General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)
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
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

1. Introduction  
   
2. Existence and Uniqueness of the Initial Value Problem  
   (a) Explicit nonlinear difference equations  
   (b) Implicit nonlinear difference equations  
3. Properties of Linear Difference Equations  
   (a) Difference equations with constant coefficients  
   (b) Difference equations with variable coefficients  
   (c) Difference equations with periodic coefficients  
4. Stability of Difference Equations  
   (a) Stability of linear difference equations  
      (i) Linear difference equations with constant coefficients  
      (ii) Linear difference equations with periodic coefficients  
      (iii) Linear difference equations with variable coefficients  
   (b) Stability of non-linear difference equations  
      (i) Stability of explicit nonlinear difference equations  
      (ii) Stability of implicit nonlinear difference equations  

Preceding Page Blank Not Filmed
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.</td>
<td>Differential Equations and Difference Equations</td>
<td>46</td>
</tr>
<tr>
<td>(a)</td>
<td>Numerical Solution of Ordinary Differential Equations</td>
<td>46</td>
</tr>
<tr>
<td>(b)</td>
<td>Numerical Solution of Linear Ordinary Differential Equations</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>Methods Proposed for Suppressing Higher Modes</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>(i) Use of viscous damping</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>(ii) Use of algorithmic damping</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>(iii) Use of temporal filtering</td>
<td>57</td>
</tr>
<tr>
<td>(c)</td>
<td>Numerical Solution of Nonlinear Ordinary Differential Equations</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Application</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>Stability of Periodic Solutions</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>Globally Unstable Solutions</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Algorithms Which Conserve Energy</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>Effect of Viscous Damping</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Effect of Viscous Damping and Additive Forces</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>Extension of Energy Conserving Algorithms to multidegree-Freedom nonlinear systems</td>
<td>87</td>
</tr>
<tr>
<td>6.</td>
<td>Application to the Dynamic Analysis of Large Space Vehicles</td>
<td>103</td>
</tr>
<tr>
<td>Appendix</td>
<td>Generalization of Theorem 8</td>
<td>110</td>
</tr>
<tr>
<td>References</td>
<td>113</td>
<td></td>
</tr>
<tr>
<td>Bibliography</td>
<td>113</td>
<td></td>
</tr>
</tbody>
</table>
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:

\[
\begin{align*}
  x_{n+1} &= f(x_n, n) \\
  x_0 &= c
\end{align*}
\]  

(1.1)

and

\[
\begin{align*}
  x_{n+1} &= f(x_n, x_{n+1}, n) \\
  x_0 &= c
\end{align*}
\]  

(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)

\[
f(x_n, n) = A(n)x_n + g(n)
\]  

(1.3)

then (1.1) becomes

\[
\begin{align*}
  x_{n+1} &= A(n)x_n + g(n) \\
  x_0 &= c
\end{align*}
\]  

(1.4)

Similarly, if in (1.2)

\[
f(x_n, x_{n+1}, n) = A(n)x_n + B(n)x_{n+1} + g(n)
\]  

(1.5)

Then (1.2) becomes

\[
\begin{align*}
  x_{n+1} &= A(n)x_n + B(n)x_{n+1} + g(n) \\
  x &= c
\end{align*}
\]  

(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:

\[
\begin{align*}
    x_{n+1} &= C(n)x_n + h(n) \\
    x_0 &= c \\
    C(n) &= [I - B(n)]^{-1}A(n) \\
    h(n) &= [I - B(n)]^{-1}g(n)
\end{align*}
\]

(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
\[ \begin{align*}
\frac{x_{n+1}}{x_n} &= f\left(x_n\right), \\
x_0 &= c
\end{align*} \tag{2.1} \]

If

(i) \( \forall x, f(x) \) is continuous in \( x \), therefore \( ||f(x)|| < \infty, \forall ||x|| < \infty \)

(ii) \( ||c|| < \infty \).

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

**Proof**

Since \( ||c|| < \infty \)

\[ \begin{align*}
||x_1|| &\leq ||f(c)|| < \infty \\
||x_2|| &\leq ||f(x_1)|| < \infty \\
&\vdots \\
| |x_n|| &\leq ||f(x_{n-1})|| < \infty
\end{align*} \]

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

\[ \begin{align*}
\frac{x_{n+1}}{x_n} &= f\left(x_n, x_{n+1}\right), \\
\frac{x_0}{c} &= c
\end{align*} \tag{2.2} \]
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 \( f(x_n, x_{n+1}) \), that there exists a unique local solution of (2.2) provided \( |c| \) is sufficiently small. In some special cases Eq. (2.2) may be inverted so that it is described by an explicit equation.

\[
\begin{align*}
  x_{n+1} &= \mathcal{F}(x_n) \\
  x_0 &= c
\end{align*}
\]  

(2.3)

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

\[
\begin{align*}
  x_{n+1} &= x_n + \epsilon f_1(x_n, x_{n+1}) \\
  x_0 &= c
\end{align*}
\]  

(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

\[
x = g(x)
\]  

(2.5)

and the iterative procedure

\[
x_{n+1} = g(x_n) \quad n = 0, 1, 2 \ldots
\]  

(2.6)

Then if \( g(x) \) satisfies the following conditions

(i) \[ ||g(x) - g(y)|| \leq \lambda ||x - y|| \] for \( \forall \ x, y \in S \) \( S: ||x-x^0|| \leq \rho \)

with \( 0 \leq \lambda \leq 1 \)  

(2.7)

(ii) There exists an \( x_0 \) \( \exists \) \[ ||g(x^0) - x^0|| \leq (1-\lambda)\rho \]  

(2.8)

then \( \forall \) iterates \( x_n \) satisfy the following conditions
(i) \[ |x_n - x_0| \leq \rho \]

(ii) \[ \lim_{n \to \infty} |x_n| = \alpha \quad \text{where} \quad \alpha = \tilde{g}(\alpha) \]  

(iii) \( \alpha \) is the only root of Eq. (2.5) in \( |x - x_0| \leq \rho \)

Proof

Since \( x_{n+1} = \tilde{g}(x_n) \)

\[ x_n = \tilde{g}(x_{n-1}) \]  

\[ \therefore \quad |x_{n+1} - x_n| = |\tilde{g}(x_n) - \tilde{g}(x_{n-1})| \]  

\[ \leq \lambda |x_n - x_{n-1}| \]  

if \( x_{n+1}, x_n \in S \)

Now \( |x_1 - x_0| \leq (1-\lambda)\rho < \rho \)

\[ \therefore \quad x_1 \in S \]

\[ |x_2 - x_1| \leq \lambda |x_1 - x_0| \leq \lambda(1-\lambda)\rho \]  

\[ \therefore \quad |x_2 - x_0| \leq |x_2 - x_1| + |x_1 - x_0| \]  

\[ \leq \lambda(1-\lambda) + (1-\lambda)\rho = (1-\lambda^2)\rho < \rho \]  

\[ \therefore \quad x_2 \in S \]

Suppose that \( x_0, x_1, \ldots, x_n \in S \)

then \[ |x_{n+1} - x_n| \leq \lambda |x_n - x_{n-1}| \]  

\[ \leq \lambda^n |x_1 - x_0| \]  

\[ \therefore \quad |x_{n+1} - x_0| \leq |x_{n+1} - x_n| + |x_n - x_{n-1}| + \ldots + |x_1 - x_0| \]  

\[ \leq (\lambda^n + \lambda^{n-1} + \ldots + \lambda + 1)(1-\lambda)\rho \]  

\[ \leq (1-\lambda^{n+1})\rho < \rho \]  

\[ \therefore \quad x_{n+1} \in S \]
\[
\begin{align*}
\lim_{n \to \infty} \|x_{n+1} - x_n\| &\leq \lim_{n \to \infty} \lambda^n p = 0 \\
\|x_{n+k} - x_n\| &\leq \|x_{n+h} - x_{n+h-1}\| + \cdots + \|x_{n+1} - x_n\| \\
&\leq (\lambda^{n+h-1} + \lambda^{n+h-2} + \cdots + \lambda^{n})(1 - \lambda)p \\
&\leq \lambda^n p
\end{align*}
\]

\(\therefore\) \(\lim_{n \to \infty} \|x_{n+k} - x_n\| = 0\) \(\text{(2.26)}\)

\(\therefore\) the sequence \(\{x_n\}\) is a Cauchy sequence and converges uniformly.

\(\therefore\) \(\lim_{n \to \infty} x_{n+1} = \alpha = \lim_{n \to \infty} g(x_n) = g(\lim_{n \to \infty} x_n) = g(\alpha)\) \(\text{(2.27)}\)

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

**Unicity**

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

\[
\begin{align*}
\alpha &= g(\alpha) \quad \alpha \in S \\
\beta &= g(\beta) \quad \beta \in S
\end{align*}
\]

\(\therefore\) \(\|\alpha - \beta\| = \|g(\alpha) - g(\beta)\| \leq \lambda \|\alpha - \beta\|\) \(\text{(2.30)}\)

\(\therefore\) \(\|\alpha - \beta\| (1 - \lambda) \leq 0 \quad \therefore \alpha = \beta\) \(\text{(2.31)}\)

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.
Theorem 3  Given the implicit nonlinear difference equation

\[ \frac{x_{n+1} - x_n}{\varepsilon} + z = f(x_n, x_{n+1}) \]

where

\[ x_0 = c \]

If

(i) \( f(x_n, x_{n+1}) \) is continuous in \( x_n \) and \( x_{n+1} \).

(ii) \( f(x_n, x_{n+1}) \) has continuous first partial derivatives with respect to \( x_{n+1} \).

(iii) \( |\varepsilon| \) is sufficiently small.

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

Proof  Let \( x = x_{n+1} = x_n + \varepsilon f(x_n, x_{n+1}) = g(x) \)

With hypothesis (i) and (ii), \( g(x) \) satisfies

\[ \|g(x) - g(y)\| = |\varepsilon| \|f(x_n, x) - f(x_n, y)\| \]

\[ \leq |\varepsilon| \|J(x)\| \cdot \|x - y\| \]

\[ \leq \lambda \|x - y\| \]  

where

\[ J(x) = \frac{f(x_n, x)}{x} \left|_{x = \bar{x}} \right. \]

\[ \bar{x} = \alpha x + (1 - \alpha)y \quad 0 < \alpha < 1 \]

For

\[ x, y \in \|x - x_n\| < \rho, \quad \|y - x_n\| < \rho \]

We can always choose \( |\varepsilon| \) sufficiently small so that
If we choose $x^0 = x_n$ as our initial iterate, then

$$||g(x^0) - x^0|| = ||ef(x_n, x_n)||$$

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

$$||g(x^0) - x^0|| \leq (1-\lambda)e$$

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 $x_n$, there exists a unique solution $x_{n+1}$ satisfying

$$x_{n+1} = x_n + \epsilon f(x_n, x_{n+1})$$

Thus, starting with $x_0 = x$ and $|\epsilon|$ fixed and sufficiently small, there exists a unique $x_1$, satisfying Eq. (2.32). If, using $x_1$ and the same value of $\epsilon$, conditions (2.37) and (2.39) are satisfied, then there exists a unique $x_2$ 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

\[ x_{n+1} = Ax_n + f_n \]
\[ x_0 = c \quad |A| \neq 0 \]

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

\[ x_1 = Ac + f_0 \]
\[ x_2 = Ax_1 + f_1 = A^2c + Af_0 + f_1 \]
\[ x_3 = Ax_2 + f_2 = A^3c + \sum_{i=0}^{2} A^{2-i}f_i \]
\[ \vdots \]
\[ x_{n+1} = A^{n+1}c + \sum_{i=0}^{n} A^{n-i}f_i \]

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

\[ X_{n+1} = AX_n \quad ; \quad X_0 = I \]

thus

\[ X_1 = A \]
\[ X_2 = A^2 \]
\[ \vdots \]
\[ X_n = A^n \]
Thus, the solution to Eq. (3.1) can be written

\[ x_{n+1} = x_{n+1} + \sum_{i=0}^{n} x_{n-i} f_i \]  

(3.10)

We note that \[ x_{n-i} = A^{-i} = A^{n+1} A^{-(i+1)} \]  

(3.11)

\[ x_{n+1} = x_{n+1} + \sum_{i=0}^{n} x_{n+1-i} f_i \]  

(3.12)

(b) **Difference Equations with Variable Coefficients**

Consider the linear difference equation

\[ x_{n+1} = A(n)x_n + f_n \]  

(3.13)

\[ x_0 = c \quad |A| \neq 0 \]

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

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

\[ x_1 = A(0)c + f_0 \]  

(3.14)

\[ x_2 = A(1)x_1 + f_1 = A(1)A(0)c + A(1)f_0 + f_1 \]  

(3.15)

\[ x_3 = A(2)x_2 + f_2 = A(2)A(1)A(0)c + A(2)A(1)f_0 + A(2)f_1 + f_2 \]  

(3.16)

\[ x_{n+1} = A(n)A(n-1) \ldots A(0) \left[ c + \sum_{i=0}^{n} \sum_{h=0}^{i} A(h)^{-1}f_i \right] \]  

(3.17)

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

\[ x_{n+1} = A(n)x_n \quad X_0 = I \]  

(3.18)
Thus
\[ X_1 = A(0) \quad (3.19) \]
\[ X_2 = A(1)A(0) \quad (3.20) \]
\[ \vdots \]
\[ X_{n-1} = A(n-1)A(n-2) \cdots A(0) = \prod_{i=0}^{n-1} A(i) \quad (3.21) \]

Thus, the solution to Eq. (3.13) can be written
\[ X_{n+1} = X_n + \sum_{i=0}^{n} X_{n+i+1} X_{i+1} \quad (3.22) \]

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
\[ X_{n+1} = A(n)X_n \]
\[ X_0 = I \quad (3.23) \]

where
\[ A(n+N) = A(n) \quad |A(n)| \neq 0 \quad \forall n \]
\[ \|A(n)\| < \alpha \quad \forall n \quad (3.24) \]

The principal matrix solution \[ X_n \] has the form
\[ X_n = Q(n)C^n \quad (3.25) \]
where $Q(n+N) = Q(n)$ is a periodic matrix and $C$ is a non-singular constant matrix.

Proof From (3.23)

\[ X_0 = I \]
\[ X_1 = A(0) \]
\[ X_2 = A(1)A(0) \]
\[ \vdots \]
\[ X_n = \prod_{i=0}^{N-1} A(i) \]

We note that $X_k = \prod_{i=0}^{k-1} A(i)$ is non-singular since $A(i)$ is non-singular for $i = 0, 1, \ldots, N-1$.

Similarly

\[ X_{N+1} = \prod_{i=0}^{N} A(i) = A(N) \prod_{i=0}^{N-1} A(i) = X_1X_N \]
\[ X_{N+2} = \prod_{i=0}^{N+1} A(i) = \prod_{i=0}^{N} A(i)X_N = X_2X_N \]
\[ X_{N+k} = \prod_{i=0}^{N+k-1} A(i) = \prod_{i=0}^{k-1} A(i)X_N = X_kX_N \]

Since $X_N$ is non-singular, we may write

\[ X_N = C^N, \quad C \text{ - a constant matrix} \]

Consider the matrix
\[ Q(n) = X_n C^{-n} \] (3.35)

Since
\[ X_0 = I, \quad Q(0) = I \] (3.36)

Thus
\[ Q^{(N+n)} = X_{N+n} C^{-(N+n)} \] (3.37)
\[ = X_n X_N C^{-N} C^{-n} \] (3.38)

But
\[ X_N C^{-N} = I \]
\[ \therefore Q^{(N+n)} = X_n C^{-n} = Q(n) \] (3.39)

\[ \therefore Q(n) \text{ is a periodic matrix with period } N. \]

Hence
\[ X_n = Q(n) C^n \] (3.40)

Thus the complete structure of \( X_n \) is known for all \( n \) 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
\[ X_n = A^n \] (3.41)

4. Stability of Difference Equations

Definition Liapunov Stability (L.S.)

Given the difference equation
where
\[ f(0, 0) = 0 \] (4.2)

The equilibrium solution \( x = 0 \) is said to be Liapunov stable if given any \( \delta > 0 \), there exists an \( \epsilon > 0 \), such that if \( \| x_0 \| < \epsilon \), then \( \| x_n \| < \delta \) for all \( n > 0 \).

Liapunov Asymptotic Stability (L.A.S.)

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

(a) Stability of Linear Difference Equations

(i) Linear Difference Equations with Constant Coefficients

Theorem 5 Given the difference equation

\[ x_{n+1} = Ax_n \] (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

\[ T \triangleq T^{-1}AT = \Lambda \]

where \( \Lambda \) is a diagonal matrix.

A has the representation

\[ A = T \Lambda T^{-1} \]

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

\[ X_n = A^n = (T \Lambda T^{-1})^n = T \Lambda^n T^{-1} \]

(4.4)

\[ \text{Sufficiency} \]

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

\[ \therefore X_n \text{ is bounded and remains bounded as } n \to \infty \]

\[ \therefore \|X_n\| \leq M < \infty \text{ , } \forall n \]

(4.5)

From (4.3)

\[ x_n = X_n x_0 \]

(4.6)

\[ \therefore \text{if } \|x_0\| < \varepsilon \]

(4.7)

\[ \therefore \|x_n\| \leq \|X_n\| \varepsilon \]

(4.8)

\[ \leq M \varepsilon \]

(4.9)

\[ \therefore \text{if } \varepsilon \leq \delta/M \]

\[ \|x_n\| \leq \delta \text{ } \forall n \]

(4.10)

\[ \therefore (4.3) \text{ is Liapunov stable at } x = 0. \]

\[ \text{Necessity} \]

If \( |\lambda_i(A)| > 1 \) for some \( i \), then \( \lambda_i^n(A) \) cannot remain bounded as
n \rightarrow \infty. Hence \( X_n \) cannot remain bounded as \( n \rightarrow \infty \).

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

\[
T^{-1}AT = \begin{bmatrix}
J_{\alpha_1} & & \\
& J_{\alpha_2} & \\
& & \ddots \\
& & & J_{\alpha_h}
\end{bmatrix} = J
\] (4.11)

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

\[
A = TJT^{-1}
\] (4.12)

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

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

where

\[
J^n = \begin{bmatrix}
J_{\alpha_1}^n & & \\
& J_{\alpha_2}^n & \\
& & \ddots \\
& & & J_{\alpha_h}^n
\end{bmatrix}
\] (4.14)

and
and

$$
J_{\alpha_i} = \begin{bmatrix}
\lambda_i^n & n\lambda_i^{n-1} & \frac{n(n-1)}{2} & \lambda_i^{n-2} & \ldots \\
\lambda_i^n & n\lambda_i^{n-1} & & \\
0 & \lambda_i^n & & \\
\vdots & & & \\
& & & & \lambda_i^n
\end{bmatrix}
$$

(4.15)

**Sufficiency** If \( |\lambda_i(A)| < 1 \), then \( J_{\alpha_i}^n \) remains bounded and tends to zero as \( n \to \infty \). \( \therefore X_n \) is bounded and tends to zero as \( n \to \infty \).

\[
\|X_n\| \leq M < \infty \forall n
\]

and

\[
\lim_{n \to \infty} \|X_n\| = 0
\]

(4.16)

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.

**Necessity** If \( |\lambda_i(A)| \geq 1 \) for some \( i \), then \( J_{\alpha_i}^n \) 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.

**Theorem 6 (Liapunov)** Given the difference equation

$$
X_{n+1} = AX_n
$$

(4.17)

\( A \) - a constant matrix

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

$$
A^T PA = P = -Q
$$

(4.18)
Proof Sufficiency

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

$$V_n = x_n^T P x_n$$  \hspace{1cm} (4.19)$$

Since $P$ is symmetric and positive definite $V_n$ is positive definite

$$V_{n+1} = x_{n+1}^T P x_{n+1}$$ \hspace{1cm} (4.20)

Using (4.17)

$$V_{n+1} = (Ax_n)^T P (Ax_n)$$ \hspace{1cm} (4.21)

$$= x_n^T (A^T P A) x_n$$ \hspace{1cm} (4.22)

$$\therefore \Delta V_n = V_{n+1} - V_n = x_n^T (A^T P A - P) x_n$$ \hspace{1cm} (4.23)

Using (4.13)

$$\Delta V_n = -x_n^T Q x_n < 0$$ \hspace{1cm} (4.24)

Thus

$$V_{n+1} < V_n < V_{n-1} < V_{n-2} \cdots < V_0$$ \hspace{1cm} (4.25)

Since $V_n$ vanishes only at the origin

$$\therefore V_n \rightarrow 0 \text{ as } n \rightarrow \infty \hspace{1cm} (4.26)$$

If we define

$$\|x_n\|_p = \sqrt[p]{V_n}$$ \hspace{1cm} (4.27)

we see that

$$\|x_n\|_p < \|x_0\|_p$$ \hspace{1cm} (4.28)
\[ \| x_0 \| \leq \epsilon = \delta \]  
\[ (4.29) \]

then

\[ \| x_n \| < \delta \quad \forall n \]  
\[ (4.30) \]

and

\[ \| x_n \| \to 0 \quad \text{as} \quad n \to \infty \]  
\[ (4.31) \]

\[ \therefore \text{Eq. (4.17)} \text{ is Liapunov asymptotically stable at the origin.} \]

**Necessity**

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

\[ (4.18) \quad A^T PA - P = -Q \]  
\[ (4.32) \]

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 \( A^2, A^3, \) etc., we obtain

\[ A^T PA - P = -Q \]
\[ A^2 T PA^2 - A^T PA = - A^T Q A \]
\[ A^3 T PA^3 - A^2 T PA^2 = - A^2 T Q A^2 \]
\[ \ldots \ldots \ldots \ldots \]
\[ A^n T PA^n - A^{n-1} T PA^{n-1} = - A^{n-1} T Q A^{n-1} \]  
\[ (4.33) \]

Adding, we obtain
\[(A^n)^T P A^n - P = - \sum_{i=0}^{n-1} (A^i)^T Q A^i \] (4.34)

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

\[ P = \sum_{i=0}^{\infty} (A^i)^T Q A^i \] (4.35)

We note that:

\[ P^T = \left( \sum_{i=0}^{\infty} (A^i)^T Q A^i \right)^T = \sum_{i=0}^{\infty} (A^i)^T Q^T A^i \] (4.36)

But \( Q^T = Q \)

\[ P^T = P \] (4.37)

\[ x^T P x = \sum_{i=0}^{\infty} (A^i x)^T Q (A^i x) \] (4.38)

But \( Q \) is positive definite

\[ (A^i x)^T Q (A^i x) > 0 \] (4.39)

provided \( A^i x \neq 0 \).

\( \therefore \) If

\[ |A| \neq 0 \quad |A^i| \neq 0 \quad \therefore \quad x \neq 0 \] (4.40)

\[ x^T P x > 0 \quad x \neq 0 \] (4.41)

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

\( \therefore \) 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 \((4.29)\) \(x \in \mathbb{R}\), the range space of \(A, P\) is positive definite.

(ii) **Linear Difference Equations with Periodic Coefficients**

**Theorem 7** Given the difference equation

\[
\begin{align*}
X_{n+1} &= A(n)X_n \\
A(n+N) &= A(n)
\end{align*}
\]

(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 \(x_0 = c\) is given by

\[
X_n = X_n c
\]

where

\[
X_n = Q(n)C^n
\]

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

If \(X_N\) is simple, there exists a similarity matrix \(T\) such that

\[
T^{-1}X_N T = \Lambda - \text{a diagonal matrix}
\]

\(X_n\) therefore has the representation
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 $x = 0$.

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

\[
X_{n+1} = A(n)X_n
\]  

(4.48)

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

\[
\begin{cases}
  i) & P(k+N) = P(k) = P^T(k) \text{ positive definite} \\
  ii) & A^T(k)P(k+1)A(k) = -Q(k) \\
  iii) & Q^T(k) = Q(k) = Q(k+N) \text{ positive definite}
\end{cases}
\]

(4.49)

**Proof.** SUFFICIENCY

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

\[
V_n = x_n^T P(n)x_n
\]

(4.50)

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

\[ V_{n+1} = x_n^T P(n+1)x_{n+1} \]  \hspace{1cm} (4.51)

Using (4.48)

\[ V_{n+1} = x_n^T A(n)P(n+1)A(n)x_n \]  \hspace{1cm} (4.52)

\[ \Delta V_n = V_{n+1} - V_n = x_n^T (A(n)P(n+1)A(n) - P(n))x_n \]  \hspace{1cm} (4.53)

Using (4.49') ii)

\[ \Delta V_n = -x_n^T Q(n)x_n < 0 \]  \hspace{1cm} (4.54)

Since \( V_n \) vanishes only at the origin

\[ \therefore \quad V_n \to 0 \quad \text{as} \quad n \to \infty \]  \hspace{1cm} (4.55)

\[ \therefore \quad \text{Equation (4.48) is Liapunov asymptotically stable at the origin.} \]

\[ \text{Necessity} \]

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

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

\[ A^T(k)P(k+1)A(k) - P(k) = -Q(k) \]  \hspace{1cm} (4.56)

similarly

\[ A^T(k+1)P(k+2)A(k+1) - P(k+1) = -Q(k+1) \]  \hspace{1cm} (4.57)

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

\[ (A(k+1)A(k))^T P(k+2)(A(k+1)A(k)) - A(k)^T P(k+1)A(k) = -A(k)^T Q(k+1)A(k) \]  \hspace{1cm} (4.58)
Similarly
\[ A^T(k+2)P(k+3)A(k+2) - P(k+2) = - Q(k+2) \]  \( (4.59) \)

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

\[
\begin{align*}
& (A(k+2)A(k+1)A(k))^T P(k+3)(A(k+2)A(k+1)A(k) \\
& - (A(k+1)A(k))^T P(k+2)(A(k+1)A(k)) \\
& = - (A(k+1)A(k))^T Q(k+2)A(k+1)A(k))
\end{align*}
\]  \( (4.60) \)

Repeating the procedure \(n\) times gives

\[
\begin{align*}
& \left( \prod_{i=0}^{n-1} A(k+i) \right)^T P(k+n) \left( \prod_{i=0}^{n-1} A(k+i) \right) - \left( \prod_{i=0}^{n-2} A(k+i) \right)^T P(k+n-1) \left( \prod_{i=0}^{n-2} A(k+i) \right) \\
& = - \left( \prod_{i=0}^{n-2} A(k+i) \right)^T Q(k+n-1) \left( \prod_{i=0}^{n-2} A(k+i) \right)
\end{align*}
\]  \( (4.61) \)

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

\[
\begin{align*}
& \left( \prod_{i=0}^{n-1} A(k+i) \right)^T P(k+n) \left( \prod_{i=0}^{n-1} A(k+i) \right) - P(k) \\
& = - \sum_{j=0}^{n-1} \left( \prod_{i=0}^{j-1} A(k+i) \right)^T Q(k+j) \left( \prod_{i=0}^{j-1} A(k+i) \right)
\end{align*}
\]  \( (4.62) \)

Now
\[
\prod_{i=0}^{n-1} A(K+i) = \Phi(k+n, k)
\]  \( (4.63) \)
where
\[
\phi(m,k) = X_m X_k^{-1}
\]  \hspace{1cm} (4.64)

satisfies the equation
\[
\phi(n+1,k) = A(n)\phi(n,k)
\]  \hspace{1cm} (4.65)
\[
\phi(k,k) = I
\]

Since \(\forall\) solutions of (4.48) tend to zero as \(n \to \infty\)
\[
\phi(m,k) \to 0 \text{ as } m \to \infty
\]  \hspace{1cm} (4.66)

Thus as \(n \to \infty\) Eq. (4.62) becomes
\[
\begin{align*}
P(k) &= \sum_{j=k}^{\infty} \phi(j,k)^T Q(j) \phi(j,k) \\
\end{align*}
\]

\begin{itemize}
\item i) \quad \sum_{j=k}^{\infty} \phi(j,k)^T Q(j) \phi(j,k) = P(k)
\item ii) \quad P(k+N) = \sum_{j=k+N}^{\infty} \phi(j,k+N)^T Q(j) \phi(j,k+N)
\item iii) \quad x^T P(h)x = \sum_{j=k}^{\infty} (\phi(j,k)x)^T Q(j) (\phi(j,k)x) > 0 \quad x \neq 0
\end{itemize}

\hspace{1cm} (4.67)

\hspace{1cm} (4.68)
Since $|\theta(j+k)| \neq 0$ if $|A(k)| \neq 0 \forall k : P(k)$ 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

$$x_{n+1} = A(n)x_n$$

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

$$A(n) = A_0(n) + B(n)$$

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

$$x_{n+1} = A_0(n)x_n + B(n)x_n \quad |A_0(n)| \neq 0$$

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

If

1) $\forall$ solutions of $x_{n+1} = A_0(n)x_n$ are bounded as $n \to \infty$

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

3) $\|x_0\| = \|c\| < \infty$

*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

\[
\theta(n) \leq C + \sum_{i=0}^{n-1} \psi(i) \theta(i)
\]

(4.72)

\[\theta, \psi, C \geq 0\]

Then

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

(4.73)

**Proof**

From (4.72)

\[
\frac{\theta(n) \psi(n)}{n-1} \leq \psi(n)
\]

(4.74)

\[C + \sum_{i=0}^{n-1} \psi(i) \theta(i)\]

\[
\therefore 1 + \frac{\theta(n) \psi(n)}{n-1} \leq [1 + \psi(n)]
\]

(4.75)

\[C + \sum_{i=0}^{n-1} \psi(i) \theta(i)\]

\[
\therefore [C + \sum_{i=0}^{n-1} \psi(i) \theta(i)] \leq [C + \sum_{i=0}^{n-1} \psi(i) \theta(i)][1 + \psi(n)]
\]

(4.76)

\[C + \sum_{i=0}^{n-1} \psi(i) \theta(i)\]

\[
\therefore [C + \sum_{i=0}^{n-1} \psi(i) \theta(i)] \leq [C + \sum_{i=0}^{n-2} \psi(i) \theta(i)][1 + \psi(n-1)]
\]

(4.77)

\[C + \sum_{i=0}^{n-1} \psi(i) \theta(i)\]

\[
\therefore [C + \sum_{i=0}^{n-1} \psi(i) \theta(i)] \leq [C + \sum_{i=0}^{n-1} \psi(i) \theta(i)][1 + \psi(n)][1 + \psi(n-1)]
\]

(4.78)

Hence

\[
[C + \sum_{i=0}^{n} \psi(i) \theta(i)] \leq C \prod_{i=0}^{n} [1 + \psi(i)]
\]

(4.79)
But \( \theta_{n+1} \leq C + \sum_{i=0}^{n} \psi(i)\theta(i) \)  \hspace{1cm} (4.80)

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

**Lemma 2**

The product series \( S_n \):

\[ S_n = \prod_{i=0}^{n} (1 + v_i) \]  \hspace{1cm} (4.82)

is convergent iff the series \( V_n \):

\[ V_n = \sum_{i=0}^{n} v_i \]  \hspace{1cm} (4.83)

is convergent.

**Proof:**

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

\[ L_n = \sum_{i=0}^{n} \ln(1 + v_i) \]  \hspace{1cm} (4.84)

is convergent. This follows immediately from the fact that

\[ S_n = \exp(S_n) = e^{L_n} \]  \hspace{1cm} (4.85)

\[ \therefore \text{If } \lim_{n \to \infty} (\ln L_n) = L, \text{ then } \lim_{n \to \infty} (S_n) = e^L = S\]  \hspace{1cm} (4.86)

2) We know that if \( S_n \) is convergent, \( v_n \to 0 \) as \( n \to \infty \). \( \therefore \) let \( N \) be such that for \( n \geq N \) \( v_n \leq \frac{1}{2} \).
Now
\[ \frac{1}{2} v_n = v_n \left(1 - \frac{1}{4} \frac{1}{1 - \frac{1}{2}}\right) = v_n \left(1 - \frac{1}{2^2} - \frac{1}{2^3} \ldots\right) \] (4.87)

But
\[ \ln(1 + v_n) = (v_n - \frac{v_n^2}{2} + \frac{v_n^2}{3} + \ldots) \] (4.88)
\[ = v_n \left(1 - \frac{v_n}{2} + \frac{v_n^2}{3} \ldots\right) \] (4.89)

If \( v_n \leq \frac{1}{2} \)

then
\[ \ln(1 + v_n) = v_n \left(1 - \frac{1}{2^2} + \frac{1}{12} - \ldots\right) \] (4.90)

\[ \frac{1}{2} v_n < \ln(1 + v_n) < v_n \left(1 + \frac{1}{2^2} + \frac{1}{2^3} + \ldots\right) \] (4.91)

\[ \frac{1}{2} v_n < \ln(1 + v_n) < \frac{3}{2} v_n \] (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)

\[ \sum_{i=0}^{\infty} \ln(1 + v_i) < \frac{3}{2} \sum_{i=0}^{\infty} v_i < \infty \] (4.94)

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

then \( \sum_{i=0}^{\infty} v_i < 2 \sum_{i=0}^{\infty} \ln(1 + v_i) < \infty \) (4.96)

Returning now to the proof of Theorem 9, using equation (3.22),
\[ x_n = X_n x_0 + \sum_{i=0}^{n-1} X_n x_{i+1}^{-1} B(i) x_i \] (4.97)
Taking norms of both sides of equation (4.97)

\[ \|x_n\| \leq \|x_n\| \|c\| + \sum_{i=0}^{n-1} \|x_n\| \|x_{i+1}^{-1}\| \|B(i)\| \|x_i\| \]  

But, by hypothesis i), \( \|x_n\| \leq M_1 \), \( \|x_n\| \|x_{i+1}^{-1}\| \leq M_2 \) \( \forall i,n \)

\[ \therefore \|x_n\| \leq M_1 \|c\| + M_2 \sum_{i=0}^{k-1} \|B(i)\| \|x_i\| \]  

Using Lemma 1 with \( C = M_1 \|c\| \), \( \theta(i) = \|x_i\| \) and

\[ \psi(i) = M_2 \|B(i)\| \]  

we have

\[ \|x_n\| \leq M_1 \|c\| \prod_{i=0}^{n-1} (1 + M_2 \|B(i)\|) \]  

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

\[ \therefore \prod_{i=0}^{\infty} (1 + M_2 \|B(i)\|) < d_0 < \infty \]

\[ \therefore \|x_n\| \leq M_1 d_0 \|c\| \forall n \]  

\[ \therefore \text{Using hypothesis iii) we see that} \]

\[ \|x_n\| < \infty \quad \forall n \]  

Thus proving the theorem.

**Theorem 10**

Given the linear difference equation

\[ x_{n+1} = A_0(n)x_n + B(n)x_n \quad |A_0(n)| \neq 0 \]  

\[ (4.104) \]
where \( A_0(n) \) is either a constant or a periodic matrix. If

i) \( A_0(n) \) is a stability matrix, i.e., solutions of \( x_{n+1} = A_0(n)x_n \) tend to zero as \( n \to \infty \)

ii) \( \|B(n)\| \leq b_0 \ \forall n \) and \( b_0 \) 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,
\[
x_n = X_n c + \sum_{i=0}^{n-1} X_n X_{i+1}^{-1} B(i)x_i
\]

Taking norms of both sides of equation (4.105)
\[
\|x_n\| \leq \|x_n\|\|c\| + \sum_{i=0}^{n-1} \|X_n\| \|X_{i+1}^{-1}\|\|B(i)\|\|x_i\|
\]

Using hypothesis i)
\[
\|X_n\| \leq M_1 \delta^n, \quad \delta < 1
\]
\[
\|X_n\| \|X_{i+1}^{-1}\| \leq M_2 \delta^{n-1}
\]

Using hypothesis ii) equation (4.106) becomes
\[
\|x_n\| \leq M_1 \|c\| \delta^n + M_2 b_0 \sum_{i=0}^{n-1} \delta^{n-i-1} \|x_i\|
\]

Multiplying both sides of equation (4.108) by \( \delta^{-n} \) and setting \( \theta(i) = \|x_i\| \delta^{-i} \)
\[
C = M_1 \|c\|; \ \psi(i) = M_2 b_0, \text{ and using Lemma 1}
\]
\[
\|x_n\| \delta^{-n} \leq M_1 \|c\| \prod_{i=0}^{n-1} \left(1 + \frac{M_2 b_0}{\delta}\right)
\]
\[
\|x_n\| \leq M_1 \|c\| \delta^n \prod_{i=0}^{n-1} \left(1 + \frac{M_2 b_0}{\delta}\right) = M_1 \|c\| \frac{\delta^n}{\delta} \left(\delta + M_2 b_0\right)^n
\]
-33-

Hence, if $b_0$ is sufficiently small,

$$\delta + M_2 b_0 < 1$$  \hfill (4.111)

$${\| x_n \|} \leq \frac{M_1 \| c \|}{\delta} \forall n$$  \hfill (4.112)

$${\| x_n \|} \to 0 \text{ as } n \to \infty$$

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

$$A_o^T(n) P(n+1) A_o(n) = P(n) = -Q(n)$$  \hfill (4.113)

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

Let $$V_n = x_n^T P(n) x_n$$  \hfill (4.114)

then $$V_{n+1} = x_{n+1}^T P(n+1) x_{n+1}$$  \hfill (4.115)

Using equation (4.90)

$$V_{n+1} = x_n^T (A_o^T(n) P(n+1) A_o(n)) x_n + x_n^T (B(n) P(n) A_o(n)) x_n$$

$$+ x_n^T (A_o(n) P(n+1) B(n)) x_n + x_n^T B(n)^T P(n+1) B(n) x_n$$  \hfill (4.116)

$$\Delta V_n = V_{n+1} - V_n = -x_n^T (A_o^T(n) P(n+1) A_o(n) - P(n)) x_n + x_n^T (S(n)) x_n$$  \hfill (4.117)
where \( S(n) = B^T(n) P(n+1) A_o(n) + A_o^T(n) P(n+1) B(n) + B^T(n) P(n+1) B(n) = S^T(n) \)

Using (4.112), equation (4.118) becomes

\[
\Delta V_n = -x_n^T Q(n)x_n + x_n^T S(n)x_n
\]

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

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

\[
\Delta V_n < 0
\]

\[
\therefore V_{n+1} < V_n < V_{n-1} \cdots < V_1 < V_0
\]

Since \( V_n \) is positive definite and vanishes only at the origin, therefore \( V_n \) and hence \( ||x_n|| \) tends to zero as \( n \to \infty \), and since \( V_n \) is bounded above by \( V_o \), \( ||x_n|| \) 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 \( r = \text{Min}_n \lambda(n) \), \( r \) is positive, since \( Q(n) \) is positive definite.

\( s = \text{Max}_n \mu(n) \), 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

\[
\Delta V_n \leq -x_n^T(r-s)x_n
\]

and hence by making \( b_o \) sufficiently small \( \Delta V_n \) 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

\[
\begin{align*}
    x_{n+1} &= A(n)x_n + f(x_n, n) \\
    x_0 &= c
\end{align*}
\]

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

If

1) \( A(n) \) is a stability matrix

2) \( \lim_{\|x\| \to 0} \|f(x, n)\| = 0 \quad \forall n \) \hspace{1cm} (4.123)

3) \( \|c\| \) is sufficiently small

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

\[
A(n)P(n+1)A(n) - P(n) = -Q(n)
\]

\[
Q(n) = Q^T(n) = Q(n+N) \text{ positive definite}
\]

Let

\[
V_n = x_n^T P(n)x_n
\]

\[
V_{n+1} = x_{n+1}^T P(n+1)x_{n+1}
\]

Making use of equations (4.122),

\[
V_{n+1} = x_n^T(A^T(n) P(n+1) A(n))x_n + x_n^T(A^T(n) P(n+1) f(x_n, n))
\]

\[
+ f(x_n, n)^T(P(n+1) A(n))x_n + f^T(x_n, n) P(n+1) f(x_n, n)
\]

\[
(4.127)
\]
Thus \( \Delta V_n = V_{n+1} - V_n \)
\[
= x_n^T(A^T(n) P(n+1) A(n) - P(n))x_n
+ x_n^T(A^T(n) P(n+1) f(x_n, n) + f^T(x_n, n)(P(n+1) A(n))x_n
+ f^T(x_n, n) P(n+1) f(x_n, n)
\]
\[(4.128)\]

Using (4.124) equation (4.128) becomes
\[
\Delta V_n = -x_n^T Q(n) x_n
+ x_n^T(A^T(n) P(n+1) f(x_n, n) + f^T(x_n, n)(P(n+1) A(n))x_n
+ f^T(x_n, n) P(n+1) f(x_n, n)
\]
\[(4.129)\]

Using hypothesis ii), \( \|f(x_n, n)\| \sim O(\|x\|^2) \) as \( \|x\| \to 0 \)

Hence
\[
a) x_n^T Q(n) x_n \sim O(\|x\|^2)
\]
\[
b) x_n^T(A^T(n) P(n+1) f(x_n, n) + f^T(x_n, n)(P(n+1) A(n))x_n \sim O(\|x_n\|^3)
\]
\[
c) f^T(x_n, n) P(n+1) f(x_n, n) \sim O(\|x_n\|^4)
\]
\[
as \|x_n\| \to 0
\]
\[(4.130)\]

Thus for \( \|x_n\| \) sufficiently small, the sign of \( \Delta V_n \) is that of the first term
\( \Delta V_n \) is negative definite. Hence,
\[
V_{n+1} < V_n < V_{n-1} < \cdots < V_1 < V_0
\]
\[(4.131)\]

Thus if \( \|x\| \) is sufficiently small,
\[
\Delta V_n < 0 , \quad \forall n
\]
\[(4.132)\]

and since \( V_n \) is positive definite and vanishes only at the origin, therefore \( V_n \to 0 \), and hence \( \|x_n\| \to 0 \) as \( n \to \infty \). Thus equation (4.122) is
Liapunov asymptotically stable at the origin.

**Theorem 12**

Given the nonlinear difference equation

\[ x_{n+1} = [A_0(n) + B(n)]x_n + f(x_n, n) \]  
\[ x_0 = c \]

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

If

1) \( A_0(n) \) is a stability matrix
2) \( \|B(n)\| \) is sufficiently small
3) \( \lim \frac{\|f(x, n)\|}{\|x\|} = 0 \quad \forall n \)
4) \( \|c\| \) is sufficiently small

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

\[ x_{n+1} = A(n)x_n + f(x_n, x_{n+1}, n) \]
\[ x_0 = c \]

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

\[ \lim_{\|x\|,\|y\| \to 0} \frac{\|f(x, y, n)\|}{\|x\| + \|y\|} = 0 \quad (4.136) \]

\[ \|c\| \text{ is sufficiently small} \]

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:

i) \( A(n)^T P(n+1) A(n) - P(n) = -Q(n) \)

ii) \( Q(n) = Q^T(n) = Q(n+n) \) positive definite

Let \( V_n = x_n^T P(n) x_n > 0 \quad x_n \neq 0 \quad (4.138) \)

\[ V_{n+1} = x_{n+1}^T P(n+1) x_{n+1} \quad (4.139) \]

Making use of (4.135),

\[ V_{n+1} = x_n^T (A^T(n) P(n+1) A(n)) x_n + x_n^T (A^T(n) P(n+1)) f(x_n, x_{n+1}, n) \]

\[ + x_n^T (x_n, x_{n+1}, n) (P(n+1) A(n)) x_n \]

\[ + x_n^T (x_n, x_{n+1}, n) (P(n+1)) f(x_n, x_{n+1}, n) \quad (4.140) \]

\[ \Delta V_n = V_{n+1} - V_n \]

\[ = x_n^T (A^T(n) P(n+1)) f(x_n, x_{n+1}, n) + x_n^T (A^T(n) P(n+1) A(n) - P(n)) x_n \]

\[ + x_n^T (x_n, x_{n+1}, n) (P(n+1) A(n)) x_n \]

\[ + x_n^T (x_n, x_{n+1}, n) (P(n+1)) f(x_n, x_{n+1}, n) \quad (4.141) \]
Making use of equation (4.137)

\[
\Delta V_n = -x_n^T Q(n)x_n + x_n^T (A^T(n)P(n+1)) f(x_n, x_{n+1}, n)
+ f^T(x_n, x_{n+1}, n)(P(n+1)A(n))x_n
+ f^T(x_n, x_{n+1}, n) P(n+1) f(x_n, x_{n+1}, n)
\]

(4.142)

From equation (4.135)

\[
\|x_{n+1}\| + \|x_n\| \leq \|I + A(n)\| \|x_n\| + \|f(x_n, x_{n+1}, n)\| \]

(4.143)

\[
\therefore \|x_{n+1}\| + \|x_n\| \leq M_1 \|x_n\| + \|f(x_n, x_{n+1}, n)\| \]

(4.144)

From (4.136ii)

\[
\|f(x_n, x_{n+1}, n)\| \leq M_2(\delta)(\|x_n\| + \|x_{n+1}\|)^2
\]

for \(\|x_n\| + \|x_{n+1}\| \leq \delta \)

(4.145)

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

From (4.144) and (4.145)

\[
\|x_{n+1}\| + \|x_n\| \leq \frac{M_1 \|x_n\|}{1 - M_2(\delta)\delta} = M_3 \|x_n\| < \delta
\]

(4.146)

if \(\|x_n\| < \frac{\delta}{M_3}\) is sufficiently small.

Thus, if \(\|x_n\|\) is sufficiently small, the first term in (4.142) is of order \(\|x_n\|^2\), while the second and third terms are of order \(\|x_n\|^3\), and the fourth term is of order \(\|x_n\|^4\). Hence, if \(\|x_n\|\) is sufficiently small,
the sign of \( \Delta V_n \) is that of the first term, therefore

\[
\Delta V_n < 0 \quad (4.147)
\]

Thus, if \( ||c|| \) is sufficiently small, \( \Delta V_n \) is negative for all \( n \), therefore

\[
V_{n+1} < V_n < V_{n-1} < \cdots < V_1 < V_0 \quad (4.148)
\]

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

Theorem 14 (Liapunov-Poincaré)

Given the nonlinear difference equations

\[
x_{n+1} = A(n)x_n + f(x_n, n)
\]

\[
x_0 = c
\]

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

If

\[
\begin{align*}
i) & \text{ there exists at least one unstable solution of the equation } x_{n+1} = A(n)x_n \\
\text{ii) } & \lim_{||x|| \to 0} \frac{||f(x, n)||}{||x||} = 0 \quad \forall n \\
\text{iii) } & ||c|| \text{ is sufficiently small}
\end{align*}
\]

Then there exist unstable solutions of equation (4.149).
Proof. Let
\[ R^N = x^{-1}_k x_{N+k} = x_N \]
\[ \theta_k = x_k R^{-k} \]
Let \[ x_k = \theta_k v_k \]
Substituting into equation (4.149)
\[ y_{n+1} = \theta_{n+1}^{-1} A(n) \theta_n y_n + \theta_{n+1}^{-1} f(\theta_n y_n, n) \]
\[ y_0 = \theta_0^{-1} c \]
Now
\[ \theta_{n+1}^{-1} A(n) \theta_n = R^{n+1} x_{n+1}^{-1} A(n) x_n R^{-n} \]
But \[ x_{n+1} = A(n) x_n \]
\[ \therefore \theta_{n+1}^{-1} A(n) \theta_n = R \]
\[ \therefore y_{n+1} = R y_n + g(y_n, n) \]
\[ y_0 = b ; \quad b = \theta_0^{-1} c ; \quad g(y_n, n) = \theta_{n+1}^{-1} f(\theta_n y_n, n) \]
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
\[ T^{-1}RT = A \]

where \[ |\lambda_i| > 1 \quad i \in (l,k) \]  \hspace{1cm} (4.157)
\[ |\lambda_j| < 1 \quad j \in (k+1,L) \]

Let \[ y_n = T z_n \]  \hspace{1cm} (4.158)

Then \[ z_{n+1} = \Lambda z_n + h(z_n, n) \]

\[ z_0 = d \]  \hspace{1cm} (4.159)

where \[ h = T^{-1}a(Tz_n, n) \]

\[ z_0 = d = T^{-1}b \]

Let \[ P = \begin{bmatrix} I_k & 0 \\ 0 & -I_{L-k} \end{bmatrix} \]  \hspace{1cm} (4.160)

\[ V_n = Z_n^* P Z_n \]

It will be observed that \[ V_n \] is sign indefinite.

\[ V_{n+1} = Z_{n+1}^* P Z_{n+1} \]  \hspace{1cm} (4.161)

Substituting from (4.145),

\[ V_{n+1} = Z_n^* \Lambda^* P \Lambda Z_n + h^*(Z_n, n)P \Lambda Z_n \]

\[ + Z_n^* \Lambda^* P h(Z_n, n) + h^*(Z_n, n)P h(Z_n, n) \]  \hspace{1cm} (4.162)
\[ \Delta V_n = V_{n+1} - V_n = z_n^* (A^* P A - P) z_n + (h^*(z_n, n) P A z_n + z_n^* A^* P h(z_n, n)) + h^*(z_n, n) P h(z_n, n) \]  

(4.163)

Now

\[ A^* P A - P = \begin{bmatrix} |\lambda_i|^2 - 1 & 0 \\ 0 & (1 - |\lambda_j|^2) \end{bmatrix} \]  

(4.164)

Since

\[ |\lambda_i| > 1 \quad i \in (1, k) \]

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

(4.165)

\[ A^* P A - P \] is positive definite Hermitian. Using (4.150)iii, the first term in (4.163) is positive and of order \( \|z_n\|^2 \), the second term is of order \( \|z_n\|^3 \), while the fourth term is of order \( \|z_n\|^4 \); thus for sufficiently small \( \|z_n\| \), the sign of \( \Delta V_n \) is that of the first term and is positive.

\[ \Delta V_n > 0 \quad \text{for} \quad \|z_n\| \text{sufficiently small.} \]  

(4.166)

Since \( V_n \) is sign indefinite we can define a set

\[ \Omega: V_n \geq 0; \quad \|z_n\| < \delta \]  

(4.167)

Clearly, the origin is a boundary point of \( \Omega \). In \( \Omega \), \( V_n > 0 \), \( \Delta V_n > 0 \), therefore starting in \( \Omega \), \( z_n \) cannot approach the origin. Since \( V_o > 0 \),
Only exit $\Omega$ through the boundary $\|z_n\| = \delta$; thus the system is unstable.

**Theorem 15**

Given the nonlinear implicit equation

$$x_{n+1} = A(n)x_n + f(x_n, x_{n+1}, n)$$

(4.168)

$$x_0 = c$$

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

If

1) there exists at least one unstable solution of the equation

$$x_{n+1} = A(n)x_n$$

2) \( \lim_{|x|, |y| \to 0} \frac{f(x, y, n)}{|x| + |y|} = 0 \quad \forall n \) (4.169)

3) $\|c\|$ 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

$$x_{n+1} = A(n)x_n + f(x_n, n)$$

(4.170)

$$x_0 = c$$
where \( A(n) \) is either a constant matrix or a periodic matrix,

If 

i) the principal matrix \( X_n \) of the linear difference equation

\[
x_{n+1} = A(n)x_n
\]

has an eigenvalue \( \pm 1 \), or a pair of complex conjugate eigenvalues of modulus unity

ii) \[
\lim_{\|x\| \to 0} \frac{\|f(x)\|}{\|x\|} = 0 \quad \forall n
\]

iii) \( \|c\| \) 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, \((A^*PA - P)\) 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 \((A^*PA - P)\) is only positive semidefinite, we see that the sign of \( \Delta V_n \) depends on the terms in \( h(x_n, n) \). 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

\[
\frac{dx}{dt} = A(t)x + f(x,t)
\]

\[
x(t_0) = c, \quad t_0 \leq t \leq T_a < \infty
\]

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

\[
y_{n+1} = B(n)y_n + g(y_n, y_{n+1}, n)
\]

\[
y_0 = c
\]

such that \( y_n = x(t_n) \), \( t_{n+1} = t_n + \Delta t \), \( n = 0, 1, 2, \ldots, M \).

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

1) They have a unique solution,

2) This solution, at least for sufficiently small \( \Delta t_n \), 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

\[
\frac{dx}{dt} = Ax + f(t) \quad 0 \leq t \leq T < \infty
\]

\[
x(0) = c \quad \text{A = a constant matrix}
\]

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

\[
\begin{align*}
Y_{n+1} &= Y_n + \frac{\Delta t}{2} A(Y_n + Y_{n+1}) + \frac{\Delta t}{2} (f_n + f_{n+1}) \\
Y_0 &= c \\
\Delta t &= T_0/M
\end{align*}
\]

Equation (5.4) may be written in explicit form:

\[
\begin{align*}
Y_{n+1} &= AY_n + B(f_n + f_{n+1}) \\
Y_0 &= c
\end{align*}
\]

where

\[
\begin{align*}
A &= [I - A \frac{\Delta t}{2}]^{-1} [I + A \frac{\Delta t}{2}] \\
B &= [I - A \frac{\Delta t}{2}]^{-1} \frac{\Delta t}{2}
\end{align*}
\]

Accuracy

Let \( x_{n+1} = x(n\Delta t) \)

Let \( e_{n+1} \) be the local truncation error defined by

\[
x_{n+1} = Ax_n + B(f_{n+1} + f_n) + e_{n+1} \Delta t
\]
Let \( e_n = (x_n - y_n) \) be the solution error \( (5.8) \)

Then subtracting (5.5) from (5.7)

\[
e_{n+1} = A e_n + I_{n+1} \Delta t
\]

Now \( e_0 = x_0 - y_0 = 0 \)

\[
\begin{align*}
\varepsilon_1 &= I_1 \Delta t \\
\varepsilon_2 &= (A_{11} + \tau_2) \Delta t \\
\vdots \\
\varepsilon_n &= (A_{n-1} I_1 + A_{n-2} I_2 + \cdots I_n) \Delta t
\end{align*}
\]

(5.10)

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

\[
T^{-1} A T = \Lambda \quad \text{and hence} \quad A = T \Lambda T^{-1}
\]

(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.

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

(5.12)

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

(5.13)

\[
\theta_j = (1 - \lambda_j \frac{\Delta t}{2})^{-1} (1 + \lambda_j \frac{\Delta t}{2})
\]

If \( \lambda_j \) is pure imaginary, say \( \lambda_j = i \omega_j \), then

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

(5.14)
If $\lambda_j$ is complex with negative real part, say $\lambda_j = -\omega_j \zeta_j + i \omega_j \sqrt{1 - \zeta_j^2}$ then

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

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

In either case we have

$$e_n = T[\theta^{n-1}\tau_{1} + \theta^{n-2}\tau_{2} + \ldots \tau_{n}] \Delta t \quad (5.16)$$

Let $T^{-1}\tau_{j} = b_{j}$

$$\|e_n\| \leq \|T\| \left[ \|\theta^{n-1}b_{1}\| + \|\theta^{n-2}b_{2}\| + \ldots \|b_{n}\| \right] \Delta t \quad (5.17)$$

But

$$\|\theta^{n-k}b_{k}\| = \sum_{j=1}^{N} \theta_j^{n-k}b_j^k \leq \sum_{j=1}^{N} b_j^k = \|b_{k}\| \quad (5.18)$$

$$\|b_{k}\| \leq \|T^{-1}\|\|\tau_{k}\| \quad (5.19)$$

If $\tau = \max_{k} \|\tau_{k}\| \quad (5.20)$

then

$$\|e_n\| \leq \|T\|\|T^{-1}\|n\Delta t \tau \quad (5.21)$$

but $n\Delta t = t_n \leq T \quad (5.22)$

$$\|e_n\| \leq \|T\|\|T^{-1}\|\tau \quad (5.23) \leq K_T \tau$$

Now
\[ x_{n+1} = x_n + \frac{dx_n}{dt} \Delta t + \frac{d^2x_n}{dt^2} \frac{\Delta t^2}{2} + \frac{d^3x_n}{dt^3} \frac{\Delta t^3}{6} \]

\[ f_{n+1} = f_n + \frac{df_n}{dt} \Delta t + \frac{d^2f_n}{dt^2} \frac{\Delta t^2}{2} + \frac{d^3f_n}{dt^3} \frac{\Delta t^3}{6} \]

Substituting into (5.7) and using (5.3)

\[ I_{n+1} = -[I - A \frac{\Delta t}{2}]^{-1} \left[ \frac{1}{12} A^3 x_n + \frac{1}{2} A^2 f_n + \frac{1}{4} A \frac{df_n}{dt} + \frac{1}{2} \frac{d^2f_n}{dt^2} \right] \Delta t^2 \] (5.25)

Then, provided \( \|x_n\|, \|f_n\|, \|\frac{df_n}{dt}\|, \|\frac{d^2f_n}{dt^2}\| \) are bounded,

\[ \|I_{n+1}\| \leq K_2 \Delta t^2 \]

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

\[ \therefore \|e_n\| \leq K_1 K_2 \Delta t^2 \quad \text{as} \quad \Delta t \to 0 \] (5.27)

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

**Application**

Consider the conservative dynamical system:

\[ M\ddot{x} + Kx = 0 \]

\[ x(0) = a, \dot{x}(0) = b \] (5.28)

where \( M = M^T \) is positive definite

\( K = K^T \) is positive definite

If \( z = \begin{pmatrix} x \\ \dot{x} \end{pmatrix} \), equation (5.28) may be written
\[
\frac{dz}{dt} = A z
\] (5.29)

\[
z(0) = c = \begin{pmatrix} a \\ b \end{pmatrix}
\]

where
\[
A = \begin{bmatrix}
0 & I \\
-M^{-1}K & 0
\end{bmatrix}
\] (5.30)

If \( A \) is simple, there exists a similarity matrix such that
\[
T^{-1} A T = \begin{bmatrix}
i \omega_1 & -i \omega_1 \\
i \omega_2 & -i \omega_2
\end{bmatrix}
\] (5.31)

The trapezoidal difference equation corresponding to (5.29) is

\[
\begin{align*}
w_{n+1} &= w_n + \frac{A t}{2} \ A(w_{n+1} + w_n) \\
w_n &= \begin{pmatrix} y_n \\ \dot{y}_n \end{pmatrix} \quad w_0 = c
\end{align*}
\] (5.32)

Alternatively,
\[
\begin{align*}
y_{n+1} &= y_n + \frac{A t}{2} \ (\dot{y}_{n+1} + \dot{y}_n) \\
\dot{y}_{n+1} &= \dot{y}_n - \frac{t}{2} \ M^{-1}K(y_{n+1} + y_n)
\end{align*}
\] (5.33)

From (5.33) we see that
\[
\frac{1}{2} \ \dot{y}_{n+1}^T \ M \ \dot{y}_{n+1} + \frac{1}{2} \ y_{n+1}^T K \ y_{n+1} = \text{constant}
\] (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
\[ \frac{1}{2} \dot{x}^T M \dot{x} + \frac{1}{2} x^T M x = \text{constant} \]  
(5.35)

From equation (5.32)
\[ w_{n+1} = A w_n \]  
(5.36)

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

Using (5.31)
\[ A = T [I - \frac{\Delta t}{2} \Lambda]^{-1} [I + \frac{\Delta t}{2} \Lambda] T^{-1} \]  
(5.38)

\[ \therefore \quad A = T \Theta T^{-1} \quad ; \quad \Theta = [\theta_j] \]  
(5.39)

\[ \theta_j = \frac{1 + i \frac{\Delta t}{2} \omega_j}{1 - i \frac{\Delta t}{2} \omega_j} \quad \forall j \quad (1,2N) \]

Hence \( |\lambda_j(A)| = 1 \quad \forall j \)  
(5.40)

Thus the eigenvalues of \( A \) 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:
\[ w_n = \begin{pmatrix} v_n \\ \dot{y}_n \end{pmatrix} = A^n c \]  
(5.41)

where \( A^n = T \Theta^n T^{-1} \)  
(5.42)
Let $\theta_j = \frac{1 + i \frac{\Delta t}{2} \omega_j}{1 - i \frac{\Delta t}{2} \omega_j} = e^{i \Delta \phi_j}$

$$\Delta \phi_j = \tan^{-1} \frac{\omega_j \Delta t}{1 - (\omega_j \frac{\Delta t}{2})^2}$$

Let $\Omega_j = \frac{\Delta \phi_j}{\Delta t}$

$$\psi_n = \begin{pmatrix} y_n \\ \dot{y}_n \end{pmatrix} = T \left[ e^{i \Omega_j t_n} \right]^{-1}$$

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

$$\mathbf{z}_n = \begin{pmatrix} x_n \\ \dot{x}_n \end{pmatrix} = T \left[ e^{i \omega_j t_n} \right]^{-1}$$

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.

**Period Error**

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

$$\epsilon_T = \frac{T_j^d - T_j}{T_j}$$

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

We note that
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 $\Delta t$ 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 $\Delta t$ 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 of 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

$$\begin{cases}
M \ddot{x} + C \dot{x} + K x = 0 \\
\dot{x}(0) = a \\
x(0) = b
\end{cases}$$

$$M = M^T > 0 ; \quad C = C^T \geq 0 ; \quad K = K^T > 0$$

If $Z = \begin{pmatrix} \frac{X}{\dot{X}} \end{pmatrix}$, equation (5.49) may be rewritten
The trapezoidal difference equation corresponding to (5.50) is:

\[
\frac{dz}{dt} = A \ z \\
\begin{bmatrix}
  z(0) = c = \frac{a}{b}
\end{bmatrix}
\]

\[\text{where } A = \begin{bmatrix}
  0 & 1 \\
  -M^{-1}K & -M^{-1}C
\end{bmatrix}\]

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

\[
T^{-1} AT = A = \begin{bmatrix}
  -\omega_1 \zeta_1 + i \omega_1 \sqrt{1 - \zeta_1^2} \\
  -\omega_1 \zeta_1 - i \omega_1 \sqrt{1 - \zeta_1^2} \\
  -w_n \zeta_n + i w_n \sqrt{1 - \zeta_n^2} \\
  -w_n \zeta_n - i w_n \sqrt{1 - \zeta_n^2}
\end{bmatrix}
\]

\[\text{The trapezoidal difference equation corresponding to (5.50) is:}\]

\[
\begin{bmatrix}
  w_{n+1} \\
  \nu_{n+1}
\end{bmatrix} = A \begin{bmatrix}
  w_n \\
  \nu_n
\end{bmatrix}
\]

\[\begin{bmatrix}
  w_0 = \underline{c}
\end{bmatrix}\]

\[\text{where } A = T \Theta T^{-1}\]

\[
\Theta_j = \begin{bmatrix}
  1 - \omega_j \zeta_j \frac{\Delta t}{2} + i \omega_j \frac{\Delta t}{2} \sqrt{1 - \zeta_j^2} \\
  1 + \omega_j \zeta_j \frac{\Delta t}{2} - i \omega_j \frac{\Delta t}{2} \sqrt{1 - \zeta_j^2}
\end{bmatrix}
\]

\[\Theta_j \text{ may be expressed as}\]
\[ \theta_j = \rho_j e^{i \Delta \phi_j} \]

\[ \rho_j = \left\{ \frac{1 - \omega_j \xi_j \Delta t + (\omega_j \frac{\Delta t}{2})^2}{1 + \omega_j \xi_j \Delta t + (\omega_j \frac{\Delta t}{2})^2} \right\} \quad (5.54) \]

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

For the lower modes \( \rho_j = 1 \). As mode order increases, \( \rho_j \) decreases initially, then starts to increase again. For the higher modes \( \rho_j \) 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

\[ \dot{w}_{n+1} = w_n + \Delta t A \left( (1 - \alpha) w_{n+1} + \alpha w_n \right) \]

\[ w_n = \begin{pmatrix} v_n \\ \dot{v}_n \end{pmatrix} \quad w_o = c \quad 0 < \alpha < 1 \]

Equation (5.36) now becomes

\[ \dot{w}_{n+1} = \mathcal{A}_\alpha w_n \quad (5.55) \]

where

\[ \mathcal{A}_\alpha = [I - (1-\alpha)\Delta t A]^{-1} \left[ I + \alpha \Delta t A \right] \quad (5.56) \]

\[ \mathcal{A}_\alpha = \mathcal{T} \Theta_\alpha \mathcal{T}^{-1} \quad \left\{ \begin{array}{l} \text{where} \\ \Theta_\alpha = \frac{1 + \Delta t i \omega_j}{1 - \Delta t (1-\alpha) \omega_j} \end{array} \right\} \quad (5.57) \]
Hence

$$\theta_{\alpha_j} = \rho_{\alpha_j} e^{i \Delta \phi_{\alpha_j}}$$

$$\rho_{\alpha_j} = \left\{ \frac{1 + (\alpha \Delta t \omega_j)^2}{1 + (1-\alpha) \Delta t \omega_j^2} \right\}$$

$$\Delta \phi_{\alpha_j} = \tan^{-1} \left( \frac{\omega_j \Delta t}{1 - \alpha(1-\alpha)(\omega_j \Delta t)^2} \right)$$

(5.58)

We note that

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

2) $$0 < \alpha < 1/2 \quad \rho_j < 1$$

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

$$||e_n|| \leq c_3 \left\{ 1 - \alpha(\Delta t) + k_1 k_2 0(\Delta t^2) \right\}$$

(5.59)

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

(iii) Use of Temporal Filtering

Let $$v_n = \frac{1}{4} [w_{n+1} + 2w_n + w_{n-1}]$$

(5.60)

where $$w_{n+1} = \mathcal{A} w_n$$

(5.61)

$$\therefore v_n = \frac{1}{4} [\mathcal{A} + 2I + \mathcal{A}^{-1}] w_n$$

(5.62)
Now, for the trapezoidal algorithm

\[ A = T \Theta T^{-1} \]  

(5.63)

where \( \Theta_j = \frac{1 + i \omega_j \frac{\Delta t}{2}}{1 - i \omega_j \frac{\Delta t}{2}} \)  

(5.64)

\[ v_n = \frac{1}{4} T [\Theta + 2I + \Theta^{-1}] T^{-1} w_n \]  

(5.65)

\[ = \frac{1}{4} T \left[ \frac{1 + i \omega_j \frac{\Delta t}{2}}{1 - i \omega_j \frac{\Delta t}{2}} + 2 + \frac{1 - i \omega_j \frac{\Delta t}{2}}{1 + i \omega_j \frac{\Delta t}{2}} \right] T^{-1} w_n \]  

(5.66)

\[ v_n = T \left[ \frac{1}{1 + (\omega_j \frac{\Delta t}{2})^2} \right] T^{-1} w_n \]  

(5.67)

We note that \( \frac{1}{1 + (\omega_j ^2 \frac{\Delta t}{2})^2} \approx 1 \) for the low modes and tends to zero for the high modes. Since \( w_n \) 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

\[ \frac{dx}{dt} = A \mathbf{x} + f(x) + g(t) \quad 0 \leq t \leq T < \infty \]

(5.68)

\[ \mathbf{x}(0) = \mathbf{c} \quad A = \text{constant matrix} \]
One technique for solving (5.68) is the use of the trapezoidal algorithm

\[
Y_{n+1} = Y_n + \frac{\Delta t}{2} \left[ A(Y_{n+1} + Y_n) + f(Y_{n+1}) + f(Y_n) + g_{n+1} + g_n \right]
\]

\[
Y_0 = C \quad \Delta t = T/M
\]

Equation (5.69) may also be written as

\[
Y_{n+1} = AY_n + g(y_{n+1}) + g(Y_n) + g_{n+1} + g_n
\]

\[
Y_0 = c
\]

**Accuracy**

Let \( x_{n+1} = x(n\Delta t) \)

Let \( \tau_{n+1} \) be the local truncation error defined by

\[
x_{n+1} = x_n + \frac{\Delta t}{2}[A(x_{n+1} + x_n) + f(x_{n+1}) + f(x_n) + g_{n+1} + g_n + 2\tau_{n+1}]
\]

(5.71)

Let \( e_n = x_n - Y_n \) be the solution error

(5.72)

Then subtracting (5.69) from (5.71)

\[
e_{n+1} = (I + A\frac{\Delta t}{2})e_n + A\frac{\Delta t}{2}e_{n+1} + (f(x_{n+1}) - f(Y_n))\frac{\Delta t}{2}
\]

\[+ (f(x_n) - f(Y_n))\frac{\Delta t}{2} + \tau_{n+1}\Delta t
\]

(5.73)

If

i) \( x_n, Y_n \) are bounded \( \forall n \in (1, M) \)

ii) \( \|f(x) - f(y)\| \leq K\|x - y\| \) \( \forall x_n, Y_n \) bounded

(5.74)

iii) \( f(x) \) continuous and continuous first and second partials
Then

\[ \|e_{n+1}\| \leq (1 + \|A\| \frac{\Delta t}{2}) \|e_n\| + \|A\| \frac{\Delta t}{2} \|e_{n+1}\| \\
+ K \frac{\Delta t}{2} \|e_{n+1}\| + k \frac{\Delta t}{2} \|e_n\| + \|\tau_{n+1}\| \Delta t \]  \hspace{1cm} (5.75)

\[ \|e_{n+1}\| \leq \frac{(1 + (\|A\| + K) \frac{\Delta t}{2}) \|e_n\| + \|\tau_{n+1}\| \Delta t}{1 - (\|A\| + K) \frac{\Delta t}{2}} \\
(1 - (\|A\| + K) \frac{\Delta t}{2}) \|e_n\| + \frac{\|\tau_{n+1}\| \Delta t}{1 - (\|A\| + K) \frac{\Delta t}{2}} \]  

\[ \|e_0\| = 0 \]

Thus

\[ \|e_1\| \leq \frac{\|\tau_1\| \Delta t}{1 - (\|A\| + K) \frac{\Delta t}{2}} \]

\[ \|e_2\| \leq \left[ \frac{1 + (\|A\| + K) \frac{\Delta t}{2}}{1 - (\|A\| + K) \frac{\Delta t}{2}} \|\tau_1\| + \|\tau_2\| \right] \frac{\Delta t}{1 - (\|A\| + K) \frac{\Delta t}{2}} \]  \hspace{1cm} (5.76)

\[ \|e_n\| \leq [\psi^{n-1} \|\tau_1\| + \psi^{n-2} \|\tau_2\| + \cdots + \|\tau_n\|] \frac{\Delta t}{1 - (\|A\| + K) \frac{\Delta t}{2}} \]

where

\[ \psi = \frac{1 + (\|A\| + K) \frac{\Delta t}{2}}{1 - (\|A\| + K) \frac{\Delta t}{2}} \]  \hspace{1cm} (5.77)

If \( \tau = \max_k \|\tau_2\| \)

\[ \|e_n\| \leq \left[\psi^{n-1}\right]^{\tau} \frac{\Delta t}{1 - (\|A\| + K) \frac{\Delta t}{2}} \]  \hspace{1cm} (5.78)

But

\[ \psi - 1 = \frac{1 + (\|A\| + K) \frac{\Delta t}{2}}{1 - (\|A\| + K) \frac{\Delta t}{2}} - 1 = \frac{\Delta t(\|A\| + K)}{1 - (\|A\| + K) \frac{\Delta t}{2}} \]  \hspace{1cm} (5.79)
\[
\begin{align*}
\text{Now} \quad e^y &= 1 + y + \frac{y^2}{2} e^{\theta y} \quad 0 < 0 < 1 \\
(1 + y) &\leq e^y \\
1 - y &\geq 1 - 2y + \frac{1}{2} (2y)^2 \quad \text{for } y \leq \frac{1}{2} \\
e^{-z} &= 1 - z + \frac{z^2}{2} = \frac{z^2}{6} e^{-\theta z} \\
\frac{1}{1 - z + \frac{z^2}{2}} &\leq e^{-z} \\
\frac{1}{1 - y} &\leq \frac{1}{1 - 2y + 2y^2} \leq e^{2y} \\
\frac{1 + y}{1 - y} &\leq e^{3y} \\
\text{Hence, if } (\|A\| + K) \frac{\Delta t}{2} < \frac{1}{2} \\
p &= \frac{1 + (\|A\| + K) \frac{\Delta t}{2}}{1 - (\|A\| + K) \frac{\Delta t}{2}} \leq e^{\frac{3}{2}(\|A\| + K) \Delta t} \\
\|e_n\| &\leq \frac{\tau}{\|A\| + K} e^{\frac{3}{2}(\|A\| + K) n\Delta t} \\
\text{But } n\Delta t = t_n \leq T \\
\|e_n\| &\leq \frac{\tau}{\|A\| + K} e^{\frac{3}{2}(\|A\| + K) T}
\end{align*}
\]
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

\[
\begin{align*}
\dot{x} + K \dot{x} + f(x) &= 0 \\
x(0) &= a \\
\dot{x}(0) &= b
\end{align*}
\]

\[
M = M^T > 0, \quad K = K^T > 0
\]

\[
f(x) = \nabla U(x) \quad U(x) > 0 \quad x \neq 0
\]

Equation (5.94) has the first integral
\[
\frac{1}{2} \hat{x}^T M \dot{x} + \frac{1}{2} \hat{x}^T K \dot{x} + U(x) = \text{const.} \quad (5.95)
\]

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

\[
\frac{dz}{dt} = A z + g(z)
\]

\[
z(0) = c
\]

where

\[
A = \begin{bmatrix} 0 & 1 \\ -M^{-1}K & 0 \end{bmatrix}, \quad g(z) = \begin{bmatrix} 0 \\ -M^{-1}f(x) \end{bmatrix}
\]

(5.96)

If \( A \) is simple, it has the representation

\[
A = T \Lambda T^{-1}
\]

(5.97)

\[
\Lambda = \begin{bmatrix} i\omega_1 & 0 \\ 0 & -i\omega_1 \\ 0 & i\omega_2 \\ 0 & -i\omega_2 \end{bmatrix}
\]

(5.98)

e etc.

The trapezoidal difference equation corresponding to (5.94) is

\[
\begin{align*}
\frac{w_{n+1} - w_n}{\Delta t} &= \frac{\Delta t}{2} A (w_{n+1} + w_n) + \frac{\Delta t}{2} \left( g(w_{n+1}) + g(w_n) \right) \\
\frac{w_n}{\dot{x}_n} &= \frac{\dot{y}_n}{y_n} \\
\end{align*}
\]

(5.99)

Alternatively,

\[
\begin{align*}
\dot{y}_{n+1} &= \dot{y}_n + \frac{\Delta t}{2} \left( \dot{y}_{n+1} + \dot{y}_n \right) \\
\ddot{y}_{n+1} &= \ddot{y}_n - \frac{\Delta t}{2} M^{-1} \left[ K(y_{n+1} + y_n) + f(y_{n+1}) + f(y_n) \right] \\
\end{align*}
\]

(5.100)

\[
y_0 = a \quad \dot{y}_0 = b
\]
From (5.100) we see that

\[ \frac{1}{2} \left[ y_{n+1}^T M \dot{y}_{n+1} - \dot{y}_n^T M \dot{y}_n \right] + \frac{1}{2} \left[ y_{n+1}^T K y_{n+1} - \dot{y}_n^T K \dot{y}_n \right] \\
+ \frac{1}{2} \left( y_{n+1} - y_n \right)^T \left( f(y_{n+1}) + f(y_n) \right) \equiv 0 \quad (5.101) \]

which can be written

\[ \frac{1}{2} y_{n+1}^T M \dot{y}_{n+1} + \frac{1}{2} y_n^T K y_{n+1} + \tilde{U}_{n+1} = \text{constant} \quad (5.102) \]

where \( \tilde{U}_{n+1} = \sum_{i=0}^{n+1} \frac{1}{2} \left( y_{i+1} - y_i \right)^T \left( f(y_{i+1}) + f(y_i) \right) + U(y_0) \) \quad (5.103a)

We note that if \( a, b \) are bounded, then as \( \Delta t \to 0 \), \( (y_{i+1} - y_i) \to 0 \) and that (5.103) becomes

\[ \lim_{\Delta t \to 0} \tilde{U}_{n+1} = \lim_{\Delta t \to 0} \sum_{i=0}^{n+1} \frac{1}{2} \left( y_{i+1} - y_i \right)^T \left( f(y_{i+1}) + f(y_i) \right) + U(y_0) = U(y_{n+1}) \quad (5.103b) \]

In this case, equations (5.102) and (5.95) are identical. In general, for \( \Delta t \) 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

\[ \begin{aligned}
    w_{n+1} &= \mathcal{A} w_n + \mathcal{B} \left( g(w_{n+1}) + g(w_n) \right) \\
    w_0 &= \mathcal{C}
\end{aligned} \quad (5.104) \]

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

\( \mathcal{B} = \left[ I - \frac{\Delta t}{2} A \right]^{-1} \frac{\Delta t}{2} \)
Since the eigenvalues of $A$ are pure imaginary, the eigenvalues of $A$ 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:

$$\ddot{x} + f(x) = 0$$

$$f(x) = x \quad |x| \leq 1$$

$$f(x) = \text{sgn} x + \mu (x - \text{sgn} x) \quad |x| > 1$$

Equation (5.105) has the following first integral

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

$$F(x) = \frac{1}{2} x^2 \quad |x| \leq 1$$

$$= \frac{1}{2} + |x| + \frac{\mu}{2} (|x|-1)^2 \quad |x| > 1$$

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:

$$y_{n+1} - y_n = n(\tilde{y}_{n+1} + \tilde{y}_n) \quad n = \Delta t/2$$

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

which may be written

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

(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

\[
\begin{align*}
    y_{n+1} &= g(z_n) \\
    \dot{y}_{n+1} &= \frac{y_n + n\dot{y}_n + g(z_n)}{n}
\end{align*}
\]

(5.109)

where

\[
    g(z) = \begin{cases} 
        \frac{1}{1+n^2}z & \text{for } |z| \leq 1 + n^2 \\
        \sgn z + \frac{1}{1+\mu n^2} (z - (1+n^2) \sgn z) & \text{for } |z| > 1 + n^2
    \end{cases}
\]

(5.110)

\[
|g(z)| \leq |z| \quad \forall z
\]

(5.111)

Let

\[
    z_{n+1} = y_{n+1} + 2n\dot{y}_{n+1} - n^2 f(y_{n+1})
\]

(5.112)

\[
    := z_{n+1} = 4g(z_n) - 2z_n = z_{n-1}
\]

(5.113)

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

a)

\[
    z_{n+1} - 2z_n + z_{n-1} = -\frac{4n^2}{1+n^2} k_3(z_n)
\]

\[
    k_3(z_n) = \begin{cases} 
        z_n & \text{for } |z_n| \leq (1+n^2) \\
        (1+n^2) \sgn z_n + \frac{n(1+n^2)}{1+\mu n^2} (z_n - (1+n^2) \sgn z_n) & \text{for } |z_n| > 1 + n^2
    \end{cases}
\]

(5.114)

From (5.114) we see that if $\mu \geq 0$, $\sgn(z_{n+1} - 2z_n + z_{n-1}) = -\sgn z_n$ (5.115)
Thus the sign of the finite difference curvature is always opposite to that of the displacement, thus the solutions are always oscillatory.

\[
b) \quad \begin{aligned}
  z_{n+1} - \frac{2(1-\mu n^2)}{(1+\mu n^2)} z_n + z_{n-1} &= - \frac{4n^2(1-\mu)}{(1+\mu n^2)} k_2(z_n) \\
  k_2(z) &= \frac{1}{1+n^2} z \quad \text{for } |z| \leq 1+n^2 \\
  &= \text{sgn } z \quad \text{for } |z| > 1+n^2
\end{aligned}
\]

Equations can also be written in the first order form,

\[
\begin{aligned}
  \theta_{n+1} &= A\theta_n + k(\theta_n) \\
  \theta_n &= \begin{pmatrix} z_n \\ z_{n-1} \end{pmatrix} \quad A = \begin{bmatrix} \frac{2(1-\mu n^2)}{1+\mu n^2} & -1 \\ 1 & 0 \end{bmatrix} \\
  k(\theta_n) &= \begin{pmatrix} -\frac{4n^2(1-\mu)}{1+\mu n^2} k_2(z_n) \\ 0 \end{pmatrix}
\end{aligned}
\]

The matrix \( A \) has eigenvalues

\[
\lambda = -\frac{1-\mu n^2}{1+\mu n^2} \pm i \sqrt{1 - \left(\frac{1-\mu n^2}{1+\mu n^2}\right)^2}
\]

\[
= e^{\pm i\phi}
\]

\[
\cos \phi = \frac{(1-\mu n^2)}{1+\mu n^2}
\]

Using equation (3.5) with \( \mu \neq 0 \)
\[ \theta_{n+1} = A^{n+1} \theta_0 + \sum_{i=0}^{n} A^{n-i} k(\theta_i) \]  

(5.119)

But \[ x = TAT^{-1} \implies x^k = T A^k T^{-1} \]  

(5.120)

where \[ |\lambda_i| = 1 \quad i = 1,2 \]

\[ \|\theta_{n+1}\| \leq \|T\| \|T^{-1}\| \|\theta_0\| + \|T\| \|T^{-1}\| \sum_{i=0}^{n} \|k(\theta_i)\| \]  

(5.121)

But \[ \|k(\theta_i)\| \leq 4n^2 \left| \frac{1 - \mu}{1 + \mu n^2} \right| \]  

(5.122)

\[ \|\theta_{n+1}\| \leq K(\|\theta_0\| + 4n^2 \left| \frac{1 - \mu}{1 + \mu n^2} \right| n) \]  

(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 \( \lambda_i = 1, \quad i = 1,2 \)

\[ A = T \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} T^{-1} \]  

(5.124)

\[ \|\theta_{n+1}\| \leq \|T\| \|T^{-1}\| \|\theta_0\| (2 + n) + 4n^2 \|T\| \|T^{-1}\| \sum_{i=0}^{n} (2i) \]  

\[ \leq K[\|\theta_0\| (2 + n) + 4n^2 (2(n+1) + \frac{n(n+1)}{2})] \]  

(5.125)

Thus, even in this case, the \( \|\theta_n\| \sim O(n^2) \) as \( n \to \infty \), as it is only weakly unstable.

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

\[ \theta_{n+1} = M(\theta_n) \]  

(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

\[ \theta^* = M(\theta^*) \quad \text{is} \quad \theta^* = 0 \quad (5.127) \]

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

\[
\begin{align*}
\theta^*(m+1) &= M^m(\theta^*(1)), \\
\theta^*(1) &= M^k(\theta^*(1)) \\
\end{align*}
\]

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

**Stability of Periodic Solutions**

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

Let \( \theta(m) = \theta^*(m) + \delta \theta(m) \) \( (5.129) \)

Then \( \delta \theta(m+1) = M \theta^*(m) \delta \theta(m) \) \( (5.130) \)

Provided \( \theta(m) \) and \( \theta^*(m) \) are on the same piecewise linear branches of \( k(\theta) \).

Thus, for \( \delta \theta(k) \) small, but not infinitesimal, and \( \theta^*(m) \) not on a corner of \( k(\theta) \),

\[
\delta \theta(m+1) = \left[ \sum_{i=1}^{m} M \theta^*(i) \right] \delta \theta(1) \quad (5.131)
\]

Hence \( \delta \theta(k+1) = \left[ \sum_{i=1}^{K} M \theta^*(i) \right] \delta \theta(1) \quad (5.132) \)

\[ = A_k \delta \theta(1) \]
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.

Example

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

$$
\begin{align*}
z_{n+1} - 2z_n + z_{n-1} &= -4k_2(z_n) \\
k_2(z) &= \begin{cases} 
\frac{z}{2} & \text{for } |z| \leq 2 \\
\text{sgn } z & \text{for } |z| > 2
\end{cases}
\end{align*}
$$

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

$$
\begin{align*}
z_{n+1} + z_{n-1} &= 0 \\
z_{n+1} &= -z_{n-1}
\end{align*}
$$

$$
\begin{align*}
z_2 &= -z_0 \\
z_3 &= -z_1 \\
z_4 &= z_0 \\
z_5 &= z_1 \\
&\vdots \\
z_{2n} &= (-1)^nz_0 \\
z_{2n+1} &= (-1)^nz_1
\end{align*}
$$

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

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

$$
\begin{align*}
\theta_{n+1} &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \theta_n
\end{align*}
$$
\[ M_{0}(\theta^{*}(i)) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \forall i, \text{ provided } \theta^{*}(i) \text{ is not close to a corner of } k_{2}(z). \text{ Thus} \]

\[ \delta \theta(K+1) = A_{K} \delta \theta(1) \]

\[ A_{K} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \]

Thus, \[ |\lambda_{4}(A_{K})| = 1 \quad i = 1,2 \]

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

\[ z_{n+1} - 2z_{n} + z_{n-1} = -4 \quad (5.137) \]

With \(z_{0} = N+1\) and \(z_{1} = 3N-1\),

\[ z_{n} = (1+N) + 2n(N-n) \quad (5.138) \]

Thus \(z_{N-1} = 3N-1 = z_{1} ; \ z_{N} = N+1 = z_{0} > 2\)

\[ z_{N+1} = -(N+1) = -z_{0} < 2 \quad (5.139) \]

If \(z_{n} < -2\), equation (5.133) becomes

\[ z_{n+1} - 2z_{n} + z_{n-1} = +4 \quad (5.140) \]
\[ z_{N+2} = -(3N-1) = -z_1 \]

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

\[ z_{2(N+2)} = (N+1) = +z_0 \quad (5.141) \]

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

**Stability**

Since each point \( \theta^*(k) \) satisfies the condition \( |z_n| > 2 \), each \( M_\theta(\theta^*(i)) \) is the same.

\[ A_K = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \quad (5.142) \]

Now
\[
\begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} = T \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} T^{-1}
\]

\[ A_K = T \begin{bmatrix} 1 & K \\ 0 & 1 \end{bmatrix} T^{-1} \quad (5.144) \]

Thus, though \( |\lambda_1(A_K)| = 1 \forall i \in (1,2) \), \( A_K \) is not simple, hence the periodic solutions are weakly unstable, and will grow until \( \theta(n) = \theta^*(n) + \delta\theta(n) \) 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

\[
\begin{align*}
    z_{n+1} - 2z_n + z_{n-1} &= -4k_2(z_n) \\
    k_2(z) &= z/2 \quad \text{for} \quad |z| \leq 2 \\
    &= \text{sgn} z \quad \text{for} \quad |z| > 2
\end{align*}
\]

(5.145)

\[
\begin{align*}
    z_0 &= 1.5N + 2.5 \quad , \quad z_1 = 3.5N + 1.5 \quad , \quad N = 1, 2, 3, \text{etc.}
\end{align*}
\]

Since \(z_0, z_1 > 2\)

\[
\begin{align*}
    z_k &= A + Bk - 2k^2, \quad \text{provided} \quad z_{n-1} > 2
\end{align*}
\]

(5.146)

Using the given initial data

\[
\begin{align*}
    z_0 &= A = 1.5N + 2.5 \\
    z_1 &= A + B - 2 = 3.5N + 1.5 \quad : \quad B = 2N + 1
\end{align*}
\]

(5.147)

\[
\begin{align*}
    z_k &= (1.5N + 2.5) + (2N + 1 - 2k)k \quad \text{for} \quad k \leq N + 2
\end{align*}
\]

(5.148)

\[
\begin{align*}
    z_{N+1} &= 0.5N + 1.5 \quad , \quad z_{N+2} = -(1.5N + 3.5)
\end{align*}
\]

(5.149)

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

\[
\begin{align*}
    z_{N+3} &= -(3N + 7) - (0.5N + 1.5) + 4 \\
    &= -(3.5N + 4.5)
\end{align*}
\]

(5.150)

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

\[
\begin{align*}
    z_{N+2+k} &= -[A_1 + B_1k - 2k^2] \quad k > 0
\end{align*}
\]

where \(A_1 = -z_{N+2} = (1.5N + 3.5)\)

\[
\begin{align*}
    z_{N+3} &= -[A_1 + B - 2] \quad : \quad B_1 = 2N + 3
\end{align*}
\]

(5.151)

\[
\begin{align*}
    z_{N+2+k} &= -[1.5N + 3.5 + (2N + 3 - 2k)k] \quad , \quad k \leq N + 2
\end{align*}
\]

(5.152)

\[
\begin{align*}
    z_{2N+4} &= -(0.5N + 1.5) \\
    z_{2N+5} &= (1.5N + 5.5) = (z_0 + 3)
\end{align*}
\]
To calculate $z_{2N+6}$ we return to equation (5.145), from which

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

Thus at the end of a complete cycle

$$T_1 = (2N+5)$$

and

$$z_{T_1} = z_0 + 3 \quad (5.154)$$

are the initial data for the next cycle

$$z_{T_1+1} = z_1 + 7$$

At the end of the next complete cycle

$$T_2 = 4N+14$$

and

$$z_{T_2} = z_0 + 2 \times 3 \quad (5.155)$$

$$z_{T_2+1} = z_1 + 2 \times 7$$

At the end of the $k$th complete cycle

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

and

$$z_{T_k} = z_0 + 3k \quad (5.156)$$

are the initial data for the next cycle

$$z_{T_k+1} = z_1 + 7k$$

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

$$\frac{1}{2} (y_{n+1}^2 - y_n^2) + (y_{n+1} - y_n)(f(y_{n+1}) + f(y_{n+1})) = 0 \quad (5.157)$$

If $y_{n+1}, y_n$ are both of the same sign and both are greater than unity, then (5.157) becomes
From equation (5.106) with $x_{n+1}$ and $x_n$ of the same sign and both greater than unity, then,

\[
\frac{1}{2} (y_{n+1}^2 - y_n^2) + \frac{1}{2} (y_{n+1} - y_n)(\text{sgn } y_{n+1} + \text{sgn } y_n) = 0 \quad (5.158)
\]

\[
\therefore \quad \frac{1}{2} (y_{n+1}^2 - y_n^2) + |y_{n+1}| - |y_n| = 0 \quad (5.159)
\]

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

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

But \[ F(x_{n+1}) - F(x_n) = |x_{n+1}| - |x_n| \]

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

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),

\[
y_{n+1} = g(z_n)
\]

\[
\dot{y}_{n+1} = -\frac{1}{2} [z_n + z_{n-1} - 2g(z_n)]
\]

(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.

\[
E_k = \frac{1}{2} \dot{y}_{k+1}^2 + |y_{k+1}|
\]

(5.163)

\[
|y_{n+1}| = |z_n - \text{sgn } z_n| = |z_n| - 1
\]
From (5.156)

\[ E_k = \frac{1}{2} (z_0 + z_1 + 10k - 2)^2 + \left[ z_1 + 7k \right] - 1 \quad (5.164) \]

As \( k \to \infty \),

\[ E_k \sim O(k^2) \quad (5.165) \]

But from (5.155)

\[ T_k = k(2N+3+2k). \]

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

Therefore, combining (5.165) and (5.166)

\[ E_k \sim T_k \quad \text{as } k \to \infty \quad (5.167) \]

Thus confirming analytically what Hughes and others had obtained numerically.

Equation (5.116) was carefully examined for \( \mu = 0 \) and \( \eta \) arbitrary; nothing essentially new was learned, except that even for \( \eta \) 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 \( \eta \).

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 \( \mu \), the effect of small \( \eta \) 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 nonlinearities lead to weak instability. In this section we shall look at several new algorithms which conserve energy.

Consider the conservation nonlinear differential equation

\[ \frac{d^2 x}{dt^2} + f(x) = 0 \quad 0 < t < T \]

where

\[ f(x) = \frac{dF}{dx}, \quad F(x) = \int_0^x f(\eta)d\eta > 0 \quad x \neq 0 \]

and

\[ xf(x) > 0 \quad x \neq 0 \]

The system (5.168) has the first integral

\[ \frac{1}{2} \dot{x}^2 + F(x) = \text{const} \] (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

\[ \frac{d^2 x}{dt^2} + \frac{dF}{dx} = 0 \] (5.170)

This immediately suggests the algorithm

\[
\begin{align*}
T_{n+1} - y_n &= \Delta t \left( \hat{T}_{n+1} + \hat{T}_n \right) \\
T_{n+1} - \hat{T}_n &= -\Delta t \frac{F(y_{n+1}) - F(y_n)}{(y_{n+1} - y_n)}
\end{align*}
\]

(5.171)
Cross multiplication immediately shows that:

\[
\frac{1}{2} \sum (y_{n+1}^2 - y_n^2) + F(y_{n+1}) - F(y_n) = 0 \quad (5.172)
\]

Or on summing

\[
\frac{1}{2} \sum y_{n+1}^2 + F(y_{n+1}) = \text{const} \quad (5.173)
\]

If in (5.168)

\[
\begin{cases}
    f(x) = w^2x + g(x) \\
    g(x) = \frac{dG}{dx}, C(x) = \int_0^x g(\eta) d\eta > 0 \quad \text{for } x \neq 0
\end{cases}
\]

\[
 w^2x^2 + g(x) > \forall x \neq 0
\]

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

\[
\begin{cases}
    y_{n+1} - y_n = \frac{\Delta t}{2} (y_{n+1}^2 - y_n^2) \\
    y_{n+1} y_n = -\frac{\Delta t}{2} w^2(y_{n+1} + y_n) - \Delta t \frac{G(y_{n+1}) - G(y_n)}{(y_{n+1} - y_n)}
\end{cases}
\]

which may also be written in the form

\[
\frac{w_{n+1}}{w_n} = w_n + \frac{\Delta t}{2} A(w_{n+1}, w_n) + x(w_{n+1}, w_n) \quad (5.176)
\]

where
\[ A = \begin{bmatrix} 0 & 1 \\ y_n^2 & 0 \end{bmatrix}, \quad w_n = \begin{pmatrix} y_n \\ y_n^2 \end{pmatrix} \]

\[ X = \left( \Delta t \frac{G(y_{n+1}) - G(y_n)}{y_{n+1} - y_n} \right) \]

\[ w_{n+1} = \begin{pmatrix} y_{n+1} \\ y_{n+1}^2 \end{pmatrix} = \begin{pmatrix} y_n \\ y_n^2 \end{pmatrix} + \begin{pmatrix} \Delta t \frac{G(y_{n+1}) - G(y_n)}{y_{n+1} - y_n} \end{pmatrix} \]

\[ \Delta t \frac{G(y_{n+1}) - G(y_n)}{y_{n+1} - y_n} = g(y_n) + g'(y_n) \frac{y_{n+1} - y_n}{2} + g''(\xi) \frac{(y_{n+1} - y_n)^2}{6} \]

\[ \xi = \theta y_n + (1-\theta)y_{n+1}, \quad 0 < \theta < 1 \]

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

\[ \left| \frac{G(y_{n+1}) - G(y_n)}{y_{n+1} - y_n} \right| < M \quad \forall n \quad (5.179) \]

\[ \therefore \quad \| X (w_{n+1}, w_n) \| < M \Delta t \quad \forall n \quad (5.180) \]

Similarly if \( y_{n+1}, y_n \) are bounded \( \forall n \)

\[ \| \frac{\Delta t}{2} A(w_{n+1} + w_n) \| < N \Delta t \quad \forall n \quad (5.181) \]

Hence from equation (5.176)

\[ \| w_{n+1} - w_n \| \leq (N+M) \Delta t \sim O(\Delta t) \quad \text{for } \Delta t \text{ small} \quad (5.182) \]

As before, we define the truncation error \( r_{n+1} \), by the equations
Using equations (5.174), (5.178) and (5.183) it is readily shown that if $x_n, \bar{x}_n$ are bounded for all n:

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

Defining the solution error $e_n = z_n - w_n$, if (5.176) is subtracted from (5.183)

$$e_{n+1} = e_n + \frac{\Delta t}{2} A(e_{n+1} + e_n) + \chi(z_{n+1}, z_n) + \chi(w_{n+1}, w_n) + \tau_{n+1} \Delta t$$  \hspace{1cm} (5.185)

$$\|e_{n+1}\| \leq \|e_n\| + \frac{\Delta t}{2} \|A\| (\|e_{n+1}\| + \|e_n\|) + \|\chi(z_{n+1}, z_n) + \chi(w_{n+1}, w_n)\|$$

$$+ \|\tau_{n+1}\| \Delta t$$  \hspace{1cm} (5.186)

Using (5.178) and (5.182)

$$\|\chi(z_{n+1}, z_n) - \chi(w_{n+1}, w_n)\| \leq \frac{\Delta t}{2} K (\|e_{n+1}\| + \|e_n\|) + \delta_{n+1} \Delta t$$

provided that $z_n, z_{n+1}, w_n, w_{n+1}$ are bounded $\forall n$ where

$$\delta_{n+1} \sim O(\Delta t^2)$$  \hspace{1cm} (5.187)

Let

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

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

$$\|e_{n+1}\| \leq \left(1 + (\|A\| + K) \frac{\Delta t}{2}\right) \|e_n\| + \delta \Delta t$$  \hspace{1cm} (5.189)
If $\|e_0\| = 0$ then as previously shown

$$\|e_n\| \leq \frac{\exp\left( (\|A\|+K) \frac{3}{2} T \right)}{\|A\| + K} \phi$$

(5.190)

Thus for $T$ fixed

$$\|e_n\| \leq K_3 \phi \sim O(\Delta t^2) \text{ as } \Delta t \to 0$$

(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

$$C(y_{n+1}) - G(y_n) = \frac{1}{2} (g(y_{n+1}) + g(y_n)) + \left[ \frac{g''(\xi_1) + g''(\xi_2) - 3g''(\xi_3)}{12} \right] (y_{n+1} - y_n)^2$$

$$\xi_i = y_{n+1} \theta_i + (1 - \theta_i) y_n$$

$$0 < \theta_i < 1$$

(5.192)

Thus if $y_n, y_{n+1}$ are bounded $\forall n$, then for $\Delta t$ tending to zero.

$$C(y_{n+1}) - G(y_n) = \frac{1}{2} [g(y_{n+1}) + g(y_n)] + M(\Delta t)^2$$

(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

$$\ddot{x} + 2z \dot{x} + \omega^2 x + g(x) = 0 \quad , \quad z > 0$$

(5.194)

Let
\[
V(x, x) = \frac{1}{2}[x^2 + 2zx + w^2x^2 + 2zx + 2G(x)] \tag{5.195}
\]

\[
V(x, x) > 0 \text{ if } x, x \neq 0
\]

\[
\dot{V}(x, x) = x \dot{x} + zx^2 + z\dot{x}x + w^2x + 2zx + 2G(x) \tag{5.196}
\]

Using equation (5.194)

\[
\dot{V} = -z(x^2 + w^2x^2 + xg(x)) \tag{5.197}
\]

But, from equation (5.174)

\[
w^2x^2 + xg(x) > 0 \quad \forall x \neq 0
\]

\[
\therefore \quad \dot{V} < 0 \tag{5.198}
\]

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)

\[
y_{n+1} - y_n = \Delta t \left( \frac{\dot{y}_{n+1} + \dot{y}_n}{2} \right) \tag{5.199}
\]

\[
\frac{\dot{y}_{n+1} - \dot{y}_n}{\Delta t} = -\frac{\Delta t}{2}[2z(y_{n+1} + y_n)x + \omega^2(y_{n+1} + y_n)] - \Delta t \frac{G(y_{n+1}) - G(y_n)}{y_{n+1} - y_n} \tag{5.200}
\]

Let

\[
V_n = \frac{1}{2}[y_n^2 + zy_n^2 + (z^2 + \omega^2)y_n^2 + 2G(y_n)] \tag{5.201}
\]

\( V_n \) is positive definite and vanishes only when \( y_n = \dot{y}_n = 0 \).

\[
\therefore \quad V_{n+1} = \frac{1}{2}[(y_{n+1} + zy_{n+1})^2 + (z^2 + \omega^2)y_{n+1}^2 + 2G(y_{n+1})] \tag{5.202}
\]
\[ \Delta V_n = V_{n+1} - V_n \]
\[ = \frac{1}{2} \left[ (y_{n+1}^* - y_n^*)^2 - y_n^* y_{n+1}^* (y_{n+1}^* + y_n^*) (y_{n+1}^* + y_n^*) \right] \]
\[ + \frac{1}{2} \left[ (\dot{y}_{n+1}^* + \dot{y}_n^*)^2 - (\dot{y}_{n+1}^* + \dot{y}_n^*) (y_{n+1}^* + y_n^*) \right] \]
\[ + \frac{1}{2} \left[ (z^2 + w^2)(y_{n+1}^* - y_n^*)^2 + \dot{G}(y_{n+1}^*) - \dot{G}(y_n^*) \right] \]  

(5.203)

Using equations (5.200)

\[ \Delta V_n = \frac{1}{2} \left[ -\Delta t \frac{G(y_{n+1}) - G(y_n)}{y_{n+1} - y_n} - \frac{z \Delta t}{2} (y_{n+1}^* + y_n^*) - \frac{\Delta t}{2} w^2 (y_{n+1}^* + y_n^*) \right] \]
\[ \times \left[ \frac{2}{\Delta t} (y_{n+1} - y_n) + z (y_{n+1} + y_n) \right] \]
\[ + \frac{1}{2} \left[ (z^2 + w^2)(y_{n+1}^* - y_n^*)^2 + \dot{G}(y_{n+1}) - \dot{G}(y_n) \right] \]  

(5.204)

\[ \Delta V_n = -z \Delta t \left[ \left( \frac{y_{n+1}^* + y_n^*}{2} \right)^2 w^2 \left( \frac{y_{n+1}^* + y_n^*}{2} \right) + \left( \frac{y_{n+1}^* + y_n^*}{2} \right) \left( \frac{G(y_{n+1}) - G(y_n)}{y_{n+1}^* - y_n^*} \right) \right] \]  

(5.205)

If \( G(y) \) is an even monotone increasing function say \( G(y) = G^*(y^2) \)

(5.206)

\( G^*(y^2) \) is also a monotone increasing function. Hence

\[ \frac{G^*(y_{n+1}^2) - G^*(y_n^2)}{y_{n+1}^2 - y_n^2} \geq 0 \]  

(5.207)

\[ \Delta V_n = -z \Delta t \left[ \left( \frac{y_{n+1}^* + y_n^*}{2} \right)^2 + \left( \frac{y_{n+1}^* + y_n^*}{2} \right) \left( \frac{G^*(y_{n+1}^2) - G^*(y_n^2)}{y_{n+1}^2 - y_n^2} \right) \right] \]

\[ \leq -z \Delta t \left[ \left( \frac{y_{n+1}^* + y_n^*}{2} \right)^2 + w^2 \left( \frac{y_{n+1}^* + y_n^*}{2} \right)^2 \right] \]  

(5.208)

From (5.208) we observe that:
(a) \( \Delta V_n \leq 0 \)

(b) Since \( y_{n+1} - y_n = \frac{\Delta t}{2} \left[ y_{n+1} + y_n \right] \), both terms in (5.208) cannot vanish simultaneously unless \( y_n = y_{n+1} = 0 \)

(c) Since \( w_{n+1} - w_n \sim O(\Delta t) \)

\[ \Delta V_n \leq -z\Delta t [ (y_n')^2 + w^2 (y_n')^2 ] + O(\Delta t^2) \text{ as } \Delta t \to 0 \]

\[ \therefore \text{The discrete system (5.199) is Liapunov asymptotically stable.} \]

**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

\[ \ddot{x} + 2z\dot{x} + w^2 x + g(x) = p(t) \quad (5.209) \]

then if \( \sup_t |p(t)| = p_o \), all solutions of (5.209) are ultimately bounded.

Let

\[ V(x, \dot{x}) = \frac{1}{2} \left[ x^2 + 2z \dot{x} x + w^2 \dot{x}^2 + 2z \dot{x}^2 x + 2G(x) \right] \quad (5.210) \]

\[ V(x, \dot{x}) > 0 \quad \forall x, \dot{x} \neq 0 \]

\[ \dot{V} = x \dddot{x} + z \ddot{x}^2 + z \dot{x} \ddot{x} + w^2 \dot{x} \ddot{x} + 2z \ddot{x} \dot{x} + g(x) \dot{x} \quad (5.211) \]

Using equation (5.209)

\[ \dot{V} = -z(x^2 + w^2 \dot{x}^2 + xg(x)) + p(t) (x + z \dot{x}) \]

\[ \therefore \text{if } xg(x) > 0 \quad \forall x \neq 0 \]

\[ \dot{V} \leq -z(x^2 + w^2 \dot{x}^2) + p_o \sqrt{ \left| \dot{x} \right| + z \left| x \right| } \quad (5.212) \]

\[ \leq -z(x^2 + w^2 \dot{x}^2) + p_o \sqrt{ 1 + \left( \frac{z}{w} \right)^2 } \sqrt{ \frac{z^2 + w^2 x^2}{x^2} } \quad (5.213) \]

Let \( S \) be the set \( x^2 + w^2 \dot{x}^2 \leq 2 \left( \frac{p_o}{x} \right)^2 \left( 1 + \left( \frac{z}{w} \right)^2 \right) \quad (5.214) \)
Outside the set $S$, $\dot{V} < 0$

Let $\Omega$ be the set $V < e$, where $e$ is such that $S$ is a proper subset of $\Omega$. Then, for points $(x, \dot{x})$ outside $\Omega$, $\dot{V} < 0$ \hspace{1cm} (5.215)

Starting outside $\Omega$, $V > 0$, $\dot{V} < 0$, therefore, $V$ decreases and the trajectory must eventually enter $\Omega$, and once inside $\Omega$, the trajectory cannot leave $\Omega$ since $\dot{V} < 0$ on $\partial\Omega$. Starting inside $\Omega$, $V > 0$, $\dot{V}$ is in general sign indefinite, therefore $V$ may increase, however, it is clear that the trajectory cannot leave $\Omega$ since $\dot{V} < 0$ on $\partial\Omega$. Thus all solutions of (5.209) are ultimate bounded in $\bar{n} = (\Omega + \partial\Omega)$.

Consider now the discrete form of equation (5.209)

$$y_{n+1} - y_n = \frac{\Delta t}{2} (\dot{y}_{n+1} + \dot{y}_n)$$

$$\dot{y}_{n+1} - \dot{y}_n = -\frac{\Delta t}{2} [2z(\dot{y}_{n+1} + \dot{y}_n) + w^2(y_{n+1} + y_n) - (p_{n+1} + p_n)]$$

$$- \Delta t \frac{G(y_{n+1}) - G(y_n)}{y_{n+1} - y_n}$$ \hspace{1cm} (5.216)

Let

$$V_n = \frac{1}{2} [(\dot{y}_n + z y_n)^2 + (z^2 + w^2)y_n^2 + 2G(y_n)]$$ \hspace{1cm} (5.217)

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

$$= \frac{1}{2} [(y_{n+1} + z y_{n+1})^2 + (z^2 + w^2)y_{n+1}^2 + 2G(y_{n+1})]$$

$$- [(\dot{y}_n + z y_n)^2 + (z^2 + w^2)y_n^2 + 2G(y_n)]$$ \hspace{1cm} (5.218)

Using equation (5.216)
\[ \Delta V_n = -z\Delta t \left[ \left( \frac{y_{n+1}^{+y_n}}{2} \right)^2 + \left( \frac{y_{n+1}^{+y_n}}{2} \right) \left( w^2 + \frac{C^*(y_{n+1}^2) - C(y_n^2)}{y_{n+1}^2 - y_n^2} \right) \right] \]

\[ + \Delta t \left[ \frac{p_{n+1}^{+p_n}}{2} \left( \frac{y_{n+1}^{+y_n}}{2} \right) + z \left( \frac{y_{n+1}^{+y_n}}{2} \right) \right] \]  

(5.219)

Using (5.207) we have:

\[ \Delta V_n \leq -z\Delta t \left[ \left( \frac{y_{n+1}^{+y_n}}{2} \right)^2 + w^2 \left( \frac{y_{n+1}^{+y_n}}{2} \right)^2 \right] \]

\[ + \Delta t p_0 \sqrt{2 \left( 1 + \left( \frac{z}{w} \right)^2 \right) \left( y_n^2 + w^2 \left( y_n^2 \right) \right)} \]  

(5.220)

Let

\[ \frac{y_{n+1}^{+y_n}}{2} = \langle y_n^2 \rangle \]  

\[ \frac{y_{n+1}^{+y_n}}{2} = \langle y_n \rangle \]

\[ \Delta V_n \leq -z\Delta t \left[ \langle y^2 \rangle + w^2 \langle y_n^2 \rangle \right] \]

\[ + \Delta t p_0 \sqrt{2 \left( 1 + \left( \frac{z}{w} \right)^2 \right) \left( \langle y_n^2 \rangle + w^2 \langle y_n^2 \rangle \right)} \]  

(5.221)

Let \( \langle S \rangle \) be the set

\[ \langle y_n^2 \rangle + w^2 \langle y_n^2 \rangle \leq 2 \left( \frac{p_0}{2} \right)^2 \left( 1 + \left( \frac{z}{w} \right)^2 \right) \]  

(5.222)

Since

\[ w_{n+1} - w_n \sim O(\Delta t) \]  

as \( t \to 0 \).

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

\[ M\ddot{x} + Kx + \nabla u(x) = 0 \]  

(5.223)

where \( M, K \) are \( N \times N \) symmetric positive definite matrices and \( u(x) \) is a positive definite potential function. For the system (5.233)

\[ V(x, \dot{x}) = \frac{1}{2} [\dot{x}^T M \dot{x} + \dot{x}^T K \dot{x} + 2u(x)] > 0 \]

\[ \dot{V} = \dot{x}^T M \ddot{x} + \dot{x}^T K \dot{x} + \dot{x}^T \nabla x u(x) = 0 \]

Algorithm A

We may write equation (5.223) in discrete form as

\[ y_{n+1} - y_n = \frac{\Delta t}{2} \left[ y_{n+1}^T + y_n^T \right] \]

\[ M [y_{n+1} \dot{y}_{n+1} - y_n \dot{y}_n] = -\Delta t K[y_{n+1} \dot{y}_{n+1} + y_n \dot{y}_n] \]

(5.225)

\[ \frac{(u(y_{n+1}) - u(y_n)) [\nabla u(y_{n+1}) + \nabla u(y_n)]}{(y_{n+1} - y_n) [\nabla u(y_{n+1}) + \nabla u(y_n)]} \]

Cross multiplication yields

\[ \frac{1}{2} \left[ y_{n+1}^T M y_{n+1} \dot{y}_{n+1} + y_{n+1}^T K y_{n+1} + 2u(y_{n+1}) \right] \]

\[ = \frac{1}{2} \left[ y_n^T M y_n \dot{y}_n + y_n^T K y_n + 2u(y_n) \right] \]

(5.226)

Thus if
\[
V_n = \frac{1}{2} \left[ y_n^T M y_n + y_n^T K y_n + 2u(y_n) \right] > 0
\]  

(5.227)

\[
\Delta V_n = V_{n+1} - V_n = 0
\]

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, \( y_n, \dot{y}_n \) are bounded for all \( n \), provided \( y_0, \dot{y}_0 \) are bounded. Thus as

\[
\Delta t \to 0, \quad y_{n+1} - y_n \sim O(\Delta t)
\]

\[
\frac{(u(y_{n+1}) - u(y_n)) [\nu(y_{n+1}) + \nu(y_n)]}{(y_{n+1} - y_n) [\nu(y_{n+1}) + \nu(y_n)]}
\]

(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 \( \Delta t \) is not small, we have not been able to prove that:

\[
(y_{n+1} - y_n)^T [\nu(y_{n+1}) + \nu(y_n)] = 0 \implies u(y_{n+1}) = u(y_n)
\]

Thus we are unable to prove that the last term in (5.225) is bounded when \( y_n, y_{n+1} \) are bounded.

**Algorithm B**

An alternative to the discrete gradient operator in (5.225) is the
operator

\[ Lu = \left\{ \frac{1}{N} \sum_{k=1}^{N} \frac{\Delta u_k}{y_{n+1} - y_n} \right\} \] (5.229)

where

\[ \Delta u_k = u(y^1, y^2, \ldots, y^{k+1}, y^n, \ldots, y_{n+1}, y^n, \ldots, y_n)^i, i'+1, \ldots, y^n) \]

\[ -u(y^1, y^2, \ldots, k, k+1, k+2, \ldots, y^n) \] (5.230)

Thus

\[ y_{n+1} = \frac{\Delta t}{2} \left[ \frac{y_{n+1} + y_n}{2} \right] \]

\[ M[\dot{y}_{n+1} - \dot{y}_n] = -\frac{\Delta t}{2} K [y_{n+1} - y_n] \]

\[ \begin{cases} \sum \Delta u_k / (y^1 - y^n) \\ \sum \Delta u_k / (y^2 - y^n) \\ \vdots \\ \sum \Delta u_k / (y^n - y^n) \end{cases} \] (5.231)

Cross multiplication yields

\[ \frac{1}{2} \left[ \dot{y}^T_{n+1} M \dot{y}_{n+1} + \dot{y}^T_{n+1} K \dot{y}_{n+1} \right] \]

\[ + \frac{1}{N} \sum_{i,k} \Delta u_k = \frac{1}{2} \left[ \dot{y}^T_n M \dot{y}_n + \dot{y}^T_n K \dot{y}_n \right] \] (5.232)

Now
\[ \frac{1}{N} \sum_{i,k} \Delta_i u_k = u(y_{n+1}) - u(y_n) \quad (5.233) \]

To see this, consider \( N = 2 \):

\[ \begin{array}{c}
\frac{1}{2} \sum_{i,k} \Delta_i u_k = \frac{1}{2} \left[ u(y_{n+1}, y_n) - u(y_n, y_{n+1}) + u(y_{n+1}, y_n) - u(y_n, y_{n+1}) \right] \\
= \frac{1}{2} \left[ 2u(y_{n+1}^1, y_{n+1}^2) - 2u(y_n^1, y_n^2) \right] \\
= u(y_{n+1}^1) - u(y_n^1)
\end{array} \quad (5.234) \]

Thus, using (5.234) in (5.232) we have:

\[ \frac{1}{2} \left[ y_{n+1}^T M y_{n+1} + y_{n+1}^T K y_{n+1} + 2u(y_{n+1}) \right] = \text{constant} \quad (5.235) \]

Thus algorithm B also conserves energy. If the potential \( u(y) \) can be expressed in the form \( u \)

\[ u(y) = \sum_{i=1}^{N} \alpha_i (y_i^1)^2 \quad (5.236) \]

where \( \alpha_i(r) \) is a positive monotone increasing function of \( r \), then it may be shown that

\[ \frac{1}{N} \sum_{i,k} \Delta_i u_k = \frac{1}{(y_{n+1}^1)^2 - (y_n^1)^2} \geq 0 \quad \forall i \quad (5.237) \]

Application of Algorithm B to System (5.223) with Viscous Damping and Additive Forces

If to equation (5.223), viscous damping and external forces are
\[ M\ddot{x} + C\dot{x} + Kx + \frac{\partial}{\partial x} u(x) = p(t) \] (5.238)

Then if \( C \) is symmetric and positive definite and \( \| p(t) \| \) is bounded, all solutions of (5.238) are ultimately bounded provided \( x^T V u(x) \geq 0 \).

Let
\[ V = \frac{1}{2} \left[ z(x^T(C-zM)x) + (x+z\dot{x})^T M(x+z\dot{x}) + 2u(x) \right] + x^T K x \] (5.239)

where \( 2z = \) smallest eigenvalue of \( \lambda M - C \equiv 0 \)

\[ \dot{V} = z(x^T(C-zM)x) + (x+z\dot{x})^T M(x+z\dot{x}) + x^T V u(x) + x^T K x \] (5.240)

Using (5.238), we have
\[ \dot{V} = z(x^T(C-zM)x) - (x+2z\dot{x})^T [(C-zM)x + Kx + u(x)] - p(t) + x^T K x + x^T V u(x) \] (5.241)

Since
\[ x^T V u(x) \geq 0 \]
\[ \dot{V} \leq -z(x^T(C-zM)x) - z(x^T K x) + \frac{1}{2}(x+z\dot{x})^T p(t) \] (5.242)

Since \( 2z \) is equal to the smallest eigenvalue of \( M^{-1} C, C-zM \) is symmetric and positive definite, hence the first two terms in (5.242) are negative definite. Since \( p(t) \) is bounded, and the third term contains \( x \) and \( \dot{x} \) linearly, there exists a set \( S: x^T M \dot{x} + x^T K x \leq K \) such that outside of \( S, \dot{V} < 0 \). Let \( \Omega \) be the set \( V \leq C \), where \( C \) is such that
$S$ is a proper subset of $\Omega$. Then outside of $\Omega \dot{\psi} < 0$, and applying the arguments used previously, we see that all solutions of (5.238) are ultimately bounded in $\Omega$.

Now consider the discrete form of equation (5.238)

$$
\frac{y_{n+1} - y_n}{\Delta t} = \frac{\Delta t}{2} \left[ \mathbf{y}_{n+1}^* + \mathbf{y}_n^* \right] - \frac{\Delta t}{2} \left[ y_{n+1}^* + y_n^* \right] - \frac{\Delta t}{2} \left[ y_{n+1}^* + y_n^* \right] C \left[ y_{n+1}^* + y_n^* \right]
$$

$$
M[y_{n+1}^* - y_n^*] = \frac{\Delta t}{2} K [y_{n+1}^* + y_n^*] - \frac{\Delta t}{2} C [y_{n+1}^* + y_n^*]
$$

$$
= \left\{ \begin{array}{c}
\sum_{k=1}^{N} \Delta_t u_k / (y_{n+1}^* - y_n^*) \\
\vdots \\
\sum_{k=1}^{N} \Delta_n u_k / (y_{n+1}^* - y_n^*)
\end{array} \right\} (5.243)
$$

Let

$$
V_n = \frac{1}{2} \left[ z(y_n^T (C-zM)y_n) + (\mathbf{y}_n^* + zy_n^*)^T M (\mathbf{y}_n^* + zy_n^*) + y_n^T K y_n + 2u(y_n) \right]
$$

Then

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

$$
= \frac{1}{2} \left[ z(y_{n+1}^T (C-zM)(y_{n+1}^* + y_n^*)) + (\mathbf{y}_{n+1}^* + zy_{n+1}^* + zy_n^*)^T M (\mathbf{y}_{n+1}^* + zy_{n+1}^* + zy_n^*) + (\mathbf{y}_{n+1} - \mathbf{y}_n)^T K (\mathbf{y}_{n+1} + \mathbf{y}_n) + 2(u(y_{n+1}) - u(y_n)) \right]
$$

(5.245)

Using equations (5.243) and (5.237)
\[ \Delta V_n \leq -\Delta t \langle \dot{y}_n^T (C-zM) \dot{y}_n^T + z \dot{y}_n^T K \dot{y}_n^T \rangle + \Delta t \left| \langle \dot{y}_n^T + z \dot{y}_n \rangle \right| \tag{5.246} \]

where
\[ \langle x_n \rangle = \frac{1}{2} [x_{n+1} + x_n] \tag{5.247} \]

Since \( w_{n+1} - w_n \sim O(\Delta t) \) as \( \Delta t \to 0 \), it is clear that the right hand side of (5.246) tends to
\[ -\Delta t \left[ \langle \dot{y}_n^T (C-zM) \dot{y}_n^T + z \dot{y}_n^T K \dot{y}_n^T \rangle + \Delta t \left| \langle \dot{y}_n^T + z \dot{y}_n \rangle \right| \right] \quad \text{as} \quad \Delta t \to 0 \tag{5.248} \]

Thus, for small \( \Delta t \), 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 \( \Delta t \to 0 \)
\[
\frac{1}{N} \left\{ \sum_{k=1}^{N} \frac{\Delta u_k}{(y_{n+1}^k - y_n^k)} \right\} = \frac{1}{2} [\nabla u(y_{n+1}) + \nabla u(y_n)] + O(\Delta t^2) \tag{5.249} \]

Thus, as \( \Delta t \to 0 \), equations (5.244) are of trapezoidal form, therefore algorithm B is second order accurate as \( \Delta t \) tends to zero.

Algorithm B conserves energy and has some nice stability.
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 \( u(y) \).

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)

Algorithm C

Consider the system of conservative nonlinear differential equations

\[
M \dddot{x} + K \dot{x} + \nabla_x u(x) = 0 \tag{5.250}
\]

where \( M, K \) are \( N \times N \) symmetric positive definite matrices and \( u(x) \) is a positive definite potential function. We know that (5.250) has the Liapunov function

\[
V(x, \dot{x}) = \frac{1}{2} [x^T M \dddot{x} + x^T K \dot{x} + 2u(x)] = \text{constant} \tag{5.251}
\]

If we write (5.250) in trapezoidal discrete form

\[
\begin{aligned}
\frac{Y_{n+1} - Y_n}{\Delta t} &= \frac{\Delta t}{2} [Y_{n+1} + Y_n] \\
M(Y_{n+1} - Y_n) &= -\frac{\Delta t}{2} [K(Y_{n+1} + Y_n) + V_{n+1} + V_n]
\end{aligned} \tag{5.252}
\]

We now wish to constrain (5.252) such that

\[
\frac{1}{2} Y_{n+1}^T M Y_{n+1} + \frac{1}{2} Y_{n+1}^T K Y_{n+1} + u_{n+1} = \text{const} \tag{5.253}
\]
Using the first of equations (5.252)

\[ y_{n+1} = \frac{2}{\Delta t} [y_{n+1} - y_n] - \ddot{y}_n \]  

(5.254)

Substituting into the second of equations (5.252)

\[ M[y_{n+1} - y_n - \Delta t \ddot{y}_n] = - \left( \frac{\Delta t}{2} \right)^2 [K(y_{n+1} + y_n) + \nu(u_{n+1} + u_n)] \]  

which may be written

\[ M[y_{n+1} + \left( \frac{\Delta t}{2} \right)^2 \nu u_{n+1}] = M[y_n + \Delta t \ddot{y}_n] - \left( \frac{\Delta t}{2} \right)^2 [K y_n + \nu u_n] \]  

(5.255)

Using (5.254), equation (5.253) may be written

\[ G(y_{n+1}) = \frac{1}{2} (y_{n+1} - y_n - \frac{\Delta t}{2} \ddot{y}_n)^T M (y_{n+1} - y_n - \frac{\Delta t}{2} \ddot{y}_n) \]

\[ + \left( \frac{\Delta t}{2} \right)^2 \left[ \frac{1}{2} y_n^T K y_{n+1} + u_{n+1} \right] \]

\[ - \frac{1}{2} y_n^T M \ddot{y}_n - \left( \frac{\Delta t}{2} \right)^2 \left[ \frac{1}{2} y_n^T K y_n + u_n \right] = 0 \]  

(5.256)

Let us now construct the functional

\[ F(y_{n+1}) = \frac{1}{2} y_{n+1}^T M y_{n+1} + \left( \frac{\Delta t}{2} \right)^2 \left( \frac{1}{2} y_{n+1}^T K y_{n+1} + u_{n+1} \right) \]

\[ - y_{n+1}^T M[y_n + \Delta t \ddot{y}_n] + y_{n+1}^T [K y_n + \nu u_n] \left( \frac{\Delta t}{2} \right)^2 \]  

(5.257)

Then

\[ \delta y_{n+1}^T \frac{\partial F}{\partial y_{n+1}} = \delta y_{n+1}^T \left[ M y_{n+1} + \left( \frac{\Delta t}{2} \right)^2 \left[ K y_{n+1} + \nu u_{n+1} \right] \right. \]

\[ \left. - M[y_n + \Delta t \ddot{y}_n] + \left( \frac{\Delta t}{2} \right)^2 [K y_n + \nu u_n] \right] \]  

(5.258)
Thus necessary and sufficient conditions that equation (5.255) hold, are that (5.259) vanish for arbitrary "variations" \( \delta y_{n+1} \).

In order to force equations (5.252) to conserve energy we combine (5.250) and (5.257) through the use of a Lagrange multiplier \( \lambda \); our new functional is

\[
F(y_{n+1} + \lambda G(y_{n+1}))
\]

(5.260)

Thus

\[
\delta y_{n+1}^T \left[ -\frac{\delta F}{\delta y_{n+1}} + \lambda \frac{\delta G}{\delta y_{n+1}} \right] + \delta \lambda \cdot G(y_{n+1}) = 0 \tag{5.261}
\]

If (5.261) is to vanish for arbitrary \( \delta y_{n+1}, \delta \lambda \) then equation (5.257) is satisfied and in addition:

\[
M y_{n+1} + \left( \frac{\Delta t}{2} \right)^2 [K y_{n+1} + \nabla u_{n+1}] - M [\dot{y}_n + \Delta t \dot{y}_n^*] + \left( \frac{\Delta t}{2} \right) [K y_{n+1} + \nabla u_n] + \lambda \cdot M (y_{n+1} - y_n - \frac{\Delta t}{2} y_n) + \left( \frac{\Delta t}{2} \right)^2 [K y_{n+1} + \nabla u_{n+1}] = 0 \tag{5.262}
\]

Thus

\[
(1+\lambda) \left[ M y_{n+1} + \left( \frac{\Delta t}{2} \right)^2 [K y_{n+1} + \nabla u_{n+1}] \right] = (1+\lambda) [M y_n] + \Delta t \left( 1 + \frac{\lambda}{2} \right) M \dot{y}_n^* - \left( \frac{\Delta t}{2} \right)^2 (K y_n + \nabla u_n) \tag{5.263}
\]

The new algorithm thus consists of the two systems of equations, (5.261) and (5.263).

When \( y_{n+1} \) has been determined, \( \dot{y}_{n+1} \) 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
we have

$$\begin{align*}
\gamma_{n+1}^i - \gamma_n^i &= \Delta \gamma_{n+1}^i ; \quad \lambda_{n+1}^i - \lambda_n^i = \Delta \lambda_n^i \\
\end{align*}$$

\hspace{1cm} (5.264)

\begin{align*}
(\lambda_{n+1}^i) \left[ M \Delta \gamma_{n+1}^i + \left( \frac{\Delta t}{2} \right)^2 B(\gamma_{n+1}^i) \Delta \gamma_{n+1}^i \right] \\
+ \Delta \lambda^i \left[ M(\gamma_{n+1}^i - \gamma_n^i) + \left( \frac{\Delta t}{2} \right)^2 (K \gamma_{n+1}^i + v u_{n+1}^i) - \frac{\Delta t}{2} \frac{\partial}{\partial \gamma_n^i} M \gamma_n^i \right] \\
+ (\lambda_{n+1}^i) \left[ M(\gamma_{n+1}^i - \gamma_n^i) + \left( \frac{\Delta t}{2} \right)^2 (K \gamma_{n+1}^i + v u_{n+1}^i) \right] \\
+ \left( \frac{\Delta t}{2} \right)^2 [K \gamma_n^i + v u_{n+1}^i] - \left( 1 + \frac{\lambda_{n+1}^i}{2} \right) \gamma_n^i \Delta t = 0
\end{align*}

\hspace{1cm} (5.265)

$$\begin{align*}
\Delta \gamma_{n+1}^i T \left[ M(\gamma_{n+1}^i - \gamma_n^i - \frac{\Delta t}{2} \gamma_n^i) + \left( \frac{\Delta t}{2} \right)^2 \Delta \gamma_{n+1}^i [K \gamma_{n+1}^i + v u_{n+1}^i] \\
\end{align*}$$

\hspace{1cm} + G(\gamma_{n+1}^i) = 0

\hspace{1cm} (5.266)

where

$$B(\gamma_{n+1}^i) = K + J(\gamma_{n+1}^i)$$

\hspace{1cm} (5.267)

$$J = \frac{\partial}{\partial \gamma_n^i} v u_{n+1}^i$$

Equations (5.266), (5.267) are of the form

$$A_{12}^i b^i = b^i$$

\hspace{1cm} (5.268)

where

$$A_i = \begin{bmatrix} A_{11}^i & A_{12}^i \\ A_{21}^i & 0 \end{bmatrix}$$

$$z^i = \left( \frac{\Delta \gamma_{n+1}^i}{\Delta \lambda_n^i} \right) \quad b^i = \begin{bmatrix} b_{1i}^i \\ b_{2i}^i \end{bmatrix}$$

\hspace{1cm} (5.269)

where
\[ A_n^i = [1 + \lambda^i] [M + (\Delta t^2/2) B(y_{n+1}^i)] \]
\[ A_{12}^i = [M(y_{n+1}^i - y_n^i - \Delta t^2/2 \dot{y}_n^i) + (\Delta t^2/2) [Ky_{n+1}^i + vu_{n+1}^i]] \]
\[ A_{21}^i = (A_{12}^i)^T \]
\[ b_1^i = (1 + \lambda^i) [M(y_{n+1}^i - y_n^i) + (\Delta t^2/2) (Ky_{n+1}^i + vu_{n+1}^i)] \]
\[ + (\Delta t^2/2) (Ky_n^i + vu_n^i) - (1 + \lambda^i) \dot{y}_n^i \Delta t \]
\[ b_2^i = G(y_{n+1}^i) \]

Rewriting equation (5.268),
\[ A_{21}^i \Delta y_{n+1}^i = b_2^i \]
\[ A_{12}^i \delta y_{n+1}^i + A_{12}^i \Delta \lambda^i = b_1^i \] (5.271)

From the second equation in (5.271),
\[ \Delta y_{n+1}^i = (A_{11}^i)^{-1} (b_1^i - A_{12}^i \Delta \lambda^i) \] (5.272)

Using the first equation in (5.271),
\[ A_{21}^i \Delta y_{n+1}^i = A_{21}^i [A_{11}^i]^{-1} b_1^i - A_{21}^i (A_{11}^i)^{-1} A_{12}^i \Delta \lambda^i = b_2^i \]
\[ \Delta \lambda^i = [(A_{12}^i)^T (A_{11}^i)^{-1} b_1^i - b_2^i] / (A_{12}^i)^T (A_{11}^i)^{-1} A_{12}^i \] (5.273)
\[ \Delta y_{n+1}^i = (A_{11}^i)^{-1} [b_1^i - \Delta \lambda^i A_{11}^i] \] (5.274)

It should be observed that \( y_{n+1}^i \) 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:

\[
\begin{align*}
\dot{y}_{n+1} & = \frac{\Delta t}{2} [\dot{y}_{n+1} + \dot{y}_n] \\
\dot{y}_{n+1} & = -\frac{\Delta t}{2} M^{-1} [K(y_{n+1} + y_n) + \nu u_{n+1} + \nu u_n] \\
& \quad + \frac{\lambda}{1+\lambda} [M^{-1}(\nu y_n + \nu u_n) \frac{\Delta t}{2} - \dot{y}_n] \\
\frac{1}{2} \dot{y}_{n+1}^T M \dot{y}_{n+1} + \frac{1}{2} [y_{n+1}^T K y_{n+1} + 2u_{n+1}] \\
& = \frac{1}{2} y_n^T M y_n + \frac{1}{2} [y_n^T K y_n + 2u_n]
\end{align*}
\]

From the first two equations

\[
\begin{align*}
\frac{1}{2} [y_{n+1}^T M y_{n+1} - y_n^T M y_n] + \frac{1}{2} [y_{n+1}^T K y_{n+1} - y_n^T K y_n] \\
& + \frac{1}{2} (y_{n+1} - y_n)^T (\nu u_{n+1} + \nu u_n) = \frac{\lambda}{1+\lambda} (y_{n+1} - y_n)^T [K y_n + \nu u_n - \frac{2}{\Delta t} M \dot{y}_n]
\end{align*}
\]

Using the third equation of (5.275)

\[
\frac{\lambda}{1+\lambda} (y_{n+1} - y_n)^T [K y_n + \nu u_n - \frac{2}{\Delta t} M \dot{y}_n] \\
= \frac{1}{2} (y_{n+1} - y_n)^T (\nu u_{n+1} + \nu u_n) + u_n - u_{n+1}
\]

Thus
\[ \frac{\lambda}{1 + \lambda} \left( (K\nu_n + \nu u_n) \frac{\Delta t}{2} - M \nu_n \right) \]
\[ = C(y_{n+1} - y_n) \]  
(5.278)

where

\[ C = \frac{1}{2} \left[ J(\xi_1) - J(\xi_2) \right] \]
\[ J = \frac{\partial}{\partial y} \nu u \]
\[ \xi_i = y_n \theta_i + y_{n+1} (1 - \theta_i) \]
\[ 0 \leq \theta_i \leq 1, i = 1, 2 \]

(5.279)

Hence, equations (5.275) may be rewritten

\[ y_{n+1} - y_n = \frac{\Delta t}{2} \left[ \dot{y}_{n+1} + \dot{y}_n \right] \]
\[ \dot{y}_{n+1} - \dot{y}_n = -\frac{\Delta t}{2} M^{-1} [K(y_{n+1} + y_n) + \nu u_{n+1} + \nu u_n] \]
\[ + \frac{\Delta t}{2} M^{-1} C(y_{n+1} - y_n) \]  
(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:

\[ ||e_n|| = ||\dot{y}_n - x_n|| + ||y_n - x_n|| \leq D \Delta t^2 , \quad \text{as} \quad \Delta t \to 0 \]  
(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.,

$$
\begin{align*}
\mathbf{v}_{n+1} - \mathbf{v}_n &= \frac{-\Delta t}{2} \left[ \mathbf{v}_{n+1}^T (\mathbf{C} - \mathbf{zM}) \mathbf{v}_{n+1} + \mathbf{v}_n^T (\mathbf{C} - \mathbf{zM}) \mathbf{v}_n \\
&\quad + \mathbf{z}(\mathbf{v}_{n+1}^T \mathbf{Kv}_{n+1} + \mathbf{v}_n^T \mathbf{Kv}_n) \\
&\quad + \mathbf{v}_{n+1}^T \mathbf{v}_{n+1} + \mathbf{v}_n^T \mathbf{v}_n \right] \\
&\quad + (\mathbf{v}_{n+1} + \mathbf{zv}_n)^T \mathbf{p}_{n+1} + (\mathbf{v}_n + \mathbf{zv}_n)^T \mathbf{p}_n \right] \\
&\quad (5.282)
\end{align*}
$$

Replacing $\mathbf{v}_{n+1}$ by $\frac{2}{\Delta t} (\mathbf{v}_{n+1} - \mathbf{v}_n) - \mathbf{v}_n$, the constraint equation becomes:

$$
\begin{align*}
&\frac{1}{2} \left[ \mathbf{zv}_{n+1}^T (\mathbf{C} - \mathbf{zM}) \mathbf{v}_{n+1} + \left( \frac{2}{\Delta t} (\mathbf{v}_{n+1} - \mathbf{v}_n) - \mathbf{v}_n \right)^T \mathbf{M} \left( \frac{2}{\Delta t} (\mathbf{v}_{n+1} - \mathbf{v}_n) - \mathbf{v}_n \right) \\
&\quad + \mathbf{v}_{n+1}^T \mathbf{Kv}_{n+1} + 2u(\mathbf{v}_{n+1}) \right] \\
&\quad - \frac{1}{2} \left[ \mathbf{zv}_n^T (\mathbf{C} - \mathbf{zM}) \mathbf{v}_n + \left( \mathbf{v}_n + \mathbf{zv}_n \right)^T \mathbf{M} \left( \mathbf{v}_n + \mathbf{zv}_n \right) \\
&\quad + \mathbf{v}_n^T \mathbf{Kv}_n + 2u(\mathbf{v}_n) \right] \\
&\quad + \frac{\Delta t}{2} \left[ \left( \frac{2}{\Delta t} (\mathbf{v}_{n+1} - \mathbf{v}_n) - \mathbf{v}_n \right)^T (\mathbf{C} - 2\mathbf{M}) \left( \frac{2}{\Delta t} (\mathbf{v}_{n+1} - \mathbf{v}_n) - \mathbf{v}_n \right) \\
&\quad + \mathbf{v}_n^T (\mathbf{C} - \mathbf{zM}) \mathbf{v}_n + \mathbf{z}(\mathbf{v}_{n+1}^T \mathbf{Kv}_{n+1} + \mathbf{v}_n^T \mathbf{Kv}_n) \\
&\quad + \mathbf{z}(\mathbf{v}_{n+1}^T \mathbf{v}_{n+1} + \mathbf{v}_n^T \mathbf{v}_n) \\
&\quad + (\mathbf{v}_{n+1} + \mathbf{zv}_n)^T \mathbf{p}_{n+1} + (\mathbf{v}_n + \mathbf{zv}_n)^T \mathbf{p}_n \right] \\
&\quad = G(\mathbf{v}_{n+1}) \equiv 0 \quad (5.283)
\end{align*}
$$
The functional $F(y_{n+1})$ of equation (5.258) is replaced by:

$$F(y_{n+1}) = \frac{1}{2} \dot{y}_{n+1}^T M y_{n+1} + \frac{\Delta t}{4} y_{n+1}^T C y_{n+1}$$

$$+ \left( \frac{\Delta t}{2} \right)^2 \left( \frac{1}{2} \dot{y}_{n+1}^T C y_{n+1} + u_{n+1} \right)$$

$$- y_{n+1}^T [M(y_n + \Delta t \dot{y}_n) + \frac{\Delta t}{2} C y_n]$$

$$+ \left( \frac{\Delta t}{2} \right)^2 y_{n+1}^T [k y_n + v u_{n+1} - p_{n+1} - p_n] \tag{5.284}$$

The variation of the functional $(F(y_{n+1}) + \lambda G(y_{n+1}))$ yields the new algorithm:

$$\frac{\partial F}{\partial y_{n+1}} + \lambda \frac{\partial G}{\partial y_{n+1}} = 0$$

$$G(y_{n+1}) = 0 \tag{5.285}$$

$$\dot{y}_{n+1} = \frac{2}{\Delta t} (y_{n+1} - y_n) - \dot{y}_n$$

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:

\[ M\dddot{x} + C\dot{x} + Kx + v(u) = p(t) \quad (6.1) \]

Here, \( M \) is an \( N \times N \) symmetric positive definite matrix, \( C \) and \( K \) are \( N \times N \) symmetric positive semi-definite matrices, and \( v(u) \) 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 \( \Delta t \), but by \( (\omega_j \Delta t) \) where \( \omega_j \) is the highest "frequency" which can be excited. In order that the higher modes be integrated accurately, \( \Delta t \) 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 \( \Delta t \) 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:

\[
\begin{align*}
\ddot{x}_1 + 2z_1\dot{x}_1 + x_1 + \mu(x_1-x_2)^3 &= p_1 \cos \omega t + p_2 \sin \omega t + p_3 \cos 3\omega t \\
\ddot{x}_2 + 2z_2\dot{x}_2 + 9x_2 + \mu(x_2-x_1)^3 &= 0
\end{align*}
\] (6.2)

where

\[
\begin{align*}
p_1 &= (1-\omega^2)A + \frac{3}{4} \mu A^3 \\
p_2 &= -2z_1\omega A \\
p_3 &= \frac{1}{4} \mu A^3
\end{align*}
\] (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

\[x_1 = A \cos \omega t\] (6.4)

If we now turn to the second equation, regarding \(x_2\) as small,

\[\ddot{x}_2 + 2z_2\dot{x}_2 + 9x_2 + \mu x_1^3 = \frac{3}{4} \mu A^3 \cos \omega t + \frac{1}{4} \mu A^3 \cos 3\omega t\] (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
\[ y = M^{1/2} x \]  \hspace{1cm} (6.6)

Let
\[ C = M^{-1/2}CM^{-1/2} \]
\[ \mathcal{K} = M^{-1/2} K M^{-1/2} \]  \hspace{1cm} (6.7)

\[ U(x) = V(y) \]
\[ M^{-1/2} p(t) = q(t) \]

Using (6.6) and (6.7) in (6.1)

\[ I\ddot{y} + C\dot{y} + \mathcal{K}y + \nabla_y V(y) = q(t) \]  \hspace{1cm} (6.8)

\[ y(0) = a \quad \dot{y}(0) = b \]

Let $T$ be the orthogonal matrix which diagonalizes $\mathcal{K}$.

Let
\[ y = T z \]
\[ T^T \mathcal{K} T = A \]
\[ V(y) = W(z) \]  \hspace{1cm} (6.9)
\[ \varphi = T^T C T \]
\[ T^T q = f(t) = \{f_i(t)\} \]
Unless (6.1) has classical normal modes, \( \mathcal{M} \) is not diagonal. Using (6.9) in (6.8),

\[
\begin{align*}
I \ddot{z} + \mathcal{M} \dot{z} + \Lambda z + \nabla z \cdot W(z) &= f(t) \\
\dot{z}(0) &= \bar{a} \\
\dot{z} &= \bar{b}
\end{align*}
\]

(6.10)

Suppose that \( |f_j(t)| < \varepsilon \) for \( j > P \), \( \varepsilon \ll 1 \).

Let

\[
\begin{align*}
Z_p &= \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_p \end{bmatrix} \\
(6.11)
\end{align*}
\]

If we suppose that the \( j \)th modes, \( j > P \) are at most weakly excited, let us set \( z_j = 0 \), \( j \notin (P+1,N) \). Then,

\[
W(z) = W(z_p)
\]

(6.12)

equation (6.10) becomes

\[
\begin{align*}
I \ddot{z}_1 &= f_1(t) \\
I \ddot{z}_2 &= \mathcal{M}_{22} \dot{z}_2 + \Lambda_2 z_2 + \nabla z_2 \cdot W(z_2) = f_2(t)
\end{align*}
\]

(6.13)

where

\[
\begin{align*}
Z_1 &= \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_6 \end{bmatrix} \text{ represents the rigid body modes} \\
Z_2 &= \begin{bmatrix} z_7 \\ z_8 \\ \vdots \\ z_p \end{bmatrix} \text{ represents the first (P-6) flexible modes.}
\end{align*}
\]
$\omega_{22} = \omega_{22}^T$ is the $(P-6) \times (P-6)$ damping metric associated with the $z_p$ modes.

$$A_2 = \begin{pmatrix}
\lambda_7 \\
\lambda_8 \\
\vdots \\
\lambda_p \\
\end{pmatrix} = \begin{pmatrix}
\omega_7 \\
\omega_8 \\
\vdots \\
\omega_p \\
\end{pmatrix} \quad (6.14)$$

If $P$ is not too large, we can select a $\Delta t$ such that $(\omega_p \Delta t) \sim 0.1$, then using any of the "energy conserving" algorithms of Section 5, equations (6.13) can be integrated with good accuracy. Having determined $z_p$, one can easily compute the physical coordinates,

$$x_p = M^{-1/2} T_p z_p$$

where $T_p$ is the $N \times P$ 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

$$\ddot{z}_j + \omega_j^2 z_j + \lambda_j z_j = - \frac{\partial W(z)}{\partial z_j} |_{z=z_2} \quad (6.15)$$

$$j \in (P+1,N)$$

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

$$|\lambda_j - \frac{p\lambda_k}{q}| \ll 1 \quad (6.16)$$

where $p$ and $q$ are integers and $\lambda_k$ is an element of $A_2$. 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 $(w_k \Delta t) \sim 0.1$ 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

$$z_3 = \begin{pmatrix} z_k \\ z_\lambda \\ \vdots \\ z_m \end{pmatrix}$$

(6.17)

where $k, \lambda, m$ are the modes of (6.15) showing significant behavior.

Let

$$z_4 = \begin{pmatrix} z_2 \\ z_3 \end{pmatrix}$$

(6.18)

Equation (6.10) may, in this case, be written

$$\begin{cases} I \ddot{z}_2 + D_{22} \dot{z}_2 + D_{23} \dot{z}_3 + V_{z_2} W(z_4) + \Lambda_2 z_2 = f_2(t) \\ I \ddot{z}_3 + D_{32} \dot{z}_2 + D_{33} \dot{z}_3 + V_{z_3} W(z_4) + \Lambda_3 z_3 = 0 \end{cases}$$

(6.19)

This first set of equations is integrated using a $\Delta t_1$ appropriate to the highest eigenvalue in $\Lambda_2$. The second set of equations is integrated using a $\Delta t_2$ appropriate to the highest eigenvalue in $\Lambda_3$, say $\Delta t_2 = \frac{1}{K} \Delta t_1$, $K$ an
integer. The values of ε₂ 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

\[ x_{n+1} = A(n)x_n \quad |A(n)| \neq 0, \quad ||A(n)|| < \infty \quad n > n_0 \quad (A1) \]

then A1 is uniformly Liapunov asymptotically stable at \( x = 0 \) iff there exists a bounded, symmetric, positive definite matrix \( P(n) \) such that,

i) \( P(n) = P^T(n) \) positive definite and bounded above & below

ii) \( A^T(n) P(n+1) A(n) - P(n) = -\Theta(n) \) \( (A2) \)

iii) \( \Theta(n) = \Theta^T(n) \) positive definite

iv) \( ||\Theta(n)|| \leq M_2(n_0) \quad \forall n > n_0 \) and \( \forall n_0 \)

Proof

Sufficiency

Suppose that there exists such a matrix P(n) satisfying A2

Let \( V_n = x_n^T P(n)x_n \quad (A3) \)

Since \( P(n) \) is positive definite and bounded,

i) \( V_n > 0 \)

ii) \( V_n \leq M_3 x_n^T x_n \quad M_3 < \infty \)

\[ V_{n+1} = x_{n+1}^T P(n+1) x_{n+1} \quad (A4) \]

Using (A1),

\[ V_{n+1} = x_n^T A^T(n) P(n+1) A(n) x_n \quad (A5) \]

\[ \Delta V_n = (V_{n+1} - V_n) = x_n^T (A^T(n) P(n+1) A(n) - P(n)) x_n \quad (A6) \]
Using (A2)

\[ \Delta V_n = -x_n^T \theta(n)x_n < 0 \]  \hfill (A7)

\[ V_{n+1} < V_n < V_{n-1} < \cdots < V_1 < V_0 \]  \hfill (A8)

Since \( P(n) \) is bounded \( V_n \), \( V_n \) is finite if \( \|x_n\| \) is, and since \( V_n \) is zero only if \( x_n = 0 \), hence \( V_n \) and therefore \( \|x_n\| \) tends to zero as \( n \) tends to infinity. Since the result is independent of \( n_0 \), the trivial solution \( x = 0 \) 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:

\[ P(n) = \sum_{j=n}^\infty \phi(j,n)^T \theta(j) \phi(j,n) \]  \hfill (A9)

Thus if (A1) is uniformly asymptotically stable at the origin

\[ \|\phi(j,n)\| \leq M_1 \delta^{j-n} \]  \hfill (A10)

\[ 0 < \delta < 1 \quad , \quad j > n \quad , \quad \forall n > n_0 \]

Using (A2iv)

i) \( \|P(n)\| \leq \sum_{j=n}^\infty H_1^2 H_2(n_0) \delta^{2(j-n)} \)  \hfill (A11)

\[ \leq \frac{H_1^2 H_2(n_0)}{1 - \delta^2} < \infty \]  \hfill (A12)

\[ P^T(n) = (\sum_{j=n}^\infty \phi^T(j,n) \theta(j) \phi(j,n))^T = P(n) \]  \hfill (A13)
iii) \( \mathbf{x}^T P(n) \mathbf{x} = \sum_{j=n}^{\infty} (\Phi(j,n) \mathbf{x})^T \Theta(j)(\Phi(j,n) \mathbf{x}) > 0 \)

if \( \Phi(j,n) \mathbf{x} \neq 0 \)

and since \( |\Phi(j,n)| \neq 0 \),

\[ \mathbf{x}^T P(n) \mathbf{x} > 0 \quad \mathbf{x} \neq 0 \]  \hspace{1cm} (A14)

Thus if (A1) is uniformly L.A.S. at \( \mathbf{x} = 0 \), 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 (A9), which requires the unknown transition matrix \( \Phi(j,n) \). Its main use is in proving theorems.
References


Bibliography

