THE NUMERICAL SOLUTION OF STIFF DIFFERENTIAL EQUATIONS

G. J. COOPER
Edinburgh, Scotland

This paper first discusses the conditions in which a set of differential equations should give stable solutions, starting with linear systems assuming that these do not differ greatly in this respect from non-linear systems. Methods of investigating the stability of particular systems are briefly discussed. Most real biochemical systems are known from observation to be stable, but little is known of the regions over which stability persists; moreover, *models* of biochemical systems may not be stable, because of inaccurate choice of parameter values.

The separate problem of stability and accuracy in numerical methods of approximating the solution of systems of non-linear equations is then treated. Stress is laid on the consistently unsatisfactory results given by explicit methods for systems containing "stiff" equations, and implicit multistep methods are particularly recommended for this class of problem, which is likely to include many biochemical model systems. Finally, an iteration procedure likely to give convergence both in multistep methods and in the steady-state approach is recommended, and areas in which improvement in methods is likely to occur are outlined.

1. Introduction

A model of a biochemical system may be constructed from assumptions concerning the interaction and behaviour of the components of the system. The usual basic assumptions are that the system is adequately represented by an interdependent set of simple chemical equations, and that these equations obey rate laws of reaction kinetics. Suppose the system has n components, with concentrations at time t denoted by $y_i = y_i(t)$, $i = 1, 2, ..., n$, and denote the derivative of y_i with respect to t by y_i'. Then such a model leads to a set of differential equations, typically of the form

$$y_i' = f_i(y_1, y_2, ..., y_n) , \quad i = 1, 2, ..., n , \qquad (1.1)$$

each derivative being defined as a known function of the concentrations. Here the functions also depend on parameters (rate constants) which are assumed known. To complete the model, conditions are imposed at some fixed time t_0,

$$y_i(t_0) = y_{i_0} , \quad i = 1, 2, ..., n , \qquad (1.2)$$

defining an initial value problem.

Some models may give a set of differential equations of order greater than unity. However, such a set of equations may be reduced to another (larger) set of equations of the form (1.1) with appropriate initial conditions (1.2) [1, p. 12] *. Again, the independent variable, t, may occur explicitly, but can be removed by the addition of an extra differential equation, $y_0' = 1$, with $y_{00} = t_0$, so that $y_0 = t$.

The validity of such a model may be tested by obtaining a solution of (1.1) which satisfies (1.2). Such a solution gives concentration values for a range of values of t, which may be compared with experimental observations. Similarly, a model may be used to predict the behaviour of a system. Of particular interest is the determination of steady states. Here typical numerical methods of solution are examined, with consideration given to the special form of the functions in (1.1), and to difficulties that arise. Indeed, the natural time constants of biochemical systems vary so considerably that the differential equations that arise are often stiff. Then the usual numerical methods fail and recourse must be made to special techniques.

It is thought that the behaviour of stiff non-linear

* References to equations are in parentheses ( ).
References to the literature are in brackets [ ].

North-Holland Publishing Company – Amsterdam

systems is similar to that of stiff linear systems. In this paper, properties of stiff linear systems are examined prior to a treatment of non-linear equations. Then a few numerical methods are considered. The failure of the usual methods is explained and other methods are described which may be useful.

2. Linear differential equations

It is appropriate to start by considering the behaviour of systems of linear differential equations and, in particular, the single differential equation,

$$y' = \lambda(y - \alpha), \quad \alpha, \lambda \text{ given constants.} \qquad (2.1)$$

This equation has a solution $y = \alpha$, while the homogeneous equation, $y' = \lambda y$, has the *general* solution

$$y = \beta\phi(t), \quad \phi(t) = e^{\lambda t}, \quad \beta \text{ any constant.}$$

The general solution of (2.1) is thus $y = \alpha + \beta\phi(t)$ and a given initial condition defines β.

Consider the initial value problem defined by (2.1) and $y(0) = \alpha$, the solution being required for $t \geqslant 0$. Let the initial condition be perturbed by an amount ϵ, such that $y_0 = \alpha + \epsilon$. The solution of this new initial value problem is $y = \alpha + \epsilon\phi(t)$ whereas the solution required is $y = \alpha$. There are two cases:

(i) *Instability.* If $\lambda > 0$ the exponential term dominates the solution for sufficiently large t, irrespective of the magnitude of ϵ. This behaviour becomes more marked as λ increases. The *initial value problem* is unstable. That is, a small perturbation of the initial condition alters the solution drastically.

(ii) *Stability.* For $\lambda < 0$ the exponential term decreases in magnitude as t increases. A small perturbation of the initial condition has little effect on the solution. Nevertheless, this case presents formidable numerical problems when $\lambda \ll 0$.

The numerical problems that arise for both cases will be examined in more detail later. However, little can be achieved for the unstable case unless a detailed analysis of the particular problem is made, when it may become possible to reformulate the problem in a stable form.

The situation is much the same for the system of independent linear differential equations,

$$y_i' = \sum_{j=1}^{n} a_{ij}(y_i - \alpha_j), \quad i = 1, 2, ..., n,$$

where α_j, a_{ij}, $i, j = 1, 2, ..., n$, are given constants. This system may be written in matrix form as

$$\mathbf{y}' = \mathbf{A}(\mathbf{y} - \boldsymbol{\alpha}), \qquad (2.2)$$

where $\mathbf{y}$, $\mathbf{y}'$, $\boldsymbol{\alpha}$ are n dimensional column vectors and $\mathbf{A}$ a square $n \times n$ non-singular matrix. This system has a solution $\mathbf{y} = \boldsymbol{\alpha}$ and it remains to consider the general solution of the homogeneous equation $\mathbf{y}' = \mathbf{A}\mathbf{y}$.

It is known [1,6], that there exist n fundamental solutions $\Phi_1(t), \Phi_2(t), ..., \Phi_n(t)$, each an n dimensional column vector, such that

$$\mathbf{y} = \sum_{i=1}^{n} \beta_i \Phi_i(t), \quad \beta_1, \beta_2, ..., \beta_n \text{ any constants,}$$

is the general solution of the homogeneous equation. The simplest case occurs when $\mathbf{A}$ has n distinct eigenvalues, $\lambda_1, \lambda_2, ..., \lambda_n$, and this restriction does not alter the present arguments appreciably. Here

$$\Phi_i(t) = (c_{i1} e^{\lambda_i t}, c_{i2} e^{\lambda_i t}, ..., c_{in} e^{\lambda_i t}),$$

$$i = 1, 2, ..., n,$$

where $c_{i1}, c_{i2}, ..., c_{in}$ are the components of the eigenvector associated with the eigenvalue λ_i. The general solution of (2.2) may now be expressed as

$$y_j = \alpha_j + \sum_{i=1}^{n} \beta_i c_{ij} e^{\lambda_i t}, \quad j = 1, 2, ..., n, \qquad (2.3)$$

the constants $\beta_1, \beta_2, ..., \beta_n$ being determined by an initial condition $\mathbf{y}(t_0)$.

The values of the eigenvalues are crucial in assessing the behaviour of the linear system. If * $\mathrm{Re}(\lambda_i) < 0$, $i = 1, 2, ..., n$, then perturbations are damped out. If

* $\mathrm{Re}(\lambda_i) \equiv$ real parts of (λ_i).

not, then perturbations may be magnified. Both cases may cause numerical difficulties if $|\mathrm{Re}(\lambda_i)| \gg 0$ for some $i = 1, 2, ..., n$. Further, the imaginary parts of the eigenvalues introduce oscillations. The magnitudes of $|\lambda_1|, |\lambda_2|, ..., |\lambda_n|$ give a measure of the *stiffness* of the system.

The case $\mathrm{Re}(\lambda_i) > 0$ is relatively stable if the required solution is itself increasing exponentially. Here, although perturbations are magnified, the solution is also increasing in magnitude and the relative accuracy may be tolerable. The initial value problem *may* be (relatively) stable for $\mathrm{Re}(\lambda_i) > 0$.

The usual numerical methods become inadequate when they fail to approximate exponential functions (accurately) over a sufficiently large range of the argument.

It is worth noting that for linear systems it is possible to solve the eigenvector problem $\mathbf{Ac} = \lambda\mathbf{c}$, and then use the initial conditions to determine the constants $\beta_1, \beta_2, ..., \beta_n$ in (2.3). For large systems this requires considerable computation.

3. Non-linear differential equations

For systems of non-linear differential equations, much the same difficulties occur as for linear systems. In addition, problems of an entirely different character occur due to the non-linearity [5]. These problems are not considered here.

A rigorous analysis of non-linear systems is complicated. However, some progress may be made by replacing

$$y_i' = f_i(\mathbf{y}) = f_i(y_1, y_2, ..., y_n), \quad i = 1, 2, ..., n,$$

by the (identical) system

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}^*) + \mathbf{J}(\mathbf{y}^*)(\mathbf{y} - \mathbf{y}^*) + \Phi(\mathbf{y}, \mathbf{y}^*),$$

where $\mathbf{y}^* = \mathbf{y}(t^*)$, and the Jacobian matrix $\mathbf{J}(\mathbf{y}^*)$ is

$$\mathbf{J}(\mathbf{y}^*) = \left\{ \frac{\partial f_i(\mathbf{y}^*)}{\partial y_j} \right\} \tag{3.1}$$

The behaviour of the new linear system

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}^*) + \mathbf{J}(\mathbf{y}^*)(\mathbf{y} - \mathbf{y}^*)$$

depends on the eigenvalues of $\mathbf{J}(\mathbf{y}^*)$. This linearization is justified, at least for an investigation of the stability of the original system, if

$$|\Phi(\mathbf{y}, \mathbf{y}^*)| < c|\mathbf{y}| = c \max \{ |y_1|, |y_2|, ..., |y_n| \},$$

for some sufficiently small c [5, §7.1].

Many biochemical processes give models of particularly simple form

$$y_i' = \sum_{j=1}^{j=n} \left[\sum_{l=1}^{l=n} k_{ijl}\, y_l \right] (y_j - \alpha_j), \quad i = 1, ..., n, \tag{3.2}$$

where the k_{ijl}, $i, j, l = 1, ..., n$, are essentially rate constants. Since the solutions are functions of t, formally:

$$y_i' = \sum_{j=1}^{n} a_{ij}(t)\,(y_j - \alpha_j), \quad i = 1, 2, ..., n, \tag{3.3}$$

where the coefficients are unknown functions of t, depending on the solution. Again, if these coefficients can be represented as constant terms plus non-linear terms appropriately bounded by $|\mathbf{y}|$, then the behaviour of (3.2) depends on the eigenvalues of a constant matrix $\mathbf{A}(t^*)$.

This suggests a possible numerical scheme. At time t_0 compute the eigenvalues and vectors of $\mathbf{A}(t_0)$. Assume that $\mathbf{A}(t)$ remains constant over a (small) interval $t_0 \leqslant t \leqslant t_0 + h$ and compute the solution of (3.3) at $t_0 + h$ from (2.3). The process is then repeated and this provides a check on the assumption. Alternatively, the Jacobian matrix could be used to solve the linearized system. Iterative methods are available for improving the approximations [7]. For large systems of equations such schemes appear to demand too much computing time to be feasible.

There remains the possibility of evaluating the eigenvalues of $\mathbf{A}(t^*)$ or $\mathbf{J}(\mathbf{y}^*)$ to examine the stiffness of the system. This may be done for a few values of t^*, irrespective of the numerical method used. Indeed, eigenvalue *bounds* are of principal concern, and there are methods of use here [8, pp. 285–286].

A detailed analysis of systems of the form (3.2) appears worthwhile. In particular such an analysis might indicate the stability properties of the system as a function of the rate constants. It certainly appears

that biochemical systems give rise to stiff differential equations requiring special numerical methods.

4. Stability of numerical methods

Consider the differential equations (1.1),

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}) = \mathbf{f}(y_1, y_2, ..., y_n)\,, \qquad (4.1)$$

expressed in vector form. Assume $h \geqslant 0$ throughout. Let

$$t_m = t_0 + mh\,, \qquad m = 0, 1, ..., N\,,$$

be a sequence of (equally) spaced points. Let $\mathbf{y}_m$ be an approximation to $\mathbf{y}(t_m)$, $m = 0, 1, ..., N$.

$$\mathbf{y}_m = \mathbf{y}(t_m) + \boldsymbol{\varepsilon}_m = [y_{1m}, y_{2m}, ..., y_{nm}]'\,,$$

where $\mathbf{y}_0$ is the given (perturbed) initial condition. In the following sections numerical methods are discussed which provide the sequence $\mathbf{y}_0, \mathbf{y}_1, ...$. These methods determine y_m using previous members of the sequence.

It is convenient to distinguish (local) accuracy and stability. Assume that $\boldsymbol{\varepsilon}_0 = \boldsymbol{\varepsilon}_1 = ... = \boldsymbol{\varepsilon}_{m-1} = \mathbf{0}$. Then the accuracy of a method is a measure of $\boldsymbol{\varepsilon}_m$ (the discretization error), and a method is of order p if †

$$\boldsymbol{\varepsilon}_m = \mathrm{O}(h^{p+1})\,, \qquad \boldsymbol{\varepsilon}_0 = \boldsymbol{\varepsilon}_1 = ... = \boldsymbol{\varepsilon}_{m-1} = \mathbf{0}\,.$$

The stability concept is more difficult. Indeed, there are a number of definitions. Of main concern here, is the propagation of discretization errors. Assume that $\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1, ...$ are not zero. A method is (absolutely) stable if the components of $\boldsymbol{\varepsilon}_m$ are bounded for all m.

Accurate methods are required so that the solutions of a system of differential equations are adequately represented locally. Stable methods are required so that local discretization errors do not accumulate. Unfortunately, it is difficult to obtain stability conditions for finite h, and is usual to consider instead asymptotic stability conditions, $h \to 0, Nh$ constant. Asymptotically stable methods are often not satisfactory for stiff systems, as will be seen.

Asymptotic stability of a method, and stability of the initial value problem, are separate concepts and should be treated as such. Consider an unstable initial value problem. Here, no numerical method (stable or unstable) can be expected to give satisfactory results. On the other hand, consider a differential equation with one increasing solution. If the initial condition is such that this is the required solution, then a stable numerical method may give satisfactory results. (Relative stability may be more appropriate here.)

For stability with h finite, conditions can be obtained at least for simple cases. These conditions relate the step length to some property of the differential equations, usually the eigenvalues of an associated matrix or a Lipschitz constant.

A more detailed discussion is given by Henrici [6] who also discusses stability with respect to rounding errors.

5. Single step methods

Single step methods compute $\mathbf{y}_m$ from section (4.1) using only the approximation $\mathbf{y}_{m-1}$ determined at the previous step of the calculation. In particular, these methods are suitable for commencing calculation.

5.1. *Explicit methods*

At each stage of an explicit method the argument $\mathbf{z}$ of $\mathbf{f}(\mathbf{z})$ is known in advance. All that is required is the direct calculation of values

$$f_i(\mathbf{z}) = f_i(z_1, z_2, ..., z_n)\,, \qquad i = 1, 2, ..., n\,.$$

The *Euler method* is the simplest of this type. It is defined by

$$\mathbf{y}_m = \mathbf{y}_{m-1} + h\,\mathbf{f}(\mathbf{y}_{m-1})\,, \qquad m = 1, 2, ..., N\,,$$

and is of order one. That is, if $\boldsymbol{\varepsilon}_{m-1} = \mathbf{0}$,

$$\mathbf{y}_m = \mathbf{y}(t_{m-1}) + h\,\mathbf{y}'(t_{m-1}) = \mathbf{y}(t_m) + \mathrm{O}(h^2)\,.$$

Higher order (more accurate) methods are available, but *all* methods of this type are inadequate for stiff systems.

This failure is exhibited by an examination of

† The notation $\epsilon = \mathrm{O}(h)$ means ϵ is of the order of h.

Euler's method applied to

$$y' = \lambda(y - \alpha), \quad y(0) = \alpha, \quad t_0 = 0. \qquad (5.1)$$

Again, let $y_m = y(t_m) + \epsilon_m$, $m = 0, 1, ..., N$, with y_0 the perturbed initial condition. Euler's method gives

$$y_m = y_{m-1} + h\lambda(y_{m-1} - \alpha), \quad m = 1, 2, ..., N.$$

By repeated application, since $y_0 = \alpha + \epsilon_0$,

$$y_m = \alpha + (1 + h\lambda)^m \epsilon_0, \quad m = 0, 1, ..., N. \qquad (5.2)$$

Now $y(t)$ is the solution of the differential equation with initial condition $y(0) = \alpha$. Thus $y(t) = \alpha$,

$$y_m - y(t_m) = \epsilon_m = (1 + h\lambda)^m \epsilon_0 .$$

This identity describes the behaviour of the error ϵ_0 propagated through the method. There are three possibilities:

(i) If $\lambda > 0$, ϵ_m increases with m and for large m dominates the required solution α. This occurs irrespective of the magnitude of $h\lambda$ (finite), though it may not be serious if λ is small. The method reflects the behaviour of the differential equation.

(ii) If $\lambda < 0$ and $h\lambda < -2$, ϵ_m increases with m (even though the differential equation is stable). Large negative values of $h\lambda$ must be avoided.

(iii) If $\lambda < 0$ and $h\lambda > -2$, ϵ_m decreases, giving a (non-asymptotic) stability condition. If the differential equation is stiff, i.e., $\lambda \ll 0$, then an extremely small value for h must be chosen.

The cause of this behaviour is shown by (5.2). The Euler method represents exponential terms by polynomials:

$$e^{h\lambda} = (1 + h\lambda) + O(h\lambda)^2 .$$

This is satisfactory if $|h\lambda|$ is small *or* if $(1 + h\lambda)$ is small *when* the exponential term is small.

A similar analysis may be given for higher order explicit methods and for (general) systems of differential equations. The non-asymptotic stability condition takes the form

$$-c_1 < h\,\mathrm{Re}(\lambda_i) < c_2 \ (= 0), \quad i = 1, 2, ..., n,$$

where $\lambda_1, \lambda_2, ..., \lambda_n$ are the eigenvalues of a matrix associated with the differential equations, and c_1, c_2 (small) non negative constants associated with the method. Conditions should also be imposed on the imaginary parts of the eigenvalues since these parts introduce oscillating components into the solution.

Even if the eigenvalues are known, for stiff systems h must be chosen so small that computation is often impractical. If the eigenvalues are not known, and the system may be stiff, other methods should be used. It is *not* sufficient to compare results for different values of h.

The analysis applies even if the initial conditions are exact, for computation introduces rounding errors and thus a perturbed initial value problem. Thus explicit single step methods (and the Taylor series method), although asymptotically stable, fail for stiff systems unless h is so small that computation becomes excessive.

5.2. *Implicit methods*

If the argument, $\mathbf{z}$, of $\mathbf{f(z)}$ is not known in advance, then a set of (non-linear) equations must be solved at each step of the calculation. Despite this complication such implicit methods have considerable advantages.

The simplest method of this type (of order one) is defined by

$$\mathbf{y}_m = \mathbf{y}_{m-1} + h\,\mathbf{f}(\mathbf{y}_m), \quad m = 1, 2, ..., N.$$

Again consider equation (5.1). Since this is linear, the method gives a linear algebraic equation for $\mathbf{y}_m$, and proceeding as before

$$y_m = y_{m-1} + h\lambda(y_{m-1} - \alpha),$$

$$\epsilon_m = \epsilon_0(1 - h\lambda)^{-m} .$$

The errors decrease provided that $|1 - h\lambda| > 1$. Thus any value of $h\lambda$ in the complex plane, exterior to the unit circle with centre unity, gives a stable method. The magnitude of h is dictated primarily by accuracy requirements. For explicit methods the magnitude of h is dictated by accuracy *and* stability requirements.

However, accuracy considerations may require h so small that this advantage is lost. This is the case for example if $\lambda \gg 0$. Then higher order methods must be used. Such methods have been developed recently by Butcher [2] and *some* seem appropriate for stiff

differential equations. In particular consider the following method (of order four)

$$\mathbf{k}_1 = \mathbf{f}(\mathbf{y}_{m-1} + \tfrac{1}{4}h\mathbf{k}_1 + \tfrac{1}{12}h(3-2\sqrt{3})\mathbf{k}_2)\,,$$

$$\mathbf{k}_2 = \mathbf{f}(\mathbf{y}_{m-1} + \tfrac{1}{4}h\mathbf{k}_2 + \tfrac{1}{12}h(3+2\sqrt{3})\mathbf{k}_1)\,,$$

$$\mathbf{y}_m = \mathbf{y}_{m-1} + \tfrac{1}{2}h(\mathbf{k}_1 + \mathbf{k}_2)\,.$$

The vectors $\mathbf{k}_1$, $\mathbf{k}_2$ are defined implicitly and have to be determined iteratively. As in the previous method it is important to choose a *suitable* iteration. This will be considered later.

For equation (5.1) this method gives linear equations for k_1, k_2, and

$$\epsilon_m = \epsilon_0 \left\{1 + \frac{h\lambda}{1 - \tfrac{1}{2}h\lambda + \tfrac{1}{12}(h\lambda)^2}\right\}^m = \{\varphi(h\lambda)\}^m\,.$$

Thus the errors decrease for all $h\lambda < 0$ and the error growth is well controlled for $h\lambda \gg 0$. A more detailed examination requires the determination of the curves $|\varphi(h\lambda)| = c$, in the complex $h\lambda$ plane.

An analysis has only been given for equation (5.1) but this suggests that implicit methods are of some use for stiff systems and merit further investigation.

6. Multistep methods

A linear k step method for solving (4.1) is defined by the vector equations

$$\sum_{i=0}^{k} \alpha_{k-i}\,\mathbf{y}_{m-i} + h\sum_{i=0}^{k} \beta_{k-i}\,\mathbf{f}(\mathbf{y}_{m-i}) = \mathbf{0}\,,$$

$$m \geqslant k\,,$$

and k starting conditions $\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_{k-1}$ and required. A method is defined by a choice of the parameters α_i, β_i, $i = 0, 1, \ldots, k$, and methods are available which are (asymptotically) stable and of high order (accurate).

If $\beta_k = 0$ the method is explicit, and implicit if $\beta_k \neq 0$. Explicit methods are not suitable for stiff systems, for reasons similar to those given in section 5. Some recent investigations have been concerned with the development of implicit methods that are appropriate for such systems.

An examination of stability is more complicated than for single step methods. Again consider equation (5.1). Then a k step method gives a non-homogeneous linear difference equation

$$\sum_{i=0}^{k} \alpha_{k-i}\,y_{m-i} + h\lambda\sum_{i=0}^{k} \beta_{k-i}\,y_{m-i} = h\lambda\alpha\sum_{i=0}^{k} \beta_i\,. \qquad (6.1)$$

Let Z_m, $m = 0, 1, \ldots, N$, be a solution of (6.1). To Z_m may be added any linear combination of solutions of the homogeneous equation ($\alpha = 0$). If $\zeta_1, \zeta_2, \ldots, \zeta_k$ are distinct zeros of the polynomial

$$\rho(\zeta) + h\lambda\sigma(\zeta) = \sum_{i=0}^{k} \alpha_i\,\zeta^i + h\lambda\sum_{i=0}^{k} \beta_i\,\zeta^i\,, \qquad (6.2)$$

then $\zeta_1^m, \zeta_2^m, \ldots, \zeta_k^m$ satisfy the homogeneous equation, and the *general* solution of (6.1) is

$$y_m = Z_m + \sum_{i=1}^{k} \gamma_i\,\zeta_i^m\,, \qquad m = 0, 1, \ldots, N, \qquad (6.3)$$

where $\gamma_1, \gamma_2, \ldots, \gamma_m$ are arbitrary constants.

Before proceeding with the examination of stability a *consistency* condition must be imposed. If $\epsilon_0 = 0$ the solution of the differential equation is $y(t) = \alpha$. For the difference equation (6.1) to be consistent, it is required that $y_m = \alpha$ if $y_{m-1} = y_{m-2} = \ldots = y_{m-k} = \alpha$, and this implies the consistency condition $\alpha_0 + \alpha_1 + \ldots \alpha_k = 0$. Thus exact starting conditions $y_0 = y_1 = \ldots = y_{k-1} = \alpha$ give the solution $y_m = Z_m = \alpha$ of (6.1).

Now the substitutions

$$y_{m-i} = y(t_{m-i}) + \epsilon_{m-i} = \alpha + \epsilon_{m-i}\,,$$

$$i = 0, 1, \ldots, k\,,$$

in (6.1) show that the errors satisfy the homogeneous linear difference equation. Thus

$$\epsilon_m = \sum_{i=1}^{k} \gamma_i\,\zeta_i^m\,,$$

the constants $\gamma_1, \gamma_2, ..., \gamma_k$ being determined by the starting conditions $\epsilon_0, \epsilon_1, ..., \epsilon_{k-1}$. For stability (asymptotic or otherwise),

$$|\zeta_i| \leqslant 1 , \qquad i = 1, 2, ..., k . \tag{6.4}$$

If the zeros of (6.2) are not all distinct, the analysis requires modification. It turns out that *multiple* zeros should have moduli strictly less than unity.

For general systems of differential equations, a similar analysis (not rigorous) may be given. Here, the zeros of a set of polynomial equations

$$\rho(\zeta) + h\lambda_i\sigma(\zeta) = 0 , \qquad i = 1, 2, ..., n ,$$

are required, where $\lambda_1, \lambda_2, ..., \lambda_n$ are the eigenvalues of the Jacobian matrix of the differential equations. These eigenvalues are functions of the required solution so that stability conditions are required for all $h\lambda$ in some region of the complex plane.

In the case of asymptotic stability, however, only the zeros of $\rho(\zeta)$ are of interest. The coefficients α_0, $\alpha_1, ..., \alpha_k$ are chosen to ensure (6.4) while $\beta_0, \beta_1, ..., \beta_k$ are chosen to give methods of high order. Asymptotic stability is not alone sufficient for stiff systems, for then, if $|h\lambda_i|$, $i = 1, 2, ..., n$, are small, h must be so small that computation is impractical.

Implicit multistep methods have been developed recently which are suitable for stiff systems. Non-asymptotic stability is established for all values of $h\lambda$ in a restricted region of the complex plane. Dahlquist [3] considered all $h\lambda$ with negative real parts (the left half complex plane). He showed that the maximum order of such methods is $p = 2$. By further restricting the region of the complex plane considered, Widlund [9] has obtained methods of order $p \geqslant 4$. More recently, Gear [4] has obtained methods of order $p \leqslant 6$, stable for various regions, and these methods seem particularly useful.

7. Computation

It is desirable to have a number of methods available for solving a given system of differential equations, and to have some knowledge of the stiffness of the system together with estimates of the local discretization errors of the methods. Then it may be possible to select a method of appropriate order p and a suitable h, and to vary these choices during computation. Many difficulties arise and no entirely satisfactory procedures are available. Henrici [6] and Gear [4] consider some of these points.

A single step method may be used to start computation before progressing to the use of a multistep method. For stiff systems all methods used must be implicit. Thus, at each stage of the computation a set of non-linear equations has to be solved if the differential equations are non-linear. The techniques used are generally simple iterative methods. Difficulties occur for stiff systems because some iterative methods fail to converge, and even if convergence takes place a large number of iterations may be required to give adequately accurate solutions.

Consider the general k step method of section 6, and assume $\alpha_0 \neq 0$. This may be expressed as

$$\mathbf{y}_m = h\beta\mathbf{f}(\mathbf{y}_m) + \Phi , \qquad \beta = -\beta_0/\alpha_0 \neq 0 ,$$

where Φ is a known vector function of $\mathbf{y}_{m-1}, \mathbf{y}_{m-2}, ..., \mathbf{y}_{m-k}$. This is a set of n non-linear equations in n unknowns. In the present context, the usual iterative method is defined by

$$\mathbf{y}_m^{(r+1)} = h\beta\mathbf{f}(\mathbf{y}_m^{(r)}) + \Phi , \qquad r = 0, 1, ... ,$$

where $\mathbf{y}_m^{(0)}$ is an initial approximation to the required solution. It may be shown that this iteration converges if all $|h\lambda_i|$ are small, where $\lambda_1, \lambda_2, ..., \lambda_n$ are the eigenvalues of $\mathbf{J}(\mathbf{y})$. For stiff systems, however, this implies that h must be chosen *extremely* small.

However, an appropriate method is the Newton-Raphson iteration defined by

$$\mathbf{y}_m^{(r+1)} = \mathbf{y}_m^{(r)} - \{\mathbf{I} - h\beta\mathbf{J}(\mathbf{y}_m^{(r)})\}^{-1} \times \{\mathbf{y}_m^{(r)} - \Phi - h\beta\mathbf{f}(\mathbf{y}_m^{(r)})\} ,$$

where $\mathbf{I}$ is the identity matrix. This scheme converges if all eigenvalues of the inverse matrix are of moduli less than unity. If $\lambda_1, \lambda_2, ..., \lambda_n$ are the eigenvalues of $\mathbf{J}$ then

$$\frac{1}{1 - h\beta\lambda_i} , \qquad i = 1, 2, ..., n ,$$

are the eigenvalues of $(\mathbf{I} - h\beta\mathbf{J})^{-1}$, and the method is satisfactory for stiff systems. Such a scheme requires considerable computation at each step, but some reduction is possible [4].

Further investigation should lead to considerably superior algorithms for solving systems of the form (3.2). The Jacobian may be evaluated easily and there should be many refinements possible. In addition new types of method may become available. Hybrid methods (combined multistep and single step methods) hold promise.

In conclusion it is noted that the steady state equations derived from (1.1), $\mathbf{f}(\mathbf{y}) = \mathbf{0}$, may also be difficult to solve if the differential equations are stiff. Again an *appropriate* iterative method should be used.

Acknowledgement

The author thanks Prof. S.Michaelson for many helpful comments.

References

[1] J.C.Burkill, The Theory of Ordinary Differential Equations (Oliver and Boyd, Edinburgh, 1956).
[2] J.C.Butcher, Implicit Runge-Kutta processes, Math. Comp. 18 (1964) 50.
[3] G.Dahlquist, A special stability problem for linear multistep methods, B.I.T. 3 (1963) 27.
[4] C.W.Gear, The automatic integration of stiff ordinary differential equations, Proc. IFIP Congress 68 (North-Holland Publishing Company, Amsterdam, 1969) to be published.
[5] A.Halany, Differential Equations, Stability, Oscillations, Time Lags (Academic Press, New York, London, 1966).
[6] P.Henrici, in: Error in Digital Computation, ed. L.B.Rall, vol. 1 (Wiley, New York, 1965) p. 185.
[7] R.S.H.Mah, S.Michaelson and R.W.H.Sargent, Dynamic behaviour of multi-component multi-stage systems. Numerical methods for the solution, Chem. Eng. Sci. 17 (1962) 619.
[8] J.Todd (ed.), A Survey of Numerical Analysis (McGraw-Hill, New York, 1962).
[9] O.B.Widlund, A note on unconditionally stable linear multistep methods, B.I.T. 7 (1967) 65.