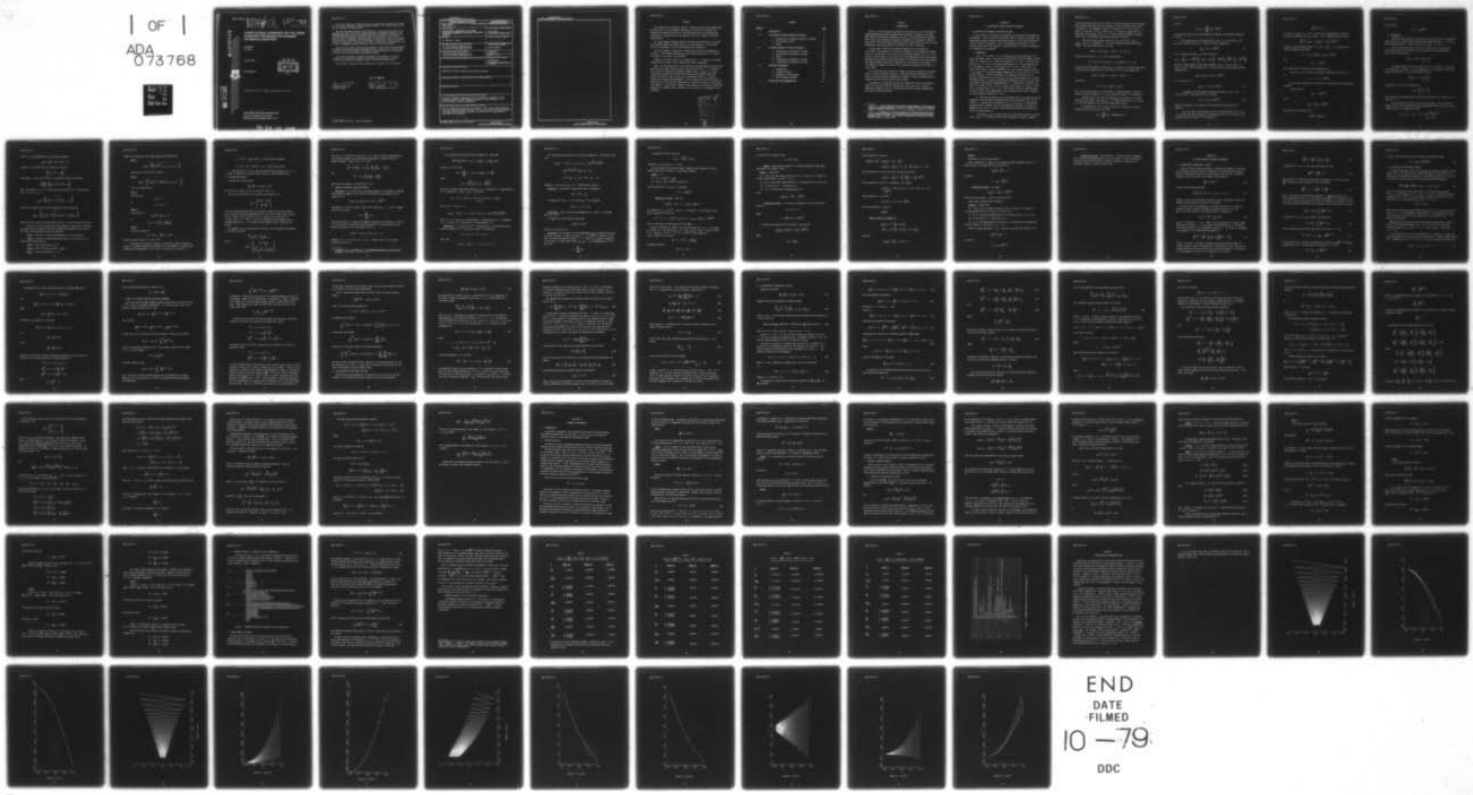
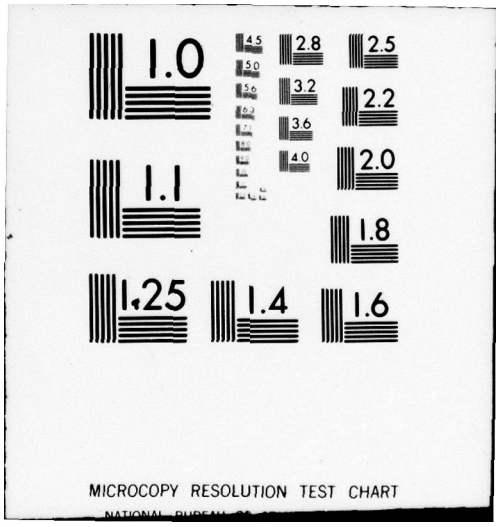
AD-A073 768 AIR FORCE WEAPONS LAB KIRTLAND AFB NM F/G 12/1
COMPUTATIONAL EXPERIMENTS ON TWO ERROR ESTIMATION PROCEDURES FO--ETC(U)
JUN 79 B EPSTEIN, D HICKS

UNCLASSIFIED AFWL-TR-78-119 SBIE-AD-E200 353 NL

| OF |
ADA
073768



END
DATE
FILMED
10 -79
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
7.1
8.0
9.0
MICROCOPY RESOLUTION TEST CHART

AFWL-TR-78-119

LEVEL II

DDC

AFWL-TR-78-119

AD-E200353

AD A073768

COMPUTATIONAL EXPERIMENTS ON TWO ERROR ESTIMATION PROCEDURES FOR ORDINARY DIFFERENTIAL EQUATIONS

B. Epstein
D. Hicks

June 1979

DDC
RECEIVED
SEP 13 1979
C

Final Report

DDC FILE COPY

Approved for public release; distribution unlimited.

AIR FORCE WEAPONS LABORATORY
Air Force Systems Command
Kirtland Air Force Base, NM 87117

79 08 20 049

AFWL-TR-78-119

This final report was prepared by the Air Force Weapons Laboratory, Kirtland Air Force Base, New Mexico, under Job Order 99910001. Dr. Bernard Epstein (AD) was the Laboratory Project Officer.

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This report has been authored by an employee of the United States Government. Accordingly, the United States Government retains a nonexclusive, royalty-free license to publish or reproduce the material contained herein, or allow others to do so, for the United States Government purposes.

This report has been reviewed by the Office of Information (OI) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Bernard Epstein

BERNARD EPSTEIN, PhD
Project Officer

David E. McIntyre

DAVID E. MCINTYRE
Chief, Computational Division

DO NOT RETURN THIS COPY. RETAIN OR DESTROY.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFWL-TR-78-119	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) COMPUTATIONAL EXPERIMENTS ON TWO ERROR ESTIMATION PROCEDURES FOR ORDINARY DIFFERENTIAL EQUATIONS		5. TYPE OF REPORT & PERIOD COVERED Final Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) B. Epstein, D. Hicks		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Air Force Weapons Laboratory (AD) Kirtland Air Force Base, NM 87117		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62601F/99910001
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Weapons Laboratory (AD) Kirtland Air Force Base, NM 87117		12. REPORT DATE June 1979 13. NUMBER OF PAGES 76
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Ordinary differential equations, partial differential equations, finite difference methods, error estimates, asymptotic error estimates, error gradient method, numerical experiments.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Two error estimation procedures are compared. One is called the asymptotic error estimation procedure; the other, the error gradient estimation procedure. The theory underlying these procedures and the results of some numerical experiments are presented.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

PREFACE

The authors wish to thank Mr. David E. McIntyre, Chief of the Computational Division (AD), Air Force Weapons Laboratory, Kirtland Air Force Base, for proposing the general direction of the investigation and for arranging the sponsorship. The authors are also grateful to him for a number of valuable conversations.

Mr. Eugene Omoda of AFWL/AD worked tirelessly and effectively in carrying out numerous computations whose results, to be presented in a later report, served to encourage us as our investigations were proceeding.

Dr. Larry Bertholf, Supervisor of Computational Physics and Mathematics, Division I of the Sandia Laboratories, kindly permitted one of us (D. Hicks) to engage in this work at Kirtland AFB.

Thanks are also due to Mrs. M. M. Madsen and Dr. T. J. Burns of the Sandia Laboratories for their critical reading and helpful suggestions.

Note: As the preparation of this report was nearing completion, we had the good fortune to meet, and hold several discussions with, Dr. H. J. Stetter of the Institut für Numerische Mathematik (INM), Vienna. He gave us a report ("The Defect Correction Principle and Discretization Methods") published recently by INM (Nr. 26/77), which is a preliminary version of a paper to appear soon in Numerische Mathematik. In this report many useful insights are furnished into the historical development of estimation of error of discretization methods. In particular, Stetter's work appears to constitute a significant advance over Lindberg's work, referred to in this report, on asymptotic error estimates.

Accession For
NTIS GRA&I
DDC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist. | Avail and/or special

CONTENTS

SECTION I

INTRODUCTION

There is a serious uncertainty about the true accuracy of certain numerical schemes employed for obtaining approximate solutions of systems of nonlinear PDEs (partial differential equations), such as in the hydrocodes and wavecodes. Until close error estimates are proved, the output of any numerical scheme is open to the criticism of being little more than a guess about the exact solution. Furthermore, even the existence of solutions is often in question.

In this report, two error estimation procedures are introduced and illustrated. One procedure is called the asymptotic error estimation procedure (ref. 1), and the other is called the error gradient procedure (ref. 2). Also included in this report are the results of some computations which were carried out with the intention of providing concrete illustrations of the two procedures. Four ODEs (ordinary differential equations) are considered in some detail. It is intended to present the results of corresponding computations for PDEs in a later report.

It is hoped that this report will render a valuable instructional function, namely, to assist numerical analysts in obtaining accurate estimates of the errors incurred by various computational schemes, so that an optimum decision among the various available options may be made.

The authors also hope that the specific computations which are presented herein will serve to stimulate further work along these lines. They will welcome any comments as well as information concerning further computational work which has been undertaken.

---

1. Lindberg, B., Error Estimation and Iterative Improvements for the Numerical Solution of Operator Equations, Report No. VIVCDCS-R-76-820, Department of Computer Sciences, University of Illinois, Champaign, Illinois, July 1976.

2. Hicks, D., A Procedure to Produce Estimates for Difference Approximations to Differential Equations, I. Introduction and Illustration of the Basic Concept, Report No. SAND-75-0487, Sandia Laboratories, Kirtland Air Force Base, New Mexico, January 1976.

SECTION II

THE ASYMPTOTIC ERROR ESTIMATION PROCEDURE

1. ILLUSTRATION ON ALGEBRAIC EQUATIONS AND ODEs

In this section we illustrate the asymptotic error estimation procedure on two examples. First, to bring out the basic ideas as clearly as possible, the comparatively simple problem of locating the zero of a real-valued function of a real variable is considered. Than an ODE problem is discussed. In the following subsection a PDE problem is discussed briefly, and then the theoretical foundations are presented.

a. Example #1.

Let $F: R \to R$; that is, F is a real-valued function defined on the whole real number system. For each sufficiently small positive number h (henceforth, all h means all h in some fixed interval $0 < h < h_0$) let the family of functions $\phi_h: R \to R$ be a p-th order approximation to F in the sense that for all real numbers x and all h the equality $\phi_h(x) = F(x) + O(h^p)$ holds. As usual, $O(h^p)$ denotes a quantity depending on x and h whose absolute value is $\leq C h^p$, where C is a positive number independent of x and h, and p is a fixed positive integer. (Here the parameter h is somewhat artificial, but in practical applications it will play the role of a mesh-increment in a finite-difference scheme.) We make the additional hypothesis that F and ϕ_h are twice continuously and boundedly differentiable; i.e., there exists a positive constant A such that $|F'(x)|$, $|F''(x)|$, $|\phi_h'(x)|$, and $|\phi_h''(x)| \leq A$ for all x and h. Also, we assume that $\phi h'(x) = F'(x) + O(h^p)$, where $O(h^p)$ is as defined above. (Of course, two or more appearances of the expression $O(h^p)$ refer in general to quite different functions of x and h.)

As indicated previously, we are interested here in illustrating how the asymptotic error estimation procedure may be used to estimate a zero of the function F. Suppose that the equations $F(x) = 0$ and $\phi_h(x) = 0$ both possess unique solutions, denoted y and $\eta(h)$ respectively, and that it is easier to determine $\eta(h)$ (for all h) than y; what can be said about the accuracy with which $\eta(h)$ approximates y?

The quantity $\phi_h(y)$ is termed the local discretization error; note that $\phi_h(y) = \phi_h(y) - F(y)$, which by the original hypothesis is $O(h^p)$. However,

this does not imply that $\eta(h)$ is close to y; roughly speaking, such an undesirable situation may occur when the graphs of F and ϕ_h are quite flat in the vicinity of y. For example, if $F(x) = x^3$ and $\phi_h(x) = x^3 - h$ (so that $p = 1$), then $y = 0$ and $\eta(h) = h^{1/3}$, so that $\eta(h) - y$ is large in comparison with h. Note that, in this case, $F'(y) = 0$. In order to make the general situation quite clear, we introduce the following concept of stability.

Let z be any real number; since $|\phi_h'(x)| \leq A$ for all x and h, $|\phi_h(\xi) - \phi_h(z)| < r$ whenever $|\xi - z| < r/A$, r being any specified positive number. This shows that, given r, the set

$$\mathcal{A}(z;r) = \left\{\xi: |\phi_h(\xi) - \phi_h(z)| < r \text{ for all } h\right\}$$

contains the interval $|\xi - z| < r/A$. We now define

$$R = R(z;r) = \min |\phi_h'(\xi)|, \ \xi \text{ in } \mathcal{A}(z,r), \ 0 < h < h_o$$

We now say that the family of functions ϕ_h is stable at z provided that there exist positive constants S and r such that, for all h, the inequalities

$$|\phi_h(\xi^1) - \phi_h(z)| < r, \ |\phi_h(\xi^2) - \phi_h(z)| < r$$

imply that

$$|\xi^1 - \xi^2| \leq S \, |\phi_h(\xi^1) - \phi_h(\xi^2)|$$

Now, if for some choice of r we find that R, as defined above, is strictly positive, then by choosing $S = R^{-1}$ we see that the family ϕ_h is stable at z.

Roughly stated, the family of functions ϕ_h is stable at z if these functions do not have a flat spot at z. If these functions have a flat spot at y then the numbers $\eta(h)$ may differ significantly from y.

Now suppose that for a certain function z(h) there exist positive integers p and M, with $p \leq M$, such that

$$z(h) = \sum_{j=p}^{M} h^j e_j, \text{ independent of } h$$

so that

$$\eta(h) = y + \sum_{j=p}^{M} h^j e_j + O\left(h^{M+1}\right)$$

then the right side of this last equation is termed a p-M asymptotic expansion of $\eta(h)$.

Now, suppose that a p-M approximation, $z(h)$, to y is available, with $M+1 \leq 2p$. (Recall that $p \leq M$.) Then it is easily shown that

$$\phi_h(y) + F(\eta) = O\left(h^{M+1}\right) \tag{1}$$

This is proven as follows: By Taylor's theorem

$$\phi_h(\eta) = \phi_h\left[y + z + O\left(h^{M+1}\right)\right] = \phi_h(y) + \phi_h(y)\left[z + O\left(h^{M+1}\right)\right] + \frac{\phi_h''(\tilde{y})}{2} \cdot \left[z + O\left(h^{M+1}\right)\right]^2$$

where $\tilde{y}$ is some suitably chosen number between y and η. Since $|\phi_h''(\tilde{y})| \leq A$, $\phi_h'(y) = F'(y) + O(h^p)$, and $2p \geq M+1$, it is now evident that the above equation simplifies to

$$\phi_h(\eta) = \phi_h(y) + F'(y)z + O\left(h^{M+1}\right)$$

or, since $\phi_h(\eta) = 0$,

$$\phi_h(y) = -F'(y)z + O\left(h^{M+1}\right) \tag{2}$$

An entirely similar argument furnishes the equality $F(\eta) = F(y) + F'(y)z + O(h^{M+1})$, and since $F(y) = 0$, we obtain

$$F(\eta) = F'(y)z + O\left(h^{M+1}\right) \tag{3}$$

Addition of equation 2 and equation 3 furnishes the desired equality of equation 1.

Thus, the local discretization error of the approximation $\phi_h(\cdot) + F(\eta)$ is of higher order than the local discretization error of $\phi_h(\cdot)$, and this fact suggests that the solution η^E of the equation

$$\phi_h\left(\eta^E\right) + F(\eta) = 0$$

is closer to y than η is. This is in fact true (asymptotically) under the additional hypothesis that ϕ_h is stable. This is demonstrated as follows:

$$\phi_h\left(\eta^E\right) - \phi_h(y) = -\left[\phi_h(y) + F(\eta)\right] = O\left(h^{M+1}\right)$$

Therefore, for sufficiently small h, $|\phi_h(\eta^E) - \phi_h(y)| < r$, and now by the stability hypothesis we obtain

$$|\eta^E - y| \leq S\ |\phi_h\left(\eta^E\right) - \phi_h(y)| = O\left(h^{M+1}\right)$$

Thus

$$\eta^E = y + O\left(h^{M+1}\right)$$

which shows that η^E constitutes an improvement over η.

We can use this last result to estimate asymptotically the error $\eta - y$, for

$$\eta - y = \eta - \eta^E + O\left(h^{M+1}\right)$$

This simple observation is the basis of the asymptotic error estimation procedure.

Observe that if

$$\phi_h^E = F + O\left(h^{M+1}\right)$$

then

$$\phi_h(y) + \phi_h^E(\eta) = O\left(h^{M+1}\right)$$

and therefore the solution $\tilde{\eta}^E$ to

$$\phi_h\left(\tilde{\eta}^E\right) + \phi_h^E(\tilde{\eta}) = 0$$

also satisfies

$$\tilde{\eta}^E = y + O\left(h^{M+1}\right)$$

b. Example #2.

The ideas which have been introduced up to this point can be illustrated quite clearly by a simple initial-value problem, namely the ODE $y' = y$ with the initial condition $y(0) = 1$ (whose solution, of course, is $y = e^x$).

Let E be the Banach space $C^1[0, 1]$, the family of all real-valued functions possessing a continuous first derivative on the closed interval $[0, 1]$, with norm defined by

$$||y||_E = |y(0)| + \max_{0 \le x \le 1} |y'(x)|$$

The Banach space E^0 will be chosen as $R \times C[0, 1]$; that is, the set of all ordered pairs $\begin{matrix}\alpha \\ f\end{matrix}$, where α is a real number and f is a continuous real-valued function defined on $[0, 1]$, the norm being defined by

$$\left|\left|\begin{matrix}\alpha \\ f\end{matrix}\right|\right|_{E^0} = |\alpha| + \max_{0 \le x \le 1} |f(x)|$$

The operator $F: E \to E^0$ will be defined by

$$F(y) = \begin{pmatrix} y(0) - 1 \\ y' - y \end{pmatrix}$$

Then, clearly, the aforementioned initial-value problem can be restated in the form $F(y) = 0$.

We discretize the problem in the following manner: For any positive integer N let $h = 1/N$, let E_h be the Banach space consisting of all (N+1)-tuples (ordered sets of N+1 numbers) with norm

$$|| y_0, y_1, \cdots, y_N || = |y_0| + \max_{0 \le j \le N-1} \frac{|y_{j+1} - y_j|}{h}$$

and let E be projected onto E_h by the obvious mapping

$$\Delta_h(y) = \left[y(0),\ y(h) - y(2h), \cdots\right]$$

Similarly, let E_h^o consist of all objects of the form

$$\left[\overset{\alpha}{\left(c_0,\ c_1, \cdots,\ c_N\right)}\right]$$

with norm $|\alpha| + \max |c_j|$, while E^o is projected onto E_h^o by the mapping

$$\Delta_h^o\begin{pmatrix}\alpha \\ f\end{pmatrix} = \left[\overset{\alpha}{\left(f(0),\ f(h),\ f(2h), \cdots\right)}\right]$$

Next, the operator $F:E \to E^o$ is replaced by the operator $\phi_h:E^o \to E_h^o$ as follows: For $y = (y_0, y_1, y_2, \cdots)$,

$$\phi_h(y) = \left(\overset{y_0 - l}{\frac{y_1 - y_0}{h} - y_0},\ \frac{y_2 - y_1}{h} - y_1, \cdots\right)$$

Finally, the higher-order discrete approximation ϕ_h^E is defined by

$$\phi_h^E(y) = \left(\overset{y_0 - l}{\frac{y_1 - y_0}{h} - \frac{y_0 + y_1}{2}},\ \frac{y_2 - y_1}{h} - \frac{y_1 + y_2}{2}, \cdots\right)$$

Note that ϕ_h^E is accurate to second order, while ϕ_h is only first-order accurate.

To sum up, we list step-by-step the procedure to be employed in deriving an asymptotic error estimate, and then illustrate the procedure in the case of the initial-value problem described in Example #2.

Step 1: Construct a discrete approximation ϕ_h to the operator F.

Step 2: Construct a second approximation ϕ_h^E to F such that ϕ_h^E is of higher order accuracy than ϕ_h.

Step 3: Find η such that $\phi_h(\eta) = 0$.

Step 4: Find η^E such that $\phi_h\left(\eta^E\right) = -\phi_h^E(\eta)$.

Step 5: The error estimate is $\eta - \eta^E$.

Referring to Example #2, the steps assume the following form:

Step 1:

$$\phi_h(\xi) = \left(\frac{\xi_{j+1} - \xi_j}{h} \xi_0 - 1 - \xi_j,\ 0 \le j \le N - 1 \right)$$

(Note that ϕ_h is of the first order.)

Step 2:

$$\phi_h^E(\xi) = \left(\frac{\xi_{j+1} - \xi_j}{h} - \frac{\xi_{j+1} - \xi_j}{2},\ 0 \le j \le N - 1 \right)$$

(ϕ_h^E is of second order.)

Step 3:

The solution to

$$\phi_h(\eta) = 0$$

is

$$\eta_j = (1 + h)^j$$

Step 4:

The solution to

$$\phi_h\left(\eta^E\right) + \phi_h^E(\eta) = 0$$

is

$$\eta_j^E = (1 + h)^j \left(1 + \frac{jh^2}{2 + 2h}\right)$$

Step 5:

The error estimate is

$$\left(\eta - \eta^E\right) j \approx - \frac{jh^2}{2} (1 + h)^{j-1}$$

while the actual error is $(1 + h)^j - e^{jh}$.

Although the next report contains a considerable amount of numerical results, it appears worthwhile to present here a few numbers illustrating the problem under consideration. If N = 20, h = 1/20, and j = 20, then

$$\left(\eta - \eta^E\right) j = -\frac{1}{40}(1.05)^{19} \approx -0.063 \text{ (error estimate)}$$

$$(1 + h)^j - e^{jh} = (1.05)^{20} - e \approx -0.065 \text{ (actual error)}$$

Note that $\eta_j^E \approx 2.716$, so that the improved error estimate, $\eta_j^E - e$, is -0.002, in contrast to the much cruder value of -0.065 obtained with η_j.

2. ILLUSTRATION ON PDEs

Consider the simple problem

$$\frac{\partial w}{\partial t} + \frac{\partial w}{\partial x} = 0,\ w(0,x) = f(x)$$

The solution is given, of course, by $w(t,x) = f(x - t)$.

Now, for each function $w(t,x)$ we define the operator F by

$$F(w) = \begin{bmatrix} w(0,x) - f(x) \\ \frac{\partial w}{\partial t} + \frac{\partial w}{\partial x} \end{bmatrix}$$

(We omit the detailed definition of the spaces E and E^0 and of their norms.) Let ϕ_h and ϕ_h^E be discrete approximations to F with ϕ_h^E of higher order accuracy than ϕ_h. Let v be the solution of $\phi_h(v) = 0$ and v_E be the solution to $\phi_h(v_E) + \phi_h^E(v) = 0$. Then $v - v_E$ is the asymptotic error estimate, that is, the approximation to $v - w$ furnished by the asymptotic error estimation procedure.

For example, let ϕ_h be determined by taking as the difference approximation to the PDE the equation

$$\frac{v_j^{n+1} - v_j^n}{\Delta t} + \frac{v_j^n - v_{j-1}^n}{\Delta x} = 0$$

that is,

$$\phi_h(v) = \begin{bmatrix} v_j^o - f(x_j) \\ \frac{v_j^{n+1} - v_j^n}{\Delta t} + \frac{v_j^n - v_{j-1}^n}{\Delta x} \end{bmatrix}$$

Note that ϕ_h is accurate of order $O(\Delta t + \Delta x)$. Next, let ϕ_h^E be determined by a more accurate scheme, say the LWR method (Richtmyer's version of the Lax-Wendroff scheme). That is (ref. 3, p. 303),

$$v_j^{n+\frac{1}{2}} = \frac{1}{2}\left(v_{j+1}^n + v_{j-1}^n\right) - \frac{\Delta t}{2\Delta x}\left(v_{j+1}^n - v_{j-1}^n\right)$$

and

$$v_j^{n+1} = v_j^n - \frac{\Delta t}{\Delta x}\left(v_{j+\frac{1}{2}}^{n+\frac{1}{2}} - v_{j-\frac{1}{2}}^{n+\frac{1}{2}}\right)$$

Note that this scheme is of order $O(\Delta t^2 + \Delta x^2)$.

3. THEORY OF ASYMPTOTIC ERROR ESTIMATION

Definition: ϕ_h is said to be a p-M approximation to the operator F provided there exists an operator F on E which is of order h^p and satisfies, for all vectors z in E,

$$\phi_h\left(\Delta_h z\right) = \Delta_h^o\left\{F(z) + F(z)\right\} + O\left(h^{M+1}\right)$$

Furthermore, if for each vector z there exist vectors $f_p, \cdots, f_M$ in E^o (independent of h) such that

$$F(z) = \sum_{\nu=p}^{M} h^\nu f_\nu$$

then we say that $\phi_h - F$ admits an asymptotic expansion to the order M. Finally, ϕ_h is said to be a Lipschitzian p-M approximation to F provided there exist positive constants L and b such that, for all h,

$$||F(y^1) - F(y^2)|| \leq h^p L\, ||y^1 - y^2||$$

whenever $||y^1 - y|| \leq b$ and $||y^2 - y|| \leq b$. (Recall that y is the unique solution to $F(y) = 0$.)

---

3. Richtmyer, R. D., and Morton, K. W., Difference Methods for Initial Value Problems, Interscience, New York, 1967.

We illustrate the above definition on Example #2. Note that

$$\frac{z(x + h) - z(x)}{h} = z'(x) + z''(x)\frac{h}{2!} + z'''(x)\frac{h^2}{3!} + O(h^3)$$

Therefore, in this case,

$$F(z) = \sum_{\nu=1}^{2} h^{\nu} f_{\nu}(z) = z''(x)\frac{h}{2!} + z'''(x)\frac{h^2}{6}$$

where

$$f_1(z) = \begin{pmatrix} 0 \\ \frac{z''}{2} \end{pmatrix} \text{ and } f_2(z) = \begin{pmatrix} 0 \\ \frac{z'''}{6} \end{pmatrix}$$

Now let us consider under what conditions ϕ_h is a Lipschitzian 1-1 approximation to F in Example #2. Recall that $y = e^x$ and that

$$||y^1 - y^2|| = |y^1(0) - y^2(0)| + \max_{0 \le x \le 1} |[y^1(x) - y^2(x)]'|$$

Thus, for M = 1 and $F = f_1$,

$$||F(y^1) - F(y^2)|| = ||f_1(y^2) - f_1(y^2)|| = \max \frac{|(y^1)'' - (y^2)''|}{2}$$

Thus, if $s \ge 2$, then ϕ_h is a Lipschitzian 1-1 approximation to F. Furthermore, if $s \ge 3$, then ϕ_h is a Lipschitzian 1-2 approximation to F.

Definition: ϕ_h is said to be h^{-k} - Lipschitzian at ξ provided there exist positive constants L and d such that the pair of inequalities

$$||\xi^2 - \xi|| \le d, \quad ||\xi^2 - \xi|| \le d$$

imply that

$$||\phi_h(\xi^1) - \phi_h(\xi^2)|| < L ||\xi^1 - \xi^2|| h^{-k}$$

We illustrate this definition in the case of Example #2. Note that in this case

$$||\phi_h(\xi^1) - \phi_h(\xi^2)|| \leq |\xi_0^1 - \xi_0^2| + \max_i \frac{|\xi_{i+1}^1 - \xi_{i+1}^2|}{h}$$

$$+ \max_i \frac{|\xi_i^1 - \xi_i^2|}{h} + \max_i |\xi_i^1 - \xi_i^2|$$

$$\leq 4h^{-1} \max_i |\xi_i^1 - \xi_i^2| = 4h^{-1} ||\xi^1 - \xi^2||$$

Therefore, in this case, ϕ_h is h^{-1} - Lipschitzian at every ξ.

Definition: The global discretization error is defined as

$$\varepsilon(h) = \eta(h) - \Delta_h y$$

In Example #2, $\eta(h)_i = (1 + h)^i$ and $\Delta_h y_i = e^{x_i}$; therefore,

$$\varepsilon(h)_i = (1 + h)^i - e^{x_i}$$

Definition: $\eta(h)$ is said to be convergent to y of order $p \geq 1$ provided that $\varepsilon(h) = O(h^p)$.

In Example #2, one can readily show that

$$|\varepsilon(h)_i| \leq hx_i\, e^{x_i}$$

and so, in this case, $p = 1$.

Definition: We say that η is a p-M approximation to y provided there exist $M \geq p$ and a vector $z = O(h^p)$ such that $\eta = y + z + O\left(h^{M+1}\right)$. Moreover, we say that the global discretization error $\varepsilon = \eta - y$ admits an asymptotic expansion to the order M if there exist vectors $e_p, e_{p+1}, \cdots, e_M$ (independent of h) in E such that

$$z = \sum_{\nu=p}^{M} h^\nu e_\nu$$

In Example #2 one can show that

$$\varepsilon(h)_i = -\frac{x_i e^{x_i} h}{2} + O(h^2)$$

Therefore, in this case, $e_1 = -xe^x/2$.

We now proceed to state three lemmas, together with indications of their proofs, which will be helpful in formulating Theorem 1.

Lemma 1. If

(1) ϕ_h is stable at ξ, and

(2) $\phi_h(\xi) + \overset{E}{h}(\eta) = O\left(h^{M+1}\right)$ (with $M \geq 0$)

then the solution η^E of $\phi_h(\eta) = 0$ satisfies

$$\eta^E = \xi + O\left(h^{M+1}\right)$$

Indication of Proof: Note that

$$||\phi_h\left(\eta^E\right) - \phi_h(\xi)|| = ||\phi_h(\xi) + \phi_h^E(\eta)||$$

By hypothesis (2), $||\phi_h\left(\eta^E\right) - \phi_h(\xi)|| < r$ whenever h is sufficiently small. Now by hypothesis (1) we have

$$||\eta^E - \xi|| < S\,||\phi_h\left(\eta^E\right) - \phi_h(\xi)|| = S\,||\phi_h(\xi) + \phi_h^E(\eta)|| = O\left(h^{M+1}\right)$$

which is equivalent to the assertion of the lemma.

We now illustrate this lemma in the case of Example #2. Let $\xi = \Delta_h y$ (where, as always, $F(y) = 0$) and let $\phi_h(\eta) = 0$. Then $\phi_h(\Delta_h y) + \phi_h^E(\eta) = O(h^2)$, so that $M = 1$. Now $\Delta_h y_i = e^{x_i}$, $\eta_i = (1 + h)^i$, so that η^E, the solution to $\phi_h(\eta^E) + \phi_h^E(\eta) = 0$ is given by

$$\eta_i^E = (1 + h)^i \left(1 + \frac{ih^2}{2 + 2h}\right)$$

We therefore obtain

$$\eta_i^E = e^{x_i} + O(h^2)$$

in contrast to the weaker result

$$\eta_i = e^{x_i} + O(h)$$

Remark: The purpose of Lemma 2 is to provide conditions under which hypothesis (2) of Lemma 1 holds.

Lemma 2. Assume that

(1) y and η are the unique solutions to $F(y) = 0$ and $\phi_h(\eta) = 0$ where η is a p-M+1 approximation to y with $1 \leq p \leq M$.

(2) ϕ_h^E is a consistent approximation to F of order M+1 with $p \leq M+1 \leq 2p$.

(3) ϕ_h and ϕ_h^E are h^{-1} - Lipschitzian at y.

(4) ϕ_h is a Lipschitzian p-M approximation to F.

Then,

$$\phi_h\left(\Delta_h y\right) = -\phi_h^E(\eta) + O\left(h^{M+1}\right)$$

Indication of Proof: In relation to hypothesis (1) we introduce the notation

$$\eta = \Delta_h\ (y + z) + \delta$$

where

$$z = O\left(h^p\right) \text{ and } \delta = O\left(h^{M+2}\right)$$

In relation to assumption (4) we introduce the notation

$$\phi_h\left(\Delta_h y\right) = \Delta_h^o\left[F(y) + F(y)\right] + O\left(h^{M+1}\right)$$

where

$$F(y) = O\left(h^p\right)$$

From assumption (1) we have

$$\phi_h\left(\Delta_h y\right) + \phi_h^E(n) = \phi_h\left(\Delta_h y\right) - \phi_h(n) + \phi_h^E(n)$$

$$= \phi_h\left(\Delta_h y\right) - \phi_h\left[\Delta_h(y + z) + \delta\right] + \phi_h^E(n)\left[\Delta_h(y + z) + \delta\right]$$

From assumptions (1) and (3) the last line may be rewritten

$$= \phi_h\left(\Delta_h y\right) - \phi_h\left[\Delta_h(y + z)\right] + \phi_h^E\left[\Delta_h(y + z)\right] + O\left(h^{M+1}\right)$$

From assumptions (2) and (4), furthermore, this becomes

$$= \Delta_h^o \left\{\left[F(y) + F(y)\right] - \left[F(y + z) + F(y + z)\right] + F(y + z)\right\} + O\left(h^{M+1}\right)$$

From assumption (1), we obtain

$$= \Delta_h^o \left\{F(y) - F(y + z)\right\} + O\left(h^{M+1}\right)$$

and from assumptions (1) and (4)

$$= O\left(h^{M+1}\right)$$

Lemma 2 applied to Example #2:

$$\phi_h\left(\Delta_h y\right)_i = e^{x_i}\left[\frac{h}{2} + O(h^2)\right]$$

$$\phi_h^E(n)_i = e^{x_i}\left[\frac{h}{2} + O(h^2)\right] = (1 + h)^i\left(\frac{h}{2}\right)$$

Therefore

$$\phi_h\left(\Delta_h y\right) + \phi_h^E(n) = O(h^2),\ M = 1$$

Lemma 3.

From Lemmas 1 and 2 we have Lemma 3.

Under assumptions (1) through (4) of Lemma 2 and the assumption that ϕ_h is stable at $\Delta_h y$, it follows that the solution η^E to

$$\phi_h\left(\eta^E\right) + \phi_h^E(\eta) = 0$$

satisfies

$$\eta^E = \Delta_h y + O\left(h^{M+1}\right)$$

Indication of Proof: By Lemma 2

$$\phi_h\left(\Delta_h y\right) = - \phi_h^E(\eta) + O\left(h^{M+1}\right)$$

and then applying Lemma 1 yields the desired result.

Lemma 3 may be generalized to theorem 1.

Theorem 1: Assume that

(1) y and η are the unique solutions to $F(y) = 0$ and $\phi_h(\eta) = 0$ where η is a p-M_o approximation to y with $M_o \geq p + k$ and $k \geq 0$.

(2) ϕ_h^E is a consistent approximation to F of order q with $1 \leq p < q \leq 2p$.

(3) ϕ_h and ϕ_h^E are h^{-j} - Lipschitzian at y.

(4) ϕ_h is a p-M approximation to F and ϕ_h is a Lipschitzian p-N_1 approximation to F where $N_1 = \min (M, M_o - K, q - 1)$.

Under the above hypotheses, if ϕ_h is stable at $\Delta_h y$, then the solution η^E of

$$\phi_h\left(\eta^E\right) + \phi_h^E(\eta) = 0$$

satisfies

$$\eta^E = \Delta_h y + O\left(h^{N_1+1}\right)$$

Indication of Proof: See reference 1. The theory of the asymptotic error estimation procedure seems to be in its infancy. There are many open questions about the existence of asymptotic expansions; some proofs are available for ODEs but almost none for PDEs.

SECTION III

THE ERROR GRADIENT ESTIMATION PROCEDURE

1. ILLUSTRATION OF VARIATION E ON ODEs

In this and the following subsection, we present two methods or variations for estimating the error incurred in employing a finite-difference scheme to obtain an approximate solution to an ODE initial-value problem. For ease in exposition we shall consider the extremely simple problem

$$\frac{dW}{dt} = W, \; W(0) = 1 \tag{4}$$

along with the difference-scheme

$$\frac{V^{k+1} - V^k}{h} = V^k, \; V^0 = 1 \; (k = 0, 1, 2, \cdots) \tag{5}$$

However, it will be evident that the basic concept is applicable without any essential complication to more difficult problems.

For each value of the parameter h, the mesh-function V^n is filled in by employing linear interpolation between successive points (nh, V^n) and $[(n + 1)\, h, V^{n+1}]$; that is, the function U(h,t) is defined by

$$U(h,t) = V^n\left(1 - \frac{\tau}{h}\right) + \frac{\tau}{h} V^{n+1} \tag{6}$$

where $n = [t/h]$ and $\tau = t - nh$. (Recall that $[x]$ denotes the largest integer which does not exceed x.) In this manner the function U(h,t) is defined for all positive values of h and nonnegative values of t. For fixed t we determine how U depends on h by differentiating equation 6:

$$\frac{\partial U(h,t)}{\partial h} = \frac{dV^n}{dh} + \left(\frac{\tau}{h}\frac{d}{dh} - \frac{t}{h^2}\right)\left(V^{n+1} - V^n\right) \tag{7}$$

(Since, for fixed t, the index n undergoes jump-discontinuities when t/h = integer, equation 7 must be appropriately interpreted for these values of h as a one-sided derivative, but this does not cause any difficulty.) Eliminating V^{n+1} from equation 7 with the aid of equation 5, one obtains

$$\frac{\partial U}{\partial h}(h,t) = \frac{dV^n}{dn} + \left(\tau \frac{d}{dh} - \frac{t}{h}\right) V^n \tag{8}$$

In particular, as $t/h \to n$, this equation assumes the form

$$\frac{\partial U}{\partial h}(h,t) = \frac{dV^n}{dh} - nV^n \tag{9}$$

The value of V^n is obtained by recurrence from equation 7, and the value of $\frac{dV^n}{dh}$ is obtained, also by recurrence, from the equation

$$\frac{dV^{k+1}}{dh} = (1 + h)\frac{dV^k}{dh} + V^k, \quad \frac{dV^0}{dh} = 0 \tag{10}$$

which is obtained by differentiating equation 5. Accepting the fact that $U(h,t)$ approaches $W(t)$ as $h \to 0$, and assuming that $\frac{\partial U}{\partial h}(h,t)$ is integrable (which is trivially true in the present case and is not difficult to establish for any first-order ODE with right-hand side satisfying very mild hypotheses), we obtain

$$U(h,t) - W(t) = \int_0^h \frac{\partial U}{\partial h'}(h',t)\, dh' \tag{11}$$

In particular, if $h = t/n$ this assumes the simpler form

$$V^n - W(t) = \int_0^h \frac{\partial U}{\partial h'}(h',t)\, dh' \tag{12}$$

Thus, one obtains the following upper bound on the error $V^n - W(t)$:

$$|V^n - W(t)| \leq h \max_{0 < h' < h} \left|\frac{\partial U}{\partial h'}(h',t)\right| \tag{13}$$

In the present case, the explicit computation indicates that $\left|\frac{\partial U}{\partial h}(h',t)\right|$ depends monotonically on h', so that equation 13 becomes (for $h = t/n$):

$$|V^n - W(t)| \leq h \left|\frac{\partial U}{\partial h}(h,t)\right| \tag{14}$$

In fact, a more accurate analysis furnishes the improved estimate

$$|V^n - W(t)| \approx \frac{h}{2} \left|\frac{\partial U(h,t)}{\partial h}\right| \tag{15}$$

2. ILLUSTRATION OF VARIATION H ON ODEs

We consider the same initial-value problem and difference scheme as in subsection III-1. The first step is to extend the domain of definition of the solution V of the difference scheme to all nonnegative values of t. To do this we introduce the solution u(h,t) to the continuum-difference equation

$$\frac{u(h,t + \delta t) - u(h,t)}{\delta t} = u(h,t), \quad u(h,0) = W(0) = 1 \tag{16}$$

Here, for convenience, we have converted h to a dimensionless parameter, $0 \leq h \leq 1$, by introducing $\delta t = h\Delta t$. For $0 < t < \delta t$ it is evidently necessary to define u(h,t). That is, first we select a function f(h,t) for $0 \leq t < \delta t$ (with f(h,0) = 1) and we define u(h,t) for $0 \leq t < \delta t$ by

$$u(h,t) = f(h,t)$$

Then the above difference scheme determines u(h,t) for $t > \delta t$; clearly $u(h,t^n) = V^n$ for h = 1. If the difference scheme equation 16 is convergent, then $\lim_{h \to 0} u(h,t)$ exists and is readily seen to equal W(t), the solution of the initial-value problem.

Observe that the solution to the initial-value problem is W(t) = exp(t); to the finite-difference scheme is $V^n = (1 + \Delta t)^n$; and to the continuum difference equation 16 is $u(h,t) = (1 + \delta t)^n f(h,\tau)$, where $n = [t/\delta t]$ and $\tau = t - n\delta t$.

It is evident that if u is to be continuous it is necessary to impose the same restriction on f; similarly for differentiability. The most natural choice of f is apparently $f(h,t) = 1 + t$, $0 \leq t < \delta t$, and so we adopt the definition

$$u(h,t) = 1 + t, \quad 0 \leq t < \delta t \tag{17}$$

From equation 16 we obtain difference equations for $\frac{\partial u}{\partial t}$ and $\frac{\partial u}{\partial h}$, namely

$$\frac{\partial u}{\partial t}(h,t+\delta t) = (1+\delta t)\frac{\partial u}{\partial t}(h,t)$$

and

$$\frac{\partial u}{\partial h}(h,t+\delta t) = (1+\delta t)\frac{\partial u}{\partial h}(h,t) - \delta t\, R(h,t)$$

where

$$R(h,t) = \frac{\partial u}{\partial t}(h,t+\delta t) - u(h,t)$$

Furthermore, from equation 17 we obtain

$$\frac{\partial u}{\partial t}(h,t) \equiv 1,\ \frac{\partial u}{\partial h}(h,t) = 0 \text{ for } 0 \leq t < \delta t$$

Let

$$\frac{\partial u}{\partial t}^{n} = \frac{\partial u}{\partial t}(1,t^{n})$$

and let

$$\frac{\partial u}{\partial h}^{n} = \frac{\partial u}{\partial h}(1,t^{n})$$

Variation H of the error gradient estimation procedure involves finding the solution of the following system of difference equations:

$$V^{n+1} = (1+\delta t)\,V^{n},\ V^{o} = 1$$

$$\frac{\partial u}{\partial t}^{n+1} = (1+\delta t)\frac{\partial u}{\partial t}^{n},\ \frac{\partial u}{\partial t}^{o} = 1$$

$$\frac{\partial u}{\partial h}^{n+1} = (1+\delta t)\frac{\partial u}{\partial h}^{n} - \delta t\, R^{n}$$

where

$$R^{n} = \frac{\partial u}{\partial t}^{n+1} - V^{n}$$

The error gradient estimate for variation H is

$$V^n - W(t^n) \approx \frac{\partial u}{\partial h}^n$$

3. THEORY OF THE ERROR GRADIENT ESTIMATION PROCEDURE

Since the derivatives $\frac{\partial u}{\partial t}$ and $\frac{\partial V}{\partial t}$ may become discontinuous at the mesh-values $t = t^n$, we take only forward t derivatives. If, similarly, $\frac{\partial}{\partial h}$ is also interpreted as forward differentiation, then

$$\frac{d}{dh} V(h,t + h) = \frac{\partial V(h,t + h)}{\partial h} + \frac{\partial V(h,t + h)}{\partial t}$$

and, similarly,

$$\frac{du(h,t + h\Delta t)}{dh} = \frac{\partial u(h,t + h\Delta t)}{\partial h} + \Delta t \frac{\partial u(h,t + h\Delta t)}{\partial t}$$

In both the E and H variations, the error estimate is based on the identity

$$u(h_1,t) - u(h_2,t) = \int_{h_2}^{h_1} \frac{\partial u(h,t)}{\partial h} \, dh$$

(under the reasonable assumption that u is absolutely continuous with respect to h). If, as we assume,

$$w(t) = \lim_{h \to 0} u(h,t)$$

the above identity yields

$$u(h,t) - w(t) = \int_0^h \frac{\partial u}{\partial h'}(h',t) \, dh'$$

Thus, the error-estimation problem becomes that of estimating the integral appearing in the last equation. In the error gradient estimation procedure, the approximation

$$\int_0^h \frac{\partial u}{\partial h'}(h',t^n)\, dh' \approx h\, \frac{\partial u}{\partial h}(h,t^n)$$

is employed. Of course, the accuracy of this estimate depends critically on the behavior of $\frac{\partial u}{\partial h}$ in the interval $[0,h]$. In particular, if $\frac{\partial u}{\partial h}$ is increasing with respect to h, the above approximation provides an upper bound on the error. When monotonicity cannot be established, one must settle for the cruder upper bound

$$h \sup_{0 < h' \leq h} \left|\frac{\partial u}{\partial h'}(h',t^n)\right|$$

Returning to the simple illustrative example, one sees that variation H requires the solution of a system of three equations, namely:

$$V^{n+1} = (1 + \Delta t)\, V^n,\ V^0 = 1$$

$$V^{n+1} = (1 + \Delta t)\, \frac{\partial u}{\partial t}^n,\ \frac{\partial u}{\partial t}^0 = 1$$

$$\frac{\partial u}{\partial t}^{n+1} = (1 + \Delta t)\, \frac{\partial u}{\partial h}^n - R^n\, \Delta t,\ \frac{\partial u}{\partial h}^0 = 0$$

In contrast to this, variation E requires only the solution of a pair of equations, namely

$$V^{n+1} = (1 + \Delta t)^n,\ V^0 = 1$$

$$\frac{dV^{n+1}}{dh} = (1 + \Delta t)\, \frac{dV^n}{dh} + V^n,\ \frac{dV^0}{dh} = 0$$

It would therefore appear that variation H will involve roughly twice the effort required by variation E in order to compute an error estimate. Thus, at first appearance, variation E would appear to be preferable. However, in variation E the error bound involves the expression $\frac{dV^n}{dh} - nV^n$, which, in the illustrative example, works out to $-t^n\,(1 + \Delta t)^{n-1}$, which is small in comparison with the separate terms $\frac{dV^n}{dh}$ and nV^n. That is, the subtraction of two large quantities which are almost equal is required, and this suggests that a noisy calculation may be risked. Thus, it is difficult to say which of the two variations is,

on the whole, preferable to the other; indeed, the answer may depend critically on the specific problem under consideration.

In the illustrative example both variations lead to the same estimate, namely:

$$h \frac{\partial u}{\partial h}(h,t^n) = - \Delta t(1 + \Delta t)^{n-1} t^n$$

that is, the error estimate amounts to

$$(1 + \Delta t)^n - \exp(t^n) \approx - \Delta t(1 + \Delta t)^{n-1} t^n$$

By employing the identity

$$\int_0^t \left[u(h,t') - w(t') - w(t')\right] dt' = \int_0^t \int_0^h \frac{\partial u}{\partial h'}(h',t') \, dh' \, dt'$$

we may make the estimate

$$\int_0^{t^N} \left[u(h,t) - w(t)\right] dt \approx h \sum_{N=0}^{N} \frac{\partial u^n}{\partial h} \Delta t$$

Similarly, in the case of PDEs, the analogous estimate

$$\int_{x_o}^{x_j} \int_0^{t^N} \left[u(h;t,x) - w(t,x)\right] dt \, dx \approx h \sum_{j=0}^{J} \sum_{m=0}^{N} \frac{\partial u^n}{\partial h_j} \Delta t \, \Delta x$$

may prove useful, particularly when u may fail to converge pointwise to w but may converge to w in the L^1 - norm; similar estimates can be written for the case of convergence in other integral norms.

4. ILLUSTRATION OF VARIATION E ON PDEs

Paralleling the considerations of subsections III-1 and III-2, we shall consider in the present and the next subsection the following problem:

$$\frac{\partial w}{\partial t} + \frac{\partial w}{\partial x} = 0,\ w(0,x) = w^0(x) \tag{18}$$

We choose positive numbers Δt and Δx, then define δt and δx, respectively, as $h\Delta t$ and $h\Delta x$, where the parameter h satisfies $0 < h \leq 1$. Now consider the difference scheme

$$\frac{V_j^{n+1} - V_j^n}{\delta t} + \frac{V_j^n - V_{j-1}^n}{\delta x} = 0,\ V_j^0 = w^0\left(x_j\right) \tag{19}$$

where $x_j = j\delta x$. Proceeding by analogy with subsection III-1, we extend the mesh-function V to the entire half-plane $0 \leq t < \infty$, $-\infty < x < \infty$ by defining, in the rectangle with vertices $(n\,\delta t, j\delta x)$, $(n\delta t, (j+1)\,\delta x)$, $((n+1)\,\delta t, j\delta x)$, and $((n+1)\,\delta t, (j+1)\,\delta x)$, the function $U(h, t, x)$ as follows (by linear interpolation):

$$U(h, t, x) = A + B\,\frac{\tau}{\delta t} + C\,\frac{\xi}{\delta t \delta x} + D\,\frac{\tau\xi}{\delta t \delta x} \tag{20}$$

where

$$\tau = t - n\delta t,\ \xi = x - j\,\delta x,\ A = V_j^n,\ B = V_j^{n+1} - V_j^n$$

$$C = V_{j+1}^n - V_j^n,\ D = V_{j+1}^{n+1} - V_j^{n+1} - V_{j+1}^n + V_j^n$$

By rewriting equation 19 in the form

$$V_j^{n+1} = rV_{j-1}^n + (1 - r)\,V_j^n,\ r = \frac{\delta t}{\delta x} = \frac{\Delta t}{\Delta x} \tag{21}$$

we immediately observe that the condition $r \leq 1$ is necessary to assure convergence of the difference scheme (eq. 19) as $h \to 0$, and it is easy to show that this condition is sufficient as well. In particular, the choice $r = 1$ furnishes the trivial difference scheme $V_j^{n+1} = V_{j-1}^n$, reflecting the fact that the exact

solution of equation 18 is given by $w(t,x) = w^0(x - t)$; that is, the solution is obtained by advancing the initial data along the lines inclined at 45° to the x-axis. In order to avoid this trivial case, we shall consider a value of r strictly less than one.

By repeated use of equation 21 we easily obtain, for any positive index k and for all j,

$$V_j^k = \sum_{i=0}^{k} \binom{k}{i} r^i (1 - r)^{k-i} V_{j-i}^0 = \sum_{i=0}^{k} \binom{k}{i} r^i (1 - r)^{k-1} w^0\left(x_{j-i}\right) \qquad (22)$$

We now seek to develop an expression for U(h, 1, 1), that is, an approximation for w(1,1). For ease in exposition we choose $r = 1/2$, $\Delta x = 0.1$, $\Delta t = 0.05$, $w^0(x) = x$; the procedure in the gereral case will be apparent from this simple illustrative example. We choose a value of h slightly less than one, so that $20\ \delta t < 1 < 21\ \delta t$, $10\ \delta x < 1 < 11\ \delta x$. [That is, the point (1,1) at which we are seeking to estimate w(t,x) lies in the interior of the rectangle with vertices at (20 δt, 10 δx), (20 δt, 11 δx), (21 δt, 10 δx), and (21 δt, 11 δx).] From equation 22 we obtain, upon making the appropriate replacements,

$$V_j^k = 2^{-k} \cdot \left(\frac{h}{10}\right) \sum_{i=0}^{k} \binom{k}{i} (j - i) \qquad (23)$$

The expression on the right can be summed explicitly, and we obtain

$$V_j^k = \frac{h}{10}\left(j - \frac{k}{2}\right) \qquad (24)$$

Thus, for the values at the vertices of the rectangle under consideration, we obtain

$$V_{10}^{20} = 0,\ V_{11}^{20} = \frac{h}{10},\ V_{10}^{21} = -\frac{h}{20},\ V_{11}^{21} = \frac{h}{20},\ V_{11}^{21} = \frac{h}{20} \qquad (25)$$

From equation 20 and the following formulas we now obtain

$$U(h, 1, 1) = 0 \qquad (26)$$

Thus, in this case the procedure of filling in the rectangles of the mesh by linear interpolation furnishes the exact result, but this is due to the simple

form of the initial data. If we complicate the problem slightly by choosing $w^0(x) = x^2$, equations 23 to 26 are replaced, respectively, by

$$V_j^k = 2^{-k} \cdot \left(\frac{h}{10}\right)^2 \sum_{i=0}^{k} \binom{k}{i} (j - i)^2 \tag{27}$$

$$V_j^k = \frac{h^2}{400} (4j^2 - 4jk + k^2 + k) \tag{28}$$

$$V_{10}^{20} = \frac{h^2}{20}, \; V_{11}^{20} = \frac{3h^2}{50}, \; V_{10}^{21} = \frac{11h^2}{200}, \; V_{11}^{21} = \frac{11h^2}{200} \tag{29}$$

$$U(h, 1, 1) = \frac{-5h^2 + 8h - 2}{20} \tag{30}$$

Upon setting $h = 1$ in equation 30 or in the first part of equation 29, we obtain the approximation

$$w(1, 1) \approx \frac{1}{20} \tag{31}$$

On the other hand, upon differentiating equation 30 and setting $h = 1$, we obtain

$$\left.\frac{\partial U}{\partial h}\right|_{h=1} = -\frac{1}{10} \tag{32}$$

and so we are led to the error estimate

$$w(1, 1) - U(1, 1, 1) \approx h \left.\frac{\partial U(h, 1, 1)}{\partial h}\right|_{h=1} = -\frac{1}{10} \tag{33}$$

In fact, since $w(1,1) = 0$, the true value of the left side is $-1/20$. As in subsection III-1, we have obtained an estimate which is double the correct value. We shall explain in section IV how this factor of 2 can be removed in the case of the ODE example, and it appears that a similar (but probably more delicate) argument could be applied in the PDE case as well.

5. ILLUSTRATION OF VARIATION H ON PDEs

Consider the problem

$$\frac{\partial w}{\partial t} + \frac{\partial w}{\partial x} = 0,\ w(0,x) = x^0(x) \tag{34}$$

together with the corresponding difference scheme

$$\frac{v_j^{n+1} - v_j^n}{\Delta t} + \frac{v_j^n - v_{j-1}^n}{\Delta x} = 0,\ v_j^0 = w^0\left(x_j\right) \tag{35}$$

where $x_j = j\Delta x$. The continuum difference equation associated with equation 35 is now written out:

$$\frac{u(h;\ t + \delta t,\ x) - u(h;\ t,\ x)}{\delta t} + \frac{u(h;\ t,\ x) - u(h;\ t,\ x - \delta x)}{\delta x} = 0 \tag{36}$$

where $u(h;\ 0,x) = w^0(x)$, $u(h;t,x) = f(h;t,x)$ for $0 < t < \delta t$.

Now for $h = 1$ and $t = t^n$ (eq. 36) should agree with equation 35, and as $h \to 0$, δt and δx should also approach zero. As before, we take $\delta t = h\Delta t$. As for the relation between δx and Δx, we reason as follows:

The ratio $\delta x/\Delta x$ must be a function $g(h)$ such that $g(1)$ and $\lim_{h\to 0} g(h) = 0$. Stability of the difference scheme (eq. 35) requires that $\Delta t/\Delta x \leq 1$. The natural choice of $g(h)$ dictated by these considerations is $g(h) = h$. Then equation 36 permits the rearrangement

$$u(h;\ t + \delta t,\ x) = (1 - r)u(h;\ t,\ x) + ru(h;\ t,\ x - \delta x) \tag{37}$$

where $r = \Delta t/\Delta x$. Equation 37 suggests how f should be prescribed.

Let

$$u(h;\ t,\ x) = (1 - r)\ w^0(x) + rw^0(x - \xi) \tag{38}$$

where $0 \leq \tau < \delta t$ and $\xi = \tau/r$.

From equation 37 we may derive difference equations for $\frac{\partial u}{\partial t}$ and $\frac{\partial u}{\partial x}$. For $\frac{\partial u}{\partial t}$, we obtain

$$\frac{\partial u}{\partial t}(h;\ t + \delta t,\ x) = (1 - r)\frac{\partial u}{\partial t}(h;\ t,\ x) + r\frac{\partial u}{\partial t}(h;\ t,\ x - \delta x) \qquad (39)$$

and from equation 38 we obtain

$$\frac{\partial u}{\partial t}(h;\ \tau,\ x) = -\frac{\partial w^0}{\partial x}(x - \xi) \text{ for } 0 \leq \tau < \delta t \qquad (40)$$

[As above, $\xi = \tau/r$ and $0 \leq \tau < \delta t$.]

Similarly,

$$\frac{\partial u}{\partial x}(h;\ t + \delta t,\ x) = (1 - r)\frac{\partial u}{\partial x}(h;\ t,\ x) + r\frac{\partial u}{\partial x}(h;\ t,\ x - \delta x) \qquad (41)$$

and again from equation 38 we obtain

$$\frac{\partial u}{\partial x}(h;\ \tau,\ x) = \frac{\partial w^0}{\partial x}(x) - r\left(\frac{\partial w^0}{\partial x}(x) - \frac{\partial w^0}{\partial x}(x - \xi)\right) \text{ for } 0 \leq \tau < \delta t \qquad (42)$$

From equation 37 we derive a difference equation for $\frac{\partial u}{\partial h}$, namely

$$\frac{\partial u}{\partial h}(h;\ t + \delta t,\ x) = (1 - r)\frac{\partial u}{\partial h}(h;\ t,\ x) + r\frac{\partial u}{\partial h}(h;\ t,\ x - \delta x) - \Delta t\ R(h;\ t,\ x) \qquad (43)$$

where

$$R(h;\ t,\ x) = \frac{\partial u}{\partial t}(h;\ t + \delta t,\ x) + \frac{\partial u}{\partial x}(h;\ t,\ x - \delta x)$$

Finally, from equation 37 we obtain

$$\frac{\partial u}{\partial h}(h;\ \tau,\ x) = 0 \text{ for } 0 \leq \tau < \delta t \qquad (44)$$

In variation H, the difference equations that need to be solved for the error gradient estimates are

$$v_j^{n+1} = (1 - r)\,v_j^n + r v_{j-1}^n,\ v_j^0 = w^0(x_j) \qquad (45)$$

$$\frac{\partial u}{\partial t}_j^{n+1} = (1 - r)\,\frac{\partial u}{\partial t}_j + r\,\frac{\partial u}{\partial t}_{j-1}, \quad \frac{\partial u}{\partial t}_j^{0} = -\frac{\partial w^0}{\partial x}\left(x_j\right) \tag{46}$$

$$\frac{\partial u}{\partial t}^{n+1} = (1 - r)\,\frac{\partial u}{\partial x}_j^{n} + r\,\frac{\partial u}{\partial x}_{j-1}^{n}, \quad \frac{\partial u}{\partial x}_j^{0} = \frac{\partial w^0}{\partial x}\left(x_j\right) \tag{47}$$

and

$$\frac{\partial u}{\partial h}_j^{n+1} = (1 - r)\,\frac{\partial u}{\partial h}_j^{n} + \frac{\partial u}{\partial h}_{j-1}^{n} - \Delta t R_j^n, \quad \frac{\partial u}{\partial h}_j^{0} = 0 \tag{48}$$

where

$$R_j^n = \frac{\partial u}{\partial t}_j^{n+1} + \frac{\partial u}{\partial x}_{j-1}^{n}$$

Note that the amount of work involved can be reduced by replacing equations 46 and 47 with the single equation

$$R_j^{n+1} = (1 - r)\,R_j^n + r\,R_{j-1}^n, \quad R_j^0 = \frac{\partial u}{\partial t}_j + \frac{\partial u}{\partial x}_{j-1}^{0} \tag{49}$$

where

$$\frac{\partial u}{\partial t}_j^{1} = (1 - r)\,\frac{\partial u}{\partial t}_j^{0} + r\,\frac{\partial u}{\partial t}_{j-1}^{0}$$

Therefore, the difference equations to be solved are reduced to equations 45, 48, and 49, and the error gradient estimate is given by

$$v_j^n - w\left(t^n, x_j\right) \approx \frac{\partial u}{\partial h}_j$$

As a second illustration of variation H of the error gradient estimation procedure, we consider the heat equation:

$$\frac{\partial w}{\partial t} = \frac{\partial^2 w}{\partial x^2}, \quad w(0,x) = w^0(x)$$

For the corresponding discrete-difference scheme we take

$$\frac{v_j^{n+1} - v_j^n}{\Delta t} = \frac{v_{j+1}^n - 2v_j^n + v_{j-1}^n}{\Delta x^2}, \quad v_j^o = w^o(x_j)$$

The difference equation may be rewritten in the form

$$v_j^{n+1} = (1 - r)\, v_j^n + r \left(\frac{v_{j+1}^n + v_{j-1}^n}{2} \right) \tag{50}$$

where $r = 2\Delta t/\Delta x^2$. A familiar analysis shows that the timestep restriction $r \leq 1$ is necessary and sufficient for stability. Therefore, we let $\delta t = h\Delta t$ and $\delta x = h^{\frac{1}{2}}\Delta x$, and the associated continuum-difference scheme is

$$u(h;\, t + \delta t,\, x) = (1 - r)u(h;\, t,\, x) + r \left[u(h;\, t,\, x + \delta\times) + u(h;\, t,\, x - \delta x) \right]/2$$

with initial condition

$$u(h;\, \tau,\, x) = (1 - r)\, w^o(x) + r \left[w^o(x + \xi) - w^o(x - \xi) \right]/2$$

where

$$0 < \tau < \delta t \text{ and } \xi = \left(2\tau/r\right)^{1/2}$$

Now, differentiating with respect to h, we obtain

$$\frac{\partial u}{\partial h}(h;\, t + \delta t,\, x) = (1 - r)\, \frac{\partial u}{\mu h}(h;\, t,\, x) + r \left[\frac{\partial u}{\mu h}(h;\, t,\, x + \delta x) + \frac{\partial u}{\partial h}(h;\, t,\, x - \delta x) \right]/2 - \Delta t\, R\,(h;\, t,\, x)$$

where

$$R(h;\, t,\, x) = \frac{\partial u}{\partial t}(h;\, t + \delta t,\, x) - \frac{\frac{\partial u}{\partial x}(h;\, t,\, x + \delta x) - \frac{\partial u}{\partial x}(h;\, t,\, x - \delta x)}{2\delta x}$$

with initial condition

$$\frac{\partial u}{\partial h}(h;\ \tau,\ 0) = 0 \text{ for } 0 \leq \tau < \delta t$$

Note that $R_j^n = R(1;\ t^n,\ x_j)$ satisfies a difference equation of the form of equation 50. Therefore, we can, as before, replace the pair of difference equations for $\frac{\partial u}{\partial t}$ and $\frac{\partial u}{\partial x}$ by a single difference equation for R.

Thus, variation H leads to three difference equations:

$$v_j^{n+1} = (1 - r)\, v_j^n + r\left(v_{j+1}^n + v_{j-1}^n\right)\Big/2,\ v_j^o = w^o\left(x_j\right)$$

$$\frac{\partial u}{\partial h_j}^{n+1} = (1 - r)\,\frac{\partial u}{\partial h_j}^n + r\left(\frac{\partial u}{\partial h}^n_{j+1} + \frac{\partial u}{\partial h}^n_{j-1}\right)\Big/2 - \Delta t\, R_j^n,\ \frac{\partial u}{\partial h}^o_j = 0$$

and

$$R_j^{n+1} = (1 - r)\, R_j^n + r\left(R_{j+1}^n + R_{j-1}^n\right)\Big/2$$

with the additional conditions

$$\frac{\partial u^1}{\partial t_j} = (1 - r)\,\frac{\partial u^o}{\partial t_j} + r\left(\frac{\partial u^o}{\partial t}_{j+1} + \frac{\partial u^o}{\partial t}_{j-1}\right)$$

$$\frac{\partial u^o}{\partial t_j} = \frac{d^2 w^o\left(x_j\right)}{dx^2},\ \frac{du^o}{dx_j} = \frac{dw^o}{dx}\left(x_j\right)$$

$$R_j^o = \frac{\partial u^1}{\partial t_j} - \left(\frac{\partial u^o}{\partial x}_{j+1} + \frac{\partial u^o}{\partial x}_{j-1}\right)\Big/2\Delta x$$

In the next example we show why and how f must be modified in order to obtain better estimates when the initial data contain discontinuities. Consider the problem

$$\frac{\partial w}{\partial t} + \frac{\partial w}{\partial x} = 0,\ w(0,x) = w^o(x)$$

The LKL (Lax-Keller-leapfrog) scheme furnishes the following discrete-difference equation:

$$v_j^{n+1} = \frac{v_{j+1}^n + v_{j-1}^n}{2} - \left(\frac{v_{j+1}^n - v_{j-1}^n}{2}\right) r$$

or

$$v_j^{n+1} = \frac{1}{2}(1 - r)\, v_{j+1}^n + \frac{1}{2}(1 + r)\, v_{j-1}^n$$

where $r = \Delta t/\Delta x$. As before, the condition $r \leq 1$ is necessary and sufficient for stability.

The associated continuum difference scheme is

$$u(h;\, t + \delta t,\, x) = \frac{1}{2}\left[u(h;\, t,\, x + \delta x) + u(h;\, t,\, x - \delta x)\right] - \frac{r}{2}\left[u(h;\, t,\, x + \delta x) - u(h;\, t,\, x - \delta x)\right]$$

with initial data $u(h;\, \tau,\, x) = f(h;\, \tau,\, x)$ for $0 \leq \tau < \delta t$. If we were to proceed as in the previous examples, we would define f by

$$f(h;\, \tau,\, x) = \frac{1}{2}\left[w^0(x + \xi) + w^0(x - \xi)\right] - \frac{r}{2}\left[w^0(x + \xi) - w^0(x - \xi)\right]$$

for $0 \leq \tau < \delta t$, with $\xi = \tau/r$. We refer to this definition as the first option for f.

Differentiating with respect to t we obtain

$$\frac{\partial f}{\partial t}(h;\, \tau,\, x) = \frac{1}{2r}\left(\frac{dw^0}{dx}(x + \xi) - \frac{dw^0}{dx}(x - \xi)\right) - \frac{1}{2}\left(\frac{dw^0}{dx}(x + \xi) + \frac{dw^0}{dx}(x - \xi)\right)$$

Next, letting $\tau \to 0$ we obtain

$$\frac{\partial f}{\partial t}(h;\, 0,\, x) = -\frac{dw^0}{dx}(x)$$

and, finally, letting $h \to 1$ and $x \to x_j$ we obtain

$$\frac{\partial u^0}{\partial t_j} = - \frac{dw^0}{dx}(x_j)$$

Similarly, differentiating f with respect to μ and letting $\tau \to 0$, $h \to 1$, and $x \to x_j$, we obtain

$$\frac{\partial u^0}{\partial \mu_j} = \frac{dw^0}{dx}(x_j)$$

Similarly, differentiating with respect to h and carrying out the same limiting operations, we obtain

$$\frac{\partial u^0}{\partial h_j} = 0$$

Now consider the difference equations to be solved:

$$\frac{\partial u}{\partial t}^{n+1}_j = \frac{1}{2}\left(\frac{\partial u}{\partial t}^n_{j+1} + \frac{\partial u}{\partial t}^n_{j-1}\right) - \frac{r}{2}\left(\frac{\partial u}{\partial t}^n_{j+1} - \frac{\partial u}{\partial t}^n_{j-1}\right)$$

$$\frac{\partial u}{\partial h}^{n+1}_j = \frac{1}{2}\left(\frac{\partial u}{\partial h}^n_{j+1} + \frac{\partial u}{\partial j}^n_{-1}\right) - \frac{r}{2}\left(\frac{\partial u}{\partial h}^n_{j+1} - \frac{\partial u}{\partial h}^n_{j-1}\right) - \Delta t\, R^n_j$$

where

$$R^n_j = \frac{\partial u}{\partial t}^{n+1}_j - \frac{1}{2r}\left(\frac{\partial u}{\partial x}^n_{j+1} - \frac{\partial u}{\partial x}^n_{j-1}\right) + \frac{1}{2}\left(\frac{\partial u}{\partial x}^n_{j+1} + \frac{\partial u}{\partial x}^n_{j-1}\right)$$

$$v^{n+1}_j = \frac{1}{2}\left(v^n_{j+1} + v^n_{j-1}\right) - \frac{r}{2}\left(v^n_{j+1} - v^n_{j-1}\right)$$

$$\frac{\partial u}{\partial u}^{n+1}_j = \frac{1}{2}\left(\frac{\partial u}{\partial x}^n_{j+1} + \frac{\partial u}{\partial x}^n_{j-1}\right) - \frac{r}{2}\left(\frac{\partial u}{\partial x}^n_{j+1} - \frac{\partial u}{\partial x}^n_{j-1}\right)$$

Note that if $\frac{\partial u}{\partial h}^n_j = \frac{\partial u}{\partial h}^0_j - \frac{\partial u}{\partial x}^0_j = 0$ for all j, then $\frac{\partial u}{\partial h}^n_j = 0$ for all indices n and j.

Now we consider a case in which the initial data w^0 has a discontinuity. In particular, let

$$w^0(x) = \begin{cases} w_\ell \text{ if } x < x_s \\ \\ w_r \text{ if } x > x_s \end{cases}$$

where w_ℓ and w_r are different constants. (The subscript s suggests that a shock is being described.) If x_s is distinct from all the numbers x_j, then the initial data for $\frac{\partial u}{\partial t}$, $\frac{\partial u}{\partial h}$, and $\frac{\partial u}{\partial x}$ are zero at all the x_j, and so the unacceptable error estimate $\frac{\partial u^n}{\partial h_j} = 0$ (for all n and j) would be obtained. To remedy this situation, we introduce a modified definition of f, which we term the second option. Let

$$f\left(h;\, 0,\, x_j\right) = w^0\left(x_j\right)$$

and

$$\frac{\partial f}{\partial x}\left(h;\, 0,\, x_j\right) = \frac{w^0\left(x_{j+1}\right) - w^0\left(x_{j-1}\right)}{2\Delta x} \quad (\text{where } x_j = j\Delta x)$$

Then spline-fit f in the interval (x_j, x_{j+1}). That is, in this interval let $f\,(h;\, 0,\, x)$ be defined as a cubic polynomial,

$$f(h;\, 0,\, x) = a_j\left(x - x_j\right)^3 + b_j\left(x - x_j\right)^2 + c_j\left(x - x_j\right) + d_j$$

where the coefficients a_j, b_j, c_j, d_j are chosen to provide a smooth fit, in the following sense:

$$f\left(h;\, 0,\, x_j\right) = w^0\left(x_j\right)$$

$$\frac{\partial f}{\partial x}\left(h;\, 0,\, x_j\right) = \left[w^0\left(x_{j+1}\right) - w^0\left(x_{j-1}\right)\right] \Big/ (2\Delta x)$$

$$f\left(h;\, 0,\, x_{j+1}\right) = w^0\left(x_{j+1}\right)$$

$$\frac{\partial f}{\partial x}\left(h;\, 0,\, x_j\right) = \left[w^0\left(x_{j+2}\right) - w^0\left(x_j\right)\right] \Big/ (2\Delta x)$$

The coefficients are easily seen to be uniquely determined by the above conditions, and are given by

$$a_j = \left(- w_{j-1} + \frac{3}{2} w_j - w_{j+1} + \frac{1}{2} w_{j+2}\right) \Big/ (\Delta x)^3$$

$$b_j = \left(\frac{5}{2} w_{j-1} - \frac{5}{2} w_j + \frac{1}{2} w_{j+1} - \frac{1}{2} w_{j+2}\right) \Big/ (\Delta x)^2$$

$$c_j = \frac{\partial f}{\partial x}(h;\, 0,\, x_j) = \left[w^o(x_{j+1}) - w^o(x_{j-1})\right] \Big/ (2\Delta x)$$

$$d_j = w^o(x_j)$$

Next, define $f(h;\, \tau,\, x)$ for $0 \leq \tau < \delta t$ by

$$f(h;\, t,\, x) = \frac{1}{2}\left[f(h;\, 0,\, x + \xi) + f(h;\, 0,\, x - \xi)\right]$$

$$- \frac{r}{2}\left[f(h;\, 0,\, x + \xi) - f(h;\, 0,\, x - \xi)\right]$$

where $\xi = \tau/r$. Taking the t-derivative of f and letting $\tau \to 0$, one obtains

$$\frac{\partial f}{\partial t}(h;\, 0,\, x) = - \frac{\partial f}{\partial x}(h;\, 0,\, x)$$

Now let $h \to 1$ and $x \to x_j$ in order to obtain the discrete initial conditions for

$$\frac{\partial u}{\partial t},\ \frac{\partial u^o}{\partial t} = -c_j$$

Similarly, differentiating f with respect to x and letting $\tau \to 0$, $h \to 1$, and $x \to x_j$ furnishes

$$\frac{\partial u^o}{\partial x_j} = c_j$$

Note that f is actually independent of h, therefore

$$\frac{du^o}{dh_j} = 0$$

Observe that the second option for f, in contrast to the first, does not lead to identically vanishing initial values for $\frac{\partial u}{\partial t}$ and $\frac{\partial u}{\partial x}$ when the initial condition is given by a step function (actually containing one or more discontinuities). Therefore, the second option is to be preferred when discontinuous data are prescribed.

An additional advantage of the second option is that the program user is not required (as in option 1) to input $\frac{dw}{dx}$, for in option 2 the program computes $[w^0(x_{j+1}) - w^0(x_{j-1})]/(2\Delta x)$ from the given $w^0(x_j)$. Thus, with option 2 the modified program (i.e., modified to furnish error estimates) does not require any information beyond that required by the unmodified program.

The next example illustrates the use of the procedure in the case of a nonlinear PDE. Consider the problem

$$\frac{\partial w}{\partial t} + \frac{\partial F(w)}{\partial x} = 0, \; w(0,x) = w^0(x)$$

where F is assumed to possess suitable smoothness properties. The LKL scheme furnishes the discrete-difference equation

$$v_j^{n+1} = \frac{v_{j+1}^n + v_{j-1}^n}{2} - r\,\frac{\left(F_{j+1}^n - F_{j-1}^n\right)}{2}$$

where $r = \Delta t/\Delta x$ and $F_j^n = F\left(v_j^n\right)$. The equation of first variation is

$$\delta_j^{n+1} = \frac{\delta_{j+1}^n + \delta_{j-1}^n}{2} - \frac{r}{2}\left(A_{j+1}^n\,\delta_{j+1}^n - A_{j-1}^n\,\delta_{j-1}^n\right)$$

where $A_j^n = F'\left(v_j^n\right)$. This can be rearranged to

$$\delta_j^{n+1} = \frac{1}{2}\left(1 - r_{j+1}^n\right)\delta_{j+1}^n + \frac{1}{2}\left(1 + r_{j-1}^n\right)\delta_{j-1}^n$$

where $r_j^n = rA_j^n$. A stability analysis shows that the condition $|r_j^n| \leq 1$ is necessary for stability. Therefore, we take $\delta t = h\Delta t$ and $\delta x = h\Delta x$.

The associated continuum-difference scheme is

$$u(h;\, t + st,\, x) = \frac{1}{2}\left[u(h;\, t,\, x + \delta x) + u(h;\, t,\, x - x)\right] - \frac{r}{2}\left[F(h;\, t,\, x + \delta x) - F(h;\, t,\, x - \delta x)\right]$$

where

$$F(h;\, t,\, x) = F[u(h;\, t,\, x)]$$

The initial condition is given by

$$u(h;\, \tau,\, x) = f(h;\, \tau,\, x) \text{ for } 0 \leq \tau < \delta t$$

We choose the second option for f:

$$f\left(h;\, 0,\, x_j\right) = w^0\left(x_j\right)$$

$$\frac{\partial f}{\partial x}(h;\, 0,\, x_j) = \left[w^0\left(x_{j+1}\right) - w^0\left(x_{j-1}\right)\right] \Big/ (2\delta x)$$

and then we spline-fit f, for x between x_j and x_{j+1}, by a cubic polynomial, exactly as in the previous example. Next we define

$$f(h;\, \tau,\, x) \text{ for } 0 \leq \tau < \delta t \text{ by } f(h;\, \tau,\, x) = \frac{1}{2}\left[f(h;0,\, x + \xi) + f(h;0,\, x - \xi)\right] - \frac{r}{2}\left[F^0(h;\, x + \xi) - F^0(h;\, x - \xi)\right]$$

where $\xi = \tau/r$ and $F^0(h;\, x) = F[f(h;\, 0,\, x)]$. Now, forming $\frac{\partial f}{\partial t}$ and letting $\tau \to 0$, we obtain

$$\frac{\partial f}{\partial t}(h;\, 0,\, x) = -\frac{\partial F^0}{\partial x}(h;\, x) = -A\left[f(h;\, 0,\, x)\right] \frac{\partial f}{\partial x}(h;\, 0,\, x)$$

where $A = F'$. Next, let $h \to 1$ and $x \to x_j$; one obtains

$$\frac{\partial u^0}{\partial t} = - A\left(w^0\left(x_j\right)\right) \frac{w^0\left(x_{j+1}\right) - w^0\left(x_{j-1}\right)}{2\Delta x}$$

Similarly, by differentiating f with respect to x and letting $\tau \to 0$, $h \to 1$, $x \to x_j$, we obtain

$$\frac{\partial u^0}{\partial x_k} = \frac{w^0\left(x_{j+1}\right) - w^0\left(x_{j-1}\right)}{2\Delta x}$$

Also, differentiating f with respect to h and letting $\tau \to 0$, $h \to 1$, $x \to x_j$ we obtain

$$\frac{\partial u^0}{\partial h_j} = \frac{dw^0\left(x_j\right)}{dx} - \frac{w^0\left(x_{j+1}\right) - w^0\left(x_{j-1}\right) x_j}{2\Delta x}$$

Observe that the foregoing procedure extends to the case when u, F, and f are vectors (in which case A becomes a matrix).

SECTION IV

NUMERICAL EXPERIMENTS

1. INTRODUCTION

The method of asymptotic error estimation and the error-gradient method were tested experimentally on four ODEs. In this section we present the numerical results in both tabular and graphical form.

2. THE ODEs AND OΔEs

In analogy with the standardized abbreviation ODE for ordinary differential equation we shall employ the abbreviation OΔE for ordinary difference equation.

The ODEs considered here are of the form $dw/dt = F(w)$. Four specific cases are discussed, namely: $F(w) = w$, $F(w) = - w^2$, $F(w) = 0.9\, w^2$, and $F(w) = - w^3$. In each of these cases the exact solution can be written explicitly, and this enables us to determine the true values of the errors incurred by the difference-scheme and thus to compare the accuracy of the several methods employed to estimate these errors. In this subparagraph IV-2, the exact solutions and the output of the difference-schemes are presented for each of the four equations; in subparagraph IV-3, the estimates furnished by the asymptotic error estimation procedure are worked out; while in subparagraph IV-4, the estimates obtained by the error gradient procedure are obtained.

In each case we take for the OΔE the Euler scheme:

$$V^{n+1} = V^n + \Delta t\, F(V^n)$$

The values of V^n are termed the exact-OΔE solution. Of course, the actual output of the computer is only an approximation to the values of V^n. Certain OΔEs will be generated whose solutions serve as estimates of the error incurred by using V^n as an approximation to the exact solution of the ODE. These are the error estimates, whose values are approximated by the computer output.

Of course, the real challenge is to obtain reliable error estimates for ODE and PDE problems whose exact solutions are not available; nevertheless, use of a proposed error-estimation procedure on some problems with known solutions is frequently employed to furnish an indication of how well the procedure will work

in more challenging cases. Furthermore, the analysis of simple cases often turns out to be valuable in debugging the programs that are written to obtain the desired output.

Case 1

$$\frac{dw}{dt} = w, \; w(0) = 1$$

The solution of the ODE problem is given by $w = e^t$, while the solution of the OΔE problem $\left[\text{namely, } V^{n+1} = (1 + \Delta t)V^n, \; V^0 = 1\right]$ is given by $V^n = (1 + \Delta t)^n$.

Remark: It is of interest to note that there exists in this case a difference scheme which yields the exact solution, namely $V^{n+1} = e^{\Delta t} V^n$. When such an exact difference scheme is known it can be useful in analyzing both the error and the error estimate. This is illustrated in paragraph 3, in the discussion of the error estimate furnished by the asymptotic error estimation procedure for Case 1.

Case 2

$$\frac{dw}{dt} = - w^2, \; w(0) = 1$$

The exact solution of the ODE problem is given by $w = 1/(1 + t)$. The OΔE problem is

$$V^{n+1} = \left(1 - \Delta t V^n\right) V^n, \; V^0 = 1$$

While the computational scheme is quite trivial, the exact solution of the OΔE problem cannot be expressed in simple form. However, one can obtain a partial representation by the following device.

Replacing Δt by $-k$ and then replacing kV^n by W^n, we convert the above OΔE into the parameter-free form

$$W^{n+1} = \left(1 + W^n\right) W^n \tag{51}$$

with initial condition $W^0 = k$. Thus, $W^1 = k + k^2$, $W^2 = (k + k^2) + (k + k^2)^2 = k + 2k^2 + 2k^3 + k^4$, $W^3 = (k + 2k^2 + 2k^3 + k^4) + (k + 2k^2 + 2k^3 + k^4)^2 = k + 3k^2 + 6k^3 + 9k^4 + 10k^5 + 8k^6 + 4k^7 + k^8$, etc. By induction, it is clear that W^n is

a polynomial of degree 2^n in k, with positive integer coefficients (except for the vanishing constant term). Symbolically, we can write

$$W^n = a_N^n, a_{N-1}^n, \cdots, a_1^n, \text{ where } N = 2^n$$

Then the recurrence relation (eq. 51) furnishes a recurrence relation for the coefficients a_m^n, namely,

$$a_m^{n+1} = a_m^n + \sum_j a_j^n a_{m-j}^n$$

where it is understood that $a_m^n = 0$ when $m < 1$ and when $m \geq N$. This is what we mean by a partial representation; it is useful for debugging purposes.

Remark: It is interesting to note that the exact solution to the OΔE problem

$$D^{n+1} = D^n\left(1 + kD^{n+1}\right), \; D^0 = 1$$

is given by

$$D^n = (1-nk)^{-1}$$

Upon replacing k by $-\Delta t$ we obtain $D^n = (1 + n\Delta t)^{-1}$, which agrees exactly with the solution of the ODE problem under discussion here. Thus, as in Case 1, a slight modification of the Euler scheme turns out to furnish the exact solution.

Case 3

$$\frac{dw}{dt} = 0.9\, w^2, \; w(0) = 1$$

The exact solution of the ODE problem is given by $w = 1/(1 - 0.9\, t)$ for $0 \leq t < 1$.

$$V^{n+1} = (1 + 0.9\, \Delta t V^n)V^n, \; V^0 = 1$$

As in Case 2, no convenient representation of V^n is available; of course, the entire discussion of the preceding case is applicable if certain minor changes are carried out.

Case 4

$$\frac{dw}{dt} = - w^3, \; w(0) = 1$$

The exact solution of the ODE problem is given by $w = (1 + 2t)^{-\frac{1}{2}}$. The O$\Delta$E problem is

$$V^{n+1} = \left(1 - \Delta t \, (V^n)^2\right) V^n, \; V^0 = 1$$

A partial representation of the solution of the OΔE problem can be obtained by following, in a rather obvious manner, the above discussion of Case 2.

3. ASYMPTOTIC ERROR ESTIMATES

In this section a brief review is given of the asymptotic error estimation procedure for ODEs, and then numerical results for each of the four specific cases listed in paragraph 2 are presented. Finally, a typical FORTRAN program for performing the required computations is given.

a. Review of the Procedure for ODEs

Given the ODE $\frac{dw}{dt} = F(w)$, let ϕ and ϕ_E be two finite-difference approximations of the operator $\frac{dw}{dt} - F$, ϕ_E being of higher order than ϕ. For example,

$$\phi(V)^n = \frac{V^{n+1} - V^n}{\Delta t} - F(V^n)$$

and

$$\phi_E(V)^n = \frac{V^{n+1} - V^n}{\Delta t} - F\left(\frac{V^n + V^{n+1}}{2}\right)$$

are first-order and second-order approximations, respectively. Let V be the solution to $\phi(V) = 0$ and let V_E be the solution to $\phi\left(V_E\right) + \phi_E(V) = 0$, each satisfying the prescribed initial condition. (The standing assumption is made that the solutions V and V_E exist and are unique.) Then $V - V_E$ is defined to

be the asymptotic error estimate. That is, $V - V_E$ is taken as an approximation to be actual error $V - w$ incurred by using the lower-order scheme $\phi(V) = 0$.

For a given choice of ϕ there is wide freedom in choosing the higher-order approximation ϕ_E to $\frac{d}{dt} - F$. In order to estimate what effect the latitude in choosing ϕ_E may actually have in practice, we have considered two specific choices in our numerical experiments:

$$\text{Option A:}\quad \phi_E^{(A)}(V)^n = \frac{V^{n+1} - V^n}{\Delta t} - F\left(\frac{V^{n+1} + V^n}{2}\right)$$

$$\text{Option B:}\quad \phi_E^{(B)}(V)^n = \frac{V^{n+1} - V^n}{\Delta t} - \frac{F\left(V^{n+1}\right) + F\left(V^n\right)}{2}$$

While the lower-order approximation ϕ was chosen as above, namely

$$\phi(V)^n = \frac{V^{n+1} - V^n}{\Delta t} - F\left(V^n\right)$$

The solution of the lower-order scheme $\phi(V) = 0$ is simply denoted as V, while the solutions of the two higher-order schemes are denoted $V_E^{(A)}$ and $V_E^{(B)}$; that is,

$$\phi(V) = 0$$

$$\phi\left(V_E^{(A)}\right) + \phi_E^{(A)}(V) = 0$$

$$\phi\left(V_E^{(B)}\right) + \phi_E^{(B)}(V) = 0$$

From the theory of asymptotic error approximations (ref. 1), the difference $V - V_E = V - w + O(\Delta t^2)$. [Therefore, $V_E^{(A)} - V_E^{(B)} = O(\Delta t^2)$.] What our numerical experiments will provide is some indication of the magnitude of the constants implied in the $O(\Delta t^2)$ terms for each of the options A and B.

In order to give a heuristic overview of the asymptotic error estimation theory, we illustrate how it works in Case 1, where all the calculations can be

worked out quite explicitly. [Recall that $F(w) = w$, $w(0) = 1$.] As pointed out previously, in this case there exists an exact difference scheme, namely $V^{n+1} = e^{\Delta t} V^n$; thus, if

$$\Delta(V)^n = \frac{V^{n+1} - e^{\Delta t} V^n}{\Delta t}$$

the difference operator Δ is, in an obvious sense, an exact discretization of $D = \frac{d}{dt}(\cdot) - (\cdot)$; that is, the operator $\frac{d}{dt}$ - identity. The existence of an exact difference scheme, while exceptional, is helpful in illustrating just what is occurring in the calculations.

Now, let ϕ_k be an $O(\Delta t^k)$ approximation to D; then

$$\phi_k(V)^N = \Delta(V)^n - K_k^n \Delta t^k$$

where $K_k^n = O(1)$, a bounded quantity. In particular, let

$$f_k(\Delta t) = 1 + \frac{\Delta t}{2!} + \frac{\Delta t}{3!} + \cdots + \frac{\Delta t^{k-1}}{k!} \quad (k = 1,2,3,\cdots)$$

and let

$$\phi_k(V)^n = \frac{V^{n+1} - V^n}{\Delta t} - f_k(\Delta t) V^n$$

Since

$$\phi_k(V^n) = \Delta(V)^n + \left(\frac{e^{\Delta t} - 1 - \Delta t f_k(\Delta t)}{\Delta t}\right) V^n$$

it follows that ϕ_k is, in fact, an $O(\Delta t^k)$ approximation to D, with

$$-K_k^n \Delta t^k = \left(\frac{e^{\Delta t} - 1 - \Delta t f_k(\Delta t)}{\Delta t}\right) V^n$$

or

$$K_k^n = k_k^n(V) = O(1) V^n + O(\Delta t)$$

From the above formulas we immediately confirm the following observation:

Lemma A: For $k = 1, 2, 3, \cdots$ the quantity $K_k^n(V)$ satisfies a Lipschitz condition modulo $o(\Delta t)$; that is, there exists a positive constant $\mathfrak{L}$ such that

$$|K_k^n(V_1) - K_k^n(V_2)| \leq \mathfrak{L}|V_1^n - V_2^n|$$

Two additional observations appear worthy of note. The proofs, which are rather straightforward, are omitted.

Lemma B: If ϕ_k is an $o(\Delta t^k)$ approximation to D, then the actual error is $o(\Delta t^k)$. That is, let $\phi_k(V) = 0$ and $D(w) = 0$; then $V^n - w(t^n) = o(\Delta t^k)$.

Lemma C. Let k and ℓ be positive integers, $k < \ell$, and let ϕ_k and ϕ be $o(\Delta t^k)$ and $o(\Delta t^\ell)$ approximations, respectively, to D. Let w be the solution to $D(w) = 0$, let V be the solution to $\phi_k(V) = 0$, and let V_E be the solution to $\phi_k(V_E) + \phi_\ell(V) = 0$. Then

$$V^n = w(t^n) + o(\Delta t^k) \tag{52-a}$$

$$V_E^n = w(t^n) + o(\Delta t^{2k}) + o(\Delta t^\ell) \tag{52-b}$$

$$V^n - V_E^n = V^n - w(t^n) + o(\Delta t^{2k}) + o(\Delta t^\ell) \tag{52-c}$$

If in Lemma C we take $\ell = 2k$, then the three conclusions assume the form

$$V^n = w(t^n) + o(\Delta t^k) \tag{53-a}$$

$$V_E^n = w(t^n) + o(\Delta t^{2k}) \tag{53-b}$$

$$V_E^n = V^n - w(t^n) + o(\Delta t^{2k}) \tag{53-c}$$

Thus, in Case 1, it becomes very clear why $V^n - V_E^n$ constitutes an estimate of $V^n - w(t^n)$ of order 2k.

Now, as indicated above, we present some numerical results for each of the four problems listed in subparagraph IV-2.

Case 1

In this case, options A and B coincide:

$$\phi_E(V)^n = \frac{V^{n+1} - V^n}{\Delta t} - \frac{V^{n+1} + V^n}{2}$$

and therefore

$$V_E^{n+1} = (1 + \Delta t)\, V_E^n + (\Delta t)^2\, V^n/2$$

By iteration it is easily shown that this equation (together with the initial condition $V_E^0 = 1$) implies

$$V_E^n = V^n + (1 + \Delta t)^{n-1} \cdot t^n \Delta t/2$$

Therefore, the exact-0ΔE error estimate produced by the asymptotic error estimation procedure in the present case (for either option) is given by

$$V^n - V_E^n = -(1 + \Delta t)^{n-1} \cdot t^n \Delta t/2$$

while the actual error is $V^n - e^{n\Delta t}$, or $(1 + \Delta t)^n - e^{n\Delta t}$, which is equal to

$$-e^{n\Delta t} \cdot t^n\, \Delta t/2 + o(\Delta t^2)$$

and so

$$V^n - V_E^n = V^n - e^{n\Delta t} + o(\Delta t^2)$$

Choosing $\Delta t = 0.1$ and $n = 10$ we obtain: $w(1) = e \approx 2.718$, $V^{10} = (1.1)^{10} \approx 2.594$, and $V_E^{10} \approx 2.712$. The actual error is therefore

$$V^{10} - w(1) \approx -0.124$$

while the asymptotic error estimate is

$$V^{10} - V_E^{10} \approx - 0.118$$

Upon reducing Δt to 0.05 and correspondingly increasing n to 20, we obtain $w(1) = e \approx 2.718$, $V^{20} \approx 2.653$, and $V_E^{20} \approx 2.716$. Thus, the actual error is now

$$V^{20} - w(1) \approx - 0.065$$

while the asymptotic error estimate becomes

$$V^{20} - V_E^{20} \approx - 0.063$$

Case 2

In this case, option A reduces to

$$\phi_E^{(A)}(V)^n = (V^n)^4 \, (\Delta t)^2/4 - (V^n)^3 \Delta t$$

while option B assumes the form

$$\phi_E^{(B)}(V)^n = (V^n)^4 \, (\Delta t)^2/2 - (V^n)^3 \Delta t$$

As in Case 1, we first choose $\Delta t = 0.1$ and $n = 10$, then $\Delta t = 0.05$ and $n = 20$. With the first choice we obtain $w(1) = 0.5$, $V^{10} \approx 0.48171$, $V_{EA}^{10} \approx 0.49976$, and $V_{EB}^{10} \approx 0.49943$. Thus, the actual error is

$$V^{10} - w(1) \approx - 0.01829$$

while option A furnishes

$$V^{10} - V_{EA}^{10} \approx - 0.01805$$

and option B furnishes

$$V^{10} - V_{EB}^{10} \approx - 0.01772$$

With the second choice of Δt and n we obtain $w(1) = 0.5$, $V^{20} \approx 0.49110$, $V_{EA}^{20} \approx 0.49994$, and $V_{EB}^{20} \approx 0.49986$, and so

$$V^{20} - w(1) \approx - 0.00890$$

$$V^{20} - V_{EA}^{20} \approx - 0.00884$$

$$V^{20} - V_{EB}^{20} \approx - 0.00876$$

Case 3

With $\Delta t = 0.1$ and $n = 10$ we obtain $w(1) = 10$, $V^{10} \approx 4.46663$, $V_{EA}^{10} \approx 9.06703$, $V_{EB}^{20} \approx 9.20051$. Thus, the actual error is

$$V^{10} - w(1) \approx - 5.53337$$

while option A furnishes the error estimate

$$V^{10} - V_{EA}^{10} \approx - 4.60040$$

and option B yields

$$V^{10} - V_{EB}^{10} \approx - 4.73388$$

When Δt is reduced to 0.05 and n is increased to 20, we obtain $w(1) = 10$, $V^{20} \approx 5.72982$, $V_{EA}^{20} \approx 13.80567$, $V_{EB}^{20} \approx 13.92144$. Thus, the actual error, option A estimate, and option B estimate are, respectively,

$$V^{20} - w(1) \approx -4.27018$$

$$V^{20} - V_{EA}^{20} \approx -8.07585$$

$$V^{20} - V_{EB}^{20} \approx -8.19162$$

The rather violent behavior of the output is perhaps to be expected in view of the rapid growth of the exact solution. However, the fact that the error estimates obtained with $\Delta t = 0.05$ are worse than those obtained with $\Delta t = 0.1$ is somewhat unexpected.

Case 4

With $\Delta t = 0.1$ and $n = 10$, we obtain $w(1) = 3^{-\frac{1}{2}} \approx 0.57735$, $V^{10} \approx 0.56044$, $V_{EA}^{10} \approx 0.57679$, $V_{EB}^{10} \approx 0.57628$. Thus, the actual error is

$$V^{10} - w(1) \approx -0.01691$$

while option A furnishes the error estimate

$$V^{10} - V_{EA}^{10} \approx -0.01635$$

and option B yields

$$V^{10} - V_{EB}^{10} \approx -0.01584$$

When Δt is reduced to 0.05 and n is increased to 20, we obtain $w(1) \approx 0.57735$, $V^{20} \approx 0.56917$, $V_{EA}^{20} \approx 0.57722$, $V_{EB}^{20} \approx 0.57709$.

Thus, the actual error, option A estimate, and option B estimate are, respectively,

$$V^{20} - w(1) \approx -0.00818$$

$$V^{20} - V_{EA}^{20} \approx -0.00805$$

$$V^{20} - V_{EB}^{20} \approx -0.00792$$

b. FORTRAN Program for Asymptotic Error Computations

For option A, with $\Delta t = 0.1$, the required computations were carried out with the very simple FORTRAN program which appears in figure 1. Reduction of Δt to 0.05 required only a change in line 11, while option B required only very simple changes in lines 19 through 22.

```
      PROGRAM ASYMAP1(INPUT,OUTPUT)
      T=0.0
      V1=1.0
      V2=1.0
      V3=1.0
      V4=1.0
      W1OLD=1.0
      W2OLD=1.0
      W3OLD=1.0
      W4OLD=1.0
      H=0.1
   23 PRINT 20,T,V1,V2,V3,V4,W1OLD,W2OLD,W3OLD,W4OLD
   20 FORMAT(F5.2,8F15.10)
      W1NEW=W1OLD+H*W1OLD
      W2NEW=W2OLD-H*W2OLD**2
      W3NEW=W3OLD+0.9*H*W3OLD**2
      W4NEW=W4OLD-H*W4OLD**3
      T=T+H
      V1=V1+H*V1-W1NEW+W1OLD+H*(W1OLD+W1NEW)/2.0
      V2=V2-H*V2**2-W2NEW+W2OLD-H*((W2OLD+W2NEW)/2.0)**2
      V3=V3+H*V3**2-W3NEW+W3OLD+0.9*H*((W3OLD+W3NEW)/2.0)**2
      V4=V4-H*V4**3-W4NEW+W4OLD-H*((W4OLD+W4NEW)/2.0)**3
      W1OLD=W1NEW
      W2OLD=W2NEW
      W3OLD=W3NEW
      W4OLD=W4NEW
      IF(T.LT.1.0)GO TO 23
      STOP
      END
```

Figure 1. FORTRAN Program for Asymptotic Error Computations

4. ERROR GRADIENT ESTIMATES

We begin this section with a very brief review of the error gradient procedure, which is the subject of section III. For simplicity in exposition, the ODE problem will be assumed to be of the form $\frac{dw}{dt} = F(w)$, $w(0) = 1$, and the finite-difference technique to be employed will be taken as the Euler method,

$$V^{n+1} = V^n + hF(V^n),\ V^0 = 1 \tag{54}$$

The discrete function V, with values V^0, V^1, V^2,... is then filled in, in the initial interval [0,h], more or less arbitrarily, by a function U(h,t) which satisfies $U(h,0) = V^0 = 1$ and $U(h,t) = V^1$. Then the definition of U(h,t) for $t > h$ is accomplished by using the obvious extension of equation 54, namely

$$U(h,t + h) = U(h,t) + hF\left(U(h,t)\right)$$

For any given value of t (for simplicity, t is chosen equal to 1 in the illustrative computations which follow), $\frac{\partial U}{\partial h}$ is determined as a function of h. [If U(h,t) lacks sufficient smoothness, $\frac{\partial U}{\partial h}$ must be suitable interpreted.] Then, for any given (positive) value H of h, one obtains

$$U(H,1) - \lim_{h \to 0} U(h,1) = \int_0^H \frac{\partial U(h,1)}{\partial h}\, dh$$

Under very mild hypotheses the limit appearing in this equation exists and equals w(1), the value of the actual solution at t = 1. Thus, the error is given by

$$w(1) - U(H,1) = -\int_0^H \frac{\partial U(h,1)}{\partial h}\, dh$$

While a rigorous justification has not been obtained, the estimate

$$-\int_0^H \frac{\partial U(h,1)}{\partial h}\, dh \approx -H \left.\frac{\partial U(h,1)}{\partial h}\right|_{h=H} \tag{55}$$

has furnished remarkably good results in the four cases which are considered in this report.

In each of the four problems given in paragraph 2, three definitions of U(t,h) were employed in the interval $0 \leq t \leq h$. First, U(t,h) was defined to be linear in this interval; second, U(t,h) was defined to be quadratic in the initial interval, the three available coefficients being so chosen that

$U(h,o) = V^o = 1$, $U(h,h) = V^1$, and $\frac{\partial U(h,t)}{\partial t}$ remains differentiable when $t = h$ (and therefore at all succeeding nodes); and, third, $U(t,h)$ was defined to be cubic in the initial interval, the four available coefficients being chosen so that, in addition to the three constraints imposed in the quadratic case, $U(h,t)$ possesses a continuous second derivative at the nodes.

Thus, 12 computations were carried out altogether; in each case $U(h,1)$ was determined for $h = \frac{10}{10\,n + k - 1}$, $10 \leq n \leq 100$, $1 \leq k \leq 10$, so that h assumed the values $\frac{10}{100}, \frac{10}{101}, \frac{10}{102}, \cdots \frac{10}{1009}$. The quantities $\frac{\partial U(h,1)}{\partial h}$, the error estimates $- h \frac{\partial U(h,1)}{\partial h}$, and the corrected estimates $U(h,1) - h \frac{\partial U(h,1)}{\partial h}$ were also computed; this final quantity was also plotted. The 12 plots (figures 3 through 14) thus obtained are included at the end of this report.*

We list here a few of the numerical results (tables 1 through 4); however, the plots give a very clear picture of the manner in which the corrected estimates approach the actual solution as $h \to 0$.

c. FORTRAN Program for Error Gradient Computations

The FORTRAN program for Case 2 ($w^1 = -w^2$, $w(0) = 1$) with quadratic interpolation appears in figure 2; this program provides for plotting the corrected estimate of the solution, namely, $U(h,1) - h \frac{\partial U(h,1)}{\partial h}$. Rather straightforward modifications of this program were used for the other 11 computations.

*The captions of the plots are taken from the titles of the computer programs: CASE1L, CASE1Q, ..., CASE4C. The numeral indicates which of the four differential equations is being solved, while the final letter refers to linear, quadratic, or cubic interpolation.

TABLE 1

$U(h,1) - h\frac{\partial U}{\partial h} - w(1)$, CASE 1 $[w(1) = e \approx 2.718282]$

h	CASE 1L*	CASE 1Q	CASE 1C
$\frac{1}{10}$	.111255	-.012256	-.010388
$\frac{1}{10.5}$	-.013551	-.008648	-.009326
$\frac{1}{11}$	-.114083 .102934	-.010292	-.008722
$\frac{1}{20}$	-.064984 .061363	-.003351	-.002838
$\frac{1}{20.5}$	-.003897	-.002462	-.002707
$\frac{1}{21}$	-.062019 .058721	-.003053	-.002585
$\frac{1}{99}$	-.013603 .013444	-.000147	-.000125
$\frac{1}{99.5}$	-.000179	-.000112	-.000123
$\frac{1}{100}$	-.013468 .013312	-.000144	-.000122

*In the case of linear interpolation $\frac{\partial U}{\partial h}$ is discontinuous when h is the reciprocal of an integer; the left-hand and right-hand limits are therefore listed.

TABLE 2

$U(h,1) - h\frac{\partial U(h,1)}{\partial h} - w(1)$, CASE 2 [w(1) = 0.5]

h	CASE 2L	CASE 2Q	CASE 2C
$\frac{1}{10}$	.026860	.002606	.001040
$\frac{1}{10.5}$	.000918	.000220	.000932
$\frac{1}{11}$	-.024940 .024246	.002113	.000846
$\frac{1}{20}$	-.013133 .012937	.000588	.000239
$\frac{1}{20.5}$	.000226	.000061	.000227
$\frac{1}{21}$	-.012477 .012301	.000531	.000216
$\frac{1}{99}$	-.002550 .002542	.000022	.000009
$\frac{1}{99.5}$	.000009	.000003	.000009
$\frac{1}{100}$	-.002524 .002517	.000022	.000009

TABLE 3

$U(h,1) - h\frac{\partial U}{\partial h} - w(1)$, CASE 3 [w(1) = 10]

h	CASE 3L	CASE 3Q	CASE 3C
$\frac{1}{10}$	-2.674869	-3.869900	-3.825690
$\frac{1}{10.5}$	-3.744252	-3.699798	-3.720758
$\frac{1}{11}$	-4.679396 -2.477286	-3.661614	-3.621111
$\frac{1}{20}$	-3.387269 -1.320952	-2.398601	-2.376689
$\frac{1}{20.5}$	-2.341970	-2.317923	-2.328550
$\frac{1}{21}$	-3.284377 -1.236982	-2.302749	-2.282015
$\frac{1}{99}$	-.896927 .168615	-.368956	-.366563
$\frac{1}{99.5}$	-.365858	-.362621	-.363807
$\frac{1}{100}$	-.888060 .170631	-.363436	-.361083

TABLE 4

$U(h,1) - h\frac{\partial U}{\partial h} = w(1)$, CASE 4 $[w(1) = 3^{-\frac{1}{2}} \approx 0.577350]$

h	CASE 4L	CASE 4Q	CASE 4C
$\frac{1}{10}$	.031889	.004662	.001207
$\frac{1}{10.5}$	.001264	-.000510	.001067
$\frac{1}{11}$	-.030592 .028680	.003723	.000970
$\frac{1}{20}$	-.015641 .015101	.000976	.000262
$\frac{1}{20.5}$	.000306	-.000092	.000248
$\frac{1}{21}$	-.014836 .014349	.000879	.000237
$\frac{1}{99}$	-.002961 .002941	.000035	.000010
$\frac{1}{99.5}$	.000012	-.000003	.000010
$\frac{1}{100}$	-.002931 .002911	.000034	.000010

```
      PROGRAM CASE2Q(INPUT,OUTPUT,TAPE99=0)
      DIMENSION V(105),DV(105),X(1,5),Y(1,5)
      L=1
      DO 10 N=10,100
      DO 10 K=1,10
      H=10.0/(10.0*N+K-1.0)
      L=L+1
      X(L)=H
      V(1)=1.0-N*H*(1.0-N*H)/(1.0-H)
      DV(1)=(2.0*H*N**2-N-(N*H)**2)/(1.0-H)**2
      M=N+1
      DO 17 I=1,M
      V(I+1)=V(I)-H*V(I)**2
   17 DV(I+1)=DV(I)-V(I)**2-2.0*H*V(I)*DV(I)
      Y(L)=V(N+1)-H*DV(N+1)
      PRINT 20,N,K,H,V(N+1),DV(N+1),H*DV(N+1),V(N+1)-H*DV(N+1)
     A,V(N+1)-H*DV(N+1)-0.5
   20 FORMAT(2I5,6F15.10)
   10 CONTINUE
      CALL META
      CALL LINPLT(X,Y,L)
      CALL DONEPL
      STOP
      END
```

Figure 2, FORTRAN Program for Error Gradient Computations

SECTION V

CONCLUSIONS AND RECOMMENDATIONS

Admittedly, the computations on the error-gradient method reported here are much more extensive than those on the asymptotic method. In Cases 1, 2, and 4, the results obtained with the asymptotic method are better than those obtained by using the error-gradient method with cubic interpolation, which in turn is superior to the error-gradient method with either linear or quadratic interpolation. However, in Case 3, reduction of the step-size from h = 1/10 to h = 1/20 results in a distinct worsening of the asymptotic results, while all three interpolation methods furnish improved values. All results in Case 3 are disappointing, but this is to be expected in view of the rapid growth of the solution.

A notable phenomenon is the bracketing effect obtained in all four cases when linear interpolation is employed (see figures 3, 6, 9, and 12 for Cases 1L, 2L, 3L, 4L, respectively). Except in Case 3, where irregular behavior is to be expected, the bracketing sets in from the very beginning (h = 1/10), and the upper and lower envelopes are essentially linear with intercept at the exact solution. In Case 3 oscillation appears immediately, but bracketing does not appear until h has diminished to 0.02, approximately. It would certainly be helpful if one could prove that, when linear interpolation is employed, the error-gradient method furnishes values which bracket the true solution for sufficiently small values of h.

It is also of interest to note that when linear interpolation is employed and the left- and right-hand limits are averaged when h is the reciprocal of an integer, a very marked improvement is obtained. Thus, in Case 1, the <u>average</u> error when h = 1/20 is $\approx$ - 0.0018, while V_E^{20} - w(1) $\approx$ - 0.002 (see paragraph 3); similarly, in Case 2 the average error when h = 1/20 is $\approx$ - 0.00020, while the errors V_{EA}^{10} - w(1) and V_{EB}^{10} - w(1) are approximately - 0.00024 and - 0.00057, respectively. In Case 4, the three corresponding numbers are - 0.00027, - 0.00056, and - 0.00107. Closely related to these observations is the fact that when h is the reciprocal of a half-integer (e.g., 1/20.5) the error is noticeably dimished; e.g., h = 1/20.5 gives an error much smaller than the errors obtained either with h = 1/20 or h = 1/21.

It is by no means clear that a satisfactory theory can be worked out, but it certainly appears that further numerical experimentation and theoretical investigations are justified.

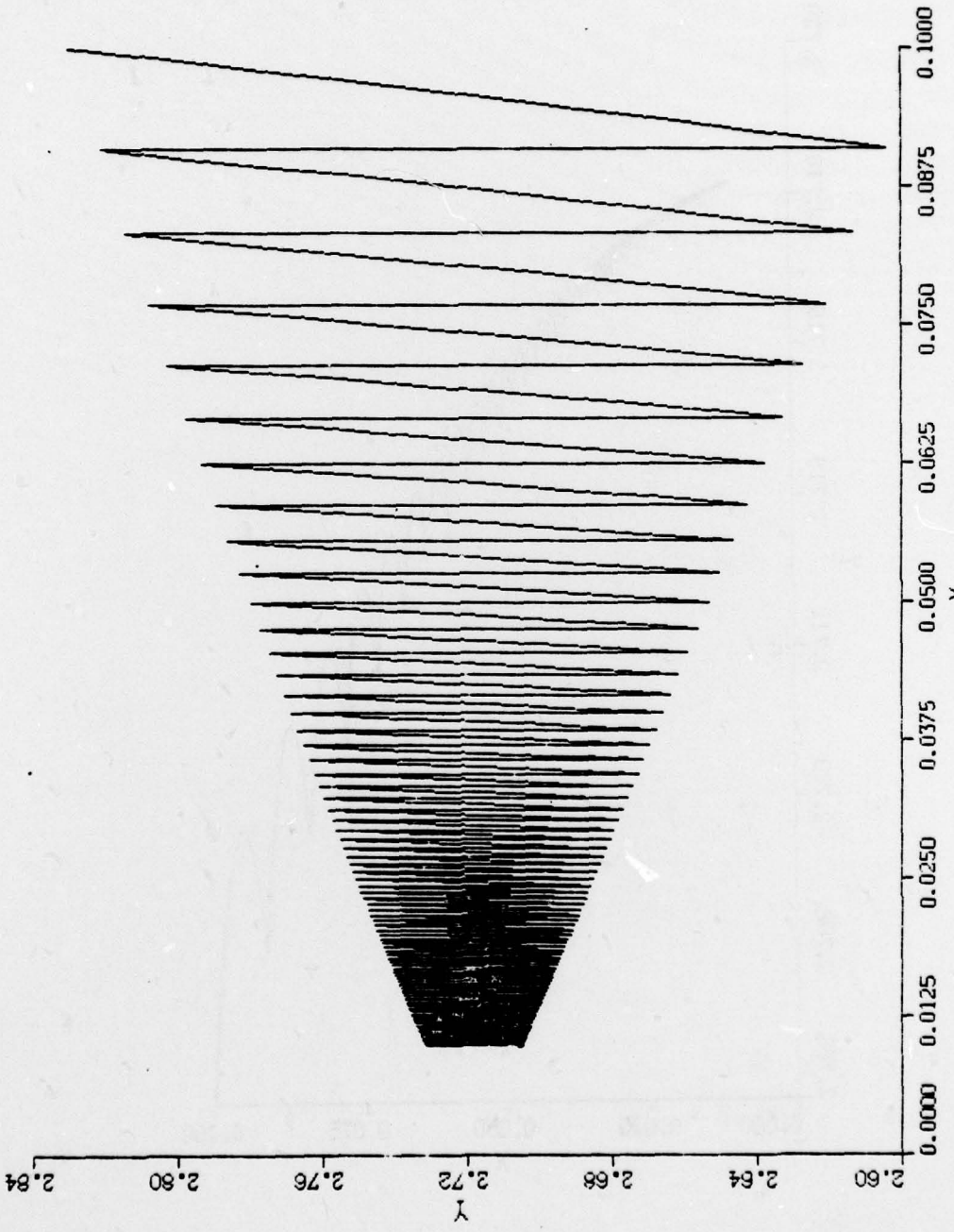


Figure 3. Case 1L

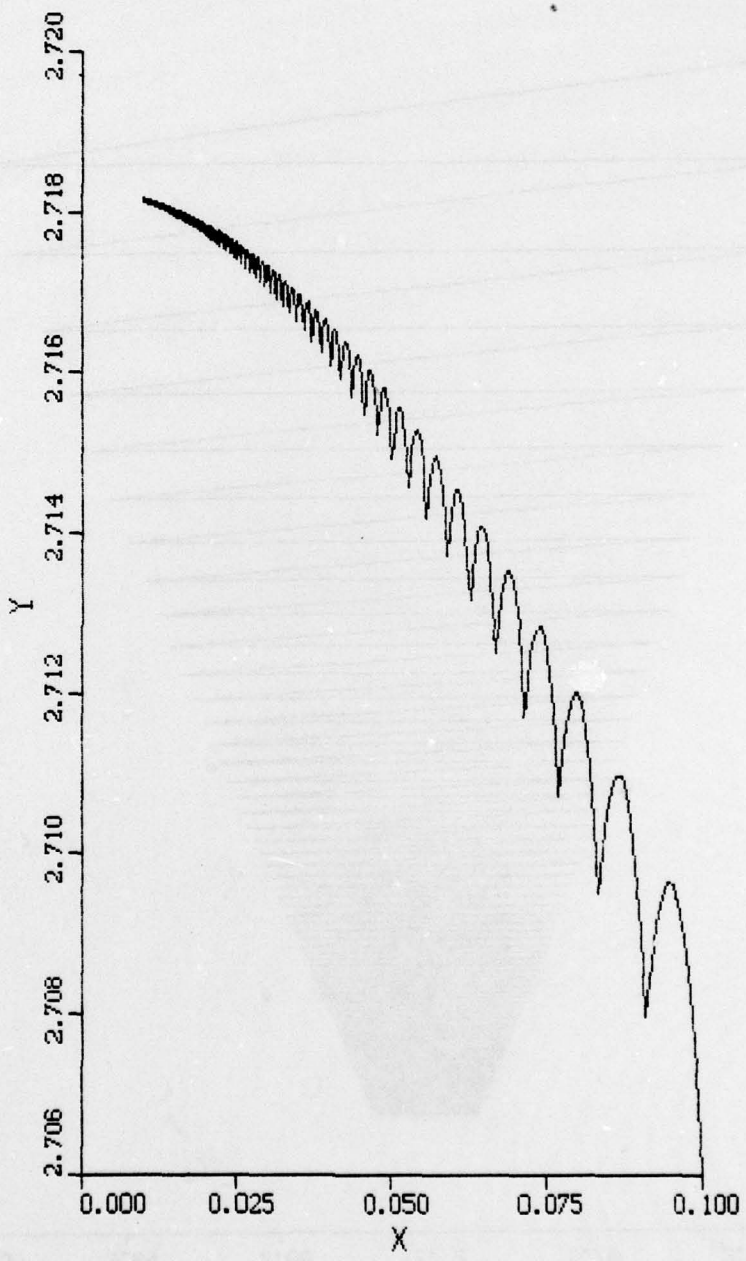


Figure 4. Case 1Q

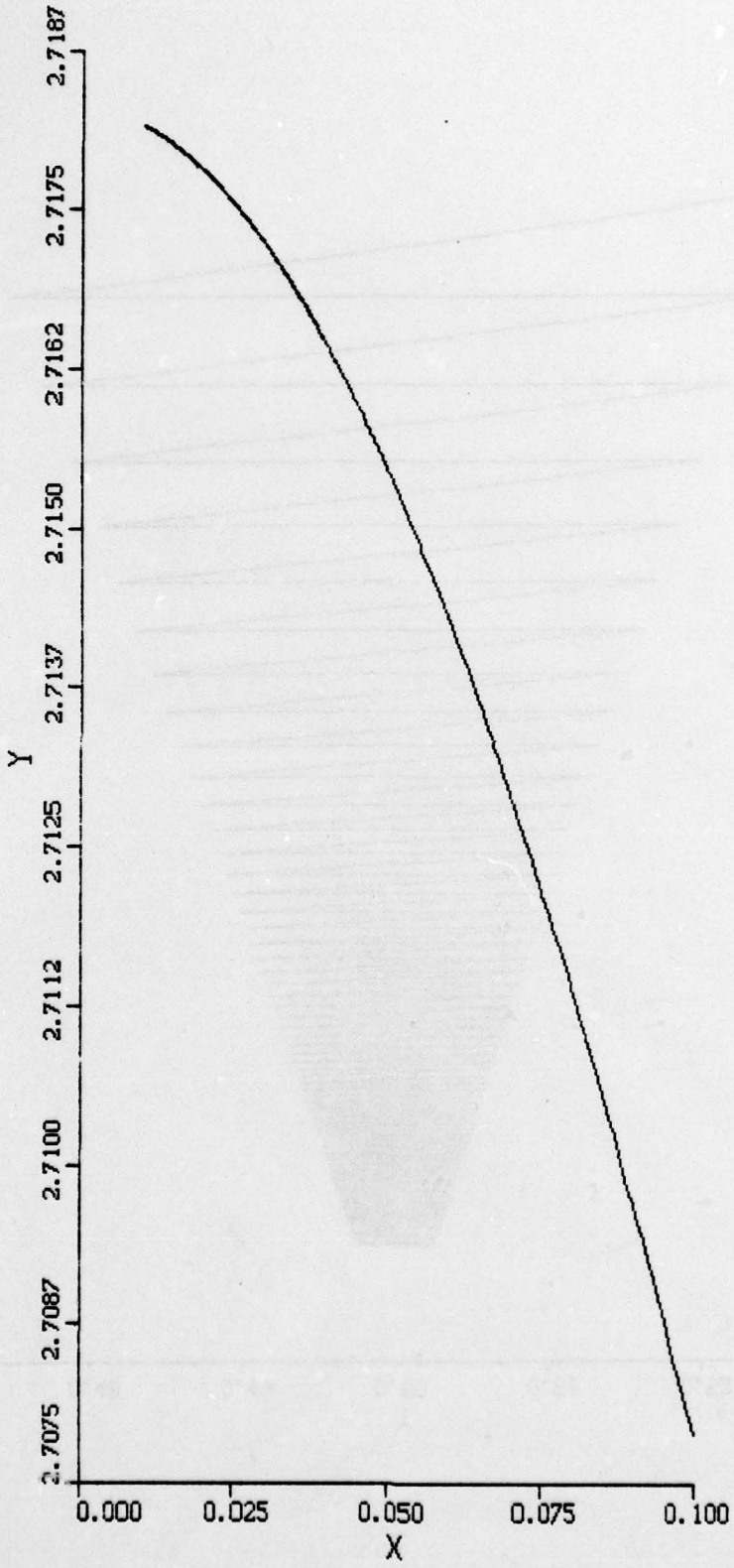


Figure 5. Case 1C

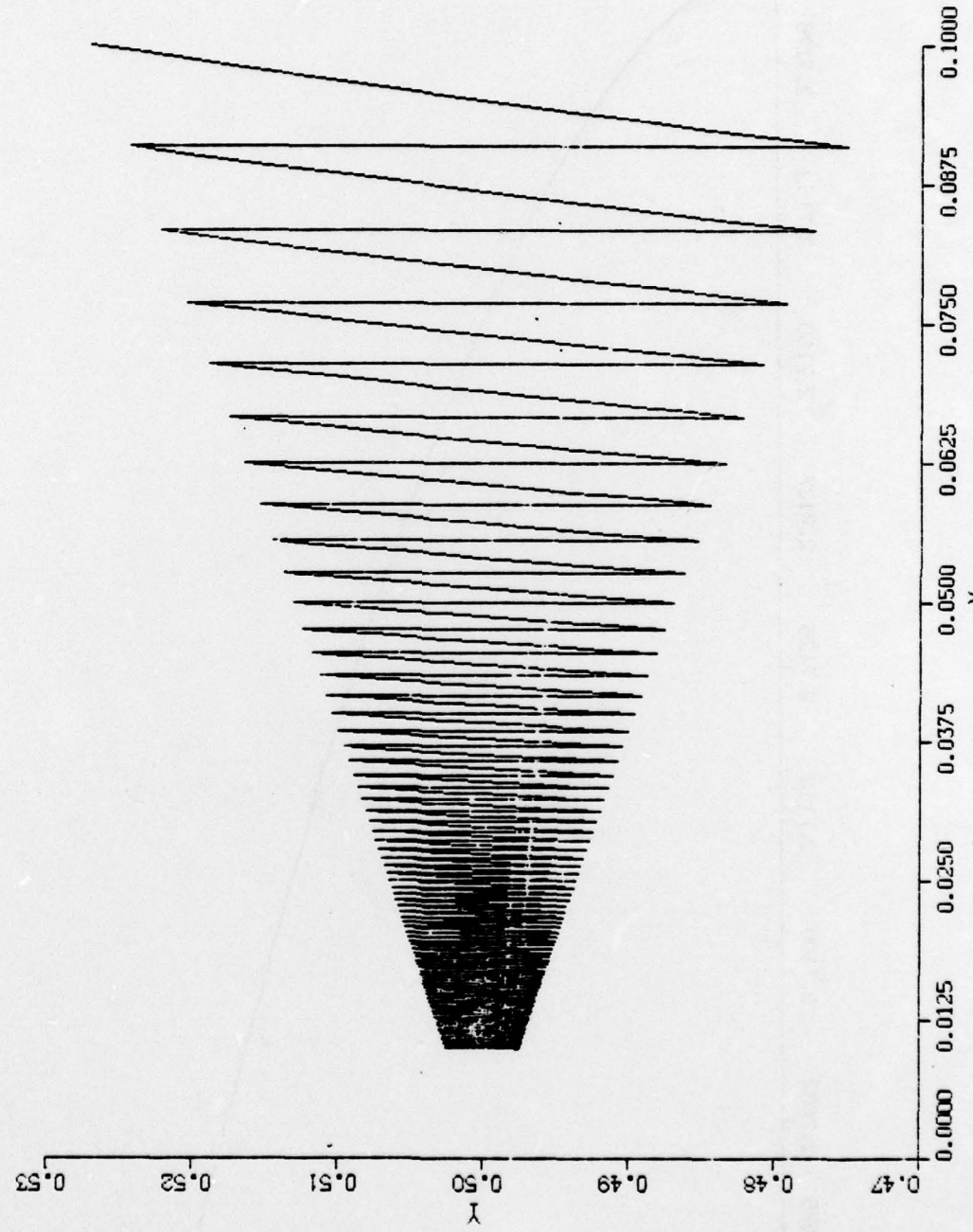


Figure 6. Case 2L

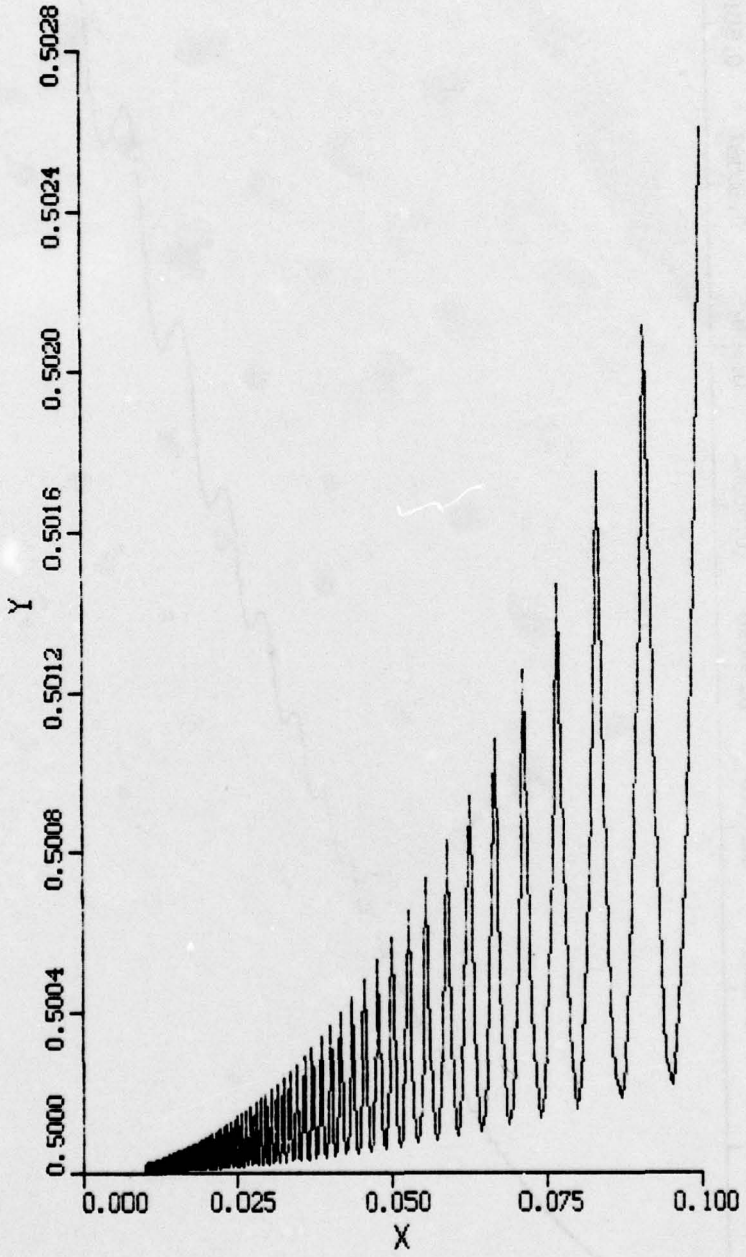


Figure 7. Case 2Q

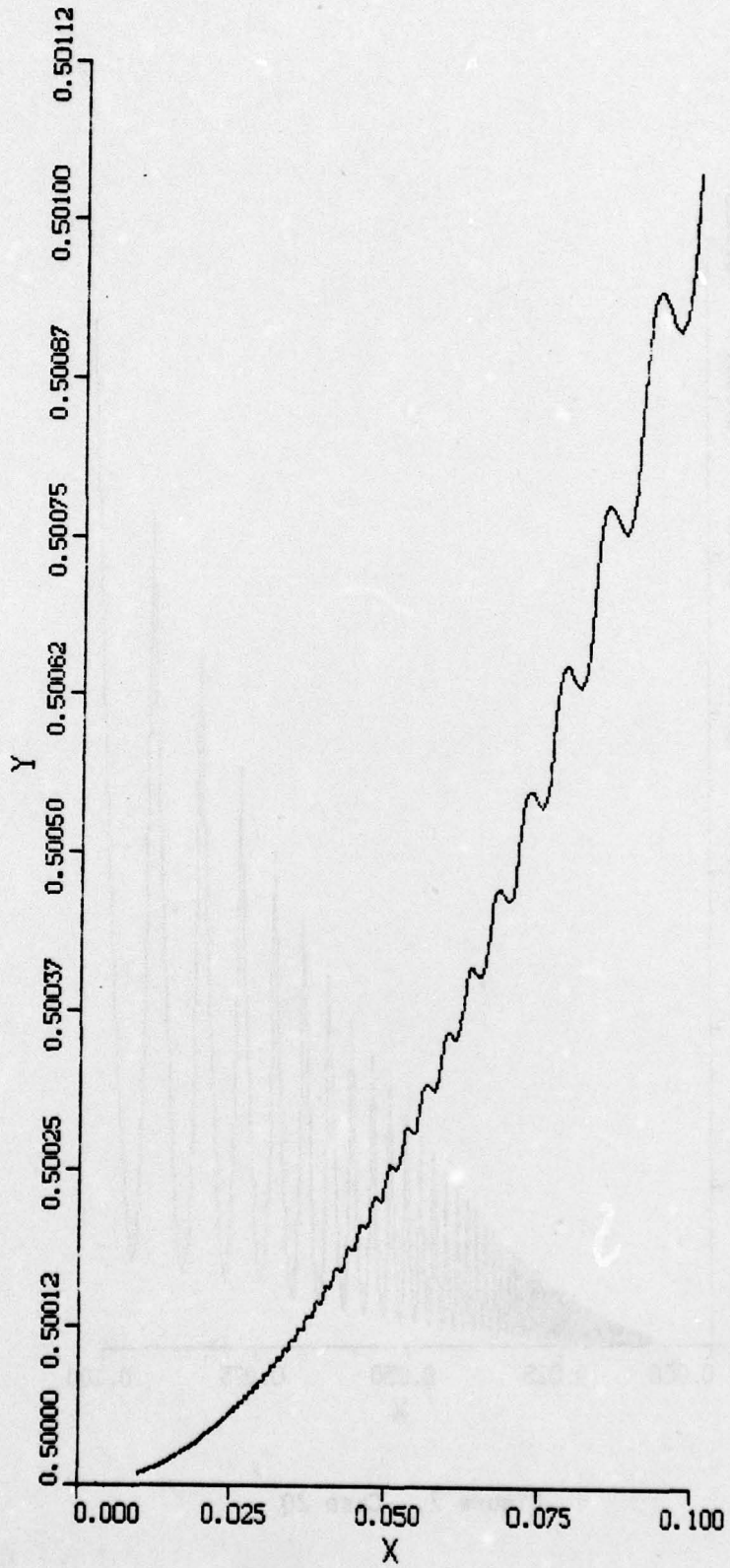


Figure 8. Case 2C

Figure 9. Case 3L

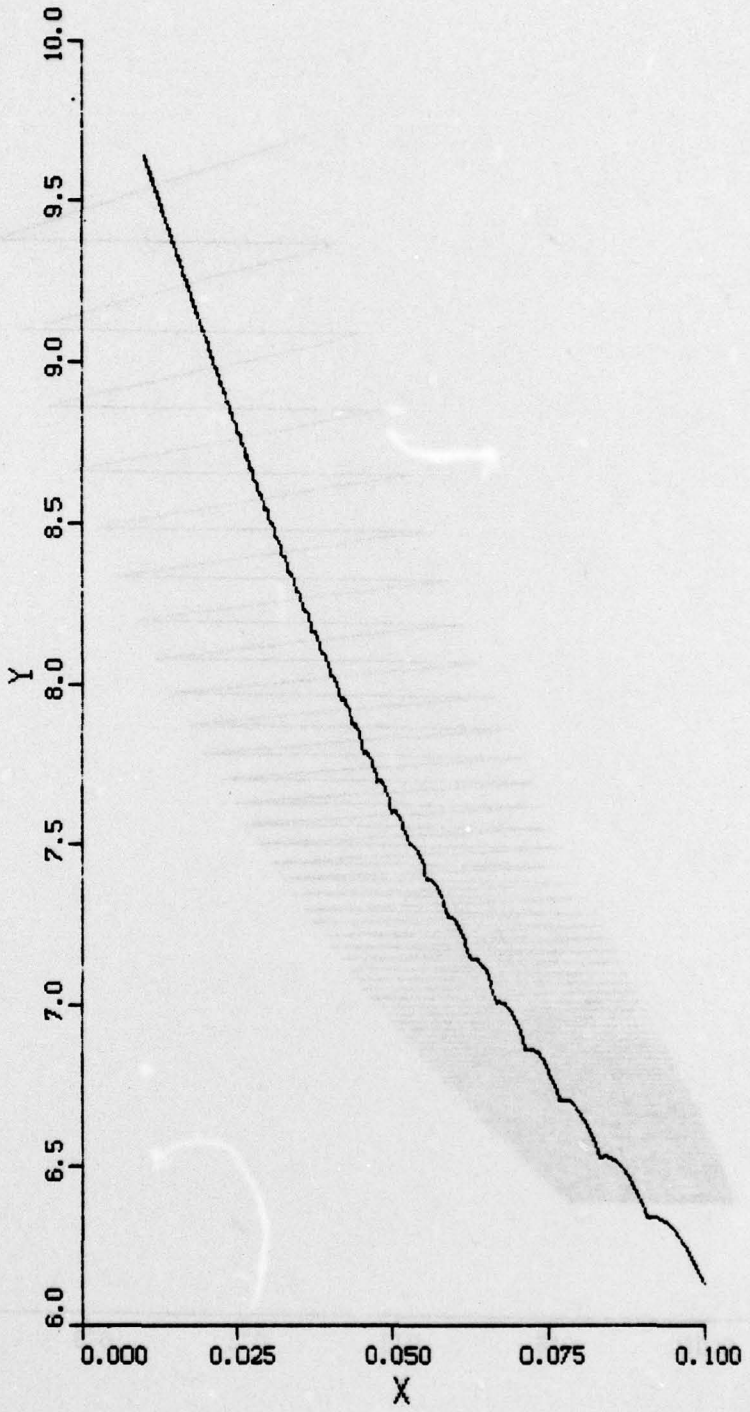


Figure 10. Case 3Q

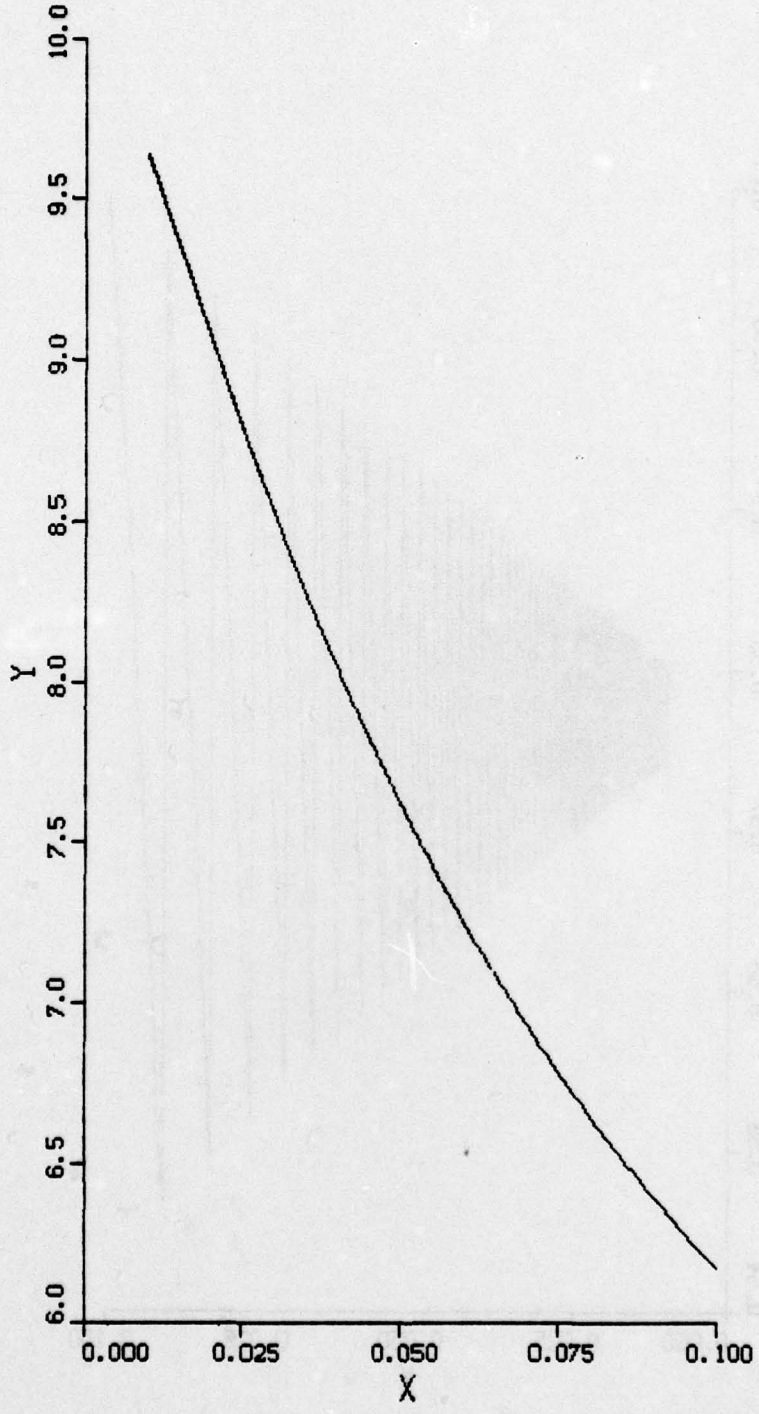


Figure 11. Case 3C

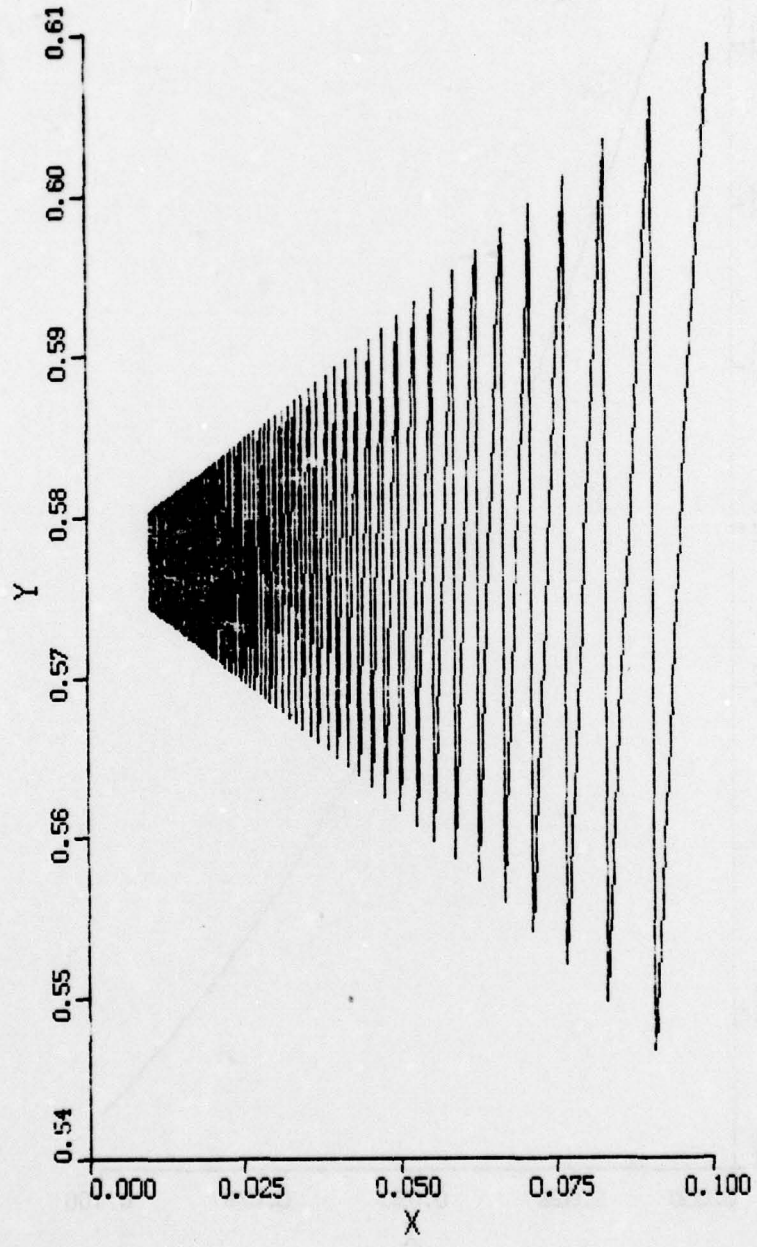


Figure 12. Case 4L

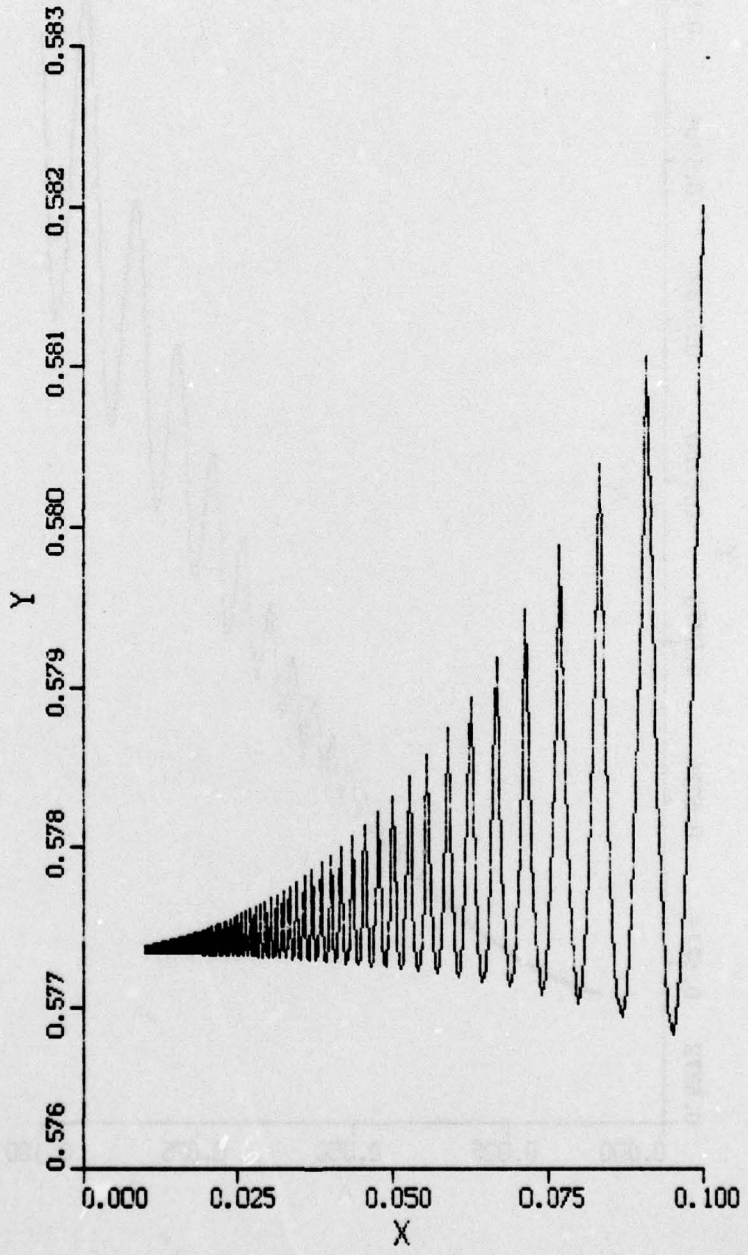


Figure 13. Case 4Q

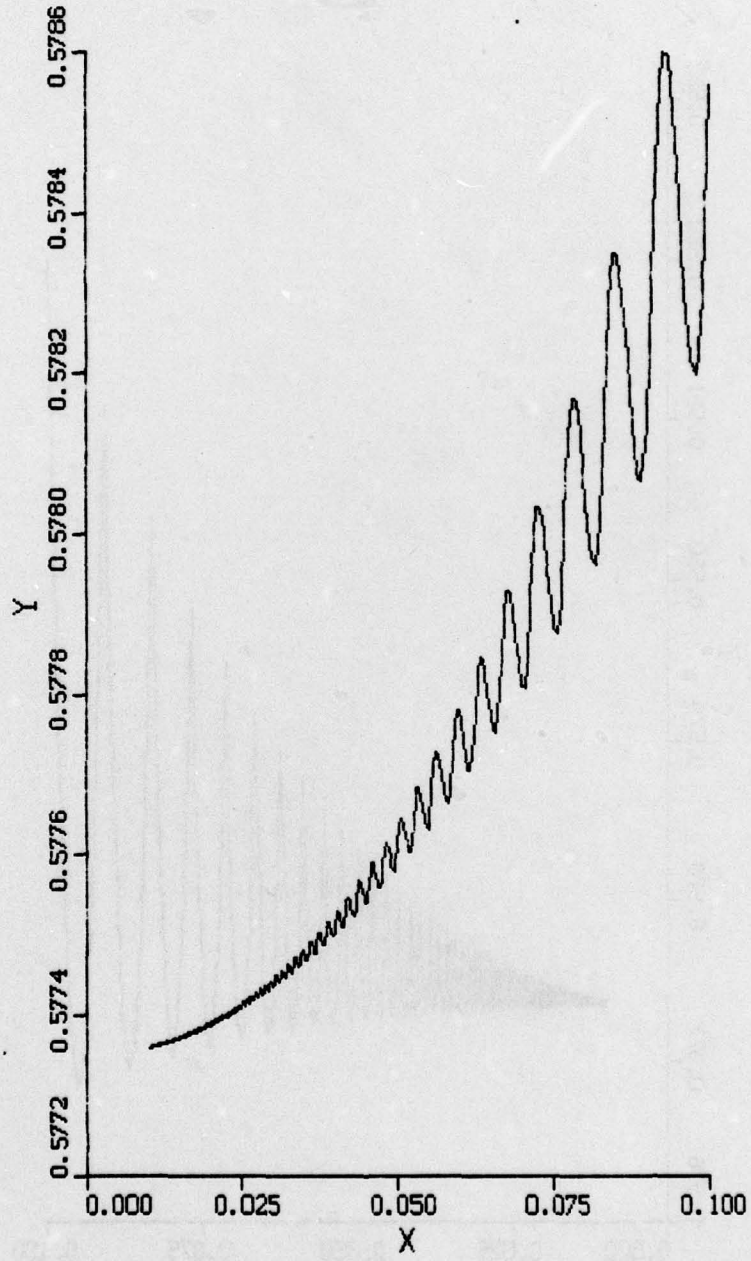


Figure 14, Case 4C