JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 63, 199–223 (1978)

Snap-Back Repellers Imply Chaos in $\mathbb{R}^n$

FREDERICK R. MAROTTO

Department of Mathematics, Boston University, Boston, Mass. 02215

Submitted by K. L. Cooke

1. INTRODUCTION

This paper is motivated by a rather surprising theorem proven recently by Li and Yorke in [1] concerning the first-order difference equation:

$$x_{k+1} = f(x_k), \qquad k = 0, 1, \ldots, \tag{1.1}$$

where $x_k \in \mathbb{R}$ and $f\colon \mathbb{R} \to \mathbb{R}$ is continuous. If we let f^n denote the composition of the function f with itself n times, then a point x is said to be a periodic point of period p if $f^p(x) = x$ but $f^k(x) \neq x$ for $1 \leqslant k < p$. In this case the collection $\{f^k(x)\}_{k=1}^{p}$ is said to constitute a p-cycle of (1.1). With this terminology, Li and Yorke have established analytically that the existence of a point of period 3 for the function f, i.e., a 3-cycle of (1.1), is a sufficient condition for the existence of:

(i) a point of period p for any positive integer p;

(ii) an uncountable, invariant set S, called the *scrambled set*, whose elements are aperiodic under f, and which satisfies

$$\limsup_{k\to\infty} |f^k(x) - f^k(y)| > 0$$

for all $x \in S$, and $y \in S$, $y \neq x$, or y any periodic point of f; and

(iii) an uncountable subset S_0 of S such that for any $x, y \in S$

$$\liminf_{k\to\infty} |f^k(x) - f^k(y)| = 0.$$

The inferences from their theorem upon the dymanics of the problem (1.1) are significant. Once a 3-cycle, or a $3n$-cycle for some positive integer n, is demonstrated, very erratic behavior can be expected for solutions of (1, 1). This includes the lack of global stability in the domain of f for any solution, and the existence of an uncountable collection of orbits which do not eventually approach any periodic pattern as $k \to \infty$. This phenomenon, which has been appropriately termed "chaos", has been observed in a wide variety of problems that are

0022-247X/78/0631-0199$02.00/0

Framed in terms of difference equations, and has thus drawn considerable interest from investigators in several mathematically related fields.

The principal impetus for the investigations that led to the theorem of [1] was provided by an interesting study conducted by Lorenz upon the solutions of a system of equations governing the convection of fluid flow in three dimensions. In [2] the original system of partial differential equations was first converted with certain simplifications into a set of ordinary differential equations, and solutions, obtained by a particular numerical integration scheme, were plotted in three-dimensional phase space. Upon recognizing the oscillatory nature of the resulting trajectories $(x(t), y(t), z(t))$, Lorenz then transformed the problem into a one-dimensional difference equation of the form (1.1) by a very clever technique, similar to a Poincare map. Letting M_k denote the maximum of the coordinate function $z(t)$ on the kth circuit of a particular trajectory around either of two equilibrium points, Lorenz noted that a scalar function f is implicitly defined by $f: M_k \to M_{k+1}$. The sequence $\{M_k\}_{k=0}^{\infty}$, it was observed, exhibits aperiodic (chaotic) behavior. In addition, although f cannot be determined explicitly, several qualitative characteristics of this function can be seen. With a furthei investigation of real valued functions that possess these same characteristics, Li and Yorke were thus able to establish rigorously the existence of the chaotic behavior first observed by Lorenz.

Although chaos was originally observed in the context of a hydrodynamical system, this phenomenon has spurred the interest primarily of mathematical biologists, particularly those in the field of population dynamics. There are basically two types of deterministic models which describe population growth in a network of n interacting species. One such model applies to networks in which population growth is a continuous process, and involves a system of ordinary differential equations. The other model governs networks of species in which the populations change only at discrete time intervals and successive generations do not overlap. This latter model can be conveniently formulated in terms of the first-order n-dimensional difference scheme:

$$X_{k+1} = F(X_k), \qquad k = 0, 1, \ldots, \tag{1.2}$$

where $X_k \in \mathbb{R}^n$ and $F: \mathbb{R}^n \to \mathbb{R}^n$ is continuous but usually nonlinear. The possibility of chaos for systems such as (1.1), a special case of (1.2), is therefore of significance to population biologists. As May suggests in [3], by virtue of this phenomenon biological processes which are described completely in terms of deterministic models in which all parameters are known may nevertheless exhibit essentially random behavior. Hence, in this manner a completely deterministic system may be effectively non-predictive.

Motivated by these considerations, a number of researchers have investigated the system (1.2) for choices of the function F appropriate to biological problems. It has been discovered that even for models in which $n = 1$, which reduces (1.2) to (1.1), and for relatively simple functions F, a wide variety of behavior can result. Two single-species models commonly accepted as discrete analogues of the logistic equation for continuous growth of a single population are:

$$N_{k+1} = N_k(1 + r(1 - N_k/C)) \tag{1.3a}$$

and

$$N_{k+1} = N_k \exp(r(1 - N_k/C)) \tag{1.3b}$$

where $r, C \in \mathbb{R}^+$. (May gives a detailed discussion of these equations in [4].) For $r < 2$ it can be easily verified that there exists a nonnegative equilibrium point in both models. However as the parameter r is increased beyond this value, May has shown numerically that the non-trivial fixed point in either model bifurcates into a stable two-point cycle. Also, in each model the stable 2-cycle eventually becomes unstable and bifurcates into a stable 4-cycle as r is further increased. Continuing in this manner, stable 2^k-cycles are produced for larger values of r and k until some critical value r^* for the parameter r is achieved. (For model (1.3a) May estimates $r^* \approx 2.570$ and for (1.3b) $r^* \approx 2.692$.) Beyond their respective values for r^*, both models possess n-point cycles for arbitrary integers n, and other solutions which are not even asymptotically periodic.

Because (1.3a) and (1.3b) are in the form (1.1), it is possible to apply the result of [1] to establish analytically the chaotic behavior of these equations for values of r exceeding the respective estimates of r^*. It can be shown that for any r greater than the critical value each equation possesses a $3n$-cycle for some positive integer n, thus implying that f^n has a point of period 3 in its domain, a condition sufficient to infer chaos for these equations. Therefore, the numerical results which May obtains for these parameter values are to be expected.

Chaos for difference schemes governing discrete population growth is by no means restricted to single-species models of the form (1.1). In [5] Guckenheimer, Oster and Ipaktchi consider the two-dimensional Leslie model:

$$\begin{aligned} x_{k+1} &= (b_1 x_k + b_2 y_k)\exp(-a(x_k + y_k)), \\ y_{k+1} &= s x_k, \end{aligned} \tag{1.4}$$

where $a, b_1, b_2, s \in \mathbb{R}^+$. They demonstrate that for certain values of these parameters solutions of (1.4) describe either stable "continuous curves" when plotted in the (x_k, y_k)-phase plane, or stable n-cycles which bifurcate into stable $2n$-cycles, etc., as these parameters are varied within specific limits. However, as with (1.3a) and (1.3b) there are other values of these parameters for which orbits of (1.4) appear to be chaotic. The same qualitative behavior is evident within the "host-parasite" scheme:

$$\begin{aligned} H_{k+1} &= H_k \exp(r(1 - H_k/N) - aP_k), \\ P_{k+1} &= bH_k(1 - \exp(-aP_k)), \end{aligned} \tag{1.5}$$

where r, a, b, $N \in \mathbb{R}^+$, which has been investigated by Beddington, Free and Lawton in [6].

Hence, the appearance of chaos in biological models framed by difference schemes of the form (1.2) is well documented. It is therefore natural to attempt to provide criteria for chaotic behavior in the multi-dimensional problem similar to those developed by Li and Yorke for a single equation. It is known that the existence of a 3-cycle for n arbitrary equations is not sufficient condition to infer such non-predictive behavior for the scheme. In fact the system (1.4) possesses for cerain choices of the parameters 3-cycles which appear numerically to be globally stable. (See [5].) The question remains then under what conditions will the general problem (1.2) behave chaotically. The analysis in [5], reminiscent of Smale's "horseshoe" argument [7], presents a partial answer to this question by establishing the existence of chaos for the two-dimensional difference scheme (1.4) for certain choices of the parameters. It is clear, however, that chaos can occur in the solutions of even two-dimensional schemes for which this type of analysis does not apply, e.g., (1.4) and (1.5) for certain parameter values.

The present paper constitutes an alternate answer to this question by providing sufficient criteria for chaos in an arbitrary n-dimensional problem. Roughly, we will show that the existence of a trajectory of (1.2) which begins arbitrarily close to an unstable fixed point of the function F, is "repelled" from this point as k increases, but suddenly "snaps back" to hit this point precisely, is sufficient to imply chaos of (1.2). This result will be developed in subsequent sections, and then applied to several problems of current interest, some having been previously described.

2. Notation and Definitions

We shall be using the following notation throughout the paper. For any function F: $\mathbb{R}^n \to \mathbb{R}^n$ and any positive integer k, let F^k represent the composition of F with itself k times, and F^{-k} the composition of the inverse of F (if it exists) with itself k times. The notions of periodic points and p-cycles of (1.2) remain the same as those previously described for the one-dimensional case. For a differentiable function F, let $DF(X)$ denote the Jacobian of F evaluated at the point $X \in \mathbb{R}^n$, and $|DF(X)|$ its determinent. Also, let $B_r(X)$ denote the closed ball in $\mathbb{R}^n$ of radius r centered at the point X, and $B_r^0(X)$ its interior. Finally, let $\|X\|$ be the usual Euclidean norm of X in $\mathbb{R}^n$.

Before stating conditions for the chaotic behavior of (1.2) it will be convenient to first present several preliminary definitions.

If V is a subset of $\mathbb{R}^n$ which is contained in the domain of F and which satisfies $F[V] \subset V$, then it is easily seen that choosing an initial point $X_0 \in V$ will guarantee a unique trajectory $\{X_k\}_{k=0}^{\infty}$ of (1.2). Under certain conditions, however, it is also possible to define a unique collection of points $\{X_k\}$ for

$-\infty < k \leqslant 0$ by iterating (1.2) "backwards". One such condition is the existence of a set $U \subset \mathbb{R}^n$ which satisfies:

$$F \text{ is 1-1 in } U \qquad \text{and} \qquad U \subset F[U]. \tag{2.1}$$

With this condition, although F may not be 1-1 everywhere in its domain, if $X \in U$ then $X \in F[U]$ and there exists a point $Y \in U$ with $F(Y) = X$. Thus we can make the following definition.

DEFINITION 2.1. For U and F satisfying (2.1) the *inverse of F in U*, denoted by $F^{-1} = F_U^{-1}$ is the function assigning to each $X \in U$ the unique $Y \in U$ with $F(Y) = X$.

In this case if we choose $X_0 \in U$ then we can obtain $\{X_k\}$ for all negative integers k via:

$$X_{k-1} = F^{-1}(X_k), \qquad k = 0, -1, \ldots, \tag{2.2}$$

where F^{-1} is the inverse of F in U.

Suppose now that the function F is differentiable in a set V and that $Z \in V$ is a fixed point of F. If all eigenvalues of $DF(Z)$ are greater than 1 in norm, then F displays the following local behavior at Z. For some $s > 1$ and $r > 0$:

$$\| F(X) - F(Y) \| > s \| X - Y \| \qquad \text{for all } X, Y \in B_r(Z). \tag{2.3}$$

This implies that F is 1-1 in $B_r(Z)$. Also, as a special case of (2.3), $\| F(X) - Z \| > \| X - Z \|$ for all $X \in B_r(Z)$. Since F is a homeomorphism in $B_r(Z)$, it must therefore be that $F[B_r(Z)] \supset B_r(Z)$. Thus F satisfies (2.1) with $U = B_r(Z)$, and so F^{-1} exists in $B_r(Z)$. But, (2.3) implies that $\| F^{-k}(X) - Z \| < s^{-k} \| X - Z \|$, and, consequently, $F^{-k}(X) \to Z$ as $k \to \infty$ for all $X \in B_r(Z)$. Hence, F^{-1} "contracts" $B_r(Z)$, or, conversely, F "expands" $B_r(Z)$. We formalize this terminology in the following.

DEFINITION 2.2. Let F be differentiable in $B_r(Z)$. The point $Z \in \mathbb{R}^n$ is an *expanding fixed point of F in $B_r(Z)$*, if $F(Z) = Z$ and all eigenvalues of $DF(X)$ exceed 1 in norm for all $X \in B_r(Z)$.

Finally, assume Z is an expanding fixed point of F in $B_r(Z)$. If F is not a 1-1 function in $\mathbb{R}^n$ then it is possible for there to exist a point $X_0 \in B_r(Z)$ with $X_0 \neq Z$ and $F^M(X_0) = Z$ for some positive integer M. Note that since F is 1-1 in $B_r(Z)$ and $F(Z) = Z$, we must have $M > 1$.

DEFINITION 2.3. Assume that Z is an expanding fixed point of F in $B_r(Z)$ for some $r > 0$. Then Z is said to be a *snap-back repeller* of F if there exists a point $X_0 \in B_r(Z)$ with $X_0 \neq Z$, $F^M(X_0) = Z$ and $| DF^M(X_0) | \neq 0$ for some positive integer M.

Note that since $X_0 \in B_r(Z)$, a collection of points $\{X_k\}_{k=-\infty}^{0}$ can be obtained by (2.2). In this case the set $\{X_k\}_{k=-\infty}^{+\infty}$ where $X_k = Z$ for all $k \geqslant M$ constitutes an orbit of (1.2) whose positive and negative limit sets consist of the point Z. Note also that the existence of a snap-back repeller is a stable property under small perturbations of F.

The condition $|DF(X_0)| \neq 0$ in the above definition provides for the existence of the inverse of F^M in a neighborhood of Z. The importance of this will become apparent in the next section. It will also be seen that functions possessing snap-back repellers exhibit a very complex type of behavior, i.e., chaos.

Remark 2.1. In [7] Smale demonstrates that for a conditionally stable fixed point of a diffeomorphism the assumption of a transverse homoclinic orbit implies the existence of an infinite number of periodic points of different periods. However, as described above, the definition of a snap-back repeller provides for the existence of a trajectory whose limit sets consist of an expanding fixed point of the function. In the next section it will be shown that this kind of "homoclinic" orbit also implies the existence of an infinite number of periodic points (and more). Thus the results presented here appear to be analogous to those of Smale for the case in which the fixed point is expanding rather than of saddle type.

Remark 2.2. It will also be seen in the next section that the assumption of a snap-back repeller Z of F provides for the existence of a sequence of compact sets $\{B_k\}_{k=-\infty}^{M}$ (homeomorphic to the unit ball in $\mathbb{R}^n$) which satisfy: (a) $B_k \to Z$ as $k \to -\infty$; (b) $F(B_k) = B_{k+1}$; (c) F is 1-1 in B_k; (d) $B_k \cap B_M = \varnothing$ for $1 \leqslant k < M$; and (e) $Z \in B_M{}^0$. It is the existence of such a sequence which in fact guarantees chaos for (1.2). Thus, for the sake of generality, the definition of a snap-back repeller could alternatively be: a fixed point Z for which such a sequence $\{B_k\}_{k=-\infty}^{M}$ exists. Under these weaker conditions the conclusions of Theorem 3.1 remain true with the assumption of continuity of F alone. In addition, for the special case $\mathbb{R}^n = \mathbb{R}$ an even weaker notion can be employed by dropping condition (c) above. (See Remark 3.1.)

3. Principal Result

In this section we establish our principal result. Let $F: \mathbb{R}^n \to \mathbb{R}^n$ be differentiable.

THEOREM 3.1. *If F possesses a snap-back repeller then* (1.2) *is chaotic. That is, there exists*

(i) *a positive integer N such that for each integer $p \geqslant N$, F has a point of period p;*

(ii) *a "scrambled set" of F, i.e., an uncountable set S containing no periodic points of F such that:*

(a) $F[S] \subset S$,

(b) *for every* $X, Y \in S$ *with* $X \neq Y$

$$\limsup_{k\to\infty} \|F^k(X) - F^k(Y)\| > 0,$$

(c) *for every* $X \in S$ *and any periodic point* Y *of* F

$$\limsup_{k\to\infty} \|F^k(X) - F^k(Y)\| > 0;$$

(iii) *an uncountable subset* S_0 *of* S *such that for every* $X, Y \in S_0$:

$$\liminf_{k\to\infty} \|F^k(X) - F^k(Y)\| = 0.$$

Before proving Theorem 3.1 we shall present two preliminary lemmas.

LEMMA 3.1. *Let* Z *be a snap-back repeller of* F. *Then for some* $s > 0$ *there exists* $Y_0 \in B_s{}^0(Z)$ *and an integer* L *such that* $F^k(Y_0) \notin B_s(Z)$ *for* $1 \leqslant k < L$ *and* $F^L(Y_0) = Z$. *Also,* $|DF^L(Y_0)| \neq 0$ *and* Z *is expanding in* $B_s(Z)$.

Proof. Since Z is a snap-back repeller, then for some $r > 0$ there exists $X_0 \in B_r(Z)$ with $X_0 \neq Z$, $F^M(X_0) = Z$ and $|DF^M(X_0)| \neq 0$. Let $X_k = F^k(X_0)$ for $k \geqslant 1$. Now since $0 < |DF^M(X_0)| \leqslant |DF^{M-k}(X_k)| \cdot |DF^k(X_0)|$, then:

$$|DF^k(X_0)| \neq 0 \qquad \text{for } 1 \leqslant k < M \tag{3.1a}$$

and

$$|DF^{M-k}(X_k)| \neq 0 \qquad \text{for } 1 \leqslant k < M. \tag{3.1b}$$

Since $X_0 \neq Z$ and $X_M = Z$, without loss of generality we can assume that $X_{M-1} \neq Z$. Otherwise, we could choose a smaller M such that this is true, and by (3.1a) we would still have $|DF^M(X_0)| \neq 0$ for this new M also.

Now, F is 1–1 in $B_r(Z)$ and $F(Z) = F(X_{M-1}) = Z$, but $X_{M-1} \neq Z$. So, we must have $X_{M-1} \notin B_r(Z)$. Also, since $X_0 \in B_r(Z)$ there must be an integer T with $0 \leqslant T < M$, and:

$$X_T \in B_r(Z) \qquad \text{and} \qquad X_{T+k} \notin B_r(Z) \qquad \text{for } 1 \leqslant k < M - T, \tag{3.2}$$

i.e., X_T is the last iterate of X_0 lying in $B_r(Z)$ before Z is hit precisely. Note that (3.1b) implies that $|DF^{M-T}(X_T)| \neq 0$.

Let $Y_0 = X_T$ and $L = M - T$. If $X_T \in B_r{}^0(Z)$ then letting $s = r$ will prove the lemma, so, assume $X_T \notin B_r{}^0(Z)$. But $X_T \in B_r(Z)$ implies $\|X_T\| = r$, and since Z is expanding in $B_r(Z)$ then by continuity of DF for some $\epsilon > 0$ Z is expanding in $B_w(Z)$ for all w satisfying $r < w \leqslant r + \epsilon$. Now (3.2) implies that for some w in the above interval $Y_0 = X_T \in B_w{}^0(Z)$ and $F^k(Y_0) = X_{T+k} \notin B_w(Z)$ for $1 \leqslant k < L$. Letting s be this w, Lemma 3.1 is proven.

LEMMA 3.2. *Let* $\{C_k\}_{k=0}^{\infty}$ *be a sequence of compact sets in* $\mathbb{R}^n$ *and let* $H: \mathbb{R}^n \to \mathbb{R}^n$ *be continuous. If* $H[C_k] \supset C_{k+1}$ *for all* $k \geqslant 0$, *then there exists a non-empty compact set* $C \subset C_0$ *with* $H^k(X) \in C_k$ *for all* $X \in C$ *and* $k \geqslant 0$.

This lemma is analogous to Lemma 1 in [1] and its proof is almost identical.

Proof of Theorem 3.1. Let Z be the snap-back repeller of F assumed by the theorem, and let $X_0 \in B_r(Z)$, $X_0 \neq Z$, $F^M(X_0) = Z$ and $|DF^M(X_0)| \neq 0$ where Z is expanding in $B_r(Z)$.

Proof of (i). Without loss of generality we can assume that:

$$X_0 \in B_r^0(Z) \qquad \text{and} \qquad F^k(X_0) \notin B_r(Z) \quad \text{for } 1 \leqslant k < M. \tag{3.3}$$

Otherwise, replacing X_0, r and M with the quantities Y_0, s and L respectively, provided by Lemma 3.1, will yield (3.3). The following analysis could then be carried out in terms of these new variables.

Now, since $F^M(X_0) = Z$ and $|DF^M(X_0)| \neq 0$, then for some $\epsilon > 0$ satisfying $0 < \epsilon < r$ there exists a continuous and 1–1 function G defined on $B_\epsilon(Z)$, with $G(Z) = X_0$ and:

$$G^{-1}(X) = F^M(X) \qquad \text{for all } X \in G[B_\epsilon(Z)]. \tag{3.4}$$

For notational convenience let Q be the compact set defined by $Q = G[B_\epsilon(Z)]$. Because of (3.3), we can assume without loss of generality that $X_0 \in Q \subset B_r(Z)$ and:

$$F^m[Q] \subset \mathbb{R}^n - B_r(Z) \qquad \text{for } 1 \leqslant m < M. \tag{3.5}$$

If not, then we could choose a smaller ϵ such that this is true.

Also, since Z is expanding in $B_r(Z)$ then F^{-1} exists in $B_r(Z)$ (see Definition 2.1), and thus $Q \subset B_r(Z)$ implies that:

$$F^{-m}[Q] \subset B_r(Z) \qquad \text{for } m \geqslant 0. \tag{3.6}$$

In addition, for any $X \in Q$, $F^{-k}(X) \to Z$ as $k \to \infty$, so there exists an integer $J = J(X) \geqslant 0$ such that $F^{-J}(X) \in B_\epsilon^0(Z)$. By continuity, therefore, we have $\delta = \delta(X) > 0$ with $F^{-J}[B_\delta^0(X)] \subset B_\epsilon(Z)$. Consider the collection of open sets: $D = \{B_\delta^0(X)\colon \text{for all } X \in Q\}$. The set D constitutes an open cover of the compact set Q, and thus a finite sub-collection D_0 of D also covers Q, where:

$$D_0 = \{B_\delta^0(X_i)\colon i = 1,\dots, L\}.$$

Letting

$$T = \max\{J(X_i)\colon i = 1,\dots, L\},$$

we have $F^{-T}(X) \in B_\epsilon(Z)$ for any $X \in Q$.

Since $\epsilon < r$, Z is also expanding in $B_\epsilon(Z)$, so:

$$F^{-k}[Q] \subset B_\epsilon(Z) \qquad \text{for all } k \geqslant T. \tag{3.7}$$

For each $k \geqslant T$ consider the function $F^{-k} \circ G$ defined for all $X \in B_\epsilon(Z)$. Since G is continuous (and 1–1) in $B_\epsilon(Z)$ and F^{-k} is continuous (and 1–1) in $G[B_\epsilon(Z)]$, then $F^{-k} \circ G$ is continuous (and (1–1) in $B_\epsilon(Z)$. So, from (3.7) and the definition of Q, $F^{-k} \circ G[B_\epsilon(Z)] \subset B_\epsilon(Z)$, and, consequently, $F^{-k} \circ G$ must have a fixed point $Y_k \in B_\epsilon(Z)$ by the Brouwer fixed point theorem. That is, $F^{-k} \circ G(Y_k) = Y_k$ for all $k \geqslant T$. Note, therefore, that $F^k(Y_k) = (F^k \circ F^{-k} \circ G)(Y_k) = G(Y_k)$. From (3.4) $F^{M+k}(Y_k) = F^M \circ G(Y_k) = G^{-1} \circ G(Y_k) = Y_k$, and Y_k is thus a fixed point of F^{M+k}.

We shall show that Y_k cannot have period less than $M + k$. From above $F^k(Y_k) = G(Y_k)$ and $Y_k \in B_\epsilon(Z)$. Hence:

$$F^k(Y_k) \in Q = G[B_\epsilon(Z)] \qquad \text{for all } k \geqslant T. \tag{3.8}$$

Taking F^{-k} of the point $F^k(Y_k)$ in (3.8) yields: $Y_k \in F^{-k}[Q]$. Letting $m = -n + k$ in (3.6), we thus obtain: $F^n(Y_k) \in F^{n-k}[Q] \subset B_r(Z)$ for all n satisfying $0 \leqslant n \leqslant k$. Also, (3.5) and (3.8) imply that $F^{m+k}(Y_k) \notin B_r(Z)$ for all m satisfying $1 \leqslant m < M$. So:

$$F^n(Y_k) \in B_r(Z) \qquad \text{for } 0 \leqslant n \leqslant k,$$
$$F^n(Y_k) \notin B_r(Z) \qquad \text{for } k + 1 \leqslant n < M + k,$$

and

$$F^{M+k}(Y_k) = Y_k .$$

It is clear, therefore, that Y_k cannot have period less than $M + k$. Hence, letting $N = M + T$ and $p = M + k$ for all $k \geqslant T$ proves (i).

Proof of (ii). Let the integers M, T and N be as in the proof of (i) and let U and V be the two compact sets defined by:

$$U = F^{M-1}[Q] \qquad \text{and} \qquad V = B_\epsilon(Z).$$

Claim 1. $U \cap V = \varnothing$.

Proof. From (3.5) $U = F^{M-1}[Q] \subset \mathbb{R}^n - B_r(Z)$. Since $\epsilon < r$, then $U \subset \mathbb{R}^n - B_\epsilon(Z)$. So, $V = B_\epsilon(Z)$ implies that $U \cap V = \varnothing$.

Claim 2. $V \subset F^N[U]$.

Proof. From the definition of U, $F[U] = F \circ F^{M-1}[Q] = F^M[Q]$. But, from (3.4) and the definition of Q, $F^M[Q] = F^M \circ G[B_\epsilon(Z)] = B_\epsilon(Z)$. So, $F[U] = B_\epsilon(Z)$, and hence $F^N[U] = F^{N-1}[F[U]] = F^{N-1}[B_\epsilon(Z)]$. Now since Z is expanding in $B_\epsilon(Z)$, $F^{N-1}[B_\epsilon(Z)] \supset B_\epsilon(Z)$, and therefore $F^N[U] \supset B_\epsilon(Z) = V$.

Claim 3. $U \subset F^N[V]$ and $V \subset F^N[V]$.

Proof. $F^N[V] = F^N[B_\epsilon(Z)] \supset B_\epsilon(Z)$ since Z is expanding in $B_\epsilon(Z)$. So, $F^N[V] \supset V$.

Also, letting $k = T + 1$ in (3.7) yields $F^{-T-1}[Q] \subset B_\epsilon(Z)$, and therefore $F^{N-T-1}[Q] \subset F^N[B_\epsilon(Z)] = F^N[V]$. But, $U = F^{M-1}[Q] = F^{N-T-1}[Q]$ and thus $U \subset F^N[V]$.

Now, let H be the function defined by: $H(X) = F^N(X)$ for all $X \in B_r(Z)$. We summerize the properties of U, V and H:

$$\inf\{\| X - Y \| : X \in U \text{ and } Y \in V\} > 0, \tag{3.9a}$$

$$V \subset H[U] \quad \text{and} \quad U, V \subset H[V]. \tag{3.9b}$$

The remainder of the proof of (ii) is essentially identical to that of the corresponding result in [1].

Let A be the set of sequences $E = \{E_n\}_{n=1}^{\infty}$ where E_n equals either U or V, and if $E_n = U$ then $E_{n+1} = E_{n+2} = V$. Let $R(E, n)$ be the number of E_i's which equal U for $1 \leqslant i \leqslant n$. For each $w \in (0, 1)$ choose $E^w = \{E_n^w\}_{n=1}^{\infty}$ to be a sequence in A satisfying:

$$\lim_{n \to \infty} \frac{R(E^w, n^2)}{n} = w.$$

If B is defined by: $B = \{E^w: w \in (0, 1)\} \subset A$, then B is uncountable. Also, from (3.9b) $H[E_n^w] \supset E_{n+1}^w$, and, therefore, by Lemma 3.2 for each $E^w \in B$ there is a point $X_w \in U \cup V$ with $H^n(X_w) \in E_n^w$ for all $n \geqslant 1$. Letting $S_H = \{H^n(X_n)$: $n \geqslant 0$ and $E^w \in B\}$, then $H[S_H] \subset S_H$, S_H contains no periodic points of H, and there exists an infinite number of n's such that $H^n(X) \in U$ and $H^n(Y) \in V$ for any $X, Y \in S_H$ with $X \neq Y$. (See [1].)

Now combining this last statement with (3.9a) implies that for any $X, Y \in S_H$ with $X \neq Y$:

$$L_1 = \limsup_{n \to \infty} \| H^n(X) - H^n(Y) \| > 0.$$

Therefore, letting $S = \{F^n(X): X \in S_H \text{ and } n \geqslant 0\}$ and recalling that $H(X) = F^N(X)$, we see that $F[S] \subset S$, S contains no periodic points of F, and for any $X, Y \in S$ with $X \neq Y$:

$$\limsup_{n \to \infty} \| F^n(X) - F^n(Y) \| \geqslant L_1 > 0.$$

We thus have (ii)a and (ii)b. In a similar manner (ii)c can be proven.

Proof of (iii). First note that since Z is expanding in $B_\epsilon(Z)$, if we define $D_n = H^{-n}[B_\epsilon(Z)]$ for all $n \geqslant 0$, then given $\delta > 0$ there exists $J = J(\delta)$ such that $\| X - Z \| < \delta$ for all $X \in D_n$ and $n > J$. The proof of (iii) again parallels the proof of the corresponding result in [1].

For any sequence $E^w = \{E_n^w\}_{n=1}^{\infty} \in A$ we shall further restrict the E_n^w in the following manner: if $E_n^w = U$ then $n = m^2$ for some integer m. Also, if $E_n^w = U$ for both $n = m^2$ and $n = (m+1)^2$ then $E_n^w = D_{2m-k}$ for $n = m^2 + k$ where $k = 1, \ldots, 2m$. For the remaining n's we shall assume $E_n^w = V$.

It can be easily checked that these sequences still satisfy $H[E_n^w] \supset E_{n+1}^w$, and thus by Lemma 3.2 there exists a point X_w with $H^n(X_w) \in E_n^w$ for all $n \geqslant 0$.

Now, defining $S_0 = \{X_w: w \in (\frac{4}{5}, 1)\}$ then S_0 is uncountable, $S_0 \subset S_H \subset S$ and for any $s, t \in (\frac{4}{5}, 1)$ there exist infinitely many m's such that $H^n(X_s) \in E_n{}^s = D_{2m-1}$ and $H^n(X_t) \in E_n{}^t = D_{2m-1}$ where $n = m^2 + 1$. But from above, given $\delta > 0$, $\| X - Z \| < \delta/2$ for all $X \in D_{2m-1}$ and m sufficiently large. Thus, for all $\delta > 0$ there exists an integer m such that $\| H^n(X_s) - H^n(X_t)\| < \delta$ where $n = m^2 + 1$. Since δ is arbitrary we have:

$$L_2 = \liminf_{n\to\infty} \| H^n(X_s) - H^n(X_t)\| = 0.$$

Therefore, for any $X, Y \in S_0$:

$$\liminf_{n\to\infty} \| F^n(X) - F^n(Y)\| \leqslant L_2 = 0$$

and (iii) is proven.

Remark 3.1. As can be seen from the proof of (i), a sequence of sets $\{B_k\}_{k=-\infty}^{M}$ was constructed, satisfying conditions (a) through (e) of Remark 2.2. In fact, $B_0 = Q$ and $B_M = B_\epsilon(Z)$. It is evident, therefore, that the existence of just such a sequence is all that is required to imply the conclusions of Theorem 3.1. Also, for the special case $\mathbb{R}^n = \mathbb{R}$, in which the B_k's are now closed intervals, it can be seen that assumption (c) of Remark 2.2 can be dropped. The reason for this lies in the fact that for any closed interval I and any continuous function $h: \mathbb{R} \to \mathbb{R}$, if $h[I] \supset I$ then $h(x_0) = x_0$ for some $x_0 \in I$. With this lemma in the proof of Theorem 3.1(i), the functions F^{M+k} need not be inverted and the Brouwer theorem need not be employed in order to demonstrate the existence of the periodic points. Thus, for the one-dimensional case we have this even weaker notion of a snap-back repeller. In fact, it can be shown that in this case the existence of a snap-back repeller of f is equivalent to the existence of a point of period 3 for the function f^n for some positive integer n. (See [8].)

4. Numerical Examples

Example 4.1. As previously described, May's numerical studies have demonstrated the chaotic behavior of solutions of (1.3a) for $r > r^* \approx 2.570$, and upon application of the result of [1] to the problem, this can be established analytically. However, in light of the previous section there is now an alternate method for verifying this behavior, i.e., we can apply Theorem 3.1 to equation (1.3a).

First note that (1.3a) can be written in the equivalent form:

$$x_{k+1} = ax_k(1 - x_k) \tag{4.1}$$

where $a = r + 1$ and $x_k = rN_k/(r + 1)\, C$. The dynamics of (4.1) are identical to those of (1.3a) and thus we should expect to observe chaos for $a > 3.570$.

Computing the non-trivial fixed point z of (4.1) yields $z = (a - 1)/a$. To apply Theorem 3.1 we must demonstrate that z is a snap-back repeller, and thus we first need to find an interval $I = [z - r, z + r]$ with $|f'(x)| > 1$ for all $x \in I$ where $f(x) = ax(1 - x)$. In particular, since $f'(z) < 0$ for $a > 3$, we must have $f'(x) < -1$ for all $x \in I$. From the shape of the function f, we know that the right hand endpoint of I is arbitrary, i.e., $f'(x) < f'(z) < -1$ for $x > z$ and $a > 3$. To estimate an acceptable left hand endpoint of I, note that $f'(z - \epsilon) = 2 - a + 2a\epsilon$, and so, $f'(z - \epsilon) < -1$ for $\epsilon < (a - 3)/2a$. If we restrict our discussion to values of $a > 3.5$, then $f'(z - \epsilon) < -1$ for all $\epsilon < 1/14 \approx .07$. Thus, letting $I = [z - .01, z - .01]$ is more than sufficient to insure that $f'(x) < -1$ for all $x \in I$ and $a > 3.5$.

Now we must find a point $z_0 \in I$ with $F^M(z_0) = z$, $z_0 \neq z$, and $(F^M)'(z_0) \neq 0$ for some positive integer M. Computing the "pre-images" of z, i.e., the points which are eventually mapped onto z under f, is not difficult for this function. By simply iterating the multi-valued "inverse" of f:

$$x_{k-1} = \frac{a \pm (a^2 - 4ax_k)^{1/2}}{2a}, \qquad k = 0, -1, \ldots, \tag{4.2}$$

with initial point $x_0 = z$, we can find all such pre-image points. Perhaps the simplest method to find z_0 satisfying the above conditions, therefore, is the following: with $x_0 = z$, we have two possible choices for each x_k, one greater than $\frac{1}{2}$ and one less, corresponding respectively to the plus or minus sign in (4.2). For $k = 0$ choosing the positive root will yield $x_{-1} = z$, which does not help in finding a z_0 with $z_0 \neq z$. Therefore, choose the negative root in (4.2) for $k = 0$. Now since we wish to find a pre-image point close to z, i.e., inside I, the optimal choice of roots in (4.2) is the positive ones for all $k \leqslant -1$. If after M iterations of (4.2) with the selection of x_k's made as described, we find that $x_{-M} \in I$ and $f'(x_{-k}) \neq 0$ for $1 \leqslant k \leqslant M$, then lettlng $z_0 = x_{-M}$ will satisfy the hypotheses of Theorem 3.1. This is evident since: $f^M(z_0) = f^M(x_{-M}) = x_0 = z$. Also, $(f^M)'(z_0) = (f^M)'(x_{-M}) = \prod_{k=1}^{M} f'(x_{-k})$, and so, $f'(x_{-k}) \neq 0$ implies that $(f^M)'(z_0) \neq 0$.

A numerical study was performed upon the inverse function (4.2) with $x_0 = z$ to find a point $z_0 = x_{-M}$ satisfying the above conditions. The selection of x_k's was made in the manner described. The results show that for all values of $a > 3.680$ a pre-image point of z lies within the interval I. In fact, $x_{-10} \in I$. In addition, none of the points x_k for $-10 \leqslant k \leqslant 0$ equals $\frac{1}{2}$, the only point x at which $f'(x) = 0$. Hence, z is a snap-back repeller and Theorem 3.1 guarantees chaos for all $a > 3.680$. For values of a less than this number it was found that in each case x_{-2} lies outside the range of f, and thus, x_{-3} is complex valued.

Note that for each iteration of (4.2) the selection of roots descibed above for x_k seems to be the most advantageous. If another sequence is generated by choosing the roots differently, it is not believed that the region of a values which allow x_{-M} to lie within I for some positive integer M will be extended below 3.680. In fact, it is more likely not even to extend the region of values this far.

Now let us perform this same type of analysis upon $g(x) = f^2(x) = f(f(x))$ It can be easily shown that the fixed points of g that are not fixed points of f are given by:

$$z_{1,2} = \frac{(a+1) \pm (a^2 - 2a - 3)^{1/2}}{2a}$$

where $f(z_i) = z_j$ for $i, j = 1, 2$, and for $a > 3$. Also, $g'(z_i) = f'(z_1)f'(z_2) < -1.25$ for $a > 3.50$. We can find an interval $J = [z_2 - r, z_2 + r]$, where z_2 is the greater of the two roots, in the following manner: for any ϵ, positive or negative, by expanding $g'(z_2 + \epsilon)$ in a Taylor series, it can be shown that:

$$g'(z_2 + \epsilon) = g'(z_2) + 2a\epsilon(f'(z_1) + f'(z_2)f'(y)) + 4a^2\epsilon^2 f'(y)$$

for some y satisfying $|z_2 - y| < \epsilon$. Now since $|f'(x)| < a$ for all $x \in (0, 1)$, then for $z_2 + \epsilon$ inside $(0, 1)$ we have: $g'(z_2 + \epsilon) \leqslant g'(z_2) + 8a^3 |\epsilon|$. Thus for $a > 3.50$, $g'(z_2 + \epsilon) \leqslant -1.25 + 8a^3 |\epsilon|$, and $g'(z_2 + \epsilon) < -1$ for $|\epsilon| \leqslant 1/32a^3$. So, choosing $r = 10^{-4}$ insures that $g'(x) < -1$ for all $x \in J$ and $3.50 < a < 4.00$.

Now letting $x_0 = z_2$, the optimal choice for the sequence $\{x_k\}$ in (4.2) was found to be the following: let x_{-1} have the positive root in (4.2), and for all $k \leqslant -2$ let the root have sign equal to $\operatorname{sgn}(-1)^k$. Numerically iterating (4.2) under these conditions, it was discovered that for all $a > 3.595$, $x_{-26} \in J$ and again $x_k \neq \frac{1}{2}$ for $-26 \leqslant k \leqslant -1$. Hence $g^{13}(x_{-26}) = f^{26}(x_{-26}) = z_2$ and $(g^{13})'(x_{-26}) \neq 0$. Thus we have a snap-back repeller of $g = f^2$ and Theorem 3.1 predicts chaos for $a > 3.595$.

In summary then, it was shown that by applying Theorem 3.1 to f, chaos is established for $a > 3.680$, and by applying it to f^2, for $a > 3.595$. It is likely that if we continue this process of investigating the pre-images of f^{2^k} for larger values of k, we shall approach the value $a \approx 3.570$, which separates stability from chaos of (4.1).

EXAMPLE 4.2. The other single-species model which May treats, (1.3b), can be equivalently written:

$$x_{k+1} = ax_k \exp(-x_k) \tag{4.3}$$

where $a = \exp(r)$ and $x_k = rN_k/C$. Investigating (4.3) in the same manner as (4.1) reveals that the non-trivial fixed point of $f(x) = (ax)\exp(-x)$ is given by $z = \ln(a)$, and an interval I with $z \in I$ and in which $f'(x) < -1$ for $x \in I$ is $I = [z - .01, z + .01]$ for $a > \exp(2.5)$. Since the multi-valued inverse of (4.3) cannot be written explicitly, pre-images of $x_0 = z$ can only be estimated numerically. For each $k \leqslant 0$, 35 iterations of the Newton method were performed to find a root of $f(x) - x_k = 0$, where x_k is known each time, thus producing an estimate for $x = x_{k-1}$. (For each k successive estimates of x_{k-1} were found to differ by less than 10^{-8} after 35 iterations of the algorithm.) Since there are again two possible values for each x_k, the manner of choosing them was similar to that of the previous example: for x_{-1} the root less than the x value producing a maximum for $f(x)$, i.e., less than 1, was selected, and for all $k \leqslant -1$, the root greater than 1 was chosen. In this case the point $x_{-12} \in I$ with $f^{12}(x_{-12}) = z$ and $f'(x_k) \neq 0$ for $-12 \leqslant k \leqslant -1$, was found for all $a > 16.999$ ($r > 2.833$ for (1.3b)).

Similarly treating the fixed points of f^2 as in Example 4.1 and applying Theorem 3.1 reveals that chaos should occur for $a > 15.250$ ($r > 2.724$). The value of r which May estimated is the dividing line between stability and chaos is $r = r^* \approx 2.692$. For (4.3), therefore, the value should be $a = \exp(r^*) \approx 14.765$. Again it is likely that finding snap-back repellers of f^{2^k} for larger integers k will yield chaos for a even closer to this estimated value.

EXAMPLE 4.3. We shall now attempt to apply Theorem 3.1 to several problems of the form (1.2) where $F: \mathbb{R}^2 \to \mathbb{R}^2$. To illustrate the technique involved, first consider the following two-dimensional generalization of the problem (4.1):

$$\begin{aligned} x_{k+1} &= (ax_k + by_k)(1 - ax_k - by_k) \\ y_{k+1} &= x_k . \end{aligned} \tag{4.4}$$

This problem possesses no special biological significance, but was selected for investigation since it can be reduced to (4.1) when $b = 0$.

Since we are primarily interested in only the positive solutions of (4.4), we shall begin by restricting the parameters a and b in the following manner: let these parameters lie in the region R of the (a, b)-plane described by $R = \{(a, b): a, b \geqslant 0 \text{ and } a + b \leqslant 4\}$. Under these conditions the set $D = \{(x, y): 0 \leqslant x, y \leqslant \frac{1}{4}\}$ is invarient under F. In order to justify an application of Theorem 3.1, let us first examine the qualitative behavior of (4.4) for $(a, b) \in R$.

The local dynamics of difference schemes in a neighborhood of an equilibrium are dependent upon the Jacobian of the function involved. Computing the two fixed points of F, we find the trivial one: $x_k = y_k = 0$, and for $a + b > 1$ the positive fixed point: $x_k = y_k = (a + b - 1)/(a + b)^2$. Also, a simple calculation shows that:

$$DF(x, y) = \begin{bmatrix} a - 2a(ax + by) & b - 2b(ax - by) \\ 1 & 0 \end{bmatrix}$$

To compute the eigenvalues λ_1, λ_2 of F at a point (x, y), therefore, we let $|DF(x, y) - \lambda I| = 0$ to obtain:

$$\lambda^2 - \lambda(a - 2a(ax + by)) - (b - 2b(ax + by)) = 0. \tag{4.5}$$

Evaluating (4.5) at $x = y = 0$, we see that for $a + b < 1$, $|\lambda_1|, |\lambda_2| < 1$ and thus $(0, 0)$ is stable in the region $R_1 = \{(a, b): a, b \geqslant 0, a + b < 1\}$. However, leaving the region R_1 across the line $a + b = 1$, one eigenvalue becomes greater than 1 making $(0, 0)$ unstable. Simultaneous with this is the appearance of the non-trivial fixed point $Z = (z, z)$ where $z = (a + b - 1)/(a + b)^2$, whose eigenvalues by (4.5) satisfy:

$$\lambda^2 + A\lambda + B = 0 \tag{4.6a}$$

where

$$A = \frac{a(a + b - 2)}{a + b} \quad \text{and} \quad B = \frac{b(a + b - 2)}{a + b}. \tag{4.6b}$$

Note that the simultaneous occurrence of an eigenvalue becoming greater than 1 with the appearance of a new fixed point is not arbitrary. In [5] this would be classified as a type (b) bifurcation of $(0, 0)$.

Solving (4.6a), it is not difficult to check that Z is stable for values of (a, b) close to the line $a + b = 1$. However, moving away from this line, there are two ways in which Z is likely to become unstable: (i) when both eigenvalues are real and one of them exceeds 1 in norm, while the other remains less, and (ii) when both eigenvalues, being complex conjugates, have norm greater than 1.

For case (i) we can find the curve in the (a, b)-plane along which both eigenvalues are real and one equals 1 in absolute value. Letting the solutions of (4.6a) equal ± 1 yields $B \pm A + 1 = 0$. Substituting the values of A and B given by (4.6b) implies either $a + b = 1$ or $b^2 - a^2 + 3a - b = 0$. The dynamics across $a + b = 1$ have already been discussed. The latter path, however, separates stability of Z from instability. The behavior across this curve will be discussed below.

For case (ii) note that if the solutions of (4.6a) are complex and equal 1 in norm then $B = 1$, and thus the path described by: $b^2 + (a - 3)b - a = 0$ also separates stability of Z from instability. Combining this with the result of case (i),

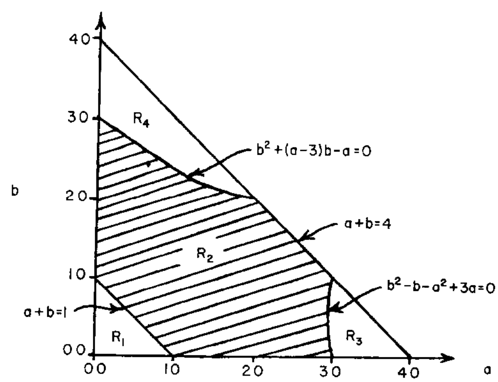

FIGURE 1

we see that Z is locally stable in the sub-region R_2 of R pictured in Fig. 1. In fact, numerical studies reveal that Z is globally stable in R_2 for $(x_0, y_0) \in D$.

We shall now investigate the dynamics of (4.4) for (a, b) in either R_3 or R_4 as pictured in Fig. 1. In the former-case crossing from R_2 into R_3, one eigenvalue of $DF(Z)$ passes from greater to less than -1. Numerical evidence suggests a type (a) bifurcation of Z, as described in [5], which doubles the period of an orbit, as a stable 2-cycle appears. Moving further to the right in R_3, the stable 2-cycle itself becomes unstable and a bifurcation into a stable 4-cycle occurs. Passing in this way through R_3, we observe successive bifurcation of 2^k-cycles into 2^{k+1}-cycles until we enter a sub-region of R_3 for which solutions of (4.4) are chaotic. Fig. 2 indicates the dynamics of (4.4) for $(a, b) \in R_3$. Note that for

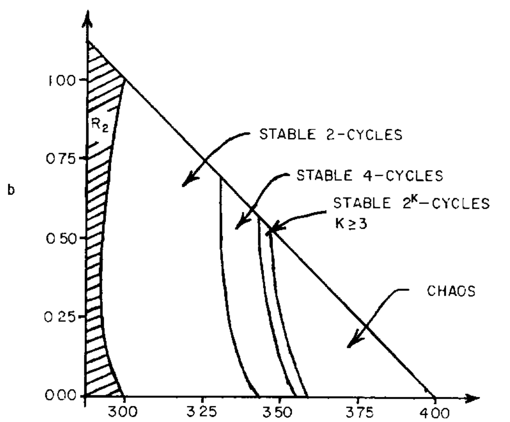

FIGURE 2

$b = 0$ the qualitative behavior of (4.4) conforms to that which May discovered for (1.3a).

Also note that it is not possible to explain the form of chaos which appears here by Theorem 3.1. This is evident by recalling that Z is unstable here only in the sense that one eigenvalue exceeds 1 in absolute value. Thus Z cannot be a snap-back repeller of F. The same is true of the fixed points of F^{2^k} for $k \geqslant 1$. Only one eigenvalue of DF^{2^k} at a point of the 2^k-cycle exceeds 1 in absolute value, and so again, Theorem 3.1 cannot be applied. We remark that the type of chaos that appears in R_3 can be explained by the same "twisted horseshoe" argument due to Guckenheimer, Oster and Ipaktchi [5].

Now let us examine (4.4) for $(a, b) \in R_4$. As previously described, the eigenvalues of $DF(Z)$ are complex and equal 1 in norm along the path separating R_2 from R_4. Thus, according to [5], we might expect a type (c) bifurcation in which a stable continuous curve or an n-cycle for some n appears in the (x_k, y_k)-phase plane. As (a, b) crosses into R_4, the eigenvalues of $DF(Z)$ cross the unit circle in the complex plane at certain angles with respect the positive real axis (one angle being the negative of the other). If these angles are irrational multiples of π, then a stable continuous curve will appear around Z, and if they are rational multiples, a stable n-cycle for some n.

Numerical indications are that this does occur, and, in addition, as (a, b) moves further upward, the "radius" of these curves or cycles around Z increases. The visual shape of these trajectories also changes as we move deeper into R_4 in the manner described in [6]. At first the curves and cycles possess well formed circular shapes, but moving further into R_4, although still remaining stable, they gradually develop several "kinked" areas. Figs. 3a and 3b illustrate the deformation a typical curve undergoes as (a, b) moves deeper into R_4.

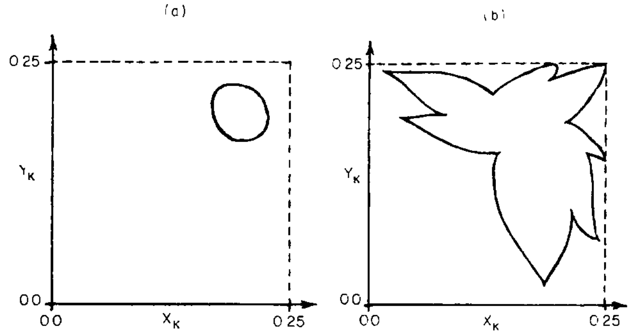

FIGURE 3

As might be expected, the stability of these curves and cycles vanishes and chaos appears, if (a, b) is moved far enough into R_4. It is interesting to note that chaos here differs visually from that encountered in R_3. In the latter case when (x_k, y_k) were plotted in phase plane, the result appeared to be one-dimensional chaos. That is, there exists a curve in the region D to which solutions of (4.4) tend as $k \to \infty$. However, on this curve there is chaotic behavior. Figure 4a depicts a typical pattern that is described for (a, b) in the chaotic

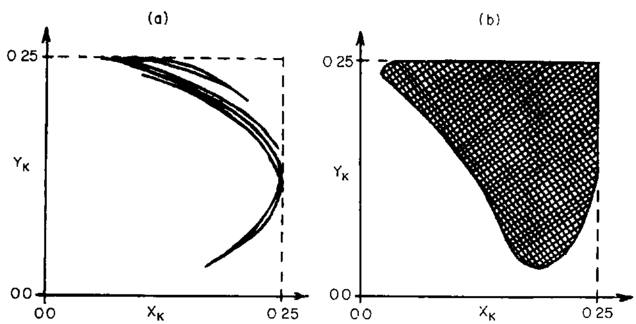

FIGURE 4

sub-region of R_3. Notice its similarity to the "strange attractor" treated in [5]. On the other hand, for (a, b) in the chaotic sub-region of R_4 the iterates (x_k, y_k) tend to fill out an entire two-dimensional subset of D as $k \to \infty$. Figure 4b typifies this kind of chaos.

The qualitative behavior of (4.4) for $(a, b) \in R_4$ is shown in Fig. 5. As can be seen from this, in addition to the dynamics thus described, there is a sub-region

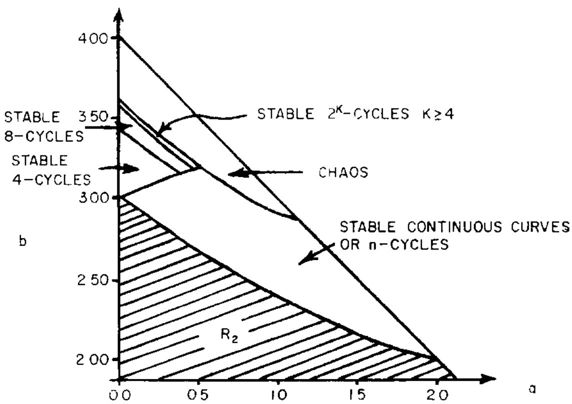

FIGURE 5

of R_4 in which the equation possesses stable 2^k-cycles for $k \geqslant 2$. Moving upward in this region, successive bifurcations (type (a)) of the 2^k-cycles into 2^{k+1}-cycles occur, until we again encounter chaos.

We shall now investigate this chaotic sub-region of R_4. Since both eigenvalues of $DF(Z)$ exceed 1 in norm, there exists a neighborhood of Z in which this point is expanding and thus there is a possibility of using Theorem 3.1 to establish rigorously the chaotic behavior observed here. As was the case with Examples 4.1 and 4.2, we need to demonstrate the existence of a pre-image point (x_{-M}, y_{-M}) of Z, not equal to but lying in a neighborhood of Z in which the eigenvalues of DF exceed 1 in norm, and $| DF^M(x_{-M}, y_{-M})| \neq 0$. Also as before, we shall first provide an acceptable neighborhood $B_r(Z)$ for all (a, b) of interest.

Let $x = z + \epsilon$ and $y = z + \delta$. We shall show that the eigenvalues of $DF(x, y)$ are complex valued and exceed 1 in norm for all ϵ and δ sufficiently small. From (4.5) the eigenvalues of $DF(x, y)$ satisfy:

$$\lambda^2 + A_1\lambda + B_1 = 0 \tag{4.7}$$

where $A_1 = A + 2a(a\epsilon + b\delta)$ and $B_1 = B + 2b(a\epsilon + b\delta)$, and A and B are given by (4.6b). It can be estimated numerically that the imaginary parts of the eigenvalues of $DF(Z)$ exceed 1 in absolute value for $(a, b) \in R_4$.

Now, since $a + b \leqslant 4$, then:

$$A_1^2 - 4B_1 < A^2 - 4B + 192(\epsilon + \delta) + 256(\epsilon + \delta)^2$$
$$< -1 + 192(\epsilon + \delta) + 256(\epsilon + \delta)^2.$$

Therefore, choosing both $|\epsilon|$ and $|\delta|$ to be less than 10^{-3} is more than sufficient to insure that the roots of (4.7) are imaginary. Thus, their norms are given by: $\|\lambda_i\| = B_1 = B + 2b(a\epsilon + b\delta)$. Now, it is not difficult to check that throughout the region of observed chaos in R_4, the norms of the eigenvalues of $DF(Z)$, which equal B, exceed 1.2 (e.g., one can solve $B = B(a, b) \geqslant 1.2$). Therefore, $\|\lambda_i\| > 1.2 - 32(\epsilon + \delta)$. Thus, choosing $|\epsilon|$ and $|\delta|$ to be less than 10^{-3} insures that the norms of the eigenvalues of $DF(x, y)$ exceed 1. So, we may take $B_r(Z)$ with $r = 10^{-3}$ to be our neighborhood.

In a manner similar to Example 4.1 the pre-images of Z can be found by letting $x_0 = y_0 = (a + b - 1)/(a + b)^2$ and inverting the function under consideration. In this case, the multi-valued inverse of (4.4) is given explicitly by:

$$x_{k-1} = y_k,$$
$$y_{k-1} = \frac{1 - 2ay_k \pm (1 - 4x_k)^{1/2}}{2b}. \tag{4.8}$$

Again, as in previous examples, selecting the positive root for y_{-1} does not help to find an appropriate pre-image point of Z. Therefore, we shall again choose the negative root for y_{-1}, and the positive roots for all $k \leqslant -1$. As before, this selection appears to be optimal.

Numerically iterating (4.8) for a wide collection of points $(a, b) \in R_4$, with the selection of y_k's made as described, reveals that the point (x_{-30}, y_{-30}) lies in $B_r(Z)$ with $r = 10^{-3}$ for those (a, b) values shown in Fig. 6. Also, the points

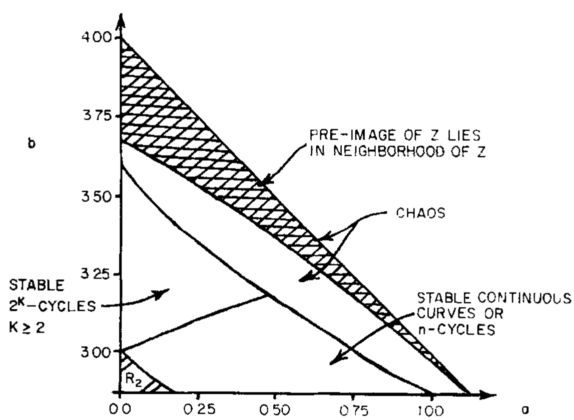

FIGURE 6

(x_k, y_k) satisfy $ax_k + by_k \neq \frac{1}{2}$ for $-30 \leqslant k \leqslant -1$. Since the line $ax + by = \frac{1}{2}$ describes the locus of points in D for which $|DF(x, y)| = 0$, we are thus guaranteed that $|DF(x_{-30}, y_{-30})| \neq 0$. Therefore, Z is a snap-back repeller here and Theorem 3.1 establishes chaos in this region of R_4.

This same type of analysis can be performed upon the fixed points of the functions F^{2^k} for $k \geqslant 2$. In addition to the region of stable 2^k-cycles in Fig. 5 and 6, there exist unstable 2^k-cycles in that part of R_4 shown in Fig. 7 (above the

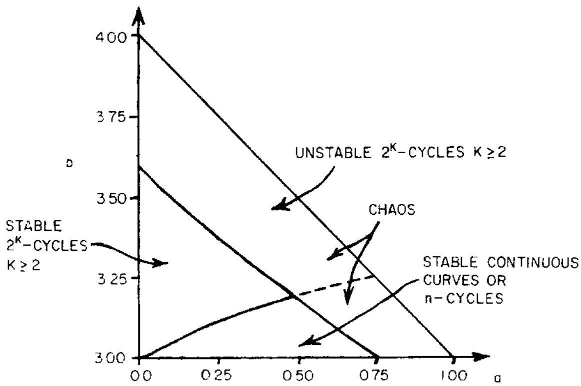

FIGURE 7

dotted line). In particular there are 4-cycles whose points are expanding under F^4. Thus there is again a possibility of applying Theorem 3.1, this time to F^4. Because of the complexity of the function involved, however, an ϵ, δ analysis of DF^4 close to a fixed point Z_4 of F^4 is virtually impossible. To investigate these points, therefore, the neighborhoods $B_r(Z_4)$ were numerically estimated for a large collection of (a, b) values covering the region of unstable 2^k-cycles. It was discovered that Z_4 is expanding under F^4 in $B_r(Z_4)$ with $r = 10^{-3}$ for all (a, b) of interest.

When graphed in the phase plane, the points of the 4-cycles are arranged roughly as the corners of a rectangle with sides parallel to the axes. Taking the element of the cycle closest to $(.25, 0)$ as the initial point (x_0, y_0), the optimal sequence of roots of (4.8) appears to be the following: let y_{-1} have the positive root, and for all $k \leqslant -2$, let y_k have the root with sign equal to $\operatorname{sgn}(-1)^m$, where m is the greatest integer less than or equal to $-(k + 2)/2$. With this selection Fig. 8 illustrates the region in which the point (x_{-40}, y_{-40}) lies in $B_r(Z_4)$ where $Z_4 = (x_0, y_0)$ and $r = 10^{-3}$. It is likely that an investigation of the functions F^{2^k} for larger k values would establish chaos for the entire region in which the unstable 2^k-cycles reside. However, in these cases selecting the appropriate sequences of roots in (4.8), which involves a great deal of guesswork, seems very unlikely.

This same type of analysis could also be performed upon the points of the

n-cycles in the chaotic sub-region of R_4 shown in Figs. 7 and 8 (below the dotted line). But, again the problems introduced in the investigation of F^n for large values of n seem insurmountable.

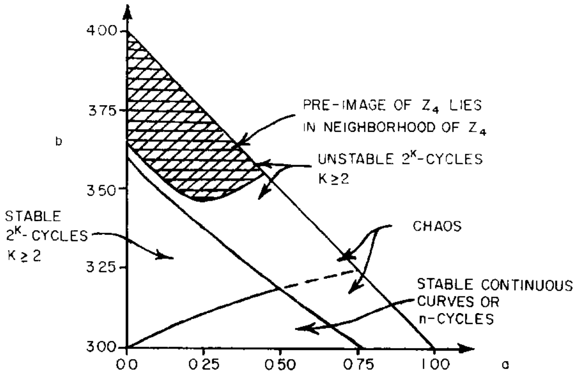

FIGURE 8

EXAMPLE 4.4. The two-dimensional difference scheme, (1.4), has been investigated in some depth by Guckenheimer, Oster and Ipaktchi. In particular they treat a special case of the equation by letting: $b_1 = b_2 = r$, $a = .1$ and $s = 1$. The resulting behavior is set out in [5]. Roughly, it is observed that, in a manner similar to the previous examples, as the parameter r is increased, the equation exhibits successive bifurcations from a stable equilibrium to a stable 3-cycle, from this 3-cycle to a stable 6-cycle, etc., until chaos is finally achieved. However, unlike previous cases, if r further increased, we observe the following pattern: regions of stability and successive bifurcations of $2^k n$-cycles into $2^{k+1}n$-cycles culminating in chaos, followed by a return of stability of $2^k n$-cycles for larger values of n. The nature of the chaos observed here, however, is similar to that obtained in the region R_3 of Example 4.3, and thus cannot be investigated with respect to Theorem 3.1. A justification of chaos for these parameter values is provided in [5] by an argument similar to Smale's "horseshoe" example [7].

Therefore, consider the following modification of (1.4):

$$\begin{aligned} x_{k+1} &= (ax_k + by_k)\exp(-ax_k - by_k) \\ y_{k+1} &= x_k \end{aligned} \tag{4.9}$$

for which the parameter values will be restricted to the region $Q = \{(a, b)$: $0 \leqslant a \leqslant 1.5, 5 \leqslant b \leqslant 25\}$. Motivation for consideration of the region Q lies in the similarity of the form of (4.9) to that of (4.4), for which Theorem 3.1 established chaos for (a, b) close to the b-axis. So again we shall look in this region.

In a manner similar to that of the previous problems, it can be shown that the non-trivial equilibrium point $Z = (z, z)$ is given by: $z = (\ln(a + b))/(a + b)$, and is (globally) stable in that part of Q lying below the curve described by: $a + b(2 - \ln(a + b)) = 0$, along which both eigenvalues of $DF(Z)$ complex and equal 1 in norm. Numerical evidence suggests that the part of Q lying above this path has regions of either stable continuous curves or n-cycles for arbitrary n, stable 2^k-cycles for $k \geqslant 2$, and chaos. See Fig. 9. Since both eigenvalues of $DF(Z)$ exceed 1 in norm, there is now a possibility of applying Theorem 3.1 to this chaotic sub-region of Q.

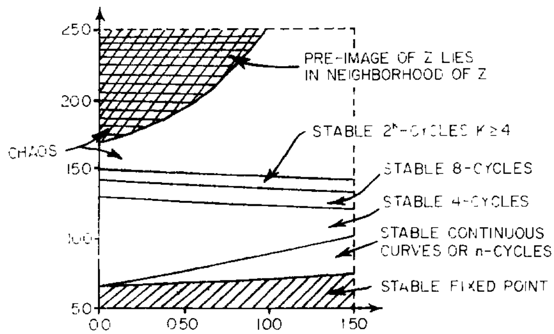

FIGURE 9

Performing an ϵ, δ analysis upon DF close to the fixed point Z of (4.9) reveals that throughout the chaotic region, the eigenvalues of $DF(z - \epsilon, z + \delta)$ are complex valued and exceed 1 in norm for $|\epsilon|, |\delta| < 10^{-3}$. Thus we may again choose $B_r(Z)$ with $r = 10^{-3}$ to be the neighborhood in which Z is expanding.

Because of the transcendental nature of (4.9), as in Example 4.2 we cannot invert our equation explicitly. Therefore, 70 iterations of the Newton method were performed for each k, to find a root of the equation: $(ax_k + by) \times \exp(-ax_k - by) - x_{k+1} = 0$, with $x_k = y_{k+1}$ and x_{k+1} known each time, thus producing an estimate for $y = y_k$. (70 iterations proved sufficient for successive estimates of y_k to differ by less than 10^{-8} for each k.) In this way, a sequence (x_k, y_k) was generated with the choice of roots determined, as in each of the previous cases, by the following method: with $x_0 = y_0 = (\ln(a + b))/(a + b)$, the smaller of the two roots was selected for y_{-1}, and thereafter the larger of the two for each y_k. Figure 9 illustrates the region for which the pre-image point (x_{-21}, y_{-21}) lies inside $B_r(Z)$ with $r = 10^{-3}$, and for which $ax_l + by_k \neq 1$ for $-21 \leqslant k \leqslant -1$, thus insuring that $|DF^{21}(x_{-21}, y_{-21})| \neq 0$.

As in Example 4.3, it is likely that similarly checking the fixed points of the unstable 2^k-cycles for $k > 2$ will extend the region for which chaos can be proven even further. However, as before, selecting the correct sequence of roots for the inverse is very difficult.

EXAMPLE 4.5. As a final example, consider the two-dimensional "host-parasite" scheme investigated by Beddington Free and Lawton:

$$\begin{aligned} x_{k+1} &= x_k \exp(r(1 - x_k/N) - ay_k), \\ y_{k+1} &= by_k(1 - \exp(-ay_k)), \end{aligned} \tag{4.10}$$

where $a, b, r, N \in \mathbb{R}^+$. In [6], (4.10) is reduced to a one-parameter family of problems in the following manner: first, letting the parameters b and N be fixed, and (x^*, y^*) represent the fixed point of (4.10), define a new quantity quantity $q = x^*/N$. Now, it can be shown that providing a value for q is equivalent to converting the parameter a into a function of b, r, N and q by:

$$a = \frac{r(1 - q)}{bNq(1 - \exp(r(q - 1))}.$$

Therefore, fixing the quantities b, N and q, we have $a = a(r)$ and (4.10) is now a family of equations parameterized by the variable r alone. Note that in addition, the fixed point (x^*, y^*) is given by: $x^* = Nq$ and $y^* = r(1 - q)/a(r)$. The particular values selected for b, N and q were: $b = 1$, $N = 10$ and $q = .4$. Thus we have $x^* = 4$, $y^* = .6r/a(r)$, and:

$$a(r) = \frac{.6r}{4(1 - \exp(-.6r))}.$$

Under these conditions (4.10) was investigated numerically in [6] for a variety of parameter values r, between $r = .5$ and $r = 2.75$. The results obtained are similar to those which we have encountered in each of the two previous examples in $\mathbb{R}^2$: for $r = .5$ the stability of the equilibrium point (x^*, y^*) can be seen, but as r is increased, a wide spectrum of behavior emerges. This includes stable continuous curves of increasing "radius" around (x^*, y^*), stable $5 \cdot 2^k$-cycles for increasing integers k, and finally for $r = 2.75$, chaos. Also, the kind of chaos which is observed for $r = 2.75$ visually resembles that which occurs for equation (4.4) in the region R_4 of Example 4.3, in which both eigenvalues of DF evaluated at the fixed point exceed 1 in norm, and to which Theorem 3.1 was applicable. Thus we may again expect this to be the case.

Numerical evidence does confirm that the eigenvalues of $DF(x^*, y^*)$ are complex conjugates and exceed 1.5 in norm for $r > 2.50$. In addition, considering the points $x^* + \epsilon$ and $y^* + \delta$, it can be shown that: $\| \lambda_i \| = \| \lambda_i(x^* + \epsilon, y^* + \delta)\| > 1$ for all $| \epsilon |$, $| \delta | < 10^{-2}$, and so (x^*, y^*) is expanding in $B_s(x^*, y^*)$ for $s = 10^{-2}$. Because of the complexity of the function, it is not as simple to compute the pre-images of (x^*, y^*). Thus the following analysis was performed: there exists a curve in the first quadrant of the (x_k, y_k)-plane along which

$|DF(x_k, y_k)| = 0$. This is very easily computed and is given by: $x_k = 10\exp(a(r)\,y_k)/r$. This path divides the quadrant into two separate regions D_1 and D_2, as pictured in Fig. 10 for the value $r = 2.75$. Note that $(x^*, y^*) \in D_1$ and that F is 1-1 in either D_1 or D_2. However, for the value $r = 2.75$, there was found to exist a pre-image point (x_{-1}, y_{-1}) of (x^*, y^*) inside D_2 by letting $(x_0, y_0) = (x^*, y^*)$ and solving the following for $k = 0$ by the two-dimensional Newton method:

$$
\begin{aligned}
x\exp(r(1 - x/10) - a(r)\cdot y) - x_k &= 0\\
x(1 - \exp(-a(r)\cdot y)) - y_k &= 0,
\end{aligned}
\tag{4.11}
$$

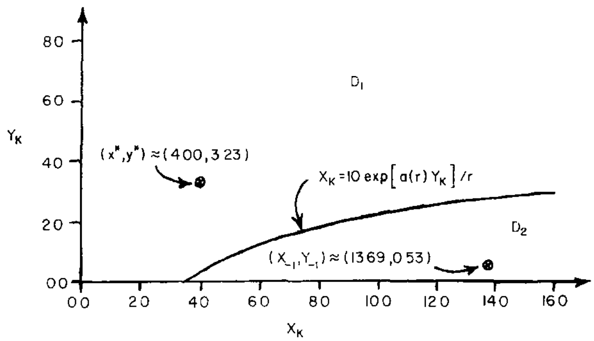

FIGURE 10

thus producing an estimate for $(x, y) = (x_{-1}, y_{-1})$. For all $k \leqslant -1$ (4.11) was successively solved, but for these cases the values of (x_k, y_k) inside D_1 were chosen. (50 iterations of the algorithm each time produced successive estimates for (x_k, y_k) differing by less than 10^{-10}.) It was discovered that with this choice of roots (x_{-28}, y_{-28}) lies inside $B_s(x^*, y^*)$ with $s = 10^{-2}$. In addition, none of the points (x_k, y_k) for $-28 \leqslant k < 0$ lies on the path separating D_1 from D_2, along which $|DF| = 0$. Hence, (x^*, y^*) is a snap-back repeller and Theorem 3.1 establishes chaos for $r = 2.75$.

This same analysis was also performed upon (4.10) for slightly smaller values of r. With the roots selected as above, the point (x_{-33}, y_{-33}) lies inside $B_s(x^*, y^*)$ with $s = 10^{-2}$ (and successive iterations of the Newton method differing by less than 10^{-10} after 50 iterations) for all $r > 2.62$. Since again the iterates did *not* satisfy: $x_k = 10\exp(a(r)\cdot y_k)/r$, for $-33 \leqslant k \leqslant -1$, then $|DF(x_{-33}, y_{-33})| \neq 0$, and chaos of (4.10) occurs for $r > 2.62$.

Because of the very complex nature of (4.10), this same type of investigation was not performed upon the functions F^n for larger values of n. It is again likely, however, that as in previous examples the range of r values producing chaos may extend even further than that which was obtained by applying Theorem 3.1 to F.

Acknowledgment

The author wishes to thank his thesis advisor, Prof. James L. Kaplan of Boston University, for suggesting the problem and for many helpful comments concerning this paper.

References

1. T.-Y. Li and J. A. Yorke, Period three implies chaos, *Amer. Math. Monthly* **82** (1975), 985–992.
2. E. N. Lorenz, Deterministic nonperiodic flow, *J. Atmos. Sci.* **20** (1963), 130–141.
3. R. M. May, Mathematical aspects of the dynamics of animal populations, to appear.
4. R. M. May, Biological populations with nonoverlapping generations: Stable points, stable cycles, and chaos, *Science* **186** (1974), 645–647.
5. J. Guckenheimer, G. Oster, and A. Ipaktchi, The dynamics of density dependent population models, to appear.
6. J. R. Beddington, C. A. Free, and J. H. Lawton, Dynamic complexity in predator prey models framed in difference equations, *Nature* **255** (1975), 58–60.
7. S. Smale, Differentiable dynamical systems, *Bull. Amer. Math. Soc.* **73** (1967), 747–817.
8. F. R. Marotto, Doctoral Thesis, Boston University, Boston, Mass.