 4 we consider the relationships between these iterative algorithms and certain other widely used methods based on singular perturbations and pseudo-steady-state approximations. The main practical difficulties associated with these last two classes of methods are discussed in some detail and it is shown how the algorithms developed in the first half of Chapter 4 overcome most of these difficulties. Finally, in the last part of Chapter 4 we consider algorithms for the solution of certain classes of second order equations which cannot easily be solved using finite difference techniques. In particular an efficient algorithm is derived for the solution of certain problems having a Jacobian matrix possessing at least one eigenvalue with a large imaginary part.

I should like to take this opportunity of expressing my thanks to the numerous people who have influenced my thinking on this subject. I was fortunate indeed to be able to spend some time with Professor Hans Stetter in Vienna and without his constant help and encouragement during the early stages of the preparation of the manuscript this book would surely never have been written. Also I should like to mention the tremendous help received from Professor F. W. J. Olver and Drs R. V. M. Zahar and J. C. P. Miller, all of whom taught me a great deal about recurrence relations. I am pleased to be able to thank Professor Richard Bellman for asking me to write this book, Professor John Whiteman for placing this book in his series and for his tremendous help in editing the manuscript, Academic Press for their technical advice during the preparation of the manuscript and Mrs Sandra Place for her excellent typing from an almost unintelligible script. Finally I would like to thank my wife, Roslyn, for her love and understanding during the time spent writing this book and (if I may steal an apt sentiment from the introduction of a book by H. J. Stetter, 1973) if there is anyone at all who will rejoice at the appearance of this text it will be her.

January 1978. J.C.

Contents

1

Some Interrelationships Between the Theories of Recurrence Relations and Ordinary Differential Equations

1.1 Introduction

Traditionally the techniques and theories relating to the solution of linear recurrence relations and to the numerical solution of initial value problems in ordinary differential equations (O.D.E.s) using linear multistep methods, although often closely related, have usually developed along slightly different lines. Algorithms for the numerical solution of linear recurrence relations are usually designed so as to be able to generate *any* desired solution which is uniquely specified in advance by means of certain initial conditions (even though this problem may be very badly posed initially). However, when solving O.D.E.s using finite difference methods it is usual to assume that the required solution is in some sense *dominant*, at least for small values of the steplength of integration. The main purposes of the present text are to examine in some detail the interrelationship between these two theories and also to develop a different approach to the numerical solution of O.D.E.s by exploiting the theory of linear recurrence relations. In particular we shall develop classes of iterative techniques which do not demand that the required solution of the system of differential equations being solved is in any sense a dominant solution of the finite difference scheme being used and which converge rapidly to the required solution for a large set of initial approximations. These techniques produce satisfactory results for the integration of stiff initial value problems using linear multistep methods even

when these are not zero-stable. This approach also enables us to develop schemes which are based on linear multistep methods and which are A-stable with order greater than two.

As an introduction this chapter starts with a brief survey of some of the relevant theory relating to the solution of linear recurrence relations and ends with an analysis of some of the stability aspects of linear multistep methods.

1.2 Some results from the theory of linear recurrence relations

Consider first the solution of the k^{th} order linear recurrence relation

$$\sum_{j=0}^{k} a_j(n) y_{n+j} = g(n), \tag{1.1}$$

where the $a_j(n)$ and $g(n)$ are given sequences of the non-negative integer variable n with $a_k(n) \neq 0$ for any n in the region of interest. Any solution, y_n, of (1.1) may be expressed in the general form

$$y_n = \sum_{i=1}^{k} c_i y_{i,n} + p_n, \tag{1.2}$$

where the $y_{i,n}$, $i = 1(1)k$, are any k linearly independent solutions of the homogeneous part of (1.1)—i.e. equation (1.1) with $g(n) \equiv 0$—p_n is any particular solution of (1.1) and the c_i are arbitrary constants. The linearly independent set $(y_{1,n}, \ldots, y_{k,n})$ will henceforth be referred to as a *basis* for the solution of (1.1), and any member $y_{i,n}$ of this set will be called a basis solution. Any basis solution $y_{1,n}$ having the property that

$$\left.\begin{aligned} &\lim_{n \to \infty} (y_{i,n}/y_{1,n}) = 0 \text{ for all } i = 2, 3, \ldots, k \\ &\text{and} \\ &\lim_{n \to \infty} (p_n/y_{1,n}) = 0 \end{aligned}\right\} \tag{1.3}$$

will be called *dominant at infinity* or simply *dominant* and it will be assumed that it is possible to choose p_n and the $y_{i,n}$ so that there is a basis solution, $y_{1,n}$, which is dominant. Most of the analysis could be extended to include the case where there are two or more independent dominant solutions, e.g. where the largest roots of the characteristic equation (in the constant coefficient case) comprise a conjugate complex pair. We shall, however, restrict ourselves to the case where there is a single dominant solution since, as will be seen later, this is

the case that is of most concern when differential equations are integrated using linear multistep methods. Any solution y_n of (1.1) containing a non-zero multiple of $y_{1,n}$ will be called a dominant solution. When this is not so, y_n will be called a non-dominant solution of (1.1). If any k consecutive values of $y_j, j = 0(1)k-1$, are given initially, the constants c_i in (1.2) are completely specified and a unique solution of (1.1) is defined. This unique solution could be generated recursively from the given initial conditions by using (1.1) in the form

$$y_{n+k} = (1/a_k(n))\left\{g(n) - \sum_{j=0}^{k-1} a_j(n) y_{n+j}\right\} \qquad n = 0, 1, 2, \ldots, \tag{1.4}$$

where the right hand side of (1.4) is known when needed for all n. Equation (1.4) will be referred to as a *direct method* for generating the solution of (1.1).

Although, as previously mentioned, the required solution of (1.1) is in theory completely specified by any given k consecutive values of y_n, an unstable build up of rounding errors can often render (1.4) completely ineffective as a numerical method for the solution of (1.1). This is invariably so if the required solution is a non-dominant one. This can be seen if it is assumed that we require to compute a solution $\bar{y}_n$ of (1.1) which is such that

$$\bar{y}_n = \sum_{i=2}^{k} c_i y_{i,n} + p_n, \qquad c_2 \neq 0.$$

The effect of rounding, which is in general inevitable on an automatic machine, means that a perturbed solution $\hat{y}_n$ of (1.1) is computed, and this has the form

$$\hat{y}_n = \sum_{i=1}^{k} \hat{c}_i y_{i,n} + p_n, \qquad \hat{c}_1 \neq 0.$$

In the early stages, the term $\hat{c}_1 y_{1,n}$ is very small and has a magnitude which depends on the accuracy to which the calculations are performed. The relative error, e_n, in $\bar{y}_n$ is given by

$$\begin{aligned} e_n &\equiv (\bar{y}_n - \hat{y}_n)/\bar{y}_n, \\ &= \left\{\sum_{i=1}^{k} (c_i - \hat{c}_i) y_{i,n}\right\} \Big/ \left\{\sum_{i=2}^{k} c_i y_{i,n} + p_n\right\}, \\ &= \left\{-\hat{c}_1 + \sum_{i=2}^{k} (c_i - \hat{c}_i) y_{i,n}/y_{1,n}\right\} \Big/ \left\{\sum_{i=2}^{k} c_i y_{i,n}/y_{1,n} + p_n/y_{1,n}\right\}. \end{aligned}$$

From (1.3) it follows that $|e_n| \to \infty$ as $n \to \infty$. From this we see that the relative error in our approximation to $\bar{y}_n$ increases without bound as $n \to \infty$. If $c_j, j > 1$, is the first non-zero coefficient appearing in (1.2) then the calculation,

using (1.4), of any $\bar{y}_n$ approximating the y_n of (1.2) will be liable to such a large accumulation of round-off error as to render the whole procedure totally ineffective. Gautschi (1967) has given some spectacular examples of such instability. Clearly there is a need for special methods for approximating y_n if any solution other than a dominant one is required. This has motivated recent research into the development of algorithms for computing non-dominant solutions of (1.1), which has been mainly directed at replacing the original initial value problem by a mathematically equivalent, well conditioned boundary value problem. Such a replacement may be done by abandoning an appropriate number, $k - m$, of the initial conditions and requiring instead that the solution should vanish at $k - m$ points sufficiently far outside the region of interest. It is very important to note that we do not require the solution which we are computing necessarily to tend to zero for large n. (Indeed, it is often possible using this approach to compute non-dominant solutions which increase without bound as n increases.)

The requirement that the desired solution should vanish outside the range of interest, in the absence of any further information, is a convenient computational procedure. However, if we do have some knowledge of how the required solution behaves for large n, such as via an asymptotic expansion, it is usually better to use this information to define new boundary conditions rather than simply to set them equal to zero. In certain circumstances it can be shown that the solution of this re-posed boundary value problem is a good approximation to the required solution of the original initial value problem over any given finite range providing the new boundary conditions are set "sufficiently far" outside the far end of this range. The case $k = 2$ of linear second order recurrence relations will be considered in the next chapter. There it will be shown that, before any solution values have been computed, the optimal positioning of the boundary conditions to obtain a given degree of precision over a specified range for any particular problem can be determined. For the case $k > 2$, however, the optimal positioning of the boundary values *ab initio* is still largely an unsolved problem. In order to recognise that direct methods of solution are unstable, and to be able to select an appropriate value for m, it is necessary in general to have some additional knowledge regarding the behaviour of the required solution. This knowledge need not necessarily be enough to enable the boundary conditions to be set more accurately, but it should be enough to enable the wrong solution to be detected during a computation. As a result, the theory for the case $k > 2$ is mainly of use in the field of ordinary differential equations. An important practical case is that in which a basis may be chosen so that $y_{i,n}$ dominates

$y_{i+1,n}$ for $i = 1, 2, \ldots, k-1$ as n increases, in the special sense that

$$|y_{i,n+1}/y_{i,n}| > |y_{i+1,n+1}/y_{i+1,n}| \qquad (i = 1, 2, \ldots, k-1; n = 1, 2, \ldots),$$

where $|y_{i+1,n}/y_{i,n}| \to 0$ as $n \to \infty$ for all i.

In this case it is usually possible to state that the re-posed problem does have a solution closely approximating that of the original problem (Oliver, 1968a). Furthermore for this situation, all the complementary functions may be ranked in a complete hierarchy of dominance with the basis solution $y_{i,n}$ being the i^{th} in the order. Equations for which we are able to pick out a dominant basis solution have important practical applications in the numerical solution of initial value problems for ordinary differential equations using linear multistep methods. The implications of this will be considered later. However, we first examine some of the analogies between the classical stability theory relating to linear multistep methods and certain parts of the basic theory of linear recurrence relations as outlined above.

1.3 Some stability aspects of linear multistep methods

We consider first of all the numerical solution, using a linear k-step method, of the system of first order ordinary differential equations

$$\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}x} = \mathbf{f}(x, \mathbf{y}), \tag{1.5}$$

with initial condition $\mathbf{y}(x_0) = \mathbf{y}_0$. The linear k-step method has the form

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h \sum_{j=0}^{k} \beta_j f_{n+j}, \tag{1.6}$$

where α_j and β_j are constants and where it is assumed that $\alpha_k \neq 0$ and that α_0 and β_0 are not both zero. Here we have used a notation consistent with common usage so that h is a positive constant called the steplength of integration, $\mathbf{y}_{n+j}$ is the approximate numerical solution obtained at the point $x_{n+j} = x_0 + (n+j)h$ and $\mathbf{f}_{n+j} = \mathbf{f}(x_{n+j}, \mathbf{y}_{n+j})$. Equation (1.6) will also be normalised in the usual way by assuming that $\alpha_k = 1$. If we now introduce the first and second characteristic polynomials, denoted by $\rho(\delta)$ and $\sigma(\delta)$ respectively, associated with (1.6) as

$$\rho(\delta) = \sum_{j=0}^{k} \alpha_j \delta^j,$$

$$\sigma(\delta) = \sum_{j=0}^{k} \beta_j \delta^j,$$