NUMERICAL METHODS FOR STIFF NONLINEAR AND
QUADRATIC DIFFERENTIAL EQUATIONS

W. Liniger

International Business Machines Corporation
T. J. Watson Research Center
P.O. Box 218
Yorktown Heights, N.Y. 10598

25 April, 1977

Final Scientific Report for Period 1 March 1975-28 February 1977

Contract No. F44620-75-C-0058

Approved for public release; distribution unlimited

Sponsored by the Air Force Office of Scientific Research
AFSC, United States Air Force
Numerical Methods for Stiff Nonlinear and Quadratic Differential Equations.

Final Scientific Report for period March 1, 1975-Feb. 28, 1977

The report summarizes the results obtained in each of these four areas, and gives references to the full accounts appearing in various research papers.
This is the final technical report covering the research carried out under this contract from March 1, 1975, through February 28, 1977. The work may be divided into these four categories:

I. Nonlinear and linear stability of multistep formulas,

II. Special high-order A-stable methods,

III. Fractional linear difference schemes for systems of quadratic differential equations, and

IV. Rotational solutions of the Josephson phase (Sine-Gordon-) equation.

Here we summarize our results obtained in these areas. A full account is given in the research papers listed in Section V and referred to hereafter by letters in brackets. Some of these papers are enclosed with this report and others are in preparation. Numbers in brackets refer to the list of references (Section VI).
I. NONLINEAR AND LINEAR STABILITY OF MULTISTEP FORMULAS

A. Nonlinear Input-output Stability of Conventional Multistep Formulas

If a linear multistep formula (LMF) is applied to a nonlinear system of differential equations, \( \dot{y} = f(y) \), then the numerical solution \( \{x_n\} \), \( n = 0,1,\ldots \), is defined by a nonlinear difference equation,

\[
N(x_n) = 0;
\]

here \( x_n \) approximates \( y_n = y(t_n) \), where \( y(t) \) is an exact solution of the differential system, \( t_n = nh, h > 0, n = 0,1,\ldots \), and

\[
N(x_n) = \sum_{j=0}^{k} a_j x_{n+j} - h \sum_{j=0}^{k} b_j f(x_{n+j}).
\]

If by \( d_n = -N(y_n) \) we denote the local truncation error, then the global error, \( e_n = x_n - y_n \), satisfies \( N(e_n) = d_n \), where \( N(e_n) = N(y_n + e_n) - N(y_n) \). The difference equation (1) is said to be input-output stable if

\[ ||e_n|| < K ||d_n|| \]

in some appropriate norm \( ||\cdot|| \), where \( K \) is a constant independent of \( n \). In order for the concept of input-output stability to be relevant for stiff systems of differential equations, the step-size \( h \) must be thought of as fixed and finite.

We have proved several theorems giving sufficient conditions for input-output stability of difference equations for various types of nonlinearities, notably for \( f \)'s satisfying the monotonicity (dissipativity) condition

\[ < f(y+e) - f(y), e > \leq \mu |e|^2 \]

for all \( y \) and \( e \) in some appropriate set and for some \( \mu < 0 \); here \( <,> \) is an
arbitrary scalar product and $|e|^2 = <e,e>$. The condition (3) implies stability (dissipation) of the differential system. We have, for example, proved the following result:

**Theorem 1:** Consider the well-known polynomials $\rho(\zeta) = \sum_{j=0}^{k} \alpha_j \zeta^j$, $\sigma(\zeta) = \sum_{j=0}^{k} \beta_j \zeta^j$ associated with the LMF and let $\Gamma$ denote its "root-locus curve", i.e., $\Gamma = \{q| q = q(\zeta), q(\zeta) = \rho(\zeta)/\sigma(\zeta), |\zeta| = 1\}$. Then, if the roots $\sigma_j$ of $\sigma(\zeta)$ satisfy $|\sigma_j| < 1$, $j = 1,\ldots,k$, and if $\mu_h < m$, where $m$ is the deepest incursion of $\Gamma$ into the left half of the $q$-plane ($m = \min |\text{Re}q(\zeta)|$, then we have input-output stability both in the $\ell_2$- and in the maximum norm.

The second condition of Theorem 1 may be interpreted by saying that the differential system has an amount of dissipation, measured by $\mu_h$, sufficient to offset the lack of A-stability [1] of the LMF, measured by $m$.

Other input-output stability theorems have been obtained for:
- nonlinearities $f$ which are monotone only for $|e| > B > 0$, $B$ sufficiently large,
- monotone $f$'s with sufficiently large negative $\mu$,
- nonlinearities $f$ which are gradients of concave, scalar functions (implying that the Jacobian matrix $\partial f/\partial y$ is symmetric and its spectrum is real),
- nonlinearities $f$ satisfying a Lipschitz condition, and
- nonlinear systems composed of loosely coupled subsystems, each of which falls into one of the previously mentioned categories.
B. Nonlinear Stability of Formulas with Second Derivatives

The nonlinear input-output stability analysis for conventional LMF described in the preceding paragraph has been extended to the LMF involving second derivatives:

\[ \sum_{j=0}^{k} a_{j} x_{n+j} - h \sum_{j=0}^{k} b_{j} \dot{x}_{n+j} + h^2 \sum_{j=0}^{k} c_{j} \ddot{x}_{n+j} = 0. \]  

(4)

The results apply to stable nonlinear systems \( \dot{x} = f(x) \) satisfying the monotonicity conditions

- \( h < f(x+e) - f(x), \quad e > -\mu_1 |e|^2, \quad (\mu_1 > 0) \)
- \( h^2 < g(x+e) - g(x), \quad e > \mu_2 |e|^2, \quad (\mu_2 > 0) \)
- \( < f(x+e) - f(x), \quad g(x+e) - g(x) > 0, \)

where \( g = f_{x}f \) and \( f_{x} \) is the Jacobian matrix \( \mathcal{J}f/\mathcal{J}x \).

First, simple sufficient conditions were derived for input-output stability of the nonlinear difference operator associated with the differential system and the integration formula. Rewrite (4) in the form

\[ \sum_{j=0}^{k} a_{j} x_{n-j} - h \sum_{j=0}^{k} b_{j} \dot{x}_{n-j} + h^2 \sum_{j=0}^{k} c_{j} \ddot{x}_{n-j} = 0 \]  

(5)

where \( a_j = a_{k-j}, b_j = b_{k-j}, c_j = \gamma_{k-j}, j = 0,1,\ldots,k, \) \( n = k, k+1,\ldots, \) and

let \( r(z) = \sum a_j z^j, s(z) = \sum b_j z^j, \) \( t(z) = \sum c_j z^j \) correspond to \( \rho(z), \) \( j=0 \)

condition is that
\[
\frac{1}{A+B\mu_1 + \mu_2} \left\{ \max_{|z|=1} \left| \frac{z^{-1}}{t(z^{-1})} A \right| \right\} + \frac{1}{B} \left\{ \max_{|z|=1} \left| \frac{z^{-1}}{t(z^{-1})} B \right| \right\} < 1, \quad (6)
\]

where A and B are free positive parameters which are at one's disposal and which are to be suitably chosen.

Second, the criterion (6) was applied to the A-stable formulas of Enright [2] with \(k=1, p=3\) and with \(k=2, p=4\), respectively. A computer search procedure was used to find values of A, B for which (6) is satisfied for various amounts of dissipation, i.e., various values of \(\mu_1\) and \(\mu_2\). For example, Enright's formula with \(k=1, p=3\),

\[
x_n - x_{n-1} - \frac{h}{3} (2\dot{x}_n + \dot{x}_{n-1}) + \frac{h^2}{6} x_n = 0, \quad (7)
\]

was found to be input-output stable with a moderate amount of dissipation:

\[
\mu_1 = \mu_2 = 1.3.
\]

C. Lyapunov-Stability of Conventional Multistep Formulas [c]

If \(\{x_n\}\) and \(\{x_n+z_n\}\) are any two solutions of the difference equation (1), then \(\{z_n\}\) satisfies the variational difference equation

\[
N(x_n) - N(x_n+z_n) = 0.
\]

The LMF giving rise to this variational equation is said to be Lyapunov-stable if, in some appropriate norm \(||\cdot||\) defined by a Lyapunov function, \(||z_n||\) \(\to 0\). It is said to be \((G,\mu)\)-stable in the sense of Dahlquist if for any \(f\) satisfying the monotonicity condition (3) it is Lyapunov-stable with respect to a norm defined by a positive definite quadratic form

\[
\sum_{i,j=0}^{k-1} \sum_{i,j=1}^{k-1} g_{ij} z_{n+k-i} z_{n+k-j}, \quad \text{where } G=(g_{ij}) \text{ is a positive}
\]


definite real symmetric $k \times k$-matrix. The LMF is said to be $G$-stable if it is $(G,0)$-stable. $A$-stability is necessary for its non-linear analogue, $G$-stability.

By a simple, explicit construction of a quadratic Liapunov function we proved the following results:

**Theorem 2:** Any member of the four-parameter family of all three-step ($k = 3$) LMF which are second-order accurate ($p = 2$) is $G$-stable if and only if it is $A$-stable.

Any three-step ($k = 3$), third-order ($p = 3$) LMF whose polynomial $\phi(\xi)$ satisfies assumption 1 of Theorem 1, as well as certain other, easily verifiable algebraic constraints between its coefficients, is $(G,\mu)$-stable for any $\mu$ below an explicitly given, negative bound which depends on the choice of the formula. For example, the formula

$$-10x_n + 42x_{n+1} - 78x_{n+2} + 46x_{n+3} - h(f_n - f_{n+1} - f_{n+2} + 25f_{n+3}) = 0,$$

a "neighbor" of the three-step backward differentiation formula, is $(G,\mu)$-stable for $\mu \leq -36/31$.

A survey of work on the stability (mostly linear) and accuracy of numerical methods for stiff differential equation [d] was presented by the principal investigator at the workshop on stiff differential equations held at the Air Force Weapons Laboratory at Kirtland Air Force Base, N.M., on May 6 and 7, 1976.

A survey of recent work on nonlinear stability of integration methods is in preparation [e].
II. SPECIAL HIGH-ORDER A-STABLE METHODS

A. Averaging [f]

High-order, A-stable numerical solutions to stiff systems of differential equations may be generated by a) calculating several A-stable solutions of order of accuracy one or two, each of which is produced by a particular member of a family of LMF depending on intrinsic parameters, and b) forming a suitable "weighted average" (linear combination) of these solutions. The recipe for forming high-order averages of low-order LMF solutions is found by considering an asymptotic expansion of the global truncation error as a function of the formula parameters.

An efficient procedure was developed for the systematic computation, to high orders in h, of an asymptotic expansion of the global truncation error of an LMF. This technique was applied for finding, to $O(h^6)$, the error expansion for the four-parameter class of all three-step, second-order LMF. From a two-parameter subclass of this class, A-stable (in fact, G-stable) methods of orders 4, 5, and 6 have been derived which require the averaging of as few as 2, 3, and 4 LMF solutions, respectively. These methods have been programmed in APL and successfully tested on linear and mildly non-linear stiff systems. The success of the numerical tests hinged partly on the use of a suitable starting procedure we developed for the step-by-step solution of the difference equations.

B. A-stable Integration Formulas with Second Derivatives [g]

Whereas for conventional linear multistep formulas (LMF) A-stability
is incompatible with orders of accuracy \( p > 2 \) [1], there do exist A-stable LMF involving second derivatives, i.e., formulas of the type (4) of as high an order of accuracy as four [2,3]. A criterion was developed for testing A-stability of formulas of class (1), which also served as a basis for the \textit{a priori} construction of a special class of A-stable formulas of this type. This criterion states that if i) the formula (4) is \( A_v \)-stable * , i.e., the roots \( \tau_i \) of the polynomial \( \tau(\xi) = \sum_{j=0}^{k} \gamma_j \xi^j \) satisfy

\[
|\tau_i| < 1, \quad i=1,\ldots,k; \text{ and ii) under the two-valued map } \xi \rightarrow q, \text{ defined by the characteristic equation}
\]

\[
\varrho(\xi) - q\varphi(\xi) + q^2\tau(\xi) = 0,
\]

the image set \( \Gamma \) of the unit-circle of the \( \xi \)-plane (sometimes referred to as the root-locus curve) satisfies \( \Gamma \subset \{ q \mid \text{Re} q \geq 0 \} \), then the formula (4) is A-stable. This criterion is based on an earlier result of ours [4] and generalizes a criterion valid for conventional LMF [5], i.e., formulas with \( \tau(\xi) \equiv 0 \). It can be implemented efficiently by using the algorithms of Routh [6] and those described by Duffin [7].

The criterion mentioned above has been used to \textit{a priori} construct the three-parameter family of all A-stable two-step \((k=2)\) formulas of type (1) † which have \( p = 4 \). The formulas of Enright [2] and those proposed by Jackson and Kenue [8] are members of this family.

---

* The present definition of \( A_v \)-stability is analogous to that for conventional LMF given in [4].

† Except possibly for some marginally A-stable ones.
C. Optimization of A-stable LMF with Respect to a Global Accuracy Criterion

If the formula (1) is applied to the test equation
\[ \dot{x} = \lambda x, \quad \lambda = \text{const.} \] (9)

the fundamental solutions of the resulting difference equation are of the form \( \{\zeta^n, n = 0, 1, 2, \ldots\} \) where \( \zeta \) is any one of the solutions of the characteristic equation (8) with \( q = h\lambda \), provided these roots are distinct. If one of them satisfies \( \zeta = e^q \), i.e., if

\[ \rho(\zeta) - \log \zeta \sigma(\zeta) + (\log \zeta)^2 \tau(\zeta) = 0, \] (10)

then, for this \( \zeta \), \( \{\zeta^n\} \) is an exact discrete solution of (9), i.e.,
\[ \zeta^n = e^{qn} = x(t_n), \quad t_n = nh, \] where \( x(t) = e^{\lambda t} \) is the exact fundamental solution of the differential equation (9). Then, as far as this "principal" root is concerned, the formula (4) is exponentially fitted [9, 10] at
\[ q = \log \zeta. \]

Let \( R(\zeta) \) denote the left side of (10). For any given \( \zeta \), \( |R(\zeta)| \), is a measure of how far the formula is from being exponentially fitted.

Thus, \( M = \int [R(\zeta)]^2 d\zeta \) is an \( L^2 \)-measure of global accuracy of the formula (4) with respect to the family of test problems \( \dot{x} = \lambda x, -\infty < \lambda < 0 \). The measure \( M \) can be computed explicitly as a function of the formula parameters.

For the particular three-parameter class of formulas defined under II.B the following problem was solved: Optimize the formula with respect to \( M \) over all three parameters, subject to the A-stability constraints. The unique optimal formula was found to be
III. FRACTIONAL LINEAR DIFFERENCE SCHEMES FOR SYSTEMS OF QUADRATIC DIFFERENTIAL EQUATIONS \( \{h, i, j\} \)

We are concerned with the numerical solution of quadratic systems of ordinary differential equations

\[ \dot{\mathbf{x}} = \mathbf{a} \mathbf{x} + \mathbf{b} \mathbf{u} + \mathbf{c}, \]

by quadratic systems of difference schemes

\[ (-R\mathbf{u}(t) + S)\mathbf{u}(t+h) = T\mathbf{u}(t)\mathbf{u}(t) + \mathbf{v}, \mathbf{u}(0) = \xi. \]

Here, \( \alpha \) and \( \beta \) are summed from 1 to \( n \), and \( i \) ranges from 1 to \( n \).

In vector notation the initial value problems for these equations are

\[ \dot{\mathbf{x}} = \mathbf{a}\mathbf{x} + \mathbf{b}\mathbf{u} + \mathbf{c}, \mathbf{x}(0) = \xi, \]

\[ (-R[u(t)] + S)u(t+h) = T[u(t)]u(t) + Vu(t) + v, u(0) = \xi. \]

In the above, \( \mathbf{x}, \mathbf{c}, \xi, \mathbf{u} \) and \( \mathbf{v} \) are vectors in \( \mathbb{R}^n \); \( \mathbf{B}, \mathbf{S} \) and \( \mathbf{U} \) are \( n \times n \) matrices; \( \ast \) is multiplication in the algebra \( \mathfrak{A}(\ast) = (a_{jk}) \); and \( R[u] \) and \( T[u] \) are \( n \times n \) matrices which are linear functions of \( u \).

When \( T[u] \equiv 0 \), the scheme (12) can be written as

\[ u(t+h) = (-R[u(t)] + S)^{-1}(Vu(t) + v), \]

which we call a fractional linear scheme. We have obtained the following results:

1) If the scheme (13) converges as \( h \to 0 \) for all \( \xi \), then it must converge to a quadratic system (11).
2) For every quadratic system (11), there exists a fractional linear scheme (13) which approximates (11) to second order and which converges to (11) as \( h \to 0 \).

3) If the coefficients of the scheme (13) satisfy

\[
R(h)[u]p = \int_0^h V^{-1}(\tau)((V(\tau)u)(V(\tau)p))d\tau,
\]

\[
S(h) = V(h)
\]

\[
U(h) = -R(h)[V(h)z(h)] + I,
\]

\[
v(h) = V(h)z(h),
\]

where \( z = z(t) \) and \( V = V(t) \) are defined by

\[
\dot{z} = z^*z + Bz + c, \quad z(0) = 0,
\]

\[
V = -V(B+2R[z]), \quad V(0) = I, \quad (R[z]x = z^*x),
\]

then (13) is a second order accurate convergent scheme for (11).

4) The scheme (13) is the exact scheme for (11) if and only if the scheme coefficients satisfy (14)–(17) and the system (11) is a P-system. That is, the solution of (11) is given by

\[
x(t,\xi) = (-R(t)[\xi]V(t)+V(t))^{-1}((-R(t)[V(t)z(t)]+I)\xi + V(t)z(t)).
\]

5) If the coefficients of (13) satisfy (14)–(17) to second order in \( h \), then (13) is also a second order accurate convergent scheme for (11).

6) The matrix Riccati equation

\[
X = XAX + BX + XC + D,
\]

which is of great practical importance, is a P-system and therefore has a fractional linear exact scheme. Further, this exact scheme can be constructed by solving a linear system of ordinary differential equations
on the interval \([0,h]\).

These results give a rather complete analytic theory of fractional linear difference schemes for quadratic systems. To test the computational effectiveness of these schemes, two computer codes are being developed implementing these schemes: one for two-dimensional systems, and one for arbitrary matrix Riccati equations.

IV. ROTATIONAL SOLUTIONS OF THE JOSEPHSON PHASE (SINE-GORDON) EQUATION\([\kappa,\varphi]\)

A. Problem Formulation

A Josephson junction consists of two superconductors separated by an extremely thin dielectric barrier. The order parameter \(\varphi\) (phase difference between the wave functions of the two superconductors) satisfies the Josephson equations \([11]\). In a mathematical limit situation where \(\varphi\) depends only on one space dimension and on time, Josephson's fundamental equations combined with Maxwell's equations give rise to a nonlinear damped wave equation for \(\varphi\) referred to as the Josephson phase equation or the "sine-Gordon" equation. In dimensionless form \(^*\) this equation is

\[
A\varphi \equiv \frac{\partial^2 \varphi}{\partial x^2} - \frac{\partial^2 \varphi}{\partial t^2} - \kappa \frac{\partial \varphi}{\partial t} = \sin \varphi.
\]

For the purpose of describing the Josephson ac-effect, one is seeking rotational solutions of (18); i.e., solutions of the form \(\varphi = \omega t + \psi\), where \(\psi\) is periodic in \(t\) of period \((2\pi/\omega)\), for some \(\omega\).

\(^*\) The dimensionless variables are those introduced in \([12]\), except that distances are measured in units of the \(L/\kappa\) where \(L\) is the junction length, rather than in units of the Josephson penetration depth \(\lambda_j\).
In a situation where a constant voltage $V_e$ is applied to one end of the junction ($x=0$), the equation (18) is to be solved subject to the boundary conditions [12]

$$\phi|_{x=0} = \omega t, \quad \omega = V_e$$

$$\frac{\partial \phi}{\partial x}|_{x=\pi} = -H_e,$$

where $H_e$ denotes an externally applied magnetic field. This case is referred to as the "voltage-driven case." Note that $\kappa = \frac{L^2}{(\pi \lambda J)^2}$ where $\lambda J$ is the Josephson penetration depth, and the $\sigma$ of [12] equals $(\pi \lambda J/L)$ times the present $\sigma$. The total current $I$ drawn by the junction is the time average (over a period) of the ac total current $I(t)$ defined by the additional constraint

$$\frac{\partial \phi}{\partial x}|_{x=\pi} - \frac{\partial \phi}{\partial x}|_{x=0} = I(t).$$

The nonlinear current-voltage characteristic of the junction is then given by $I = I(\omega)$.

If the junction is driven by a constant total imposed current $I_e$, one is looking for rotational solutions of (18) with unknown $\omega$ satisfying the boundary conditions

$$\frac{\partial \phi}{\partial x}|_{x=0} = -(H_e + I_e),$$

$$\frac{\partial \phi}{\partial x}|_{x=\pi} = -H_e.$$

Then the voltage $V$ across the junction at $x = 0$ is the time average of the ac voltage $V(t)$ defined by $V(t) = (\partial \phi/\partial t)|_{x=0}$. This case is referred to as the "current-driven case."
B. Perturbation Solutions

a) Analytic solutions valid for small nonlinearities (i.e., small values of \( \kappa \), corresponding to junction lengths \( L \) which are small compared to \( \lambda_j \)) were obtained by perturbation methods in both the voltage and the current driven cases. They exhibit the typical "resonance" behavior of the I-V-characteristics observed experimentally and predicted numerically [14].

b) A constructive perturbation procedure, similar to those used in bifurcation theory and valid in the limit of strong dissipation (\( \sigma \rightarrow \infty \)), was developed for the voltage driven case.

C. Existence, Uniqueness, and Stability Results

If one writes \( \phi = \phi_0 + \psi \), where

\[
\phi_0 = \omega t - kx + \rho x^2
\]
satisfies the formal limit equation \( \Lambda \phi_0 = 0 \) as \( \kappa \rightarrow 0 \), and where \( k = \sigma \omega + H_e \), \( \rho = \sigma \omega/2 \), and \( \psi \) is periodic in \( t \), then in the voltage driven case \( \psi \) satisfies

\[
\Lambda \psi = \kappa \sin(\phi_0 + \psi),
\]

\[
\psi\big|_{x=0} = \frac{3\psi}{3x}\big|_{x=\pi} = 0.
\]

The following results were proved in this case:

a) For any \( \sigma \neq 0 \) there exist \((2\pi/\omega)\)-periodic solutions of the problem (19). Hence, for \( \sigma \neq 0 \), there exist (possibly multiply branched [13]) I-V-characteristics for the Josephson junction.

b) For moderate amounts of dissipation relative to the strength of the nonlinearity (i.e., moderately large values of \( \sigma \) relative to \( \kappa \)), the
periodic solution of (19) is unique and globally asymptotically stable. Hence solutions with arbitrary initial data tend - exponentially in $t$ - to $\phi = \phi_0 + \psi$ with $\psi$ periodic. For example, if $\kappa = \frac{1}{8}$ (corresponding to $L/\lambda_j \approx 9$) then $\sigma \geq 9/16$ is sufficient for uniqueness; similarly, with $\kappa = 1/4\pi^2$ (corresponding to $L/\lambda_j = \frac{1}{2}$) uniqueness is assured for $\sigma \geq .06$. These results validate the perturbation procedure discussed under IV.B.a in the limit $\kappa \to 0$. 
V. LIST OF RESEARCH PAPERS


VI. REFERENCES


Yorktown Heights, April 1977

Werner Liniger, Principal Investigator