MATHEMATICAL CONTROL THEORY
BANACH CENTER PUBLICATIONS, VOLUME 14
PWN – POLISH SCIENTIFIC PUBLISHERS
WARSAW 1985

CHAOTIC DYNAMICS AND NONLINEAR FEEDBACK CONTROL

J. BAILLIEUL

Scientific Systems, Inc., Cambridge, Mass., U.S.A.

1. Introduction

Many exciting ideas have appeared in the applied mathematics literature in the last few years concerning chaotic motions in dynamical systems. A somewhat surprising development has been the discovery that the familiar models used by engineers to design and analyze electric energy systems and power conversion networks may also exhibit chaos. (See Baillieul *et al.* [1].) In this paper we shall give a detailed description of a certain class of nonlinear feedback systems displaying chaos. We shall also determine certain cases in which this class of systems may be studied using statistical methods. For the purposes of this exposition the following is the basic definition.

DEFINITION 1.1. A difference equation will be said to display *chaos* if:

(i) there is an infinite family of periodic trajectories such that for any arbitrarily chosen positive integer N there is a trajectory whose minimal period exceeds N, and

(ii) there is an uncountable family of bounded aperiodic trajectories with the property that if $x(\cdot)$ and $y(\cdot)$ are distinct members of the family, there exists some $\varepsilon > 0$ such that $\|x(k)-y(k)\| > \varepsilon$ for arbitrarily large values of k.

2. Symbolic dynamics

Symbolic dynamics provides a mechanism through which it is possible to "keep track" of trajectories of dynamical systems. Although the main ideas in this area were around in the nineteenth century, the first occasion (of which we are aware) on which they were systematically used was in two early works by Morse ([11], [12]). Using this theory Morse was able to show that geodesic motions on manifolds of negative curvature were recurrent (i.e., returned infinitely often to an arbitrarily small neighborhood of an initial point) without being periodic.

DEFINITION 2.1. For each positive integer m let $S_m = \{0, 1, \ldots, m\}$ This is called the *symbol set*. Any sequence $\{\gamma_k\}_{k=0}^{\infty}$ taking values in S_m will be called a *symbol sequence*.

To employ this concept in the study of discrete dynamical systems

$$x(k+1) = F\big(x(k)\big),$$

where $F\colon \mathbf{R}^n \to \mathbf{R}^n$, we partition $\mathbf{R}^n$ into $m+1$ mutually disjoint and exhaustive subsets $R_0, R_1, \ldots, R_m$. We shall say that a trajectory $x(\cdot)$ *realizes* a symbol sequence $\{\gamma_k\}$ or equivalently that $\{\gamma_k\}$ *describes* $x(\cdot)$ if $\gamma_k = j$ implies and is implied by $x(k) \in R_j$.

In what follows our main interest will be in the two element symbol set $S_1 = \{0, 1\}$.

DEFINITION 2.2. We shall say a symbol sequence $\{\gamma_k\}$ is *admissible* if (i) $\gamma_0 = 0$, (ii) $\gamma_i\gamma_{i+1} = 0$ for all positive integers i, and (iii) there are arbitrarily large integers i such that $\gamma_i = 1$.

DEFINITION 2.3. A symbol sequence will be said to be of *class* k if it is an admissible sequence and (i) the number of zeros adjacent to any given zero is even (possibly 0) and (ii) the longest string of adjacent zeros is $2k+1$.

Symbol sequences of class k may be constructed as follows. First, define a finite set of finite symbol sequences;

$$\begin{aligned} a_0 &= 01, \\ a_1 &= 0001, \\ &\vdots \\ a_k &= \underbrace{00\ldots0}_{2k+1 \text{ zeros}}1. \end{aligned}$$

Then consider the symbol set $S = \{a_0, a_1, \ldots, a_m\}$. A symbol sequence of class k is any symbol sequence associated to S, viewed as a sequence of zeros and ones.

3. Nonlinear feedback systems. The scalar case

The main results of this paper will deal with feedback systems of the form

$$x(k+1) = Ax(k) + bf\big(cx(k)\big), \tag{1}$$

where A, b, c are real matrices of dimensions $n\times n$, $n\times 1$ and $1\times n$, respectively, and f is a (nonlinear) scalar function. To study such systems we shall partition R^n into $R_0=\{x: cx\leqslant 0\}$ and $R_1=\{x: cx>0\}$, and to begin we shall consider the special case in which $n=1$, $A=a$ (some real number), $b=c=1$ and $f(y)=\beta|y|-1$ for some real number β such that $\beta-a\geqslant\sqrt{2}$. Under these assumptions it is not difficult to show that (1) will have a solution trajectory of minimum period 2 if and only if $\beta^2-a^2\geqslant -1$, and there will be a solution trajectory of minimum period 4 if and only if $\beta^2-a^2\geqslant 1$. More generally the following facts are known:

LEMMA 1. *Consider the special case*

$$x(k+1)=ax(k)+\beta|x(k)|-1, \tag{2}$$

where $\beta-a\geqslant\sqrt{2}$. *If*

$$\beta^2-a^2>\sum_{k=0}^{m}(a-\beta)^{-2k},$$

then any symbol sequence of class $k\leqslant m$ *can be realized by a trajectory. Moreover, all such trajectories are confined to the set* $[-1,l)\cup(r,\beta-a-1]$, *where*

$$l=(a-\beta)^{-1}+(a-\beta)^{-2}+(a-\beta)^{-2}(a+\beta)^{-1}+\ldots+(a-\beta)^{-2k-3}(a+\beta)^{-1},$$

$$r=(a+\beta)^{-1}+(a+\beta)^{-1}(a-\beta)^{-1}+(a+\beta)^{-2}(a-\beta)^{-1}+$$
$$+(a+\beta)^{-2}(a-\beta)^{-2}+\ldots+(a+\beta)^{-2}(a-\beta)^{-2k-2}.$$

The interval $[-1,\beta-a-1]$ *is invariant under the motion of this system*

Proof. Define a sequence of intervals inductively by means of $a_1=-1$, $b_1=0$, $c_1=\beta-a-1$ and $(a-\beta)a_{i+1}-1=a_i$, $(a-\beta)b_{i+1}-1=b_i$ and $(a-\beta)c_{i+1}-1=c_i$. Let I_k denote the interval with endpoints b_k and c_k. Letting

$$g(x)=ax+\beta|x|-1,$$

we find that for $k>1$ g maps the interval I_k bijectively onto I_{k-1} (via the affine formula $x\mapsto(a-\beta)x-1$. The function g also maps the interval I_1 bijectively onto $[-1,g(c_1)]$ via the affine formula $x\mapsto(a+\beta)x-1$. Thus for each $k=1,2,\ldots$ g defines an affine mapping g_k: $I_k\to[-1,g(c_1)]$, which we write explicitly as

$$g_k(x)=(a-\beta)^{k-1}(a+\beta)x-(a-\beta)^{k-2}(a+\beta)-\ldots-(a-\beta)(a+\beta)-(a+\beta)-1.$$

Note that if k is even and $k\leqslant 2m+2$, then under the hypothesis of the lemma we have $I_k\subset[-1,g(c_1)]$.

Let $\{\gamma_j\}$ be a symbol sequence of class $k\leqslant j$, and let n_i denote the location in this sequence of the ith "1". Let $n_0=0$ and for each positive integer i define $m_i = n_i - n_{i-1}$. Note that since $\{\gamma_j\}$ is of class k, $m_i \leqslant 2k+1$. For each positive integer j define

$$J_j = g_{m_1}^{-1} g_{m_2}^{-1} \cdots g_{m_j}^{-1}[-1, g(c_1)].$$

Since for any even positive integer ν, such that $\nu \leqslant 2m+2$, $g_\nu^{-1}[-1, g(c_1)] \subset [-1, g(c_1)]$, it follows that $J_j \subset J_{j-1} \subset [-1, g(c_1)]$ for $j \geqslant 2$. Moreover, it follows by direct calculation using the explicit formula given above for g_j, that the length of J_j is

$$(\beta - a)^{-h}(a+\beta)^{-j+1}(\beta - a - 1), \qquad \text{where} \qquad h = \sum_{i=1}^{j}(m_i - 1).$$

(Note that $a+\beta$, $\beta - a - 1$ are positive.) Now, from the assumed properties of the sequence $\{\gamma_i\}$ it follows that $m_i \geqslant 2$ for each positive integer i. Hence $h \geqslant j$ and the length of J_j is less than or equal to

$$(\beta - a)^{-j}(a+\beta)^{-j+1}(\beta - a - 1) = (\beta^2 - a^2)^{-j}(a+\beta)(\beta - a - 1).$$

Since $0 < (\beta^2 - a^2)^{-1} < 1$, the length of J_j approaches zero as j approaches infinity. Hence there is precisely one point $x(0)$ common to all the nested subintervals. Taking this as our initial condition for (2) we obtain a trajectory, all of whose points lie in $[-1, \beta - a - 1]$, and which realizes the given symbol sequence.

To prove the second part of the lemma, let $c_1 = \beta - a - 1$ and let d satisfy

$$(a+\beta)d - 1 = \sum_{j=1}^{2k+1} (a-\beta)^{-j}.$$

First, we verify that $d < c_1$. To do this, note that it follows from the hypothesis that

$$\beta^2 - a^2 > \sum_{j=0}^{k}(a-\beta)^{-2j} > 0.$$

This implies

$$a+\beta > -\sum_{j=0}^{k}(a-\beta)^{-2j-1} > 0.$$

From this inequality we find

$$\begin{aligned}(a+\beta)c_1 - 1 &= (a+\beta)(\beta - a - 1) - 1 \\ &> -\sum_{j=0}^{k}(a-\beta)^{-2j-1}(\beta - a - 1) - 1 = \sum_{j=1}^{2k+1}(a-\beta)^{-j} = (a+\beta)d - 1.\end{aligned}$$

Comparing the first and last terms verifies the claim.

Next we claim that no point on the trajectory we have constructed lies in $[d, c_1]$. Recall the above construction of subintervals I_j. Two features of this construction are (1) the right-hand endpoint of I_{2j} coincides with the left-hand endpoint of I_{2j+2} for $j = 1, 2, \ldots$, and (2) the left-hand endpoint of each odd numbered interval lies to the right of the right-hand endpoint of every even numbered interval. Now suppose $x(\cdot)$ is our trajectory constructed above, and consider the consequences of assuming $x(j) \in [d, c_1]$ for some j. Then

$$x(j+1) = (a+\beta)x(j)-1 \geqslant (a+\beta)d-1 = \sum_{j=1}^{2k+1} (a-\beta)^{-i}.$$

But this last quantity is just the right-hand endpoint of I_{2k+2}. There are exactly three possible cases in which $x(j+1)$ can lie to the right of $I_{2(k+1)}$.

I. $x(j+1) > 0$. In this case the associated symbol sequence would have $\gamma_j = \gamma_{j+1} = 1$, which violates the hypothesis that $\{\gamma_j\}$ is an admissible symbol sequence of class k.

II. $x(j+1) \in I_{2i+1}$ *for some positive integer* i. In this case there would have to be an even number of adjacent zeros in the associated symbol sequence, again violating our a priori restrictions.

III. $x(j+1) \in I_{2i}$ *for* $i \geqslant k+1$. In this case the number of consecutive zeros exceeds $2k+1$ which also violates our original assumption that $\{\gamma_j\}$ is a symbol sequence of class k.

In any case we have shown that no point on the trajectory can lie in the interval $[d, c_1]$. But $[l, r] = g^{-2}[d, c_1]$, and hence no point on our trajectory can lie within $[l, r]$. This completes the proof of our lemma.

In analyzing the stability of the constructed family of trajectories it will be useful to have estimates of the closest possible approach of any trajectory to zero and also the maximum distance from zero attained by any trajectory. (Recall zero is the corner point of the nonlinear function g.) Hence we let

$$\eta_0 = \min\{|l|, r\}$$

and

$$\eta_1 = \max\{1, \beta-a-1\}.$$

Closed-form expressions for η_0 and η_1 in terms of a and β may be easily calculated. In the next section we shall be interested in the function

$$k_1(a, \beta) = \frac{\eta_0}{\eta_0+2\eta_1}.$$

Consider a trajectory of period k from the family constructed in the proof of the above lemma, and define a mapping L from the space of periodic sequences of period k to itself by means of the input-output relation

$$x(k+1) = ax(k)+\beta|x(k)|-1+u(k). \tag{3}$$

LEMMA 2. *The incremental l_∞ gain of the mapping L computed about any periodic trajectory constructed in the proof of Lemma 1 is bounded by* ([1])

$$\|L\| \leqslant k_2(a, \beta),$$

where

$$k_2(a, \beta) = \frac{(\beta-a)(\beta-a+1)}{\beta^2-a^2-1} + \frac{1}{\beta-a-1}.$$

Moreover, if $\bar{x}$ is a period k sequence generated by (3) *with $u \equiv 0$ and u is a period k input with $\|u\| < \eta_0/k_2(a, \beta)$, then the corresponding period k output x has the property that*

$$\operatorname{sgn} x(j) = \operatorname{sgn} \bar{x}(j) \quad \textit{for} \quad j = 0, 1, \ldots, k-1.$$

Proof. The period k solutions to (3) (for $u \equiv 0$) in the proof of Lemma 1 may also be obtained as solutions to the affine equation $x = Dx-\xi$, where ξ is a k-vector of 1's and D is the $k\times k$ matrix with $a+\beta$ in the $(1, k)$ entry, $a-\beta$ and $a+\beta$ alternating according to the associated symbol sequence down the subdiagonal and 0's elsewhere. The mapping L is given by $(I-D)^{-1}$. To show that this L satisfies the stated bounds we consider separately the two cases $a+\beta>1$ and $a+\beta\leqslant 1$.

In the former case we find $\|D^{-1}\| = \max\{1/(\beta-a),\ 1/(a+\beta)\} < 1$. From this we estimate $\|L\|$ from the following sequence of inequalities:

$$\begin{aligned}\|L\| &= \|(I-D)^{-1}\| = \|(D^{-1}-I)^{-1}D^{-1}\| \\ &\leqslant \|(I-D^{-1})^{-1}\|\,\|D^{-1}\| \leqslant \frac{\|D^{-1}\|}{1-\|D^{-1}\|} \leqslant \frac{1}{\beta-a-1}.\end{aligned}$$

The next to last inequality is a standard result from operator theory. Since $k_2(a, \beta) > 1/(\beta-a-1)$, the corollary is proved in this case. One may establish the same result when $a+\beta\leqslant 1$ by explicitly computing the inverse of $I-D$. The calculations involved here are somewhat tedious and we omit the details.

Finally, let $\bar{x}$ be the (unique!) solution of the equation $x = Dx-\xi$. Let $u = \big(u(k-1), u(0), \ldots, u(k-2)\big)^T$ and suppose x satisfies $x = Dx-\xi+u$. Then $x-\bar{x} = Lu$. Since $\|x-\bar{x}\| \leqslant \|L\|\,\|u\|$, if $\|u\| < \eta_0/k_2(a, \beta)$, then $\operatorname{sgn} x(j) = \operatorname{sgn}\bar{x}(j)$ for $j = 0, 1, \ldots, k-1$.

([1]) All vector norms in this paper are l_∞ norms, and norms of matrices and linear operators are the associated operator norms.

Intuitive insight into the quantities η_0, η_1 and $k_1(\alpha, \beta)$ may be gleaned from the graph of $F(x) = \alpha x + \beta |x| - 1$ displayed in Figure 1 (where $e = \sum_{j=1}^{2m+1} (\alpha - \beta)^{-j}$ and l, r, d and c_1 are described in Lemma 1 and its proof). $k_2(\alpha, \beta)$ gauges how sensitive our family of periodic trajectories is to perturbations in the parameters of the system. One would expect that in cases where $k_2(\alpha, \beta)$ is small the equation (3) could be changed by a relatively large amount and still have the qualitative properties of our family of periodic trajectories preserved.

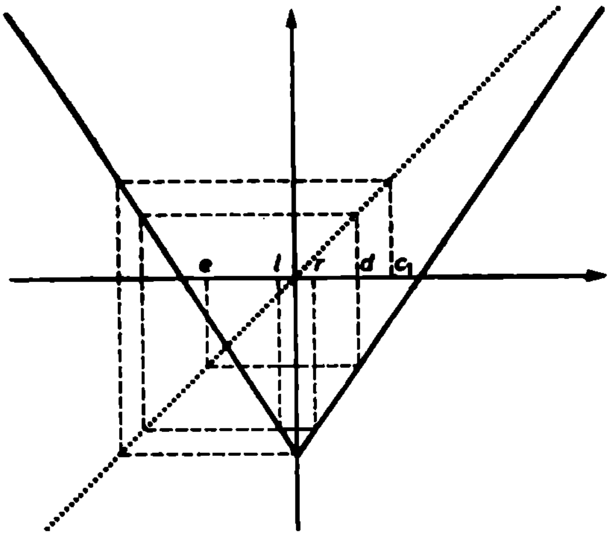

Fig. 1

4. Nonlinear feedback systems in R^n. Main results

The theory of chaotic behavior of difference equations in higher dimensions is not as completely developed as in the one-dimensional case. In [9] a characterization of chaotic dynamics has been given in terms of "snap-back repellers", but since snap-back repellers can be found easily only in very special cases, these results at present remain difficult to apply. In this section we present more easily verified conditions which imply the existence of an infinite family of periodic trajectories as well as a certain uncountable family of aperiodic trajectories. Our main result in this regard is a frequency domain characterization of the existence of chaotic solutions to vecotor difference equations of the form (1), where f is a nonlinear function satisfying a certain Lipschitz-type condition.

THEOREM 1. *Suppose the transfer function of the minimal triple* (A, b, c) *has the expression*

$$g(z) = c(Iz - A)^{-1}b = \frac{1}{-\alpha + z + h(z)},$$

where $h(\cdot)$ is a proper rational function having all its poles in the interior of the unit disk and a, β are numbers which satisfy $\beta - a \geqslant \sqrt{2}$ and

$$\beta^2 - a^2 > \sum_{k=0}^{m} (a-\beta)^{-2k}.$$

Suppose also that $f(y) = \beta|y| - 1 + \varphi(y)$, where φ satisfies a uniform Lipschitz condition with constant $\mu \geqslant 0$. Then if(2)

$$k_2(a, \beta)(\mu + \|h\|_\infty) < k_1(a, \beta),$$

any symbol sequence of class k (for $0 \leqslant k \leqslant m$) may be realized by a trajectory of (1). If $m > 0$, then there are infinitely many periodic trajectories together with an uncountably infinite family of aperiodic trajectories with the property that if $x^1(\cdot)$ and $x^2(\cdot)$ are distinct members of this family,

$$\|x^1(j) - x^2(j)\| \geqslant \eta_0$$

for arbitrarily large values of j.

Remark. Recall we are considering the symbol set $S = \{0, 1\}$ and the corresponding partition of $\boldsymbol{R}^n$ is into

$$R_0 = \{x\colon\ cx \leqslant 0\},\ R_1 = \{x\colon\ cx > 0\}.$$

Proof. Pick a minimal triple $(\hat{A}, \hat{b}, \hat{c})$ to realize $h(z)$ so that

$$g(z) = \frac{1}{-a + z - \hat{c}(Iz - \hat{A})^{-1}\hat{b}}.$$

One possible choice of A, b, and c having this transfer function is

$$A = \begin{bmatrix} a & \hat{c} \\ \hat{b} & \hat{A} \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad c = (1, 0, \ldots, 0). \tag{4}$$

One easily checks that this is minimal, and by the state space isomorphism theorem it can differ from the original triple defining (1) only by a linear change of basis, given, say, by a nonsingular matrix P. Since the system (4) will satisfy the conclusion of the theorem if and only if the corresponding system defined by PAP^{-1}, Pb and cP^{-1} does (with possibly a different value of η_0), we may as well assume from the start that A, b and c have the form (4).

(2) $h(\cdot)$ defines a linear transformation from the space of all bounded input sequences to the space of output sequences. $\|h\|_\infty$ is used to denote the l_∞ operator norm of this mapping.

Let $\bar{y}(\cdot)$ be any period ν trajectory of the scalar system (2) corresponding to a symbol sequence (of the same period) of class $k \leqslant m$. Represent y as a ν-vector and let $\bar{w}(\cdot)$ be the corresponding periodic solution of $w(j+1) = \hat{A}w(j)+\hat{b}y(j)$. Write the equations for these period ν solutions in extensive form

$$\begin{bmatrix}\bar{y}\\ \bar{w}\end{bmatrix}\left[\begin{array}{c|c} D & 0\\ \hline \tilde{b} & \tilde{A}\end{array}\right]\begin{bmatrix}\bar{y}\\ \bar{w}\end{bmatrix}-\begin{bmatrix}\xi\\ 0\end{bmatrix}, \tag{5}$$

where D and ξ have been defined in the proof of Lemma 2, and

$$\tilde{b} = \begin{bmatrix} 0 & 0 & 0 & . & . & . & 0 & \hat{b}\\ \hat{b} & 0 & 0 & . & . & . & 0 & 0\\ 0 & \hat{b} & 0 & & . & . & 0 & 0\\ 0 & 0 & \hat{b} & & & & 0 & 0\\ \vdots & \vdots & \vdots & & & & \vdots & \vdots\\ 0 & 0 & 0 & & & & \hat{b} & 0 \end{bmatrix}$$

and

$$\tilde{A} = \begin{bmatrix} 0 & 0 & 0 & . & . & . & 0 & \hat{A}\\ \hat{A} & 0 & 0 & . & . & . & 0 & 0\\ 0 & \hat{A} & 0 & . & . & . & 0 & 0\\ 0 & 0 & \hat{A} & . & . & . & 0 & 0\\ \vdots & \vdots & \vdots & & & & \vdots & \vdots\\ 0 & 0 & 0 & . & & . & \hat{A} & 0 \end{bmatrix}.$$

To find a period k solution to (1) we consider the related equation

$$\begin{bmatrix}y\\ w\end{bmatrix}=\left[\begin{array}{c|c} D & \tilde{c}\\ \hline \tilde{b} & \tilde{A}\end{array}\right]\begin{bmatrix}y\\ w\end{bmatrix}-\begin{bmatrix}\xi\\ 0\end{bmatrix}+\begin{bmatrix}\Phi(y)\\ 0\end{bmatrix}, \tag{6}$$

where

$$\Phi\big((y(0), y(1), \ldots, y(\nu-1))^{\mathrm{T}}\big) = \big(\varphi(y(\nu-1)), \varphi(y(0)), \ldots, \varphi(y(\nu-2))\big)^{\mathrm{T}},$$

and where

$$\tilde{c} = \begin{bmatrix} 0 & 0 & 0 & . & . & . & 0 & \hat{c}\\ \hat{c} & 0 & 0 & . & . & & 0 & 0\\ 0 & \hat{c} & 0 & . & . & . & 0 & 0\\ 0 & 0 & \hat{c} & . & . & . & 0 & 0\\ \vdots & \vdots & \vdots & & & & \vdots & \vdots\\ 0 & 0 & 0 & . & . & . & \hat{c} & 0 \end{bmatrix}.$$

The hypothesis of the theorem implies that $I-\hat{A}$ is invertible and this in turn implies the invertibility of $I-\tilde{A}$. Hence we may write $w = (I-\tilde{A})^{-1}\tilde{b}y$, and we obtain an equation for y from (6)

$$[I-D-\tilde{c}(I-\tilde{A})^{-1}\tilde{b}]y = \Phi(y)-\xi.$$

We rewrite this equation as

$$L^{-1}[I-L\tilde{c}(I-\tilde{A})^{-1}\tilde{b}]y = \Phi(y)-\xi. \tag{7}$$

We know that if $\|L\|\,\|\tilde{c}(I-\tilde{A})^{-1}\tilde{b}\|<1$, then the matrix $I-L\tilde{c}(I-\tilde{A})^{-1}\tilde{b}$ has an inverse. By Lemma 2 $\|L\|<k_2(\alpha,\beta)$, and from the definitions of $\tilde{A}$, $\tilde{b}$ and $\tilde{c}$ it is a straightforward calculation to show that

$$\|\tilde{c}(I-\tilde{A})^{-1}\tilde{b}\| = \sum_{j=0}^{\nu-1}|\hat{c}\hat{A}^j(I-A^\nu)^{-1}\hat{b}|$$

but

$$\sum_{j=0}^{\nu-1}|\hat{c}\hat{A}^j(I-A^\nu)^{-1}\hat{b}| = \sum_{j=0}^{\nu-1}\left|\hat{c}\hat{A}^j\sum_{i=0}^{\infty}\hat{A}^\nu\hat{b}\right| \leqslant \sum_{j=0}^{\infty}|\hat{c}\hat{A}^j\hat{b}|.$$

This last quantity is just the l_∞ gain of $h(\cdot)$ and it follows from the hypothesis that it, and a fortiori $\|\tilde{c}(I-\tilde{A})^{-1}b\|$, are less than $k_1(\alpha,\beta)/k_2(\alpha,\beta)$. Thus we may invert the coefficient of y on the left-hand side of (7) to obtain

$$y = [I-L\tilde{c}(I-\tilde{A})^{-1}\tilde{b}]^{-1}L\Phi(y)-[I-L\tilde{c}(I-\tilde{A})^{-1}\tilde{b}]^{-1}L\xi. \tag{8}$$

This is an equation of the form $y = F(y)$, where a straightforward calculation shows that F satisfies a uniform Lipschitz condition with constant

$$k_2(\alpha,\beta)\mu/[1-k_1(\alpha,\beta)+k_2(\alpha,\beta)\cdot\mu].$$

Since this is less than 1, it follows from the Banach fixed-point theorem that (8) has a unique solution, and from this we obtain a solution to (6).

It remains only to check that this solution we have obtained to (6) actually represents a solution to (1). That is, we must check that $\operatorname{sgn} y(j) = \operatorname{sgn}\bar{y}(j)$ for $j = 0,1,\dots,\nu-1$. To see this we write

$$\begin{aligned} y-\bar{y} &= D(y-\bar{y})+\tilde{c}(I-\tilde{A})^{-1}\tilde{b}y+\Phi(y) \\ &= [I-D-\tilde{c}(I-\tilde{A})^{-1}b]^{-1}[\Phi(y)-\Phi(\bar{y})+\tilde{c}(I-\tilde{A})^{-1}b\bar{y}+\Phi(\bar{y})]. \end{aligned}$$

From this it follows that

$$\|y-\bar{y}\| < \frac{k_2(\alpha,\beta)}{1-k_2(\alpha,\beta)\|\tilde{c}(I-\tilde{A})^{-1}\tilde{b}\|}\left(\mu\|y-\bar{y}\|+\|\tilde{c}(I-\tilde{A})^{-1}b\|\,\|\bar{y}\|+\mu\|\bar{y}\|\right).$$

Hence

$$\|y-\bar{y}\| < \frac{k_2(\alpha,\beta)\,(\mu+\|\tilde{c}(I-\tilde{A})^{-1}\tilde{b}\|)}{1-k_2(\alpha,\beta)\,(\mu+\|\tilde{c}(I-\tilde{A})^{-1}\tilde{b}\|)}\|\bar{y}\|.$$

But we have seen that $\|\tilde{c}(I-\tilde{A})^{-1}\tilde{b}\| = \|h(\cdot)\|_\infty$ and it then follows from the hypothesis of the theorem that

$$\|y-\bar{y}\| < \frac{k_1(\alpha,\beta)}{1-k_1(\alpha,\beta)}\,\|\bar{y}\| \leqslant \frac{\eta_0}{2}.$$

From the definition of η_0 this implies that $\operatorname{sgn} y(j) = \operatorname{sgn} \bar{y}(j)$ for $j=0, 1,\ldots,\nu-1$.

We shall next show that to each of the aperiodic trajectories for the scalar system $y(j+1) = \alpha y(j)+\beta|y(j)|-1$ constructed in the proof of Lemma 1 there corresponds an aperiodic trajectory for (1) having the stated properties. Let $\{\gamma_j\}$ be a symbol sequence of class k ($k\leqslant m$). As before, for each positive integer i let n_i denote the position of the ith "1" in the sequence $\{\gamma_j\}$. Let $\big(y_i(\cdot), w_i(\cdot)\big)$ denote the periodic sequence of period η_i constructed above with periodic symbol sequence defined by $\gamma_0\gamma_1\cdots\gamma_{n_i-1}$. Then $\big(y_i(0), w_i(0)\big)$ is a bounded sequence in $\boldsymbol{R}^n$ and by the Bolzano–Weierstrass Theorem there is an accumulation point, $\big(y(0), w(0)\big)$, and a subsequence converging to it. Suppose we write, by renumbering if necessary, $\lim_{i\to\infty}\big(y_i(0), w_i(0)\big) = \big(y(0), w(0)\big)$. Then by continuity and a simple induction argument $\lim_{i\to\infty}\big(y_i(k), w_i(k)\big) = \big(y(k), w(k)\big)$, where $\big(y(k), w(k)\big)$ is defined in terms of $\big(y(0), w(0)\big)$ by (1). It is now not difficult to see that (1) and the initial condition $\big(y(0), w(0)\big)$ define a bounded aperiodic trajectory which realizes the symbol sequence $\{\gamma_j\}$.

It remains only to prove the final statement of the theorem. If $m>0$, then any symbol sequence formed by concatenating 01 and 0001 in any order may be realized by a trajectory. Clearly there are infinitely many trajectories (both periodic and aperiodic) of this type. Define an equivalence relation on symbol sequences of class m by saying that $\{\gamma_j\}\sim\{\delta_j\}$ if and only if there exist integers $N>0$ and M such that $\gamma_{j+M}=\delta_j$ whenever $j\geqslant N$. We shall show how to explicitly construct the equivalence classes for this equivalence relation, and in the process prove that each equivalence class is a countable set. Let $\{\gamma_j\}$ be an arbitrary symbol sequence of class m. For each positive integer ν define an equivalent sequence $\{\gamma_j^\nu\}$ by writing $\gamma_j^\nu = \gamma_{j+\nu}$ for $j=0,1,\ldots$ From this collection of sequences discard all those which are not admissible. (i.e., discard all those which initiate with a "1".) Let $\mathscr{S}$ denote the set of all finite symbol sequences of class m. (The elements of $\mathscr{S}$ consist of finite concatenations of symbol sequences of the form

$$a_0 = 01,\ a_1 = 0001, \ldots, a_m = \underbrace{00\ldots0}_{2m+1\text{ zeros}}1.)$$

$\mathscr{S}$ is a countable set. For each positive integer ν let $\mathscr{S}^\nu$ denote the set of sequences formed by the concatenation of members of $\mathscr{S}$ with the sequence $\{\gamma_i^\nu\}$. I.e., $\mathscr{S}^\nu$ is the set of all sequences of the form $a_{i_1} a_{i_2} \ldots a_{i_j} \gamma_0^\nu \gamma_1^\nu \gamma_2^\nu \ldots$ Then it is clear that the equivalence class of the symbol sequence $\{\gamma_j\}$ which we started with is the union of all sets of sequences $\mathscr{S}^\nu$ such that $\{\gamma_j^\nu\}$ is an admissible sequence of class m. Clearly this is a countable set (it is the countable union of countable sets), and this proves our claim.

Now choose a representative from each equivalence class we have just constructed and let the set of these representatives be denoted by $\mathscr{T}$. One may show that $\mathscr{T}$ is an uncountably infinite set by observing that the set $\mathscr{U}$ of all symbol sequences of class m is uncountably infinite, that $\mathscr{U} = \bigcup_{t \in \mathscr{T}} [t]$ (where $[t]$ = equivalence class of t) and that if $\mathscr{T}$ were countable, this would be a representation of $\mathscr{U}$ as the countable union of countable sets. The uncountable set to which we referred in the statement of the theorem is the set of trajectories corresponding to $\mathscr{T}$.

Let $\{\gamma_j^1\}$ and $\{\gamma_j^2\}$ now denote any two distinct symbol sequences in $\mathscr{T}$, and let $x^1(\cdot) = (y^1(\cdot), w^1(\cdot))$ and $x^2(\cdot) = (y^2(\cdot), w^2(\cdot))$ denote the corresponding trajectories. We claim that $|y^i(j)| > \eta_0/2$ for $i = 1, 2$ and $j = 0, 1, \ldots$ To prove this, recall that we showed that any trajectory corresponding to a class m symbol sequence could be approximated pointwise by a family of periodic trajectories. Thus, let $\bar{x}^i(\cdot) = (\bar{y}^i(\cdot), \bar{w}^i(\cdot))$ denote a family of periodic trajectories whose symbol sequences are of class m and such that $\lim_{i \to \infty} \bar{x}^i(j) = x^1(j)$ for each $j = 0, 1, \ldots$ By our construction of periodic sequences corresponding to symbol sequences of class m we know $|\bar{y}^i(j)| > \eta_0$ for all $i = 0, 1, 2, \ldots$ and $j = 0, 1, 2, \ldots$ Since $\lim |\bar{y}^i(j)| = |y^1(j)|$, we may for each j choose i sufficiently large that $|\bar{y}^i - y^1(j)| < \eta_0/2$. From this it follows that $|y^1(j)| > \eta_0/2$. Now from our construction of the set $\mathscr{T}$ it follows that there are arbitrarily large values of j such that $\gamma_j^1 = 0$ and $\gamma_j^2 = 1$ or vice versa. Hence there are arbitrarily large values of j such that $y^1(j) < \eta_0/2$ and $y^2(j) > \eta_0/2$ or vice versa. Hence for arbitrarily large values of j

$$\|x^1(j) - x^2(j)\| \geqslant |y^1(j) - y^2(j)| > \eta_0.$$

This completes the proof of the theorem.

Remark. Viewing $\lambda = \beta^2 - a^2$ as a parameter in these models, it is apparent that as λ increases the set of symbol sequences which can be realized becomes increasingly rich. Indeed, when $\beta^2 - a^2 > a + \beta + 1$ (note that under the stated assumptions that $\beta - a \geqslant \sqrt{2}$, $a + \beta > 0$ we have $a + \beta + 1 > \sum_{k=0}^{m} (a - \beta)^{-2k}$ for all integers $m \geqslant 0$) it was shown in [1] than *any* admissible symbol sequence can be realized by a bounded trajectory of (1). Surprisingly, if $\beta^2-\alpha^2$ is sufficiently large, despite the existence of this rich family of bounded trajectories, almost all trajectories of (1) do not in general remain bounded. This observation is made precise in the following theorem.

THEOREM 2. *Under the hypotheses of Theorem* 1, *if in addition we assume*

$$\mu \leqslant \alpha+\beta, \quad \beta^2-\alpha^2<2\beta,$$
$$2\max\{\alpha+\beta+\mu,\ \beta-\alpha+\mu\}\|h\|_\infty<\alpha^2-\beta^2-\mu^2-2\beta\mu+2(\beta+\mu),$$

then there is a set of positive but finite Lebesgue measure in $\boldsymbol{R}^n$ *which is positively invariant under the motion of* (1). *If* $\beta^2-\alpha^2>2\beta$, *there will not generally exist an invariant set of positive but finite Lebesgue measure.*

Remarks. (1) This theorem is proved in [1].

(2) Also in [1] it is shown that when $\beta^2-\alpha^2>2\beta$ the maximal bounded subset of $\boldsymbol{R}$ which is positively invariant under the motion of the scalar system (2), is a subset of a Cantor set. Thus no finite invariant interval exists in this case.

5. A statistical approach to chaotic feedback systems

Much attention in the dynamical systems literature has been devoted to the development of statistical characterizations of deterministic chaotic motion. (See, for instance, [2], [5], [6], [7], [8], [13], and [14], to list but a few.) In this section we shall discuss the feasibility of such characterizations for systems of the form (1), and in certain special scalar cases we shall write down invariant densities explicitly. The evolution equations (1) in these cases turn out to be related to the so-called Markov maps defined by Bowen (see [3]). We point out that in our development these maps are not assumed to be almost everywhere expanding.

Suppose $U\subset \boldsymbol{R}^n$ and F: $U\to U$ defines a dynamical system by

$$x(k+1)=F\big(x(k)\big). \tag{9}$$

We shall say that a function ϱ: $U\to\{0,1\}$ is a *density for the process* F *at* x if for each measurable set $E\subset U$

$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=0}^{k-1}\chi_E\big(F^j(x)\big)$$

exists and equals

$$\int_E \varrho(s)\,ds.$$

Such a density will be called *invariant* under F if

$$\int_E \varrho(s)\,ds = \int_{F^{-1}(E)} \varrho(s)\,ds \tag{10}$$

for every measurable E. Let $g\colon U \to \mathbf{R}$. It follows from the Birkhoff ergodic theorem that if there exists a unique density ϱ, invariant under F, then for almost all (Lebesgue measure) $x \in U$

$$\lim_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k-1} g\big(F^j(x)\big) = \int_U g(s)\varrho(s)\,ds$$

whenever the integral exists.

In many cases it is possible to determine an invariant density by explicitly solving the integral equation (10). Suppose, for instance, that F is a k-to-one mapping and U is partitioned into k disjoint subsets, $U = \bigcup_{j=1}^{k} U_j$, such that $F\big|_{U_i^{\mathrm{int}}}$ is a diffeomorphism of U_i^{int} onto U^{int}. Let

$$T_i\colon\ U^{\mathrm{int}} \to U_i^{\mathrm{int}}$$

be the inverse of this restriction of F. If a piecewise continuous invariant density ϱ exists on U, it must then satisfy the functional equation

$$\varrho(x) = \sum_{i=1}^{k} \left|\frac{\partial T_i}{\partial x}\right| \varrho\big(T_i(x)\big),$$

where $|\partial T_i/\partial x|$ denotes the Jacobian determinant of T_i. In the special case of the system (1), where $f(y) = |y|-1$, this implies the following

PROPOSITION. *Let $f(y) = |y|-1$ and suppose the determinants $|A-bc| \neq 0$ and $|A+bc| \neq 0$. Then in any rectangular set $E \subset U$ such that $F^{-1}(E) \subset U$, a piecewise continuous invariant density for* (1) *must satisfy the functional equation*

$$\varrho(x) = \frac{1}{|A-bc|}\,\varrho[(A-bc)^{-1}(x+b)] + \frac{1}{|A-bc|}\,\varrho[(A+bc)^{-1}(x+b)].$$

For the scalar system (2) *this functional equation becomes*

$$\varrho(x) = \frac{1}{\beta-a}\,\varrho\left[\frac{x+1}{a-\beta}\right] + \frac{1}{a+\beta}\,\varrho\left[\frac{x+1}{a+\beta}\right] \tag{11}$$

if $\beta^2-\alpha^2 \geqslant 2\beta$, *and*

(12) $$\varrho(x) = \begin{cases} \dfrac{1}{\beta-\alpha}\,\varrho\left[\dfrac{x+1}{\alpha-\beta}\right]+\dfrac{1}{\alpha+\beta}\,\varrho\left[\dfrac{x+1}{\alpha+\beta}\right] & \text{if} \quad -1 \leqslant x \leqslant \beta^2-\alpha^2-\alpha-\beta-1, \\ \dfrac{1}{\beta-\alpha}\,\varrho\left[\dfrac{x+1}{\alpha-\beta}\right] & \text{if} \quad \beta^2-\alpha^2-\alpha-\beta-1 < x \leqslant \beta-\alpha-1 \end{cases}$$

if $\beta^2-\alpha^2 < 2\beta$.

The possibility of studying a deterministic system from a statistical point of view will require, at least, the existence of an invariant density. This existence is by no means assured for systems of the form (1). In the case of the scalar system (2), for example, we have remarked that when $\beta^2-\alpha^2 > 2\beta$ the set of initial conditions leading to bounded trajectories has measure zero. The following result thus comes as no surprise.

THEOREM 3. *Suppose* $\beta^2-\alpha^2 > 2\beta$. *Then*

(i) *there is no finite interval which is positively invariant for* (2).

(ii) *If* $\varrho(\cdot)$ *is a piecewise continuous invariant density for* (2) *it must satisfy* (11) *on* $[-1, \beta-\alpha-1]$.

(iii) *There is no bounded nontrivial non-negative solution to* (11) *defined on* $[-1, \beta-\alpha-1]$.

Proof. (i) is proved in [1] (Theorem 4.1). (ii) follows from the above proposition.

(iii): Suppose $\varrho(\cdot)$ is a solution of (11) and $x_0 \in [-1, \beta-\alpha-1]$ is such that $\varrho(x_0) > 0$. There are two values $(x_0+1)/(\alpha-\beta)$ and $(x_0+1)/(\alpha+\beta)$ in $[-1, \beta-\alpha-1]$ which are "antecedents" of x_0 under the motion of (2). ϱ must take on a value $\geqslant (\beta^2-\alpha^2)/2\beta\cdot\varrho(x_0) > \varrho(x_0)$ at one of these antecedents, which we label x_1. (Otherwise,

$$\begin{aligned} \varrho(x_0) &= \frac{1}{\beta-\alpha}\,\varrho\left[\frac{x_0+1}{\alpha-\beta}\right]+\frac{1}{\alpha+\beta}\,\varrho\left[\frac{x_0+1}{\alpha+\beta}\right] \\ &< \frac{1}{\beta-\alpha}\,\frac{\beta^2-\alpha^2}{2\beta}\,\varrho(x_0)+\frac{1}{\alpha+\beta}\,\frac{\beta^2-\alpha^2}{2\beta}\,\varrho(x_0) = \varrho(x_0), \end{aligned}$$

which is obviously impossible.) Now a simple induction argument shows there is a sequence of points $x_1, x_2, \ldots$ in $[-1, \beta-\alpha-1]$ such that

$$\alpha x_{i+1}+\beta|x_{i+1}|-1 = x_i \text{ (sic)} \quad \text{and} \quad \varrho(x_i) \geqslant [(\beta^2-\alpha^2)/2\beta]^i\varrho(x_0).$$

Since $(\beta^2-\alpha^2)/2\beta > 1$, this proves (iii).

It is easy to verify that when $\beta^2-\alpha^2 = 2\beta$ the uniform density on $[-1, \beta-\alpha-1]$ solves (11). When $\beta^2-\alpha^2 < 2\beta$, there may exist limit cycles for (2) (i.e., stable periodic trajectories). It can be shown that limit cycles provide severe restrictions on the form of any invariant density. (See the work of Misiurewicz [10], or [4], on this point.) Nevertheless, we may determine explicit solutions to (12) in a number of interesting cases meeting the hypotheses of Lemma 1.

THEOREM 4. *Suppose that $\beta-\alpha \geqslant \sqrt{2}$ and let*

$$\beta^2-\alpha^2 = \sum_{j=0}^{k-1} (\alpha-\beta)^{-2j}.$$

Then there is an invariant density ϱ for the system (2), *and this is given explicitly by*

$$\varrho(x) = \begin{cases} a & \text{if } -1 < x < (\alpha-\beta)^{-1}, \\ \left[(\beta^2-\alpha^2)^{-1}\sum_{j=0}^{k-l}(\alpha-\beta)^{-2j}\right]a & \text{if } \sum_{j=1}^{2l-3}(\alpha-\beta)^{-j} < x < \sum_{j=1}^{2l-1}(\alpha-\beta)^{-j} \\ & \qquad \text{for } l = 2, \dots, k, \\ \left[(\beta-\alpha)^{-1}(\beta^2-\alpha^2)^{-1}\sum_{j=0}^{k-l}(\alpha-\beta)^{-2j}\right]a & \text{if } \sum_{j=1}^{2(l-1)}(\alpha-\beta)^{-j} < x < \sum_{j=1}^{2(l-2)}(\alpha-\beta)^{-j} \\ & \qquad \text{for } l = 3, \dots, k, \\ \left[(\beta-\alpha)^{-1}(\beta^2-\alpha^2)^{-1}\sum_{j=0}^{k-2}(\alpha-\beta)^{-2j}\right]a & \text{if } (\alpha-\beta)^{-1}+(\alpha-\beta)^{-2} < x < 0, \\ \dfrac{a}{\beta-\alpha} & \text{if } 0 < x < \beta-\alpha-1, \\ 0 & \text{elsewhere}, \end{cases}$$

where a is a normalization factor chosen so that

$$\int_{-1}^{\beta-\alpha-1} \varrho(x)\,dx = 1.$$

Proof. While this result severely taxes the notation, its proof is a straightforward verification that (12) is satisfied.

Remarks. (1) An interesting feature of this result is that the systems which satisfy the hypothesis of Theorem 4 do not meet the hypotheses of the general existence result of Lasota and Yorke [8]. Specifically, if

$F(y) = ay+\beta|y|-1$, then $|F'(y)| = a+\beta < 1$ for $x > 0$ and thus F is not almost everywhere expanding. (Misiurewicz has pointed out that if $F(\beta-a-1) = \beta^2-a^2-a-\beta-1 < 0$, then the second iterate of F is everywhere expanding, and the Lasota–Yorke existence result applies to F^2.)

(2) If we let $S = \{x_0, x_1, \ldots, x_{2k+1}\}$ be the discontinuity points of ϱ, then $F(S) \subset S$, and F is monotonic on each (x_i, x_{i+1}). Such functions are called ***Markov maps***, and these have been studied by Bowen, [3].

(3) Note that the density is nonzero almost everywhere on $[-1, \beta-a-1]$ except on the subinterval (a, b), where

$$a = \sum_{j=1}^{2k-1} (a-\beta)^{-j}, \qquad b = \sum_{j=1}^{2k-2} (a-\beta)^{-j}.$$

One can show that any point in this subinterval is mapped into $(-1, a)\cup(b, \beta-a-1)$ by a suitable iterate of F. Moreover, $(-1, a)\cup(b, \beta-a-1)$ is positively invariant, so that all trajectories initiating in (a, b) leave and never return.

EXAMPLE. Suppose $k = 2$, $a = -11/16$, $\beta = 21/16$. Then according to Lemma 1, (2) displays chaos. From Theorem 4 we have

$$\varrho(x) = \begin{cases} 5/6, & -1 < x < 1/2, \\ 2/3, & -1/2 < x < -3/8, \\ 1/3, & -1/4 < x < 0, \\ 5/12, & 0 < x < 1, \\ 0 & \text{elsewhere}. \end{cases}$$

The fact that $\varrho(x) = 0$ on $(-3/8, -1/4)$ reflects the fact that almost all trajectories initiating in this set sooner or later enter the set $(-1, -3/8)\cup(-1/4, 1)$.

Acknowledgements

Many of the results included in this paper were developed with the aid of conversations with Professor R. W. Brockett, and it is a pleasure to acknowledge his valuable input. I would also like to thank Professor J. Zabczyk for the opportunity to speak about some related material at the XVI Semester on Mathematical Control Theory at the Banach Center in Warsaw. Finally, I wish to acknowledge the support of this research by the Basic Energy Sciences Division of the U.S. Department of Energy under Contract DE-AC05-80ER10778.

References

[1] J. Baillieul, R. W. Brockett and R. B. Washburn, *Chaotic motions in non-linear feedback systems*, IEEE Trans. Circuits Systems CAS-27 (1980), 990–997.

[2] G. D. Birkhoff and B. O. Koopman, *Recent contributions to the ergodic theory*, in: *Collected Mathematical Papers of George David Birkhoff*, vol. 2, Dover, New York 1968, 462–465.

[3] R. Bowen, *On Axiom A diffeomorphisms*, Regional Conf. Series in Math. 35, CBMS: Amer. Math. Soc. Providence, R. I. 1978.

[4] P. Collet and J. P. Eckmann, *Iterated Maps on the Interval as Dynamical Systems*, Progress in Physics Series, Birkhäuser, Boston 1980.

[5] T. Erben, P. Everett and P. Johnson, *The simulation of random processes on digital computers with Čebyšev mixing transformations*, J. Computational Phys. 32 (1979), 168–211.

[6] R. E. Kalman, *Nonlinear aspects of sampled-data control systems*, in: Proceedings of Symposium on Nonlinear Circuit Analysis, Polytechnic Institute of Brooklyn, 1956, 273–313.

[7] A. Lasota, *Invariant measures and functional equations*, Aequationes Math. 9 (1973), 193–200.

[8] A. Lasota and J. A. Yorke, *On the existence of invariant measures for piecewise monotonic transformations*, Trans. Amer. Math. Soc. 186 (1973), 481–488.

[9] A. Marotto, *Snap-back repellers imply chaos in R^n*, J. Math. Anal. Appl. 63 (1978), 199–223.

[10] M. Misiurewicz, *Absolutely continuous measures for certain maps of an interval*, Publ. Math. IHES, 1980.

[11] M. Morse, *A one-to-one representation of geodesics on a surface of negative curvature*, Amer. J. Math. 43 (1921), 33–51.

[12] —, *Recurrent geodesics on a surface of negative curvature*, Trans. Amer. Math. Soc. 22 (1921), 84–110.

[13] B. Schweizer and A. Sklar, *Probabilistic metric spaces determined by measure preserving transformations*, Z. Wahrscheinlichkeitstheorie verw. Gebiete 26 (1973), 235–239.

[14] J. von Neumann, *Proof of the Quasi-Ergodic Hypothesis*, Proc. National Acad. Sci. USA 18 (1932), 70–82.