NGR-40-002-015
Brown U.

Some Aspects of Difference Equations

by

James Hurt

B.S., University of Iowa, 1961

M.S., University of Iowa, 1963

Thesis

submitted in partial fulfillment of the requirements for the

Degree of Doctor of Philosophy in the Division of

Applied Mathematics at Brown University

June, 1967

NGR-40-002-015
Brown U.

Some Aspects of Difference Equations

by

James Hurt

B.S., University of Iowa, 1961

M.S., University of Iowa, 1963

Thesis

submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in the Division of Applied Mathematics at Brown University

June, 1967

This thesis by James Hurt

is accepted in its present form by the Division of
Applied Mathematics as satisfying the
thesis requirement for the degree of Doctor of Philosophy

Date: . . . .

Recommended to the Graduate Council

Date: . . . .

. . . .

Date: . . . .

. . . .

Approved by the Graduate Council

Date: . . . .

Vita

The author of this thesis was born on [redacted] He graduated from University High School in Iowa City in June of 1957. He attended the University of Iowa, Iowa City, Iowa, from 1957 until 1963, receiving a Bachelor of Science degree (with distinction) in Mechanical Engineering in June of 1961 and a Master of Science degree in Engineering Mechanics in August of 1963. While attending the University of Iowa, he was elected to membership in the Pi Tau Sigma and Tau Beta Pi national honorary fraternities and in the Society of the Sigma Xi.

Durning the 1963-1964 school year, the author was a full time instructor at the University of Iowa. During this period he co-authored with Royce Beckett the book Numerical Calculations and Algorithms, published by McGraw-Hill, New York, 1967.

In the fall of 1964, the author entered Brown University on a Brown University fellowship and received NASA Traineeships for the following two years while working for his Ph.D. degree. During this time he wrote a paper, "Some Stability Theorems for Difference Equations", which has been submitted to the SIAM Journal for Numerical Analysis for publication.

Acknowledgement

The author of this thesis wishes to thank all those who have helped him so greatly in the preparation of this thesis. Special thanks go to his thesis advisor, Dr. Jack K. Hale, for the countless hours of patient listening and for all the pertinent advice he gave to the author. Thanks are also due to Drs. J. P. LaSalle and E. F. Infante for their critizisms and helpful comments on the contents of this thesis. The Traineeship program of the National Aeronautics and Space Administration and Brown University are thanked for their contribution of money and office space.

Thanks are also due in no small measure the KSue Brinson for her diligence in typing the manuscript and to my wife, Gretchen, for her forbearance.

Table of Contents

Abstract

In recent years there has been a great interest in the study of methods for computing the solution of an initial value problem numerically. Most methods used form a difference analog to the differential equation and use the solution of the difference analog as an approximation for the solution of the differential equation. Most of the work done to date is concerned with estimating the difference between the true solution and the computed solution on a finite interval. Such problems as determining stability properties of the differential equation or finding a periodic solution are intrinsically involved with unbounded intervals. When a difference analog is used in these cases, the question becomes not one of how accurate is the approximation on a finite interval but do the solutions of the difference analog display the same properties as do the solutions of the differential equation.

The questions concerning stability properties of the difference analog led to research which is reported in a paper by the author entitled "Some Stability Theorems for Difference Equations". This paper has been submitted to the SIAM Journal on Numerical Analysis for publication. The theorems in this paper have application to other areas of Numerical Analysis, notably iteration theory, and to sampled data control systems.

In this thesis, the problem of finding periodic, almost periodic, and bounded solutions of a difference analog to a differential equation are considered. In the introduction, several examples are given to show how the behavior of a difference analog can be different from that of the differential equation. The problem of periodic solutions of the difference analog is discussed and, in order to have the period depend continuously on the parameters

in the difference analog, the solutions must be extended from the integers to the reals. Definitions of and theorems related to solutions on the reals are given.

The properties of solutions of Linear and Affine difference equations are discussed and a noncritical case is treated in some detail. The noncritical case considered is when the linear part has a constant matrix which has no eigenvalues on the unit circle. A theorem is presented concerning the existence of a periodic, almost periodic, or bounded solution of a noncritical difference equation when the nonlinearity has a small parameter.

The simplest critical case, where the matrix of the linear part is the identity matrix, is discussed and a theorem is given concerning the existence of a T-periodic solution where T is a rational number. Finally, an averaging method and an invariant manifold method are given, each with an application which proves that, under certain conditions, certain difference analogs do display the same behavior as the corresponding differential equation. The Crude Euler analog of $\dot{x} = \epsilon f(t,x)$ has an almost periodic solution for ϵ and h small enough under the same conditions which insure the existence of an almost periodic solution of the differential equation. The central difference analog to the Van der Pol equation $\ddot{x} + 2\epsilon\dot{x}(x^2-1)+x = 0$ has an invariant curve of radius approximately 2 for $\epsilon > 0$ and small.

INTRODUCTION.

Consider the initial value problem given in (1) where x and

$$\dot{x} = f(t,x) \qquad x(t_0) = x_0 \tag{1}$$

$f(t,x)$ are n-vectors (n-dimensional vectors), t is a real number called time and $\dot{x} = dx/dt$. In recent years there has been a great interest in the study of methods for computing the solution to (1) numerically [1,2]. Most numerical methods divide time into increments with spacing h such that $t_k = t_0 + kh$ and form the difference analog (2).

$$x(t_{k+1}) = g(t_k, x(t_k), x(t_{k-1}), \ldots, x(t_{k-N}); h) \tag{2}$$

$$x(t_k) = x_k \quad , \quad 0 \leq k \leq N$$

The solution to the difference equation (2) is then taken as an approximation to the solution of (1). However, most of the work to date is concerned with computing the solution on a finite interval, say for $t_0 \leq t \leq t_1$. Such problems as determining stability properties of (1) or finding periodic solutions of (1) are intrinsically involved with the unbounded intervals $t_0 \leq t < \infty$ or $-\infty < t < \infty$. When the difference analog (2) is used in these cases, the question becomes not one of how accurate is the approximation on a finite interval but do the solutions of (2) display the same properties as the solutions of (1)?

Several very simple cases will be covered here as a partial answer to this question. Consider first the difference equation (3).

$$x(k+1) = Ax(k) \tag{3}$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad x(k) = \begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix}$$

If the eigenvalues of the real matrix A are λ_1 and λ_2, $\lambda_1 \neq \lambda_2$, then all solutions of (3) are of the form

$$x(k) = c\lambda_1^k + d\lambda_2^k$$

where c and d are constant 2-vectors. The origin $x(k) = 0$ is also a solution. The behavior of the solutions near the origin is completely determined by the values of λ_1 and λ_2. If $|\lambda_1| < 1$ and $|\lambda_2| < 1$, then all solutions approach the origin as $k \to \infty$. If λ_1 and λ_2 are real, then the origin is said to be a stable node (SN). If λ_1 and λ_2 are complex, then the origin is said to be a stable focus (SF). If $|\lambda_1| = |\lambda_2| = 1$, then all solutions remain bounded and the origin is said to be a center (C). If $|\lambda_1| > 1$ and $|\lambda_2| > 1$, then all solutions become unbounded as $k \to \infty$. If λ_1 and λ_2 are real, the origin is said to be an unstable node (UN) and if the eigenvalues are complex, the origin is said to be an unstable focus (UF). If $|\lambda_1| > 1$ and $|\lambda_2| < 1$, then the origin is said to be a saddle point (SP). The terms stable node, stable focus, center, etc., come from differential equations and are used here because the behavior of the solutions of the difference equation is very similar to the behavior of solutions of a differential equation.

For example, consider the difference equation (3) when the matrix A is given by

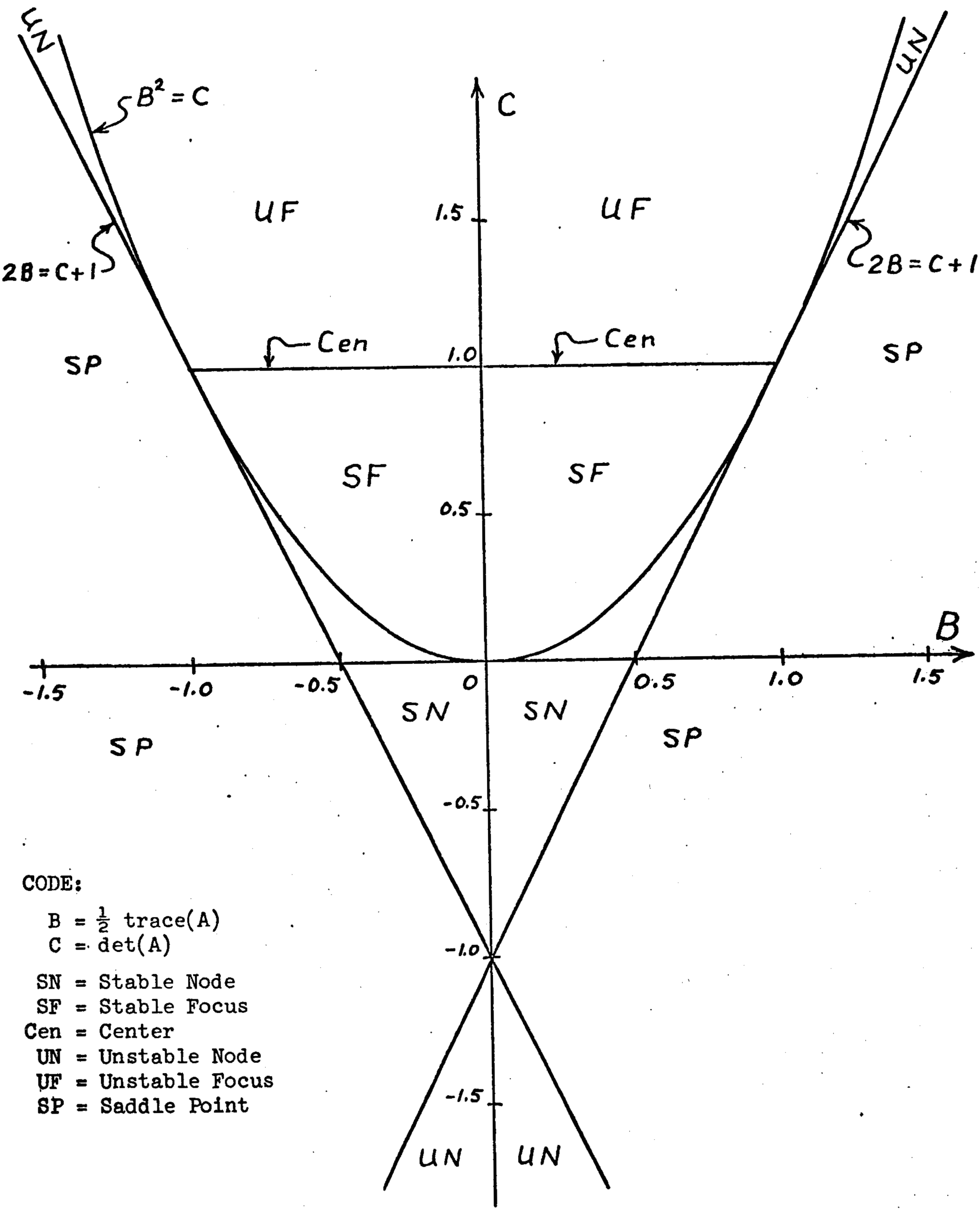
$$A = \begin{pmatrix} 0 & 1 \\ -1 & 2\cos\omega \end{pmatrix} . \tag{3a}$$

The eigenvalues of A are $\lambda_1 = e^{i\omega}$ and $\lambda_2 = e^{-i\omega}$, and so the origin is a center. Consider the functional $V(x)$

$$V(x) = x_1^2 - 2\cos\omega\, x_1 x_2 + x_2^2 .$$

We see immediately that $V(x(k+1)) = V(x(k))$ for any solution $x(k)$ and hence that any solution remains on a level surface of V. A level surface of V is an ellipse and hence the name center for this type of behavior.

The eigenvalues of A depend entirely on two numbers, $B = +\frac{1}{2}\,\text{trace}\,(A) = +\frac{1}{2}(a_{11} + a_{22})$ and $C = \det\,(A) = a_{11}a_{22} - a_{12}a_{21}$. Figure 1 is a graph showing the various regions of behavior in the B,C plane. It should be noted that, on the line $B^2 = C$, $\lambda_1 = \lambda_2$ and there is the possibility of A not having simple elementary divisors. However, since this line is always a transition line between two different types of behavior, this raises no serious question at this time.



CODE:

$B = \frac{1}{2}$ trace(A)
C = det(A)
SN = Stable Node
SF = Stable Focus
Cen = Center
UN = Unstable Node
UF = Unstable Focus
SP = Saddle Point

Figure 1

Regions of Behavior for x(k+1) = Ax(k)

Using the results shown in Figure 1, several difference analogs of the differential equation (4) are studied.

$$\ddot{x} + 2d\dot{x} + x = 0 \tag{4}$$

In most cases, the differential equation (4) will be written in the vector form of (4').

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 - 2dx_2 \end{aligned} \tag{4'}$$

The behavior of the solutions of (4) depend only on the parameter d. The origin is a center if $d = 0$, a stable focus for $0 < d < 1$ and a stable node for $1 < d < \infty$. The origin is an unstable focus for $-1 < d < 0$ and an unstable node for $-\infty < d < -1$. All the difference analogs will have an additional parameter, the spacing h, and the regions of the d,h plane for each type of behavior for the difference analog should be compared to this ideal. The variations will be noted.

The first numerical method to be considered is the central difference analog where the derivatives of x in (4) are replaced by their central difference approximations.

$$\frac{1}{h^2}\left(x(t_{k+1}) - 2x(t_k) + x(t_{k-1})\right) + \frac{2d}{2h}\left(x(t_{k+1}) - x(t_{k-1})\right) + x(t_k) = 0$$

Letting $y_1(k) = x(t_{k-1})$ and $y_2(k) = x(t_k)$, this becomes the difference equation (5)

$$y(k+1) = Ay(k) \tag{5}$$

where

$$y(k) = \begin{pmatrix} y_1(k) \\ y_2(k) \end{pmatrix} \qquad A = \begin{pmatrix} 0 & 1 \\ -\frac{1-hd}{1+hd} & \frac{2-h^2}{1+hd} \end{pmatrix} \tag{6}$$

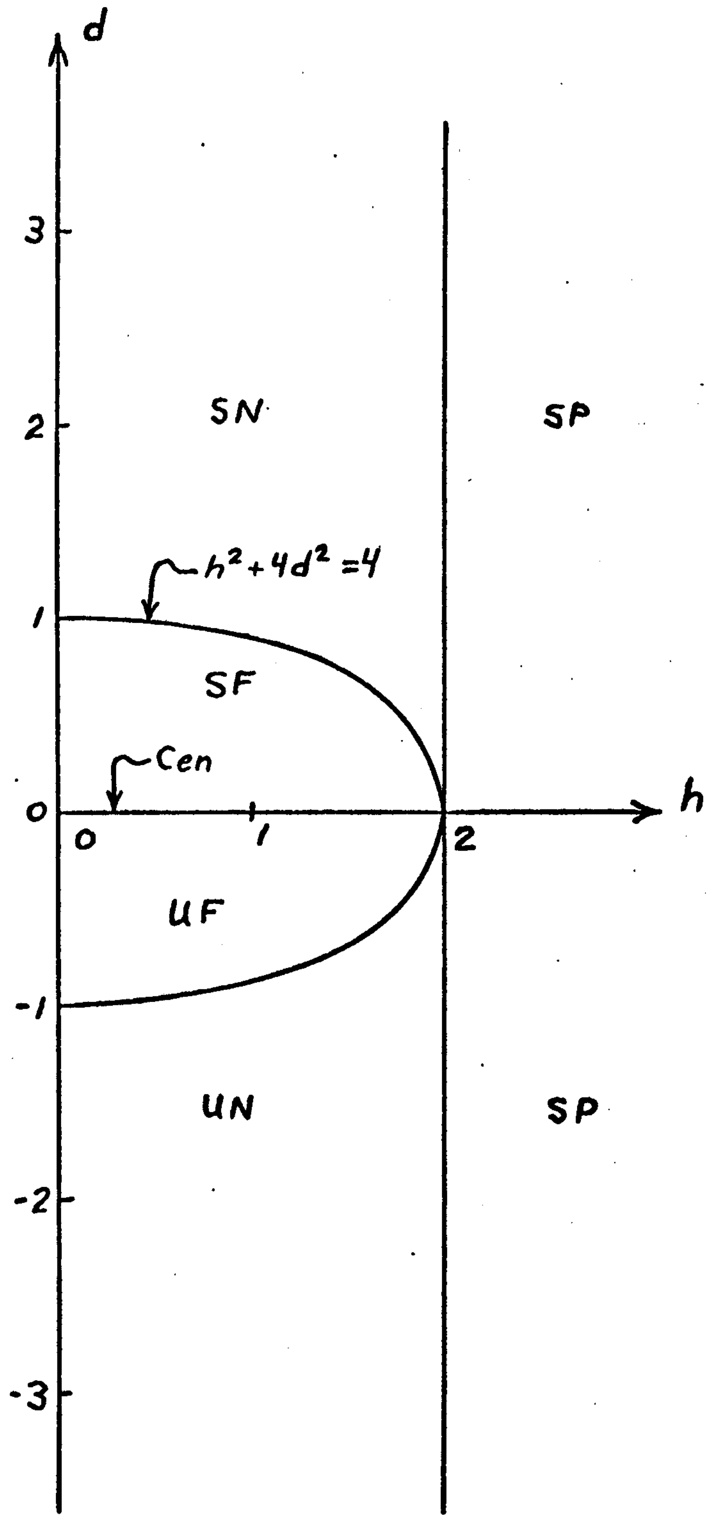
The parameters are

$$B = \frac{1}{2}\,\frac{2-h^2}{1+hd} \qquad \text{and} \qquad C = \frac{1-hd}{1+hd} .$$

The regions of behavior are shown in Figure 2. Notice that for $h > 2$, the behavior is always that of a saddle point, a behavior which does not at all resemble any of the behaviors of solutions of the differential equation. However, the main region of interest is for h very small. Figure 2 shows the behavior of solutions of (5) for large h because this gives a clearer picture of what happens for small h.

For small h, we notice that the origin is a center when $d = 0$, is stable when $d > 0$ and unstable when $d < 0$ --- in agreement with the origin of the differential equation. However, note that the transition from Focal behavior to Nodel behavior does not occur at $d = \pm 1$ but on the ellipse $4d^2 + h^2 = 4$. Still, from this point of view, the central difference analog is a good method for computing a solution to (4) numerically.

The next method considered is the Crude Euler method [3]. This method uses equation (4') and is described by the difference equation (5)



CODE:

h = Step size
d = Damping coefficient
SN = Stable Node
SF = Stable Focus
Cen = Center
UN = Unstable Node
UF = Unstable Focus
SP = Saddle Point

Figure 2

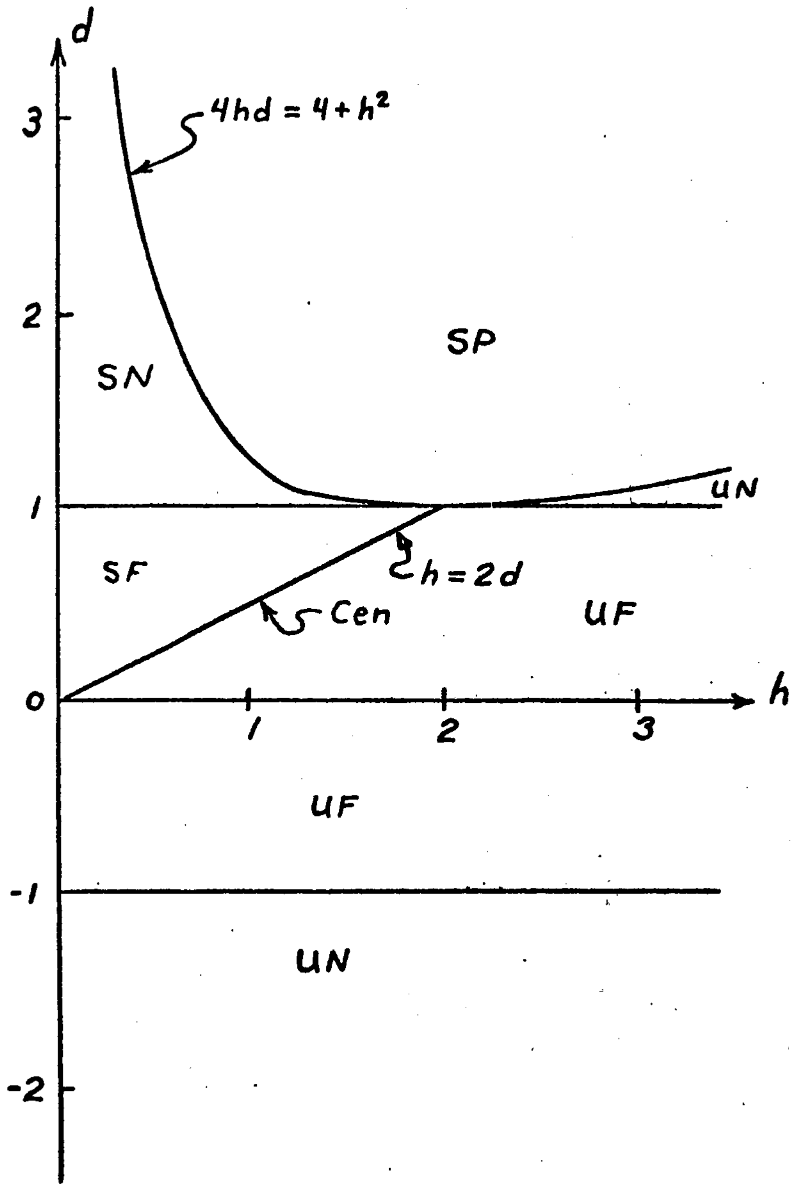
Regions of Behavior - Central Difference Analog

where

$$y(k) = \begin{pmatrix} x_1(t_k) \\ x_2(t_k) \end{pmatrix} \qquad A = \begin{pmatrix} 1 & h \\ -h & 1-2hd \end{pmatrix} \qquad (7)$$

The parameters are $B = 1-hd$ and $C = 1-2hd+h^2$. The regions of behavior are shown in Figure 3. Notice that, for $h > 2$, the behavior is always unstable. Also, for small h and $d = 0$, the behavior is not that of a center but that of an unstable focus. If h is small and $0 < d < \frac{1}{2}h$, the behavior of the difference equation is that of an unstable focus instead of a stable focus. For large d, it is possible to get unstable behavior of the difference equation by having h large enough to be in the saddle point region even though it is still small. In order to get the desired behavior in this case, we must take h much smaller than $2d$ if d is small and much smaller than $1/d$ if d is large. Notice that, for $d = 0$, the Crude Euler method behaves like an unstable focus for any $h > 0$ and not like the center for the differential equation. From this point of view, this method is not very acceptable. However, other considerations must be used in passing judgement on any given method.

In the Corrected Euler method [3], the value of $y(k+1)$ obtained from the Crude Euler method is used as a first approximation to the solution at t_{k+1}. This first approximation is used along with the Trapezoidal rule of integration to obtain a second approximation which is used as the numeric solution. This difference analog of (4') is given by (5) where



CODE:

h = Step size
d = Damping coefficient
SN = Stable Node
SF = Stable Focus
Cen = Center
UN = Unstable Node
UF = Unstable Focus
SP = Saddle Point

Figure 3

Regions of Behavior - Crude Euler Method

$$y(k) = \begin{pmatrix} x_1(t_k) \\ x_2(t_k) \end{pmatrix} \qquad A = \begin{pmatrix} 1-\frac{1}{2}h^2 & h(1-hd) \\ -h(1-hd) & 1-\frac{1}{2}h^2-2hd(1-hd) \end{pmatrix} \qquad (8)$$

The parameters are $B = 1-\frac{1}{2}h^2-hc(1-hc)$ and $C = (1-\frac{1}{2}h^2)^2-2hc(1-hc)(1-\frac{1}{2}h^2)+h^2(1-hc)^2$. The regions of behavior are shown in Figure 4. Notice that, while the line of centers is no longer a straight line and comes much closer to the h-axis, this figure is not much different from Figure 3 for the Crude Euler method and that the same comments apply.

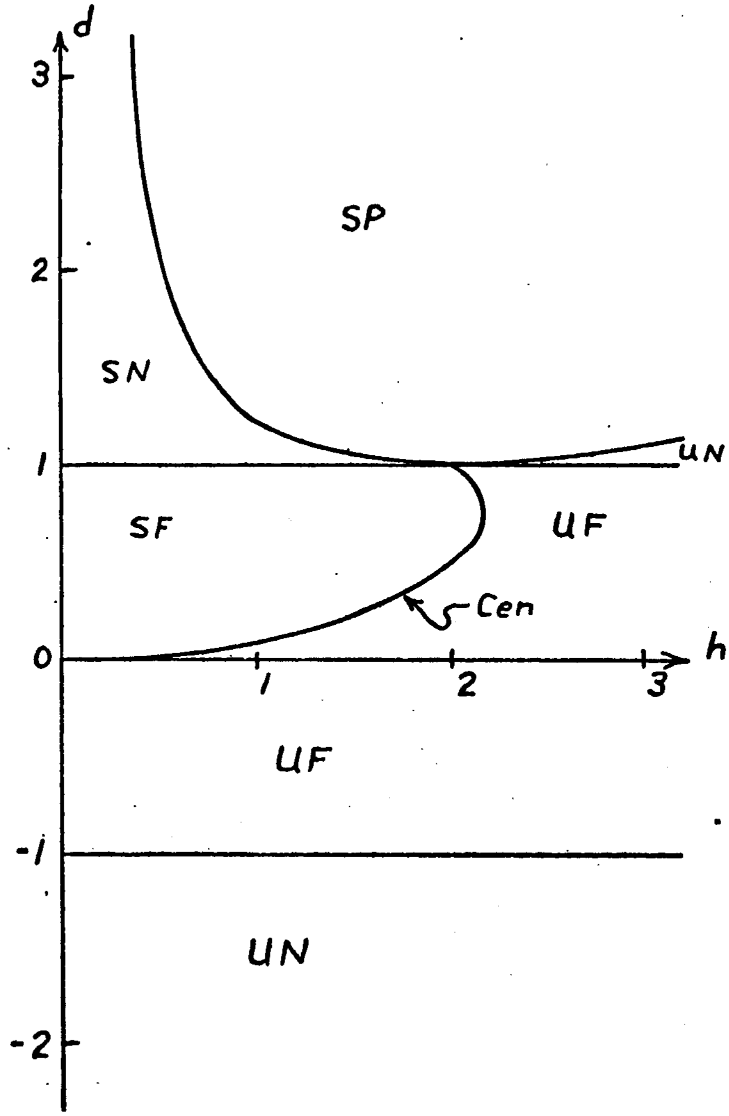
The final method discussed here is the Iterated Euler method [3], sometimes called the Modified Euler method. In this method, the trapezoidal rule of integration is used to obtain second, third, etc., approximations to the solution at $t = t_{k+1}$ until these iterations converge. The difference analog of (4') is given by equation (5) where

$$y(k) = \begin{pmatrix} x_1(t_k) \\ x_2(t_k) \end{pmatrix} \qquad A = \begin{pmatrix} \frac{4+4hd-h^2}{4+4hd+h^2} & \frac{4h}{4+4hd+h^2} \\ -\frac{4h}{4+4hd+h^2} & \frac{4-4hd-h^2}{4+4hd+h^2} \end{pmatrix} \qquad (9)$$

The parameters are

$$B = \frac{4-h^2}{4+4hd+h^2} \qquad \text{and} \qquad C = \frac{(4+h^2)^2-16h^2d^2}{(4+4hd+h^2)^2} .$$

The regions of behavior are shown in Figure 5. The only discrepancy between this figure and the ideal figure is the saddle point region where there should be an unstable node. However, for $d < -1$, if we take h much smaller than

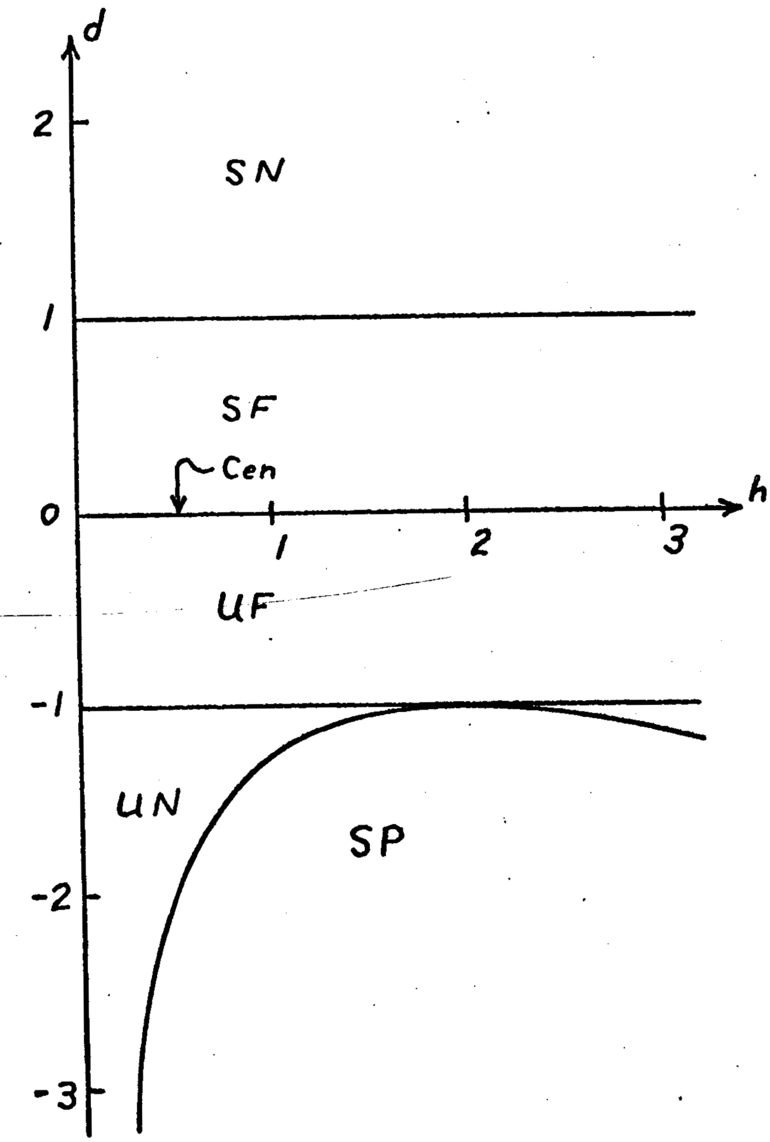


CODE:

h = Step size
d = Damping coefficient
SN = Stable Node
SF = Stable Focus
Cen = Center
UN = Unstable Node
UF = Unstable Focus
SP = Saddle Point

Figure 4

Regions of Behavior - Corrected Euler Method



CODE:

h = Step size
d = Damping coefficient
SN = Stable Node
SF = Stable Focus
Cen = Center
UN = Unstable Node
UF = Unstable Focus
SP = Saddle Point

Figure 5

Regions of Behavior - Iterated Euler Method

$-1/d$, then the behavior will be correctly that of an unstable node.

The significance of these examples is that, in most cases, the behavior of solutions of a difference analog to a differential equation is not the same as the behavior of solutions of the differential equation. In the case of a more general differential equation, this variation of behavior can be detected only by computing the solution twice, once with $h = h_o$ and once with a considerably different h, say $h = \frac{1}{2}h_o$ or $h = \frac{1}{10} h_o$. If the behavior of the solutions does not change drastically from one computed solution to another, then one may assume that the behavior of solutions of the difference analog and of the differential equation are the same. If the behavior does change drastically, even if the basic type of behavior doesn't change, then one must assume that the behavior of solutions of the difference analog and of the differential equation are different and the solution should be recomputed with another value of h. For example, if $d = 0$ in equation (4) and we were using the Crude Euler method to compute a solution, we would note that the computed solution is spiralling outward like an unstable focus. When the solution is recomputed with half the original spacing, we still see the outward spiral but its rate of spiralling out will be reduced by about half, leading to the conclusion that the behavior of the differential equation is not that of an unstable focus.

A greater analysis of the center of a difference equation is in order. Consider the difference equation (3) where the matrix A is as given in (3a). This happens to be the matrix A given for the central difference analog in equation (6) with $d = 0$ and $2 \cos \omega = 2-h^2$, but we shall use it in the form given in (3) and (3a). As noted before, all solutions remain on a level surface of $V(x) = x_1^2 - 2 \cos \omega \, x_1 x_2 + x_2^2$. One such level surface, for

$\omega = 2\pi/9$ and $x_1(0) = x_2(0) = 1$, is the ellipse shown in Figure 6. The tic-marks on the ellipse show the points of the solution starting at $x_1(0) = x_2(0) = 1$. Since $x(9) = x(0)$, this solution is periodic with period 9. Any point on this ellipse can be used as the starting point $x(0)$ and will lead to a 9-periodic solution.

Suppose now that ω was changed slightly, say from $\omega = 2\pi/9 = 10\pi/45$ to $\omega = 10\pi/46$. The ellipse $V(x) =$ constant will change very slightly but each solution will now take on 46 distinct points on the ellipse instead of 9 and we will have $x(46) = x(0)$, leading us to say that the solution is 46-periodic. A change in the parameter ω of $\pi/207 \doteq .0015$ changes the period by a factor of 5. In fact, if ω is an irrational multiple of π, there will not be any integer k such that $x(k) = x(0)$ and so the solution is not periodic. Yet, each point $x(k)$ of the solution will be on the level surface of $V(x) = V(x(0))$ and the points $x(k)$ will move around the ellipse in the same manner as shown in Figure 6.

Assume that the eigenvalues of A are $e^{i\omega}$ and $e^{-i\omega}$ where $e^{i\omega} \neq e^{-i\omega}$. Then each solution of (3) can be written as in (10)

$$x(k) = a \cos \omega k + b \sin \omega k \qquad (10)$$

where a and b are constant 2-vectors which are completely determined by $x(0)$. Letting $k = 0$, we get $a = x(0)$. We determine b from the equation $x(1) = Ax(0)$.

$$b = \frac{(A - \cos \omega\, I)}{\sin \omega}\, x(0)$$

Since $e^{i\omega} \neq e^{-i\omega}$, $\sin \omega \neq 0$ and b is uniquely determined.

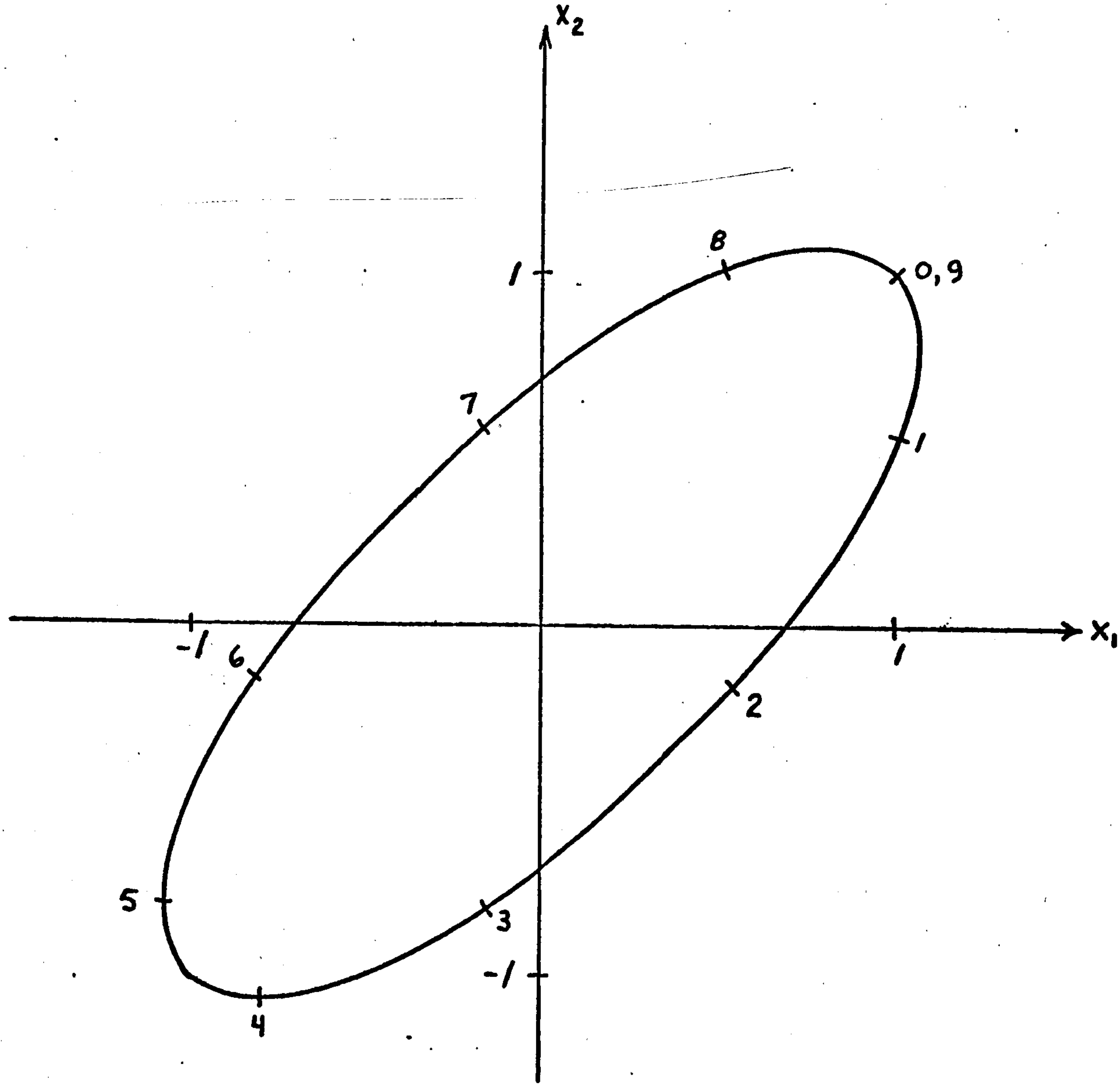


Figure 6

A Level Surface of $V(x) = x_1^2 - 2\cos(2\pi/9)x_1x_2 + x_2^2$

Now, $x(k) = x(0)$ only when $\cos \omega k = 1$, i.e., when ωk is some integer multiple of 2π. Then $\omega = 2\pi/9$, then $9\omega = 2\pi$ and the solution is 9-periodic. When $\omega = 10\pi/46$, then $46\omega = 10\pi$ and the solution is 46-periodic. Notice that, in both cases, the period k is the lowest integer such that $x(k) = x(0)$. When ω is an irrational multiple of 2 , then ωk is never equal to an integer multiple of 2π and the solution is not periodic. Yet equation (10) gives the formula for a $2\pi/\omega$-periodic function. Why can't we say that the solution (10) is $2\pi/\omega$-periodic for any ω? If we could, then the period would be continuous in the parameter ω (except when $e^{i\omega} = e^{-i\omega}$).

The problem arises because the solution is defined only on the integers. What is meant by the statement, "x is T-periodic" when x is defined only on the integers and T is not an integer? Since $x(k+T)$ is not defined whenever k is an integer and T is not, we cannot say x is T-periodic whenever $x(k+T) = x(k)$. If we could define equation (10) as a solution of (3) for all real k, then we could say x is T-periodic whenever $x(k+T) = x(k)$. However, (10) is a solution of (3) not only for a and b constant 2-vectors, but also for a and b any 1-periodic functions, as in (10').

$$x(k) = a(k) \cos \omega k + b(k) \sin \omega k \qquad (10')$$

It is easily confirmed by substitution that a and b are no longer determined by $x(0)$ but that $x(k)$ is arbitrary for $0 \leq k < 1$. If $x(k)$ is defined (as initial value) for $0 \leq k < 1$, the $a(k)$ (hence $b(k)$) is completely determined and (10') represents the solution with the initial values. Notice that, by defining the initial values for the solution on an interval instead of at a point, we get the solution defined for all real k instead of only on the integers and we can logically talk about T-periodic solutions for non-integer T. In (10') the vector $b(k)$ is given in terms of $a(k)$ by

$$b(k) = \frac{(A - \cos \omega\, I)}{\sin \omega}\, a(k) \qquad 0 \leq k < 1 \; .$$

Just what does this "new" definition of a solution mean in terms of the simple difference equation (3)? Two different solutions of (3) for $\omega = 2\pi/9$ are shown in Figure 7. The solid line is the solution which happens to have $a(k)$ and $b(k)$ constant and the dotted line has $a(k)$ and $b(k)$ non-constant. Notice that, on the integers (marked by the tics), both solutions look the same and that both solutions are 9-periodic.

This definition of a solution of a difference equation is formalized here. To emphasis that the solution is defined on some real interval and not on only the integers, t is used as the independent variable instead of k. The difference equation under consideration is given by equation (11) where each x and $f(t,x)$ is in E^n, an n-dimensional vector space.

$$x(t+1) = f(t, x(t)) \tag{11}$$

DEFINITION: A function $x(t) = x(t; t_o, x_o)$ where t and t_o are real numbers, $t_o \leq t < t_o + T$, some $T \geq 1$, and x_o is a n-vector valued function on $[0,1)$, is called a solution of the difference equation (11) if, for some $T \geq 1$,

a) $x(t) = x_o(t - t_o)$ for $t_o \leq t < t_o + 1$

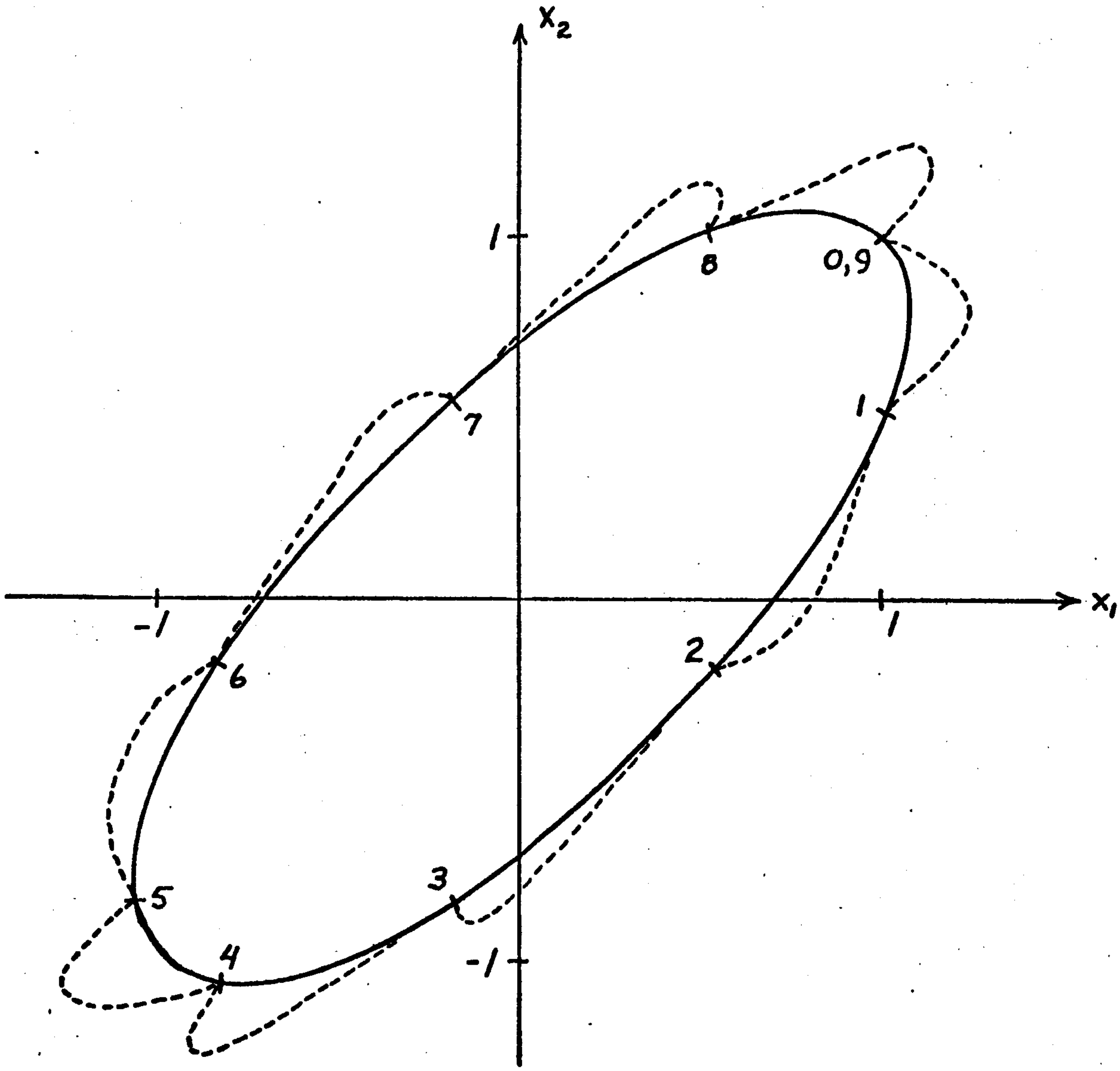


Figure 7

Two Solutions of $x(k+1) = Ax(k)$

$$\text{b) } x(t+1) = f(t,x(t)) \quad \text{for} \quad t_o \leq t < t_o + T-1 \ .$$

The difference between this definition of a solution and the usual definitions is that the initial value x_o is now a function instead of a vector. The function x_o defines the solution on the unit interval $[t_o, t_o+1)$ and the difference equation (11) is used to extend this definition to the right.

Notice that the solution is defined for all real t in the interval $[t_o, t_o+T)$ instead of restricting $t-t_o$ to be an integer. Notice also that the solution space, which is finite dimensional when the solution is defined only on the integers, is now infinite dimensional. An initial function must be defined instead of an initial point.

For any real number t, the integer part of t, $[t]$, is the largest integer $n \leq t$. The fractional part of t, $\langle t \rangle$, is given by the formula $\langle t \rangle = t-[t]$. Thus, as examples, $[3.14] = 3$, $\langle 3.14 \rangle = .14$, $[-1.62] = -2$, $\langle -1.62 \rangle = .38$, $[0] = 0$, and $\langle 0 \rangle = 0$. Let R be the real line, R^+ be the positive real half-line, and E^n be an n-dimensional normed vector space. For each vector $x \in E^n$, denote the norm of x by $|x|$. The most common norm used here will be the Euclidean norm $|x| = (x^*x)^{\frac{1}{2}}$ (x^* denotes the complex conjugate transpose of x) but discussion will not necessarily be limited to this norm.

For any function x mapping R (or R^+) into E^n, let x_t be the function defined by $x_t(\theta) = x(t+\theta)$ for $0 \leq \theta < 1$. Let X_1 be a Banach space of n-vector valued functions defined on $[0,1)$. For each $x_t \in X_1$, denote the norm of x_t by $\|x_t\|$. The usual space for X_1 will be $L_2[0,1)$, but discussion will not necessarily be limited to this space.

Let E be the operator defined by $Ex(t) = x(t+1)$. Then the difference equation (11) can be written in the somewhat shorter notation of

(11'). The arguments on x and Ex are understood to be t.

$$Ex = f(t,x) \tag{11'}$$

THEOREM 1 On existence and uniqueness of solutions. Suppose t_o and an $x_{t_o} \in X_1$ are given. If there are open, possibly unbounded, sets $G(t) \subset E^n$ for all $t \geq t_o$ such that $f(t,x) \in G(t+1)$ for all $x \in G(t)$, all $t \geq t_o$ and $x_{t_o}(\theta) \in G(t_o+\theta)$ for $0 \leq \theta < 1$, then the solution $x(t) = x(t;t_o,x_{t_o})$ exists for all $t \geq t_o$. If $f(t,x)$ is single valued, the solution is unique.

PROOF: The proof is by induction. The solution exists and is unique for $t_o \leq t < t_o + 1$ (defined by the initial function x_{t_o}). For each integer $N > 0$, assume that the solution $x(t)$ exists, is unique and is in $G(t)$ for $t_o \leq t < t_o + N$. Then the difference equation (11) gives the solution on $t_o \leq t < t_o + N+1$. By assumptions, $x(t+1) = f(t,x(t)) \in G(t+1)$ since $x(t) \in G(t)$ for $t_o + N-1 \leq t < t_o + N$ and $x(t+1)$ is defined and unique for $t_o + N-1 \leq t < t_o + N$. By induction, the solution exists and is unique for all $t \geq t_o$.

It should be noted here that existence comes from the assumption that $f(t,x) \in G(t+1)$ for all $x \in G(t)$ and uniqueness comes from the (trivial?) assumption that $f(t,x)$ is single valued for all $x \in G(t)$. The sets $G(t)$ are open and possibly unbounded. The most common sets used in this work are $G(t) = E^n$ for all $t \geq t_o$, but discussion is not necessarily limited to these sets. The fact that the sets are open play no vital role in this theorem. Note also that the theorem holds for each x_{t_o} which satisfies the conditions and is not limited to a specific given x_{t_o}.

For the function $f(t,x)$, let $f_t(y)$, where y is any function from $[0,1)$ into E^n, be the function defined by

$$f_t(y)(\theta) = f(t+\theta, y(\theta)) \ , \quad 0 \leqq \theta < 1 \ .$$

The following three theorems each show that $x_t \in X_1$ under proper conditions on x_{t_o} and f for three different choices of the space X_1. The statements of the theorems and their proofs are almost identical. The difference in these proofs, and the difficulty for a general Banach function space, is the proof that, if $x_{t_o} \in X_1$ and $x_{t_o+1} \in X_1$, then $x_t \in X_1$ for $t_o \leqq t \leqq t_o+1$. In Theorem 2a, for example, this is not true unless x_{t_o} satisfies a certain boundary condition.

<u>THEOREM 2a</u> If, in addition to the assumptions of Theorem 1, the space X_1 is the space of all continuous functions $C[0,1)$ with the uniform norm, $f_t(y) \in C[0,1)$ for each $y \in C[0,1)$ such that $y(\theta) \in G(t+\theta)$, $0 \leqq \theta < 1$, and each $t \geqq t_o$, and the initial function x_{t_o} is in $C[0,1)$ and satisfies the boundary condition

$$\lim_{\theta \to 1} x_{t_o}(\theta) = f(t_o, x_{t_o}(0))$$

then the solution $x(t) = x(t;t_o,x_{t_o})$ satisfies $x_t \in C[0,1)$ for each $t \geqq t_o$.

PROOF: The boundary condition on x_{t_o} insures that the solution $x(t)$ is continuous at $t = t_o+1$. Since $x_{t_o+1} = f_{t_o}(x_{t_o})$ is in $C[0,1)$, the solution is continuous for $t_o \leqq t < t_o+2$, i.e., $x_t \in C[0,1)$ for $t_o \leqq t < t_o+1$.

If $x_t \in C[0,1)$ for $t_o \leq t < t_o+N$, then, using the fact that $x_{t+1} = f_t(x_t)$ is in $C[0,1)$ whenever $x_t \in C[0,1)$, we get that $x_t \in C[0,1)$ for $t_o \leq t < t_o+N+1$, and the theorem is proven by induction.

<u>THEOREM 2b</u> If, in addition to the assumptions of Theorem 1, the space X_1 is the space of all p-th power integrable functions $L_p[0,1)$ with the usual norm, some $p \geq 1$, $f_t(y) \in L_p[0,1)$ for each $y \in L_p[0,1)$ such that $y(\theta) \in G(t+\theta)$, $0 \leq \theta < 1$, and each $t \geq t_o$, and $x_{t_o} \in L_p[0,1)$, then the solution $x(t) = x(t;t_o,x_{t_o})$ satisfies $x_t \in L_p[0,1)$ for each $t \geq t_o$.

PROOF: We have $x_{t_o} \in L_p[0,1)$ and $x_{t_o+1} = f_{t_o}(x_{t_o}) \in L_p[0,1)$. Then, for any t, $t_o \leq t < t_o+1$,

$$\|x_t\|^p = \int_t^{t+1} |x(s)|^p ds = \int_t^{t_o+1} |x(s)|^p ds + \int_{t_o+1}^{t+1} |x(s)|^p ds$$

$$\leq \int_{t_o}^{t_o+1} |x(s)|^p ds + \int_{t_o+1}^{t_o+2} |x(s)|^p ds = \|x_{t_o}\|^p + \|x_{t_o+1}\|^p$$

and hence $x_t \in L_p[0,1)$ for $t_o \leq t < t_o+1$.

If $x_t \in L_p[0,1)$ for $t_o \leq t < t_o+N$, then, using the fact that $x_{t+1} = f_t(x_t)$ is in $L_p[0,1)$ whenever $x_t \in L_p[0,1)$, we get that $x_t \in L_p[0,1)$ for $t_o \leq t < t_o+N+1$ and the theorem is proven by induction.

Krasnosel'skii [8] gives a necessary and sufficient condition for $f_t(y)$ to be in $L_p[0,1)$ for all y in $L_p[0,1)$ such that $y(\theta) \in G(t+\theta)$, $0 \leq \theta < 1$. This condition is given here without proof, $f_t(y)$ is in $L_p[0,1)$ if and only if there exists a non-negative valued function $a(\theta)$, $0 \leq \theta < 1$, with $\int_0^1 a(\theta)^p d\theta < \infty$, and a non-negative constant b such that

$$|f(t+\theta,x)| \leq a(\theta) + b|x|$$

for all $x \in G(t+\theta)$, $0 \leq \theta < 1$. This condition is interesting for the case where $G(t+\theta)$ is unbounded because, while any linear function of x satisfies it, there are large classes of non-linear functions of x which do not. For the case where $G(t+\theta)$ is bounded for all θ, $0 \leq \theta < 1$, this condition reduces to the condition that $|f(t+\theta,x)|$ be bounded for all $x \in G(t+\theta)$, $0 \leq \theta < 1$.

THEOREM 2c If, in addition to the assumptions of Theorem 1, the space X_1 is the space of all functions of bounded variation $BV[0,1)$ with the usual norm, $f_t(y) \in BV[0,1)$ for each y in $BV[0,1)$ such that $y(\theta) \in G(t+\theta)$, $0 \leq \theta < 1$, each $t \geq t_o$, and $x_{t_o} \in BV[0,1)$, then the solution $x(t) = x(t;t_o,x_{t_o})$ satisfies $x_t \in BV[0,1)$ for each $t \geq t_o$.

PROOF: Let $V(t_1,t_2)$ be the variation of the solution x on the interval $[t_1,t_2)$. Then, if $t_3 \geq t_2$,

$$V(t_1,t_3) = V(t_1,t_2) + V(t_2,t_3) .$$

Thus, for $t_o \leq t < t_o+1$,

$$\begin{aligned} \|x_t\| &= |x(t)| + V(t,t+1) \\ &= |x(t)| + V(t,t_o+1) + V(t_o+1,t+1) \\ &\leq |x(t)| + |x(t_o)| + V(t_o,t_o+1) + |x(t_o+1)| + V(t_o+1,t_o+2) \\ &\leq |x(t)| + \|x_{t_o}\| + \|x_{t_o+1}\| \end{aligned}$$

and $x_t \in BV[0,1)$ for $t_o \leq t < t_o+1$ since x_{t_o} and x_{t_o+1} are in $BV[0,1)$.

If $x_t \in BV[0,1)$ for $t_o \leq t < t_o+N$, then, using the fact that $x_{t+1} = f_t(x_t)$ is in $BV[0,1)$ whenever $x_t \in BV[0,1)$, we get that $x_t \in BV[0,1)$ for $t_o \leq t < t_o+N+1$ and the theorem is proven by induction.

The solution $x(t)$, while written as a function of t, t_o, x_{t_o}, really depends only on $t, t_o + \langle t-t_o \rangle$ and $x_{t_o}(\langle t-t_o \rangle)$. Since $t = t_o + \langle t-t_o \rangle + n$ where $n = [t-t_o]$ is an integer, the value of the solution at $t_o + \langle t-t_o \rangle$ determines the solution at $t_o + \langle t-t_o \rangle + k$ for $k = 0,1,2, \ldots, n$, and thus determines the solution at t.

<u>THEOREM 3</u> If $f(t,x)$ is a continuous function of x for all $x \in G(t)$ and all $t \geq t_o$ and the assumptions of Theorem 1 hold, then the solution $x(t) = x(t;t_o,x_{t_o})$ is a continuous function of $x_{t_o}(\langle t-t_o \rangle)$.

PROOF: If $t_o \leq t < t_o+1$, then $x(t) = x_{t_o}(\langle t-t_o \rangle)$ and this is obviously continuous in $x_{t_o}(\langle t-t_o \rangle)$.

Now $x(t+1) = f(t,x(t))$, a continuous function of a continuous function, is also continuous in $x_{t_o}(\langle t-t_o \rangle)$. Thus, the theorem is proven by induction for all $t \geq t_o$.

<u>THEOREM 4</u> Under the conditions of Theorem 1 and 2b, the solution x_t considered as a function of x_{t_o} is a continuous function of the initial data.

The proof of this theorem is very difficult and it is given on pages 20-26 of Krasnosel'skii [8].

In looking for periodic solutions of a difference equation, it would be nice to be able to use only continuous functions, i.e., functions which are continuous in t. However, the following example illustrates that this is not always possible.

For $x(t)$ a real number, consider the difference equation (12).

$$Ex = -x^{1/3} \tag{12}$$

One solution of (12) is the function $x(t)$ given by (12')

$$x(t) = \begin{cases} 1 & 0 \leqq t < 1 \\ -1 & 1 \leqq t < 2 \end{cases} \tag{12'}$$

and $x(t+2) = x(t)$. This solution is 2-periodic so there exists a non-trivial 2-periodic solution of (12). However, this solution is not continuous in t.

Assume that there is a non-trivial 2-periodic solution of (12) which is continuous in t. Then there is a t_1 such that $x(t_1) > 0$ and $x(t_1+1) < 0$. Since x is continuous there is a t_2 such that $t_1 < t_2 < t_1+1, x(t_2) = 0$, and $0 < x(t_2-\epsilon) < 1$ for some $\epsilon > 0$. Then $x(t_2-\epsilon+2) = -x(t_2-\epsilon+1)^{1/3} = (x(t_2-\epsilon))^{1/9} > x(t_2-\epsilon)$. This is a contradiction since x is 2-periodic, i.e., $x(t_2-\epsilon+2) = x(t_2-\epsilon)$. Thus, any non-trivial 2-periodic solution of (12) must be discontinuous in t.

LINEAR DIFFERENCE EQUATIONS

A linear difference equation is an equation of the form (13) where $A(t)$ is an n by n matrix (real or complex).

$$Ex = A(t)x \tag{13}$$

This equation is called linear because solutions of (13) obey the superposition rule.

LEMMA 1 Superposition rule. If $x_1(t)$ and $x_2(t)$ are any two solutions of (13), then $x(t) = \alpha_1(t)x_1(t) + \alpha_2(t)x_2(t)$ is also a solution for any scalars α_1, α_2 which are 1-periodic.

PROOF:
$$\begin{aligned} A(t)x(t) &= \alpha_1(t)A(t)x_1(t) + \alpha_2(t)A(t)x_2(t) \\ &= \alpha_1(t)x_1(t+1) + \alpha_2(t)x_2(t+1) \\ &= \alpha_1(t+1)x_1(t+1) + \alpha_2(t+1)x_2(t+1) = x(t+1). \end{aligned}$$

In linear difference equations 1-periodic scalars play a role very analogous to that played by constants in linear differential equations. Thus, the set of functions $x_1(t), x_2(t), \ldots, x_N(t)$ are said to be linearly dependent at t over the integers if there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_N$, not all zero, such that

$$\sum_{k=1}^{N} \alpha_k x_k(t+m) = 0$$

identically for all integers m. Clearly, these scalars $\alpha_1, \alpha_2, \ldots, \alpha_N$ are functions of t and can be chosen to be 1-periodic in t. If the functions $x_1(t), x_2(t), \ldots, x_N(t)$ are linearly dependent at t over the integers for all t, then they are said to be linearly dependent over the integers or merely linearly dependent. Repeating, the functions $x_1(t),x_2(t),\dots,x_N(t)$ are linearly dependent (over the integers) if there exists 1-periodic scalar functions $\alpha_1(t),\alpha_2(t),\dots,\alpha_N(t)$ such that

$$\sum_{k=1}^{N}|\alpha_k(t)|^2 > 0 \quad \text{for all } t$$

and

$$\sum_{k=1}^{N}\alpha_k(t)x_k(t) = 0 \quad \text{for all } t .$$

The functions $x_1(t),x_2(t),\dots,x_N(t)$ are linearly dependent if the only 1-periodic scalars satisfying this equation are all identically zero. The fact that the scalars $\alpha_1(t),\alpha_2(t),\dots,\alpha_k(t)$ must be 1-periodic is important. If each $x_k(t)$ is in E^n and $N \geq n+1$, then there always are scalars which satisfy this equation but they may not be 1-periodic.

The following Lemmas are not new with this paper but are included here for completeness.

LEMMA 2 Linear dependence of solutions -- Any $n+1$ solutions $x_1(t),x_2(t),\dots,x_{n+1}(t)$ defined for $t \geq t_o$ of (13) are linearly dependent on $[t_o,\infty)$.

PROOF: For each t, $t_o \leq t < t_o+1$, there exist scalars $\alpha_1(t),\alpha_2(t),\dots,\alpha_{n+1}(t)$ such that

$$\sum_{k=1}^{n+1}|\alpha_k(t)|^2 > 0$$

and

$$\sum_{k=1}^{n+1} \alpha_k(t)x_k(t) = 0$$

identically in t on this interval $(t_o \leq t < t_o+1)$. This is because, for each t, this equation is the relationship depicting the linear dependence of $n+1$ n-vectors. Operating on this equation with $A(t)$, we get

$$0 = A(t)\sum_{k=1}^{n+1}\alpha_k(t)x_k(t) = \sum_{k=1}^{n+1}\alpha_k(t)A(t)x_k(t)$$
$$= \sum_{k=1}^{n+1}a_k(t)x_k(t+1) .$$

If we extend $\alpha_k(t)$ to the right by letting $\alpha_k(t+1) = \alpha_k(t)$, we see that we have a set of 1-periodic scalars which satisfy the condition for linear dependence.

<u>LEMMA 3</u> If there are n linearly independent solutions of (13) $x_1(t), x_2(t), \ldots, x_n(t)$ defined for $t \geq t_o$, then every solution $x(t)$ for $t \geq t_o$ can be uniquely expressed as the sum

$$x(t) = \sum_{k=1}^{n} \alpha_k(t)x_k(t) \tag{14}$$

where each $\alpha_k(t)$ is a 1-periodic scalar.

PROOF: For each t in the interval $t_o \leq t < t_o+1$, there exists scalars $\alpha_1(t), \alpha_2(t), \ldots, \alpha_n(t)$ which satisfy (14). These scalars are uniquely defined. If we extend these scalars to the right by using the relation $\alpha_k(t+1) = \alpha_k(t)$, then the right-hand side of (14) represents a solution to (13) which agrees with $x(t)$ on $t_o \leq t < t_o+1$. Since the solution with this initial function is unique, we must have equation (14) holding for all $t \geq t_o$.

If we let $X(t)$ be the matrix formed by the linearly independent solutions $x_1(t), x_2(t), \ldots, x_n(t)$ ($x_k(t)$ is the k-th column of $X(t)$) and $y(t)$ be the vector with elements $\alpha_k(t)$ ($\alpha_k(t)$ is the k-th element of $y(t)$), then $y(t+1) = y(t)$ and equation (14) can be rewritten as (14').

$$x(t) = X(t)y(t) \tag{14'}$$

The matrix $X(t)$ is called a <u>fundamental matrix solution</u> of (13) and satisfies the matrix equation $X(t+1) = A(t)X(t)$. If $X(t) = I$, the unit matrix, for $t_o \leq t < t_o+1$, then $X(t)$ is the <u>principal matrix solution</u> of (13) and is written $X(t,t_o)$. We note that, with the principal matrix solution, the function $y(t)$ in (14') becomes the initial function $x_{t_o}(\langle t-t_o \rangle)$ and hence the solution $x(t)$ can be written as (14").

$$x(t) = X(t,t_o)x_{t_o}(\langle t-t_o \rangle) \tag{14"}$$

<u>LEMMA 4</u> Existence of n linearly independent solutions -- If $\det(A(t)) \neq 0$ for all $t \geq t_o$, then there exists n linearly independent solutions of (13) for $t \geq t_o$.

PROOF: Let $X(t)$ be a matrix solution of (13) such that $\det(X(t)) \neq 0$ for $t_o \leq t < t_o+1$. Then, since $\det(X(t+1)) = \det(A(t)X(t)) = \det(A(t))\det(X(t))$, we get by induction that $\det(X(t)) \neq 0$ for all $t \geq t_o$. The columns of $X(t)$ thus represent n linearly independent solutions of (13).

In everything that follows, I assume that $\det(A(t)) \neq 0$ for all $t \geq t_o$ and hence that Lemma 4 holds.

Since $X(t)$ is non-singular for all $t \geq t_o$, $X^{-1}(t)$ must exist for all t. From the identity $X^{-1}(t)X(t) = I$, we get that

$$X^{-1}(t+1)X(t+1) = X^{-1}(t+1)A(t)X(t) = X^{-1}(t)A(t)$$

and, since $X(t)$ is non-singular, $X^{-1}(t)$ satisfies the difference equation (15).

$$X^{-1}(t) = X^{-1}(t+1)A(t) \tag{15}$$

For $z(t)$ a row vector (a 1 by n matirix), the difference equation (15') is called the <u>adjoint</u> to (13).

$$z(t) = z(t+1)A(t) \quad \text{or} \quad z(t+1) = z(t)A^{-1}(t) \tag{15'}$$

A fundamental matrix solution of (15') is the inverse of some fundamental matrix solution of (13). In particular, the principal matrix solution of (15') is the inverse of the principal matrix solution of (13).

For each vector norm $|x|$, define the norm of a matrix M by $|M|$.

$$|M| = \inf \{b : |Mx| \leq b|x| \quad \text{for all} \quad x\}$$

This matrix norm satisfies all the usual properties of a norm [6].

(i) $|M| \geq 0$ for all M and $|M| = 0$ only when $M = 0$.

(ii) $|M_1 + M_2| \leq |M_1| + |M_2|$

(iii) $|\alpha M| = |\alpha|\ |M|$ for any scalar α

(iv) $|M_1M_2| \leq |M_1|\ |M_2|$

(v) $|Mx| \leq |M|\ |x|$ for any vector x

LEMMA 5 Boundedness of the Principal Matrix Solution -- Let $a(t)$ be such that $|A(t)| \leq a(t)$ for all $t \geq t_o$ and $\xi(t,t_o)$ be defined by

$$\xi(t,t_o) = 1, \qquad t_o \leq t < t_o+1$$

$$\xi(t,t_o) = \prod_{j=1}^{[t-t_o]} a(t-j) \quad \text{for} \quad t \geq t_o+1.$$

Then $|X(t,t_o)| \leq \xi(t,t_o)$ for all $t \geq t_o$.

PROOF: The proof will proceed by induction. Since $X(t,t_o) = I$ for $t_o \leq t < t_o+1$, we get that $|X(t,t_o)| = 1 = \xi(t,t_o)$ on this interval. Assume that the inequality $|X(t,t_o)| \leq \xi(t,t_o)$ holds for $t_o \leq t < s$, some $s \geq t_o+1$. Then, since $X(t,t_o)$ satisfies the difference equation $X(t+1,t_o) = A(t)X(t,t_o)$, we get

$$|X(t+1,t_o)| \leq |A(t)||X(t,t_o)| \leq a(t)\xi(t,t_o) = \xi(t+1,t_o)$$

for all t, $t_o \leq t < s$. Thus, if the inequality is true for $t_o \leq t < s$, then it is true for $t_o \leq t < s+1$ and, by induction, is true for all $t \geq t_o$.

COROLLARY If $a(t) \leq a$ for all $t \geq t_o$, then $\xi(t,t_o) \leq a^{[t-t_o]} \leq k(a)a^{t-t_o}$ where $k(a) = 1$ if $a \geq 1$ and $k(a) = a^{-1}$ if $a < 1$.

PROOF:

$$\xi(t,t_o) = \prod_{j=1}^{[t-t_o]} a(t-j) \leq \prod_{j=1}^{[t-t_o]} a = a^{[t-t_o]} .$$

Since $[t-t_o] = t-t_o - \langle t-t_o \rangle$,

$$\xi(t,t_o) \leqq a^{[t-t_o]} = a^{t-t_o-\langle t-t_o\rangle} = (a^{-\langle t-t_o\rangle})a^{t-t_o} .$$

If $a \geqq 1$, $a^{-\langle t-t_o\rangle} \leqq 1 = k(a)$. If $a < 1$, $a^{-\langle t-t_o\rangle} \leqq a^{-1} = k(a)$ and the corollary is proven.

<u>LEMMA 5'</u> On Boundedness of the inverse of the principal matrix solution -- Let $a_1(t)$ be such that $|A^{-1}(t)| \leqq a_1(t)$ for all $t \geqq t_o$ and $\xi_1(t,t_o)$ be defined by

$$\xi_1(t,t_o) = 1 \quad \text{for} \quad t_o \leqq t < t_o+1$$

$$\xi_1(t,t_o) = \prod_{j=1}^{[t-t_o]} a_1(t-j) \quad \text{for} \quad t \geqq t_o+1$$

then $|X^{-1}(t,t_o)| \leqq \xi_1(t,t_o)$ for all $t \geqq t_o$.

<u>COROLLARY</u> If $a_1(t) \leqq a_1$ for all $t \geqq t_o$, then

$$\xi_1(t,t_o) \leqq a_1^{[t-t_o]} \leqq k(a_1)a_1^{t-t_o}$$

where $k(a_1) = 1$ if $a_1 \geqq 1$ and $k(a_1) = a_1^{-1}$ if $a_1 < 1$.

The proofs are almost identical to the proofs for Lemma 5 and its corollary.

Notice that, under the conditions of the corollaries, the principal matrix solution (and hence any solution) and its inverse are both bounded above by an exponential. If, in the corollary to Lemma 5, $a < 1$, then all solutions approach zero exponentially as $t \to \infty$.

LINEAR AUTONOMOUS DIFFERENCE EQUATIONS

A linear autonomous difference equation is of the form (13') where A is some constant n by n matrix (real or complex).

$$Ex = Ax \tag{13'}$$

In order to completely characterize the solutions of (13) we will need the following notation.

For any matrix B and some integer $k \geq 0$, Null $(B)^k$ is the set of all vectors x such that $B^k x = 0$. Null $(B)^k$ forms a linear subspace of X. The distinct complex numbers $\lambda_1, \lambda_2, \ldots, \lambda_r$ such that $\mathrm{Null}(A-\lambda_i I)^1$ contains non-zero elements are called the eigenvalues of A. With each eigenvalue λ_i of A there is associated an integer $m_i \geq 1$ which is the smallest integer such that $\mathrm{Null}(A-\lambda_i I)^{m_i} = \mathrm{Null}(A-\lambda_i I)^{m_i+1}$. The set of all non-zero vectors x in $\mathrm{Null}(A-\lambda_i I)^{m_i}$ are called the generalized eigenvectors of A corresponding to the eigenvalue λ_i. It can be proven [4, Theorem 2 on page 113-114] that any n-vector x can be written uniquely as the sum

$$x = \sum_{k=1}^{r} x_k$$

where each $x_k \in \mathrm{Null}(A-\lambda_k I)^{m_k}$.

For each integer n, the factorial polynomial of degree n, written $t^{(n)}$, is defined by

$$\begin{aligned} t^{(n)} &= (t(t+1)\ldots(t(n+1)))^{-1} \quad \text{if} \quad n < 0 \\ t^{(0)} &= 1 \\ t^{(n)} &= t(t-1)\ldots(t-(n-1)) \quad \text{if} \quad n > 0 . \end{aligned}$$

The factorial polynomial satisfies the very useful equality [5]

$$(t+1)^{(n)} = t^{(n)} + nt^{(n-1)}.$$

Any complex number $\lambda \neq 0$ can be written as $\lambda = re^{i\omega}$ for some real $r > 0$ and real ω. For real t, we can define the exponential function λ^t to be $\lambda^t = r^t e^{i\omega t}$. However, we can also write $\lambda = re^{i(\omega+2\pi k)t}$, and we have a different definition for λ^t for each integer k. However, noting that $e^{i2\pi kt}$ is 1-periodic for every integer k, we can chose ω such that $-\pi < \omega \leq \pi$ and absorb all the other definitions into an arbitrary 1-periodic coefficient. With the knowledge that there are many possible definitions of λ^t but that their differences can be absorbed by a 1-periodic coefficient, we will use the definition that $\lambda^t = r^t e^{i\omega t}$ where $-\pi < \omega \leq \pi$ in all of the following.

<u>LEMMA 6</u> Linear independence of the factorial polynomials -- The factorial polynomials $t^{(k)}$, $k = 0,1,2,..$, are linearly independent over $[t_o, t_o+s)$ where s is any number, $s > 1$.

PROOF: The proof will proceed by induction. $t^{(0)} = 1$ by definition, and hence $t^{(0)}$ by itself is linearly independent. Assume now that $t^{(0)}, t^{(1)}, \ldots, t^{(n-1)}$ are linearly independent but that $t^{(0)}, t^{(1)}, \ldots, t^{(n-1)}, t^{(n)}$ are linearly dependent for some integer $n \geq 1$. Then we can write

$$t^{(n)} = \sum_{k=0}^{n-1} \alpha_k(t) t^{(k)}$$

where each $\alpha_k(t)$ is 1-periodic. Then

$$(t+1)^{(n)} = t^{(n)} + nt^{(n-1)}$$

$$= \sum_{k=0}^{n-1} \alpha_k(t+1)(t+1)^{(k)}$$

$$= \sum_{k=0}^{n-1} \alpha_k(t)(t^{(k)} + kt^{(k-1)})$$

$$= \sum_{k=0}^{n-1} \alpha_k(t)t^{(k)} + \sum_{k=0}^{n-1} \alpha_k(t)_{kt}{}^{(k-1)}$$

$$= t^{(n)} + \sum_{k=0}^{n-1} \alpha_k(t)kt^{(k-1)} .$$

Thus, we get that

$$nt^{(n-1)} = \sum_{k=0}^{n-1} \alpha_k(t)kt^{(k-1)} .$$

This is possible only if $\alpha_k(t) = 0$, $k = 0,1,2,..,n-2$, and $\alpha_{n-1}(t)(n-1) = n$, i.e.,

$$\alpha_{n-1}(t) = \frac{n}{n-1} .$$

Then

$$t^{(n)} = \frac{n}{n-1} t^{(n-1)} .$$

From the definition of the factorial polynomials, we get

$$t^{(n)} = (t-(n-1))t^{(n-1)} .$$

Hence $\alpha_{n-1}(t)$ as above is not possible and we have a contradiction. Thus, the set $t^{(0)},t^{(1)},\ldots,t^{(n)}$ is linearly independent whenever $t^{(0)},t^{(1)},\ldots,$ $t^{(n-1)}$ is. The lemma is proven by induction.

LEMMA 6' Linear independence of the exponential functions -- If $\lambda_1, \lambda_2, \ldots, \lambda_r$ are r distinct scalars (complex numbers), none of which are zero, then the exponential functions $\lambda_1^t, \lambda_2^t, \ldots, \lambda_r^t$ are linearly independent.

PROOF: Let $\alpha_1(t), \alpha_2(t), \ldots, \alpha_r(t)$ be 1-periodic functions such that

$$\sum_{k=1}^{r} \alpha_k(t)\lambda_k^t = 0 \quad \text{for all } t.$$

Then, substituting $t+j$ for t, we get

$$\sum_{k=1}^{r} \alpha_k(t)_k^{t+j} = 0 \quad \text{for } j = 0,1,2,\ldots,r-1 .$$

The determinant of the coefficients on the α_k's is $W(t)$.

$$W(t) = \det \begin{pmatrix} \lambda_1^t & \lambda_2^t & \cdots & \lambda_r^t \\ \lambda_1^{t+1} & \lambda_2^{t+1} & \cdots & \lambda_r^{t+1} \\ \vdots & & & \\ \lambda_1^{t+r-1} & \lambda_2^{t+r-1} & \cdots & \lambda^{t+r-1} \end{pmatrix}$$

$$= \lambda_1^t \lambda_2^t \cdots \lambda_r^t \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_r \\ \vdots & & & \\ \lambda_1^{r-1} & \lambda_2^{r-1} & \cdots & \lambda_r^{r-1} \end{pmatrix}$$

Let there be constants $\beta_0, \beta_1, \ldots, \beta_{r-1}$ such that $p(\lambda) = \sum_{k=0}^{r-1} \beta_k \lambda^k$ satisfies $p(\lambda_1) = p(\lambda_2) = \ldots = p(\lambda_r) = 0$. This $p(\lambda)$ is a polynomial in λ of degree less than r which has at least r roots. The only possible polynomial with this property is the zero polynomial, i.e., $\beta_0 = \beta_1 = \ldots = \beta_{r-1} = 0$. This in turn says that the rows of the last determinant above are linearly independent and hence that determinant is non-zero. Since none of the λ_k's are zero, each factor λ_k^t in $W(t)$ is non-zero, and so $W(t) \neq 0$ for all t.

Since $W(t) \neq 0$ for all t, the only possible solution of the equations for the α_k's is $\alpha_1(t) = \alpha_2(t) = \ldots = \alpha_r(t) = 0$ for all t, and the exponential functions $\lambda_1^t, \lambda_2^t, \ldots, \alpha_r^t$ are linearly independent.

Suppose now that A has only one eigenvalue $\lambda \neq 0$ and that $m_1 = m$ (or that we restrict our attention to the space $\text{Null}(A-\lambda I)^m$). If x_0 is a non-trival solution of the equation $(A-\lambda I)x_0 = 0$, then one solution is of the form $x(t) = x_0\lambda^t$. However, this is only one of the m linearly independent solutions which exist. Even if there were several linearly independent x_0's which satisfy the equation $(A-\lambda I)x_0 = 0$, in general there will not be m of them. Since the factorial polynomials $t^{(0)}, t^{(1)}, t^{(2)}, \ldots,$ form a linearly independent set on $[t_0, \infty)$, try the solution $x(t)$

$$x(t) = \left(\sum_{k=0}^{\infty} t^{(k)} x_k\right)\lambda^t$$

where each x_k is a constant vector. Then

$$x(t+1)-Ax(t) = \left(\sum_{k=0}^{\infty}(t+1)^{(k)}x_k\right)\lambda^{t+1} - \left(\sum_{k=0} t^{(k)}Ax_k\right)\lambda^t$$

$$= \sum_{k=0}^{\infty}\left(\lambda\left(t^{(k)}+kt^{(k-1)}\right)x_k - t^{(k)}Ax_k\right)\lambda^t$$

$$= \sum_{k=0}^{\infty} t^{(k)}(\lambda x_k + (k+1)\lambda x_{k+1} - Ax_k)\lambda^t.$$

From the requirement that $x(t+1)-Ax(t) = 0$ for all t, we get that x_o is arbitrary in $\text{Null}(A-\lambda I)^m$ and

$$x_{k+1} = \frac{1}{(k+1)\lambda}(A-\lambda I)x_k$$

for $k = 0,1,2,\ldots$. By repeatedly applying this identity, we get that

$$x_k = \frac{1}{k!\,\lambda^k}(A-\lambda I)^k x_o.$$

Since $x_o \in \text{Null}(A-\lambda I)^m$, we see that $x_k = 0$ for all $k \geq m$. Thus, the sum for $x(t)$ has only a finite number of non-zero terms and there is no worry about convergence. The solution is

$$x(t) = x(t;\lambda,x_o) = \sum_{k=0}^{m-1} \frac{t^{(k)}}{k!}(A-\lambda I)^k x_o \qquad (16)$$

for any $x_o \in \text{Null}\ (A-\lambda I)^m$. Note that this solution is a polynomial in t times the exponential function λ^t and is a linear function of x_o.

In the more general case where A has the r distinct eigenvalues $\lambda_1,\lambda_2,\ldots,\lambda_r$, we can find n linearly independent solutions in the following manner. Pick n linearly independent vectors $x_1,x_2,\ldots,x_n$ such that each x_k is in one of the spaces $\text{Null}\ (A-\lambda_j I)^{m_j}$, $j = 1,2,\ldots,r$. This can always be done [4]. If $x_k \in \text{Null}\ (A-\lambda_j I)^{m_j}$, form the solution $x_k(t)$ in (17) using the solution in (16).

$$x_k(t) = x(t;\lambda_j, x_k) \quad \text{for} \quad k = 1,2,\ldots,n \tag{17}$$

LEMMA 7 The n solutions (17) are linearly independent.

Let $x(t)$ be the solution of (13') with the initial condition $x_0(t) = x(t)$ for $0 \leq t < 1$. Write

$$x_0(t) = \sum_{k=1}^{r} x_k(t) \qquad 0 \leq t < 1$$

where $x_k(t) \in \text{Null}\,(A-\lambda_k I)^{m_k}$ for each k. Then the solution $x(t)$ is of the form (18) using the solutions in (16).

$$x(t) = \sum_{k=0}^{r} x([t];\lambda_k, x_k(\langle t\rangle)) \tag{18}$$

Thus, any solution can be expressed as a linear combination of solutions of the form of (16), i.e., as a linear combination of exponential functions times factorial polynomials.

The behavior of each solution (16) depends on the eigenvalue λ_i under consideration. If $|\lambda_i| > 1$, then the solution (16) grows exponentially as $t \to \infty$. Equation (16) can be used as the definition of a solution for all t, and we see that the solution approaches zero exponentially as $t \to -\infty$. Let X^1 be the subspace of X corresponding to all the eigenvalues λ_i which satisfy $|\lambda_i| > 1$. Then any solution which starts in X^1 (i.e., $x_{t_0}(\theta) \in X^1$ for $0 \leq \theta < 1$) remains in X^1 and grows exponentially as $t \to \infty$. The only solution in X^1 which is bounded for all t in $(-\infty,\infty)$ is the trivial solution $x(t) = 0$.

If $|\lambda_i| < 1$, then the solution (16) approaches zero exponentially as $t \to \infty$ and grows exponentially as $t \to -\infty$. Let X^2 be the subspace of X corresponding to all the eigenvalues λ_i which satisfy $|\lambda_i| < 1$. Then any solution which starts in X^2 (i.e., $x_{t_0}(\theta) \in X^2$ for $0 \leq \theta < 1$) remains in X^2 and approaches zero exponentially as $t \to \infty$. The only solution in X^2 which is bounded for all t in $(-\infty, \infty)$ is the trivial solution $x(t) = 0$.

If $|\lambda_i| = 1$, then the behavior of the solution depends on the values of $(A-\lambda_i I)^k x_0$ in equation (16). If x_0 is such that $(A-\lambda_i I)x_0 = 0$, then the solution is bounded for all t and is, in fact, periodic. Suppose $\lambda_i = e^{i2\pi\omega}$ for some ω, $0 < \omega \leq 1$. Then the solution (16) is given by

$$x(t) = e^{i2\pi\omega t} x_0$$

which is $(1/\omega)$-periodic. But, since $e^{i2\pi kt}$, k an integer, is a 1-periodic function, $e^{i2\pi kt}x(t) = e^{i2\pi(\omega+k)t}x_0$ is also a solution of (13), and it is periodic with period $1/(\omega+k)$. Since each T-periodic function is also mT-periodic (m integer, $m \neq 0$), the set S of periods for solutions of (13) is $S_\omega = \{m/(\omega+k);\ m \neq 0, \text{ all } k\}$. This set S_ω is dense in the real line, in contrast to the case in differential equations where S_ω (the set of periods of periodic solutions of $\dot{x} = Ax$) consists of a countable set of distinct points.

If $|\lambda_i| = 1$ and x_0 is such that $(A-\lambda_i I)^m x_0 = 0$ but $(A-\lambda_i I)x_0 \neq 0$, then the solution (16) grows like a power of t as $t \to \infty$. The only bounded solution is the trivial solution $x(t) = 0$.

AFFINE DIFFERENCE EQUATIONS

An affine difference equation is an equation of the form of (19)

$$Ex = A(t)x + f(t) \tag{19}$$

where $A(t)$ is an n by n matrix and $f(t)$ is an n-vector. Associated with each affine difference equation (19) is a corresponding linear difference equation (13).

$$Ex = A(t)x \tag{13}$$

Let $X(t,t_o)$ be the principal matrix solution of (13) and make the change of variables (20).

$$x(t) = X(t,t_o)y(t) \tag{20}$$

Then

$$\begin{aligned} x(t+1) &= X(t+1,t_o)y(t+1) \\ &= A(t)X(t,t_o)y(t) + f(t) \\ &= X(t+1,t_o)y(t) + f(t) \end{aligned}$$

and so we get that $y(t)$ must satisfy the difference equation

$$y(t+1) = y(t) + X^{-1}(t+1,t_o)f(t) .$$

For $t_o \leq t < t_o+1$, we have $X(t,t_o) = I$ and so $y(t) = x_{t_o}(\langle t-t_o \rangle)$.

The solution for $y(t)$ is, for $t \geq t_o+1$,

$$y(t) = x_{t_o}(\langle t-t_o \rangle) + \sum_{j=1}^{[t-t_o]} X^{-1}(t+1-j,t_o)f(t-j)$$

and the solution $x(t)$ is given by (21).

$$x(t) = X(t,t_o)x_{t_o}(\langle t-t_o \rangle) + \sum_{j=1}^{[t-t_o]} X(t,t_o)X^{-1}(t+1-j,t_o)f(t-j) \qquad (21)$$

Equation (21) is called the _variation of constants formula_ for (19).

Another form for the variation of constants formula can be obtained by a repeated application of equation (19).

$$\begin{aligned} x(t) &= A(t-1)x(t-1) + f(t-1) \\ &= A(t-1)A(t-2)x(t-2) + A(t-1)f(t-2) + f(t-1) \\ &= \ldots \\ &= A(t-1)\ldots A(t-n)x(t-n) + f(t-1) + A(t-1)f(t-2) \\ &\quad + A(t-1)A(t-2)f(t-3)+\ldots+A(t-1)\ldots A(t-n+1)f(t-n) \; . \end{aligned}$$

If we let $n = [t-t_o]$, then $t-n = t_o+\langle t-t_o \rangle$ and

$$x(t) = X(t,t_o)x_{t_o}(\langle t-t_o \rangle) + \sum_{j=1}^{[t-t_o]} X(t,t_o+j)f(t_o+\langle t-t_o \rangle+j-1)$$

This equation is more informative of the nature of the coefficients on x_{t_o} and f in the variation of constants formula.

Let $x^o(t)$ be the solution of (19) with zero initial function. Then $x_{t_o}(\langle t-t_o \rangle) = 0$ for all t and, for $t \geq t_o+1$,

$$x^{o}(t) = x^{o}(t;f) = \sum_{j=1}^{[t-t_o]} X(t,t_o)X^{-1}(t+1-j,t_o)f(t-j)$$

We see that x^o is a linear function of the function f. Let $x^1(x) = x^1(t;x_{t_o})$ be the solution of the homogeneous equation (13) with initial value x_{t_o}. Then x^1 is a linear function of the function x_{t_o} and the solution $x(t)$ in (21) can be rewritten as in (21').

$$x(t) = x^1(t;x_{t_o}) + x^o(t;f) \qquad (21')$$

Suppose now that $f \in C[t_o,t]$ and $\|f\|_t = \max \{|f(s)|,\ t_o \leq s \leq t\}$. Then

$$|x^o(t;f)| = \left| \sum_{j=1}^{[t-t_o]} X(t,t_o+j)f(t_o+ t-t_o +j-1)\right|$$

$$\leq \sum_{j=1}^{[t-t_o]} |X(t,t_o+j)|\,|f(t_o+ t-t_o +j-1)|$$

$$\leq \sum_{j=1}^{[t-t_o]} \xi(t,t_o+j)\|f\|_t = m(t,t_o)\|f\|_t .$$

Hence, x^o is a continuous linear function of f. Thus there exists a matrix valued function $\beta(t,s)$ such that each element, considered as a function of s (holding t fixed), is of bounded variation on $[t_o,t]$ and such that

$$x^o(t;f) = \int_{t_o}^{t} d_s\beta(t,s)f(s)$$

See [7], for example.

BOUNDED SOLUTIONS -- NONCRITICAL CASE

Suppose that, in equation (19), where A is a constant matrix,

$$Ex = Ax + f(t) \tag{19}$$

the function $f(t)$ is defined for all t in $(-\infty,\infty)$ and is bounded for all t in $(-\infty,\infty)$. Under what conditions is there a bounded solution $x(t)$ of (19)?

Let $|x|$ be any vector norm for $x \in E^n$. A function $f: (-\infty,\infty) \to E^n$ is said to be bounded it there is some number $F \geq 0$ such that $|f(t)| \leq F$ for all t in $(-\infty,\infty)$. Let B be the set of all bounded functions and define addition and scalar multiplication on elements of B in the usual sense. For each $f \in B$, let

$$\|f\| = \sup \{|f(t)| : \text{all } t\} .$$

With this norm, B becomes a Banach space.

<u>LEMMA 8</u> Let $\{u_n\}$ be a sequence of bounded functions, $u_n \in B$, which converge to $u_o \in B$. If each u_n is (uniformly) continuous in t, then u_o is (uniformly) continuous in t. If each u_n is almost periodic uniformly in n, then u_o is almost periodic.

This lemma is a fairly standard one since convergence in B means pointwise convergence uniformly in t. For the continuous and uniformly continuous proof, see any advanced calculus book, for example [16]. For the almost periodic proof, see, for example, [17].

The first case handled is the non-critical case where none of the eigenvalues of A lie on the unit circle. The critical case is where one or more of the eigenvalues of A lie on the unit circle. For the non-critical case, we assume that A has the following form

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

where A_1 and A_2 are square matrices such that all the eigenvalues of A_1 lie inside the unit circle and all the eigenvalues of A_2 lie outside the unit circle. If the matrix A is not in this form, it can be arranged to this form by a transformation of co-ordinates. The assumption is that such a change of co-ordinates has already been performed. Let $x = \text{col}(x_1, x_2)$ and $f(t) = \text{col}(f_1(t), f_2(t))$. Then the difference equation (19) becomes the difference equation (22).

$$\begin{aligned} Ex_1 &= A_1 x_1 + f_1(t) \\ Ex_2 &= A_2 x_2 + f_2(t) \end{aligned} \tag{22}$$

Let $x_1(t)$ be a solution of (22). Then

$$\begin{aligned} x_1(t) &= A_1 x_1(t-1) + f(t-1) = A_1^2 x_1(t-2) + A_1 f_1(t-2) + f_1(t-1) \\ &= \dots = A_1^n x_1(t-n) + \Sigma_{j=1}^{n} A_1^{j-1} f_1(t-j) . \end{aligned}$$

Since all the eigenvalues of A_1 are inside the unit circle, the matrices A_1^{j-1} approach zero exponentially as $j \to \infty$. Since f is bounded, this means that the sum converges as $n \to \infty$. If x_1 is bounded, then $A_1^n x_1(t-n) \to 0$ as

$n \to \infty$ and

$$x_1(t) = \sum_{j=1}^{\infty} A_1^{j-1} f_1(t-j) \tag{23}$$

There are numbers $a_1 > 0$ and b_1, $0 \leq b_1 < 1$, such that

$$|A_1^n| \leq a_1 b_1^n \qquad \text{for all integers } n \geq 0$$

and so

$$|x_1(t)| \leq \sum_{j=1}^{\infty} |A_1^{j-1} f_1(t-j)| \leq \sum_{j=1}^{\infty} |A_1^{j-1}||f_1(t-j)|$$

$$\leq \sum_{j=1}^{\infty} a_1 b_1^{j-1} \|f_1\| = \frac{a_1}{1-b_1} \|f_1\| .$$

Hence the series for $x_1(t)$ converges absolutely. This inequality along with equation (23) shows that $x_1(t)$ does exist and that it is a continuous linear function of f.

If there were two bounded solutions to (22), say x_1^1 and x_1^2, then their difference $x_1^1 - x_1^2$ would be bounded and satisfy the homogeneous difference equation $Ex_1 = A_1 x_1$. The only bounded solution of this difference equation is the trivial solution $x_1 = 0$, hence $x_1^1 = x_1^2$ and the bounded solution is unique. For future purposes, we write this solution as (23')

$$x_1 = L_1(f_1) \tag{23'}$$

where $L_1\colon B \to B$ is the continuous linear operator defined by the right-hand side of equation (23).

In a similar manner, we get the unique bounded solution $x_2(t)$ as in equation (24).

$$x_2(t) = L_2(f_2)(t) = -\sum_{j=0}^{\infty} A_2^{-j-1} f_2(t+j)$$

Since all the eigenvalues of A_2 are outside the unit circle, there are numbers $a_2 > 0$ and b_2, $0 \leq b_2 < 1$, such that

$$|A_2^{-n}| \leq a_2 b_2^{+n} \quad \text{for all integers } n \geq 0$$

and so

$$|x_2(t)| \leq + \sum_{j=0}^{\infty} a_2 b_2^{j+1} \|f_2\| = \frac{a_2 b_2}{1-b_2} \|f_2\| .$$

The series converges absolutely and $L_2 \colon B \to B$ is a continuous linear operator.

We can write the unique bounded solution of (19) as in equation (25)

$$x(t) = L(f)(t) \tag{25}$$

where $L \colon B \to B$ is defined by

$$L(f) = \begin{pmatrix} L_1(f_1) \\ L_2(f_2) \end{pmatrix}$$

and is a continuous linear operator.

<u>THEOREM 5</u> If the matrix A does not have any eigenvalues on the unit circle, then there is a continuous linear operator $L : B \to B$ such that $x = Lf$ is the unique bounded solution of (19).

$$Ex = Ax + f(t) \tag{19}$$

This operator L has the following properties.

1) If f is uniformly continuous in t, then so is Lf.

2) If f is T-periodic in t, then so is Lf.

3) If f is almost periodic in t, then so is Lf.

PROOF: The existence and continuity of L and the uniqueness of the solution $x = Lf$ are proven in the preceeding discussion. All that remains is to prove that L possesses these three properties.

The proof of each depends on the following fact. Let h be some real number and $f \in B$. Define $f_1 \in B$ by $f_1(t) = f(t+h)$. Let $x = Lf$ and defined $x_1 \in B$ by $x_1(t) = x(t+h)$. Then $x_1 = Lf_1$ since, in equations (23) and (24), t can be replaced by t+h on both sides of the equations without disturbing the equalities.

Since $L : B \to B$ is a continuous linear operator, there is a number k, $0 \leq k < \infty$, such that $\|Lf\| \leq k\|f\|$ for all $f \in B$.

1) If f is uniformily continuous in t, then, for each $\epsilon > 0$, there is a $\delta = \delta(\epsilon) > 0$ such that $|f(t+h)-f(t)| \leq \epsilon$ for all h, $|h| \leq \delta$. For any such h, let $f_1(t) = f(t+h)$. Then $f_1 \in B$, $\|f_1\| = \|f\|$, and $\|f_1-f\| = \sup\{|f_1(t)-f(t)| : \text{all } t\} \leq \epsilon$. Let $x = Lf$ and $x_1(t) = x(t+h)$. Then $x_1 = Lf_1$ and

$$\|x_1-x\| = \|Lf_1-Lf\| \leq k\|f_1-f\| \leq k\epsilon .$$

Thus we have, for each $\epsilon > 0$, $|x(t+h)-x(t)| \leq \|x_1-x\| \leq k\epsilon$ for all $h, |h| \leq \delta(\epsilon)$. Hence $x = Lf$ is uniformly continuous in t.

2) If f is T-periodic, then $f(t+T) = f(t)$. Let $f_1(t) = f(t+T)$. Then $f_1 \epsilon B$, $\|f_1\| = \|f\|$, and $f_1 - f = 0$. Let $x = Lf$ and $x_1(t) = x(t+T)$. Then $x_1 = Lf_1$ and $x_1 - x = Lf_1 - Lf = L(f_1 - f) = 0$, implying that $x = Lf$ is T-periodic.

3) If f is almost periodic, then, for any $\epsilon > 0$, there is an $\ell(\epsilon) > 0$ such that, in any interval of length $\ell(\epsilon)$, there is a T such that

$$|f(t+T)-f(t)| \leq \epsilon$$

for all t. For this ϵ, $\ell(\epsilon)$, and T, let f_1 be the function defined by $f_1(t) = f(t+T)$. Then $\|f_1 - f\| \leq \epsilon$. Let $x = Lf$ and $x_1(t) = x(t+T) = Lf_1(t)$. Then

$$\|x_1 - x\| \leq k\|f_1 - f\| \leq k\epsilon$$

and we see that $x = Lf$ is almost periodic.

As an example of the application of this theorem, consider the following interesting problem. Given the following differential equation where $f \epsilon B$, find a solution $x \epsilon B$

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = -x_1 + g(t)$$

numerically by using the Crude Euler method. For this differential equation and a given spacing $h > 0$ in t, the Crude Euler method is given by the following difference equation.

$$Ex_1 = x_1 + h\dot{x}_1 = x_1 + hx_2$$
$$Ex_2 = x_2 + h\dot{x}_2 = -hx_1 + x_2 + hg(js) \qquad (26)$$

Here, the independent variable is changed to s where $t = hs$ in order to bring this difference equation into the form of equation (19). In fact, equation (19) becomes equation (26) if we let

$$A = \begin{pmatrix} 1 & h \\ -h & 1 \end{pmatrix} \qquad f(s) = \begin{pmatrix} 0 \\ hg(hs) \end{pmatrix}$$

The eigenvalues of A are $\lambda_1 = 1+ih$ and $\lambda_2 = 1-ih$. Neither λ_1 nor λ_2 are on the unit circle, so Theorem 5 holds. Indeed, since $|\lambda_1| > 1$ and $|\lambda_2| > 1$, equation (24) gives the bounded solution of (26) for any bounded function g. Since the differential equation has no bounded solution for $g(t) = \sin t$, we see that the behavior of the difference equation (26) in this case is considerably different from the differential equation.

Let x be the bounded solution of (26). Then the bound on x is given by

$$|x(t)| \leq \frac{(1+h^2)^{-\frac{1}{2}}}{1-(1+h^2)^{-\frac{1}{2}}} \|f\| \leq \frac{1+(1+h^2)^{\frac{1}{2}}}{h^2} \|f\| = \frac{1+(1+h^2)^{\frac{1}{2}}}{h} \|g\|$$

and we see that, while $h > 0$, if h is small then the bound on $|x(t)|$ may be large indeed; in the limit as $h \to 0$, the bound becomes infinite (if $\|g\|>0$). For $g(t) = \sin \omega t$, we get the solution

$$x_1(t) = \alpha_1 \cos \omega t + \alpha_2 \sin \omega t$$
$$x_2(t) = \alpha_3 \cos \omega t + \alpha_4 \sin \omega t$$

where

$$d = ((1-\cos \omega h)^2-(\sin \omega h)^2+h^2)^2+ 4(\sin \omega h)^2(1-\cos \omega h)^2$$

$$\alpha_1 = 2h^2\sin \omega h(1-\cos \omega h)/d$$

$$\alpha_2 = h^2((1-\cos \omega h)^2-(\sin \omega h)^2+h^2)/d$$

$$\alpha_3 = -h \sin \omega h((1-\cos \omega h)^2+ (\sin \omega h)^2-h^2)/d$$

$$\alpha_4 = -h(1-\cos \alpha h)((1-\cos \omega h)^2+ (\sin \omega h)^2+h^2)/d .$$

We notice that $|x_1(t)|^2 \leq \alpha_1^2 + \alpha_2^2 = h^4/d$, that $x_2(t)$ is approximately orthogonal to $x_1(t)$, and so the term h^4/d will give us a good approximation to the upper bound of $|x(t)|^2$. Expanding $\sin \omega h$ and $\cos \omega h$ in their Taylor series, we get

$$1-\cos \omega h = \tfrac{1}{2}\omega^2h^2+ \ldots$$

$$\sin \omega h = \omega h - \frac{1}{6}\omega^3h^3+ \ldots$$

and, after some simplification,

$$d = (1-\omega^2)^2h^4+ \frac{1}{6}\omega^4(7-\omega^2)h^6+ \ldots .$$

The bound we desire is

$$\frac{h^4}{d} = \frac{1}{(1-\omega^2)^2+ \frac{1}{6}\omega^4(7-\omega^2)h^2+ \ldots} .$$

For $\omega \neq 1$, this is close to the desired amplitude. The bounded solution of the differential equation satisfies

$$|x(t)|^2 \leq 1/(1-\omega^2)^2$$

and we see that, if we ignore terms of order h^2, this bound is h^4/d. However, when $\omega = 1$, there is no bounded solution of the differential equation and the bound on the bounded solution of the difference analog is

$$|x(t)|^2 \leq \frac{h^4}{d} = \frac{1}{h^2 + \ldots}$$

This shows that the bound for $\|x\|$ can be reached. While the behavior of the bounded solution of the difference analog (26) is considerably different than the behavior of the bounded solution of the differential equation near $\omega = 1$, this fact can be easily detected by computing the solution twice, once with $h = h_o$ and once with $h = \frac{1}{2}h_o$. If the two computed solutions vary greatly, then h_o is in a region where the behavior of the difference analog differs greatly from the behavior of the differential equation. The solution is to use either a much smaller h or a different numerical method.

THEOREM 6: If the matrix A is such that none of its eigenvalues lie on the unit circle, then there is some $\epsilon_1 > 0$ such that there is a bounded solution $x^*(t)$ of (27) for each ϵ, $|\epsilon| < \epsilon_1$,

$$Ex = Ax + \epsilon f(t,x,\epsilon) \tag{27}$$

where $f(t,x,\epsilon)$ is bounded for each fixed bounded x, each ϵ, $|\epsilon| < \epsilon_1$, and all t, and f is uniformly Lipschitz continuous with Lipschitz constant F for all t in $(-\infty,\infty)$ and all x

with $|x| \leq M$, some $M > 0$. This bounded solution $x^*(t)$ has the following properties.

1) If $f(t,x(t),\epsilon)$ is continuous in t for fixed ϵ, continuous x, then so is $x^*(t)$.

2) If $f(t,x(t),\epsilon)$ is T-periodic in t for fixed ϵ, T-periodic x, then so is $x^*(t)$.

3) If $f(t,x(t),\epsilon)$ is almost periodic for any almost periodic $x(t)$, then so is $x^*(t)$.

This result differs from that of Halanay [10] in that, in this paper, Halanay only considers the solution to be defined on the integers. Thus, when Halanay deals with T-periodicity, T must be an integer. His work with almost periodic functions is for functions almost periodic on the integers. Thus, this result is a generalization of Halanay's result in [10].

If all the eigenvalues of A are inside the unit circle, then this theorem is a specialization of the result of Halanay in [11]. At the same time, this result is more general than that of Halanay in [11] since this result allows A to have eigenvalues outside the unit circle.

PROOF: Let x be any bounded function with $\|x\| \leq M$. Let N be the positive number such that $|f(t,x,\epsilon)| \leq N$ for all t, all x with $|x| \leq M$, and $|\epsilon| < \epsilon_o$, some $\epsilon_o > 0$. For this x, define the operator L^* by

$$L^*(x) = \epsilon L(f(\cdot,x(\cdot),\epsilon))$$

where L is the linear operator of Theorem 5. Then

$$\begin{aligned}\|L^*(x)\| &= |\epsilon| \, \|L(f(\cdot,x(\cdot),\epsilon))\| \\ &\leq |\epsilon| K \|f(\cdot,x(\cdot),\epsilon)\| \leq |\epsilon| KN\end{aligned}$$

where K is the bound on L from Theorem 5. Thus, L* maps bounded functions into bounded functions. Furthermore, if $|\epsilon| KN \leq M$, we get that $\|L^*(x)\| \leq M$.

If x^* is a fixed point of L^*, i.e., $x^* = L^*(x^*)$, then x^* is the bounded solution of

$$Ex = Ax + \epsilon f(t, x^*(t), \epsilon)$$

and hence is the desired solution. We will show that such a fixed point exists by showing that L* is a contradiction mapping for $|\epsilon|$ small enough.

Let x_1, x_2 be bounded solutions with $\|x_1\| \leq M$ and $\|x_2\| \leq M$. Then

$$\|L^*(x_1) - L^*(x_2)\| = |\epsilon| \|L(f(\cdot, x_1(\cdot), \epsilon)) - L(f(\cdot, x_2(\cdot), \epsilon))\|$$

$$= |\epsilon| \|L(f(\cdot, x_1(\cdot), \epsilon) - f(\cdot, x_2(\cdot), \epsilon))\|$$

$$\leq |\epsilon| K \|f(\cdot, x_1(\cdot), \epsilon) - f(\cdot, x_2(\cdot), \epsilon)\|$$

$$\leq |\epsilon| KF \|x_1(\cdot), -x_2(\cdot)\| \rightarrow \|x_1(\cdot) - x_2(\cdot)\|$$

since f is uniformly Lipschitz continuous with Lipschitz constant F.

If ϵ is such that $|\epsilon| KN \leq M$ and $|\epsilon| KF < 1$, then L* is a contraction mapping on the set $\|x\| \leq M$, and L* has one and only one fixed point x^* with $\|x^*\| \leq M$. This is the case for all ϵ such that $|\epsilon| < \epsilon_1$ and

$$\epsilon_1 = \min \{\epsilon_0, \frac{M}{KN}, \frac{1}{KF}\}$$

1) If $f(t,x,\epsilon)$ is continuous in t for fixed x,ϵ, then $f(t,x(t),\epsilon)$ is continuous in t whenever x is continuous in t, thus $L^*(x)$ is continuous in t. In the iterations used to prove the existence and uniqueness of x^*, the initial function x_1 is chosen continuous in t, then each iterate x_K will be continuous in t and so the limit x^* will be continuous in t.

2) If $f(t,x,\epsilon)$ is T-periodic in t for fixed x,ϵ, then $f(t,x(t),\epsilon)$ is T-periodic in t for any T-periodic x and $L^*(x)$ is T-periodic will lead to the limiting function x^* being T-periodic.

3) If $f(t,x(t),\epsilon)$ is almost periodic in t for any almost periodic x, then $L^*(x)$ will be almost periodic and choosing an almost periodic initial function x_1 will lead to an almost periodic limiting function x^*.

AFFINE DIFFERENCE EQUATIONS -- SIMPLEST CRITICAL CASE

The problem considered here is to find a T-periodic solution of the difference equation (28).

$$Ex = x + f(t) \tag{28}$$

where $f(t)$ is T-periodic for some $T \geqq 1$ and f is integrable.

This case is called critical because the homogeneous equation has non-trivial T-periodic solutions. It is the simplest critical case because the matrix A is the simplest possible matrix, the unit matrix I.

We use the notation that $[t]$ is the largest integer less than or equation to t and $t = [t] + \langle t \rangle$. We call $[t]$ the <u>integer part of t</u> and $\langle t \rangle$ the <u>fractional part of t</u> since $[t]$ is an integer and $0 \leqq \langle t \rangle < 1$.

For each integer $n > 0$, the solution $x(t)$ is given by

$$x(t) = x(t-n) + \sum_{j=1}^{n} f(t-j) \ .$$

Letting $n = [t]$, we get (29) for $t \geqq 1$

$$x(t) = x(\langle t \rangle) + \sum_{j=1}^{[t]} f(t-j) \tag{29}$$

where $x(\langle t \rangle) = x_o(\langle t \rangle)$ is the initial function. The problem is to determine under what conditions of f will there be an x_o such that $x(t)$ given in (29) is T-periodic. From (29) we get, using $f(t+T) = f(t)$,

$$x(t+T)-x(t) = x(\langle t+T \rangle)-x(\langle t \rangle) + \sum_{j=[t]+1}^{[t+T]} f(t+T-j)$$

$$= x(\langle t+T\rangle)-x(\langle t\rangle) + \sum_{j=1}^{[t+T]-[t]} f(\langle t\rangle-j) \tag{30}$$

While equation (30) was derived using the solution defined only for $t \geq 1$, it can be easily shown that equation (30) hold for all t in $(-\infty,\infty)$. Thus, if we can find an initial function x_o such that the right hand side of (30) is zero identically in t, then that initial function will give a T-periodic solution.

The first, and simplest, case considered is when the period T is an integer ≥ 1. Then $\langle t+T\rangle = \langle t\rangle$ and $[t+T] = T+[t]$ and (30) becomes (31)

$$x(t+T)-x(t) = \sum_{j=1}^{T} f(\langle t\rangle-j) \tag{31}$$

and $x(t)$ is T-periodic if and only if $f(t)$ satisfies (32) for all t.

$$\sum_{j=1}^{T} f(\langle t\rangle-j) = 0 \tag{32}$$

Whenever the solution is defined only for t integer, then (32) is the condition that is necessary and sufficient for a T-periodic solution of (28) to exist [10].

For any real T, let x_1 be a T-periodic solution of the homogeneous equation $Ex = x$. Since all solutions of the homogeneous equation are 1-periodic, x_1 is both 1-periodic and T-periodic. If T is irrational, then the only possibility is $x_1(t)$ = constant. If $T = m/n$ where m and n are relatively prime integers, then any $1/n$-periodic function will do for $x_1(t)$.

If T is an integer, then any 1-periodic function will do for $x_1(t)$.

Suppose that

$$I = \int_0^T x_1(t+s)^* f(t+s)ds \neq 0$$

where $x_1(t+s)^*$ denotes the complex conjungate transpose of $x_1(t+s)$. This integral is independent of t. For any solution $x(t)$ of (28), consider

$$\gamma(t) = \int_0^T x_1(t+s)^* x(t+s)ds .$$

Since $x_1(t+s+1) = x_1(t+s)$, we have

$$\gamma(t+1) = \int_0^T x_1(t+s+1)^* x(t+s+1)ds$$

$$= \int_0^T x_1(t+s)^*(x(t+s)+f(t+s))ds = \gamma(t)+I$$

and hence

$$\gamma(t) = \gamma(\langle t\rangle) + [t]I .$$

Since $|\gamma(t)| \to \infty$ as $t \to \infty$, we have that $x(t)$ must be unbounded as $T \to \infty$ and hence cannot be periodic. This condition $I \neq 0$ is the condition of resonance and is exactly analogous to resonance in differential equations.

LEMMA 9 A necessary condition for a T-periodic solution of (28) to exist is that

$$\int_0^T x_1(s)^*f(s)ds = 0. \tag{33}$$

for every T-periodic solution x_1 of the homogeneous equation $Ex = x$.

If T is an irrational number, then condition (33) becomes simply

$$\int_0^T f(s)ds = 0 \quad .$$

If $T = m/n$ where m and n are relatively prime integers, then condition (33) becomes

$$\int_0^T x_1^*(s)f(s)ds = 0$$

for all $1/n$-periodic functions $x_1(t)$.

In the case where T is an integer, all 1-periodic functions x_1 are permissible and this condition becomes

$$I = \int_0^T x_1(s)^*f(s)ds = \sum_{j=1}^{T} \int_{T-j}^{T-j+1} x_1(s)^*f(s)ds$$

$$= \sum_{j=1}^{T} \int_0^1 x_1(s)^*f(s-j)ds$$

$$= \int_0^1 x_1(s)^* \sum_{j=1}^{T} f(s-j)ds = 0 \ .$$

One candidate for $x_1(s)$ is

$$x_1(s) = \sum_{j=1}^{T} f(\langle s\rangle - j)$$

which is 1-periodic. Using this $x_1(s)$ and letting $|x|^2 = x^*x$ for any n-vector x, we get equation (34).

$$I = \int_0^1 \left| \sum_{j=1}^{T} f(s-j) \right|^2 ds = 0 \tag{34}$$

Comparing this to condition (32), we see that condition (32) implies condition (34) and also that condition (34) implies condition (32) almost everywhere in t. If we identify all functions which are equal except on sets of measure zero as the same function, then conditions (32) and (34) are equivalent. In the work that follows, this identification of functions shall always be made. Thus, when we write $f(t) = g(t)$, we mean that they agree for all t except possibly on a set of measure zero.

Let $X_1 = L_2[0,1)$. For a given T, let S be the set of all T-periodic functions f which are square integrable over one period. For each $f \in S$, define $\|f\|$ by

$$\|f\|^2 = \frac{1}{T} \int_0^T |f(s)|^2 ds .$$

The functions $\phi_n \in S$,

$$\phi_n(t) = \exp(i\omega nt) \qquad n = 0, \pm 1, \pm 2, \ldots$$

where $\omega = 2\pi/T$, form a complete orthogonal basis in this Hilbert space S. That is, each f in S can be written as (35)

$$f(t) = \sum_n f_n \phi_n(t) \tag{35}$$

where the sum on n is over all the integers, positive and negative, and the f_n are the Fourier coefficients of f.

$$f_n = \frac{1}{T}\int_0^T \phi_n(-s)f(s)ds \qquad n = 0,\pm1,\pm2,\ldots \tag{35'}$$

Let S_1 be the set of all square integrable T-periodic solutions of the homogeneous equation $Ex = x$. Then S_1 is a subspace of S and consists of functions which are both 1-periodic and T-periodic. If T is an irrational number, then S_1 consists only of constant functions and, if $x \in S_1$, then $x(t) = x_o\phi_o(t)$ for some constant vector x_o. If T is a rational number and $T = M/N$ where M and N are relatively prime integers, then S_1 consists of all square integrable $1/N$-periodic functions and, if $x \in S_1$, then

$$x(t) = \sum_n x_{Mn}\phi_{Mn}(t)$$

where x_{Mn} is the Mn-th Fourier coefficient of x.

Let P be the projection operator which maps S onto S_1. That is, if $x \in S$, then $Px \in S_1$ and, if T is irrational

$$Px(t) = x_o\phi_o(t) = \frac{1}{T}\int_0^T x(s)ds \tag{36}$$

and, if $T = M/N$ as before, then

$$Px(t) = \sum_n x_{Mn}\phi_{Mn}(t) \tag{36'}$$

Let S_2 be the orthogonal complement of S_1, i.e., all the functions x in S which satisfy $Px = 0$.

If $x \in S$, then $Ex \in S$ and

$$Ex(t) = x(t+1) = \Sigma_n x_n \phi_n(t+1) = \Sigma_n x_n \phi_n(1)\phi_n(t)$$

and, if T is irrational

$$(PEx)(t) = \frac{1}{T}\int_0^T x(s+1)ds = \frac{1}{T}\int_0^T x(s)ds = (Px)(t) \ .$$

If $T = M/N$ as before

$$(PEx)(t) = \Sigma_n x_{Mn}\phi_{Mn}(1)\phi_{Mn}(t) = \Sigma_n x_{Mn}\phi_{Mn}(t) = (Px)(t)$$

since $\phi_{Mn} \in S_1$ (hence $\phi_{Mn}(1) = 1$) for $n = 0, \pm 1, \pm 2, \ldots$. In either case, we get the important fact that $PEx = Px$ for any $x \in S$.

<u>LEMMA 10</u> The condition $Pf = 0$ is necessary for a T-periodic solution of the difference equation (28) to exist.

$$Ex = x + f(t) \tag{28}$$

This condition $Pf = 0$ is the same as conditions (32) and (33).

Lemma 10 and Lemma 9 are the same. Both are included since they are proven using different terminology. The condition $Pf = 0$ is the same as condition (33) but the former condition is easier to picture geometrically and is easier to check. P is a projection operator which projects any T-periodic f in S onto the subspace S_1 of T-periodic solutions of $Ex = x$. Thus, the condition $Pf = 0$ is that f must be orthogonal to the subspace S_1. P has the explicit representation given in equation (36), so we can check if $Pf = 0$ merely by checking certain Fourier coefficients of f.

PROOF: Suppose x is a T-periodic solution of (28). Using this x, operate with the projection P on equation (28).

$$PEx = Px + Pf.$$

Since $PEx = Px$, we get $Pf = 0$.

<u>LEMMA 11</u> If $x \in S$ is a T-periodic solution of (28), then there is a unique T-periodic solution $\hat{x}$ of (28) which satisfies $P\hat{x} = 0$. This $\hat{x}$, considered as a function of $f \in S$, $Pf = 0$, is a linear function of f. If the T-periodic solution x can be chosen in a manner such that it is a continuous function of f, then $\hat{x}: S_2 \to S$ is a continuous linear function of f and there is some k, $0 \leq k < \infty$, such that $\|\hat{x}(f)\| \leq k\|f\|$.

PROOF: For any $x \in S$, the function $\hat{x} = (1-P)x$ is in S and satisfies $P\hat{x} = 0$.

Let $\hat{\hat{x}}$ be another T-periodic solution of (28) satisfying $P\hat{\hat{x}} = 0$. Then $\hat{x}-\hat{\hat{x}}$ is a solution of the homogeneous equation $Ex = x$ and thus is an element of S_1. So

$$\hat{x}-\hat{\hat{x}} = P(\hat{x}-\hat{\hat{x}}) = P\hat{x}-P\hat{\hat{x}} = 0$$

and $\hat{x} = \hat{\hat{x}}$. Hence the T-periodic solution $\hat{x}$ satisfying $P\hat{x} = 0$ exists and is unique. Denote this as a function of f by $\hat{x} = \hat{x}(f)$.

Since $\alpha\hat{x}(f)$ is a solution of (28) with f replaced by αf for any scalar α, $\alpha\hat{x}(f)$ is T-periodic and satisfies $P(\alpha\hat{x}(f)) = \alpha P\hat{x}(f) = 0$, we see that $\alpha\hat{x}(f) = \hat{x}(\alpha f)$.

Since $\hat{x}(f_1) + \hat{x}(f_2)$ is a solution of (28) with f replaced by f_1+f_2 for any $f_1 \epsilon S$, $f_2 \epsilon S$, $Pf_1 = Pf_2 = 0$, $\hat{x}(f_1)+\hat{x}(f_2)$ is T-periodic and satisfies $P(\hat{x}(f_1)+\hat{x}(f_2)) = P\hat{x}(f_1)+P\hat{x}(f_2) = 0$, we see that $\hat{x}(f_1)+\hat{x}(f_2) = \hat{x}(f_1+f_2)$. Hence, $\hat{x}(f)$ is linear in f.

If x is chosen in a manner such that it is a continuous function of f, then, since $\hat{x} = (I-P)x$ is a continuous function of x, $\hat{x}$ is a continuous function of f. The inequality $\|\hat{x}(f)\| \leq k\|f\|$ follows from the fact that $\hat{x}$ is a continuous linear function of f [7].

This preceding lemma shows that, if a T-periodic solution exists, then there is a unique solution $\hat{x}$ which has all the desired properties, notably that $P_o\hat{x} = 0$. The problem now is one of finding a T-periodic solution to (28).

For the case when T is a rational number, the solution $\hat{x} = \hat{x}(f)$ can be found by comparing Fourier series directly.

THEOREM 7 If $f \in S$ satisfies $Pf = 0$ and $T = M/N$ where M and N are relatively prime integers, then there is a continuous linear operator $L: S_2 \to S_2$ such that $\hat{x} = L(f)$ is the unique T-periodic solution of (28) which

$$Ex = x + f(t) \tag{28}$$

satisfies $P\hat{x} = 0$.

PROOF: If there is such an $\hat{x} \in S_2$, then it can be represented by its Fourier Series.

$$\hat{x}(t) = \sum_n x_n \phi_n(t)$$

where the x_n, $n = 0, \pm 1, \pm 2, \ldots,$ are the Fourier coefficients of $\hat{x}$. Since $f \in S_2$, it has the Fourier series

$$f(t) = \sum_n f_n \phi_n(t)$$

The condition $Pf = 0$ is the condition that $f_n = 0$ whenever n/T is an integer. Then

$$\hat{x}(t+1) - \hat{x}(t) - f(t) = \sum_n ((\phi_n(1)-1)x_n - f_n)\phi_n(t) .$$

If $\hat{x}$ is to be a solution of (28), then each of the coefficients must be zero, i.e.,

$$(\phi_n(1)-1)x_n = f_n \qquad n = 0, \pm 1, \pm 2, \ldots$$

If n/T is an integer then $f_n = 0$, $\phi_n(1) = 1$, and x_n is arbitrary. The condition $P\hat{x} = 0$ specifies $x_n = 0$ whenever n/T is an integer. If n/T is not an integer, then $\phi_n(1) \neq 1$ and x_n is given by (37)

$$x_n = \frac{f_n}{\phi_n(1)-1} \tag{37}$$

Thus, $\hat{x}$ is a linear function of f. This preceding work is valid if we can show that this $\hat{x}$ is square integrable.

Since T is a rational number and $\phi_n(t) = \exp(i\omega nt)$, $T = M/N$ where M and N are relatively prime integers and

$$|e^{i\omega n}-1|^2 = 4\sin^2\tfrac{1}{2}\omega n \geq 4\sin^2\frac{\pi}{M} = \epsilon^2 > 0$$

for all n such that $n/T \neq$ an integer. Then

$$\sum_n |x_n|^2 = \sum_n{}' \frac{|f_n|^2}{|e^{i\omega n}-1|^2} \leq \frac{1}{\epsilon^2}\sum_n |f_n|^2$$

where $\sum_n'$ means the sum over all the integers n such that $n/T \neq$ an integer. This shows that $\hat{x}$ is in S_2 and is a continuous linear function of f. The linear operator L is defined by

$$\hat{x}(t) = (Lf)(t) = \sum_n{}' \frac{f_n}{(\phi_n(1)-1)} \phi_n(t)$$

where f_n is the n-th Fourier coefficient of f and $\sum_n'$ has the same meaning as before. By Lemma 11, we know that this is the only such solution.

If T is an irrational number, then the Fourier coefficients of $\hat{x}$ would still be defined by (37) (except for x_o which is zero). The problem then is one of showing that

$$\sum_n |x_n|^2 = \|\hat{x}\|^2$$

is bounded or under what conditions is $\hat{x}$ a continuous linear function of f.

The permissible functions f will probably have to be restricted to be some subset of S_2. For one solution of this problem, see Moser [11].

If there is a continuous linear operator $L:S_2 \rightarrow S_2$ such that $\hat{x} = Lf$ is the unique solution of (28) which satisfies $P\hat{x} = 0$, then a class of non-linear problems could be studied in a manner similar to Hale [12]. However, there are problems in determining the meaning of the fact that the non-linear term must be bounded and Lipschitz continuous in the norm of S, i.e. the $L_2[0,T]$ norm.

As an example, consider the Crude Euler difference analog to the differential equation

$$\dot{x} = f(t)$$

where $f(t+T) = f(t)$. Letting $t = hs$, the analog is

$$Ex = x + hf(hs) = x + g(s) .$$

If we chose h such that $T = hN$ for some integer N, then this difference analog is like equation (28). In this case, the condition $Pg = 0$, which is necessary and sufficient for an N-periodic solution of the difference analog to exist, is that

$$\int_0^N e^{i2\pi ns} g(s)ds = 0 \qquad n = 0, \pm 1, \pm 2, \ldots$$

or, in terms of f and t

$$\int_0^T e^{i\frac{2\pi nt}{h}} f(t)dt = \int_0^T e^{i\frac{2\pi}{T}Nnt} f(t)dt = 0 \quad n = 0, \pm 1, \pm 2, \ldots$$

If

$$\int_0^T f(t)dt = 0$$

then a T-periodic solution to the differential equation does exist. However, we see that this is only one of the conditions needed to have a N-periodic solution of the difference analog to exist. In fact, every N-th Fourier coefficient of f must be zero before there exists an N-periodic solution to the difference analog. If one or more of these Fourier coefficients are not zero, then there does not exist an N-periodic solution to the difference analog.

If $T/h = N$ is an irrational number, then the condition $Pg = 0$ reduces to the condition which is necessary and sufficient for a T-periodic solution of the differential equation to exist. However, I have not been able to show that the condition $Pg = 0$ is sufficient for an N-periodic solution of the difference analog to exist when N is an irrational number.

The aim of all this work is to develop a bifurcation theory analogous to that in Hale [12]. The first step in this theory is to show that there exists a continuous linear mapping $L: S_2 \rightarrow S_2$ such that $\hat{x} = Lf$ is the unique T-periodic solution of (28) which satisfies $P\hat{x} = 0$. This has yet to be done for the case when T is an irrational number. Yet, from the above example, it would appear that this would be the most interesting case.

AN AVERAGING METHOD WITH AN APPLICATION

Halanay [10] gives the following theorems for a system of almost-periodic difference equations. While all his work is concerned with the solution being defined only on the integers, his proofs are easily extended to include the solution being defined on the reals. In [10], Halanay first gives an approximation Lemma which is stated here.

LEMMA 12 If $f(t)$ is such that

a) $|f(t)| \leq F$ for all t, some $F \geq 0$

b) $\left|\frac{1}{N} \sum_{k=1}^{N} f(t-k)\right| \leq \epsilon(N)$ for all t, some $\epsilon(N) \geq 0$ and $\epsilon(N) \to 0$ as $N \to \infty$. For some η, $0 < \eta < 1$, let f_η be the function defined by

$$f_\eta(t) = \sum_{k=1}^{\infty} (1-\eta)^{k-1} f(t-k).$$

Then there exists a continuous function $\delta(\eta)$ with $\delta(\eta) \to 0$ as $\eta \to 0$ such that

$$\eta |f_\eta(t)| \leq \delta(\eta) \quad \text{for all } t$$

$$|f_\eta(t+1) - f_\eta(t) - f(t)| \leq \delta(\eta) \quad \text{for all } t.$$

That is, f_η is an approximate solution to the difference equation (28) which can be made as accurate as desired by taking η small enough.

$$Ex = x + f(t) \tag{28}$$

Condition b) is the natural one to use for difference equations since the solution of (28) is a sum of $f(t-k)$. However, this condition is usually very hard to verify. When this lemma is applied to the numerical solution of differential equations, we will have $f(t) = g(ht)$ for some $h > 0$ and

$$\left|\frac{1}{T}\int_0^T g(t-s)ds\right| \leq \epsilon_1(T) \quad \text{for all} \quad t$$

where $\epsilon_1(T) \to 0$ as $T \to \infty$.

The following identity, a modification of the Euler-Maclaurin summation formula, is easily verified by integration by parts.

$$\sum_{k=1}^{N} f(t-k) = \int_0^N f(t-s)ds - \int_0^N f'(t-s)\langle s\rangle ds$$

where $f'(t)$ is the first derivative of f with respect to T. Making the change of variables $f(t) = g(ht)$, we get

$$\frac{1}{N}\sum_{k=1}^{N} f(t-k) = \frac{1}{T}\int_0^T g(ht-s)ds - \frac{h}{T}\int_0^T g'(ht-s)\langle\frac{u}{h}\rangle du$$

where $T = Nh$. In general, this last integral will not go to zero as $T \to \infty$ even though the first one does. However, under fairly general conditions, e.g., $|g'(t)|$ bounded for all t, we can bound this last integral by hG.

$$\left|\frac{h}{T}\int_0^T g'(ht-s)\langle\frac{s}{h}\rangle ds\right| \leq hG \quad \text{for all} \quad t$$

Then

$$\left|\frac{1}{N}\sum_{k=1}^{N} f(t-k)\right| \leq \epsilon_1(T) + hG = \epsilon_1(Nh) + hG .$$

In general, we get this sum going to zero only by letting h go to zero. However, if we look at the proof of Lemma 12, we see that the critical inequality is

$$|f_\eta(t)| \leq \frac{N\ \epsilon(N)}{1-(1-\eta)^N} + NF$$

and we get the desired $\delta(\eta)$ by letting $N = N(\eta)$ be chosen as a function of η such that

$$\epsilon(N) \leq 1-(1-\eta)^N$$

Then we get $\delta(\eta) = \eta N(\eta)(1+F)$.

If, instead of the sum of f went to zero, we knew only that the integral of g went to zero where $f(t) = g(ht)$, then this crucial inequality becomes:

$$|f_\eta(t)| \leq \frac{N(\epsilon_1(Nh)+hG)}{1-(1-\eta)^N} + NF$$

We now let N and h be chosen as functions of η such that

$$2\ h\ G \leq 1-(1-\eta)^N$$
$$2\epsilon_1(Nh) \leq 1-(1-\eta)^N$$

These functions $h = h(\eta)$ and $N = N(\eta)$ satisfy $h(\eta) \to 0$ as $\eta \to 0$, $h(\eta)N(\eta) \to \infty$ as $\eta \to 0$, and $\eta N(\eta) \to 0$ as $\eta \to 0$. Then we get

$$\delta(\eta) = \eta N(\eta)(1+F)$$

Thus, Lemma 12 is still true. However, this solution has the undesirable property that we must let $h \to 0$ as $\eta \to 0$. This means that, in order to have $\delta(\eta)$ small, i.e., have $f_\eta(t)$ be an accurate solution of (28), then we must take both N large and h small.

There is one exception to this problem of $h(\eta) \to 0$ as $\eta \to 0$. If $g(t)$ is P-periodic and P/h is an irrational number, then the theory of equidistributed sequences [14,15] gives us that

$$\left|\frac{1}{N} \sum_{k=1}^{N} g(ht-hk) - \frac{1}{P} \int_0^P g(s)ds\right| \to 0 \quad \text{as} \quad N \to \infty .$$

In this case, if

$$\frac{1}{P} \int_0^P g(s)ds = 0$$

then we have the existence of some $\epsilon_3(N)$ such that

$$\left|\frac{1}{N} \sum_{k=1}^{N} g(ht-hk)\right| \le \epsilon_3(N) \quad \text{for all} \quad t$$

and $\epsilon_3(N) \to 0$ as $N \to \infty$. Thus, in this case, the conditions of Lemma 12 are satisfied. If P/h is rational, then we can construct a g such that this sume does not go to zero as $N \to \infty$, and we have to pick h small in order to get the desired accuracy.

Consider the difference equation (38) where $A(t)$, $A^{-1}(t)$, and $f(t,x,\epsilon)$ are almost periodic in t uniformly in x and ϵ.

$$Ex = A(t)x + \epsilon f(t,x,\epsilon) \tag{38}$$

Let $X(t)$ be the principle matrix solution of the homogeneous equation (39) with $X(\theta) = I$, $0 \leq \theta < 1$, and let $X(t)$ and $X^{-1}(t)$ be almost periodic in t.

$$Ex = A(t)x \tag{39}$$

If A and f are periodic in t, then they are also almost periodic. The problem in requiring $X(t)$ to be periodic is that this choice of the initial function $X(\theta) = I$, $0 \leq \theta < 1$, will not in general lead to a periodic solution even if there exists a fundamental matrix solution of (39) which is periodic. Determining that the principle matrix solution is almost periodic is an easier task than finding a fundamental matrix solution which is periodic. For example, if $A(t) = A$, a constant 2 by 2 matrix, both the eigenvalues of A have simple elementary divisors, and the eigenvalues of A lie on the unit circle, then the principle matrix solution $X(t)$ is almost periodic. A very careful choice of $X(\theta)$, $0 \leq \theta < 1$, will have to be made before $X(t)$ will be periodic. In this sense, it seems more natural to look for almost periodic solutions of difference equations than to look for periodic solutions.

By the change of variables $x = X(t)y$, the difference equation (38) becomes

$$Ey = y + \epsilon g(t,y,\epsilon) \tag{40}$$

where

$$g(t,y,\epsilon) = X(t+1)^{-1}f(t,X(t)y,\epsilon)$$

is almost periodic in t uniformly in y and ϵ. Let $g_o(y,\epsilon)$ be

$$g_o(y,\epsilon) = \lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} g(t-k,y,\epsilon) \tag{41}$$

The average value of $g(t,y,\epsilon)$ (since $g(t,y,\epsilon)$ is almost periodic, the limit exists uniformly in y and ϵ and is independent of t). Let y_o be a solution of $g_o(y,0) = 0$ and let the matrix H be

$$H = \frac{\partial g_o}{\partial y}(y_o,0)$$

The following theorem is a restatement of Theorem 5 in Halanay [10].

<u>THEOREM 8</u> The principle of averaging: If the real parts of all the eigenvalues of H are less than zero, then there exists an almost periodic solution $x(t,\epsilon)$ of (38) for each ϵ, $0 < \epsilon < \epsilon_o$, some $\epsilon_o > 0$, which reduces to $x(t,0) = X(t)y_o$ as $\epsilon \to 0$.

Consider as an example of the application of Theorem 8 the Crude Euler difference analog for (42)

$$\dot{x} = \epsilon f(t,x,\epsilon) \tag{42}$$

where $f(t,x,\epsilon)$ is almost periodic in t uniformly in x and ϵ. Let the average value of $f(t,x,\epsilon)$ be $f_o(x,\epsilon)$.

$$f_o(x,\epsilon) = \lim_{T \to \infty} \frac{1}{T} \int_o^T f(t-s,x,\epsilon)ds$$

The Crude Euler analog is given in (43)

$$Ex = x + \epsilon g(t,x,\epsilon) \tag{43}$$

where

$$g(t,x,\epsilon) = hf(ht,x,\epsilon).$$

This equation already is in the form of equation (40) so the change of variables $x = X(t)y$ does not have to be performed. The average value of $g(t,x,\epsilon)$, $g_1(x,\epsilon)$, is given by

$$g_1(x,\epsilon) = \lim_{T \to \infty} \frac{1}{T} \int_o^T g(t-s,x,\epsilon)ds = hf_o(x,\epsilon)$$

This is not the average value $g_o(x,\epsilon)$ used in Theorem 8 but, considering the discussion following Lemma 12, if we choose h as a function of ϵ properly, then $h = h(\epsilon) \to 0$ as $\epsilon \to 0$ and the conclusions of Theorem 8 still hold when $g_1(x,\epsilon)$ replaces $g_o(x,\epsilon)$. The matrix H

$$H = \frac{\partial g_1}{\partial x}(x_o,0) = h\frac{\partial f_o}{\partial x}(x_o,0)$$

goes to zero as $h \to 0$ but this does not create any problems since $g(t,x,\epsilon) \to 0$ as $h \to 0$ and the effects of these two phenomena cancel each other.

If x_o is a solution of $f_o(x,0) = 0$ and the real parts of all the eigenvalues of H are negative for $h > 0$, then there exists an almost periodic solution $x(t,\epsilon)$ of (42) for each ϵ, $0 < \epsilon < \epsilon_o$, some $\epsilon_o > 0$, which reduces to $x(t,0) = x_o$ as $\epsilon \to 0$. See, for example, [12] for a proof of this statement. The significance of this application of Theorem 8 is that, under the same conditions that insure the existence of an almost periodic solution of the differential equation (42), and the step size h is chosen small enough, there is an almost periodic solution of the difference analog (43) and, furthermore, these two solutions reduce to the same solution $x(t,0) = x_o$ as $\epsilon \to 0$.

While this is a case where the differential equation and its difference analog display the same qualitative behavior, i.e., the existence of an almost periodic solution, the above analysis leaves several questions unanswered. For example, do the two almost periodic solutions display the same stability properties? How close is the almost periodic solution of the difference analog to the almost periodic solution of the differential equation?

AN INVARIANT MANIFOLD METHOD WITH AN APPLICATION

Consider the Van der Pol equation (44) with $\epsilon > 0$.

$$\ddot{x} - 2\epsilon\dot{x}(1-x^2) + x = 0 \tag{44}$$

With a constant spacing $h > 0$ and a change of independent variable from t to $s = ht$, the central difference analog to (44) becomes (45)

$$\begin{aligned} Ey &= x \\ Ex &= (2-h^2)x - y + \epsilon h f(x,y) \end{aligned} \tag{45}$$

where

$$f(x,y) = f(x,y,h,\epsilon) = \frac{(1-x^2)((2-h^2)x-2y)}{1-\epsilon h(1-x^2)}$$

It is well known that the Van der Pol equation (44) has a stable limit cycle which is a periodic solution with amplitude approximately 2 for small ϵ [12]. It would seem reasonable to expect the difference analog (45) to display this same type of behavior, at least for the spacing h small enough. Yet the examples given at the beginning of this paper show that such blanket assumptions about the behavior of the difference analog reflecting the behavior of the differential equation cannot be made. The difference analog may display behavior completely different from that displayed by the differential equation. The purpose of this section is to show that the qualitative behavior of the difference analog (45) is the same as the qualitative behavior of the differential equation (44) for ϵ small.

The central difference analog is chosen for the following reasons. It involves no more than the two variables x and y and so solutions can be displayed on a piece of paper without undue difficulty. Also, at $\epsilon = 0$, both the differential equation (44) and its difference analog (45) have the behavior of a center. This desirable property is not shared by such other methods as the Crude Euler and the Corrected Euler. The Central Difference analog is chosen over the Iterated Euler analog because the difference equation for the Central Difference analog is much simpler.

For $\epsilon = 0$, we know that the solutions of (45) lie on a level curve of $V(x,y)$ where

$$V(x,y) = \frac{1}{4}(x+y)^2 + \alpha^2(x-y)^2$$

and

$$\alpha^2 = (4-h^2)/(4h^2) \quad , \quad \alpha > 0 .$$

A level curve of $V(x,y)$ is an ellipse as shown in Figure 6 where $2\cos\omega = 2-h^2$. For h small, these ellipses are very long and narrow. If we make a change of variables to u and v where

$$u = \tfrac{1}{2}(x+y) \qquad x = u+(v/2\alpha)$$
$$v = \alpha(x-y) \qquad y = u-(v/2\alpha) \tag{46}$$

then $V(x,y) = u^2 + v^2$ and, in the u,v plane, level surfaces of V are circles. The difference equation (45) becomes the difference equation (47)

$$Eu = \cos\omega\, u + \sin\omega\, v + \epsilon h U(u,v)$$
$$Ev = -\sin\omega\, u + \cos\omega\, v + \epsilon h 2\alpha U(u,v) \tag{47}$$

where

$$U(u,v) = U(u,v,h,\epsilon) = \tfrac{1}{2}f(u+(v/2\alpha),\ u-(v/2\alpha),h,\epsilon)\ .$$

The following identities are used in deriving the following equations.

$$2\cos\omega = 2-h^2, \quad \sin\omega = h^2\alpha, \quad \sin\tfrac{1}{2}\omega = \tfrac{1}{2}h, \quad \cos\tfrac{1}{2}\omega = \alpha h\ .$$

Since all solutions lie on circles in the u,v plane when $\epsilon = 0$, it is natural to change coordinates again to polar coordinates.

$$\begin{aligned} u &= r\cos\theta \qquad & x &= r\cos(\theta-\tfrac{1}{2}\omega)/(\alpha h) \\ v &= r\cos\theta \qquad & y &= r\cos(\theta+\tfrac{1}{2}\omega)/(\alpha h) \end{aligned} \tag{48}$$

If we let $w = r^2$, then the difference equation ~~(47) becomes~~

$$\begin{aligned} Ew &= w + \epsilon hR(w,\theta) \\ E\theta &= \theta - w + \epsilon h\Theta(w,\theta) \end{aligned} \tag{49}$$

where R and Θ are 2π-periodic in θ. For the Van der Pol equation (44), we have

$$\begin{aligned} R(r^2,\theta) &= R(r^2,\theta,h,\epsilon) = \frac{1}{h^2}\{(2-h^2)x-2y+\epsilon hf(x,y,h,\epsilon)\}f(x,y,h,\epsilon) \\ \Theta(r^2,\theta) &= \Theta(r^2,\theta,h,\epsilon) = \frac{1}{\epsilon h}\,\mathrm{Arctan}\left\{\frac{\epsilon f(x,y,h,\epsilon)\cos(\theta-\frac{1}{2}\omega)}{r+\epsilon f(x,y,h,\epsilon)\sin(\theta-\frac{1}{2}\omega)}\right\} \end{aligned} \tag{50}$$

where the arctangent is taken such that $-\frac{1}{2}\pi \leq \mathrm{Arctan}\ (x) \leq \frac{1}{2}\pi$. Notice that Θ

is continuous and bounded in ϵ as $\epsilon \to 0$ even though there is a $1/\epsilon$ coefficient on the arctangent. It is also easy to show that both R and Θ are continuous and remain bounded as $h \to 0$ despite the $1/h^2$ and $1/h$ coefficients since the formulas for x and y in terms of r, θ, and h are given by (48).

In the difference equation (49) we can look for either of two things:

1) a solution for w and θ which satisfies $w(t+T) = w(t)$ and $\theta(t+T) = \theta(t) + 2\pi$. This represents a T-periodic solution of the difference equations (45) and (47). The great difficulty with this method is that T is not known a priori since it will be, in general, a function of ϵ and h.

2) a parametric representation for w in terms of θ which is 2π-periodic and continuous in θ such that one solution of (49) is $\theta = \theta(t)$ and $w = w(\theta(t))$, i.e., an invariant manifold which is a closed curve in the u,v or the x,y planes. This is what is used in practice since, in computing solutions of Van der Pol's equation (44) numerically, a periodic solution is "found" when the numeric solution displays such a closed curve. The problem here is to show that such an invariant manifold exists and to compare it to the limit cycle of (44).

The difference equation (49) looks a lot like the difference equation studied by Moser [11] if we write Θ as the sum of its average value Θ_0 and the remainder $\Theta_1 = \Theta - \Theta_0$.

$$\Theta_0(w,h,\epsilon) = \frac{1}{2\pi}\int_0^{2\pi} \Theta(w,\theta,h,\epsilon)d\theta \ .$$

Unfortunately, one condition Moser requires is that $\partial\Theta_0/\partial w \geqq 1$ and, since $f(x,y,h,\epsilon)$ is odd in x and in y, we have $\Theta_0 = 0$ identically in w, h, ϵ.

Thus, the work done by Moser does not apply here.

Halanay [9] developes a theorem which can be applied to this problem after a slight modification. This theorem, Theorem 4 in Halanay [9], is restated here without proof.

THEOREM 9: On Existence of an Invariant Manifold. Consider the difference equation (51) where $y \in E^n$, θ is a real number, and Y_1 and Θ are 2π-periodic in θ.

$$\begin{aligned} Ey &= Y_o(t,y) + \epsilon Y_1(t,y,\theta,\epsilon) \\ E\theta &= \theta + \alpha(t) + \epsilon\Theta(t,y,\theta,\epsilon) \end{aligned} \tag{51}$$

Assume there exists positive constants H, K_1, μ, and q with $0 < q < 1$ such that, for all $t \geq t_o$, $|y| \leq H$, and $0 \leq \theta \leq 2$, Y_o, Y_1 and Θ have continuous first partial derivatives with respect to y and θ and

(1) $\|\frac{\partial Y_o}{\partial y}(t,y)\| \leq K_1$, $\|\frac{\partial \Theta}{\partial y}(t,y,\theta,\epsilon)\| \leq K_1$, $\|\frac{\partial \Theta}{\partial \theta}(t,y,\theta,\epsilon)\| \leq K_1$

(2) $\|\frac{\partial Y_o}{\partial y}(t,y_1) - \frac{\partial Y_o}{\partial y}(t,y_2)\| \leq K_1\|y_1-y_2\|^\mu$

(3) $\|\frac{\partial Y_1}{\partial y}(\theta_1,y_1,\epsilon)-\frac{\partial Y_1}{\partial y}(\theta_2,y_2,\epsilon)\| \leq K_1(\|y_1-y_2\|^\mu+\|\theta_1-\theta_2\|^\mu)$

(4) $\|\frac{\partial \Theta}{\partial \theta}(\theta_1,y_1,\epsilon)-\frac{\partial \Theta}{\partial \theta}(\theta_2,y_2,\epsilon)\| \leq K_2(\|y_1-y_2\|^\mu+\|\theta_1-\theta_2\|^\mu)$

(5) $Y_o(t,0) = 0$

(6) Let $A(t) = \frac{\partial Y_o}{\partial y}(t,0)$ and, if $z(t)$ is any solution of $Ez = A(t)z$, then

$$\|z(t)\| \leq K_1 q^{t-t_o} \|z(t_o)\|$$

for all $t \geq t_o$.

Then, for $|\epsilon|$ small enough there exists a function $p(t,\theta)$, constants $\ell = \ell(\epsilon)$ and $L = L(\epsilon)$, continuous and monotone in ϵ, such that $\ell(0) = L(0)$ and $\ell(\epsilon) > 0$, $L(\epsilon) > 0$ for $\epsilon \neq 0$ and

a) $\|p(t,\theta)\| \leq \ell(\epsilon)$

b) $\|p(t,\theta_1)-p(t,\theta_2)\| \leq L(\epsilon)|\theta_1-\theta_2|$

c) If $\|y_o\| \leq \ell(\epsilon)$ then

$$\|y(t;t_o,y_o,\theta_o)-p(t,\theta(t;t_o,y_o,\theta_o))\| \leq K_1 q^{t-t_o}\|y_o-p(t_o,\theta_o)\|$$

d) If $y_o = p(t_o,\theta_o)$, then $y(t;t_o,y_o,\theta_o) = p(t,\theta(t;t_o,y_o,\theta_o))$

e) This $p(t,\theta)$ is unique and 2π-periodic in θ

f) If Y_o, Y_1, and Θ are T-periodic in t, then so is p.

g) If Y_o, Y_1, and Θ are independent of t, then so is p.

The set of points (t,y,θ) where $y = p(t,\theta)$ is an invariant manifold of the system (56) (result d), it is never very far from the set $(t,0,\theta)$ (result a), it is Lipschitzian in θ (result b), and solutions near $y = 0$ approach the manifold exponentially as $t \to \infty$ (result c).

For the difference equation (49), let

$$R_o(w) = \frac{1}{2\pi}\int_o^{2\pi} R(w,\theta)d\theta$$

and let $R_1(w,\theta) = R(w,\theta)-R_o(w)$. Let w_o be a solution of $R_o(w) = 0$ and let $w = w_o + y$. Let $R_o'(w)$ denote the first derivative of R_o with respect to w. Then the difference equation (49) becomes (52)

$$\begin{aligned} Ey &= (1+\epsilon h R_o'(w_o))y + \epsilon h R_2(y,\theta) \\ E\theta &= \theta - \omega + \epsilon h \Theta(w_o + y,\theta) \end{aligned} \tag{52}$$

where

$$R_2(y,\theta) = R_o(w_o + y)-R_o'(w_o)y + R_1(w_o + y,\theta).$$

This equation (52) is of the form of equation (51) with the matrix $A(t) = 1 + \epsilon h R_o'(w_o)$. If $\epsilon h R_o'(w_o) < 0$, then the conditions of Theorem 9 seem to be satisfied. The problem is that the estimate q is given by $q = 1+\epsilon h R_o'(w)$ and $q \to 1$ as $\epsilon \to 0$ or as $h \to 0$. Since the proof of Theorem 9 depends on forming a contraction mapping where the contraction constant is $\alpha = \epsilon K/(1-q)$ for some $K > 0$, we see that this theorem cannot be really applied to this problem.

Try letting $w = p + \epsilon h u(p,\theta,\epsilon)$. Then the difference equation (49) for w becomes (53)

$$\begin{aligned} Ep-p+\epsilon(u(Ep,E\theta,\epsilon)-u(p,E\theta,\epsilon)) &= \epsilon h R_o(p,0)-\epsilon h(u(p,E\theta,\epsilon)-u(p,\theta,\epsilon)- \\ &\qquad R_1(p,\theta,\epsilon)) \\ &\quad + \epsilon h(R_o(p+\epsilon u,\epsilon)-R_o(p,0))+\epsilon h(R_1(p+\epsilon u,8,\epsilon)-R_1(p,\theta,\epsilon)) \end{aligned} \tag{53}$$

where $E\theta$ is the function of θ, ϵ, and $w = p+\epsilon u$ given in the second equation of (49). Since R_o and R_1 are continuous in all their arguments, there is some continuous monotone function $\delta(\epsilon)$ with $\delta(0) = 0$ and $\delta(\epsilon) > 0$ for $\epsilon > 0$ such that

$$|R_o(p+\epsilon u, \epsilon) - R_o(p, 0)| \leq \delta(\epsilon)$$

$$|R_1(p+\epsilon u, \theta, \epsilon) - R_1(p, \theta, \epsilon)| \leq \delta(\epsilon)$$

Furthermore, we can choose h as a function of p, θ, ϵ such that (see Lemma 12)

$$u(p, E\theta, \epsilon) - u(p, \theta, \epsilon) - R_1(p, \theta, \epsilon) = -\epsilon u(p, \theta, \epsilon) \tag{54}$$

and

$$\epsilon |u(p, \theta, \epsilon)| \leq \delta(\epsilon) .$$

Thus, (53) becomes (55)

$$\begin{aligned} Ep - p + \epsilon(u(Ep, E\theta, \epsilon) - u(p, E\theta, \epsilon) \\ = \epsilon h R_o(p, 0) + \epsilon h R_2(p, \theta, \epsilon) \end{aligned} \tag{55}$$

where

$$|R_2(p, \theta, \epsilon)| \leq 3\delta(\epsilon) \to 0 \quad \text{as} \quad \epsilon \to 0.$$

Since u satisfies (54), we get that

$$\frac{\partial u}{\partial p}(p, E\theta, \epsilon) - \frac{\partial u}{\partial p}(p, \theta, \epsilon) - \frac{\partial R_1}{\partial p}(p, \theta, \epsilon) = -\epsilon \frac{\partial u}{\partial p}(p, \theta, \epsilon)$$

and that

$$\epsilon\left|\frac{\partial u}{\partial p}(p,\theta,\epsilon)\right| \leq \delta(\epsilon) \to 0 \quad \text{as} \quad \epsilon \to 0 .$$

Thus, by the implicit function theorem, see [18] for example, there is a function $H = H(p,y,\theta,\epsilon)$ such that $H(p,0,\theta,\epsilon) = 0$ and

$$Ep-p = \epsilon hR_o(p,0) + \epsilon hR_3(p,\theta,\epsilon) \tag{56}$$

is the solution of (55) for ϵ small enough where

$$R_3(p,\theta,\epsilon) = R_2(p,\theta,\epsilon) + H(p,\epsilon hR_o(p,0)+\epsilon hR_o(p,\theta,\epsilon),E\theta,\epsilon)$$

and

$$|R_3(p,\theta,\epsilon)| \leq \delta_1(\epsilon) \to 0 \quad \text{as} \quad \epsilon \to 0. \tag{56'}$$

Let p_o be a root of $R_o(p,0) = 0$ and let $p = p_o + v$. Then the difference equation (56) becomes, in terms of v,

$$Ev = (I+\epsilon hR_o'(p_o,0))v + \epsilon hR_4(v) + \epsilon hR_3(p_o + v,\theta,\epsilon) \tag{57}$$

where

$$R_4(v) = R_o(p_o + v,0) - R_o'(p_o,0)v \tag{57'}$$

and there is some $\delta_2(|v|) > 0$ for $|v| > 0$, $\delta_2(|v|) \to 0$ as $|v| \to 0$, and

$$|R_4(v)| \leq |v|\,\delta_2(|v|) .$$

If none of the eigenvalues of $R_o'(p_o,0)$ lie on the imaginary axis, then there is some $\epsilon_1 h_1 > 0$ such that, for $0 < \epsilon h < \epsilon_1 h_1$, none of the eigenvalues of $(I+\epsilon h R_o'(p_o,0))$ lie on the unit circle and, from Theorem 5, there is a continuous linear operator $L: B \to B$ such that $v = Lf$ is the unique bounded solution of $Ev = (I+\epsilon h R_o'(p_o,0))v + f$. There is some constant $K > 0$ such that

$$\|Lf\| \leq \frac{K}{h\epsilon} \|f\| .$$

We are interested in a fixed point of the non-linear operator $N: B \to B$

$$N(v) = L(\epsilon h R_4(v) + \epsilon h R_3(p_o + v, \theta, \epsilon))$$

We see that, since we have the inequalities (56') and (57'), we can choose the set $S = \{v \in B: \|v\| \leq \delta_3\}$, i.e., δ_3, and h as functions of ϵ for ϵ small enough so that N is a contraction mapping of S, hence N has a fixed point v. Since R_3 is 2π-periodic in θ, this fixed point will be 2π-periodic in θ.

These results are summarized in the following theorem.

<u>THEOREM 10</u> Consider the difference equation (49).

$$\begin{aligned} Ew &= w + \epsilon h R_o(w,\epsilon) + \epsilon h R_1(w,\theta,\epsilon) \\ E\theta &= \theta - \omega + \epsilon h \Theta(w,\theta,\epsilon) \end{aligned} \qquad (49)$$

where R_o and R_1 have continuous first partial derivatives with respect to w and

$$\frac{1}{2\pi}\int_0^{2\pi} R_1(w,\theta,\epsilon)d\theta = 0.$$

If w_o is a solution of $R_o(w_o,0) = 0$ and none of the eigenvalues of the matrix $R_o'(w_o,0)$ lie on the imaginary axis, then there is an $\epsilon_1 > 0$, and functions $h(\epsilon)$ and $w(\theta,\epsilon,h)$ such that, if $0 < \epsilon < \epsilon_1$, $0 < h < h(\epsilon)$, then $w(\theta,\epsilon,h)$ represents an invariant curve of (49). Also, $h(\epsilon) \to 0$ as $\epsilon \to 0$ and $w(\theta,\epsilon,h) \to w_o$ as $\epsilon \to 0$.

This is a generalization of Halanay's theorem [9] which is given here as Theorem 9.

If the application to Van der Pol's equation (44), we have

$$R_o(w,0) = \tfrac{1}{2}w(4-h^2-w)$$

$$R_o'(w,0) = \tfrac{1}{2}(4-h^2-2w)$$

The two roots of $R_o(w,0)$ are $w_1 = 0$ and $w_2 = 4-h^2$. For $w = w_1$, $R_o'(w_1,0) = \frac{1}{2}(4-h^2) > 0$ and we have an invariant curve $w(\theta,\epsilon,h) = 0$. For $w = w_2$, $R_o'(w_2,0) = -\frac{1}{2}(4-h^2) < 0$, and we have an invariant curve $w(\theta,\epsilon,h) \doteq 4-h^2$. Thus we see that the central difference analog (45) to the Van der Pol equation (44) has the same qualitative behavior (w is the square of the amplitude).

SUGGESTIONS FOR FUTURE RESEARCH

The theory of difference equations as presented here is very incomplete. More work needs to be done with finding periodic solutions in critical cases, i.e., in cases when the homogeneous difference equation also has periodic solutions. This includes the case for $Ex = x + f$, considered here but when the period T is an irrational number. This also includes the case for $Ex = Ax + f$ where the matrix A has eigenvalues on the unit circle which may or may not have simple elementary divisiors or where the matrix A is a T-periodic function of t.

More work also needs to be done in studying the behavior of various difference analogs to differential equations which are not considered here. For example, what would a graph of the regions of behavior of the Kutta-Simpson analog for $\ddot{x} + 2c\dot{x} + x = 0$ plotted as in Figure 3 look like? How do the discreprencies affect the usefullness of these methods for finding periodic solutions of a differential equation. If the differential equation has a periodic solution which is exponentially stable, can Theorem 10 be applied to the difference analog to show that it has an invariant manifold which is close to this periodic solution? If an invariant manifold of the difference analog is found which is exponentially stable, does this imply that the differential equation has a periodic or almost periodic solution near this manifold, These are all questions which need to be answered. There are many more questions not listed here which need to be answered.

BIBLIOGRAPHY

[1] Peter Henrici, Discrete Variable Methods in Ordinary Differential Equations, Wiley, New York, 1962. This book has a very good bibliography for the work done before 1962.

[2] Peter Henrici, Error Propagation for Difference Methods, Wiley, New York, 1963.

[3] Royce Beckett and James Hurt, Numerical Calculations and Algorithms, McGraw-Hill, New York, 1967.

[4] Paul R. Halmos, Finite-Dimensional Vector Spaces, D. Van Nostrand, New York, 1958.

[5] Tomlinson Fort, Finite Differences and Difference Equations in the Real Domain, Clarendon Press, London, 1948.

[6] V.N. Faddeeva, Computational Methods of Linear Algebra (translated by Curtis D. Benster), Dover, New York, 1959.

[7] Riesz, F., and B. Sz. Nagy, Functional Analysis, Frederick Ungar Publ., New York, 1952.

[8] M.A. Krasnosel'skii, Topological Methods in the Theory of Nonlinear Integral Equations (translated by A.H. Armstrong and J. Burlak), Macmillan, New York, 1964.

[9] A. Halanay, "Invariant Manifolds for Discrete Systems", to appear in Proceedings of the Second Czeckoslovakian Conference on Differential Equations and their Applications (Equadiff II), Bratislava, Czeckoslavakia, 1966.

[10] A. Halanay, "Periodic and Almost Periodic Solutions of Systems of Finite Difference Equations", Archives of Rational Mechanics and Analysis, Vol. 12(1963), pp. 134-149.

[11] Jürgen Moser, "Perturbation Theory for Almost Periodic Solutions for Undamped Nonlinear Differential Equations", in Nonlinear Differential Equations and Nonlinear Mechanics, edited by J.P. LaSalle and S. Lefschetz, Academic Press, 1963, pp. 71-79.

[12] Jack K. Hale, Oscillations in Nonlinear Systems, McGraw-Hill, New York, 1963.

[13] N. Bogolinbov and Y.A. Mitropolski, "Asymptotic Methods in the Theory of Nonlinear Oscillations", translated by Gordon and Breach, 1962. Goz. Iz. Fiz. Mat. Lit.

[14] Philip J. Davis and [redacted]ation, Blaisdell, New York, 1967.

[15] Philip J. Davis, Interpolation and Approximation, Blaisdell, New York, 1963.

[16] John M. H. Olmsted, Advanced Calculus, Appleton-Century-Crofts, 1956, 1959.

[17] A.S. Besicovitch, Almost Periodic Functions, Dover, New York, 1954.

[18] Wendell H. Fleming, Functions of Several Variables, Addision-Wesley, Reading, Massachusetts, 1966.