The effect of variable mesh size on the stability of multistep methods
/ by C. W. Gear [and K. W. Tu].
Gear, C. William (Charles William), 1935-
Urbana : Dept. of Computer Science, University of Illinois at Urbana-Champaign, 1973.

https://hdl.handle.net/2027/uuiuo.ark:/13960/t3dz1n819
The person charging this material is responsible for its return to the library from which it was withdrawn on or before the Latest Date stamped below.

Theft, mutilation, and underlining of books are reasons for disciplinary action and may result in dismissal from the University.

UNIVERSITY OF ILLINOIS LIBRARY AT URBANA-CHAMPAIGN

MAR 22 1978
DEC 14 1976
DEC 12 RECD

L161—O:1096
THE EFFECT OF VARIABLE MESH SIZE ON
THE STABILITY OF MULTISTEP METHODS

by

C. W. Gear
K. W. Tu

April 1973
THE EFFECT OF VARIABLE MESH SIZE ON
THE STABILITY OF MULTISTEP METHODS*

by

C. W. Gear
K. W. Tu

April 1973

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
URBANA, ILLINOIS 61801

* Supported in part by grant US AEC AT(ll-1)1469.
1. INTRODUCTION

This paper is concerned with the stability over the finite interval $t \in [0, T]$ of multistep formulas subject to varying step sizes. Although practical algorithms vary the step size in order to use as large a step as possible consistent with a local control of the estimated error, the existing literature--with the exceptions noted below--deals only with the stability of fixed step formulas.

The standard derivations of multistep formula, e.g. Henrici [3], are based on equal intervals. There are two common ways of handling variable steps for multistep formula. One is to interpolate through the computed points to get values at equally spaced points for use in a standard equal interval formula. The other is to derive multistep formula directly based on unequal intervals. These techniques do not have the same stability characteristics. In 1962 Nordsieck [4] noted an apparent instability in his interpolation version of Adams formula if the step size was varied too frequently. In 1969 Piowtrowski [5] proved that a variable step form of the Adams-Moulton formula was stable and convergent. Brayton, et.al. [1] and Gear [2] (pp. 145-146) point out that the two techniques for changing step size are not equivalent, and Brayton, et.al. give test examples that show that both can cause instability, but that the technique of step variation used by Piotrowski might be more stable.

The stability of a method is affected by three factors: the underlying multistep formula, the technique used to handle variable steps, and the scheme used to select step sizes—the formula, technique, and scheme for short. Method will be used to refer to the combination of a formula and a technique. This paper introduces the idea of a step selection scheme, and defines the stability and convergence of a method
with respect to a scheme. A stability condition is derived. It is necessary for stability, and, together with consistency, implies convergence and stability.

It is easy to generate consistent formulas. The more difficult task is to determine whether a combination of formula, technique, and scheme is stable. The two common techniques for step changing are examined, and it is proved that the variable step technique is superior to the interpolation technique for Adams formula. Specifically, it is proved that the variable step technique used with Adams formula is stable and convergent with respect to any step selection scheme, while the interpolation method can be unstable if the steps are allowed to change too much. However, it is proved that the interpolation technique with the k-step Adams formula is stable if the step selection scheme guarantees that the step is constant for at least k-steps after each change.

The investigation is extended to all multistep formulas, whether explicit, implicit, or predictor-corrector. It is proved that if the underlying formula is stable, then both step changing techniques lead to stable methods if the rate of change of the step size is small. It is also shown that common error control algorithms lead to small rates of change in the step size.
2. BACKGROUND

The two step changing techniques will be called interpolation and variable step, respectively. Intuitively, we define them as follows. Suppose we are using a k-step formula for integration and suppose that an approximation \( y_{n-1} \) to the solution is known at k values of the independent variable \( t \), say \( t_{n-i} \), \( k > i > 0 \). The interpolation technique for changing step size to \( h \) consists of the following two steps:

1. Interpolate through the known points to get the values of the solution at the points \( \hat{t}_{n-i} = t_n - ih \), \( k > i > 0 \).
2. Use the fixed step integration formula on these new points \( \hat{t}_{n-i} \) to find the value at \( t_{n+1} = t_n + h \).

In the variable step technique we start with \( k \) unequally spaced points \( t_{n-i} \), \( k > i > 0 \) and compute the coefficients of the multistep formula

\[
y_{n+1} = \alpha_1,n \ y_n + \ldots + \alpha_{k,n} \ y_{n-k+1} + \beta_{0,n} \ h_n \ y'_n + \ldots + \beta_{k,n} \ h_{n-k} \ y'_{n-k+1}
\]

so that it is exact to the required order, where \( h_n = t_{n+1} - t_n \). Since the order cannot exceed \( 2[k/2]+2 \) if the formula is to be stable for fixed step sizes, additional restrictions must be imposed. These consist of specifying the values of some of the coefficients \( \alpha \) and \( \beta \). The fixed step formula underlying the variable step technique is said to be the formula that is obtained from (1) when equally spaced points \( t_{n-i} \) are used. For example, if we required that \( \alpha_{j,n} = 0 \), \( j > 2 \), and that (1) have order \( k+1 \), then Adams-Moulton is the underlying fixed step formula. If we also ask that \( \beta_{0,n} = 0 \) and lower the order to \( k \), we get a variable step method based on the Adams-Bashforth formula.
The stability of an integration formula subject to a step changing technique may depend on the way in which the steps are varied. Therefore, we must define stability with respect to a particular step selection scheme. First, therefore, we state

**Definition 1** A step selection scheme is a function \( \theta \) such that

\[
h_n = h\theta(t_n, h)
\]

where, for all \( h > 0, 0 \leq t \leq T \)

\[
1 \geq \theta(t, h) \geq \Delta > 0
\]

**NOTE:** As \( h \) is reduced to zero, the maximum step size is reduced to zero. The lower bound \( \Delta \) serves to prevent zero step sizes for any non-zero \( h \); these would not give useful numerical integration schemes!

In this paper we will be concerned only with differential equations \( y' = f(y, t) \) for which \( f(y, t) \) is Lipschitz continuous in the strip \(|y| < \infty, 0 \leq t \leq T < \infty\). Small perturbations to the solutions of such equations at a given point cause small perturbations later on. Stability means that small numerical disturbances at one point \( t_n \) cause small perturbations in the numerical solution at later times for all small \( h \). Convergence means that the numerical solution can be made arbitrarily accurate as \( h \to 0 \).

**Definition 2** A method is stable with respect to a scheme \( \theta \) if there exists a constant \( M < \infty \) (dependent on the differential equation only) such that

\[
|y_m - \bar{y}_m| < M |y_n - \bar{y}_n|
\]

for all \( 0 \leq t_n < t_m \leq T \), where \( y_i \) and \( \bar{y}_i \) are two numerical solutions.
Definition 3  A method is convergent with respect to a scheme \( \theta \) if the computed solution \( y_n \) converges to \( y(t_n) \) for any \( 0 \leq t_n \leq T \) as \( h \to 0 \) and the starting errors tend to zero.

We will only discuss a single differential equation. The extension to systems of equations is straightforward, although notationally complex. We will also assume that \( f(y,t) \) and \( y(t) \) are differentiable as often as necessary to save stating all the conditions each time. The results are valid if \( f \) is only Lipschitz continuous, and can be obtained by replacing \( f_y = \partial f/\partial y \) by the Lipschitz constant \( L \) whenever \( f_y \) appears.
3. NOTATIONAL DEVELOPMENT

By developing a uniform notation for all the formulas and step changing techniques to be discussed, we will be able to give general proofs in Section 4 that can be applied to all cases. Some of this notation has been developed in [2]. It is summarized briefly here and suitably extended.

Consider the integration formula (1) with $\beta_{0,n} = 0$. It is an explicit formula which can be used directly. It can also be used as a predictor for a corrector formula. We will write the corrector formula as

$$y_{n+1}^{(m+1)} = \alpha_{l,n}^* y_n + \cdots + \alpha_{k,n}^* y_{n-k+1} + \beta_{0,n}^* h_n f(y_{n+1}^{(m)}, t_{n+1})$$

$$+ \beta_{l,n}^* h_{n-1} y_n' + \cdots + \beta_{k,n}^* h_{n-k} y_n'^{k+1}$$

A predictor-corrector formula is one in which a predicted value $y_{n+1}^{(0)}$ is computed by the explicit equation

$$y_{n+1}^{(0)} = \alpha_{l,n}^* y_n + \cdots + \alpha_{k,n}^* y_{n-k+1}$$

$$+ \beta_{l,n}^* h_{n-1} y_n' + \cdots + \beta_{k,n}^* h_{n-k} y_n'^{k+1}$$

and the equation (2) is used to compute a fixed finite number of corrected values $y_{n+1}^{(1)}, \ldots, y_{n+1}^{(M)}$. The approximation $y_{n+1}^{(M)}$ is defined as $y_{n+1}^{(M)}$ and $y_{n+1}^{(M)}$ is set to either $f(y_{n+1}^{(M)})$ or $f(y_{n+1}^{(M-1)})$ depending on whether or not a final function evaluation is needed.

An implicit formula is one in which equation (2) is solved for $y_{n+1}^{(m+1)} = y_n^{(m)}$. One way of doing this is to iterate a predictor-corrector formula until successive iterates are approximately equal. This converges for $h$ sufficiently small but may diverge for large $h$. Other methods, such as Newton-Raphson, can also be used to solve (2). The stability of a method based on an implicit formula is independent of the process used to solve equation (2), so we do not need to specify that process.
During the computation, a number of previous values of the function $y_{n-i}$ and the scaled derivatives $h_{n-i-1}y'_{n-i}$ must be saved. Let us collect these together in a column vector $y_n$ defined by 

$$y_n = [y_n, y_{n-1}, \ldots, y_{n-k+1}, h_{n-1}y', h_{n-2}y', \ldots, h_{n-k}y'_{n-k+1}]^T$$

for a $k$-step formula, where $T$ is the transpose operator. (If some of these values are not needed, we can drop them from the vector. For example, in Adams formula, $y_n$ would be $[y_n, h_{n-1}y', \ldots, h_{n-k}y'_{n-k+1}]^T$.)

A numerical integration method is a method for computing $y_{n+1}$ from $y_n$. We will first describe it for the variable step technique.

Let $y_{n+1}^{(m)}$ be defined as 

$$[y_{n+1}^{(m)}, y_n, \ldots, y_{n-k+2}, h_n y'_{n+1}^{(m)}, h_{n-1} y', \ldots, h_{n-k+1} y'_{n-k+2}]^T$$

where $h_n y'_{n+1}^{(m)}$ is $h_n f(y_{n+1}^{(m)}, t_{n+1})$ for $m \geq 1$,

$$h_n y_{n+1}^{(0)} = y_{1,n} y_n + y_{2,n} y_{n-1} + \ldots + y_{k,n} y_{n-k+1} + \delta_{1,n} h_{n-1} y' + \delta_{2,n} h_{n-2} y'_{n-1} + \ldots + \delta_{k,n} h_{n-k} y'_{n-k+1}$$

when $m = 0$, and the coefficients $\gamma$ and $\delta$ are given by

$$\gamma_{i,n} = (a_{i,n} - a_{i,n})/\beta_{0,n}$$

and

$$\delta_{i,n} = (\beta_{i,n} - \beta_{i,n})/\beta_{0,n}$$

From these definitions we can write

$$y_{n+1}^{(0)} = A_n y_n$$

(4)
where \( A_n = \)

\[
\begin{bmatrix}
\alpha_{1,n} & \alpha_{2,n} & \ldots & \alpha_{k-1,n} & \alpha_{k,n} \\
\beta_{1,n} & \beta_{2,n} & \ldots & \beta_{k-1,n} & \beta_{k,n} \\
1 & 0 & \ldots & 0 & 0 \\
0 & 1 & \ldots & 0 & 0 \\
& & \ldots & \ldots & \ldots \\
0 & 0 & \ldots & 1 & 0
\end{bmatrix}
\]

and

\[
\gamma_{1,n} \gamma_{2,n} \ldots \gamma_{k-1,n} \gamma_{k,n} \delta_{1,n} \delta_{2,n} \ldots \delta_{k-1,n} \delta_{k,n}
\]

\[
\begin{bmatrix}
\beta_{1,n} & \beta_{2,n} & \ldots & \beta_{k-1,n} & \beta_{k,n} \\
1 & 0 & \ldots & 0 & 0 \\
0 & 1 & \ldots & 0 & 0 \\
& & \ldots & \ldots & \ldots \\
0 & 0 & \ldots & 1 & 0
\end{bmatrix}
\]

\[
Y_{n+1} = Y_n + \gamma_n \left[ h_n f(y_{n+1}, t_{n+1}) - h_n y'_{n+1} \right]
\]

Equation (5) represents the predictor—the evaluation of \( y_{n+1} \) and \( y'_{n+1} \) as linear combinations of the known information plus the shifting over of the saved information and the discarding of \( y_{n-k+1} \) and \( y'_{n-k+1} \). Equation (5) represents the corrector—the evaluation of \( f(y_{n+1}, t_{n+1}) \) and the updating of \( y_{n+1} \) and \( y'_{n+1} \). For convenience, we define

\[
P_n(y_{n+1}) = h_n f(y_{n+1}, t_{n+1}) - h_n y'_{n+1}
\]
Then equation (5) can be written

\[ \nu_{n+1}^{(m+1)} = \nu_{n+1}^{(m)} + \ell_n^{-1} F_n \{ \nu_{n+1}^{(m)} \} \]  
\[ (8) \]

If a predictor only formula is used, we still formally must apply one step of (8) with \( \ell_n = \ell_n^P = [0, \ldots, 0, 1, 0, \ldots, 0]^T \). (In this case the value of \( a^* \) and \( b^* \) are unimportant.) If a final function evaluation is used, we apply (8) \( M \) times using the \( \ell_n \) given in (6), and then apply one further step of (8) using \( \ell_n = \ell_n^P \).

Now let us consider a fixed step method in preparation for formalizing the interpolation technique. In this case, the \( a, b, c, d \) and \( \delta \) in \( A_n \) and \( \ell_n \) are independent of \( n \). Suppose we have just completed the step from \( t_{n-1} \) to \( t_n \) using step \( h_{n-1} \). The values saved in \( y_n \) are the computed approximations at the points \( t_n - ih_{n-1}, 0 \leq i < k \). Before stepping to \( t_{n+1} = t_n + h_n \) we must interpolate to the points \( t_{n-i} = t_n - ih_n \). This is a linear process and can be represented by first premultiplying \( y_n \) by an interpolating matrix \( C_n \). Consequently, we write the predictor step for the interpolation process as

\[ y_{n+1}^{(0)} = A C_n y_n \]  
\[ (9) \]

The corrector step (8) is unchanged except that \( \ell_n = \ell_n^* \) is independent of \( n \).

Both for theoretical development and computational ease, it is convenient to perform a transformation based on Nordsieck's [4] form of Adams formula for the interpolation technique. This consists of computing and saving \( Q y_n \) rather than \( y_n \) where \( Q \) is a nonsingular matrix. The particular form of \( Q \) used is obtained as follows: The set of values of \( y_{n-i} \) and \( y'_{n-i}, 0 \leq i < k \) determine a unique \((2k-1)\)-th degree polynomial.
The equation (2) can be rewritten as:

$$\frac{1}{2}(2n) = (n)$$

In a matrix with one column to make an "x" matrix, we have:

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & n \end{pmatrix}$$

The determinant of this matrix is:

$$\text{det}(A) = \frac{1}{n}$$

This determinant is used in the calculation of the determinant of the inverse of the matrix.

The determinant of the inverse of a matrix is:

$$\text{det}(A^{-1}) = \frac{1}{\text{det}(A)}$$

The matrix inverse is:

$$A^{-1} = \frac{1}{\text{det}(A)} \cdot \text{adj}(A)$$

where \(\text{adj}(A)\) is the adjugate of the matrix.

The adjugate of the matrix is:

$$\text{adj}(A) = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}$$

The determinant of the inverse of the matrix is:

$$\text{det}(A^{-1}) = \frac{1}{\text{det}(A)}$$

The characteristic polynomial of the matrix is:

$$\det(A - \lambda I) = 0$$

The eigenvalues are:

$$\lambda_1, \lambda_2, \ldots, \lambda_n$$

The eigenvectors are:

$$v_1, v_2, \ldots, v_n$$

The eigenspace is:

$$\text{span}\{v_1, v_2, \ldots, v_n\}$$

The matrix is diagonalizable if:

$$A = PDP^{-1}$$

where \(D\) is a diagonal matrix and \(P\) is the matrix of eigenvectors.

The trace of the matrix is:

$$\text{tr}(A) = \sum_{i=1}^{n} \lambda_i$$

The determinant of the matrix is:

$$\text{det}(A) = \prod_{i=1}^{n} \lambda_i$$

The rank of the matrix is:

$$\text{rank}(A)$$

The nullity of the matrix is:

$$\text{null}(A)$$

The determinant is a scalar value that can be associated with a square matrix. It is a function that is defined only for square matrices.

The determinant of a matrix is an important concept in linear algebra. It is used in various applications such as solving systems of linear equations, finding the inverse of a matrix, and in the study of eigenvalues and eigenvectors.

The determinant of a matrix is defined as:

$$\text{det}(A) = \sum_{\sigma} \text{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}$$

where \(\sigma\) is a permutation of \(\{1, 2, \ldots, n\}\), \(\text{sgn}(\sigma)\) is the sign of the permutation, and \(a_{i\sigma(i)}\) is the entry in the \(i\)th row and \(\sigma(i)\)th column of \(A\).

The determinant is used in various applications such as solving systems of linear equations, finding the inverse of a matrix, and in the study of eigenvalues and eigenvectors.
Represent that polynomial instead by its value at $t_n$ and the values of its first $2k-1$ scaled derivatives $h_{n-1}^j y^{(j)}/j!$, $1 \leq j < 2k$. (This leads to what is called a $2k$-value method in Gear [2].) If $y_n$ is the set of saved values of $y$ and $h y'$ based on step $h_{n-1}$, and $a_n$ is the set of values of $[y_n, h_{n-1} y'_n, h_{n-1}^2 y''/2, \ldots, h_{n-1}^{2k-1} y^{(2k-1)}/(2k-1)!]^T$ representing the same polynomial of degree $2k-1$, then $Q$ is defined by $a_n \equiv Q y_n$. $Q$ is independent of $y$ and $h_{n-1}$. If we apply this transformation to (9) and (8) we get

\[ Q y_{n+1}^{(0)} = (QAQ^{-1})(QC_n Q^{-1})Q y_n \]  
\[ Q y_{n+1}^{(m+1)} = Q y_{n+1}^{(m)} + Q Q_n F_n (Q^{-1} (Q y_n^{(m)})) \]  

To simplify the presentation we will write $y_n$ rather than $a_n$, $A$ rather than $QAQ^{-1}$, $C_n$ rather than $QC_n Q^{-1}$, $\xi_n$ rather than $Q \xi_n$, and $F(c)$ rather than $F(Q^{-1}c)$ when we are talking about the transformed scheme. Consequently, equations (10) and (11) are identical to (9) and (8). The advantage of making the transformation is that

\[ C_n = \text{diag}[1, h_n/h_{n-1}, \ldots, (h_n/h_{n-1})^{2k-1}] \]  

so that the interpolation process is simplified. This was Nordsieck’s motivation. Finally, we note that (9) is identical to (4) if $A_n$ is interpreted as $A C_n$ for the interpolation technique. Consequently, equations (4) and (8) represent the predictor only formula or the prediction-correction formula in either step changing technique.

The corrector only formula can sometimes be implemented by iterating the corrector to convergence and treated theoretically by setting $y_{n+1} = \lim_{m \to \infty} y_{n+1}^{(m)}$. However, this process may not converge (as in stiff equations), in which case the corrector must be solved by other
means. We note that the corrector equation (8) always adds on a multiple of $\frac{h}{n}$ to the current value. Consequently, the final value has the form

$$V_{n+1} = V_{n+1}^{(0)} + \frac{h}{n} = A_n V_n + \frac{h}{n}$$

where $\omega$ is a scalar. The value of $\omega$ is such that

$$F(Y_{n+1}^{(0)} + \frac{h}{n}) = 0$$

(Although we have used predictor coefficients in $A_n$, some straightforward arithmetic will show that the final value of $V_{n+1}$ is independent of these coefficients when the corrector equation is solved by (13) and (14).)

We must now develop the error propagation formula for equations (4) and either (6) or (13) and (14). The global error $\varepsilon_n$ is defined by

$$\varepsilon_n = V_n - V(t_n)$$

where $V(t_n)$ is the value of the components of the vector $V_n$ as evaluated along the desired solution of the differential equation. The truncation error $d_n$ is the error introduced by one step of the method, starting from the solution of the differential equation. Thus, $d_n$ is defined by the equations

$$\varphi_n^{(0)} = A_n V(t_n)$$

$$\varphi_n^{(m+1)} = \varphi_n^{(m)} + \frac{h}{n} F_n(\varphi_n^{(m)})$$

either

$$\varphi_n^{(M)}$$

if $M$ corrector iterations are used, or

$$\varphi_n^{(0)} = \varphi_n^{(0)} + \frac{h}{n} = A_n V(t_n) + \frac{h}{n},$$

(16)
\[ N + \frac{1}{2} \beta \cdot A = \left[ \frac{A}{B} \right] \cdot \left[ \frac{0}{B+1} \right] = \left[ \frac{0}{B} \right] \]

\[ 0 \cdot \left[ \frac{A}{B} \right] = \left[ \frac{0}{B} \right] \]

\[ N + \frac{1}{2} \beta \cdot A = \left[ \frac{A}{B} \right] \cdot \left[ \frac{0}{B+1} \right] = \left[ \frac{0}{B} \right] \]

\[ 0 \cdot \left[ \frac{A}{B} \right] = \left[ \frac{0}{B} \right] \]

\[ 0 \cdot \left[ \frac{A}{B} \right] = \left[ \frac{0}{B} \right] \]

\[ N + \frac{1}{2} \beta \cdot A = \left[ \frac{A}{B} \right] \cdot \left[ \frac{0}{B+1} \right] = \left[ \frac{0}{B} \right] \]

\[ 0 \cdot \left[ \frac{A}{B} \right] = \left[ \frac{0}{B} \right] \]
where \( F(\vec{y}_{n+1}) = 0 \) if the corrector is solved exactly, and

\[
\frac{d}{n} = \vec{y}_{n+1} - \vec{y}(t_{n+1})
\]

From the definition of \( \varepsilon_{n+1} \) and \( \frac{d}{n} \)

\[
\varepsilon_{n+1} = \vec{y}_{n+1} - \vec{y}(t_{n+1})
= \vec{y}_{n+1} - \vec{y}_{n+1} + \vec{y}_{n+1} - \vec{y}(t_{n+1})
= \vec{y}_{n+1} - \vec{y}_{n+1} + \frac{d}{n}
\]  

(Note that if round-off errors are to be considered, they can be included in \( \frac{d}{n} \) at this point, so that the computation of \( \vec{y}_{n+1} \) can be considered to be exact.) If \( M \) corrector iterations are used, (17) can be written as

\[
\varepsilon_{n+1} = \vec{y}(M)_{n+1} - \vec{y}(M)_{n+1} + \frac{d}{n}
= \vec{y}(M)_{n+1} + \frac{\partial F}{\partial y} \varepsilon (M)_{n+1} - \vec{y}(M-1)_{n+1} - \vec{y}(M-1)_{n+1} + \frac{d}{n}
= \left[ I + \frac{\partial F}{\partial y} \right] \varepsilon (M)_{n+1} - \vec{y}(M-1)_{n+1} + \frac{d}{n}
\]

by the mean value theorem. \( \frac{\partial F}{\partial y} \) is the row vector of partial derivatives of \( F_n \) with respect to the components of \( \vec{y} \) evaluated somewhere in the interval \((\vec{y}_{n+1}, \vec{y}_{n+1})\). From equation (7) we see that

\[
\frac{\partial F}{\partial y} = \left[ h_0, f(1), 0, \ldots, 0, -1, 0, \ldots, 0 \right]
\]

where \( f(1) \) is \( \partial f/\partial y \) evaluated in the interval \((\vec{y}_{n+1}, \vec{y}_{n+1})\). The process leading to (18) can be repeated to get

\[
\varepsilon_{n+1} = \left[ I + \frac{\partial F}{\partial y} \right] \cdots \left[ I + \frac{\partial F}{\partial y} \right] \left[ \vec{y}(0) - \vec{y}(0) + \frac{d}{n} \right]
= \left[ I + \frac{\partial F}{\partial y} \right] \cdots \left[ I + \frac{\partial F}{\partial y} \right] A_n \varepsilon_n + \frac{d}{n}
\]
We write this in the form
\[ \mathbf{e}_{n+1} = \mathbf{S}_n \mathbf{e}_n + \mathbf{d}_n \]  
(20)

Let \( \mathbf{e}_i^T \) be the unit row vector consisting of a one in the \( i \)-th position and zeros elsewhere. Thus, \( \mathbf{e}_i^T \) is \([1, 0, \ldots, 0] \). Let subscript \( d \) represent the component corresponding to the \( h_{n-1} y_n' \) entry in \( y_n \). (\( d \) is \( k+1 \) for the untransformed method, 2 for the transformed method.) Then
\[ \frac{\partial f^{(m)}}{\partial y} = h_n f_y \mathbf{e}_i^T - \mathbf{e}_d^T \]  
(21)

Therefore, from (19), (20), and (21)
\[ \mathbf{S}_n = \prod_{m=1}^{M} \left[ (I - \frac{1}{n} \mathbf{e}^T_d) + h_n f_y \frac{1}{n} \mathbf{e}_i^T \right] \mathbf{A}_n \]  
(22)

\[ = \left[ (I - \frac{1}{n} \mathbf{e}^T_d) \right]^M + \sum \text{Products of } M \text{ terms of form} \]
\[ h_n f_y \frac{1}{n} \mathbf{e}_i^T \text{ or } (I - \frac{1}{n} \mathbf{e}^T_d) \]
\[ \text{with at least one of the former} \right] \mathbf{A}_n \]

We define
\[ \mathbf{S}_n = (I - \frac{1}{n} \mathbf{e}^T_d)^M \mathbf{A}_n \]  
(23)

Then we can write (22) as
\[ \mathbf{S}_n = \mathbf{\hat{S}}_n + h_n \mathbf{S}_n \]  
(24)

where \( \mathbf{\hat{S}}_n \) is a matrix whose elements are polynomials in \( h_n, f_y^{(m)} \) and entries from \( \mathbf{A}_n \) and \( \frac{1}{n} \).
Lemma 1 If α, β, α* and β* are uniformly bounded or if $A_n$ and $\ell_n$ are uniformly bounded so is $||\tilde{S}_n||$ for all $|h| \leq u$.

Proof The elements of $\tilde{S}_n$ are polynomials in $h_n$ whose coefficients are polynomials in $f_y^{(m)}$ and the entries of $A_n$ and $\ell_n$. $|f_y^{(m)}|$ is bounded by the Lipschitz constant, while entries in $A_n$ and $\ell_n$ are uniformly bounded or if $A_n$ and $\ell_n$ are uniformly bounded so is $||\tilde{S}_n||$ for all $|h| \leq u$.

Proof The elements of $\tilde{S}_n$ are polynomials in $h_n$ whose coefficients are polynomials in $f_y^{(m)}$ and the entries of $A_n$ and $\ell_n$. $|f_y^{(m)}|$ is bounded by the Lipschitz constant, while entries in $A_n$ and $\ell_n$, except possibly for the d-th row of $A_n$, are bounded if $a$, $b$, $a*$ and $b*$ are bounded. The d-th row of $A_n$ contains elements of the form $(a-a*)/b*_{0,n}$, and could be unbounded if $b*_{0,n}$ can approach 0. However, from (22) and (24), the only way the d-th row of $A_n$ enters $S_n$ is when $A_n$ is premultiplied by $(I - \ell_n e_d^T)$ which zeros the d-th row and changes the first row from $a$ to $a*$. Consequently, the coefficients of $h_n$ in $\tilde{S}_n$ are bounded, so for any $u > 0$, $||\tilde{S}_n||$ is bounded for all $0 < h \leq u$. If $A_n$ and $\ell_n$ are uniformly bounded, the result follows similarly.

Q.E.D.

The behavior of the error $e_n$ is determined by equation (20). When $||\tilde{S}_n||$ is bounded, we can use equation (24) to relate the stability of the method to the properties of $\tilde{S}_n$. The following condition will be shown to be necessary for a method to the stable:
Definition 4 A method satisfies the stability condition with respect to a step selection scheme if there exists a \( C < \infty \), such that

\[
\beta_n^m = \beta_{m-1} \beta_{m-2} \ldots \beta_n
\]

satisfies

\[
||\beta_n^m|| \leq \epsilon
\]

for all \( 0 \leq t_n < t_m \leq T \).

If an implicit method is used, the corrector is solved exactly, and we must derive equations (20) and (24) by a different technique. Substitute (13) and (16) in (17) to get

\[
\varepsilon_{n+1} = A_n \varepsilon_n + (\omega - \bar{\omega}) \frac{\partial F}{\partial \varepsilon} + d_n
\]

(The subscript \( n \) has been omitted from \( \varepsilon \) for clarity later.) We know

\[
0 = F(y_{n+1}) - F(y_{n+1})
\]

\[
= \frac{\partial F}{\partial y} (y_{n+1} - y_{n+1})
\]

\[
= \frac{\partial F}{\partial y} (A_n \varepsilon_n + (\omega - \bar{\omega}) \varepsilon)
\]

Hence,

\[
(\omega - \bar{\omega}) = \left( \frac{\partial F}{\partial y} \right)^{-1} \left( \frac{\partial F}{\partial y} A_n \varepsilon \right)
\]

Note that both quantities enclosed in braces are scalars. Substituting in (27) we get

\[
\varepsilon_{n+1} = A_n \varepsilon_n - \left( \frac{\partial F}{\partial y} \right)^{-1} \left( \frac{\partial F}{\partial y} A_n \varepsilon \right) + d_n
\]

\[
= (I - \left( \frac{\partial F}{\partial y} \right)^{-1} \left( \frac{\partial F}{\partial y} A_n \varepsilon \right) A_n \varepsilon + d_n
\]

\[
= S_n \varepsilon + d_n
\]
(\text{where } a, b, c, d \text{ are constants})
where

\[ S_n = \left( I - \frac{\partial F}{\partial y_1} \right)^{-1} \frac{\partial F}{\partial y_n} A_n \]

\[ = \left( I - \frac{\partial h}{\partial y_1} f_y - \ell_d \right)^{-1} \frac{\partial (h_n f_y e^T_{y1} - e^T_d)}{\partial y_n} A_n \]

Here, \( \ell_i \) is the i-th component of \( \ell \).

If we choose \( u \) such that for \( 0 < h_n < u, h_n \ell f_y - \ell_d \) is bounded away from zero (\( \ell_d \) is one in the methods discussed), then we can write

\[ \left( h_n \ell f_y - \ell_d \right)^{-1} \text{ as } 2^{-1} \{ 1 + \ell h_n f_y \ell / \ell_d + 0(h_n^2) \}. \]

Hence,

\[ S_n = \left( I - \ell_d^{-1} y \frac{\ell}{\ell_d} \right) A_n + h_n \tilde{S}_n \]

\[ = \tilde{S}_n + h_n \tilde{S}_n \]

The form of \( \tilde{S}_n \) is the same as the previous form with \( M = 1 \) and \( \ell \) scaled so that \( \ell_d = 1 \), as it always is in methods discussed in this paper.

We also need to extend Lemma 1 to this case.

**Lemma 1a** \( \| \tilde{S}_n \| \) defined in (28) is bounded if \( \alpha, \beta, \alpha^* \) and \( \beta^* \) or \( A_n \) and \( \ell_n \) are uniformly bounded.

**Proof** From (28) and the preceding equation

\[ \tilde{S}_n = -\left( \ell h f_y \ell \right)^{-1} \frac{\ell}{\ell_d} \frac{\ell}{\ell_d} A_n \]

\[ + \left[ \left( \ell h f_y \ell \right)^{-1} + \left( \ell_d^{-1} \ell \frac{\ell}{\ell_d} A / h_n \right) \right] \frac{\ell D}{\ell_d} A_n \]

\[ = -\left( \ell h f_y \ell \right)^{-1} \frac{\ell}{\ell_d} \frac{\ell}{\ell_d} A_n \]

\[ + \ell D^{-1} \left( \ell h f_y \ell \right)^{-1} \ell f y \ell - \ell D^{-1} A_n \]
All but possibly the d-th row of $A_n$ are bounded. The first term does not involve the d-th row of $A_n$, and is, therefore, bounded for $0 \leq h \leq u$. The second term involves the d-th row of $A_n$, but scales it by $\lambda = \beta^*_0,n$. Hence, it is also bounded for $0 \leq h \leq u$. Hence, $||\hat{S}_n||$ is bounded. If $A_n$ and $\lambda_n$ are bounded, the result follows directly.

Q.E.D.
4. STABILITY AND CONVERGENCE

Theorem 1  If a method is stable with respect to a step selection scheme $\theta$, it satisfies the stability condition (26) with respect to that scheme.

Proof  Consider the differential equation $y' = 0$ for which $\hat{S}_n$, defined by (24), is zero. Evidently, the difference between two numerical solutions satisfies

$$y_{n+1} - \hat{y}_{n+1} = \hat{S}_n(y_n - \hat{y}_n)$$

If the $|S_n^m|$ are not uniformly bounded, there exist $t_m$ and $t_n$ and bounded $y_n - \hat{y}_n$ such that

$$y_m - \hat{y}_m = S_n^m(y_n - \hat{y}_n)$$

is arbitrarily large, implying that the method is not stable with respect to $\theta$. Hence stability implies the stability condition.

Q.E.D.

The convergence theorem depends on the following lemma.

Lemma 2  If, for the difference equation,

$$x_{n+1} = \hat{S}_n x_n + h\hat{S}_n x_n + h_n x_n$$

there exist constants $k_0$, $k_1$, and $k_2$ such that
\[ ||\tilde{S}_n|| \leq k_1 \quad \forall n, \]
\[ ||\lambda_n|| \leq k_2 \quad \forall n, \]

and
\[ ||\tilde{S}_m|| \leq k_0 \quad \forall m \geq n \]

where
\[ \tilde{S}_n = 1 \]

and \( \tilde{S}_n^m \) is defined by equation (25),

then, if \( k_1 > 0 \)
\[ ||x_n|| \leq [k_0 ||x_0|| + \frac{k_2}{k_1}] e^{k_0 k_1 (t_n - t_0)} - \frac{k_2}{k_1} \]  \hspace{1cm} (30)

or if \( k_1 = 0 \)
\[ ||x_n|| \leq k_0 ||x_0|| + k_0 k_2 (t_n - t_0) \]  \hspace{1cm} (31)

**Proof** From (29) we have
\[ x_n = \tilde{S}_0 x_0 + \sum_{n=0}^{N-1} \tilde{S}_n^{n+1} h_n (\tilde{S}_n x_n + \lambda_n) \]

Therefore,
\[ ||x_n|| \leq k_0 ||x_0|| + \sum_{n=0}^{N-1} k_0 h_n (k_1 ||x_n|| + k_2) \]  \hspace{1cm} (32)

Consider the case \( k_1 > 0 \) first. We show (30) by induction. It is evidently true for \( N = 0 \) since \( k_0 \geq ||\tilde{S}_n|| = 1 \). Assuming its validity for \( n < N \), we substitute in (32) to get
\[ \|x_N\| \leq k_0 \|x_0\| + \sum_{n=0}^{N-1} k_0 k_1 h_n [k_0 \|x_0\| + \frac{k_2}{k_1}] e^{k_0 k_1 (t_n - t_0)} \]

\[ = [k_0 \|x_0\| + \frac{k_2}{k_1}] [1 + \sum_{n=0}^{N-1} k_0 k_1 h_n e^{k_0 k_1 (t_n - t_0)}] - \frac{k_2}{k_1} \]

\[ \leq [k_0 \|x_0\| + \frac{k_2}{k_1}] [1 + \sum_{n=0}^{N-1} (e^{k_0 k_1 h_n - 1}) e^{k_0 k_1 (t_n - t_0)}] - \frac{k_2}{k_1} \]

\[ = [k_0 \|x_0\| + \frac{k_2}{k_1}] e^{k_0 k_1 (t_N - t_0)} - \frac{k_2}{k_1} \]

This completes the inductive proof of (30). If \( k_1 = 0 \), equation (31) follows from (32) directly.

Q.E.D.

The convergence theorem is now straightforward.

**Theorem 2**  If a method satisfies the stability condition with respect to a step selection scheme \( \theta \), if \( a, a^*, L \) and \( \beta^* \) or \( A_n \) and \( L_n \) are uniformly bounded, and if the truncation error \( d_n \) satisfies

\[ \|d_n\| \leq h_n e(h) \quad n \]

where \( e(h) \rightarrow 0 \) as \( h \rightarrow 0 \), then the method is stable and converges with respect to \( \theta \), and there exist constants \( k_3 \) and \( k_4 \) such that the global error \( e_N \) satisfies

\[ \|e_N\| \leq k_3 \|e_0\| + k_4 e(h) \text{ for } 0 \leq t_N \leq T \]

Notes: 1. The condition on \( d_n \) is a consistency condition. If the method has order \( r \), then \( e(h) = O(h^r) \). However, a more useful way of looking at this result is that if we control the step size \( h_n \) so that the...
truncation error is bounded by $\varepsilon h_n$, the global error is bounded by $k_3 \|\varepsilon_0\| + k_4 \varepsilon$, so that changes in the truncation error control are proportionately reflected in the global error bound.

2. The condition that $a, a*, \beta$ and $\beta^*$ or $A_n$ and $\ell_n$ be uniformly bounded is also a form of consistency. It rules out the following type of case:

The explicit formula

$$y_{n+1} = \frac{1}{h_{n-2}h_{n-1}} \left[ \left( \frac{h_n + h_{n-1}}{h_{n-2} + h_{n-1}} \right)^2 \frac{h_n}{h_{n-2} + h_{n-1}} y_n + \left( \frac{h_n + h_{n-1}}{h_{n-2} + h_{n-1}} \right)^2 \frac{h_n}{h_{n-2} + h_{n-1}} y_{n-2} \right. \right.$$  

$$- \frac{h_n (h_n + h_{n-1} + h_{n-2})}{h_{n-2}} y_n \right]$$

is second order provided that $h_{n-2} \neq h_{n-1}$. However, as $h_{n-2} \to h_{n-1}$, the coefficients and the error term blow up. If we used this formula with a step selection scheme which kept $h_{n-2}/h_{n-1}$ bounded away from one, we would get convergence if the method was stable, but if the step selection scheme let $h_{n-1}/h_{n-2} \to 1$ as $h \to 0$, we may not get convergence.

**Proof** From (20) and (24) the error equation for a method is given by

$$\varepsilon_{n+1} = S_n \varepsilon_n + d_n$$  

$$= \hat{S}_n \varepsilon_n + h_n \hat{S}_n \varepsilon_n + d_n$$

Consistency and Lemma 1 imply that $\|\hat{S}_n\| \leq k_1$ for some $k_1 < \infty$ provided that $h_n \leq u$. Lemma 2 can be applied to (33) with $k_2 = e(h)$ since the stability hypothesis guarantees the existence of a constant $k_0$ such that $\|\hat{S}_n^m\| \leq k_0$.
Therefore,

\[
||\varepsilon_n|| \leq [k_0 ||\varepsilon_0|| + \frac{e(h)}{k_1} e^{k_0 k_1 (t_n - t_0)} - \frac{e(h)}{k_1}] 
\]

Hence, \( ||\varepsilon_n|| \to 0 \) if \( ||\varepsilon_0|| \) and \( e(h) \to 0 \) and we have convergence.

The difference between two numerical solutions satisfies

\[
\varphi_{n+1} - \hat{\varphi}_{n+1} = S_n (\varphi_n - \hat{\varphi}_n) 
\]

The same analysis (with \( e(h) = 0 \)) yields

\[
||\varphi_n - \hat{\varphi}_n|| \leq k_3 ||\varphi_n - \hat{\varphi}_n|| 
\]

Therefore, the method is stable.

Q.E.D.

The difficult problem left is to verify the stability condition for various methods and step selection schemes. First, we will deal with the Adams methods in Theorems 3, 4, 5, and 6. Theorem 5 is an extension of the result of Piotrowski [5].

**Theorem 3** \( A_n \) and \( \frac{1}{n} \) are bounded for the interpolation technique using a step selection scheme 0.

**Proof** For the interpolation technique \( \frac{1}{n} \) is independent of \( n \), and \( A_n = A C_n \) where \( C_n \), given by equation (12), is bounded by \( A^{-2k+1} \) from definition 1. Therefore, \( A_n \) is bounded.

Q.E.D.
Theorem 4 \( \alpha, \alpha^*, \beta \) and \( \beta^* \) are bounded for the variable step Adams method using a step selection scheme \( \theta \).

Proof The variable step Adams method has the form
\[
y_{n+1} = y_n + \beta_{0,n} h_n y_{n+1} + \ldots + \beta_{k,n} h_{n-k} y_{n-k+1} \tag{35}
\]
where the \( \beta_{i,n} \) are determined from the requirement that the method have order \( k+1 \). (For an explicit method, \( \beta_{0,n} = 0 \) and the order is \( k \).) A necessary and sufficient condition for the order to be \( k+1 \) is that (35) be exact for the functions \( y(t) = (t - t_{n+1})^r \), \( 0 \leq r \leq k+1 \). This leads to a system of \( k+1 \) linear equations for the \( \beta_{i,n} \), namely
\[
\sum_{i=0}^{k} \beta_{i,n} h_{n-i} r(t_{n+1-i} - t_{n+1})^{r-1} = -h_n^r, \ r = 1, 2, \ldots, k+1 
\]
Define
\[
\omega_i = \frac{t_{n+1-i} - t_{n+1}}{h_n + h_{n-1} + \ldots + h_{n+1-i}} 
\]
From definition 1
\[
\Delta \leq |\omega_i| < i, \quad i \geq 1 \tag{37}
\]
\[0 = \omega_0
\]
and
\[
|\omega_i - \omega_j| \geq \Delta \quad \text{if} \ i \neq j \tag{38}
\]
Divide (36) by \( rh^r \) to get
\[
\sum_{i=0}^{k} \beta_{i,n} (\omega_i - \omega_{i+1}) \omega_i^{r-1} = -\frac{\omega_i^r}{r} \tag{39}
\]
The right hand side of (39) and the coefficients of $\beta^*_{i,n}$ on the left hand side are bounded uniformly in $n$ by (37). The determinant $D$ of the coefficients is a simple multiple of the Vandermonde determinant, namely

$$D = \prod_{k^2+i>j>0} (\omega_i - \omega_j) \prod_{i=0}^{k} (\omega_i - \omega_{i+1})$$

(40)

By (38) this is bounded away from zero, hence the solution of (39) is bounded uniformly in $n$.

Q.E.D.

**Theorem 5** The variable step Adams-Bashforth (AB), Adams-Bashforth-Moulton predictor-corrector (ABM), and Adams-Moulton (AM) methods are stable and convergent for any step selection scheme.

**Proof** The vector of saved values $y_n$ is $[y_n, hy'_n, \ldots, hy'_{n-k+1}]^T$ so the matrix $A_n$ is given by

$$A_n = \begin{bmatrix}
1 & \beta_{1,n} & \beta_{2,n} & \ldots & \beta_{k,n} \\
0 & \delta_{1,n} & \delta_{2,n} & \ldots & \delta_{k,n} \\
0 & 1 & \ldots & 0 \\
\vdots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & 1 & 0
\end{bmatrix}$$

while $\hat{y}_n$ is $[0, 1, 0, \ldots, 0]^T$ for AB and $[\beta^*_{0,n}, 1, 0, \ldots, 0]^T$ for ABM and AM.
Hence, from equation (23),

\[ \hat{S}_n = (I - \frac{e_n}{n} e_n^T)^M A_n \]

where

\[ \eta_{i,n} = \begin{cases} \beta_{i,n} & \text{for } AB \\ \beta_{i,n}^* & \text{for } ABM \text{ and } AM \end{cases} \]

From Theorem 4, \( \eta_{i,n} \) are uniformly bounded, therefore, \( \hat{S}_n \) is uniformly bounded. Hence the \( ||\hat{S}_n^{n+m}|| \) are uniformly bounded for \( 0 \leq m \leq k \). From the structure of \( \hat{S}_n \), we see that for \( m > k \)

\[ \hat{S}_n^{n+m} = \hat{S}_n^{n+k} = \begin{bmatrix} 1 & x_1 & x_2 & \cdots & x_k \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \]  

Thus, the three conditions of Theorem 2 are satisfied.

Q.E.D.
Theorem 6  If the step selection scheme is such that at least $k$ steps of constant size are taken between step changes, then the $k$-step interpolation $AB$, $ABM$, and $AM$ methods are stable and convergent.

Proof  Once again, we show that the hypothesis of Theorem 2 are satisfied. We will use the Nordsieck vector to express the method as a $k+1$ value method. Then the matrix $S_n$ is given by

$$\hat{S}_n = [I - \frac{h}{\Delta} e_2^T] A C_n$$

$$= R C_n$$

The matrix $R$ has the important properties that the simple eigenvalue one corresponds to the right eigenvector $e_1$, and that all other eigenvalues are zero (Gear [2], pp. 138-140). Consequently, $R^k x = \mu_x e_1$ where $|\mu_x|/||x||$ is uniformly bounded for all $x$ (Tu [6], p. 20). The eigenvalue one of $C_n$ corresponds to the eigenvector $e_1$. When $h_n = h_{n-1}$, $C_n = I$. From the hypothesis of the theorem, it is not possible for $C_i \neq I \neq C_j$ if $|i-j| < k$. Consequently, the product

$$\hat{S}_m = R C_{m-1} R C_{m-2} \ldots R C_n$$

either contains less than $2k+2$ matrices, or can be reduced to

$$\hat{S}_m = R C_{m-1} \ldots R C_q R^k C_{q-k} R^{q-k-n}$$

where

$$0 \leq q-k-n < k$$

Since the ratio of adjacent steps is bounded by $A^{-1}$, $||C_m||$ is uniformly bounded. Hence, $||C_{q-k} R^{q-k-n}|| \leq k_2$ uniformly. Therefore, if $\hat{S}_m$ contains $2k+2$ or more matrices
\[ \frac{S^m}{n} z = \sum_{R} C_{m-1} \ldots R C_{q} R^k C_{q-k} \ldots = R C_{m-1} \ldots R C_{q} R^k x \]

where \( ||x|| \leq k_5 ||z|| \).

Hence,

\[ \frac{S^m}{n} z = \sum_{R} C_{m-1} \ldots R C_{q} \mu_x e_{1} \]

so that

\[ ||\frac{S^m}{n} z|| = ||\mu_x|| \leq k_6 ||x|| \quad \text{(as } \mu_x/||x|| \text{ is bounded)} \]

\[ \leq k_5 k_6 ||z|| \]

Therefore,

\[ ||\frac{S^m}{n}|| \leq k_5 k_6 \]

If \( \frac{S^m}{n} \) contains less than 2k+2 matrices, \( ||\frac{S^m}{n}|| \) is bounded by the finite product of the norms of each matrix, hence the method satisfies the stability condition.

Theorem 3 shows that \( A_n \) and \( \hat{A}_n \) are bounded. \( \tilde{A}_n \) is bounded by the same argument as used in Theorem 5. Hence Theorem 2 proves convergence and stability.

Q.E.D.

Tu [6] points out that the interpolation form of the Adams formula is not necessarily stable under any step selection scheme. He shows that for the three step ABM formula with

\[ h_{2i} = \omega h_{2i+1} = h_{2i+2} = \omega h_{2i+3} = \ldots \]
the matrix $S_{2i}^{2i+2m}$ has an eigenvalue of $[(\omega-1)^2/4\omega]^m$. For sufficiently large $\omega$ this is unbounded. Examples of instability when $\omega = 10$ are given in that thesis along with examples showing that the variable step technique is stable for the same step selection scheme. This case points out that the variable step technique is more stable in some cases.

Another example in Tu's report substantiates Brayton, et al.'s claim that the variable step technique is more stable. The problem $y' = -y$ with $h_{2i} = .05$ and $h_{2i+1} = .005$ was integrated using a three step backward differentiation formula. It was apparently stable for the variable step technique and apparently unstable for the interpolation technique. It is hypothesized that the variable step method can also cause instabilities, but no examples have yet been found.

In practice, we do not vary the step in an extreme way. When the step changes are controlled to be small, both techniques are stable as shown below.

\textbf{Theorem 7} If a method satisfies the stability condition for fixed steps, it also does so with respect to a step selection scheme that produces step size changes small in the sense that

$$\frac{h_{n+1}}{h_n} = 1 + O(h) \text{ as } h \to 0$$

(44)

\textbf{Note:} If $\theta(t, h) \equiv \theta(t)$ and $\theta$ is a boundedly differentiable function of $t$, then the hypotheses of the theorem are satisfied since

$$\frac{h_{n+1}}{h_n} = \frac{\theta(t_{n+1})}{\theta(t_n)} = 1 + h_n \frac{\theta'(\xi_n)}{\theta(t_n)}$$

$$= 1 + h \theta'(\xi)$$

$$= 1 + O(h)$$
Proof. We will use the hypotheses of the theorem to decompose \( \hat{S}_n \) into the form

\[
\hat{S}_n = \hat{S} + h_n \overline{S}_n
\]  

(45)

where \( \hat{S} \) is the form of \( \hat{S}_n \) when constant steps are used and \( \overline{S}_n \) can be bounded. (This step will be delayed until Lemma 3 below.) Then we can examine \( ||\hat{S}_n^m|| \) as follows. For fixed \( n \), set

\[
x_m = \hat{S}_n^m x_n \quad m \geq n
\]  

(46)

Then

\[
x_{m+1} = \hat{S}_m x_m
\]

\[
= \hat{S} x_m + h_m \overline{S}_m x_m
\]  

(47)

If the constant step method satisfies the stability condition, then \( ||\hat{S}_n^{m-n}|| \leq k_0 \). Lemma 3 will show that \( ||\overline{S}_n^m|| \leq k_1 \). Therefore, Lemma 2 can be applied to (47) with \( k_2 = 0 \). Hence,

\[
||x_m|| = ||\hat{S}_n^m x_n|| \leq k_0 e^{k_0 k_1 T} ||x_n||
\]

Hence,

\[
||\hat{S}_n^m|| \leq k_0 e^{k_0 k_1 T}
\]

so the method satisfies the stability condition.

Q.E.D.

Lemma 3. If a step selection scheme satisfies the hypothesis of Theorem 7, \( \hat{S}_n \) can be written as \( \hat{S} + h_n \overline{S}_n \) where \( \hat{S} \) is independent of \( h \) and \( n \), and \( ||\overline{S}_n^m|| \) is uniformly bounded.
Proof

Part I: Interpolation Technique

From (9) and (23)

$$\hat{S}_n = (I - \frac{\ell}{n}e_2^T)^M A C_n$$

where $\frac{\ell}{n}$ is independent of $n$. Now

$$C_n = \text{diag}(1, h_{n}/h_{n-1}, (h_{n}/h_{n-1})^2, ...)$$

$$= I + \text{diag}(0, h_{n}/h_{n-1} - 1, (h_{n}/h_{n-1})^2 - 1, ...)$$

$$= I + O(h)$$

$$= I + O(h_n)$$

since $h_{n}/h_{n-1} = 1 + O(h)$ and $h \leq h_{n}^{-1}$.

Therefore,

$$\hat{S}_n = (I - \frac{\ell}{n}e_2^T)^M A(I + O(h_n)) = \hat{S}[I + O(h_n)]$$

Hence,

$$\tilde{S}_n = \hat{S} \frac{O(h_n)}{h_n}$$

is uniformly bounded.

Part II: Variable Step Technique

It is shown in Tu [6], p. 52, that the coefficients $\alpha$ and $\beta$ have the form

$$\alpha_{i,n} = \alpha_i + R_{\alpha}(\nu, \ldots, \nu_{n-k+2})$$

$$\beta_{i,n} = \beta_i + R_{\beta}(\nu, \ldots, \nu_{n-k+2})$$

where

$$\nu_i = \frac{h_i - h_{i-1}}{h_{i-1}}$$
where the R's are rational functions that vanish at $v = v_{n-1} = \ldots = v_{n-k+2} = 0$
(when the step sizes are equal), and where the $\alpha_i$ and $\beta_i$ are the corresponding
fixed step coefficients. (This result follows directly from the nonsingular
system of equations that determine $\alpha$ and $\beta$.) From the hypothesis of
Theorem 7, $v = O(h)$, hence $R = O(h) = O(h_n)$. Therefore, we can express $S_n$
which involves the coefficients $\alpha$, $\beta$, $\alpha^*$ and $\beta^*$, as $S + h_n S_n$, where $S$ is the
matrix obtained by setting the $\alpha_{i,n}$ equal to $\alpha_i$, etc., and $S_n$ includes the
terms $R$ divided by $h$. These can be uniformly bounded.

Q.E.D.
5. PRACTICAL IMPLICATIONS

It can be pointed out that a common step selection scheme satisfies the hypotheses of Theorem 7. If the step $h_n$ is controlled so that the local error estimate is equal to $\epsilon (or \epsilon h_n)$, and if the truncation error estimate is $\phi(t_n) h_n^{r+1} + O(h_n^{r+2})$ where $\phi(t)$ is the principle error function for the method, we have

$$\phi(t_n) h_n^{r+1} + O(h_n^{r+2}) = \epsilon (or \epsilon h_n)$$

Hence,

$$h_n = \left( \frac{\epsilon}{\phi(t_n)} \right)^{1/q} + O(h_n^2)$$

where $q = r+1$ or $r$, provided that $\phi(t_n)$ is not zero. If $\phi(t)$ is bounded away from zero and $\phi'(t)$ is bounded above, then

$$\frac{h_{n+1}}{h_n} = \left( \frac{\phi(t_n)}{\phi(t_{n+1})} \right)^{1/q} + O(h_n^2)$$

$$= 1 + O(h_n)$$

and $h_{\min}/h_{\max} > \Delta > 0$. Consequently, we can expect either method to be stable if the fixed step method is stable.

However, in practice we do not send $h$ to the limit and we often do not bother to adjust the step if the change is small. Therefore, there is reason to prefer the variable step approach. Brayton, et.al. [1] point out that the amount of work in evaluating $A_n \psi_n$ is less in the variable step approach than in the interpolation/Nordsieck approach, but that the coefficients in $A_n$ and $\psi_n$ must be recomputed for each step. For a large system of equations, this is not important as $A_n$ is the same for each equation in the system. For a few equations this is a serious consideration, which leads us to recommend variable step methods for large systems of equations and interpolation methods with the Nordsieck vector for small systems.
LIST OF REFERENCES

[1] Brayton, R. K., Gustavson, F. G., and Hachtel, R. D. "A New Efficient Algorithm for Solving Differential-Algebraic Systems Using Implicit Backward Difference Formulas," Proceedings of the IEEE, 60, #1, pp. 98-108, 1972.

[2] Gear, C. W. NUMERICAL INITIAL VALUE PROBLEMS IN ORDINARY DIFFERENTIAL EQUATIONS, Prentice-Hall, Inc., New Jersey, 1971.

[3] Henrici, P. DISCRETE VARIABLE METHODS FOR ORDINARY DIFFERENTIAL EQUATIONS, Wiley, New York, 1962.

[4] Nordsieck, A. "On the Numerical Integration of Ordinary Differential Equations," Math. Comp., 16, pp. 22-49, 1962.

[5] Piotrowski, P. "Stability, Consistency and Convergence of Variable K-Step Methods," Conference on the Numerical Solution of Differential Equations, 109, Springer, Berlin, 1969.

[6] Tu, K. W. "Stability and Convergence of General Multistep and Multivalue Methods with Variable Step Size," Department of Computer Science Report 526, University of Illinois at Urbana-Champaign, July 1972 (Ph.D. Thesis).
1. AEC REPORT NO.  
   C00-1469-0220  
2. TITLE  
   THE EFFECT OF VARIABLE MESH SIZE ON THE STABILITY OF MULTISTEP METHODS  

3. TYPE OF DOCUMENT (Check one):  
   a. Scientific and technical report  
   □ b. Conference paper not to be published in a journal:  
      Title of conference  
      Date of conference  
      Exact location of conference  
      Sponsoring organization  
   □ c. Other (Specify)  

4. RECOMMENDED ANNOUNCEMENT AND DISTRIBUTION (Check one):  
   a. AEC's normal announcement and distribution procedures may be followed.  
   □ b. Make available only within AEC and to AEC contractors and other U.S. Government agencies and their contractors.  
   □ c. Make no announcement or distribution.  

5. REASON FOR RECOMMENDED RESTRICTIONS:  

6. SUBMITTED BY: NAME AND POSITION (Please print or type)  
   C. W. Gear  
   Professor and Principal Investigator  
   Organization  
   Department of Computer Science  
   University of Illinois  
   Urbana, Illinois 61801  
   Signature  
   Date April 1973  

7. AEC CONTRACT ADMINISTRATOR'S COMMENTS, IF ANY, ON ABOVE ANNOUNCEMENT AND DISTRIBUTION RECOMMENDATION:  

8. PATENT CLEARANCE:  
   □ a. AEC patent clearance has been granted by responsible AEC patent group.  
   □ b. Report has been sent to responsible AEC patent group for clearance.  
   □ c. Patent clearance not required.
The effects of two different techniques for implementing variable mesh sizes in multistep formulas are investigated. It is proved that one is more stable than the other for some cases, but that both are stable when the step changes are small. The practical implications of these results are discussed.
