JPL PUBLICATION 79-50

Nonlinear Ordinary Difference Equations

Thomas K. Caughey
California Institute of Technology

(NASA-CR-162106) NONLINEAR ORDINARY DIFFERENCE EQUATIONS (Jet Propulsion Lab.) 118 p HC A06/MF A01 CSCL 12A
N79-30986
Unclas
G3/64 31790

June 1, 1979

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

JPL PUBLICATION 79-50

Nonlinear Ordinary Difference Equations

Thomas K. Caughey
California Institute of Technology

June 1, 1979

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

TABLE OF CONTENTS

PRECEDING PAGE BLANK NOT FILMED

ACKNOWLEDGMENTS

The research in this publication was carried out under the auspices of the Applied Mechanics Division of the Jet Propulsion Laboratory, California Institute of Technology under NASA Contract No. NAS 7-100. The effort was supported by Dr. A. Amos, Materials and Structures Division, Office of Aeronautics and Space Technology.

Abstract

This note is part of a continuing study of future problem areas in structural dynamics of space vehicles, conducted by the author for the Jet Propulsion Laboratory.

The motivation for this particular piece of work is the conviction that future space vehicles will be relatively large and flexible, and that active control will be necessary to maintain geometrical configuration. While the stresses and strains in these new space vehicles are not expected to be excessively large, their cumulative effects will cause significant geometrical nonlinearities to appear in the equations of motion, in addition to the nonlinearities caused by material properties. Since the only effective tool for the analysis of such large complex structures is the digital computer, it will be necessary to gain a better understanding of the nonlinear ordinary difference equations which result from the time discretization of the semi-discrete equations of motion for such structures.

1. Introduction

Equations of the type:

$$\left.\begin{aligned} \underline{x}_{n+1} &= \underline{f}\,(\underline{x}_n, n) \\ \underline{x}_o &= \underline{c} \quad . \end{aligned}\right\} \tag{1.1}$$

$$\left.\begin{aligned} \underline{x}_{n+1} &= \underline{f}(\underline{x}_n, \underline{x}_{n+1}, n) \\ \underline{x}_o &= \underline{c} \end{aligned}\right\} \tag{1.2}$$

are known as nonlinear ordinary difference equations or point mappings. Equation (1.1) is known as an explicit nonlinear difference equation, while Eq. (1.2) is known as an implicit nonlinear difference equation.

If in Eq. (1.1)

$$\underline{f}\,(\underline{x}_n, \underline{n}) = A(n)\underline{x}_n + \underline{g}(n) \tag{1.3}$$

then (1.1) becomes

$$\left.\begin{aligned} \underline{x}_{n+1} &= A(n)\underline{x}_n + \underline{g}(n) \\ x_o &= \underline{c} \end{aligned}\right\} \tag{1.4}$$

Similarly, if in (1.2)

$$\underline{f}(\underline{x}_n, \underline{x}_{n+1}, n) = A(n)\underline{x}_n + B(n)\underline{x}_{n+1} + \underline{g}(n) \quad . \tag{1.5}$$

Then (1.2) becomes

$$\left.\begin{aligned} \underline{x}_{n+1} &= A(n)\underline{x}_n + B(n)\underline{x}_{n+1} + g(n) \\ \underline{x} &= \underline{c} \end{aligned}\right\} \tag{1.6}$$

Equation (1.4) is known as a linear explicit difference equation, while (1.6) is known as an implicit linear difference equation.

Since Eq. (1.6) can be rewritten as:

$$\left.\begin{aligned} \underline{x}_{n+1} &= C(n)\underline{x}_n + \underline{h}(n) \\ \underline{x}_0 &= \underline{c} \\ C(n) &= [I-B(n)]^{-1}A(n) \\ \underline{h}(n) &= [I-B(n)]^{-1}\underline{g}(n) \end{aligned}\right] \quad . \tag{1.7}$$

Thus, there is no difference, in theory, between explicit linear difference equations and implicit linear difference equations. Unfortunately the same is not true, in general, for nonlinear difference equations.

Difference equations arise in a variety of scientific and engineering disciplines, for example:

(a) In biology; population genetics and dynamics are described by nonlinear difference equations.

(b) In control theory; sampled data control system are described by either linear or nonlinear difference equations.

(c) In numerical analysis; in order to solve a differential equation on a digital computer, the independent variable must be discretized and the differential equation becomes a difference equation. In particular, nonlinear differential equations become nonlinear difference equations.

It is to this last class of problem that this note is addressed.

2. Existence and Uniqueness of a Solution of the Initial Value Problem

a) Explicit Nonlinear Difference Equations

Theorem 1 Given the explicit nonlinear difference equation

$$\left.\begin{aligned} \underline{x}_{n+1} &= \underline{f}(\underline{x}_n) \\ \underline{x}_o &= \underline{c} \end{aligned}\right\} \qquad (2.1)$$

If (i) $\forall \underline{x}$, $\underline{f}(\underline{x})$ is continuous in $\underline{x}$, therefore $\|\underline{f}(\underline{x})\| < \infty$, $\forall \|\underline{x}\| < \infty$

(ii) $\|\underline{c}\| < \infty$.

Then there exists a unique solution of the initial value problem (2.1)

Proof

Since $\|\underline{c}\| < \infty$

$$\therefore \quad \|\underline{x}_1\| \le \|\underline{f}(\underline{c})\| < \infty$$

$$\|\underline{x}_2\| \le \|\underline{f}(\underline{x}_1)\| < \infty$$

$$\vdots \qquad \vdots$$

$$\|\underline{x}_n\| \le \|\underline{f}(\underline{x}_{n-1})\| < \infty \quad .$$

Therefore there exists a solution to Eq. (2.1), satisfying the initial data. Since the process of generating the solution is explicit, then there exists one and only one solution of (2.1) satisfying the initial data, therefore the solution of (2.1) is unique. It will be noted that for explicit nonlinear difference equations, the question of existence and uniqueness of a solution is trivially answered in comparison with the same question for nonlinear differential equations.

b) Implicit Nonlinear Difference Equations

Let us consider now the implicit nonlinear difference equation

$$\left.\begin{aligned} \underline{x}_{n+1} &= \underline{f}(\underline{x}_n, \underline{x}_{n+1}) \\ \underline{x}_o &= \underline{c} \end{aligned}\right\}. \qquad (2.2)$$

In the general case, we can say relatively little about the existence of a solution to Eq. (2.2). The implicit function theorem guarantees, under weak restrictions on $\underline{f}(\underline{x}_n, \underline{x}_{n+1})$, that there exists a unique local solution of (2.2) provided $||\underline{c}||$ is sufficiently small. In some special cases Eq. (2.2) may be inverted so that it is described by an explicit equation.

$$\left.\begin{aligned} &\underline{x}_{n+1} = \underline{F}(\underline{x}_n) \\ &\underline{x}_0 = \underline{c} \end{aligned}\right\} . \tag{2.3}$$

In the case of most practical importance, Eq. (2.2) has the structure

$$\left.\begin{aligned} &\underline{x}_{n+1} = \underline{x}_n + \epsilon \underline{f}_1(\underline{x}_n, \underline{x}_{n+1}) \\ &\underline{x}_0 = \underline{c} \end{aligned}\right\} \tag{2.4}$$

where $|\epsilon|$ is frequently a small quantity.

Before proving the existence of a unique solution of Eq. (2.4) we will establish the following theorem.

Theorem 2 Given the implicit equation

$$\underline{x} = \underline{g}(\underline{x}) \tag{2.5}$$

and the iterative procedure

$$\underline{x}_{n+1} = \underline{g}(\underline{x}_n) \qquad n = 0, 1, 2 \ . \ . \ . \quad . \tag{2.6}$$

Then if $\underline{g}(x)$ satisfies the following conditions

(i) $||\underline{g}(x) - \underline{g}(y)|| \le \lambda \, ||\underline{x} - \underline{y}||$ for $\forall \ \underline{x}, \underline{y} \in S \quad S: ||z - \underline{x}^o|| \le \rho$ with $0 \le \lambda \le 1$ (2.7)

(ii) There exists an $\underline{x}_o \ni ||\underline{g}(\underline{x}^o) - x^o|| \le (1-\lambda)\rho$ (2.8)

then $\forall$ iterates $\underline{x}_n$ satisfy the following conditions

$$\left.\begin{array}{l}\text{(i)}\quad \|\underline{x}_n-\underline{x}_o\| \le \rho \\ \text{(ii)}\quad \lim_{n\to\infty} \|\underline{x}_n\| = \underline{\alpha} \qquad \text{where} \quad \underline{\alpha} = \underline{g}(\underline{\alpha}) \\ \text{(iii)}\quad \underline{\alpha} \text{ is the only root of Eq. (2.5) in } \|\underline{x}-\underline{x}_o\| \le \rho\end{array}\right\} \tag{2.9}$$

Proof

Since $\underline{x}_{n+1} = \underline{g}(\underline{x}_n)$ (2.10)

$$\underline{x}_n = \underline{g}(\underline{x}_{n-1}) \tag{2.11}$$

$$\therefore \quad \|\underline{x}_{n+1} - \underline{x}_n\| = \|g(\underline{x}_n)-\underline{g}(x_{n-1})\| \tag{2.12}$$

$$\le \lambda \|\underline{x}_n - \underline{x}_{n-1}\| \tag{2.13}$$

if $\underline{x}_{n+1}, \underline{x}_n \in S$

Now $\|\underline{x}_1 - \underline{x}_o\| \le (1-\lambda)\rho < \rho$ (2.14)

$\therefore \quad \underline{x}_1 \in S$

$$\|\underline{x}_2-\underline{x}_1\| \le \lambda \|\underline{x}_1 - \underline{x}_o\| \le \lambda(1-\lambda)\rho \tag{2.15}$$

$$\therefore \quad \|\underline{x}_2 - \underline{x}_o\| \le \|\underline{x}_2 - \underline{x}_1\| + \|\underline{x}_1 - \underline{x}_o\|$$

$$\le [\lambda(1-\lambda) + (1-\lambda)]\rho = (1-\lambda^2)\rho < \rho \tag{2.16}$$

$\therefore \quad \underline{x}_2 \in S$

Suppose that $\underline{x}_o, \underline{x}_1 \; . \; . \; . \; \underline{x}_n \in S$

then $\|\underline{x}_{n+1} - \underline{x}_n\| \le \lambda \|\underline{x}_n - \underline{x}_{n-1}\|$ (2.17)

$$\le \lambda^n \|\underline{x}_1 - \underline{x}_o\| \tag{2.18}$$

$$\therefore \quad \|\underline{x}_{n+1} - \underline{x}_o\| \le \|\underline{x}_{n+1} - \underline{x}_n\| + \|\underline{x}_n - \underline{x}_{n-1}\| + \; . \; . \; . \; \|x_1 - \underline{x}_o\| \tag{2.19}$$

$$\le (\lambda^n + \lambda^{n-1} + \; . \; . \; . \; \lambda + 1)(1-\lambda)\rho \tag{2.20}$$

$$\le (1-\lambda^{n+1})\rho < \rho \tag{2.21}$$

$\therefore \quad \underline{x}_{n+1} \in S$

$$\lim_{n\to\infty} \|\underline{x}_{n+1} - \underline{x}_n\| \le \lim_{n\to\infty} \lambda^n \rho \equiv 0 \tag{2.22}$$

$$\|\underline{x}_{n+k} - \underline{x}_n\| \le \|\underline{x}_{n+h} - \underline{x}_{n+h-1}\| + \; . \; . \; . \; . \; + \|\underline{x}_{n+1} - \underline{x}_n\| \tag{2.23}$$

$$\le (\lambda^{n+h-1} + \lambda^{n+h-2} + \; . \; . \; . \; \lambda^n)(1-\lambda)\rho \tag{2.24}$$

$$\le \lambda^n \rho \tag{2.25}$$

$$\therefore \lim_{n\to\infty} \|\underline{x}_{n+k} - \underline{x}_n\| = 0 \tag{2.26}$$

$\therefore$ the sequence $\{\underline{x}_n\}$ is a Cauchy sequence and converges uniformly.

$$\therefore \lim_{n\to\infty} \underline{x}_{n+1} = \underline{\alpha} = \lim_{n\to\infty} \underline{g}(\underline{x}_n) = \underline{g}(\lim_{n\to\infty} x_n) = g(\underline{\alpha}) \tag{2.27}$$

$\therefore$ the sequence $\{\underline{x}_n\}$ converges uniformly to a limit, $\underline{\alpha} \in S$, which is a solution of Eq. (2.5).

Unicity

If $\underline{\alpha}$ and $\underline{\beta}$ are solutions of Eq. (2.5) which both belong to the set S. Then

$$\underline{\alpha} = \underline{g}(\underline{\alpha}) \qquad \underline{\alpha} \in S \tag{2.28}$$

$$\underline{\beta} = \underline{g}(\underline{\beta}) \qquad \underline{\beta} \in S \tag{2.29}$$

$$\therefore \|\underline{\alpha} - \underline{\beta}\| = \|\underline{g}(\underline{\alpha}) - \underline{g}(\underline{\beta})\| \le \lambda \|\underline{\alpha} - \underline{\beta}\| \tag{2.30}$$

$$\therefore \|\underline{\alpha} - \underline{\beta}\|(1-\lambda) \le 0 \qquad \therefore \underline{\alpha} \equiv \underline{\beta} \; . \tag{2.31}$$

Thus, under the hypothesis of Theorem 2 there exists a unique solution of Eq. (2.5).

Returning to the question of the existence and uniqueness of a solution of the initial value problem for an implicit nonlinear difference equation, we have Theorem 3.

<u>Theorem 3</u> Given the implicit nonlinear difference equation

$$\left.\begin{aligned} \underline{x}_{n+1} &= \underline{x}_n + \epsilon \underline{f}(\underline{x}_n, \underline{x}_{n+1}) \\ \underline{x}_o &= \underline{c} \end{aligned}\right\} \qquad (2.32)$$

If

(i) $\underline{f}(\underline{x}_n, \underline{x}_{n+1})$ is continuous in $\underline{x}_n$ and $\underline{x}_{n+1}$.

(ii) $\underline{f}(\underline{x}_n, \underline{x}_{n+1})$ has continous first partial derivatives with respect to $\underline{x}_{n+1}$.

(iii) $|\epsilon|$ is sufficiently small.

Then there exists a unique solution to Eq. (2.32) on some finite interval $0 \leq n < N$.

<u>Proof</u> Let $\underline{x} = \underline{x}_{n+1} = \underline{x}_n + \epsilon \underline{f}(\underline{x}_n, \underline{x}_{n+1}) = \underline{g}(\underline{x})$ (2.33)

With hypothesis (i) and (ii) $\underline{g}(\underline{x})$ satisfies

$$\left.\begin{aligned} \|\underline{g}(\underline{x}) \cdot \underline{g}(\underline{y})\| &= |\epsilon| \, \|\underline{f}(\underline{x}_n, \underline{x}) - \underline{f}(\underline{x}_n, \underline{y})\| \\ &\leq |\epsilon| \, \|J(\xi)\| \cdot \|\underline{x} - \underline{y}\| \\ &\leq \lambda \, \|\underline{x} - \underline{y}\| \end{aligned}\right\} \qquad (2.34)$$

where

$$\left.\begin{aligned} & J(\underline{\xi}) = \underline{f},_{\underline{x}}(\underline{x}_n, \underline{x}) \big|_{\underline{x} = \underline{\xi}} \\ & \underline{\xi} = \alpha \underline{x} + (1-\alpha)\underline{y} \qquad 0 < \alpha < 1 \end{aligned}\right\} \qquad (2.35)$$

For

$$\underline{x}, \underline{y} \ni \|\underline{x} - \underline{x}_n\| < \rho \quad , \quad \|\underline{y} - \underline{x}_n\| < \rho \qquad (2.36)$$

We can always choose $|\epsilon|$ sufficiently small so that

$$\lambda = |\epsilon| \, \| J(\underline{\xi}) \| < 1 \tag{2.37}$$

If we choose $\underline{x}^o = \underline{x}_n$ as our initial iterate, then

$$\| \underline{g}(\underline{x}^o) - \underline{x}^o \| = \| \epsilon \underline{f}(\underline{x}_n, \underline{x}_n) \| \tag{2.38}$$

Since $\underline{f}(\underline{x}, \underline{y})$ is continuous in $\underline{x}$ and $\underline{y}$ we can always choose $|\epsilon|$ sufficiently small so that

$$\| \underline{g}(\underline{x}^o) - \underline{x}^o \| \leq (1-\lambda)\rho \tag{2.39}$$

Thus, given the hypothesis (i) and (ii) we can always choose $|\epsilon|$ sufficiently small so that the conditions of Theorem 2 are satisfied. Thus given an $\underline{x}_n$, there exists a unique solution $\underline{x}_{n+1}$ satisfying

$$\underline{x}_{n+1} = \underline{x}_n + \epsilon \underline{f}(\underline{x}_n, \underline{x}_{n+1}) \tag{2.40}$$

Thus, starting with $\underline{x}_o = \underline{c}$ and $|\epsilon|$ fixed and sufficiently small, there exists a unique $\underline{x}_1$, satisfying Eq. (2.32). If, using $\underline{x}_1$ and the same value of ϵ, conditions (2.37) and (2.39) are satisfied, then there exists a unique $\underline{x}_2$ satisfying Eq. (2.32). Proceeding in this way, we check at each step to see if conditions (2.37) and (2.39) are satisfied. If they are satisfied at each step, the solution can be continued indefinitely into the future. If they are not satisfied after a finite number of steps the solution may cease to exist or go to infinity. Thus, given condition (i), (ii) and (iii) there exists a unique solution to Eq. (2.32), at least on some finite interval $0 < n < N$.

Theorems 1 and 3 deal with autonomous equations that is, equations which do not contain n explicitly. The hypothesis of Theorems 1 and 2 can be relaxed to include explicit dependence on n, in addition to domain dependent continuity properties.

3. Properties of Linear Difference Equations

(a) Difference Equations with Constant Coefficients

Consider the linear difference equations

$$\left.\begin{aligned} \underline{x}_{n+1} &= A\underline{x}_n + \underline{f}_n \\ \underline{x}_o &= \underline{c} \qquad |A| \neq 0 \end{aligned}\right\} \tag{3.1}$$

where A is a constant matrix with $\|A\| = a < \infty$, $\|\underline{f}(n)\| < \infty$. The solution of Eq. (3.1) is easily formed by elementary methods

$$\underline{x}_1 = A\underline{c} + \underline{f}_o \tag{3.2}$$

$$\underline{x}_2 = A\underline{x}_1 + \underline{f}_1 = A^2\underline{c} + A\underline{f}_o + \underline{f}_1 \tag{3.3}$$

$$\underline{x}_3 = A\underline{x}_2 + \underline{f}_2 = A^3\underline{c} + \sum_{i=o}^{2} A^{2-i}\underline{f}_i \tag{3.4}$$

$$\therefore \underline{x}_{n+1} = A^{n+1}\underline{c} + \sum_{i=o}^{n} A^{n-i}\underline{f}_i \tag{3.5}$$

Alternatively we can write this solution in terms of the principal matrix solution X_n, where

$$X_{n+1} = AX_n \quad ; \quad X_o = I \tag{3.6}$$

thus

$$X_1 = A \tag{3.7}$$

$$X_2 = A^2 \tag{3.8}$$

.

.

.

$$X_n = A^n \tag{3.9}$$

Thus, the solution to Eq. (3.1) can be written

$$\underline{x}_{n+1} = X_{n+1}\underline{c} + \sum_{i=o}^{n} X_{n-i}\,\underline{f}_i \qquad (3.10)$$

We note that $\quad X_{n-i} = A^{n-i} = A^{n+1}A^{-(i+1)} \qquad (3.11)$

$$\therefore\ \underline{x}_{n+1} = X_{n+1}\underline{c} + \sum_{i=o}^{n} X_{n+1}X_{i+1}^{-1}\,\underline{f}_i \qquad (3.12)$$

(b) <u>Difference Equations with Variable Coefficients</u>

Consider the linear difference equation

$$\left.\begin{aligned} \underline{x}_{n+1} &= A(n)\underline{x}_n + \underline{f}_n \\ \underline{x}_o &= \underline{c} \qquad |A| \neq 0 \end{aligned}\right\} \qquad (3.13)$$

where $A(n)$ is a step dependent matrix, with $\|A(n)\| < \infty,\ \forall n$ and $\|\underline{f}(n)\| < \infty,\ \forall n$.

The solution of Eq. (3.13) is also easily formed by elementary methods.

$$\underline{x}_1 = A(0)\underline{c} + \underline{f}_o \qquad (3.14)$$

$$\underline{x}_2 = A(1)\underline{x}_1 + \underline{f}_1 = A(1)A(0)\underline{c} + A(1)f_o + \underline{f}_1 \qquad (3.15)$$

$$\underline{x}_3 = A(2)\underline{x}_2 + \underline{f}_2 = A(2)A(1)A(0)\underline{c} + A(2)A(1)f_o + A(2)\underline{f}_1 + \underline{f}_2 \qquad (3.16)$$

$$\underline{x}_{n+1} = A(n)A(n-1)\ .\ .\ .\ A(0)\left[\underline{c} + \sum_{i=o}^{n}\prod_{h=o}^{i} A(h)^{-1}f_i\right] \qquad (3.17)$$

Alternatively we can write this solution in terms of the principal matrix solution X_n, where

$$X_{n+1} = A(n)X_n \quad , \quad X_o = I \qquad (3.18)$$

Thus

$$X_1 = A(0) \tag{3.19}$$

$$X_2 = A(1)A(0) \tag{3.20}$$

.

.

.

.

$$X_n = A(n-1)A(n-2)\ .\ .\ .A(0) = \prod_{i=o}^{n-1} A(i) \tag{3.21}$$

Thus, the solution to Eq. (3.13) can be written

$$\underline{x}_{n+1} = X_{n+1}\underline{c} + \sum_{i=o}^{n} X_{n+1} X_{i+1}^{-1} \underline{f}_i \tag{3.22}$$

In general we can say very little about the structure of the solution in the case of variable coefficients. There is, however, one special case, the case of a difference equation with periodic coefficients.

(c) Difference Equation with Periodic Coefficients

Theorem 4 Consider the homogeneous difference equation

$$\left.\begin{aligned} X_{n+1} &= A(n)\,X_n \\ X_o &= I \end{aligned}\right\} \tag{3.23}$$

where

$$\left.\begin{aligned} &A(n+N) = A(n) \qquad |A(n)| \neq 0 \qquad \forall n \\ &\|A(n)\| < \alpha \qquad \forall n \end{aligned}\right\} \tag{3.24}$$

The principal matrix solution X_n has the form

$$X_n = Q(n)C^n \tag{3.25}$$

where $Q(n+N) = Q(n)$ is a periodic matrix and C is a non-singular constant matrix.

Proof From (3.23)

$$X_o = I \tag{3.26}$$

$$X_1 = A(0) \tag{3.27}$$

$$X_2 = A(1)A(0) \tag{3.28}$$

.

.

.

$$X_n = \prod_{i=o}^{N-1} A(i) \tag{3.29}$$

We note that $X_k = \prod_{i=o}^{k-1} A(i)$ is non-singular since $A(i)$ is non-singular $\forall i$.

$$X_{N+1} = \prod_{i=o}^{N} A(i) = A(N)\prod_{i=o}^{N-1} A(i) = X_1 X_N \tag{3.30}$$

$$X_{N+2} = \prod_{i=o}^{N+1} A(i) = \prod_{i=o}^{1} A(i)X_N = X_2 X_N \tag{3.31}$$

$$X_{N+k} = \prod_{i=o}^{N+k-1} A(i) = \prod_{i=o}^{k-1} A(i)X_N = X_k X_N \tag{3.32}$$

Similarly

$$X_{\ell N+k} = X_k X^{\ell}(N) \tag{3.33}$$

Since X_N is non-singular, we may write

$$X_N = C^N \quad , \quad C \text{ - a constant matrix} \tag{3.34}$$

Consider the matrix

$$Q(n) = X_n C^{-n} \tag{3.35}$$

Since

$$X_o = I \quad , \quad Q(0) = I \tag{3.36}$$

Thus

$$Q^{(N+n)} = X_{N+n} C^{-(N+n)} \tag{3.37}$$

$$= X_n X_N C^{-N} C^{-n} \tag{3.38}$$

But

$$X_N C^{-N} = I$$

$$\therefore Q(N+n) = X_n C^{-n} = Q(n) \tag{3.39}$$

$\therefore Q(n)$ is a periodic matrix with period N.

Hence

$$X_n = Q(n) C^n \tag{3.40}$$

Thus the complete structure of X_n is known for $\forall n$ if X_h is known for $0 \leq h \leq N$.

We note in passing that if A is a constant matrix, then A is a periodic matrix of period $N = 1$, hence, difference equations with constant coefficients are a special case of difference equations with periodic coefficients and in this case the matrix $C = A$ and Eq. (3.40) becomes

$$X_n = A^n \tag{3.41}$$

4. Stability of Difference Equations

Definition Liapunov Stability (L.S.)

Given the difference equation

$$\underline{x}_{n+1} = \underline{f}(\underline{x}_n, \underline{x}_{n+1}) \tag{4.1}$$

where

$$\underline{f}(\underline{0}, \underline{0}) = \underline{0} \tag{4.2}$$

The equilibrium solution $\underline{x} = 0$ is said to be Liapunov stable if given any $\delta > 0$, there exists an $\epsilon > 0$, such that if $\|\underline{x}_o\| < \epsilon$, then $\|\underline{x}_n\| < \delta$ for all $n > 0$.

Liapunov Asymptotic Stability (L.A.S.)

The equilibrium solution $\underline{x} = 0$ is said to be Liapunov asymptotically stable if (a) it is Liapunov stable and (b) $\|\underline{x}_n\| \to 0$ as $n \to \infty$.

(a) Stability of Linear Difference Equations

(i) Linear Difference Equations with Constant Coefficients

Theorem 5 Given the difference equation

$$\underline{x}_{n+1} = A\underline{x}_n \tag{4.3}$$

A - a constant matrix

(i) If A is non-defective (i.e.has a full complement of ordinary eigenvector) necessary and sufficient conditions for Liapunov stability are that the eigenvalues of A should be less than or equal to unity in modulus.

(ii) If A is defective (i.e. does not have a full complement of ordinary eigenvectors) necessary and sufficient conditions for Liapunov stability are that the eigenvalues of A should be less than unity in modulus.

Proof

(i) If A is simple, i.e. non-defective there exists a similarity matrix

$$T \ni T^{-1}AT = \Lambda$$

where Λ is a diagonal matrix.

A has the representation

$$A = T\Lambda T^{-1}$$

As previously shown, the principal matrix solution of (4.3) is:

$$X_n = A^n = (T\Lambda T^{-1})^n = T\Lambda^n = T\Lambda^n T^{-1} \tag{4.4}$$

Sufficiency

If $|\lambda_i(A)| \leq 1$, then $\lambda_i^n(A)$ remains bounded as $n \to \infty$

$\therefore X_n$ is bounded and remains bounded as $n \to \infty$

$$\therefore \|X_n\| \leq M < \infty \quad , \quad \forall n \tag{4.5}$$

From (4.3)

$$\underline{x}_n = X_n \underline{x}_o \tag{4.6}$$

$$\therefore \text{if} \qquad \|\underline{x}_o\| < \epsilon \tag{4.7}$$

$$\therefore \qquad \|\underline{x}_n\| \leq \|X_n\| \epsilon \tag{4.8}$$

$$\leq M\epsilon \tag{4.9}$$

$\therefore$ if $\quad \epsilon \leq \delta / M$

$$\|\underline{x}_n\| \leq \delta \quad \forall n \tag{4.10}$$

$\therefore$ (4.3) is Liapunov stable at $\underline{x} = 0$.

Necessity

If $|\lambda_i(A)| > 1$ for some i, then $\lambda_i^n(A)$ cannot remain bounded as

$n \to \infty$. Hence X_n cannot remain bounded as $n \to \infty$.

(ii) If A is defective, it cannot be reduced to diagonal form, however, there is a similarity matrix $T \ni$

$$T^{-1}AT = \begin{bmatrix} J_{\alpha_1} & & & & & \\ & J_{\alpha_2} & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & J_{\alpha_h} \end{bmatrix} = J \tag{4.11}$$

where the J_{α_i} are Jordan blocks associated with the eigenvalues $\lambda_i(A)$ $i \in (1,h)$ A has the representation

$$A = TJT^{-1} \tag{4.12}$$

As previously shown, the principal matrix solution of (4.3) is:

$$X_n = A^n = (TJT^{-1})^n = TJ^nT^{-1} \tag{4.13}$$

where

$$J^n = \begin{bmatrix} J^n_{\alpha_1} & & & & & \\ & J^n_{\alpha_2} & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & J^n_{\alpha_h} \end{bmatrix} \tag{4.14}$$

and

and

$$J_{\alpha i} = \begin{bmatrix} \lambda_i^n & n\lambda_i^{n-1} & \frac{n(n-1)}{\lfloor 2} & \lambda_i^{n-2} & \cdots \\ & \lambda_i^n & n\lambda_i^{n-1} & & \\ 0 & & \lambda_i^n & & \\ & & & \ddots & \\ & & & & \lambda_i^n \end{bmatrix} \tag{4.15}$$

<u>Sufficiency</u> If $|\lambda_i(A)| < 1$, then $J^n_{\alpha_i}$ remains bounded and tends to zero as $n \to \infty$. $\therefore$ X_n is bounded and tends to zero as $n \to \infty$.

$$\therefore \quad \|X_n\| \le M < \infty \ \forall n \tag{4.16}$$

$$\text{and} \quad \lim_{n\to\infty} \|X_n\| \to 0$$

From which we immediately deduce that if $|\lambda_i(A)| < 1 \ \forall i$, the system (4.3) is not only Liapunov stable, but is asymptotically stable.

<u>Necessity</u> If $|\lambda_i(A)| \ge 1$ for some i, then $J^n_{\alpha_i}$ cannot remain bounded as $n \to \infty$, hence X_n cannot remain bounded as $n \to \infty$.

Alternatively use can be made of Liapunov's Theorem.

<u>Theorem 6</u> (Liapunov) Given the difference equation

$$\underline{x}_{n+1} = A\underline{x}_n \tag{4.17}$$

A - a constant matrix

Then (4.17) is Liapunov asymptotically stable at $\underline{x} = 0$ iff there exists a symmetric positive definite matrix P such that

$$A^T PA - P = -Q \tag{4.18}$$

Proof Sufficiency

Suppose that there exists a matrix P satisfy (4.18) let

$$V_n = \underline{x}_n^T P \underline{x}_n \tag{4.19}$$

Since P is symmetric and positive definite V_n is positive definite

$$V_{n+1} = \underline{x}_{n+1}^T P \underline{x}_n \tag{4.20}$$

Using (4.17)

$$V_{n+1} = (A\underline{x}_n)^T P(A\underline{x}_n) \tag{4.21}$$

$$= \underline{x}_n^T A^T PA\underline{x}_n \tag{4.22}$$

$$\therefore \quad \Delta V_n = V_{n+1} - V_n = \underline{x}_n^T (A^T PA - P)\underline{x}_n \tag{4.23}$$

Using (4.13)

$$\Delta V_n = -\underline{x}_n^T Q \underline{x}_n < 0 \tag{4.24}$$

Thus

$$V_{n+1} < V_n < V_{n-1} < V_{n-2} \ldots < V_o \tag{4.25}$$

Since V_n vanishes only at the origin

$$\therefore \quad V_n \to 0 \quad \text{as} \quad n \to \infty \tag{4.26}$$

If we define

$$\|\underline{x}_n\|_p = \sqrt{V_n} \tag{4.27}$$

we see that

$$\|\underline{x}_n\|_p < \|\underline{x}_o\|_p \tag{4.28}$$

∴ If

$$\|\underline{x}_o\| \le \epsilon = \delta \tag{4.29}$$

then

$$\|\underline{x}_n\|_p < \delta \quad \forall n \tag{4.30}$$

and

$$\|\underline{x}_n\|_p \to 0 \quad \text{as} \quad n \to \infty \tag{4.31}$$

∴ Eq. (4.17) is Liapunov asymptotically stable at the origin.

Necessity

Let A be a stability matrix i.e. $|\lambda_i(A)| < 1 \ \forall i$. Let P satisfy (4.18) i.e.

$$A^T PA - P = -Q \tag{4.32}$$

We wish to show that P is symmetric and positive definite.

If we premultiply (4.20) by A^T and post multiply by A, then by A^{2T} and A^2, etc., we obtain

$$\left.\begin{aligned}
&A^T PA - P = -Q \\
&A^{2T} PA^2 - A^T PA = -A^T QA \\
&A^{3T} PA^3 - A^{2T} PA^2 = -A^{2T} QA^2 \\
&\cdots\cdots\cdots\cdots \\
&\cdots\cdots\cdots\cdots \\
&A^{n^T} PA^n - A^{n-1^T} PA^{n-1} = -A^{n-1^T} QA^{n-1}
\end{aligned}\right\} \tag{4.33}$$

Adding, we obtain

$$(A^n)^T P A^n - P = -\sum_{i=o}^{n-1} (A^i)^T Q A^i \tag{4.34}$$

Since A is a stability matrix $A^n \to [0]$ as $n \to \infty$

$$\therefore \qquad P = \sum_{i=o}^{\infty} (A^i)^T Q A^i \tag{4.35}$$

We note that:

$$(1) \quad P^T = \left(\sum_{i=o}^{\infty} (A^i)^T Q A^i\right)^T = \sum_{i=o}^{\infty} (A^i)^T Q^T A^i \tag{4.36}$$

But $Q^T = Q$

$$\therefore \qquad P^T = P \tag{4.37}$$

$$(2) \quad \underline{x}^T P \underline{x} = \sum_{i=o}^{\infty} (A^i \underline{x})^T Q (A^i \underline{x}) \tag{4.38}$$

But Q is positive definite

$$\therefore \qquad (A^i \underline{x})^T Q (A^i \underline{x}) > 0 \tag{4.39}$$

provided $A^i \underline{x} \neq 0$.

⊗ If

$$|A| \neq 0 \qquad |A^i| \neq 0 \qquad \therefore \quad \underline{x} \neq 0 \tag{4.40}$$

$$\therefore \qquad \underline{x}^T P \underline{x} > 0 \qquad \underline{x} \neq 0 \tag{4.41}$$

Thus, if A is a stability matrix there exists a P, symmetric and positive definite such that Eq. (4.32) is satisfied.

⊗ Note If $|A| = 0$, it appears that (4.41) is not satisfied. However, if $|A| = 0$, then A has one or more zero eigenvalues, and the displacements in these modes vanish after one step, thus the problem is really one in (N-h) dimensions, where h is the multiplicity of the zero eigenvalue. Thus, if in (4.29) $\underline{x} \in R$, the range space of A, P is positive definite.

(ii) Linear Difference Equations with Periodic Coefficients

Theorem 7 Given the difference equation

$$\left.\begin{aligned} \underline{x}_{n+1} &= A(n)\underline{x}_n \\ A(n+N) &= A(n) \end{aligned}\right\} \tag{4.42}$$

(i) If the principal matrix solution X_N is simple, necessary and sufficient conditions for Liapunov stability are that the eigenvalues of X_N should be less than or equal to unity in modulus.

(ii) If the principal matrix solution X_N is defective , necessary and sufficient conditions for stability are that the eigenvalues of X_N should be less than unity in modulus.

Proof As previously shown, the solution of (4.42) with initial data $\underline{x}_o = \underline{c}$ is given by

$$\underline{x}_n = X_n \underline{c} \tag{4.43}$$

where

$$X_n = Q(n)C^n \tag{4.44}$$

$Q(n+N) = Q(n)$ is a periodic matrix and $C = X_N^{1/N}$ is a constant matrix.

If X_N is simple, there exists a similarity matrix $T \ni$

$$T^{-1}X_N T = \Lambda \text{ -- a diagonal matrix} \tag{4.45}$$

X_n therefore has the representation

$$X_N = T\Lambda T^{-1} \tag{4.46}$$

$$\therefore \qquad C^n = T\Lambda^{\frac{n}{N}} T^{-1} \tag{4.47}$$

If $|\lambda(X_N)| \leq 1$, clearly C^n and hence X_n remains bounded as $n \to \infty$, therefore (4.42) is Liapunov stable at $\underline{x} = 0$.

The remainder of the proof closely follow that of Theorem 5 and will not be repeated here.

Theorem 8 Theorem 6 can be generalized to the case of linear difference equations with periodic coefficients.

Given the difference equation

$$\left.\begin{array}{l} \underline{x}_{n+1} = A(n)\underline{x}_n \\ A(n+N) = A(n) \quad , \quad |A(n)| \neq 0 \quad \|A(n)\| < \infty \ \forall n \end{array}\right\} \tag{4.48}$$

Then (4.46) is Liapunov asymptotically stable at $\underline{x} = 0$ iff there exists a symmetric positive definite periodic matrix $P(k)$ such that

$$\left.\begin{array}{ll} \text{i)} & P(k+N) = P(k) = P^T(k) \ \text{ positive definite} \\ \text{ii)} & A^T(k)P(k+1)A(k) - P(k) = -Q(k) \\ \text{iii)} & Q^T(k) = Q(k) = Q(k+N) \ \text{ positive definite} \end{array}\right\} \tag{4.49}$$

$$\forall k$$

Proof Sufficiency

Suppose that there exists a matrix $P(k)$ satisfying (4.49). Let

$$V_n = \underline{x}_n^T P(n)\underline{x}_n \tag{4.50}$$

Since $P(n)$ is symmetric and positive definite for all n, V_n is positive definite.

$$V_{n+1} = \underline{x}_{n+1}^T P(n+1)\underline{x}_{n+1} \quad (4.51)$$

Using (4.48)

$$V_{n+1} = \underline{x}_n^T A^T(n)P(n+1)A(n)\underline{x}_n \quad (4.52)$$

$$\therefore \quad \Delta V_n = V_{n+1} - V_n = \underline{x}_n^T(A^T(n)P(n+1)A(n) - P(n))\underline{x}_n \quad (4.53)$$

Using (4.49) ii)

$$\Delta V_n = -\underline{x}_n^T Q(n)\underline{x}_n < 0 \quad (4.54)$$

Since V_n vanishes only at the origin

$$\therefore \quad V_n \to 0 \quad \text{as} \quad n \to \infty \quad (4.55)$$

∴ Equation (4.48) is Liapunov asymptotically stable at the origin.

Necessity

Let A be a stability matrix so that ∀ solutions of (4.48) tend to zero as $t \to \infty$.

Let $P(k)$ satisfy (4.49) i.e.

$$A^T(k)P(k+1)A(k) - P(k) = -Q(k) \quad (4.56)$$

similarly

$$A^T(k+1)P(k+2)A(k+1) - P(k+1) = -Q(k+1) \quad (4.57)$$

If (4.57) is premultiplied by $A^T(k)$ and post multiplied by $A(k)$ we obtain

$$(A(k+1)A(k))^T P(k+2)(A(k+1)A(k)) - A(k)^T P(k+1)A(k) = -A(k)^T Q(k+1)A(k) \quad (4.58)$$

Similarly

$$A^T(k+2)P(k+3)A(k+2) - P(k+2) = - Q(k+2) \qquad (4.59)$$

If (4.59) is premultiplied by $(A(k+1)A(k))^T$ and postmultiplied by $A(k+1)A(k)$, we obtain

$$\begin{aligned}(A(k+2)A(k+1)A(k))^T P(k+3)(A(k+2)A(k+1)A(k) \\ - (A(k+1)A(k)^T P(k+2)(A(k+1)A(k)) \\ = - (A(k+1)A(k)^T Q(k+2)A(k+1)A(k))\end{aligned} \qquad (4.60)$$

Repeating the procedure n times gives

$$\left(\prod_{i=o}^{n-1} A(k+i)\right)^T P(k+n)\left(\prod_{i=o}^{n-1} A(k+i)\right) - \left(\prod_{i=o}^{n-2} A(k+i)\right.^T P(k+n-1)\prod_{i=o}^{n-2} A(k+i)\Bigg)$$

$$= - \left(\prod_{i=o}^{n-2} A(k+i)\right)^T Q(k+n-1)\prod_{i=o}^{n-2} A(k+i)\Bigg) \qquad (4.61)$$

If these n equations are added, we find that just as in (4.33) we obtain cancellation in pairs and finally we have:

$$\left(\prod_{i=o}^{n-1} A(k+i)\right)^T P(k+n) \left(\prod_{i=o}^{n-1} A(k+i)\right) - P(k)$$

$$= - \sum_{j=o}^{n=1} \left(\prod_{i=o}^{j-1} A(k+i)\right)^T Q(k+j)\left(\prod_{i=o}^{j-1} A(k+i)\right) \qquad (4.62)$$

Now

$$\prod_{i=o}^{n-1} A(K+i) = \Phi(k+n, k) \qquad (4.63)$$

where

$$\Phi(m,k) = X_m X_k^{-1} \tag{4.64}$$

satisfies the equation

$$\begin{aligned} &\Phi(n+1,k) = A(n)\Phi(n,k) \\ &\Phi(k,k) = I \end{aligned} \tag{4.65}$$

Since $\forall$ solutions of (4.48) tend to zero as $n \to \infty$

$$\Phi(m,k) \to 0 \quad \text{as} \quad m \to \infty \tag{4.66}$$

Thus as $n \to \infty$ Eq. (4.62) becomes

$$P(k) = \sum_{j=k}^{\infty} \Phi(j,k)^T Q(j)\Phi(j,k) \tag{4.67}$$

$$\left.\begin{aligned}
&\therefore\ \text{i)} \quad P(k)^T = \sum_{j=k}^{\infty} \Phi(j,k)^T Q(j)\Phi(j,k))^T = P(k) \\
&\text{ii)} \quad P(k+N) = \sum_{j=k+n}^{\infty} \Phi(j,k+n)^T Q(j)\Phi(j,k+n) \\
&\qquad = \sum_{j=k}^{\infty} \Phi(j,k)^T Q(j)\Phi(j,k) = P(k) \quad \text{since } Q(j+N)=Q(j) \\
&\text{iii)} \quad \underline{x}^T P(h)\underline{x} = \sum_{j=k}^{\infty} (\Phi(j,k)\underline{x})^T Q(j)(\Phi(j,k)\underline{x}) \\
&\qquad > 0 \qquad \underline{x} \neq 0
\end{aligned}\right\} \tag{4.68}$$

Since $|\Phi(j+k)| \neq 0$ if $|A(k)| \neq 0 \; \forall \; k \therefore P(k)$ is symmetric, periodic of period N, and positive definite. This completes the proof of the theorem.

We note in passing that Theorem 6 is a special case of Theorem 8 when $N = 1$.

(iii) Linear Difference Equations with Variable Coefficients

Given the difference equation

$$\underline{x}_{n+1} = A(n)\, \underline{x}_n \tag{4.69}$$

we can say very little* about the stability of equation (4.69) for the general case of arbitrary step varying matrices $A(n)$. If the matrix $A(n)$ can be represented as

$$A(n) = A_0(n) + B(n) \tag{4.70}$$

where $A_0(n)$ is either a constant or a periodic matrix, then in a number of cases we can develop sufficient, but not necessary conditions for stability.

Theorem 9

Given the linear difference equation

$$\underline{x}_{n+1} = A_0(n)\, \underline{x}_n + B(n)\, \underline{x}_n \qquad |A_0(n)| \neq 0 \tag{4.71}$$

where $A_0(n)$ is either a constant or a periodic matrix,

If i) $\forall$ solutions of $\underline{x}_{n+1} = A_0(n)\, \underline{x}_n$ are bounded as $n \to \infty$

ii) $\sum_{i=0}^{\infty} \|B(i)\| = b_0 < \infty$

iii) $\|\underline{x}_0\| = \|\underline{c}\| < \infty$

---

*See Appendix 1.

Then $\forall$ solutions of (4.71) are bounded for $\forall n$. Before proving Theorem 9 we shall establish two important lemmas.

Lemma 1 (Discrete Form of Bellman-Gronwall's Lemma)

If
$$\left.\begin{aligned} &\theta(n) \le C + \sum_{i=0}^{n-1} \psi(i)\,\theta(i) \\ &\theta(i), \psi(i), C \ge 0 \end{aligned}\right\} \tag{4.72}$$

Then
$$\theta(n) \le C \prod_{i=0}^{n-1} [1 + \psi(i)] \tag{4.73}$$

Proof

From (4.72)

$$\frac{\theta(n)\ \psi(n)}{C + \sum_{i=0}^{n-1} \psi(i)\ \theta(i)} \le \psi(n) \tag{4.74}$$

$$\therefore \quad 1 + \frac{\theta(n)\ \psi(n)}{C + \sum_{i=0}^{n-1} \psi(i)\ \theta(i)} \le [1 + \psi(n)] \tag{4.75}$$

$$\therefore \quad [C + \sum_{i=0}^{n} \psi(i)\ \theta(i)] \le [C + \sum_{i=0}^{n-1} \psi(i)\ \theta(i)][1 + \psi(n)] \tag{4.76}$$

$$\therefore \quad [C + \sum_{i=0}^{n-1} \psi(i)\ \theta(i)] \le [C + \sum_{i=0}^{n-2} \psi(i)\ \theta(i)][1 + \psi(n-1)] \tag{4.77}$$

$$\therefore \quad [C + \sum_{i=0}^{n} \psi(i)\ \theta(i)] \le [C + \sum_{i=0}^{n-2} \psi(i)\ \theta(i)][1 + \psi(n)][1 + \psi(n-1)] \tag{4.78}$$

Hence
$$[C + \sum_{i=0}^{n} \psi(i)\ \theta(i)] \le C \prod_{i=0}^{n} [1 + \psi(i)] \tag{4.79}$$

But $$\theta_{n+1} \leq C + \sum_{i=0}^{n} \psi(i)\theta(i) \qquad (4.80)$$

$$\therefore \quad \theta(n) \leq C \prod_{i=0}^{n-1} [1 + \psi(i)] \qquad (4.81)$$

Lemma 2

The product series S_n:

$$S_n = \prod_{i=0}^{n} (1 + v_i) \qquad v_i \geq 0 \qquad (4.82)$$

is convergent iff the series V_n

$$V_n = \sum_{i=0}^{n} v_i \qquad (4.83)$$

is convergent.

Proof:

1) The product series (4.82) is convergent if the series L_n

$$L_n = \sum_{i=0}^{n} \ln(1 + v_i) \qquad (4.84)$$

is convergent. This follows immediately from the fact that

$$S_n = e^{\ln S_n} = e^{L_n} \qquad (4.85)$$

$$\therefore \text{ If } \lim_{n\to\infty}(\ln L_n) = L, \text{ then } \lim_{n\to\infty}(S_n) = e^{L} = S \qquad (4.86)$$

2) We know that if S_n is convergent, $v_n \to 0$ as $n \to \infty$. ∴ let N be such that for $n \geq N$ $v_n \leq \frac{1}{2}$.

Now

$$\frac{1}{2}v_n = v_n(1 - \frac{1}{4}\frac{1}{(1-\frac{1}{2})}) = v_n(1 - \frac{1}{2^2} - \frac{1}{2^3}\cdots) \quad (4.87)$$

But $$\ell n(1+v_n) = (v_n - \frac{v_n^2}{2} + \frac{v_n^2}{3} + \cdots) \quad (4.88)$$

$$= v_n(1 - \frac{v_n}{2} + \frac{v_n^2}{3}\cdots) \quad (4.89)$$

If $v_n \leq \frac{1}{2}$

then $$\ell n(1+v_n) = v_n(1 - \frac{1}{2^2} + \frac{1}{12} - \cdots) \quad (4.90)$$

$$\therefore \quad \frac{1}{2}v_n < \ell n(1+v_n) < v_n(1 + \frac{1}{2^2} + \frac{1}{2^3} + \cdots) \quad (4.91)$$

$$\therefore \quad \frac{1}{2}v_n < \ell n(1+v_n) < \frac{3}{2}v_n \quad (4.92)$$

Thus

i) If $\sum_{i=0}^{\infty} v_i < \infty$ a) $\sum_{i=0}^{N} v_i < \infty$ and $\sum_{k+1}^{\infty} v_i < \infty$ (4.93)

$$\therefore \quad \sum_{i=N+1}^{\infty} \ell n(1+v_i) < \frac{3}{2}\sum_{i=N+1}^{\infty} v_i < \infty \quad (4.94)$$

ii) If $\sum_{i=0}^{\infty} \ell n(1+v_i) < \infty$ (4.95)

then $$\sum_{i=0}^{\infty} v_i < 2\sum_{i=0}^{\infty} \ell n(1+v_i) < \infty \quad (4.96)$$

Returning now to the proof of Theorem 9, using equation (3.22),

$$\underline{x}_n = X_n\underline{x}_0 + \sum_{i=0}^{n-1} X_n X_{i+1}^{-1} B(i)\underline{x}_i \quad (4.97)$$

Taking norms of both sides of equation (4.97)

$$\|\underline{x}_n\| \leq \|X_n\| \; \|\underline{c}\| + \sum_{i=0}^{n-1} \|X_n\| \; \|X_{i+1}^{-1}\| \; \|B(i)\| \, \|\underline{x}_i\| \tag{4.98}$$

But, by hypothesis i), $\|X_n\| \leq M_1$, $\|X_n\| \; \|X_{i+1}^{-1}\| \leq M_2 \quad \forall i,n$

$$\therefore \quad \|\underline{x}_n\| \leq M_1 \|c\| + M_2 \sum_{i=0}^{k-1} \|B(i)\| \; \|\underline{x}_i\| \tag{4.99}$$

Using Lemma 1 with $C = M_1 \|\underline{c}\|$, $\theta(i) = \|\underline{x}_i\|$ and

$$\psi(i) = M_2 \|B(i)\| \tag{4.100}$$

we have

$$\|\underline{x}_n\| \leq M_1 \|\underline{c}\| \prod_{i=0}^{n-1} (1 + M_2 \|B(i)\|) \tag{4.101}$$

But by hypothesis ii) $\sum_{i=0}^{\infty} \|B(i)\| = b_0 < \infty$

$\therefore$ By Lemma 2, $\prod_{i=0}^{\infty} (1 + M_2 \|B(i)\|) < d_0 < \infty$

$$\therefore \quad \|\underline{x}_n\| \leq M_1 d_0 \|\underline{c}\| \quad \forall n \tag{4.102}$$

$\therefore$ Using hypothesis iii) we see that

$$\|\underline{x}_n\| < \infty \qquad \forall n \tag{4.103}$$

Thus proving the theorem.

Theorem 10

Given the linear difference equation

$$\underline{x}_{n+1} = A_0(n)\underline{x}_n + B(n)\underline{x}_n \qquad |A_0(n)| \neq 0 \tag{4.104}$$

where $A_o(n)$ is either a constant or a periodic matrix. If

i) $A_o(n)$ is a stability matrix, i.e., solutions of $\underline{x}_{n+1} = A_o(n)\underline{x}_n$ tend to zero as $n \to \infty$

ii) $\|B(n)\| \leq b_o \;\forall n$ and b_o sufficiently small.

Then $\forall$ solutions of (4.104) tend to zero as $n \to \infty$, and the origin is Liapunov asymptotically stable.

Proof. As before,

$$\underline{x}_n = X_n\underline{c} + \sum_{i=0}^{n-1} X_n X_{i+1}^{-1} B(i)\underline{x}_i \tag{4.105}$$

Taking norms of both sides of equation (4.105)

$$\|\underline{x}_n\| \leq \|\underline{x}_n\| \; \|\underline{c}\| + \sum_{i=0}^{n-1} \|X_n\| \; \|X_{i+1}^{-1}\| \|B(i)\| \|\underline{x}_i\| \tag{4.106}$$

Using hypothesis i)

$$\left.\begin{aligned} \|X_n\| &\leq M_1\delta^n , \qquad \delta < 1 \\ \|X_n\| \; \|X_{i+1}^{-1}\| &\leq M_2 \, \delta^{n-c-1} \end{aligned}\right\} \tag{4.107}$$

Using hypothesis ii) equation (4.106) becomes

$$\|\underline{x}_n\| \leq M_1 \|\underline{c}\| \delta^n + M_2 b_o \sum_{i=0}^{n-1} \delta^{n-i-1} \|\underline{x}_i\| \tag{4.108}$$

Multiplying both sides of equation (4.108) by δ^{-n} and setting $\theta(i) = \|\underline{x}_i\| \, \delta^{-i}$; $C = M_1 \|\underline{c}\|$; $\psi(i) = M_2 b_o$, and using Lemma 1

$$\|\underline{x}_n\| \, \delta^{-n} \leq M_1 \|c\| \prod_{i=0}^{n-1} \left(1 + \frac{M_2 b_o}{\delta}\right) \tag{4.109}$$

$$\therefore \quad \|\underline{x}_n\| \leq M_1 \|\underline{c}\| \, \delta^n \prod_{i=0}^{n-1} \left(1 + \frac{M_2 b_o}{\delta}\right) = \frac{M_1 \|c\|}{\delta} \left(\delta + M_2 b_o\right)^n \tag{4.110}$$

Hence, if b_o is sufficiently small,

$$\delta + M_2 b_o < 1 \tag{4.111}$$

$$\therefore \quad \left.\begin{array}{l} \|\underline{x}_n\| \leq \dfrac{M_1 \|\underline{c}\|}{\delta} \quad \forall n \\[2ex] \|\underline{x}_n\| \to 0 \quad \text{as} \quad n \to \infty \end{array}\right\} \tag{4.112}$$

Therefore, the trivial solution of (4.104) is L.A.S.

Theorem 10a

Given the same hypotheses as Theorem 10, we can prove the theorem using Liapunov's direct approach.

Proof. Since $A_0(n)$ is a stability matrix, we know that there exists a symmetric, positive definite, periodic matrix $P(n)$ such that

$$A_0^T(n)\ P(n+1)\ A_0(n) - P(n) = -Q(n) \tag{4.113}$$

$$Q(n) = Q(n)^T = Q(n+N) \quad \text{positive definite}$$

Let $\quad V_n = \underline{x}_n^T\ P(n)\underline{x}_n \qquad (4.114)$

then $\quad V_{n+1} = \underline{x}_{n+1}^T\ P(n+1)\underline{x}_{n+1} \qquad (4.115)$

Using equation (4.90)

$$V_{n+1} = \underline{x}_n^T(A_0^T(n)\ P(n+1)\ A_0(n))\underline{x}_n + \underline{x}_n^T(B^T(n)\ P(n+1)\ A_0(n))\underline{x}_n + \underline{x}_n^T(A_0^T(n)\ P(n+1)B(n))\ \underline{x}_n + \underline{x}_n^T\ B(n)^T\ P(n+1)\ B(n)\underline{x}_n \tag{4.116}$$

$$\therefore \quad \Delta V_n = V_{n+1} - V_n = -\underline{x}_n^T(A_0^T(n)\ P(n+1)\ A_0(n) - P(n))\underline{x}_n + \underline{x}_n^T(S(n))\underline{x}_n \tag{4.117}$$

where $S(n) = B^T(n)\ P(n+1)\ A_o(n) + A_o^T(n)\ P(n+1)\ B(n)$

$$+ B^T(n)\ P(n+1)\ B(n) = S^T(n) \tag{4.118}$$

Using (4.112), equation (4.118) becomes

$$\Delta V_n = -\underline{x}_n^T\ Q(n)\underline{x}_n + \underline{x}_n^T\ S(n)\underline{x}_n \tag{4.119}$$

Since Q(n) is positive definite for $\forall$ n, it is clear that by making $||B(n)||$, and hence the elements of B(n), sufficiently small, ΔV_n can be made negative definite.

Hence for $||B(n)||$ sufficiently small,

$$\Delta V_n < 0 \tag{4.120}$$

$$\therefore \quad V_{n+1} < V_n < V_{n-1} \cdot \cdot \cdot \cdot < V_1 < V_o \tag{4.121}$$

Since V_n is positive definite and vanishes only at the origin, therefore V_n and hence $||\underline{x}_n||$ tends to zero as $n \to \infty$, and since V_n is bounded above by V_o, $||\underline{x}_n||$ is bounded for all n. Therefore the trivial solution of (4.104) is Liapunov stable.

Note: Let $\lambda(n)$ be the smallest eigenvalue of Q(n) and $\mu(n)$ be the largest eigenvalue of S(n) in absolute value.

Let $r = \underset{n}{\text{Min}}\ \lambda(n)$, r is positive, since Q(n) is positive definite.

$s = \underset{n}{\text{Max}}\ \mu(n)$, we note that since S(n) tends to zero as b_o tends to zero, s may be made arbitrarily small by making b_o sufficiently small. Now

$$\Delta V_n \leq -\underline{x}_n^T(r-s)\underline{x}_n$$

and hence by making b_o sufficiently small ΔV_n can be made negative definite.

---

b) Stability of Nonlinear Difference Equations

(i) Stability of Explicit Nonlinear Difference Equations

Theorem 11 (Liapunov-Poincaré)

Given the nonlinear difference equation

$$\left.\begin{aligned} \underline{x}_{n+1} &= A(n)\underline{x}_n + \underline{f}(\underline{x}_n,n) \\ \underline{x}_0 &= \underline{c} \end{aligned}\right\} \tag{4.122}$$

where A(n) is either a constant matrix or a periodic matrix,

$$\left.\begin{aligned} &\text{If} \quad \text{i)} \quad A(n) \text{ is a stability matrix} \\ &\text{ii)} \quad \lim_{\|\underline{x}\| \to 0} \frac{\|\underline{f}(\underline{x},n)\|}{\|\underline{x}\|} = 0 \qquad \forall n \\ &\text{iii)} \quad \|\underline{c}\| \text{ is sufficiently small} \end{aligned}\right\} \tag{4.123}$$

then $\forall$ solution of equations (4.122) are Liapunov asymptotically stable.

Proof. If A(n) is a stability matrix, then by Theorem 8 there exists a symmetric, positive definite, periodic matrix P(n) such that

$$\left.\begin{aligned} &A(n)^T P(n+1)\, A(n) - P(n) = -Q(n) \\ &Q(n) = Q^T(n) = Q(n+N) \text{ positive definite} \end{aligned}\right\} \tag{4.124}$$

Let

$$V_n = \underline{x}_n^T P(n)\underline{x}_n \tag{4.125}$$

$$V_{n+1} = \underline{x}_{n+1}^T P(n+1)\, \underline{x}_{n+1} \tag{4.126}$$

Making use of equations (4.122),

$$\begin{aligned} V_{n+1} &= \underline{x}_n^T(A^T(n)\, P(n+1)\, A(n))\underline{x}_n + \underline{x}_n^T(A^T(n)\, P(n+1)\, \underline{f}(\underline{x}_n n)) \\ &\quad + \underline{f}(\underline{x}_n,n)^T(P(n+1)\, A(n))\underline{x}_n + \underline{f}^T(\underline{x}_n,n)\, P(n+1)\, \underline{f}(\underline{x}_n,n) \end{aligned} \tag{4.127}$$

Thus $\Delta V_n = V_{n+1} - V_n$

$$= \underline{x}_n^T(A^T(n)\ P(n+1)\ A(n) - P(n))\underline{x}_n$$

$$+ \underline{x}_n^T(A^T(n)\ P(n+1)\ \underline{f}(\underline{x}_n,n) + \underline{f}^T(\underline{x}_n,n)(P(n+1)\ A(n))\underline{x}_n$$

$$+ \underline{f}^T(\underline{x}_n,n)\ P(n+1)\ \underline{f}(\underline{x}_n,n) \qquad (4.128)$$

Using (4.124) equation (4.128) becomes

$$\Delta V_n = -\underline{x}_n^T\ Q(n)\underline{x}_n$$

$$+ \underline{x}_n^T(A^T(n)\ P(n+1)\ \underline{f}(\underline{x}_n,n) + f^T(\underline{x}_n,n)(P(n+1)\ A(n))\underline{x}_n$$

$$+ \underline{f}^T(\underline{x}_n,n)\ P(n+1)\ \underline{f}(\underline{x}_n,n) \qquad (4.129)$$

Using hypothesis ii), $||\underline{f}(x,n)|| \sim O(||\underline{x}||^2)$ as $||\underline{x}|| \to 0$

Hence a) $\underline{x}_n^T\ Q(n)\underline{x}_n^T \sim O(||\underline{x}||^2)$

b) $\underline{x}_n^T(A^T(n)\ P(n+1)\ \underline{f}(\underline{x}_n,n) + \underline{f}^T(\underline{x}_n,n)(P(n+1)\ A(n))\underline{x}_n \sim O(||\underline{x}_n||^3)$

c) $\underline{f}^T(\underline{x}_n,n)\ P(n+1)\ \underline{f}(\underline{x}_n,n) \sim O(||\underline{x}_n||^4)$

$$\text{as} \quad ||\underline{x}_n|| \to 0 \qquad (4.130)$$

Thus for $||\underline{x}_n||$ sufficiently small, the sign of ΔV_n is that of the first term $\therefore \Delta V_n$ is negative definite. Hence,

$$V_{n+1} < V_n < V_{n-1} < \cdot \cdot \cdot < V_1 < V_0 \qquad (4.131)$$

Thus if $||\underline{c}||$ is sufficiently small,

$$\Delta V_n < 0\ , \quad \forall n \qquad (4.132)$$

and since V_n is positive definite and vanishes only at the origin, therefore $V_n \to 0$, and hence $||\underline{x}_n|| \to 0$ as $n \to \infty$. Thus equation (4.122) is

Liapunov asymptotically stable at the origin.

Theorem 12

Given the nonlinear difference equation

$$\left.\begin{aligned} \underline{x}_{n+1} &= [A_0(n) + B(n)]\underline{x}_n + \underline{f}(\underline{x}_n, n) \\ \underline{x}_0 &= \underline{c} \end{aligned}\right\} \qquad (4.133)$$

where $A_0(n)$ is either a constant matrix or a periodic matrix.

$$\left.\begin{aligned} &\text{If i) } A_0(n) \text{ is a stability matrix} \\ &\quad\text{ii) } \|B(n)\| \text{ is sufficiently small} \\ &\quad\text{iii) } \lim_{\|\underline{x}\| \to 0} \frac{\|\underline{f}(\underline{x},n)\|}{\|\underline{x}\|} = 0 \quad \forall n \\ &\quad\text{iv) } \|\underline{c}\| \text{ is sufficiently small} \end{aligned}\right\} \qquad (4.134)$$

Then $\forall$ solutions of equation (4.133) are Liapunov asymptotically stable.

Proof. The proof follows along exactly the same lines as Theorem 10a and Theorem 11, and will not be repeated here.

ii) Stability of Implicit Nonlinear Difference Equations

Theorem 13

Given the implicit nonlinear difference equations

$$\left.\begin{aligned} \underline{x}_{n+1} &= A(n)\underline{x}_n + \underline{f}(\underline{x}_n, \underline{x}_{n+1}, n) \\ \underline{x}_0 &= \underline{c} \qquad |A(n)| \neq 0 \\ & \qquad\qquad \|A(n)\| < \infty \end{aligned}\right\} \forall n \qquad (4.13.5)$$

where $A(n)$ is either a constant matrix or a periodic matrix. If

$$\left.\begin{array}{ll} \text{i)} & A(n) \text{ is a stability matrix} \\ \text{ii)} & \lim\limits_{\|\underline{x}\|,\|y\|\to 0} \dfrac{\|\underline{f}(\underline{x},\underline{y},n)\|}{\|x\|+\|y\|} = 0 \\ \text{iii)} & \|\underline{c}\| \text{ is sufficiently small} \end{array}\right\} \tag{4.136}$$

Then $\forall$ solutions of equation (4.135) are Liapunov asymptotically stable.

Proof. Since $A(n)$ is a stability matrix, then by Theorem 8 there exists a symmetric, positive definite, periodic matrix $P(n)$ such that:

$$\left.\begin{array}{ll} \text{i)} & A(n)^T P(n+1) A(n) - P(n) = -Q(n) \\ \text{ii)} & Q(n) = Q^T(n) = Q(N+n) \text{ positive definite} \end{array}\right\} \tag{4.137}$$

Let $$V_n = \underline{x}_n^T P(n)\underline{x}_n > 0 \qquad \underline{x}_n \neq 0 \tag{4.138}$$

$$V_{n+1} = \underline{x}_{n+1}^T P(n+1)\underline{x}_{n+1} \tag{4.139}$$

Making use of (4.135),

$$\begin{aligned} V_{n+1} = {} & \underline{x}_n^T(A^T(n)\ P(n+1)\ A(n))\underline{x}_n + \underline{x}_n^T(A^T(n)\ P(n+1))\ \underline{f}(\underline{x}_n,\underline{x}_{n+1},n) \\ & + \underline{f}^T(\underline{x}_n,\underline{x}_{n+1},n)(P(n+1)\ A(n))\underline{x}_n \\ & + \underline{f}^T(\underline{x}_n,\underline{x}_{n+1},n)(P(n+1))\ \underline{f}(\underline{x}_n,\underline{x}_{n+1},n) \end{aligned} \tag{4.140}$$

$$\begin{aligned} \therefore \quad \Delta V_n &= V_{n+1} - V_n \\ &= \underline{x}_n^T(A^T(n)\ P(n+1))\ \underline{f}(\underline{x}_n,\underline{x}_{n+1},n) + \underline{x}_n^T(A^T(n)P(n+1)A(n)-P(n))\underline{x}_n \\ &+ \underline{f}^T(\underline{x}_n,\underline{x}_{n+1},n)(P(n+1)\ A(n))\underline{x}_n \\ &+ \underline{f}^T(\underline{x}_n,\underline{x}_{n+1},n)(P(n+1))\ \underline{f}(\underline{x}_n,\underline{x}_{n+1},n) \end{aligned} \tag{4.141}$$

Making use of equation (4.137)

$$\Delta V_n = -\underline{x}_n^T Q(n)\underline{x}_n + \underline{x}_n^T(A^T(n)P(n+1))\,\underline{f}(\underline{x}_n,\underline{x}_{n+1},n)$$

$$+ \underline{f}^T(\underline{x}_n,\underline{x}_{n+1},n)(P(n+1)A(n))\underline{x}_n$$

$$+ \underline{f}^T(\underline{x}_n,\underline{x}_{n+1},n)\; P(n+1)\; \underline{f}(\underline{x}_n,\underline{x}_{n+1},n) \qquad (4.142)$$

From equation (4.135)

$$||\underline{x}_{n+1}|| + ||\underline{x}_n|| \le ||I + A(n)||\; ||\underline{x}_n|| + ||\underline{f}(\underline{x}_n,\underline{x}_{n+1},n)|| \qquad (4.143)$$

$$\therefore \quad ||\underline{x}_{n+1}|| + ||\underline{x}_n|| \le M_1\,||\underline{x}_n|| + ||\underline{f}(\underline{x}_n,\underline{x}_{n+1},n)|| \qquad (4.144)$$

From (4.136ii)

$$||\underline{f}(\underline{x}_n,\underline{x}_{n+1},n)|| \le M_2(\delta)(\,||\underline{x}_n|| + ||\underline{x}_{n+1}||\,)^2$$

$$\text{for } ||\underline{x}_n|| + ||\underline{x}_{n+1}|| \le \delta \qquad (4.145)$$

where $M_2(\delta) \sim O(1)$ as $\delta \to 0$

From (4.144) and (4.145)

$$||\underline{x}_{n+1}|| + ||\underline{x}_n|| \le \frac{M_1\,||\underline{x}_n||}{1 - M_2(\delta)\delta} = M_3||\underline{x}_n|| < \delta \qquad (4.146)$$

if $||\underline{x}_n|| < \frac{\delta}{M_3}$ is sufficiently small.

Thus, if $||\underline{x}_n||$ is sufficiently small, the first term in (4.142) is of order $||\underline{x}_n||^2$, while the second and third terms are of order $||\underline{x}_n||^3$, and the fourth term is of order $||\underline{x}_n||^4$. Hence, if $||\underline{x}_n||$ is sufficiently small, the sign of ΔV_n is that of the first term, therefore

$$\Delta V_n < 0 \tag{4.147}$$

Thus, if $||\underline{c}||$ is sufficiently small, ΔV_n is negative for all n, therefore

$$V_{n+1} < V_n < V_{n-1} < \cdot \cdot \cdot < V_1 < V_0 \tag{4.148}$$

Since V_n is positive definite and vanishes only at $\underline{x} = \underline{0}$, therefore $V_n \to 0$ as $n \to \infty$, and equation (4.135) is Liapunov asymptotically stable at $\underline{x} = 0$, provided the initial data are sufficiently small.

<u>Theorem 14</u> (Liapunov-Poincaré)

Given the nonlinear difference equations

$$\begin{aligned} \underline{x}_{n\_1} &= A(n)\underline{x}_n + \underline{f}(\underline{x}_n, n) \\ \underline{x}_0 &= \underline{c} \end{aligned} \tag{4.149}$$

where A(n) is either a constant matrix or a periodic matrix,

If i) there exists at least one unstable solution of the equation

$$\underline{x}_{n+1} = A(n)\underline{x}_n$$

ii) $$\lim_{||\underline{x}|| \to 0} \frac{||\underline{f}(\underline{x},n)||}{||\underline{x}||} = 0 \qquad \forall n$$

iii) $||\underline{c}||$ is sufficiently small

(4.150)

Then there exist unstable solutions of equation (4.149).

Proof. Let

$$\left.\begin{aligned} R^N &= X_k^{-1} X_{N+k} = X_N \\ \theta_k &= X_k R^{-k} \end{aligned}\right\} \qquad (4.151)$$

Let $\underline{x}_k = \theta_k \underline{y}_k$ (4.152)

Substituting into equation (4.149)

$$\left.\begin{aligned} \underline{y}_{n+1} &= \theta_{n+1}^{-1} A(n)\, \theta_n \underline{y}_n + \theta_{n+1}^{-1} \underline{f}(\theta_n y_n, n) \\ y_o &= \theta_o^{-1} \underline{c} \end{aligned}\right\} \qquad (4.153)$$

Now

$$\theta_{n+1}^{-1} A(n)\theta_n = R^{n+1} X_{n+1}^{-1} A(n) X_n R^{-n} \qquad (4.154)$$

But $X_{n+1} = A(n) X_n$

$$\therefore \quad \theta_{n+1}^{-1} A(n)\theta_n = R \qquad (4.155)$$

$$\therefore \quad \left.\begin{aligned} \underline{y}_{n+1} &= R\underline{y}_n + \underline{g}(\underline{y}_n, n) \\ \underline{y}_o &= \underline{b} \;;\quad \underline{b} = \theta_o^{-1} \underline{c} \;;\quad \underline{g}(\underline{y}_n, n) = \theta_{n+1}^{-1} \underline{f}(\theta_n \underline{y}_n, n) \end{aligned}\right\} \qquad (4.156)$$

Since $R = X_N^{1/N}$, therefore from (4.150i), R must have at least one eigenvalue of modulus greater than unity.

Suppose that R is simple, and that the first k eigenvalues have modulus greater than unity, suppose that the remaining (L-k) eigenvalues have modulus less than unity. Since R is simple, there exists a similarity matrix T such that

$$T^{-1}RT = \Lambda$$

where $|\lambda_i| > 1 \quad i \in (1,k)$ (4.157)

$$|\lambda_j| < 1 \quad j \in (k+1,L)$$

Let $\underline{y}_n = T\,\underline{z}_n$ (4.158)

Then $\underline{z}_{n+1} = \Lambda\,\underline{z}_n + \underline{h}(\underline{z}_n,n)$

$$\underline{z}_0 = \underline{d} \tag{4.159}$$

where $\underline{h} = T^{-1}\underline{g}(T\underline{z}_n,n)$

$$\underline{z}_0 = \underline{d} = T^{-1}\,\underline{b}$$

Let $P = \left[\begin{array}{c|c} I_k & 0 \\ \hline 0 & -I_{L-k} \end{array}\right]$

$$V_n = \underline{Z}_n^* P \underline{Z}_n \tag{4.160}$$

It will be observed that V_n is sign indefinite.

$$V_{n+1} = \underline{Z}_{n+1}^*\,P\,\underline{Z}_{n+1} \tag{4.161}$$

Substituting from (4.145),

$$\underline{V}_{n+1} = \underline{z}_n^*\,\Lambda^*\,P\,\Lambda\,\underline{z}_n + \underline{h}^*(\underline{z}_n,n)P\,\Lambda\,\underline{z}_n$$

$$+\,\underline{z}_n^*\,\Lambda^*\,P\,\underline{h}(\underline{z}_n,n) + \underline{h}^*(\underline{z}_n,n)\,P\,\underline{h}(z_n,n) \tag{4.162}$$

$$\therefore \quad \Delta V_n = V_{n+1} - V_n$$
$$= \underline{z}_n^*[\Lambda^* P \Lambda - P]\underline{z}_n$$
$$+ (\underline{h}^*(\underline{z}_n,n)P\,\Lambda \underline{z}_n + \underline{z}_n^* \Lambda^* P\,\underline{h}(\underline{z}_n,n))$$
$$+ \underline{h}^*(\underline{z}_n,n)P\,h(\underline{z}_n,n) \qquad (4.163)$$

Now

$$\Lambda^* P \Lambda - P = \left[\begin{array}{c|c} (|\lambda_i|^2 - 1 & 0 \\ \hline 0 & (1 - |\lambda_j|^2) \end{array}\right] \qquad (4.164)$$

Since

$$|\lambda_i| > 1 \qquad i \in (1,k)$$
$$|\lambda_j| < 1 \qquad j \in (k+1,L) \qquad (4.165)$$

$\Lambda^* P \Lambda - P$ is positive definite Hermitian. Using (4.150) iii, the first term in (4.163) is positive and of order $\|\underline{z}_n\|^2$, the second term is of order $\|\underline{z}_n\|^3$, while the fourth term is of order $\|\underline{z}_n\|^4$; thus for sufficiently small $\|\underline{z}_n\|$, the sign of ΔV_n is that of the first term and is positive.

$$\therefore \quad \Delta V_n > 0 \qquad \text{for } \|\underline{z}_n\| \text{ sufficiently small.} \qquad (4.166)$$

Since V_n is sign indefinite we can define a set

$$\Omega: \quad V_n \geq 0; \quad \|z_n\| < \delta \qquad (4.167)$$

Clearly, the origin is a boundary point of Ω. In Ω, $V_n > 0$, $\Delta V_n > 0$, therefore starting in Ω, $\underline{z}_n$ cannot approach the origin. Since $V_0 > 0$,

$\underline{z}_n$ only exit Ω through the boundary $||\underline{z}_n|| = \delta$; thus the system is unstable.

Theorem 15

Given the nonlinear implicit equation

$$\underline{x}_{n+1} = A(n)\underline{x}_n + \underline{f}(\underline{x}_n, \underline{x}_{n+1}, n)$$
$$\underline{x}_0 = \underline{c} \qquad (4.168)$$

where A(n) is either a constant or a periodic matrix,

If i) there exists at least one unstable solution of the equation

$$\underline{x}_{n+1} = A(n)\underline{x}_n$$

ii) $$\lim_{||x||, ||y|| \to 0} \frac{\underline{f}(\underline{x}, \underline{y}, n)}{||\underline{x}|| + ||y||} = 0 \qquad \forall n \qquad (4.169)$$

iii) $||\underline{c}||$ is sufficiently small

then there exist unstable solutions to equation (4.168)

Proof. The proof follows along the same lines as that of Theorems 13 and 14 and will not be repeated here.

Theorem 16 (Liapunov-Poincaré)

Given the nonlinear difference equation

$$\underline{x}_{n+1} = A(n)\underline{x}_n + \underline{f}(\underline{x}_n, n)$$
$$\underline{x}_0 = \underline{c} \qquad (4.170)$$

where $A(n)$ is either a constant matrix or a periodic matrix,

If i) the principal matrix X_N of the linear difference equation

$$\underline{x}_{n+1} = A(n)\underline{x}_n$$

has an eigenvalue ± 1, or a pair of complex conjugate eigenvalues of modulus unity

ii) $$\lim_{||x|| \to 0} \frac{||f(\underline{x},n)||}{||\underline{x}||} = 0 \qquad \forall n$$

iii) $||\underline{c}||$ sufficiently small

(4.171)

then the stability of equation (4.156) cannot be decided from the stability of the linearized equation.

Proof. If we repeat the proof of Theorem 14, we see that in this case, $(\Lambda^*P\Lambda - P)$ is only positive semidefinite, having a zero eigenvalue corresponding to $\lambda = \pm 1$, or a pair of zero eigenvalues corresponding to $|\lambda| = 1$.

Since the matrix $(\Lambda^*P\Lambda - P)$ is only positive semidefinite, we see that the sign of ΔV_n depends on the terms in $\underline{h}(\underline{z}_n,n)$. Thus the stability is not determined by the stability of the linearized equations.

Theorem 17 Theorem 16 is easily generalized to the case of implicit nonlinear difference equations.

Theorems 16 and 17 cover what are known as the "critical cases," that is, those cases in which the stability is not determined by the stability of the linearized equations.

5. DIFFERENTIAL EQUATIONS AND DIFFERENCE EQUATIONS

(a) Numerical Solution of Ordinary Differential Equations

As pointed out in the introduction, one of the more important sources of difference equations occurs in the numerical solution of ordinary differential equations.

Given the system of differential equations

$$\left.\begin{aligned} \frac{d\underline{x}}{dt} &= A(t)\underline{x} + \underline{f}(\underline{x},t) \\ \underline{x}(t_0) &= \underline{c} \qquad t_0 \le t \le T_a < \infty \end{aligned}\right\} \qquad (5.1)$$

we wish to approximate the solution of equation (5.1) by the solution of the difference equation

$$\left.\begin{aligned} \underline{y}_{n+1} &= B(n)\underline{y}_n + \underline{g}(\underline{y}_n, y_{n+1}, n) \\ \underline{y}_0 &= \underline{c} \end{aligned}\right\} \qquad (5.2)$$

such that $\underline{y}_n \simeq \underline{x}(t_n) \qquad t_{n+1} = t_n + \Delta t \qquad n = 0,1,2,\cdots,M$

The natural requirements for the approximating difference equations are that for any function $\underline{f}(\underline{x},t)$ in some class of sufficiently differentiable functions

1) They have a unique solution,
2) This solution, at least for sufficiently small Δt_n, should be close to the exact solution of equation (5.1),
3) This solution should be effectively computable.

These three points are examined in detail in books on numerical analysis and will not be pursued at length in this note.

(b) Numerical Solution of Linear Ordinary Differential Equations

Consider the system of differential equations

$$\left.\begin{aligned} \frac{d\underline{x}}{dt} &= A\underline{x} + \underline{f}(t) \qquad 0 \le t \le T_0 < \infty \\ \underline{x}(0) &= \underline{c} \qquad A = \text{a constant matrix} \end{aligned}\right\} \tag{5.3}$$

One technique for solving (5.3) is the use of the trapezoidal algorithm

$$\left.\begin{aligned} \underline{y}_{n+1} &= \underline{y}_n + \frac{\Delta t}{2} A(\underline{y}_n + y_{n+1}) + \frac{\Delta t}{2}(\underline{f}_n + \underline{f}_{n+1}) \\ \underline{y}_0 &= \underline{c} \qquad \Delta t = T_0/M \end{aligned}\right\} \tag{5.4}$$

Equation (5.4) may be written in explicit form:

$$\left.\begin{aligned} \underline{y}_{n+1} &= \mathcal{A}\underline{y}_n + B(\underline{f}_n + \underline{f}_{n+1}) \\ \underline{y}_0 &= \underline{c} \end{aligned}\right\} \tag{5.5}$$

where

$$\left.\begin{aligned} \mathcal{A} &= [I - A\tfrac{\Delta t}{2}]^{-1}[I + A\tfrac{\Delta t}{2}] \\ \mathcal{B} &= [I - A\tfrac{\Delta t}{2}]^{-1}\tfrac{\Delta t}{2} \end{aligned}\right\} \tag{5.6}$$

Accuracy

Let $\underline{x}_{n+1} = \underline{x}(n\Delta t)$

Let $\underline{\tau}_{n+1}$ be the local truncation error defined by

$$\underline{x}_{n+1} = \mathcal{A}\underline{x}_n + \mathcal{B}(\underline{f}_{n+1} + \underline{f}_n) + \underline{\tau}_{n+1}\,\Delta t \tag{5.7}$$

Let $\underline{e}_n = (\underline{x}_n - \underline{y}_n)$ be the solution error (5.8)

Then subtracting (5.5) from (5.7)

$$\underline{e}_{n+1} = A\underline{e}_n + \underline{\tau}_{n+1}\,\Delta t \tag{5.9}$$

Now $\underline{e}_0 = \underline{x}_0 - \underline{y}_0 = 0$

$$\therefore \quad \left.\begin{aligned} \underline{e}_1 &= \underline{\tau}_1\,\Delta t \\ \underline{e}_2 &= (A\underline{\tau}_1 + \tau_2)\,\Delta t \\ &\vdots \\ \underline{e}_n &= (A^{n-1}\underline{\tau}_1 + A^{n-2}\underline{\tau}_2 + \cdots \underline{\tau}_n)\,\Delta t \end{aligned}\right\} \tag{5.10}$$

If the matrix A is simple, there exists a similarity matrix T such that

$$T^{-1}AT = \Lambda \qquad \text{and hence} \qquad A = T\Lambda T^{-1} \tag{5.11}$$

Hence if the homogeneous solutions of (5.3) are stable we know from theory that $\lambda(A)$ must either be pure imaginary or have negative real parts.

$$A = [I - A\,\frac{\Delta t}{2}]^{-1}[I + A\,\frac{\Delta t}{2}] = T[I - \Lambda\,\frac{\Delta t}{2}]^{-1}[I + \Lambda\,\frac{\Delta t}{2}]\,T^{-1} \tag{5.12}$$

$$\therefore \quad A = T\,\Theta\,T^{-1} \qquad \text{where} \qquad \Theta = [\theta_j]$$

$$\bar{\theta}_j = (1 - \lambda_j\,\frac{\Delta t}{2})^{-1}\,(1 + \lambda_j\,\frac{\Delta t}{2}) \tag{5.13}$$

If λ_j is pure imaginary, say $\lambda_j = i\omega_j$, then

$$|\theta_j| = \left[\frac{1 + (\omega_j\,\frac{\Delta t}{2})^2}{1 + (\omega_j\,\frac{\Delta t}{2})^2}\right]^{1/2} = 1 \tag{5.14}$$

If λ_j is complex with negative real part, say $\lambda_j = -\omega_j\zeta_j + i\omega_j\sqrt{1-\zeta_j^2}$ then

$$|\theta_j| = \left[\frac{1-\omega_j\zeta_j\Delta t + (\omega_j \frac{\Delta t}{2})^2}{1+\omega_j\zeta_j\Delta t + (\omega_j \frac{\Delta t}{2})^2}\right]^{1/2} \leq 1 \tag{5.15}$$

Hence the homogeneous solutions of (5.6) are Liapunov stable by Theorem 5.

In either case we have

$$\underline{e}_n = T[\Theta^{n-1}T^{-1}\underline{\tau}_1 + \Theta^{n-2}T^{-1}\underline{\tau}_2 + \cdots T^{-1}\underline{\tau}_n]\,\Delta t \tag{5.16}$$

Let $T^{-1}\underline{\tau}_i = \underline{b}_i$

$$\therefore \quad ||\underline{e}_n|| \leq ||T||\,[\,||\Theta^{n-1}\underline{b}_1|| + ||\Theta^{n-2}\underline{b}_2|| + \cdots ||\underline{b}_n||\,]\,\Delta t \tag{5.17}$$

But $$||\Theta^{n-k}\underline{b}_k|| = \sum_{j=1}^{N} |\theta_j|^{n-k}|b_j^k| \leq \sum_{j=1}^{N} |b_j^k| = ||\underline{b}_k|| \tag{5.18}$$

$$||\underline{b}_k|| \leq ||T^{-1}||\,||\underline{\tau}_k|| \tag{5.19}$$

$\therefore$ If $$\tau = \max_k ||\underline{\tau}_k|| \tag{5.20}$$

then $$||\underline{e}_n|| \leq ||T||\,||T^{-1}||\,n\Delta t\,\tau \tag{5.21}$$

but $$n\Delta t = t_n \leq T \tag{5.22}$$

$$\therefore \quad ||\underline{e}_n|| \leq T\,||T||\,||T^{-1}||\,\tau \leq K_1\tau \tag{5.23}$$

Now

$$\left.\begin{aligned}\underline{x}_{n+1} &= \underline{x}_n + \frac{d\underline{x}_n}{dt}\Delta t + \frac{d^2\underline{x}_n}{dt^2}\frac{\Delta t^2}{2} + \frac{d^3\underline{x}_s}{dt^3}\frac{\Delta t^3}{6} \\ \underline{f}_{n+1} &= \underline{f}_n + \frac{d\underline{f}_n}{dt}\Delta t + \frac{d^2\underline{f}_n}{dt^2}\frac{\Delta t^2}{2} + \frac{d^3\underline{f}_s}{dt^3}\frac{\Delta t^3}{6}\end{aligned}\right\} \quad (5.24)$$

Substituting into (5.7) and using (5.3)

$$\underline{\tau}_{n+1} = -[I - A\frac{\Delta t}{2}]^{-1}\left[\frac{1}{12}A^3\underline{x}_n + \frac{1}{2}A^2\underline{f}_n + \frac{1}{4}A\frac{d\underline{f}_n}{dt} + \frac{1}{2}\frac{d^2\underline{f}_n}{dt^2}\right]\Delta t^2 \quad (5.25)$$

Then, provided $\|\underline{x}_n\|$, $\|\underline{f}_n\|$, $\|\frac{d\underline{f}_n}{dt}\|$, $\|\frac{d^2\underline{f}_n}{dt^2}\|$ are bounded,

$$\|\underline{\tau}_{n+1}\| \le K_2\Delta t^2$$

$$\therefore \quad \tau = K_2\Delta t^2 \quad \text{as} \quad \Delta t \to 0 \quad (5.26)$$

$$\therefore \quad \|\underline{e}_n\| \le K_1K_2\,\Delta t^2 \quad \text{as} \quad \Delta t \to 0 \quad (5.27)$$

The trapezoidal scheme is second order accurate as $\Delta t \to 0$.

Application

Consider the conservative dynamical system:

$$\left.\begin{aligned} &M\ddot{\underline{x}} + K\underline{x} = 0 \\ &\underline{x}(0) = \underline{a},\ \dot{\underline{x}}(0) = b \\ &\text{where } M = M^T \text{ is positive definite} \\ &\qquad K = K^T \text{ is positive definite}\end{aligned}\right. \quad (5.28)$$

If $\underline{z} = \begin{pmatrix}\underline{x} \\ \dot{\underline{x}}\end{pmatrix}$, equation (5.28) may be written

$$\frac{dz}{dt} = A\underline{z}$$

$$\underline{z}(0) = \underline{c} = \binom{\underline{a}}{\underline{b}} \tag{5.29}$$

$$\text{where} \quad A = \begin{bmatrix} 0 & I \\ -M^{-1}K & 0 \end{bmatrix} \tag{5.30}$$

If A is simple, there exists a similarity matrix such that

$$T^{-1}AT = \quad = \begin{bmatrix} i\omega_1 & & & & & & \\ & -i\omega_1 & & & & & \\ & & i\omega_2 & & & & \\ & & & -i\omega_2 & & & \\ & & & & \cdot & & \\ & & & & & \cdot & \\ & & & & & & \cdot \end{bmatrix} \tag{5.31}$$

The trapezoidal difference equation corresponding to (5.29) is

$$\left.\begin{aligned} \underline{w}_{n+1} &= \underline{w}_n + \frac{\Delta t}{2} A(\underline{w}_{n+1} + \underline{w}_n) \\ \underline{w}_n &= \begin{pmatrix} \underline{y}_n \\ \dot{\underline{y}}_n \end{pmatrix} \qquad \underline{w}_o = \underline{c} \end{aligned}\right\} \tag{5.32}$$

Alternatively,

$$\left.\begin{aligned} \underline{y}_{n+1} &= \underline{y}_n + \frac{\Delta t}{2} (\dot{\underline{y}}_{n+1} + \dot{\underline{y}}_n) \\ \dot{\underline{y}}_{n+1} &= \dot{\underline{y}}_n - \frac{t}{2} M^{-1}K(\underline{y}_{n+1} + \underline{y}_n) \end{aligned}\right\} \tag{5.33}$$

From (5.33) we see that

$$\frac{1}{2}\, \dot{\underline{y}}_{n+1}^T\, M\, \dot{\underline{y}}_{n+1} + \frac{1}{2}\, \underline{y}_{n+1}^T\, K\, \underline{y}_{n+1} = \text{constant} \tag{5.34}$$

Thus the difference equations (5.32) or (5.33) conserve energy in exactly the same way as equation (5.28) whose first integral is

$$\frac{1}{2}\,\dot{\underline{x}}^T M \dot{\underline{x}} + \frac{1}{2}\,\underline{x}^T M \underline{x} = \text{constant} \tag{5.35}$$

From equation (5.32)

$$\begin{aligned} \underline{w}_{n+1} &= \mathcal{A}\,\underline{w}_n \\ \underline{w}_0 &= \underline{c} \end{aligned} \tag{5.36}$$

where $\mathcal{A} = [I - \frac{\Delta t}{2} A]^{-1}\ [I + \frac{\Delta t}{2} A]$ (5.37)

Using (5.31)

$$\mathcal{A} = T[I - \frac{\Delta t}{2} \Lambda]^{-1}\ [I + \frac{\Delta t}{2} \Lambda]\ T^{-1} \tag{5.38}$$

$$\therefore \quad \mathcal{A} = T\Theta T^{-1} \quad ; \quad \Theta = [\theta_j]$$

$$\theta_j = \frac{1 + i\frac{\Delta t}{2}\omega_j}{1 - i\frac{\Delta t}{2}\omega_j} \qquad \forall j \quad (1,2N) \tag{5.39}$$

Hence $|\lambda_j(\mathcal{A})| = 1 \qquad \forall j$ (5.40)

Thus the eigenvalues of $\mathcal{A}$ all have modulus equal to unity and equation (5.36) is Liapunov stable. This property is exemplified in the fact that the energy is conserved.

Using equation (3.5), the solution of equation (5.36) is:

$$\underline{w}_n = \begin{pmatrix} \underline{y}_n \\ \dot{y}_n \end{pmatrix} = \mathcal{A}^n\,\underline{c} \tag{5.41}$$

where $\mathcal{A}^n = T\,\Theta^n\,T^{-1}$ (5.42)

Now

$$\left.\begin{aligned} \theta_j &= \frac{1 + i\,\frac{\Delta t}{2}\,\omega_j}{1 - i\,\frac{\Delta t}{2}\,\omega_j} = e^{i\Delta\phi_j} \\ \text{where}\quad \Delta\phi_j &= \tan^{-1} \frac{\omega_j \Delta t}{1 - (\omega_j \frac{\Delta t}{2})^2} \end{aligned}\right\} \tag{5.43}$$

$$\text{Let}\quad \Omega_j = \frac{\Delta\phi_j}{\Delta t} \tag{5.44}$$

$$\therefore \quad \underline{w}_n = \begin{pmatrix} \underline{y}_n \\ \dot{\underline{y}}_n \end{pmatrix} = T \begin{bmatrix} e^{i\Omega_j t_n} \end{bmatrix} T^{-1} \tag{5.45}$$

The solution of equation (5.29) at $t_n = n\Delta t$ is:

$$\underline{z}_n = \begin{pmatrix} \underline{x}_n \\ \dot{\underline{x}}_n \end{pmatrix} = T \begin{bmatrix} e^{i\omega_j t_n} \end{bmatrix} T^{-1} \tag{5.46}$$

We see that the solution of the differential equation (5.29) and the corresponding difference equation (5.36) have the same structure, however, in general the time dependence is different.

<u>Period Error</u>

Let $T_j^d = \frac{2\pi}{\Omega_j}$ be the period of the j^{th} mode of the difference equation. Let $T_j = 2\pi/\omega_j$ be the period of the j^{th} mode of the differential equation. Then

$$\varepsilon_T = \frac{T_j^d - T_j}{T_j} \tag{5.47}$$

is the period error of the j^{th} mode of the difference equation.

We note that

$$\lim_{\Delta t \to 0} \Omega_j = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \tan^{-1} \frac{\omega_j \Delta t}{1 - (\omega_j \frac{\Delta t}{2})^2} \equiv \omega_j \qquad (5.48)$$

Thus in the limit as $\Delta t \to 0$, the period error vanishes and equations (5.45) and (5.46) are identical.

From a practical computing standpoint, we cannot let Δt go to zero. While $(\omega_j \Delta t)$ can be made acceptably small for the lower modes of a complex structure, it is not possible to make $(\omega_j \Delta t)$ small for the highest modes. Thus by making Δt sufficiently small, equation (5.45) will give an accurate representation of the low mode behavior, however, higher mode behavior will not be accurately modeled. In most problems in structural dynamics, only low mode behavior is of real significance, therefore if high mode behavior can somehow be suppressed, equation (5.45) will give a reasonably accurate representation of the response of a complex structure.

Methods Proposed for Suppressing the Higher Modes

(i) Use of Viscous Damping

By analogy with continuous time systems it might appear that the use damping could be used to suppress the higher modes. As we shall now show, the method is ineffectual in suppressing the higher modes in discrete systems. If in equation (5.28) we add viscous damping, then the equation becomes

$$\left.\begin{aligned} &M\,\underline{\ddot{x}} + C\,\underline{\dot{x}} + K\,\underline{x} = 0 \\ &\underline{x}(0) = \underline{a} \qquad \underline{\dot{x}}(0) = \underline{b} \\ &M = M^T > 0\,; \quad C = C^T \geq 0\,; \quad K = K^T > 0 \end{aligned}\right\} \qquad (5.49)$$

If $Z = \begin{pmatrix} \underline{x} \\ - \\ \underline{\dot{x}} \end{pmatrix}$, equation (5.49) may be rewritten

$$\left.\begin{aligned} &\frac{dz}{dt} = A\,\underline{z} \\ &\underline{z}(0) = \underline{c} = \frac{a}{b} \\ &\text{where} \quad A = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{bmatrix} \end{aligned}\right\} \tag{5.50}$$

If A is simple, there exists a similarity matrix T such that

$$T^{-1}AT = \Lambda = \begin{bmatrix} -\omega_1\zeta_1 + i\omega_1\sqrt{1-\zeta_1^2} & & & \\ & -w_1\zeta_1 - iw_n\sqrt{1-\zeta_1^2} & & \\ & & -w_n\zeta_n + iw_n\sqrt{1-\zeta_n^2} & \\ & & & -w_n\zeta_n - iw_n\sqrt{1-\zeta_n^2} \end{bmatrix} \tag{5.51}$$

The trapezoidal difference equation corresponding to (5.50) is:

$$\left.\begin{aligned} &\underline{w}_{n+1} = \begin{pmatrix} \underline{y}_{n+1} \\ \dot{\underline{y}}_n \end{pmatrix} = A\,\underline{w}_n \\ &\underline{w}_o = \underline{c} \end{aligned}\right\} \tag{5.52}$$

where

$$\left.\begin{aligned} &A = T\Theta T^{-1} \\ &\theta_j = \frac{1 - \omega_j\zeta_j\frac{\Delta t}{2} + i\omega_j\frac{\Delta t}{2}\ \ 1-\zeta_j^2}{1 + \omega_j\zeta_j\frac{\Delta t}{2} - i\omega_j\frac{\Delta t}{2}\ \ 1-\zeta_j^2} \end{aligned}\right\} \tag{5.53}$$

θ_j may be expressed as

$$\theta_j = \rho_j \, e^{i\Delta\phi_j}$$

$$\rho_j = \left\{ \frac{1 - \omega_j\zeta_j\Delta t + (\omega_j \frac{\Delta t}{2})^2}{1 + \omega_j\zeta_j\Delta t + (\omega_j \frac{\Delta t}{2})^2} \right\} \tag{5.54}$$

$$\Delta\phi_j = \tan^{-1}\left[\frac{\omega_j\Delta t \sqrt{1-\zeta_j^2}}{1 - (\frac{\omega_j\Delta t}{2})^2} \right]$$

For the lower modes $\rho_j \simeq 1$. As mode order increases, ρ_j decreases initially, then starts to increase again. For the higher modes ρ_j tends to unity. Thus we see that viscous damping is not effective in suppressing the higher modes.

(ii) Use of Algorithmic Damping

If equation (5.32) is modified to read

$$\underline{w}_{n+1} = \underline{w}_n + \Delta t A((1-\alpha)\underline{w}_{n+1} + \alpha\underline{w}_n)$$

$$\underline{w}_n = \begin{pmatrix} \underline{y}_n \\ \dot{\underline{y}}_n \end{pmatrix} \qquad \underline{w}_o = \underline{c} \qquad 0 < \alpha < 1$$

Equation (5.36) now becomes

$$\underline{w}_{n+1} = \mathcal{A}_\alpha \, \underline{w}_n \tag{5.55}$$

where

$$\mathcal{A}_\alpha = [I - (1-\alpha)\Delta t A]^{-1} \, [I + \alpha\Delta t A] \tag{5.56}$$

$$\left.\begin{aligned} \therefore \quad & \mathcal{A}_\alpha = T \, \Theta_\alpha T^{-1} \\ \text{where} \quad & \theta_{\alpha j} = \frac{1 + \Delta t \, i\omega_j}{1 - \Delta t(1-\alpha)\,\omega_j} \end{aligned}\right\} \tag{5.57}$$

Hence

$$\left.\begin{aligned} \theta_{\alpha_j} &= \rho_{\alpha_j} e^{i\Delta\phi\alpha_j} \\ \rho_{\alpha_j} &= \left\{\frac{1 + (\alpha\Delta t\omega_j)^2}{1 + (1-\alpha)\Delta t\omega_j)^2}\right\} \\ \Delta\phi_{\alpha_j} &= \tan^{-1} \frac{\omega_j\Delta t}{1 - \alpha(1-\alpha)(\omega_j\Delta t)^2} \end{aligned}\right\} \tag{5.58}$$

We note that

i) $\Delta\phi_{\alpha_j} = \Delta\phi_j \qquad \rho_j = 1 \qquad \text{when} \quad \alpha = 1/2$

ii) $0 < \alpha < 1/2 \qquad \rho_j \leq 1$

Thus in case ii) $\rho_j \simeq 1$ for the lower modes, while $\rho_j \simeq \frac{\alpha}{1-\alpha} \leq 1$ for the higher modes. Unfortunately, if $\alpha \neq 1/2$, it is easily shown that

$$||\underline{e}_n|| \leq c_3\left|\frac{1}{2} - \alpha\right| O(\Delta t) + K_1K_2\, O(\Delta t^2) \tag{5.59}$$

Thus, if $\alpha \neq \frac{1}{2}$, the modified trapezoidal altorithm (5.54) is only of first order accuracy as $\Delta t \to 0$.

(iii) Use of Temporal Filtering

Let $\quad \underline{v}_n = \frac{1}{4}[\underline{w}_{n+1} + 2\underline{w}_n + \underline{w}_{n-1}] \qquad (5.60)$

where $\quad \underline{w}_{n+1} = \mathcal{A}\underline{w}_n \qquad (5.61)$

$$\therefore \quad \underline{v}_n = \frac{1}{4}[\mathcal{A} + 2I + \mathcal{A}^{-1}]\underline{w}_n \tag{5.62}$$

Now, for the trapezoidal algorithm

$$\mathcal{A} = T \Theta T^{-1} \tag{5.63}$$

where $$\theta_j = \frac{1 + i\omega_j \frac{\Delta t}{2}}{1 - i\omega_j \frac{\Delta t}{2}} \tag{5.64}$$

$$\therefore \quad \underline{v}_n = \frac{1}{4} T[\Theta + 2I + \Theta^{-1}]T^{-1} \underline{w}_n \tag{5.65}$$

$$= \frac{1}{4} T \left[\frac{1 + i\omega_j \frac{\Delta t}{2}}{1 - i\omega_j \frac{\Delta t}{2}} + 2 + \frac{1 - i\omega_j \frac{\Delta t}{2}}{1 + i\omega_j \frac{\Delta t}{2}} \right] T^{-1} \underline{w}_n \tag{5.66}$$

$$\therefore \quad \underline{v}_n = T \left[\frac{1}{1 + (\omega_j \frac{\Delta t}{2})^2} \right] T^{-1} \underline{w}_n \tag{5.67}$$

We note that $\frac{1}{1 + (\omega_j \frac{\Delta t}{2})^2} \simeq 1$ for the low modes and tends to zero for the high modes. Since $\underline{w}_n$ is second order accurate as $\Delta t \to 0$, the filter, which is also second order accurate, still retains second order accuracy. Thus, unlike algorithmic damping, the use of the temporal filter does not affect the accuracy of the computational scheme.

There exist many more sophisticated algorithms for solving problems such as equation (5.28), however, the author's experience has been that for linear problems, the trapezoidal algorithm with post-filtering does as good a job as the more sophisticated schemes when applied to large complex structures.

(c) Numerical Solution of Nonlinear Ordinary Differential Equations

Consider the system of nonlinear differential equations

$$\left.\begin{array}{ll} \frac{d\underline{x}}{dt} = A\,\underline{x} + \underline{f}(\underline{x}) + \underline{g}(t) & 0 \leq t \leq T < \infty \\ \underline{x}(0) = \underline{c} & A = \text{constant matrix} \end{array}\right\} \tag{5.68}$$

One technique for solving (5.68) is the use of the trapezoidal algorithm

$$\left.\begin{aligned} \underline{y}_{n+1} &= \underline{y}_n + \frac{\Delta t}{2}[A(\underline{y}_{n+1} + \underline{y}_n) + \underline{f}(\underline{y}_{n+1}) + f(\underline{y}_n) + \underline{g}_{n+1} + \underline{g}_n] \\ \underline{y}_0 &= \underline{c} \qquad \Delta t = T/M \end{aligned}\right\} \quad (5.69)$$

Equation (5.69) may also be written as

$$\left.\begin{aligned} \underline{y}_{n+1} &= \mathcal{A}\underline{y}_n + \mathcal{B}(\underline{f}(\underline{y}_{n+1}) + f(\underline{y}_n) + \underline{g}_{n+1} + \underline{g}_n) \\ \underline{y}_0 &= \underline{c} \end{aligned}\right\} \quad (5.70)$$

<u>Accuracy</u>

Let $\underline{x}_{n+1} = \underline{x}(n\Delta t)$

Let $\underline{\tau}_{n+1}$ be the local truncation error defined by

$$\underline{x}_{n+1} = \underline{x}_n + \frac{\Delta t}{2}[A(\underline{x}_{n+1} + \underline{x}_n) + \underline{f}(\underline{x}_{n+1}) + \underline{f}(\underline{x}_n) + g_{n+1} + g_n + 2\tau_{n+1}] \quad (5.71)$$

Let $\underline{e}_n = \underline{x}_n - \underline{y}_n$ be the solution error (5.72)

Then subtracting (5.69) from (5.71)

$$\underline{e}_{n+1} = (I + A\frac{\Delta t}{2})\underline{e}_n + A\frac{\Delta t}{2}\underline{e}_{n+1} + (\underline{fx}_{n+1}) - \underline{f}(\underline{y}_{n+1}))\frac{\Delta t}{2}$$

$$+ (\underline{f}(\underline{x}_n) - \underline{f}(\underline{y}_n)\frac{\Delta t}{2} + \tau_{n+1}\Delta t \quad (5.73)$$

If
i) $\underline{x}_n, \underline{y}_n$ are bounded $\forall n \in (1,M)$

ii) $||\underline{f}(\underline{x}) - \underline{f}(\underline{y})|| \leq K||\underline{x} - \underline{y}|| \quad \forall \underline{x}_n, \underline{y}_n$ bounded

iii) $\underline{f}(\underline{x})$ continuous and continuous first and second partials

(5.74)

Then

$$\|\underline{e}_{n+1}\| \le (1 + \|A\| \frac{\Delta t}{2}) \|\underline{e}_n\| + \|A\| \frac{\Delta t}{2} \|\underline{e}_{n+1}\|$$

$$+ K \frac{\Delta t}{2} \|\underline{e}_{n+1}\| + k \frac{\Delta t}{2} \|\underline{e}_n\| + \|\tau_{n+1}\| \Delta t \tag{5.75}$$

$$\therefore \quad \|\underline{e}_{n+1}\| \le \frac{(1 + (\|A\| + K) \frac{\Delta t}{2}}{(1 - (\; A \; + K) \frac{t}{2})} \|\underline{e}_n\| + \frac{\|\underline{\tau}_{n+1}\| \Delta t}{1 - (\|A\| + K) \frac{\Delta t}{2}}$$

$$\|\underline{e}_0\| = 0$$

Thus $\|\underline{e}_1\| \le \dfrac{\|\underline{\tau}_1\| \; \Delta t}{1 - (\|A\| + K) \frac{\Delta t}{2}}$

$$\|e_2\| \le \left[\left[\frac{1 + (\|A\| + K) \frac{\Delta t}{2}}{1 - (\|A\| + K) \frac{\Delta t}{2}} \right] \|\tau_1\| + \|\tau_2\| \right] \frac{\Delta t}{1 - (\|A\| + K) \frac{\Delta t}{2}} \tag{5.76}$$

$$\therefore \quad \|\underline{e}_n\| \le [\psi^{n-1} \|\underline{\tau}_1\| + \psi^{n-2} \|\tau_2\| + \cdots \|\tau_n\|] \frac{\Delta t}{1 - (\|A\| + K) \frac{\Delta t}{2}}$$

where

$$\psi = \frac{1 + (\|A\| + K) \frac{\Delta t}{2}}{1 - (\|A\| + K) \frac{\Delta t}{2}} \tag{5.77}$$

If $\quad \tau = \underset{k}{\text{Max}} \|\underline{\tau}_2\|$

$$\|\underline{e}_n\| \le \frac{[\psi^n - 1]}{[\psi - 1]} \tau \frac{\Delta t}{1 - (\|A\| + K) \frac{\Delta t}{2}} \tag{5.78}$$

But $\quad \psi - 1 = \dfrac{1 + (\|A\| + K) \frac{\Delta t}{2}}{1 - (\|A\| + K) \frac{\Delta t}{2}} - 1 = \dfrac{\Delta t(\|A\| + K)}{1 - (\|A\| + K) \frac{\Delta t}{2}} \qquad (5.79)$

$$\therefore \qquad ||\underline{e}_n|| \quad \frac{\tau}{||A|| + K} \; [\psi^n - 1] \tag{5.80}$$

Now $$e^y = 1 + y + \frac{y^2}{2} e^{\theta y} \qquad 0 < \theta < 1 \tag{5.81}$$

$$\therefore \qquad (1 + y) \leq e^y \tag{5.82}$$

$$1 - y \geq 1 - 2y + \frac{1}{2}(2y)^2 \qquad \text{for} \quad y \leq \frac{1}{2} \tag{5.83}$$

$$e^{-z} = 1 - z + \frac{z^2}{2} = \frac{z^2}{6} e^{-\theta z} \tag{5.84}$$

$$\therefore \qquad \frac{1}{1 - z + \frac{z^2}{2}} = \frac{1}{e^{-z} + \frac{z^3}{6} e^{-\theta z}} \leq e^z \tag{5.85}$$

$$\therefore \qquad \frac{1}{1-y} \leq \frac{1}{1 - 2y + 2y^2} \leq e^{2y} \tag{5.86}$$

$$\therefore \qquad \frac{1+y}{1-y} \leq e^{3y} \tag{5.87}$$

Hence, if $(||A|| + K)\frac{\Delta t}{2} < \frac{1}{2}$

$$\psi = \frac{1 + (||A|| + K)\frac{\Delta t}{2}}{1 = (||A|| + K)\frac{\Delta t}{2}} \leq e^{\frac{3}{2}(||A|| + K)\Delta t} \tag{5.88}$$

$$\therefore \qquad ||\underline{e}_n|| \leq \frac{\tau}{||A|| + K} \, e^{\frac{3}{2}(||A|| + K)n\Delta t} \tag{5.89}$$

But $n\Delta t = t_n \leq T$

$$\therefore \qquad ||\underline{e}_n|| \leq \frac{\tau}{||A|| + K} \, e^{\frac{3}{2}(||A|| + K)T} \tag{5.90}$$

Returning to equation (5.71),

$$\left.\begin{aligned} \underline{x}_{n+1} &= \underline{x}_n + \frac{dx_n}{dt}\Delta t + \frac{d^2\underline{x}_s}{dt^2}\frac{\Delta t^2}{2} \\ \underline{f}(\underline{x}_{n+1}) &= \underline{f}(\underline{x}_n) + \underline{J}(\underline{x}_n)\frac{dx_n}{dt}\Delta t + \cdots \\ \underline{g}_{n+1} &= \underline{g}_n + \frac{dg_n}{dt}\Delta t + \text{etc.} \end{aligned}\right\} \quad (5.91)$$

Then using (5.71) and (5.68) it may be shown that

$$\|\underline{\tau}_{n+1}\| \leq K_3(\Delta t)^2 \quad \text{as} \quad \Delta t \to 0 \quad (5.92)$$

$$\left.\begin{aligned} \therefore \quad \|\underline{e}_n\| &\leq \frac{K_3(\Delta t)^2\, e^{\frac{3}{2}(\|A\| + K)T}}{\|A\| + K} \\ &\leq K_4(\Delta t)^2 \quad \text{as} \quad \Delta t \to 0 \end{aligned}\right\} \quad (5.93)$$

Thus the trapezoidal scheme is second order accurate; unfortunately, unlike the situation for linear systems, the trapezoidal difference equations for nonlinear differential equations are not guaranteed to be globally stable.

Application

Consider the conservative dynamical system

$$\left.\begin{aligned} &M\,\ddot{\underline{x}} + K\,\underline{x} + \underline{f}(\underline{x}) = 0 \\ &\underline{x}(0) = \underline{a} \qquad \dot{\underline{x}}(0) = \underline{b} \\ &M = M^T > 0, \quad K = K^T > 0 \\ &\underline{f}(\underline{x}) = \nabla U(\underline{x}) \qquad U(\underline{x}) > 0 \qquad \underline{x} \neq 0 \end{aligned}\right\} \quad (5.94)$$

Equation (5.94) has the first integral

$$\frac{1}{2}\,\dot{\underline{x}}^T M\,\dot{\underline{x}} + \frac{1}{2}\,\underline{x}^T K\,\underline{x} + U(\underline{x}) = \text{const.} \tag{5.95}$$

If $\underline{z} = \begin{pmatrix} \underline{x} \\ \dot{\underline{x}} \end{pmatrix}$ equation (5.94) may be written

$$\left.\begin{aligned} &\frac{d\underline{z}}{dt} = A\,\underline{z} + \underline{g}(\underline{z}) \\ &\underline{z}(0) = \underline{c} \\ \text{where}\quad &A = \begin{bmatrix} 0 & I \\ -M^{-1}K & 0 \end{bmatrix} \qquad \underline{g}(\underline{z}) = \begin{pmatrix} 0 \\ -M^{-1}\underline{f}(\underline{x}) \end{pmatrix} \end{aligned}\right\} \tag{5.96}$$

If A is simple, it has the representation

$$A = T\Lambda T^{-1} \tag{5.97}$$

$$\Lambda = \begin{bmatrix} i\omega_1 & & & & \\ & -i\omega_1 & & & \\ & & i\omega_2 & & \\ & & & -i\omega_2 & \\ & & & & \text{etc.} \end{bmatrix} \tag{5.98}$$

The trapezoidal difference equation corresponding to (5.94) is

$$\left.\begin{aligned} &\underline{w}_{n+1} = \underline{w}_n + \frac{\Delta t}{2} A(\underline{w}_{n+1} + \underline{w}_n) + \frac{\Delta t}{2}\left(\underline{g}(\underline{w}_{n+1}) + \underline{g}(\underline{w}_n)\right) \\ &\underline{w}_n = \begin{matrix} \underline{y}_n \\ \hline \dot{\underline{y}}_n \end{matrix} \qquad \underline{w}_0 = \underline{c} \end{aligned}\right\} \tag{5.99}$$

Alternatively,

$$\left.\begin{aligned} &\underline{y}_{n+1} = \underline{y}_n + \frac{\Delta t}{2}\left(\dot{\underline{y}}_{n+1} + \dot{\underline{y}}_n\right) \\ &\dot{\underline{y}}_{n+1} = \dot{\underline{y}}_n - \frac{\Delta t}{2} M^{-1}\left[K(\underline{y}_{n+1} + \underline{y}_n) + \underline{f}(\underline{y}_{n+1}) + \underline{f}(y_n)\right] \\ &\underline{y}_0 = \underline{a} \qquad \dot{\underline{y}}_0 = \underline{b} \end{aligned}\right\} \tag{5.100}$$

From (5.100) we see that

$$\frac{1}{2}[\dot{\underline{y}}_{n+1}^T M \dot{\underline{y}}_{n+1} - \dot{\underline{y}}_n^T M \dot{\underline{y}}_n] + \frac{1}{2}[\underline{y}_{n+1}^T K \underline{y}_{n+1} - \underline{y}_n^T K \underline{y}_n^T]$$

$$+ \frac{1}{2}(\underline{y}_{n+1} - \underline{y}_n)^T (\underline{f}(y_{n+1}) + \underline{f}(\underline{y}_n)) \equiv 0 \qquad (5.101)$$

which can be written

$$\frac{1}{2}\dot{\underline{y}}_{n+1}^T M \dot{\underline{y}}_{n+1} + \frac{1}{2}\underline{y}_{n+1}^T K \underline{y}_{n+1} + \tilde{U}_{n+1} = \text{constant} \qquad (5.102)$$

where $\tilde{U}_{n+1} = \sum_{i=0}^{n+1} \frac{1}{2}(\underline{y}_{i+1} - \underline{y}_i)^T (\underline{f}(\underline{y}_{i+1}) + \underline{f}(\underline{y}_i)) + U(\underline{y}_o)$ (5.103a)

We note that if $\underline{a},\underline{b}$ are bounded, then as $\Delta t \to 0$, $(\underline{y}_{i+1} - \underline{y}_i) \to 0$ and that (5.103) becomes

$$\lim_{\Delta t \to 0} \tilde{U}_{n+1} = \lim_{\Delta t \to 0} \sum_{i=0}^{n+1} \frac{1}{2}(\underline{y}_{i+1} - \underline{y}_i)^T (\underline{f}(\underline{y}_{i+1}) + \underline{f}(\underline{y}_i)] \qquad (5.103b)$$

$$+ U(y_o) = U(\underline{y}_{n+1})$$

In this case, equations (5.102) and (5.95) are identical. In general, for Δt finite, $\tilde{U}_{n+1}$ cannot be guaranteed to be positive, in which case equations (5.99) and 5.100) are not guaranteed to be stable. It should be pointed out, that equation (5.99) can be rewritten as

$$\left.\begin{aligned} &\underline{w}_{n+1} = \mathcal{A}\underline{w}_n + \mathcal{B}(\underline{g}(w_{n+1}) + \underline{g}(\underline{w}_n)) \\ &\underline{w}_o = \underline{c} \\ \text{where } &\mathcal{A} = [I = \tfrac{\Delta t}{2} A]^{-1} [I + \tfrac{\Delta t}{2} A] \\ &\mathcal{B} = [I - \tfrac{\Delta t}{2} A]^{-1} \tfrac{\Delta t}{2} \end{aligned}\right\} \qquad (5.104)$$

Since the eigenvalues of A are pure imaginary, the eigenvalues of $\mathcal{A}$ all have modulus unity, thus equation (5.104) is one of the "critical cases" in Liapunov stability theory as discussed in Theorem 16.

To illustrate these problems let us consider the following scalar problem:

$$\left.\begin{aligned} &\ddot{x} + f(x) = 0 \\ &f(x) = x \qquad |x| \le 1 \\ &f(x) = \operatorname{sgn} x + \mu(x - \operatorname{sgn} x) \qquad |x| > 1 \end{aligned}\right\} \tag{5.105}$$

Equation (5.105) has the following first integral

$$\left.\begin{aligned} &\frac{1}{2}\dot{x}^2 + F(x) = \text{const.} \\ &F(x) = \frac{1}{2}x^2 \qquad |x| \le 1 \\ &\qquad = \frac{1}{2} + |x| + \frac{\mu}{2}(|x| - 1)^2 \; ; \; |x| > 1 \end{aligned}\right\} \tag{5.106}$$

Since $F(x) > 0 \quad x \neq 0$, equation (5.105) is globally Liapunov stable with respect to the origin.

The trapezoidal difference equation corresponding to equation (5.105) is:

$$\left.\begin{aligned} &y_{n+1} - y_n = \eta(\dot{y}_{n+1} + \dot{y}_n) \quad ; \qquad \eta = \Delta t/2 \\ &\dot{y}_{n+1} - \dot{y}_n = -\eta(f(y_{n+1}) + f(y_n)) \end{aligned}\right\} \tag{5.107}$$

which may be written

$$\eta^2 f(y_{n+1}) + y_{n+1} = z_n = y_n + 2\eta\dot{y}_n - \eta^2 f(y_n) \tag{5.108}$$

Since f(y) is piece-wise linear, equation (5.108) may be inverted to give y_{n+1} in terms of z_n.

Thus
$$\left.\begin{aligned} y_{n+1} &= g(z_n) \\ \dot{y}_{n+1} &= -\frac{y_n + \eta\,\dot{y}_n + g(z_n)}{\eta} \end{aligned}\right\} \tag{5.109}$$

where
$$\left.\begin{aligned} g(z) &= \frac{1}{1+\eta^2}\,z \qquad \text{for} \quad |z| \le 1+\eta^2 \\ &= \operatorname{sgn} z + \frac{1}{1+\mu\eta^2}\left(z - (1+\eta^2)\operatorname{sgn} z\right) \quad \text{for} \quad |z| > 1+\eta^2 \end{aligned}\right\} \tag{5.110}$$

$$|g(z)| \le |z| \qquad \forall z \tag{5.111}$$

Let
$$z_{n+1} = y_{n+1} + 2\eta\,\dot{y}_{n+1} - \eta^2 f(y_{n+1}) \tag{5.112}$$

$$\therefore \qquad z_{n+1} = 4g(z_n) - 2z_n = z_{n-1} \tag{5.113}$$

Equation (5.113) may be written in several different forms, two of which are given below:

a)
$$\left.\begin{aligned} z_{n+1} - 2z_n + z_{n-1} &= -\frac{4\eta^2}{1+\eta^2}\,k_3(z_n) \\ k_3(z_n) &= z_n \qquad \text{for} \quad |z_n| \le (1+\eta^2) \\ &= (1+\eta^2)\operatorname{sgn} z_n + \frac{\mu(1+\eta^2)}{1+\mu\eta^2}\left(z_n - (1+\eta^2)\operatorname{sgn} z_n\right) \\ &\qquad \text{for} \quad |z_n| > 1+\eta^2 \end{aligned}\right\} \tag{5.114}$$

From (5.114) we see that if $\mu \ge 0$, $\operatorname{sgn}(z_{n+1} - 2z_n + z_{n-1}) = -\operatorname{sgn} z_n$ (5.115)

Thus the sign of the finite difference curvature is always opposite to that of the displacement, thus the solutions are always oscillatory.

b)
$$\left.\begin{aligned} z_{n+1} - \frac{2(1-\mu\eta^2)}{(1+\mu\eta^2)} z_n + z_{n-1} &= - \frac{4\eta^2(1-\mu)}{(1+\mu\eta^2)} k_2(z_n) \\ k_2(z) &= \frac{1}{1+\eta^2} z \qquad \text{for } |z| \leq 1+\eta^2 \\ &= \operatorname{sgn} z \qquad \text{for } |z| > 1+\eta^2 \end{aligned}\right\} \tag{5.116}$$

Equations can also be written in the first order form,

$$\left.\begin{aligned} \underline{\theta}_{n+1} &= \mathcal{A}\underline{\theta}_n + \underline{k}(\underline{\theta}_n) \\ \underline{\theta}_n &= \begin{pmatrix} z_n \\ z_{n-1} \end{pmatrix} \qquad \mathcal{A} = \begin{bmatrix} \dfrac{2(1-\mu\eta^2)}{1+\mu\eta^2} & -1 \\ 1 & 0 \end{bmatrix} \\ \underline{k}(\theta_n) &= \begin{pmatrix} -\dfrac{4\eta^2(1-\mu)}{1+\mu\eta^2} k_2(z_n) \\ 0 \end{pmatrix} \end{aligned}\right\} \tag{5.117}$$

The matrix $\mathcal{A}$ has eigenvalues

$$\begin{aligned} \lambda &= - \frac{1-\mu\eta^2}{1+\mu\eta^2} \pm i \sqrt{1 - \left(\frac{1-\mu\eta^2}{1+\mu\eta^2}\right)^2} \\ &= e^{\pm i\phi} \\ \cos\phi &= \left(\frac{1-\mu\eta^2}{1+\mu\eta^2}\right) \end{aligned} \tag{5.118}$$

Using equation (3.5) with $\mu \neq 0$

$$\underline{\theta}_{n+1} = A^{n+1}\,\underline{\theta}_0 + \sum_{i=0}^{n} A^{n-i}\,\underline{k}(\theta_i) \tag{5.119}$$

But $\mathcal{A} = T\Lambda T^{-1} \quad \therefore \mathcal{A}^k = T\Lambda^k T^{-1}$ (5.120)

where $|\lambda_i| = 1 \quad i = 1,2$

$$\therefore \quad \|\underline{\theta}_{n+1}\| \le \|T\|\,\|T^{-1}\|\,\|\underline{\theta}_0\| + \|T\|\,\|T^{-1}\|\sum_{i=0}^{n}\|k(\underline{\theta}_i)\| \tag{5.121}$$

But $\|\underline{k}(\theta_i)\| \le 4n^2\left|\frac{1-\mu}{1+\mu n^2}\right|$ (5.122)

$$\therefore \quad \|\underline{\theta}_{n+1}\| \le K(\|\underline{\theta}_0\| + 4n^2\left|\frac{1-\mu}{1+\mu n^2}\right| n) \tag{5.123}$$

Thus, even though (5.117) may be unstable, it is only weakly unstable, with at most linear divergence. Now using (3.5) with $\mu = 0$, in this case $\lambda_i = 1, \; i = 1,2$

$$A = T\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} T^{-1} \tag{5.124}$$

$$\therefore \quad \|\underline{\theta}_{n+1}\| \le \|T\|\,\|T^{-1}\|\,\|\underline{\theta}_0\|\,(2+n) + 4n^2\,|T|\,\|T^{-1}\|\sum_{i=0}^{n}(2+i)$$

$$\le K[\,\|\underline{\theta}_0\|\,(2+n) + 4n^2(2(n+1) + \frac{n(n+1)}{2})] \tag{5.125}$$

Thus, even in this case, the $\|\underline{\theta}_n\| \sim O(n^2)$ as $n \to \infty$, as it is only weakly unstable.

Equation (5.117) defines a continuous mapping $M(\cdot)$ such that:

$$\underline{\theta}_{n+1} = M(\underline{\theta}_n) \tag{5.126}$$

Therefore, by the Brower Fixed-Point theorem, there exists at least one fixed point, or equilibrium solution. In the case of equation (5.117), it is easily seen that the only solution of

$$\underline{\theta}^* = \underline{M}(\underline{\theta}^*) \quad \text{is} \quad \underline{\theta}^* = 0 \tag{5.127}$$

Next, let M^k, k an integer, denote the mapping M applied k times. Then there may exist a sequence of distinct points $\underline{\theta}^*(1), \underline{\theta}^*(2), \cdots, \underline{\theta}^*(k)$ such that

$$\left.\begin{aligned} \underline{\theta}^*(m+1) &= \underline{M}^m(\underline{\theta}^*(1)), \qquad m = 1,2,\cdots,(k-1) \\ \underline{\theta}^*(1) &= \underline{M}^k(\theta^*(1)) \end{aligned}\right\} \tag{5.128}$$

Clearly, this sequence constitutes a periodic solution of period k.

Stability of Periodic Solutions

Let $\underline{\theta}^*(m)$; m $\in$ (1,k-1) be a k periodic solution of equation (5.117).

Let $$\underline{\theta}(m) = \underline{\theta}^*(m) + \delta\underline{\theta}(m) \tag{5.129}$$

Then $$\delta\underline{\theta}(m+1) = \underline{M}_{,\underline{\theta}}(\underline{\theta}^*(m))\, \delta\underline{\theta}(m) \tag{5.130}$$

Provided $\underline{\theta}(m)$ and $\underline{\theta}^*(m)$ are on the same piecewise linear branches of $\underline{k}(\underline{\theta})$. Thus, for $\delta\underline{\theta}(k)$ small, but not infinitesimal, and $\underline{\theta}^*(m)$ not on a corner of $\underline{k}(\underline{\theta})$,

$$\therefore \quad \delta\underline{\theta}(m+1) = [\sum_{i=1}^{m} \underline{M}_{,\underline{\theta}}(\underline{\theta}^*(i))]\delta\underline{\theta}(1) \tag{5.131}$$

Hence $$\delta\underline{\theta}(K+1) = [\sum_{i=1}^{K} \underline{M}_{,\underline{\theta}}(\theta^*(i))]\, \delta\underline{\theta}(1) \tag{5.132}$$

$$= A_K \quad \delta\underline{\theta}(1)$$

If $|\lambda_i(A_k)| < 1$, $i \in (1,2)$, the periodic solution is stable.

If $|\lambda_i(A_k)| > 1$, $i \in (1,2)$, the periodic solution is unstable.

If $|\lambda_i(A_k)| = 1$, $i \in (1,2)$, the periodic solution is stable, provided A_k is simple. Otherwise unstable. This is a property which is special to piecewise linear systems.

<u>Example</u>

If in equation (5.116) we set $\mu = 0$ and $\eta = 1$, we have

$$\left.\begin{aligned} &z_{n+1} - 2z_n + z_{n-1} = -4k_2(z_n) \\ &k_2(z) = \frac{z}{2} \qquad \text{for } |z| \le 2 \\ &\qquad\;\; = \operatorname{sgn} z \quad \text{for } |z| > 2 \end{aligned}\right\} \tag{5.133}$$

(a) If $|z_0|,|z_1| < 2$, then

$$\left.\begin{aligned} &z_{n+1} + z_{n-1} = 0 \qquad z_{n+1} = -z_{n-1} \\ \therefore\quad &z_2 = -z_0 , \qquad z_3 = -z_1 \\ &z_4 = z_0 , \qquad z_5 = z_1 \\ &\quad\vdots \\ &z_{2n} = (-1)^n z_0 \qquad z_{2n+1} = (-1)^n z_1 \end{aligned}\right\} \tag{5.134}$$

Since $|z_0|,|z_1| < 2$, it follows that $|z_k| < 2 \quad \forall\, k$

From equation (5.117) with $\mu = 0$, $\eta = 1$,

$$\underline{\theta}_{n+1} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \underline{\theta}_n \tag{5.135}$$

$$M_{,\underline{\theta}}(\underline{\theta}^*(i)) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \forall i$$, provided $\underline{\theta}^*(i)$ is not close to a corner of $k_2(z)$. Thus

$$\left.\begin{aligned} &\delta\underline{\theta}(K+1) = A_K\, \delta\theta(1) \\ &A_K = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^K \\ \therefore \quad &|\lambda_i(A_K)| = 1 \qquad i = 1,2 \end{aligned}\right\} \tag{5.136}$$

Thus as long as the initial perturbations are small, not necessarily infinitesimally small, the periodic solutions are stable

(b) If $|z_0|, |z_1| > 2$, there exist periodic solutions with $|z_n| > 2 \; \forall n$. In particular, If $z_0 = N+1$, $z_1 = 3N-1$, then $T_N = 2(N+1)$

Proof. If $z_n > 2$, equation (5.133) becomes

$$z_{n+1} - 2z_n + z_{n-1} = -4 \tag{5.137}$$

With $z_0 = N+1$ and $z_1 = 3N-1$

$$z_n = (1+N) + 2n(N-n) \tag{5.138}$$

Thus $z_{N-1} = 3N-1 = z_1$; $z_N = N+1 = z_0 > 2$

$$z_{N+1} = -(N+1) = -z_0 < 2 \tag{5.139}$$

If $z_n < -2$, equation (5.133) becomes

$$z_{n+1} - 2z_n + z_{n-1} = +4 \tag{5.140}$$

$$\therefore \quad z_{N+2} = -(3N-1) = -z_1$$

Thus for $N < n < 2N+1$, the solution for $0 < n < N$ is repeated with the negative sign.

$$\underline{z}_{2(N+2)} = (N+1) = +z_0 \tag{5.141}$$

Thus there exists a periodic solution with initial data $z_0 = N+1$, $z_1 = 3N-1$, with $|z_n| > 2 \quad \forall n$ and period $T_N = 2(N+1)$. Clearly, there exists an infinity of such solutions.

Stability

Since each point $\underline{\theta}^*(k)$ satisfies the condition $|z_n| > 2$, each $\underline{M}_{,\underline{\theta}}(\theta^*(i))$ is the same.

$$\therefore \quad A_K = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}^K \tag{5.142}$$

$$\text{Now} \quad \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} = T \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} T^{-1} \tag{5.143}$$

$$\therefore \quad A_K = T \begin{bmatrix} 1 & K \\ 0 & 1 \end{bmatrix} T^{-1} \tag{5.144}$$

Thus, though $|\lambda_i(A_K)| = 1 \; \forall i \in (1,2)$, A_K is not simple, hence the periodic solutions are weakly unstable, and will grow until $\underline{\theta}(n) = \underline{\theta}^*(n) + \delta\theta(n)$ reaches a corner of $k_2(z)$, at which point the nature of the stability will change. As shown in (5.125), the global rate of growth is limited to $O(n^2)$ as $n \to \infty$, which is still a rather weak type of instability.

Globally Unstable Solutions

Hughes (1) and others have exhibited numerically the instability of equation (5.107) and hence of equation (5.133). The instability of equation (5.133) can also be exhibited analytically. Consider equation (5.133) with prescribed initial data

$$\left.\begin{aligned} &z_{n+1} - 2z_n + z_{n-1} = -4k_2(z_n) \\ &k_2(z) = z/2 \quad \text{for} \quad |z| \leq 2 \\ &\qquad\; = \text{sgn } z \quad \text{for} \quad |z| > 2 \\ &z_0 = 1.5N + 2.5 \;, \quad z_1 = 3.5N + 1.5 \;, \quad N = 1,2,3,\text{etc.} \end{aligned}\right\} \tag{5.145}$$

Since $z_0, z_1 > 2$

$$z_k = A + Bk - 2k^2, \text{ provided } z_{n-1} > 2 \tag{5.146}$$

Using the given initial data

$$\left.\begin{aligned} &z_0 = A = 1.5N + 2.5 \\ &z_1 = A + B - 2 = 3.5N + 1.5 \quad \therefore \; B = 2N+1 \end{aligned}\right\} \tag{5.147}$$

$$\therefore \qquad z_k = (1.5N + 2.5) + (2N+1 - 2k)k \quad \text{for} \quad k \leq N+2 \tag{5.148}$$

$$\therefore \qquad z_{N+1} = 0.5N + 1.5 \;, \quad z_{N+2} = -(1.5N + 3.5) \tag{5.149}$$

To determine z_{N+3}, we use equation (5.145)

$$\begin{aligned} \therefore \qquad z_{N+3} &= -(3N+7) - (0.5N + 1.5) + 4 \\ &= -(3.5N + 4.5) \end{aligned} \tag{5.150}$$

Since z_{N+2}, z_{N+3} are less than minus two, we can write

$$\left.\begin{aligned} &z_{N+2+k} = -[A_1 + B_1 k - 2k^2] \qquad k > 0 \\ &\text{where } A_1 = -z_{N+2} = (1.5N + 3.5) \\ &z_{N+3} = -[A_1 + B - 2] \qquad \therefore B_1 = 2N+3 \end{aligned}\right\} \tag{5.151}$$

$$\left.\begin{aligned} \therefore \qquad &z_{N+2+k} = -[1.5N + 3.5 + (2N+3 - 2k)k] \;, \quad k \leq N+2 \\ &z_{2w+4} = -(0.5N + 1.5) \\ &z_{2w+5} = (1.5N + 5.5) = (z_0 + 3) \end{aligned}\right\} \tag{5.152}$$

To calculate z_{2N+6} we return to equation (5.145), from which

$$z_{2N+6} = (3.5N+8.5) = z_1 + 7 \tag{5.153}$$

Thus at the end of a complete cycle

$$\left.\begin{array}{l} T_1 = (2N+5) \\ \text{and} \quad \left.\begin{array}{l} z_{T_1} = z_0 + 3 \\ z_{T_1+1} = z_1 + 7 \end{array}\right\} \text{are the initial data for the next cycle} \end{array}\right\} \tag{5.154}$$

At the end of the next complete cycle

$$\left.\begin{array}{l} T_2 = 4N+14 \\ \text{and} \quad \left.\begin{array}{l} z_{T_2} = z_o + 2 \times 3 \\ z_{T_2+1} = z_1 + 2 \times 7 \end{array}\right\} \end{array}\right\} \tag{5.155}$$

At the end of the k^{th} complete cycle

$$\left.\begin{array}{l} T_k = k(2N+3+2k) \\ \text{and} \quad \left.\begin{array}{l} z_{T_k} = z_o + 3k \\ z_{T_{k+1}} = z_1 + 7k \end{array}\right\} \text{are the initial data for the next cycle} \end{array}\right\} \tag{5.156}$$

Returning for a moment to equation (5.101) with $\eta = 1$, then

$$\frac{1}{2}(\dot{y}_{n+1}^2 - \dot{y}_n^2) + (y_{n+1} - y_n)(f(y_{n+1}) + f(y_{n+1})) = 0 \tag{5.157}$$

If y_{n+1}, y_n are both of the same sign and both are greater than unity, then (5.157) becomes

$$\frac{1}{2}(\dot{y}_{n+1}^2 - \dot{y}_n^2) + \frac{1}{2}(y_{n+1} - y_n)(\text{sgn } y_{n+1} + \text{sgn } y_n) = 0 \tag{5.158}$$

$$\therefore \quad \frac{1}{2}(\dot{y}_{n+1}^2 - \dot{y}_n^2) + |y_{n+1}| - |y_n| = 0 \tag{5.159}$$

From equation (5.106) with x_{n+1} and x_n of the same sign and both greater than unity, then,

$$\frac{1}{2}(\dot{x}_{n+1}^2 - \dot{x}_n^2) + F(x_{n+1}) - F(x_n) = 0 \tag{5.160}$$

But $\quad F(x_{n+1}) - F(x_n) = |x_{n+1}| - |x_n|$

$$\therefore \quad \frac{1}{2}(\dot{x}_{n+1}^2 - \dot{x}_n^2) + |x_{n+1}| - |x_n| = 0 \tag{5.161}$$

Thus the trapezoidal algorithm preserves the energy identity (5.161) if y_{n+1} and y_n are both on the same nonlinear saturated branch. If y_{n+1} and y_n are both on the linear branch of the curve, energy is again conserved. If y_{n+1} and y_n are not on the linear branch or not on the same saturated nonlinear branch, then in general energy is not conserved.

Returning to equations (5.109) and (5.112),

$$\left.\begin{aligned} y_{n+1} &= g(z_n) \\ \dot{y}_{n+1} &= -\frac{1}{2}[z_n + z_{n-1} - 2g(z_n)] \end{aligned}\right\} \tag{5.162}$$

Since $|z_n| > 2 \quad |y_{n+1}| > 1$, thus we need only look at the "energy" at the beginning of each half-cycle to see how it is growing.

$$\begin{aligned} E_k &= \frac{1}{2}\dot{y}_{k+1}^2 + |y_{k+1}| \\ |y_{n+1}| &= |z_n - \text{sgn } z_n| = |z_n| - 1 \end{aligned} \tag{5.163}$$

From (5.156)

$$E_k = \frac{1}{2}(z_0 + z_1 + 10k - 2)^2 + |z_1 + 7k| - 1 \tag{5.164}$$

As $k \to \infty$,

$$E_k \sim O(k^2) \tag{5.165}$$

But from (5.155)

$$T_k = k(2N+3+2k)$$

$$\therefore \quad \text{As} \quad k \to \infty, \quad T_k \sim O(k^2) \tag{5.166}$$

Therefore, combining (5.165) and (5.166)

$$E_k \sim T_k \quad \text{as} \quad k \to \infty \tag{5.167}$$

Thus confirming analytically what Hughes and others had obtained numerically.

Equation (5.116) was carefully examined for $\mu = 0$ and η arbitrary; nothing essentially new was learned, except that even for η very small, but not zero, weak instability will still occur if the initial data are large enough.

Equation (5.116) was carefully examined for $|\mu| > 0, \eta$ both arbitrary; for $|\mu|$ sufficiently small the system behaves in very much the same way as for $\mu = 0$. It is true that the system appears to have bounded solutions, however, the bound is of the order of $(1/\mu)$ and hence can become very large for certain ranges of η.

Since the results of this section are essentially negative, we shall not report all the work that was done to investigate the effect of nonzero μ, the effect of small η and the effect of damping. Instead, we refer the interested reader to the PhD thesis of my student, B. D. Westermo (2).

Algorithms which Conserve Energy

As shown in the last section the trapezoidal algorithm, which was found to be very useful for linear problems, can for a certain class of non-linearities lead to weak instability. In this section we shall look at several new algorithms which conserve energy.

Consider the conservation nonlinear differential equation

$$\left.\begin{aligned} &\frac{d^2x}{dt^2} + f(x) = 0 \qquad 0 < t \leq T \\ &\text{where} \\ &f(x) = \frac{dF}{dx} \quad , \quad F(x) = \int_0^x f(\eta)d\eta > 0 \quad x \neq 0 \\ &\text{and} \\ &xf(x) > 0 \quad x \neq 0 \end{aligned}\right\} \tag{5.168}$$

The system (5.168) has the first integral

$$\frac{1}{2}\dot{x}^2 + F(x) = \text{const} \tag{5.169}$$

and since $F(x)$ is positive definite, (5.168) is Liapunov stable with respect to the origin.

If we rewrite equation (5.168) in the form

$$\frac{d^2x}{dt^2} + \frac{dF}{dx} = 0 \tag{5.170}$$

This immediately suggests the algorithm

$$\left.\begin{aligned} y_{n+1} - y_n &= \frac{\Delta t}{2}(\dot{y}_{n+1} + \dot{y}_n) \\ \dot{y}_{n+1} - \dot{y}_n &= -\Delta t \frac{F(y_{n+1}) - F(y_n)}{(y_{n+1} - y_n)} \end{aligned}\right\} \tag{5.171}$$

Cross multiplication immediately shows that:

$$\frac{1}{2}(\dot{y}_{n+1}^2 - \dot{y}_n^2) + F(y_{n+1}) - F(y_n) = 0 \tag{5.172}$$

Or on summing

$$\frac{1}{2}\dot{y}_{n+1}^2 + F(y_{n+1}) = \text{const} \tag{5.173}$$

If in (5.168)

$$\left.\begin{aligned} & f(x) = \omega^2 x + g(x) \\ & g(x) = \frac{dG}{dx}, G(x) = \int_0^x g(\eta)d\eta > 0 \quad x \neq 0 \\ & \omega^2 x^2 + g(x) > \quad \forall\, x \neq 0 \end{aligned}\right\} \tag{5.174}$$

g(x) continuous with continuous first and second derivatives then (5.171) becomes

$$\left.\begin{aligned} & y_{n+1} - y_n = \frac{\Delta t}{2}(\dot{y}_{n+1} + \dot{y}_n) \\ & \dot{y}_{n+1} - \dot{y}_n = -\frac{\Delta t}{2}\omega^2(y_{n+1} + y_n) - \Delta t\frac{G(y_{n+1}) - G(y_n)}{(y_{n+1} - y_n)} \end{aligned}\right\} \tag{5.175}$$

which may also be written in the form

$$\underline{w}_{n+1} = \underline{w}_n + \frac{\Delta t}{2} A(\underline{w}_{n+1} + \underline{w}_n) + \chi(\underline{w}_{n+1}, \underline{w}_n) \tag{5.176}$$

where

$$\left.\begin{aligned} A &= \begin{bmatrix} 0 & 1 \\ \omega^2 & 0 \end{bmatrix} \qquad \underline{w}_n = \begin{pmatrix} y_n \\ \dot{y}_n \end{pmatrix} \\ \underline{\chi} &= \begin{pmatrix} 0 \\ \Delta t \dfrac{G(y_{n+1}) - G(y_n)}{y_{n+1} - y_n} \end{pmatrix} \end{aligned}\right\} \tag{5.177}$$

Accuracy

Now

$$\left.\begin{aligned} \frac{G(y_{n+1})-G(y_n)}{(y_{n+1}-y_n)} &= g(y_n)+g'(y_n)\frac{y_{n+1}-y_n}{2} + g''(\xi)\frac{(y_{n+1}-y_n)^2}{6} \\ \xi &= \theta y_n + (1-\theta)y_{n+1} \qquad 0 < \theta < 1 \end{aligned}\right\} \tag{5.178}$$

Thus if y_n, y_{n+1} are bounded $\forall n$

$$\left| \frac{G(y_{n+1}) - G(y_n)}{y_{n+1} - y_n} \right| < M \qquad \forall n \tag{5.179}$$

$$\therefore \qquad \|\underline{\chi}(\underline{w}_{n+1}, \underline{w}_n)\| < M\Delta t \ \forall n \tag{5.180}$$

Similarly if $\dot{y}_{n+1}$, y_n are bounded $\forall n$

$$\left\| \frac{\Delta t}{2} A(\underline{w}_{n+1} + \underline{w}_n) \right\| < N\Delta t \qquad \forall n \tag{5.181}$$

Hence from equation (5.176)

$$\|\underline{w}_{n+1} - \underline{w}_n\| \le \quad (N+M)\Delta t \sim O(\Delta t) \quad \text{for } \Delta t \text{ small} \tag{5.182}$$

As before, we define the truncation error $\underline{\tau}_{n+1}$, by the equations

$$\underline{z}_{n+1} = \underline{z}_n + \frac{\Delta t}{2} A(\underline{z}_{n+1} + \underline{z}_n) + \underline{\chi}(\underline{z}_{n+1}, \underline{z}_n) + \underline{\tau}_{n+1} \Delta t$$

$$\underline{z}_n = \begin{pmatrix} x_n \\ \dot{x}_n \end{pmatrix} \tag{5.183}$$

Using equations (5.174), (5.178) and (5.183) it is readily shown that if x_n, $\dot{x}_n$ are bounded for all n:

$$\|\tau_{n+1}\| \sim O(\Delta t^2) \tag{5.184}$$

Defining the solution error $\underline{e}_n = \underline{z}_n - \underline{w}_n$, if (5.176) is subtracted from (5.183)

$$\underline{e}_{n+1} = \underline{e}_n + \frac{\Delta t}{2} A(\underline{e}_{n+1} + \underline{e}_n) + \underline{\chi}(\underline{z}_{n+1}, \underline{z}_n) - \underline{\chi}(\underline{w}_{n+1}, \underline{w}_n) + \underline{\tau}_{n+1} \Delta t \tag{5.185}$$

$$\therefore \quad \|\underline{e}_{n+1}\| \leq \|\underline{e}_n\| + \frac{\Delta t}{2} \|A\|(\|\underline{e}_{n+1}(1 + \|\underline{e}_n\|) + \|\underline{\chi}(\underline{z}_{n+1}, \underline{z}_n) - \chi(\underline{w}_{n+1}, \underline{w}_n)\| + \|\underline{\tau}_{n+1}\| \Delta t \tag{5.186}$$

Using (5.178)and (5.182)

$$\|\underline{\chi}(\underline{z}_{n+1}, \underline{z}_n) - \chi(\underline{w}_{n+1}, \underline{w}_n)\| \leq \frac{\Delta t}{2} K(\|\underline{e}_{n+1}\| + \|\underline{e}_n\|) + \delta_{n+1} \Delta t$$

provided that $\underline{z}_n$, $\underline{z}_{n+1}$, $\underline{w}_n$, $\underline{w}_{n+1}$ are bounded $\forall n$ where (5.187)

$$\delta_{n+1} \sim O(\Delta t^2)$$

Let

$$\phi = \text{Max}(\delta_{n+1} + \|\tau_{n+1}\|) \sim O(\Delta t^2) \quad \text{as} \quad \Delta t \to 0 \tag{5.188}$$

Using (5.186), (5.187) and (5.188)

$$\|e_{n+1}\| \leq \left(\frac{1 + (\|A\| + K) \frac{\Delta t}{2}}{1 - (\|A\| + K) \frac{\Delta t}{2}} \right) \|\underline{e}_n\| + \phi \; \Delta t \tag{5.189}$$

If $\|\underline{e}_o\| = 0$ then as previously shown

$$\|\underline{e}_n\| \le \frac{\exp\left((\|A\|+K)\frac{3}{2}T\right)}{\|A\|+K}\,\phi \tag{5.190}$$

Thus for T fixed

$$\|\underline{e}_n| \le K_3\phi \sim O(\Delta t^2) \text{ as } \Delta t \to 0 \tag{5.191}$$

Thus this algorithm is second order accurate; in addition it is Liapunov stable with respect to the origin.

We note in passing that equation (5.178) may also be written

$$\frac{G(y_{n+1})-G(y_n)}{y_{n+1}-y_n} = \frac{1}{2}(g(y_{n+1})+g(y_n)) + \left[\frac{g''(\xi_1)+g''(\xi_2)-3g''(\xi_3)}{12}\right](y_{n+1}-y_n)^2$$

$$\xi_i = y_{n+1}\theta_i + (1-\theta_i)y_n$$

$$0 < \theta_i < 1 \tag{5.192}$$

Thus if y_n, y_{n+1} are bounded $\forall n$, then for Δt tending to zero.

$$\frac{G(y_{n+1})-G(y_n)}{y_{n+1}-y_n} = \frac{1}{2}[g(y_{n+1})+g(y_n)] + M(\Delta t)^2 \tag{5.193}$$

Thus this present algorithm may be considered to be a modified "trapezoidal" algorithm.

Effect of Viscous Damping

If viscous damping is added to equation (5.168), then using (5.174) we have

$$\ddot{x} + 2z\dot{x} + \omega^2 x + g(x) = 0 \quad , \qquad z > 0 \tag{5.194}$$

Let

$$V(x,\dot{x}) = \frac{1}{2}[\dot{x}^2 + 2zx\dot{x} + \omega^2x^2 + 2z^2x^2 + 2G(x)] \qquad (5.195)$$

$$V(x,\dot{x}) > 0 \quad \text{if} \quad x,\dot{x} \neq 0$$

$$\dot{V}(x,\dot{x}) = \dot{x}\ddot{x} + z\dot{x}^2 + zx\ddot{x} + \omega^2x\dot{x} + 2z^2x\dot{x} + g(x)\dot{x} \qquad (5.196)$$

Using equation (5.194)

$$\dot{V} = -z(\dot{x}^2 + \omega^2x^2 + xg(x)) \qquad (5.197)$$

But, from equation (5.174)

$$\omega^2x^2 + xg(x) > 0 \qquad \forall x \neq 0$$

$$\therefore \qquad \dot{V} < 0 \qquad (5.198)$$

Hence V is a Liapunov function and the system (5.194) is Liapunov asymptotically stable at the origin.

Consider now the discrete form of equation (5.194)

$$y_{n+1} - y_n = \frac{\Delta t}{2}(\dot{y}_{n+1} + \dot{y}_n) \qquad (5.199)$$

$$\dot{y}_{n+1} - \dot{y}_n = -\frac{\Delta t}{2}[2z(\dot{y}_{n+1} + \dot{y}_n) + \omega^2(y_{n+1} + y_n)] - \Delta t\,\frac{G(y_{n+1}) - G(y_n)}{y_{n+1} - y_n} \qquad (5.200)$$

Let

$$V_n = \frac{1}{2}[(\dot{y}_n + zy_n)^2 + (z^2 + \omega^2)y_n^2 + 2G(y_n)] \qquad (5.201)$$

V_n is positive definite and vanishes only when $y_n = \dot{y}_n = 0$.

$$\therefore \qquad V_{n+1} = \frac{1}{2}[(\dot{y}_{n+1} + zy_{n+1})^2 + (z^2 + \omega^2)y_{n+1}^2 + 2G(y_{n+1})] \qquad (5.202)$$

$$\therefore \quad \Delta V_n = V_{n+1} - V_n$$

$$= \frac{1}{2}[(\dot{y}_{n+1} - \dot{y}_n + zy_{n+1} - zy_n)(\dot{y}_{n+1} + \dot{y}_n + zy_{n+1} + zy_n)]$$

$$+ \frac{1}{2}[(z^2 + \omega^2)(y_{n+1}^2 - y_n^2)] + G(y_{n+1}) - G(y_n) \qquad (5.203)$$

Using equations (5.200)

$$\Delta V_n = \frac{1}{2}\left[-\Delta t \frac{G(y_{n+1}) - G(y_n)}{y_{n+1} - y_n} - \frac{z\Delta t}{2}(\dot{y}_{n+1} + \dot{y}_n) - \frac{\Delta t}{2}\omega^2(y_{n+1} + y_n)\right] \times$$

$$\left[\frac{2}{\Delta t}(y_{n+1} - y_n) + z(y_{n+1} + y_n)\right]$$

$$+ \frac{1}{2}[z^2 + \omega^2][y_{n+1}^2 - y_n^2] + G(y_{n+1}) - G(y_n) \qquad (5.204)$$

$$\therefore \quad \Delta V_n = -z\Delta t\left[\left(\frac{\dot{y}_{n+1} + \dot{y}_n}{2}\right)^2 + \omega^2\left(\frac{y_{n+1} + y_n}{2}\right)^2 + \left(\frac{y_{n+1} + y_n}{2}\right)\left(\frac{G(y_{n+1}) - G(y_n)}{y_{n+1} - y_n}\right)\right] \qquad (5.205)$$

If $G(y)$ is an even monotone increasing function say

$$G(y) = G^*(y^2) \qquad (5.206)$$

$G^*(y^2)$ is also a monotone increasing function. Hence

$$\frac{G^*(y_{n+1}^2) - G^*(y_n^2)}{y_{n+1}^2 - y_n^2} \geq 0 \qquad (5.207)$$

$$\therefore \quad \Delta V_n = -z\Delta t\left[\left(\frac{\dot{y}_{n+1} + \dot{y}_n}{2}\right)^2 + \left(\frac{y_{n+1} + y_n}{2}\right)^2\left(\omega^2 + \frac{G^*(y_{n+1}^2) - G^*(y_n^2)}{y_{n+1}^2 - y_n^2}\right)\right]$$

$$\leq -z\Delta t\left[\left(\frac{\dot{y}_{n+1} + \dot{y}_n}{2}\right)^2 + \omega^2\left(\frac{y_{n+1} + y_n}{2}\right)^2\right] \qquad (5.208)$$

From (5.208) we observe that:

(a) $\Delta V_n \leq 0$

(b) Since $y_{n+1} - y_n = \frac{\Delta t}{2} [\dot{y}_{n+1} + \dot{y}_n]$, both terms in (5.208) cannot vanish simultaneously unless $y_n = y_{n+1} = 0$

(c) Since $\underline{w}_{n+1} - \underline{w}_n \sim O(\Delta t)$

$\therefore$ $\Delta V_n \leq -z\Delta t[(\dot{y}_n)^2 + \omega^2 (y_n)^2] + O(\Delta t^2)$ as $\Delta t \to 0$

$\therefore$ The discrete system (5.199)is Liapunov asymptotically stable.

Effect of Viscous Damping and Additive Forces

If to equation (5.168), viscous damping and external forces are added, then using (5.174) we have

$$\ddot{x} + 2z\dot{x} + \omega^2 x + g(x) = p(t) \tag{5.209}$$

then if $\sup_t |p(t)| = p_o$, all solutions of (5.209) are ultimately bounded.

Let

$$V(x,\dot{x}) = \frac{1}{2} [\dot{x}^2 + 2zx\dot{x} + \omega^2 x^2 + 2z^2 x^2 + 2G(x)] \tag{5.210}$$

$$V(x,\dot{x}) > 0 \qquad x, \dot{x} \neq 0$$

$$\dot{V} = \dot{x}\ddot{x} + z\dot{x}^2 + zx\ddot{x} + \omega^2 x\dot{x} + 2z^2 x\dot{x} + g(x)\dot{x} \tag{5.211}$$

Using equation (5.209)

$$\dot{V} = -z(\dot{x}^2 + \omega^2 x^2 + xg(x)) + p(t)(\dot{x} + zx)$$

$\therefore$ if $xg(x) > 0 \qquad x \neq 0$

$$\dot{V} \leq -z(\dot{x}^2 + \omega^2 x^2) + p_o(|\dot{x}| + z|x|) \tag{5.212}$$

$$\leq -z(\dot{x}^2 + \omega^2 x^2) + p_o \sqrt{2\left(1 + \left(\frac{z}{\omega}\right)^2\right)} \sqrt{\dot{x}^2 + \omega^2 x^2} \tag{5.213}$$

Let S be the set $\dot{x}^2 + \omega^2 x^2 \leq 2 \left(\frac{p_o}{z}\right)^2 \left(1 + \frac{z}{\omega}\right)^2)$ (5.214)

Outside the set S, $\dot{V} < 0$

Let Ω be the set $V \le e$, where e is such that S is a proper subset of Ω. Then, for points $(x, \dot{x})$ outside Ω, $\dot{V} < 0$ (5.215)

Starting outside Ω, $V > 0$, $\dot{V} < 0$, therefore, V decreases and the trajectory must eventually enter Ω, and once inside Ω, the trajectory cannot leave Ω since $\dot{V} \le 0$ on $\partial\Omega$. Starting inside Ω, $V > 0$, $\dot{V}$ is in general sign indefinite, therefore V may increase, however, it is clear that the trajectory cannot leave Ω since $\dot{V} \le 0$ on $\partial\Omega$. Thus all solutions of (5.209) are ultimate bounded in $\bar{\Omega} = (\Omega + \partial\Omega)$.

Consider now the discrete form of equation (5.209)

$$y_{n+1} - y_n = \frac{\Delta t}{2} (\dot{y}_{n+1} + \dot{y}_n)$$

$$\dot{y}_{n+1} - \dot{y}_n = -\frac{\Delta t}{2} [2z(\dot{y}_{n+1} + \dot{y}_n) + \omega^2 (y_{n+1} + y_n) - (p_{n+1} + p_n)] - \Delta t \frac{G(y_{n+1}) - G(y_n)}{y_{n+1} - y_n} \qquad (5.216)$$

Let

$$V_n = \frac{1}{2} [(\dot{y}_n + z y_n)^2 + (z^2 + \omega^2) y_n^2 + 2G(y_n)] \qquad (5.217)$$

$$\Delta V_n = V_{n+1} - V_n$$

$$= \frac{1}{2} [(\dot{y}_{n+1} + z y_{n+1})^2 + (z^2 + \omega^2) y_{n+1}^2 + 2G(y_{n+1}) - (\dot{y}_n + z y_n)^2 - (z^2 + \omega^2) y_n^2 - 2G(y_n)] \qquad (5.218)$$

Using equation (5.216)

$$\Delta V_n = -z\Delta t\left[\left(\frac{\dot{y}_{n+1}+\dot{y}_n}{2}\right)^2 + \left(\frac{y_{n+1}+y_n}{2}\right)^2\left(\omega^2 + \frac{G^*(y_{n+1}^2)-G^*(y_n^2)}{y_{n+1}^2-y_n^2}\right)\right]$$

$$+ \Delta t\left[\left(\frac{p_{n+1}+p_n}{2}\right)\left(\frac{\dot{y}_{n+1}+\dot{y}_n}{2}\right) + z\left(\frac{y_{n+1}+y_n}{2}\right)\right] \tag{5.219}$$

Using (5.207) we have:

$$\Delta V_n \le -z\Delta t\left[\left(\frac{\dot{y}_{n+1}+\dot{y}_n}{2}\right)^2 + \omega^2\left(\frac{y_{n+1}+y_n}{2}\right)^2\right]$$

$$+\Delta t\, p_o\left(\left|\frac{\dot{y}_{n+1}+\dot{y}_n}{2}\right| + z\left|\frac{y_{n+1}+y_n}{2}\right|\right) \tag{5.220}$$

Let

$$\frac{\dot{y}_{n+1}+\dot{y}_n}{2} = \langle\dot{y}_n\rangle \quad , \quad \frac{y_{n+1}+y_n}{2} = \langle y_n\rangle$$

$$\Delta V_n \le -z\Delta t[\langle\dot{y}\rangle^2 + \omega^2\langle y_n\rangle^2]$$

$$+ \Delta t\, p_o\sqrt{2\left(1+\left(\frac{z}{\omega}\right)^2\right)}\sqrt{\langle\dot{y}_n\rangle^2+\omega^2\langle y_n\rangle^2} \tag{5.221}$$

Let $\langle S\rangle$ be the set

$$\langle\dot{y}_n\rangle^2 + \omega^2\langle y_n\rangle^2 \le 2\left(\frac{P_o}{2}\right)^2\left(1+\left(\frac{z}{\omega}\right)^2\right) \tag{5.222}$$

Since

$$\underline{w}_{n+1} - \underline{w}_n \sim O(\Delta t) \quad \text{as} \quad t \to 0.$$

Clearly $\langle S\rangle \to S$ as $\Delta t \to 0$ and the previous arguments can be used to show that the solutions of the discrete equations (5.216) are ultimately bounded, just as the solutions of the continuous equations (5.209).

Extension of Energy Conserving Algorithms to Multidegree-Freedom Nonlinear Systems

Consider the system of conservative nonlinear differential equations

$$M\ddot{\underline{x}} + K\underline{x} + \nabla_{\underline{x}} u(\underline{x}) = 0 \tag{5.223}$$

where M, K are $N \times N$ symmetric positive definite matrices and $u(\underline{x})$ is a positive definite potential function. For the system (5.233)

$$\left.\begin{aligned} V(\underline{x},\dot{\underline{x}}) &= \frac{1}{2}[\dot{\underline{x}}^T M \dot{\underline{x}} + \underline{x}^T K \underline{x} + 2u(\underline{x})] > 0 \\ \dot{V} &= \dot{\underline{x}}^T M \ddot{\underline{x}} + \dot{\underline{x}}^T K \underline{x} + \dot{\underline{x}}^T \nabla_{\underline{x}} u(\underline{x}) = 0 \end{aligned}\right\} \tag{5.224}$$

Algorithm A

We may write equation (5.223) in discrete form as

$$\underline{y}_{n+1} - \underline{y}_n = \frac{\Delta t}{2}[\dot{\underline{y}}_{n+1} + \dot{\underline{y}}_n]$$

$$M[\dot{\underline{y}}_{n+1} - \dot{\underline{y}}_n] = -\frac{\Delta t}{2} K[\underline{y}_{n+1} + \underline{y}_n] - \Delta t \frac{(u(\underline{y}_{n+1}) - u(\underline{y}_n))[\nabla u(\underline{y}_{n+1}) + \nabla u(\underline{y}_n)]}{(\underline{y}_{n+1} - \underline{y}_n)^T[\nabla u(\underline{y}_{n+1}) + \nabla u(\underline{y}_n)]} \tag{5.225}$$

Cross multiplication yields

$$\frac{1}{2}[\dot{\underline{y}}_{n+1}^T M \dot{\underline{y}}_{n+1} + \underline{y}_{n+1}^T K \underline{y}_{n+1} + 2u(\underline{y}_{n+1})] = \frac{1}{2}[\dot{\underline{y}}_n^T M \dot{\underline{y}}_n + \underline{y}_n^T K \underline{y}_n + 2u(\underline{y}_n)] \tag{5.226}$$

Thus if

$$\left.\begin{aligned} V_n &= \frac{1}{2}\left[\dot{\underline{y}}_n^T M \dot{\underline{y}}_n + \underline{y}_n^T K \underline{y}_n + 2u(\underline{y}_n)\right] > 0 \\ \Delta V_n &= V_{n+1} - V_n = 0 \end{aligned}\right\} \qquad (5.227)$$

Hence the discrete equations (5.225) conserve energy in exactly the same way as the continuous time equations (5.223).

Accuracy

Since energy is conserved, $\underline{y}_n, \dot{\underline{y}}_n$, are bounded for all n, provided $\underline{y}_o, \dot{\underline{y}}_o$ are bounded. Thus as

$$\Delta t \to 0 \quad , \quad \underline{y}_{n+1} - \underline{y}_n \sim O(\Delta t)$$

$$\left.\begin{aligned} &\frac{(u(\underline{y}_{n+1}) - u(\underline{y}_n))\,[\nabla u(\underline{y}_{n+1}) + \nabla u(\underline{y}_n)]}{(\underline{y}_{n+1} - \underline{y}_n)^T[\nabla u(\underline{y}_{n+1}) + \nabla u(\underline{y}_n)} \\ &\qquad = \frac{1}{2}\,[\nabla u(\underline{y}_{n+1}) + \nabla u(\underline{y}_n] + O(\Delta t^2) \end{aligned}\right\} \qquad (5.228)$$

Hence, as $\Delta t \to 0$, the discrete equation of algorithm A became of trapezoidal form, and hence this algorithm is second order accurate as $\Delta t \to 0$. While this algorithm conserves energy, it has two defects.

a) It is difficult to use, that is, it is not readily computable.

b) If Δt is not small, we have not been able to prove that:

$$(\underline{y}_{n+1} - \underline{y}_n)^T[\nabla u(\underline{y}_{n+1}) + \nabla u(\underline{y}_n)] = 0 \quad \text{implies that,}$$

$$u(\underline{y}_{n+1}) = u(\underline{y}_n)$$

Thus we are unable to prove that the last term in (5.225) is bounded when $\underline{y}_n, \underline{y}_{n+1}$ are bounded.

Algorithm B

An alternative to the discrete gradient operator in (5.225) is the operator

$$Lu = \left\{ \frac{\frac{1}{N}\sum_{k=1}^{N} \Delta_i u_k}{y_{n+1}^i - y_n^i} \right\} \tag{5.229}$$

where

$$\Delta_i u_k = u(y_{n+1}^1, y_{n+1}^2 \; . \; . \; . y_{n+1}^k, \; y_n^{k+1}, \; . \; . \; . y_n^{i-1} y_{n+1}^i, \; y_n^{i+1} \; . \; . \; . y_n^n)$$

$$-u(y_{n+1}^1, y_{n+1}^2 . \; . \; . \; y_{n+1}^k, y_n^{k+1}, y_n^{k+2} \; . \; . \; . y_n^n) \tag{5.230}$$

Thus

$$\left.\begin{aligned} &\underline{y}_{n+1} - \underline{y}_n = \frac{\Delta t}{2} [\dot{\underline{y}}_{n+1} + \dot{\underline{y}}_n] \\ &M[\dot{\underline{y}}_{n+1} - \dot{\underline{y}}_n] = -\frac{\Delta t}{2} K [\underline{y}_{n+1} + \underline{y}_n] \\ &\qquad -\frac{\Delta t}{N} \left\{ \begin{matrix} \Sigma \Delta_1 u_k / (y_{n+1}^1 - y_n^1) \\ \Sigma \Delta_2 u_k / (y_{n+1}^2 - y_n^2) \\ \vdots \\ \Sigma \Delta_n u_k / (y_{n+1}^n - y_n^n) \end{matrix} \right\} \end{aligned}\right\} \tag{5.231}$$

Cross multiplication yields

$$\frac{1}{2} [\dot{\underline{y}}_{n+1}^T M \dot{\underline{y}}_{n+1} + \underline{y}_{n+1}^T K \underline{y}_{n+1}]$$

$$+ \frac{1}{N} \sum_{i,k}^{N} \Delta_i u_k = \frac{1}{2} [\dot{\underline{y}}_n^T M \dot{\underline{y}}_n + \underline{y}_n^T K \underline{y}_n] \tag{5.232}$$

Now

$$\frac{1}{N}\sum_{i,k}^{N} \Delta_i u_k = u(\underline{y}_{n+1}) - u(\underline{y}_n) \tag{5.233}$$

To see this, consider $N = 2$

$$\begin{aligned}\therefore \quad \frac{1}{2}\sum_{i,k}^{2} \Delta_i u_k &= \frac{1}{2}\,[u(y^1_{n+1}, y^2_n) - u(y^1_n, y^2_n) + u(y^1_{n+1}, y^2_{n+1}) - u(y^1_n, y^2_{n+1}) \\ &\qquad u(y^1_n, y^2_{n+1}) - u(y^1_n, y^2_n) + u(y^1_{n+1}, y^2_{n+1}) - u(y^1_{n+1}, y^2_n) \\ &= \frac{1}{2}\,[2u(y^1_{n+1}, y^2_{n+1}) - 2u(y^1_n, y^2_n)] \\ &= u(\underline{y}_{n+1}) - u(\underline{y}_n)\end{aligned} \tag{5.234}$$

Thus, using (5.234) in (5.232) we have:

$$\frac{1}{2}\,[\dot{\underline{y}}^T_{n+1} M \dot{\underline{y}}_{n+1} + \underline{y}^T_{n+1} K \underline{y}_{n+1} + 2u(\underline{y}_{n+1})] = \text{constant} \tag{5.235}$$

Thus algorithm B also conserves energy. If the potential $u(\underline{y})$ can be expressed in the form u

$$u(\underline{y}) = u^*\left(\sum_{i=1}^{N} \alpha_i (y^i)^2\right) \tag{5.236}$$

where $u^*(r)$ is a positive monotone increasing function of r, then it may be shown that

$$\frac{\frac{1}{N}\sum_{k=1}^{N} \Delta_i u_k}{(y^i_{n+1})^2 - (y^i_n)^2} \geq 0 \quad \forall i \tag{5.237}$$

<u>Application of Algorithm B to System (5.223) with Viscous Damping and Additive Forces</u>

If to equation (5.223), viscous damping and external forces are added, we have

$$M\ddot{\underline{x}} + C\dot{\underline{x}} + K\underline{x} + \nabla_{\underline{x}} u(\underline{x}) = \underline{p}(t) \tag{5.238}$$

Then if C is symmetric and positive definite and $\|\underline{p}(t)\|$ is bounded, all solutions of (5.238) are ultimately bounded provided $\underline{x}^T \nabla_{\underline{x}} u(\underline{x}) \geq 0$.

Let

$$V = \frac{1}{2}[z(\underline{x}^T(C-zM)\underline{x}) + (\dot{\underline{x}}+z\underline{x})^T M(\dot{\underline{x}}+z\underline{x}) + 2u(\underline{x})] > 0$$
$$+ \underline{x}^T K \underline{x} \tag{5.239}$$

where $2z$ = smallest eigenvalue of $|\lambda M - C| \equiv 0$

$$\dot{V} = z\underline{x}^T(C-zM)\dot{\underline{x}} + (\dot{\underline{x}}+z\underline{x})^T M(\ddot{\underline{x}}+z\dot{\underline{x}}) + \dot{\underline{x}}^T \nabla_{\underline{x}} u(\underline{x}) + \dot{\underline{x}}^T K\underline{x} \tag{5.240}$$

Using (5.238), we have

$$\dot{V} = z\underline{x}^T(C-zM)\dot{\underline{x}} - (\dot{\underline{x}}+2\underline{x})^T[(C-zM)\dot{\underline{x}} + K\underline{x} + \nabla_{\underline{x}} u(x) - \underline{p}(t)]$$
$$+ \dot{\underline{x}}^T K\underline{x} + \dot{\underline{x}}^T \nabla_{\underline{x}} u(\underline{x})$$
$$= -\dot{\underline{x}}^T(C-zM)\dot{\underline{x}} - z(\underline{x}^T K\underline{x} + \underline{x}^T \nabla_{\underline{x}} u(\underline{x})) - (\dot{\underline{x}} + z\underline{x})^T \underline{p}(t) \tag{5.241}$$

Since

$$\underline{x}^T \nabla_{\underline{x}} u(\underline{x}) \geq 0$$

$$\dot{V} \leq -\dot{\underline{x}}^T(C-zM)\dot{\underline{x}} - z(\underline{x}^T K\underline{x}) + |(\dot{\underline{x}}+z\underline{x})^T \underline{p}(t)| \tag{5.242}$$

Since $2z$ is equal to the smallest eigen value of $M^{-1}C$, $C-zM$ is symmetric and positive definite, hence the first two terms in (5.242) are negative definite. Since $\underline{p}(t)$ is bounded, and the third term contains $\underline{x}$ and $\dot{\underline{x}}$ linearly, there exists a set S: $\dot{\underline{x}}^T M\dot{\underline{x}} + \underline{x}^T K\underline{x} \leq K$ such that outside of S, $\dot{V} < 0$. Let Ω be the set $V \leq C$, where C is such that

S is a proper subset of Ω. Then outside of $\Omega \dot{V} < 0$, and applying the arguments used previously, we see that all solutions of (5.238) are ultimately bounded in Ω.

Now consider the discrete form of equation (5.238)

$$\left.\begin{aligned} &\underline{y}_{n+1} - \underline{y}_n = \frac{\Delta t}{2}[\dot{\underline{y}}_{n+1} + \dot{\underline{y}}_n] \\ &M[\dot{\underline{y}}_{n+1} - \dot{\underline{y}}_n] = \frac{\Delta t}{2} K[\underline{y}_{n+1} + \underline{y}_n] - \frac{\Delta t}{2} C[\dot{\underline{y}}_{n+1} + \dot{\underline{y}}_n] \\ &\qquad - \frac{\Delta t}{N} \begin{Bmatrix} \sum_{k=1}^{N} \Delta_1 u_k/(y_{n+1}^1 - y_n^1 \\ \vdots \qquad \vdots \\ \sum_{k=1}^{N} \Delta_N u_k / y_{n+1}^N - y_n^N \end{Bmatrix} \\ &\qquad + \frac{\Delta t}{2}(\underline{p}_{n+1} + \underline{p}_n) \end{aligned}\right\} \tag{5.243}$$

Let

$$V_n = \frac{1}{2}[z(\underline{y}_n^T(C - zM)\underline{y}_n) + (\dot{\underline{y}}_n + z\underline{y}_n)^T M(\dot{\underline{y}}_n + z\underline{y}_n) + \underline{y}_n^T K \underline{y}_n + 2u(\underline{y}_n)] \tag{5.244}$$

Then

$$\begin{aligned} \Delta V_n &= V_{n+1} - V_n \\ &= \frac{1}{2}[z(\underline{y}_{n+1} - \underline{y}_n)^T (C - zM)(\underline{y}_{n+1} + \underline{y}_n) \\ &\qquad + (\dot{\underline{y}}_{n+1} - \dot{y}_n + z(\underline{y}_{n+1} + \underline{y}_n)^T M(\dot{\underline{y}}_{n+1} + \dot{\underline{y}}_n + z(\underline{y}_{n+1} + \underline{y}_n)) \\ &\qquad + (\underline{y}_{n+1} - \underline{y}_n)^T K(\underline{y}_{n+1} + \underline{y}_n) + 2(u(\underline{y}_{n+1}) - u(\underline{y}_n))] \end{aligned} \tag{5.245}$$

Using equations (5.243) and (5.237)

$$\Delta V_n \leq -\Delta t[\langle \dot{\underline{y}}_n \rangle^T (C - zM) \langle \dot{\underline{y}}_n \rangle + z \langle \underline{y}_n \rangle^T K \langle y_n \rangle + \Delta t \, |\langle \dot{\underline{y}}_n + z\underline{y}_n \rangle^T \langle \underline{p}_n \rangle| \quad (5.246)$$

where

$$\langle \underline{x}_n \rangle = \frac{1}{2} [\underline{x}_{n+1} + \underline{x}_n] \quad (5.247)$$

Since $\underline{w}_{n+1} - \underline{w}_n \sim O(\Delta t)$ as $\Delta t \to 0$, it is clear that the right hand side of (5.246) tends to

$$-\Delta t \, [\dot{\underline{y}}_n^T (C - zM) \dot{\underline{y}}_n + z \underline{y}_n^T K \underline{y}_n] + \Delta t \, |(\dot{\underline{y}}_n + z \underline{y}_n)^T \underline{p}_n| \quad \text{as} \quad \Delta t \to 0 \quad (5.248)$$

Thus, for small Δt, the continuous time system (5.238) and the discrete time system (5.243) behave in essentially the same way and the solutions of equations (5.243) are ultimately bounded.

Accuracy

Using the fact that the solutions of (5.244) are ultimately bounded, it is easily shown that the discrete gradient operator used in (5.244) has the following form as $\Delta t \to 0$

$$\frac{1}{N} \left\{ \begin{array}{c} \sum_{k=1}^{N} \Delta_1 u_k / (y_{n+1}^1 - y_n^1) \\ \cdots \cdots \\ \sum_{k=1}^{N} \Delta_N u_k / (y_{n+1}^N - y_n^N) \end{array} \right\} = \frac{1}{2} [\nabla u(\underline{y}_{n+1}) + \nabla u(\underline{y}_n)] + O(\Delta t^2) \quad (5.249)$$

Thus, as $\Delta t \to 0$, equations (5.244) are of trapezoidal form, therefore algorithm B is second order accurate as Δt tends to zero.

Algorithm B conserves energy and has some nice stability

2-2

properties; it unfortunately has two major defects

a) It is difficult to use, that is, it is not readily computable.

b) The property (5.237), which is necessary in order to prove the ultimate boundedness properties of equations (5.243), is valid only for a restricted class of potential functions $u(\underline{y})$.

For this reason, we now turn to an alternate formulation using Lagrange multipliers. This formulation was suggested by my colleague Dr. T. J. R. Hughes and was developed jointly with him and his student Mr. W.K. Liu.(3)

<u>Algorithm C</u>

Consider the system of conservative nonlinear differential equations

$$M\ddot{\underline{x}} + K\underline{x} + \nabla_{\underline{x}} u(x) = 0 \tag{5.250}$$

where M, K are $N \times N$ symmetric positive definite matrices and $u(\underline{x})$ is a positive definite potential function. We know that (5.250) has the Liapunov function

$$V(\underline{x}, \dot{\underline{x}}) = \frac{1}{2} [\dot{\underline{x}}^T M \dot{\underline{x}} + \underline{x}^T K \underline{x} + 2u(\underline{x})] = \text{constant} \tag{5.251}$$

If we write (5.250) in trapezoidal discrete form

$$\left.\begin{aligned} \underline{y}_{n+1} - \underline{y}_n &= \frac{\Delta t}{2} [\dot{\underline{y}}_{n+1} + \dot{\underline{y}}_n] \\ M(\dot{\underline{y}}_{n+1} - \dot{\underline{y}}_n^2) &= -\frac{\Delta t}{2} [K(\underline{y}_{n+1} + \underline{y}_n) + \nabla u_{n+1} + \nabla u_n] \end{aligned}\right\} \tag{5.252}$$

We now wish to constrain (5.252) such that

$$\frac{1}{2} \dot{\underline{y}}_{n+1}^T M \dot{\underline{y}}_{n+1} + \frac{1}{2} \underline{y}_{n+1}^T K \underline{y}_{n+1} + u_{n+1} = \text{const} \tag{5.253}$$

Using the first of equations (5.252)

$$\underline{y}_{n+1} = \frac{2}{\Delta t}\,[\underline{y}_{n+1} - \underline{y}_n] - \dot{\underline{y}}_n \tag{5.254}$$

Substituting into the second of equations (5.252)

$$M[\underline{y}_{n+1} - \underline{y}_n - \Delta t\dot{\underline{y}}_n] = -\left(\frac{\Delta t}{2}\right)^2 [K(\underline{y}_{n+1} + \underline{y}_n) + \nabla(u_{n+1} + u_n)] \tag{5.255}$$

which may be written

$$M\underline{y}_{n+1} + \left(\frac{\Delta t}{2}\right)^2 [K\underline{y}_{n+1} + \nabla u_{n+1}] = M[\underline{y}_n + \Delta t\dot{\underline{y}}_n] - \left(\frac{\Delta t}{2}\right)^2 [K\underline{y}_n + \nabla u_n] \tag{5.256}$$

Using (5.254), equation (5.253) may be written

$$\begin{aligned} G(\underline{y}_{n+1}) = {} & \frac{1}{2}\left(\underline{y}_{n+1} - \underline{y}_n - \frac{\Delta t}{2}\dot{\underline{y}}_n\right)^T M\left(\underline{y}_{n+1} - \underline{y}_n - \frac{\Delta t}{2}\dot{\underline{y}}_n\right) \\ & + \left(\frac{\Delta t}{2}\right)^2 \left[\frac{1}{2}\underline{y}_{n+1}^T K\underline{y}_{n+1} + u_{n+1}\right] \\ & - \frac{1}{2}\dot{\underline{y}}_n^T M\dot{\underline{y}}_n - \left(\frac{\Delta t}{2}\right)^2 \left[\frac{1}{2}\underline{y}_n^T K\underline{y}_n + u_n\right] \equiv 0 \end{aligned} \tag{5.257}$$

Let us now construct the functional

$$\begin{aligned} F(\underline{y}_{n+1}) = {} & \frac{1}{2}\underline{y}_{n+1}^T M\underline{y}_{n+1} + \left(\frac{\Delta t}{2}\right)^2 \left(\frac{1}{2}\underline{y}_{n+1}^T K\underline{y}_{n+1} + u_{n+1}\right) \\ & - \underline{y}_{n+1}^T M[\underline{y}_n + \Delta t\dot{\underline{y}}_n] + \underline{y}_{n+1}^T [K\underline{y}_n + \nabla u_n]\left(\frac{\Delta t}{2}\right)^2 \end{aligned} \tag{5.258}$$

Then

$$\begin{aligned} \delta\underline{y}_{n+1}^T \frac{\partial F}{\partial \underline{y}_{n+1}} = \delta\underline{y}_{n+1}^T \Big[& M\underline{y}_{n+1} + \frac{\Delta t^2}{2}[K\underline{y}_{n+1} + \nabla u_{n+1}] \\ & - M[\underline{y}_n + \Delta t\dot{\underline{y}}_n] + \left(\frac{\Delta t}{2}\right)^2 [K\underline{y}_n + \nabla u_n]\Big] \end{aligned} \tag{5.259}$$

Thus necessary and sufficient conditions that equation (5.255) hold, are that (5.259) vanish for arbitrary "variations" $\delta \underline{y}_{n+1}$.

In order to force equations (5.252) to conserve energy we combine (5.250) and (5.257) through the use of a Lagrange multiplier λ; our new functional is

$$F(\underline{y}_{n+1})+\lambda G(\underline{y}_{n+1}) \tag{5.260}$$

Thus

$$\delta \underline{y}_{n+1}^T \left[\frac{\partial F}{\partial \underline{y}_{n+1}} + \lambda \frac{\partial G}{\partial \underline{y}_{n+1}}\right] + \delta\lambda\, G(\underline{y}_{n+1}) = 0 \tag{5.261}$$

If (5.261) is to vanish for arbitrary $\delta \underline{y}_{n+1}$, $\delta\lambda$ then equation (5.257) is satisfied and in addition:

$$M\underline{y}_{n+1} + \left(\frac{\Delta t}{2}\right)^2 [K\underline{y}_{n+1} + \nabla u_{n+1}] - M[\underline{y}_n + \Delta t \dot{\underline{y}}_n] + \left(\frac{\Delta t}{2}\right)^2 [K\underline{y}_{n+1} + \nabla u_n]$$
$$+ \lambda \left[M(\underline{y}_{n+1} - \underline{y}_n - \frac{\Delta t}{2}\dot{\underline{y}}_n) + \left(\frac{\Delta t}{2}\right)^2 [K\underline{y}_{n+1} + \nabla u_{n+1}] \right] \equiv 0 \tag{5.262}$$

Thus

$$(1+\lambda)\left[M\underline{y}_{n+1} + \left(\frac{\Delta t}{2}\right)^2 [K\underline{y}_{n+1} + \nabla u_{n+1})\right]$$
$$= (1+\lambda)[M\underline{y}_n] + \Delta t\left(1+\frac{\lambda}{2}\right) M\dot{\underline{y}}_n - \left(\frac{\Delta t}{2}\right)^2 (K\underline{y}_n + \nabla u_n) \tag{5.263}$$

The new algorithm thus consists of the two systems of equations, (5.261) and (5.263).

When $\underline{y}_{n+1}$ has been determined, $\dot{\underline{y}}_{n+1}$ can be calculated from equation (5.254).

The new algorithm is solved using a variant of the Newton-Raphson method.

If i denotes the iteration number then if we define

$$\underline{y}_{n+1}^{i+1} - \underline{y}_{n+1}^{i} = \Delta \underline{y}_{n+1}^{i} \quad ; \quad \lambda^{i+1} - \lambda^{i} = \Delta\lambda^{i} \tag{5.264}$$

we have

$$\begin{aligned}
&(\lambda^{i}+1)\left[M\Delta\underline{y}_{n+1}^{i} + \left(\frac{\Delta t}{2}\right)^{2} B(\underline{y}_{n+1}^{i})\Delta\underline{y}_{n+1}^{i}\right] \\
&+\Delta\lambda^{i}\left[M(\underline{y}_{n+1}^{i} - \underline{y}_{n}) + \left(\frac{\Delta t}{2}\right)^{2}(K\underline{y}_{n+1}^{i} + \nabla u_{n+1}^{i}) - \frac{\Delta t}{2} M\dot{\underline{y}}_{n}\right] \\
&+(\lambda^{i}+1)\left[M(\underline{y}_{n+1}^{i} - \underline{y}_{n}) + \left(\frac{\Delta t}{2}\right)(K\underline{y}_{n+1}^{i} + \nabla u_{n+1}^{i})\right] \\
&+\left(\frac{\Delta t}{2}\right)^{2}[K\underline{y}_{n} + \nabla u_{n}] - \left(1+\frac{\lambda^{i}}{2}\right)\dot{\underline{y}}_{n}\Delta t = 0
\end{aligned} \tag{5.265}$$

$$\begin{aligned}
&\Delta\underline{y}_{n+1}^{i\,T} M\left(\underline{y}_{n+1}^{i} - \underline{y}_{n} - \frac{\Delta t}{2}\dot{\underline{y}}_{n}\right) + \left(\frac{\Delta t}{2}\right)^{2}\Delta\underline{y}_{n+1}^{i\,T}[K\underline{y}_{n+1}^{i} + \nabla u_{n+1}^{i}] \\
&\qquad + G(\underline{y}_{n+1}^{i}) = 0
\end{aligned} \tag{5.266}$$

where

$$\left.\begin{aligned}
B(\underline{y}_{n+1}^{i}) &= K + J(\underline{y}_{n+1}^{i}) \\
J &= \frac{\partial}{\partial \underline{y}_{n+1}^{i}} \nabla u_{n+1}^{i}
\end{aligned}\right\} \tag{5.267}$$

Equations (5.266), (5.267) are of the form

$$A^{i}\underline{z}^{i} = \underline{b}^{i} \tag{5.268}$$

where

$$A^{i} = \left[\begin{array}{c|c} A_{11}^{i} & A_{12}^{i} \\ \hline A_{21}^{i} & 0 \end{array}\right]$$

$$\underline{z}^{i} = \left(\frac{\Delta\underline{y}_{n+1}^{i}}{\Delta\lambda^{i}}\right) \qquad \underline{b}^{i} = \left(\frac{\underline{b}_{1}^{i}}{b_{2}^{i}}\right) \tag{5.269}$$

where

$$\left.\begin{aligned}
A_n^i &= [1+\lambda^i][M + (\tfrac{\Delta t}{2})^2 \, B(\underline{y}_{n+1}^i)] \\
A_{12}^i &= [M(\underline{y}_{n+1}^i - \underline{y}_n - \tfrac{\Delta t}{2}\,\dot{y}_n) + (\tfrac{\Delta t}{2})^2[K\underline{y}_{n+1}^i + \nabla u_{n+1}^i)] \\
A_{21}^i &= (A_{12}^i)^T \\
\underline{b}_1^i &= (1+\lambda^i)[M(\underline{y}_{n+1}^i - \underline{y}_n) + (\tfrac{\Delta t}{2})^2(K\underline{y}_{n+1}^i + \nabla u_{n+1}^i)] \\
&\quad + (\tfrac{\Delta t}{2})^2 \, (K\underline{y}_n + \nabla u_n) - (1 + \tfrac{\lambda^i}{2})\dot{\underline{y}}_n \, \Delta t \\
\underline{b}_2^i &= G(\underline{y}_{n+1}^i)
\end{aligned}\right\} \tag{5.270}$$

Rewriting equation (5.268),

$$\left.\begin{aligned}
A_{21}^i \, \Delta\underline{y}_{n+1}^i &= b_2^i \\
Au^i \, \Delta\underline{y}_{n+1}^i + A_{12}^i \, \Delta\lambda^i &= \underline{b}_1^i
\end{aligned}\right\} \tag{5.271}$$

From the second equation in (5.271),

$$\Delta\underline{y}_{n+1}^i = (A_{11}^i)^{-1}(\underline{b}_1^i - A_{12}^i \, \Delta\lambda^i) \tag{5.272}$$

Using the first equation in (5.271),

$$A_{21}^i \, \Delta\underline{y}_{n+1} = A_{21}^i(A_{11}^i)^{-1} \underline{b}_1^i - A_{21}^i(A_{11}^i)^{-1} A_{12}^i \, \Delta\lambda^i = b_2^i$$

$$\therefore \quad \Delta\lambda^i = [(A_{12}^i)^T(A_{11}^i)^{-1} \, \underline{b}_1^i - b_2^i] \Big/ (A_{12}^i)^T(A_{11}^i)^{-1} A_{12}^i \tag{5.273}$$

$$\Delta y_{n+1}^i = (A_{11}^i)^{-1}[b_1^i - \Delta\lambda^i A_{11}^i] \tag{5.274}$$

It should be observed that $\underline{y}_{n+1}^i$ and $\Delta\lambda^i$ can be obtained with only one factorization of A_{11}^i and two forward reductions/back substitutions. Thus one additional forward reduction/back substitution is required when compared with the Newton-Raphson implementation of the trapezoidal algorithm. Thus, unlike algorithms A and B, algorithm C is readily computable.

Accuracy

Using equations (5.252) and (5.263), the new algorithm may be rewritten as:

$$\left.\begin{aligned} \underline{y}_{n+1} - \underline{y}_n &= \frac{\Delta t}{2}[\dot{\underline{y}}_{n+1} + \dot{\underline{y}}_n] \\ \dot{\underline{y}}_{n+1} - \dot{\underline{y}}_n &= -\frac{\Delta t}{2} M^{-1}[K(\underline{y}_{n+1} + \underline{y}_n) + \nabla u_{n+1} + \nabla u_n] \\ &\quad + \frac{\lambda}{1+\lambda}[M^{-1}(K\underline{y}_n + \nabla u_n)\frac{\Delta t}{2} - \dot{\underline{y}}_n] \\ \frac{1}{2}\dot{\underline{y}}_{n+1}^T M\dot{\underline{y}}_{n+1} + \frac{1}{2}[\underline{y}_{n+1}^T K\underline{y}_{n+1} + 2u_{n+1}] &\\ = \frac{1}{2}\dot{\underline{y}}_n^T M\dot{\underline{y}}_n + \frac{1}{2}[\underline{y}_n^T K\underline{y}_n + 2u_n] &\end{aligned}\right\} \tag{5.275}$$

From the first two equations

$$\frac{1}{2}[\dot{\underline{y}}_{n+1}^T M\dot{\underline{y}}_{n+1} - \dot{\underline{y}}_n^T M\dot{\underline{y}}_n] + \frac{1}{2}[\underline{y}_{n+1}^T K\underline{y}_{n+1} - \underline{y}_n^T K\underline{y}_n]$$
$$+ \frac{1}{2}(\underline{y}_{n+1} - \underline{y}_n)^T(\nabla u_{n+1} + \nabla u_n) = \frac{\lambda}{1+\lambda}(\underline{y}_{n+1} - \underline{y}_n)^T[K\underline{y}_n + \nabla u_n - \frac{2}{\Delta t}M\dot{\underline{y}}_n] \tag{5.276}$$

Using the third equation of (5.275)

$$\frac{\lambda}{1+\lambda}(\underline{y}_{n+1} - \underline{y}_n)^T[K\underline{y}_n + \nabla u_n - \frac{2}{\Delta t}M\dot{\underline{y}}_n]$$
$$= \frac{1}{2}(\underline{y}_{n+1} - \underline{y}_n)^T(\nabla u_{n+1} + \nabla u_n) + u_n - u_{n+1} \tag{5.277}$$

Thus

$$\frac{\lambda}{1+\lambda}\left((K\underline{y}_n + \nabla u_n)\frac{\Delta t}{2} - M\dot{\underline{y}}_n\right)$$

$$= C(\underline{y}_{n+1} - \underline{y}_n) \qquad (5.278)$$

where

$$\left.\begin{array}{ll} C = \frac{1}{2}[J(\xi_1) - J(\xi_2)] & \\ J = \frac{\partial}{\partial \underline{y}} \nabla u & \xi_i = \underline{y}_n\theta_i + \underline{y}_{n+1}(1-\theta_i) \\ & 0 \leq \theta_i \leq 1,\ i = 1,2 \end{array}\right\} \qquad (5.279)$$

Hence, equations (5.275) may be rewritten

$$\left.\begin{array}{l} \underline{y}_{n+1} - \underline{y}_n = \frac{\Delta t}{2}[\dot{\underline{y}}_{n+1} + \dot{\underline{y}}_n] \\ \dot{\underline{y}}_{n+1} - \dot{\underline{y}}_n = -\frac{\Delta t}{2} M^{-1}[K(\underline{y}_{n+1} + \underline{y}_n) + \nabla u_{n+1} + \nabla u_n] \\ \qquad + \frac{\Delta t}{2} M^{-1} C(\underline{y}_{n+1} - \underline{y}_n) \end{array}\right\} \qquad (5.280)$$

If u has continuous first and second partial derivatives, then $||C|| \leq C_0$ is bounded and we may apply standard techniques to (5.280) and show that:

$$||\underline{e}_r|| = ||\dot{\underline{y}}_n - \dot{\underline{x}}_n|| + ||\underline{y}_n - \underline{x}_n|| \leq D\,\Delta t^2\,, \qquad \text{as} \quad \Delta t \to 0 \qquad (5.281)$$

Thus algorithm C is also second order accurate as $\Delta t \to 0$.

The Lagrange multiplier technique is clearly superior to the other techniques, and while in the present analysis the constraint was that of conservation of energy, the technique can be used with any appropriate constraint. For example, if the technique is applied to equation (5.238) and we wish to ensure that the solutions of the discrete equations will be ultimately bounded, given that the solutions of (5.238) are ultimately bounded, the appropriate constraint would be the discrete form of (5.241) obtained by integrating from t_n to $t_n + \Delta t$, i.e.,

$$v_{n+1} - v_n = -\frac{\Delta t}{2}[\dot{\underline{y}}_{n+1}^T(C - zM)\dot{\underline{y}}_{n+1} + \dot{\underline{y}}_n^T(C - zM)\dot{\underline{y}}_n$$

$$+ z(\underline{y}_{n+1}^T K\underline{y}_{n+1} + \underline{y}_n^T K\underline{y}_n$$

$$+ \underline{y}_{n+1}^T \nabla u_{n+1} + y_n^T \nabla u_n)$$

$$+ (\dot{y}_{n+1} + z\underline{y}_{n+1})^T \underline{p}_{n+1} + (\dot{\underline{y}}_n + z\underline{y}_n)^T \underline{p}_n] \qquad (5.282)$$

Replacing $\dot{\underline{y}}_{n+1}$ by $\frac{2}{\Delta t}(\underline{y}_{n+1} - \underline{y}_n) - \dot{\underline{y}}_n$, the constraint equation becomes:

$$\frac{1}{2}[z\underline{y}_{n+1}^T(C - zM)\underline{y}_{n+1} + (\frac{2}{\Delta t}(\underline{y}_{n+1} - \underline{y}_n) - \dot{\underline{y}}_n)^T M(\frac{2}{t}(\underline{y}_{n+1} - \underline{y}_n) - \dot{\underline{y}}_n)$$

$$+ \underline{y}_{n+1}^T K\underline{y}_{n+1} + 2u(\underline{y}_{n+i})]$$

$$-\frac{1}{2}[z\underline{y}_n^T(C - zM)\underline{y}_n + (\dot{\underline{y}}_n + z\underline{y}_n)^T M(\dot{\underline{y}}_n + z\underline{y}_n)$$

$$+ \underline{y}_n^T K\underline{y}_n + 2u(\underline{y}_n)]$$

$$+\frac{\Delta t}{2}[(\frac{2}{\Delta t}(\underline{y}_{n+1} - \underline{y}_n) - \dot{\underline{y}}_n)^T (C - 2M)(\frac{2}{t}(\underline{y}_{n+1} - \underline{y}_n) - \dot{\underline{y}}_n)$$

$$+ \dot{\underline{y}}_n^T(C - zM)\dot{\underline{y}}_n + z(\underline{y}_{n+1}^T K\underline{y}_{n+1} + \underline{y}_n^T K\underline{y}_n)$$

$$+ z(\underline{y}_{n+1}^T \nabla u_{n+1} + \underline{y}_n^T \nabla u_n)$$

$$+ (\frac{2}{\Delta t}(\underline{y}_{n+1} - \underline{y}_n) - \dot{\underline{y}}_n)^T \underline{p}_{n+1} + (\dot{\underline{y}}_n + z\underline{y}_n)^T \underline{p}_n]$$

$$= G(\underline{y}_{n+1}) \equiv 0 \qquad (5.283)$$

The functional $F(\underline{y}_{n+1})$ of equation (5.258) is replaced by:

$$F(\underline{y}_{n+1}) = \frac{1}{2}\,\underline{y}_{n+1}^T\,M\underline{y}_{n+1} + \left(\frac{\Delta t}{4}\right)\,y_{n+1}^T\,C\underline{y}_{n+1}$$

$$+ \left(\frac{\Delta t}{2}\right)^2 \left[\frac{1}{2}\,\underline{y}_{n+1}^T\,K\underline{y}_{n+1} + u_{n+1}\right]$$

$$- \underline{y}_{n+1}^T\left[M(\underline{y}_n + \Delta t\,\dot{\underline{y}}_n) + \frac{\Delta t}{2}\,C\underline{y}_n\right]$$

$$+ \left(\frac{\Delta t}{2}\right)^2\,\underline{y}_{n+1}^T\left[K\underline{y}_n + \nabla u_{n+1} - p_{n+1} - p_n\right] \qquad (5.284)$$

The variation of the functional $(F(\underline{y}_{n+1}) + \lambda G(\underline{y}_{n+1}))$ yields the new algorithm:

$$\frac{\partial F}{\partial \underline{y}_{n+1}} + \lambda\,\frac{\partial G}{\partial \underline{y}_{n+1}} = 0$$

$$G(\underline{y}_{n+1}) = 0 \qquad (5.285)$$

$$\dot{\underline{y}}_{n+1} = \frac{2}{\Delta t}\,(\underline{y}_{n+1} - \underline{y}_n) - \dot{\underline{y}}_n$$

It may easily be shown that (5.285) is second order accurate as $\Delta t \to 0$.

6. Application to the Dynamical Analysis of Large Space Vehicles

Consider the system of differential equations obtained by applying finite element techniques (or other techniques) to some complex space vehicle. The equations are likely to be of the form:

$$M\ddot{\underline{x}} + C\dot{x} + K\underline{x} + \nabla_{\underline{x}} u = \underline{p}(t) \tag{6.1}$$

M is an $N \times N$ symmetric positive definite matrix, C and K are $N \times N$ symmetric positive semi-definite matrices, and $u(\underline{x})$ is a positive semi-definite potential function. (In the present analysis we shall neglect terms in (6.1) which arise due to steady rotation of the vehicle.)

Since one of the primary objectives of any structural analysis is to determine the stresses in the vehicle, it is desirable to make N, the number of coordinates, as large as possible so that stresses may be determined accurately. The number, R, of modes of the structure exhibiting significant response is usually much smaller than N. This poses a serious difficulty for the direct numerical integration of (6.1), since, as we know, the accuracy of any numerical scheme is determined not by the time step Δt, but by $(\omega_j \Delta t)$ where ω_j is the highest "frequency" which can be excited. In order that the higher modes be integrated accurately, Δt may have to be very small, much smaller than is either practical or economically feasible. In the case of linear systems, the use of algorithmic damping or post-filtering successfully overcomes this difficulty by suppressing the higher modes, which are inaccurately integrated when a moderately small value of Δt is used, thus resulting in reasonably accurate representation of the lower modes.

Many analysts use algorithmic damping or post-filtering to achieve the same result for nonlinear systems. Care must be exercised when doing this, since nonlinear systems can exhibit internal resonance, a phenomenon in which higher modes, though not excited by external forces, can be excited by nonlinear coupling to the lower modes. To illustrate this phenomenon, consider the following problem:

$$\left.\begin{aligned} &\ddot{x}_1 + 2z_1\dot{x}_1 + x_1 + \mu(x_1-x_2)^3 = P_1\cos\omega t + P_2\sin\omega t + P_3\cos 3\omega t \\ &\ddot{x}_2 + 2z_2\dot{x}_2 + 9x_2 + \mu(x_2-x_1)^3 = 0 \end{aligned}\right\} \quad (6.2)$$

where

$$\left.\begin{aligned} P_1 &= (1-\omega^2)A + \frac{3}{4}\mu A^3 \\ P_2 &= -2z_1\omega A \\ P_3 &= \frac{1}{4}\mu A^3 \end{aligned}\right\} \quad (6.3)$$

Since the second equation is not excited externally, it may seem reasonable to set $x_2 = 0$; the first equation then has the solution

$$x_1 = A\cos\omega t \quad (6.4)$$

If we now turn to the second equation, regarding x_2 as small,

$$\ddot{x}_2 + 2z_2\dot{x}_2 + 9\,x_2 \doteq \mu x_1^3 = \frac{3\mu}{4}A^3\cos\omega t + \frac{\mu}{4}A^3\cos\,3\omega t \quad (6.5)$$

If $\omega \sim O(1)$, the first term on the RHS of (6.5) causes no trouble, however, the second term will cause resonance, and if z_2 is small, will cause significant response in x_2. We therefore see that even though the second equation in (6.2) is not externally excited, it can still be driven by the first coordinate x_1. This is a simple example of internal resonance.

Quite frequently as a preliminary to performing a nonlinear dynamic analysis, a modal analysis of the linearized system will be carried out. The modal analysis can be a very useful tool in structuring the nonlinear problem for dynamical analysis.

As a preliminary, let us first put equations (6.1) into canonical form:

Let $$\underline{y} = M^{1/2}\,\underline{x} \tag{6.6}$$

Let $$\left.\begin{aligned} \mathcal{C} &= M^{-1/2} C M^{-1/2} \\ \mathcal{K} &= M^{-1/2}\,K M^{-1/2} \\ U(\underline{x}) &= V(\underline{y}) \\ M^{-1/2}\underline{p}(t) &= \underline{q}(t) \end{aligned}\right\} \tag{6.7}$$

Using (6.6) and (6.7) in (6.1)

$$\begin{aligned} &I\underline{\ddot{y}} + \mathcal{C}\underline{\dot{y}} + \mathcal{K}\underline{y} + \nabla_{\underline{y}}V(\underline{y}) = \underline{q}(t) \\ &\underline{y}(0) = \underline{a} \qquad \underline{\dot{y}}(0) = \underline{b} \end{aligned} \tag{6.8}$$

Let T be the orthogonal matrix which diagonalizes $\mathcal{K}$.

Let $$\begin{aligned} \underline{y} &= T\,\underline{z} \\ T^T\mathcal{K}T &= \Lambda \\ V(\underline{y}) &= W(\underline{z}) \\ \mathcal{D} &= T^T\mathcal{C}T \\ T^T\underline{q} &= \underline{f}(t) = \{f_i(t)\} \end{aligned} \tag{6.9}$$

Unless (6.1) has classical normal modes, $\mathcal{D}$ is not diagonal. Using (6.9) in (6.8),

$$\begin{aligned} & I\,\underline{\ddot{z}} + \mathcal{D}\underline{\dot{z}} + \Lambda\,\underline{z} + \nabla_{\underline{z}}\,W(\underline{z}) = \underline{f}(t) \\ & \underline{z}(0) = \underline{\tilde{a}} \qquad \underline{\dot{z}} = \underline{\tilde{b}} \end{aligned} \tag{6.10}$$

Suppose that $|f_j(t)| < \varepsilon$ for $j > P$, $\varepsilon << 1$.

Let

$$\underline{z}_p = \begin{Bmatrix} z_1 \\ z_2 \\ \vdots \\ z_p \end{Bmatrix} \tag{6.11}$$

If we suppose that the j^{th} modes, $j > P$ are at most weakly excited, let us set $z_j = 0$, $j \in (P+1,N)$. Then,

$$W(\underline{z}) = W(\underline{z}_p) \tag{6.12}$$

equation (6.10) becomes

$$\left.\begin{aligned} & I\,\underline{\ddot{z}}_1 = \underline{f}_1(t) \\ & I\,\underline{\ddot{z}}_2 = \mathcal{D}_{22}\underline{\dot{z}}_2 + \Lambda_2\underline{z}_2 + \nabla_{\underline{z}_2}\,W(\underline{z}_2) = \underline{f}_2(t) \end{aligned}\right\} \tag{6.13}$$

where

$$\underline{z}_1 = \begin{Bmatrix} z_1 \\ z_2 \\ \vdots \\ z_6 \end{Bmatrix}$$ represents the rigid body modes

$$\underline{z}_2 = \begin{Bmatrix} z_7 \\ z_8 \\ \vdots \\ z_p \end{Bmatrix}$$ represents the first (P-6) flexible modes.

$\mathcal{D}_{22} = \mathcal{D}_{22}^T$ is the (P-6) x (P-6) damping metric associated with the $\underline{z}_2$ modes.

$$\Lambda_2 = \begin{Bmatrix} \lambda_7 \\ \lambda_8 \\ \vdots \\ \lambda_p \end{Bmatrix} = \begin{Bmatrix} \omega_7^2 \\ \omega_8^2 \\ \vdots \\ \omega_p^2 \end{Bmatrix} \tag{6.14}$$

If P is not too large, we can select a Δt such that $(\omega_p \Delta t) \sim 0.1$, then using any of the "energy conserving" altorithms of Section 5, equations (6.13) can be integrated with good accuracy. Having determined $\underline{z}_p$, one can easily compute the physical coordinates,

$$\underline{x}_p = M^{-1/2}\, T_p\, \underline{z}_p$$

where T_p is the N x P matrix having as its columns the first P eigenvectors of the linearized problem. To check if there is any significant response in the neglected modes, due perhaps to internal resonance, we approximate the remaining modes by the system of uncoupled equations

$$\ddot{z}_j + \mathcal{D}_{jj}\dot{z}_j + \lambda_j z_j = -\left.\frac{\partial W(z)}{\partial z_j}\right|_{\underline{z}=\underline{z}_2} \tag{6.15}$$

$$j \in (P+1, N)$$

We note in passing that equations (6.15) will be exact if the linearized part of (6.1) has classical normal modes. Not all the modes in (6.15) need be examined; only those for which

$$\left|\lambda_j - \frac{p\lambda_k}{q}\right| \ll 1 \tag{6.16}$$

where p and q are integers and λ_k is an element of Λ_2. If no mode of (6.15) shows significant behavior, we can be reasonably sure that the solution of equations (6.13) will give a reasonably accurate representation of the solution to equations (6.1).

If any mode of (6.15) shows significant behavior, we can be reasonably sure that the solution of equations (6.13) will not give an accurate representation of the solution to (6.1). In this case, the modes of (6.15) which show significant behavior must be included in the solution of the problem. This presents a serious problem in the general case, since we require that $(\omega_k \Delta t) \sim 0.1$ for accurate integration of the system. If only a few modes of (6.15) show significant behavior, it may be possible to treat the problem in an efficient manner.

Let $$\underline{z}_3 = \begin{pmatrix} z_k \\ z_\ell \\ \vdots \\ z_m \end{pmatrix} \tag{6.17}$$

where k, ℓ, m are the modes of (6.15) showing significant behavior.

Let $$\underline{z}_4 = \begin{pmatrix} \underline{z}_2 \\ \underline{z}_3 \end{pmatrix} \tag{6.18}$$

Equation (6.10) may, in this case, be written

$$\left.\begin{aligned} I\,\ddot{\underline{z}}_2 + \mathcal{D}_{22}\dot{z}_2 + \mathcal{D}_{23}\dot{\underline{z}}_3 + \nabla_{\underline{z}_2} W(\underline{z}_4) + \Lambda_2 \underline{z}_2 &= \underline{f}_2(t) \\ I\,\ddot{\underline{z}}_3 + \mathcal{D}_{23}^T\dot{z}_2 + \mathcal{D}_{33}\dot{\underline{z}}_3 + \nabla_{\underline{z}_3} W(\underline{z}_4) + \Lambda_3 \underline{z}_3 &= 0 \end{aligned}\right\} \tag{6.19}$$

This first set of equations is integrated using a Δt_1 appropriate to the highest eigenvalue in Λ_2. The second set of equations is integrated using a Δt_2 appropriate to the highest eigenvalue in Λ_3, say $\Delta t_2 = \frac{1}{K}\Delta t_1$, K an integer. The values of $\underline{z}_2$ appearing in the second set of equations can be obtained by interpolation from the solutions of the first set of equations.

Internal resonance occurs most often in systems where the eigenvalues of the linearized system are integrally related, and where the nonlinear system is subjected to a steady state single frequency excitation; fortunately these two situations do not appear to arise too frequently in the space vehicle problem. Nevertheless, such situations can arise, and the analyst should be aware of them.

Appendix 1 - Generalization of Theorem 8

Theorem. Given the linear difference equation

$$\underline{x}_{n+1} = A(n)\underline{x}_n \quad |A(n)| \neq 0, \; \|A(n)\| < \infty \qquad n > n_o \tag{A1}$$

then A1 is uniformly Liapunov asymptotically stable at $\underline{x} = 0$ iff there exists a bounded, symmetric, positive definite matrix P(n) such that,

$$\begin{aligned} &\text{i)} \quad P(n) = P^T(n) \quad \text{positive definite and bounded above \& below} \\ &\text{ii)} \quad A^T(n)\, P(n+1)\, A(n) - P(n) = -\theta(n) \\ &\text{iii)} \quad \theta(n) = \theta^T(n) \quad \text{positive definite} \\ &\text{iv)} \quad \|\theta(n)\| \leq M_2(n_o) \qquad \forall\, n > n_o \text{ and } \forall\, n_o \end{aligned} \tag{A2}$$

Proof Sufficiency

Suppose that there exists such a matrix P(n) satisfying A2

Let $$V_n = \underline{x}_n^T P(n)\underline{x}_n \tag{A3}$$

Since P(n) is positive definite and bounded,

i) $V_n > 0$

ii) $V_n \leq M_3\, \underline{x}_n^T\, \underline{x}_n \qquad M_3 < \infty$

$$V_{n+1} = \underline{x}_{n+1}^T\, P(n+1)\, \underline{x}_{n+1} \tag{A4}$$

Using (A1),

$$V_{n+1} = \underline{x}_n^T\, A^T(n)\, P(n+1)\, A(n)\underline{x}_n \tag{A5}$$

$$\therefore \quad \Delta V_n = (V_{n+1} - V_n) = \underline{x}_n^T(A^T(n)\, P(n+1)\, A(n) - P(n))\underline{x}_n \tag{A6}$$

Using (A2)

$$\Delta V_n = - \underline{x}_n^T \, \theta(n) \underline{x}_n < 0 \tag{A7}$$

$$\therefore \qquad V_{n+1} < V_n < V_{n-1} < \cdots < V_1 < V_o \tag{A8}$$

Since P(n) is bounded $\forall n$, V_n is finite if $\|x_n\|$ is, and since V_n is zero only if $\underline{x}_n = 0$, hence V_n and therefore $\|\underline{x}_n\|$ tends to zero as n tends to infinity. Since the result is independent of n_0, the trivial solution $\underline{x} = 0$ is therefore uniformly Liapunov asymptotically stable.

Necessity As in the proof of necessity for Theorem 8, it is easily shown that P(n) satisfies equation (4.53), thus:

$$P(n) = \sum_{j=n} \Phi(j,n)^T \, \theta(j) \, \Phi(j,n) \tag{A9}$$

Thus if (A1) is uniformly asymptotically stable at the origin

$$\|\Phi(j,n)\| \le M_1 \, \delta^{(j-n)}$$

$$0 < \delta < 1 \quad , \qquad j > n \, , \qquad \forall n > n_0 \tag{A10}$$

Using (A2iv)

$$\text{i)} \quad \|P(n)\| \le \sum_{j=n}^{\infty} M_1^2 \, M_2(n_o) \, \delta^{2(j-n)} \tag{A11}$$

$$\le \frac{M_1^2 \, M_2(n_o)}{1 - \delta^2} < \infty \tag{A12}$$

$$\text{ii)} \quad P^T(n) = \left(\sum_{j=n}^{\infty} \Phi^T(j,n) \, \theta(j) \, \Phi(j,n)\right)^T = P(n) \tag{A13}$$

iii) $$\underline{x}^T P(n)\underline{x} = \sum_{j=n}^{\infty} (\Phi(j,n)\underline{x})^T \theta(j)(\Phi(j,n)\underline{x}) > 0$$

if $\Phi(j,n)\underline{x} \neq 0$

and since $|\Phi(j,n)| \neq 0$,

$$\underline{x}^T P(n)\underline{x} > 0 \qquad \underline{x} \neq 0 \tag{A14}$$

Thus if (A1) is uniformly L.A.S. at $\underline{x} = 0$, there exists a matrix P(n), symmetric, positive definite, and bounded which satisfies (A2ii).

Note: It is clear that P(n) is difficult to compute, except through the use of (A 9), which requires the unknown transition matrix $\Phi(j,n)$. Its main use is in proving theorems.

References

1. Hughes, T. J. R. "Stability, Convergence, and Growth and Decay of Energy of Average Acceleration Method in Nonlinear Structural Dynamics"Computers and Structures, Vol. 6, 1976, pp. 313-324.

2. Westermo, B. D. "The Solutions of a Nonlinear Difference Equation Found in Numerical Analysis" Ph. D. Thesis, Calif. Inst. of Tech., June 1978. Also Report DYNL-78-01, April 1978.

3. Hughes, T. J. R., Caughey, T. K. and Liu, W. K. "Finite-Element Methods for Nonlinear Elastodynamics which Conserve Energy", Trans. A.S. M. E. Jour. of Applied Mechanics, pp. 366-370, Vol. 45, June 1978.

Bibliography

1. Analysis of Numerical Methods — Eugene Isaacson and Herbert Bishop Keller, John Wiley & Sons, Inc. New York 1966.

2. Point Mapping Stability — J. Bernussou, Pergamon Press, Oxford 1977.

3. The Stability of Dynamical Systems, J. P. LaSalle and Z. Arstein, SIAM, Philadelphia 1976.