Chaos and Linearity

David Sanders Bennett
Honors Thesis
Mathematics
May 1991

Introduction

There are many ways to define *chaos* in everyday language. For example, synonyms of *chaos* include *disorder* and *turmoil.* Certainly, the term also connotes randomness and complexity. Mathematically, one may define *chaos* in relation to the behavior of *discrete dynamical systems*, i.e., the behavior of operators (functions) under iteration. R. Devaney, in his book *An Introduction to Chaotic Dynamical Systems* [1], gives a rigorous definition for chaos in such systems. Devaney's definition, as we shall see, preserves the everyday notions of chaos. That is, a *chaotic system* by Devaney's definition is characterized by complex, turbulent behavior.

In contrast, when the word *linear* describes mathematical structure, one almost certainly assumes that structure has a simple, orderly character. Indeed, the words *simple* and *orderly* describe the dynamics of a linear operator on a Euclidean space of finite dimension. Devaney himself describes such operators as having "extremely simple dynamics" [1, p. 190]. Moreover, in much of the literature describing chaos and chaotic operators, one often encounters the statement "nonlinearity gives rise to chaos."

It would seem, then, that chaos and linearity are incompatible. Not so! In fact, linear operators on infinite dimensional spaces may be chaotic (although, as the reader will shortly see, linear operators on finite dimensional spaces must fail to be chaotic). I first learned that chaotic linear operators exist at a colloquium talk given by J. H. Shapiro of Michigan State University. Although much of the talk was beyond my understanding at the time, the apparent discrepancy between the simplicity of linearity and the richness of chaotic behavior was intriguing. In the following paper, I will present an example of a chaotic linear operator on an infinite dimensional vector space.

Section 1: Preliminaries

Before moving into a rigorous discussion of chaos, we first introduce some elementary concepts from the study of discrete dynamical systems. Let (X, d) be a metric space (that is, a set on which there is a distance function d, called a *metric*). Let T be an operator mapping X into X. Our study will be concerned with the behavior of the *orbits* of points in X under T.

Definition. Let $x \in X$. The *orbit of x under T*, denoted $\mathcal{O}_T(x)$, is given by $\mathcal{O}_T(x) = \{ T^n x : n = 0,1,2,3, \ldots \}$.

For example, if $S: \mathbf{R} \to \mathbf{R}$ is given by $Sx = -\frac{1}{2}x$, then $\mathcal{O}_S(2) = \{ 2, -1, \frac{1}{2}, -\frac{1}{4}, \frac{1}{8}, \ldots \}$.

We will also consider points of X whose orbits under T are *periodic*; we call such points *periodic points*.

Definition. Let $x \in X$. We say x is a *periodic point (for T)* provided that $T^n x = x$, for some positive integer n. In such a case, we say x *has period n under T*.

For example, the real numbers 0 and -1 are periodic points of period 2 under the operator $Tx = x^2 - 1$. With these basic definitions in place, we can present a rigorous definition for chaos in the context of an operator on a metric space.

We adopt R. Devaney's definition of chaos [1, p. 50]. Let X be a metric space with metric d, and let $T : X \to X$. In the following definition and in the sequel, for a given $x \in X$ and $\varepsilon > 0$, $b(x,\varepsilon)$ denotes the open ball centered at x of radius ε: $b(x,\varepsilon) = \{y \in X: d(x,y) < \varepsilon \}$.

Definition. The operator T is said to be *chaotic on X*, or simply *chaotic*, provided

(a) T has *sensitive dependence on initial conditions*, meaning that there exists a $\delta > 0$ such that, for any $x \in X$ and $\varepsilon > 0$, there exists a $y \in b(x,\varepsilon)$ with $d(T^n x, T^n y) > \delta$, for some n;

(b) the set of periodic points for T is *dense* in X, meaning that given any open ball b in X contains a periodic point for T;

(c)* X contains a *universal vector for T*, meaning that there is an $x \in X$ such that $\mathcal{O}_T(x)$ is dense in X.

The reader can verify, for example, that the operator T on $[0, 2\pi)$ given by $T(x)=2x \bmod 2\pi$ is in fact chaotic [1, p. 50]. Recall that a metric space is *separable* if it has a countable dense subset. The set of real numbers, taken with the standard Euclidian metric, is an example of a separable metric space; the rational numbers constitute a countable dense subset of the reals. Note that if a metric space admits an operator with a universal vector, then the metric space is necessarily separable--the orbit of the universal vector is a countable dense set.

In this paper, we will focus on operators on a special type of metric space, called a Hilbert space. Before defining a Hilbert space, we must first define an *inner product.*

Definition. Let L be a vector space over the real numbers. The function I: $L \times L \to \mathbf{R}$ is an *inner product* on L provided

(a) $I(ag + bg, h) = aI(g, h) + bI(g, h)$, for g,h in L and a,b in **R**;

(b) $I(g, h) = I(h,g)$, for g,h in L;

(c) $I(g,g) \geq 0$ for g in L, and $I(g,g) = 0$ if and only if g is the zero vector.

(for convenience, we often express $I(g,h)$ simply as (g,h))

---

* Requirement (c) given here is equivalent to Devaney's requirement of "Topological Transitivity".

A vector space equipped with an inner product is called an *inner product space.* The reader is already familiar with one example of an inner product space, namely ($\mathbf{R}^n$, •), the n-dimensional Euclidean space with the standard dot or scalar product. Another example is the space C[0,1] of continuous real-valued functions on the unit interval, with inner product $\int_0^1 fg\,dx$, for f and g in C[0,1]. In inner product spaces, the inner product induces an additional structure, called a *norm.*

Definition. A *norm* on a vector space L is a function $\|\ \| : L \to \mathbf{R}$ satisfying:

(a) $\| g \| \geq 0$ for g in L, with $\| g \| = 0$ if and only if g is the zero vector;

(b) $\| g + h \| \leq \| g \| + \| h \|$, for g, h in L;

(c) $\| ag \| = | a | \| g \|$, for g in L and a in **R**.

(Property (b) above is often referred to as the *triangle inequality.*)

As mentioned above, an inner product on a space L induces a norm $\|\ \|$ on L, given by $\| g \| = (g,g)^{1/2}$ for all g in L. The reader can easily verify that this function is in fact a norm (the triangle inequality follows from the Cauchy - Schwartz Inequality: $|(f,g)| \leq \| f \| \| g \|$). We are now ready to introduce the definition of a *Hilbert space.*

Definition. An inner product space X is called a *Hilbert space* provided X is *complete* in the metric d, given by $d(x,y) = \| x - y \|$, where $\|\ \|$ is the norm induced by the inner product. (Recall that a metric space is *complete* provided any Cauchy sequence in the space converges to an element of the space.)

The reader may verify that ($\mathbf{R}^n$, •) is in fact a Hilbert space; the metric here is precisely the standard Euclidean metric. The space C[0,1] with the inner product mentioned above is not a Hilbert space, as completeness fails.

Another example of a Hilbert space is ℓ^2, the collection of all sequences of real numbers whose terms are square summable. Specifically, ℓ^2 is given by

$$\ell^2 = \{(\, x_i\,) : \ x_i \in \mathbf{R} \text{ for } i = 0,1,2,\ldots, \text{ and } \sum_{i=0}^{\infty} x_i^2 < \infty \ \}.$$

The inner product is given by

$$((\, x_i\,), (\, y_i\,)) = \sum_{i=0}^{\infty} (x_i\, y_i).$$

The norm on ℓ^2 is given by $\|\,(\, x_i\,)\,\| = [\sum_{i=0}^{\infty} x_i^2]^{1/2}$, for all $(\, x_i\,)$ in ℓ^2. When convenient, we adopt *functional* notation for a sequence $\mathcal{x} = (\, x_n\,)$; that is x_k, the kth term of $\mathcal{x}$, may be expressed as $\mathcal{x}(k)$. The space ℓ^2 is of particular interest to us since it will be the domain and target space for the chaotic linear operator presented later.

At this point, it is necessary to go deeper into the structure of Hilbert spaces.

Definition. Let S be a subset of a Hilbert space H. S is *orthonormal* provided that, given any two elements s and t in S,

$$(s,t) = \begin{cases} 1 & \text{if } s = t \\ 0 & \text{if } s \neq t \end{cases} .$$

S is said to be a *maximal orthonormal* subset of H provided that if $S \subseteq W \subseteq H$ and W is orthonormal, then S = W.

Definition. Let H be a Hilbert space ($\neq \{0\}$). A subset S of H is an *orthonormal basis for H* provided that S is a *maximal orthonormal* subset of H.

It can be shown that every nonzero Hilbert space has an orthonormal basis (see [2], p. 75). Just as for finite dimensional spaces, the cardinality (order) of any two orthonormal bases of a Hilbert space H is the same. Moreover, the dimension of H, denoted dim H, is equal to the cardinality (order) of any orthonormal basis of H. The dimension of $\mathbf{R}^n$ is n, since an orthonormal basis for $\mathbf{R}^n$ is the set $\{\mathbf{v}_i : i = 1, \ldots, n\}$,where the kth coordinate of $\mathbf{v}_i$, $\mathbf{v}_i(k)$, is given by

$$\mathbf{v}_i(k) = \begin{cases} 1 & \text{if } k = i \\ 0 & \text{if } k \neq i \end{cases}, \text{ for } k = 1, 2, \ldots, n.$$

On the other hand, the dimension of ℓ^2 is countably infinite. To see this, consider the set $\mathcal{B} = \{e_i \in \ell^2 : i = 0,1,2,3, \ldots\}$, where $e_i(k)$ is given by

$$e_i(k) = \begin{cases} 1 & \text{if } k = i \\ 0 & \text{if } k \neq i \end{cases}, \text{ for } k = 0,1,2,3, \ldots.$$

It is clear that $\mathcal{B}$ is a countably infinite set, and it's easy to verify that $\mathcal{B}$ is in fact an orthonormal basis for ℓ^2.

Finally, we must define a *linear operator*, and also present some elementary facts about linear operators on Hilbert spaces.

Definition. Let T be an operator on a vector space X. We say that T is *linear* provided that

(a) $T(x + y) = Tx + Ty$

(b) $T(ax) = aTx$

for any x and y in X and any scalar a.

Here are some examples of linear operators. Let A be any m by n matrix. Representing elements of $\mathbf{R}^n$ and $\mathbf{R}^m$ as column matrices, we may define an operator f mapping $\mathbf{R}^n$ into $\mathbf{R}^m$ by $f(\mathbf{x}) = A\mathbf{x}$. It is easy to verify that f is linear. Now consider the operator B: $\ell^2 \to \ell^2$ defined by $[B\mathfrak{w}](k) = \mathfrak{w}(k+1)$. The operator B is called the *backward shift*: clearly $B(\mathfrak{w}(0), \mathfrak{w}(1), \mathfrak{w}(2), \ldots) = (\mathfrak{w}(1), \mathfrak{w}(2), \ldots)$. Also, we may define a *forward shift* operator F: $\ell^2 \to \ell^2$ by $F(w(0), w(1), w(2), \ldots) = (0, w(0), w(1), \ldots)$ for any w in ℓ^2. That both B and F are linear is also easily verified.

We will show, in a theorem to follow shortly, that the concepts of *continuity* and *boundedness* are equivalent for linear operators on Hilbert spaces. These terms are defined as follows.

Definition. Let H be a Hilbert space. Let T: H $\to$ H be a linear operator.

(a) T is *continuous at a point x in H* provided that given $\varepsilon > 0$, there exists a $\delta > 0$ such that if y is in H and $\| x - y \| < \delta$, then $\| Tx - Ty \| < \varepsilon$. If T is continuous at every pt in H, then we say T is *continuous (on H).*

(b) T is *bounded* provided that there exists a constant M such that $\| Tx \| \leq M \| x \|$ for all x in H. Moreover, for a bounded T, the *norm of T*, denoted $\| T \|$, is defined by $\| T \| = \sup\{\frac{\| Tx \|}{\| x \|}: x \neq 0\}$, or equivalently $\| T \| = \sup\{\| Tx \|: \| x \| = 1\}$.

It should be noted that part (b) the definition above implies that if $\| T \|$ exists, then T is necessarily bounded.(Also, these definitions may be extended to operators taking one Hilbert space into another simply by changing the respective norms.) It is easily verified that $\| B \| = \| F \| = 1$, where B and F are the backward and forward shift operators mentioned above.

Theorem. 1.1. Let H be a Hilbert space. Let T: H $\to$ H be a linear operator. The following statements are equivalent:

(a) T is continuous;

(b) T is continuous at zero;

(c) T is bounded;

PROOF. That (a) implies (b) is clear. We show (b) implies (c): Fix $\varepsilon = 1$. Since T is continuous at 0 and maps 0 to 0, there exists a $\delta > 0$ such that if $\| x \| < \delta$, then $\| Tx \| < 1$. Let x be nonzero. Then, by the linearity of T,

$$\| Tx \| = \| T(\frac{\delta}{2}\frac{x}{\|x\|}) \| (\frac{2}{\delta}\| x \|) < \frac{2}{\delta}\| x \|;$$

the last inequality follows since $\| \frac{\delta}{2}\frac{x}{\| x \|} \| = \frac{\delta}{2} < \delta$. Hence T is bounded (this inequality clearly holds for $x = 0$). Showing (c) implies (a) is easy: Let $x \in H$ be arbitrary. Given $\varepsilon > 0$, choose $\delta = \frac{\varepsilon}{\| T \|}$. Then, if $y \in H$ and $\| y - x \| < \delta$, we have the following:

$$\| Ty - Tx \| = \| T(y - x) \| \leq \| T \| \| y - x \| < \| T \| \delta = \varepsilon. \blacksquare$$

If, in theorem 1.1, H is $\mathbf{R}^n$, an even stronger result is possible. We state it below in the form of a theorem.

Theorem. 1.2. If $T:\mathbf{R}^n \to \mathbf{R}^n$ is linear, then T is continuous.

PROOF. Let $V = \{\mathbf{v}_i: i = 1, \ldots, n\}$ be an orthonormal basis for $\mathbf{R}^n$. Let $\mathbf{z} \in \mathbf{R}^n$ such that $\| \mathbf{z} \| = 1$. We know that $\mathbf{z} = \sum_{i=1}^{n} a_i\mathbf{v}_i$, where $a_i \in \mathbf{R}$ for all i. Then by definition of $\| \;\|$ and orthonormality of V, we have $(\sum_{i=1}^{n} a_i^2)^{1/2} = 1$. Now,

$$\| T\mathbf{z} \| = \| \sum_{i=1}^{n} a_i T\mathbf{v}_i \| \leq \sum_{i=1}^{n} | a_i | \| T\mathbf{v}_i \|.$$

Thus, $\|T\mathbf{z}\| \leq \max\{\|T\mathbf{v}_i\|: i = 1, \ldots, n\}(\sum_{i=1}^{n} | a_i |)$. Applying the Cauchy - Schwarz inequality, we have:

$$\|T\mathbf{z}\| \leq \max\{\|T\mathbf{v}_i\|: i = 1, \ldots, n\}(\sum_{i=1}^{n} a_i^2)^{1/2}(\sum_{i=1}^{n} 1^2)^{1/2} = \sqrt{n}\, \sup\{\|T\mathbf{v}_i\|: i = 1, \ldots, n\}\|\mathbf{z}\|.$$

Hence T is bounded and therefore continuous, by Theorem 1.1. ■

Since *any* finite dimensional Hilbert space is isomorphic to $\mathbf{R}^n$ for some n (establish a one-to-one correspondence between orthonormal basis elements and extend linearly), the result above may be restated as follows.

Theorem. 1.3. If H is a finite dimensional Hilbert space and T: H → H is linear, then T is continuous.

PROOF. H is isomorphic to $\mathbf{R}^n$ for some n. Apply Theorem 1.2. ■

We have now completed our preliminary discussion, and move on to our discussion of chaos and linearity in the context of Hilbert spaces.

Section 2: Chaos and Linearity on Finite Dimensional Spaces

In this section we shall prove that a linear operator on a finite dimensional Hilbert space cannot be chaotic. As mentioned above, any finite dimensional Hilbert space is isomorphic to a Euclidean space $\mathbf{R}^n$, for some n. Hence we restrict our attention to $\mathbf{R}^n$, and move on to our first result.

Proposition 2.1. Any subspace of $\mathbf{R}^n$ is a (topologically) closed subset of $\mathbf{R}^n$.

PROOF. Let Y be a subspace of $\mathbf{R}^n$. The Gram-Schmidt Theorem guarantees the existence of an orthonormal basis $\mathcal{B} = \{b_i : i = 1,2,3,\ldots,m\}$ for Y (note $m \le n$). Let p be any cluster point of Y. There exists a sequence (y_i) in Y such that $y_i \to p$ as $i \to \infty$. It follows that (y_i) is a Cauchy sequence in $\mathbf{R}^n$. Hence, given any $\varepsilon > 0$ there exists an N such that if $j \ge N$ and $k \ge N$, then $\| y_k - y_j \| < \varepsilon$. But (y_i) is a sequence in Y, so, for all i, $y_i = \sum_{l=1}^{m} a_{l\,i}\, b_l$, where each $a_{l\,i}$ is a real number. Thus

$$\| y_k - y_j \| = \| \sum_{l=1}^{m} a_{l\,k} b_l - \sum_{l=1}^{m} a_{l\,j}\, b_l \| = \| \sum_{l=1}^{m} (a_{l\,k} - a_{l\,j}) b_l \| = [\sum_{l=1}^{m} (a_{l\,k} - a_{l\,j})^2\,]^{1/2} < \varepsilon,$$

where the last equality follows from the characterization of the norm in terms of the inner product on $\mathbf{R}^n$. It follows from the inequality above that each $(a_{l\,i})_{i=1}^{\infty}$ is a Cauchy sequence of real numbers (hence convergent). For each l, let a_l be the limit of $(a_{l\,i})$. We claim that $(y_i) \to \sum_{l=1}^{m} a_l b_l$. To see this, for each l choose n_l such that if $j \ge n_l$ then $| a_{l\,j} - a_l | < \frac{\varepsilon}{\sqrt{m}}$. Let $N_0 = \max\{n_1, n_2, n_3, \ldots, n_m\}$. Then if $j \ge N_0$, we have

$$\| y_j - \sum_{l=1}^{m} a_l b_l \| = \| \sum_{l=1}^{m} a_{lj} b_l - \sum_{l=1}^{m} a_l b_l \| = \| \sum_{l=1}^{m} (a_{lj} - a_l) b_l \|$$

$$= [\sum_{l=1}^{m} (a_{lj} - a_l)^2]^{1/2} < [\sum_{l=1}^{m} (\frac{\varepsilon}{\sqrt{m}})^2]^{1/2} = (\varepsilon^2)^{1/2} = \varepsilon .$$

Thus, $(y_i) \to \sum_{l=1}^{m} a_l b_l$ and, by the uniqueness of limits, $p = \sum_{l=1}^{m} a_l b_l \in Y$. So Y contains all its cluster points, and hence is closed. ■

The main result of this section--linear operators on a finite dimensional Hilbert space must fail to be chaotic--is an easy consequence of the following theorem. The theorem will show that if a linear operator on a finite dimensional Hilbert space has a dense collection of periodic points, then in fact all points in the space will be periodic. Once again, we only consider operators on $\mathbf{R}^n$.

Theorem 2.2. Let $T: \mathbf{R}^n \to \mathbf{R}^n$ be a linear operator. Suppose T has a dense collection of periodic points; then there exists a number m such that $T^m = I$, where I is the identity operator on $\mathbf{R}^n$.

PROOF. Let $P = \{p \in \mathbf{R}^n : p \text{ is a periodic point under } T\}$. Note that P is a subspace of $\mathbf{R}^n$ (the reader can verify that the zero element is in P, and also that P is closed under addition and scalar multiplication). By Proposition 2.1, P is a closed subset of $\mathbf{R}^n$. But the closure of a dense subset of a space is the space itself, so $P = \mathbf{R}^n$ (P is dense by hypothesis). Hence each $x \in \mathbf{R}^n$ is a periodic point under T. Let $\mathcal{B} = \{b_i : i = 1,2,3,\dots,n\}$ be a basis for $\mathbf{R}^n$. By the above, each b_i is a periodic point under T ; let each b_i have period m_i. Let m be the least common multiple of $\{m_i : i = 1,2,\dots, n\}$. We show $T^m = I$. Let $x \in \mathbf{R}^n$. Then $x = \sum_{i=1}^{n} a_i b_i$, where each a_i is a scalar. Note $T^m b_i = T^{m_i n_i} b_i = b_i$ (here $n_i = \frac{m}{m_i}$). Thus, by the linearity of T,

$$T^m x = \sum_{i=1}^{n} a_i T^m b_i = \sum_{i=1}^{n} a_i b_i = x. \text{ Thus } T^m = I. \blacksquare$$

As mentioned above, a corollary of this theorem is that a linear operator on a finite dimensional Hilbert space (or Euclidean space) cannot be chaotic.

Corollary 2.3. A linear operator on $\mathbf{R}^n$ cannot be chaotic.

PROOF. Let $T : \mathbf{R}^n \to \mathbf{R}^n$ be linear. Suppose T has a dense collection of periodic points under T. Let $x \in \mathbf{R}^n$. By Theorem 2.2 above, x is a periodic point for T, and $\mathcal{O}_T(x)$ is a finite set. Since no finite set may be dense in $\mathbf{R}^n$, x cannot have dense orbit. Since x was an arbitrary vector in $\mathbf{R}^n$, T cannot have a universal vector; thus T cannot be chaotic. ■

Once again, since any finite dimensional Hilbert space is essentially equivalent to the Euclidean space $\mathbf{R}^n$, we have obtained the promised result. It is possible to show directly that a linear operator on a finite dimensional Hilbert space may not satisfy condition (c) of the definition of chaotic, i.e., such an operator cannot have a vector with dense orbit under T. A proof which shows the incompatibility of conditions (b) and (a) of the definition, rather than (b) and (c) as given above, is also possible. We now move on to an example of a chaotic linear operator on an infinite dimensional Hilbert space.

Section 3: A Chaotic Linear Operator

In this section we show the operator 2B, *twice the backward shift,* is chaotic on ℓ^2. (For any sequence w in ℓ^2, $(2B)w$ is the sequence given by $[(2B)w](k) = 2[w(k+1)]$.) Note that 2B is linear and bounded ($\| 2B \| = 2$); we now proceed to show it is chaotic.

Part 1. Sensitive Dependence on Initial Conditions.

To be called chaotic, 2B must have *sensitive dependence on initial conditions,* part (a) of the definition of a chaotic operator. However, we shall show that 2B actually satisfies a stronger condition. We shall show that given any sequence w in ℓ^2 and $\varepsilon > 0$, there exists a sequence w^* in ℓ^2 such that $\| w - w^* \| < \varepsilon$ and $\lim_{n\to\infty} \| (2B)^n w - (2B)^n w^* \| = \infty$.

First, consider the sequence v given by $v(k) = (\frac{2}{3})^k$. Note that

$$\sum_{k=0}^{\infty} | v(k) |^2 = \sum_{k=0}^{\infty} (\tfrac{2}{3})^{2k} = \sum_{k=0}^{\infty} (\tfrac{4}{9})^k = \tfrac{9}{5},$$

so $v \in \ell^2$. Next, observe that $[(2B)v](k) = 2[v(k+1)] = 2[(\frac{2}{3})^{n+1}] = \frac{4}{3}(\frac{2}{3})^n = \frac{4}{3}[v(k)]$ for all k. Hence $(2B)v = \frac{4}{3}v$, so v is an *eigenvector* for 2B, with *eigenvalue* $\frac{4}{3}$. Also, by the linearity of $(2B)^n$ and the properties of eigenvectors, $(2B)^n(av) = a[(2B)^n v] = a(\frac{4}{3})^n v$ for all positive integers n and scalars a.

Now, let $w \in \ell^2$ and $\varepsilon > 0$ be arbitrary. Choose k such that $\frac{1}{k}\|v\| = \|\frac{1}{k}v\| < \varepsilon$. Next, define $w^* = w - \frac{1}{k}v$ (since ℓ^2 is a vector space $w^* \in \ell^2$). Note $\| w - w^* \| = \| \frac{1}{k} v \| < \varepsilon$. Moreover, for any n,

$$\| (2B)^n w - (2B)^n w^* \| = \| (2B)^n (w - w^*) \| = \| (2B)^n (\frac{1}{k} v) \|$$
$$= \| \frac{1}{k} (\frac{4}{3})^n v \| = | \frac{1}{k} (\frac{4}{3})^n | \, \| v \|.$$

Hence, $\lim_{n \to \infty} \| (2B)^n w - (2B)^n w^* \| = \infty$. Thus 2B has *sensitive dependence on initial conditions.*

Part 2. A Dense Collection of Periodic Points.

Let v_2 be a sequence of period two under 2B. Then $(2B)^{2m} v_2 = v_2$ for all positive integers m. Thus terms of even index satisfy the equation $v_2(0) = 2^{2m} v_2(2m)$, and terms of odd index satisfy the equation $v_2(1) = 2^{2m} v_2(2m+1)$. It follows that v_2 must be of the form

$$v_2 = (v_0, v_1, \frac{v_0}{2^2}, \frac{v_1}{2^2}, \frac{v_0}{2^4}, \frac{v_1}{2^4}, \ldots),$$

where v_0 and v_1 are real numbers. Also, $\| v_2 \|^2 = [\, |v_0|^2 + |v_1|^2] \sum_{i=0}^{\infty} \frac{1}{2^{4i}}$. It is natural to guess that a sequence of period three must be of the form

$$v_3 = (v_0, v_1, v_2, \frac{v_0}{2^3}, \frac{v_1}{2^3}, \frac{v_2}{2^3}, \frac{v_0}{2^6}, \frac{v_1}{2^6}, \frac{v_2}{2^6}, \ldots).$$

Theorem 3.1 below will show that this guess is correct.

We introduce some notation to facilitate the statement of Theorem 3.1. Let m and n be non-negative integers. Define the *quotient of n by m*, denoted q[n/m], to be the largest integer k such that $mk \le n$*. Also define the *remainder of n by m*, denoted r[n/m], by

---

* the function q on **R** corresponds to the standard *greatest integer function*, which is often denoted [x] for x in **R**.

r[n/m] = n - m(q[n/m]). For example, q[13/5] = 2, and r[13/5] = 3. Note that r[n/m] must be an integer between 0 and m-1 (inclusive); additionally, n = m(q[n/m]) + r[n/m] for any positive integer n. The reader may also verify that q[(m + n)/m] = q[n/m] + 1, and that r[(m+n)/m] = r[n/m]. We are now ready to present Theorem 3.1.

Theorem 3.1. Let $\mathbf{v} = (v_0, v_1, \ldots, v_{m-1})$ be any m-tuple of real numbers. Let $\boldsymbol{v}_m \in \ell^2$ be the sequence corresponding to **v** given by

$$\boldsymbol{v}_m(k) = \frac{v_{r[k/m]}}{2^{m(q[k/m])}},$$

for k = 0,1,2,3, ... Then $\boldsymbol{v}_m$ is a periodic point with period m under 2B. Moreover, if $\boldsymbol{w}$ is a periodic point of ℓ^2 under 2B having period m, then $\boldsymbol{w} = \boldsymbol{v}_m$ for some m-tuple **v** of real numbers.

PROOF. Let n be arbitrary. Observe that

$$[(2B)^m\boldsymbol{v}_m](k) = 2^m\boldsymbol{v}_m(m + k) = 2^m\frac{v_{r[(k+m)/m]}}{2^{mq[(m+k)/m]}} = 2^m\frac{v_{r[k/m]}}{(2^m)^{q[k/m]+1}}$$

$$= \frac{v_{r[k/m]}}{2^{mq[k\,/\,m]}} = \boldsymbol{v}_m(k).$$

Hence $\boldsymbol{v}_m$ is periodic under 2B, with period m. It's easily to see that $\boldsymbol{v}_m$ is in ℓ^2. In fact,

$$\|\boldsymbol{v}_m\|^2 = \sum_{k=0}^{\infty} [\boldsymbol{v}_m(k)]^2 = \sum_{i=0}^{\infty} [\sum_{j=0}^{m-1} (\frac{v_j}{2^{im}})^2] = \sum_{i=0}^{\infty} [\frac{1}{2^{2im}}\sum_{j=o}^{m-1} (v_j)^2]$$

$$= \sum_{i=0}^{\infty} \frac{M}{2^{2im}} = M\,[\sum_{i=0}^{\infty} (\frac{1}{2^{2m}})^i\,] = \frac{M}{1 - \frac{1}{2^{2m}}},$$

where $M = \sum_{j=0}^{m-1} (v_j)^2$.

Now let $\boldsymbol{w}$ be a periodic point of period m under 2B. Note we must have $(2B)^{mn}\boldsymbol{w} = [2B]^m\boldsymbol{w} = \boldsymbol{w}$, for all positive integers n. Now, let k be a fixed positive integer. By the above, $w(r[k/m]) = [(2B)^{mq[k/m]} w](r[k/m])$. But, $[(2B)^{mq[k/m]} w](r[k/m]) = 2^{mq[n/m]} w(mq[k/m] + r[m/k]) = 2^{mq[k/m]} w(k)$. Combining these results, we see that for arbitrary n,

$$w(k) = \frac{w_{r[k/m]}}{2^{m(q[k/m])}},$$

where $w_{r[k/m]} = w(r[k/m])$. ■

Having characterized the period points for 2B, we now proceed to show that they constitute a dense subset of ℓ^2.

Theorem 3.2. Let P be the set of periodic points of ℓ^2 under 2B. P is dense in ℓ^2.

PROOF. Let $w \in \ell^2$ and $\varepsilon > 0$ be arbitrary. We show that there exists a sequence v_N in P such that $v_N \in b(w, \varepsilon)$. First, for all m let v_m be the sequence given by

$$v_m(k) = \frac{w(r[k/m])}{2^{m(q[k/m])}}.$$

By Theorem 3.1, v_m is periodic of period m for each m. Now, since the first m terms of v_m are precisely the first m terms of w, from the proof of Theorem 3.1, we have

$$\|v_m\|^2 = \frac{2^{2m}}{2^{2m}-1} \sum_{k=0}^{m-1} w(k)^2 \leq \frac{2^{2m}\| w \|^2}{2^{2m}-1}.$$

Since v_m is periodic of period m,

$$\| B^m v_m \|^2 = \| \frac{1}{2^m}([2B]^m v_m) \|^2 = \frac{1}{2^{2m}} \| [2B]^m v_m \|^2 = \frac{1}{2^{2m}} \| v_m \|^2 \leq \frac{\|w\|}{2^{2m}-1}.$$

Hence, there exists N such that $\| B^N v_N \| < \frac{\varepsilon}{2}$; making N larger if necessary, we may also assume that $\| B^N w \| = (\sum_{k=N}^{\infty} [w(k)]^2)^{1/2} < \frac{\varepsilon}{2}$. Recall that for k = 0, 1, . . ., N, $v_N(k) = w(k)$. Thus,

$$\| w - v_N \| = \| B^N(w - v_N) \| = \| B^N w - B^N v_N \| \leq \| B^N w \| + \| B^N v_N \| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence $v_N \in b(w, \varepsilon)$, and P is dense in ℓ^2. ■

Part 3. A Universal Vector.

Having shown that 2B has a dense collection of periodic points, as well as sensitive dependence on initial conditions, we come to the final requirement necessary to call 2B chaotic in Devaney's sense. We shall show that there exists a sequence w in ℓ^2 which is universal for 2B. It appears that this result was first proved by Rolewicz [5] in 1969. We establish the existence of w by applying a more general theorem due to Kitai [4], and, independently, Gethner and Shapiro [3, Thm. 2.2]. Their result is the following.

Theorem 3.3. Suppose T is a continuous linear operator on a separable Hilbert space H. Suppose there exists a dense subset D of H and a right inverse S for T (TS = identity for H) such that $\|T^n x\| \to 0$ and $\|S^n x\| \to 0$ for all $x \in D$. Then H has universal vectors for T.

REMARK. Theorem 3.3 is a weaker version of Theorem 2.2 in [3].

Before presenting a proof of theorem 3.3, we use the theorem to show that there is a vector $w \in \ell^2$ whose orbit under 2B is in fact dense in ℓ^2. We have noted that 2B is bounded and hence continuous, so our first task is to establish that ℓ^2 is separable. To this end, we construct a set $D \subset \ell^2$ which is both dense and countable. For $i \geq 1$, let $V_i = \{ v \in \ell^2 : v(n)$ is rational for $n < i$, and $v(n) = 0$ for $n \geq i \}$. Now, let $D = \bigcup_{i=1}^{\infty} V_i$.

The countability of D is a result of the countability of the rationals, and the fact that countable unions of countable sets are countable. To see that D is dense, let $w \in \ell^2$ and $\varepsilon > 0$ be arbitrary. Choose k large enough that $\sum_{i=k}^{\infty} [w(i)]^2 < \frac{\varepsilon}{2}$. Next, for $0 \leq n < k$ choose rational numbers q_n such that $|w(n) - q_n| < (\frac{\varepsilon}{2k})^{1/2}$. Let $v \in D$ be given by

$$v(n) = \begin{cases} q_n & \text{if } n < k \\ 0 & \text{if } n \geq k \end{cases} .$$

Then

$$\| w - v \|^2 = \sum_{i=0}^{\infty} |w(i) - v(i)|^2 = \sum_{i=0}^{k-1} |w(i) - q_i|^2 + \sum_{i=k}^{\infty} |w(i) - 0|^2$$

$$< \left(\sum_{i=0}^{k-1} |(\frac{\varepsilon}{2k})^{1/2}|^2 \right) + \frac{\varepsilon}{2} = k(\frac{\varepsilon}{2k}) + \frac{\varepsilon}{2} = \varepsilon.$$

Thus D is dense in ℓ^2.

Twice the backward shift should have *one half the forward shift* as its right inverse: clearly $[(2B)(\frac{1}{2} F)]x = x$ for any $x \in \ell^2$, as desired. Now note that $\|(2B)^n v\| \to 0$ for any v in D since v has only finitely many nonzero terms. Also note that $\| (\frac{1}{2} F)^n v \| = (\frac{1}{2})^n \|v\|$ for all v in ℓ^2 so clearly $\| (\frac{1}{2} F)^n v \| \to 0$ for any v in ℓ^2 (and hence for all v in D). Thus,accepting the validity of Theorem 3.3, the existence of a universal vector for 2B is established.

Hence Theorem 3.3 guarantees that 2B satisfies the third condition necessary for chaos. As the other conditions, i.e. *sensitive dependence on initial conditions* and *the existence of a dense collection of periodic points*, have already been established, the claim that 2B is chaotic hinges on the verification of Theorem 3.3.

Shapiro and Gethner's proof of Theorem 3.3--the proof we present--is based on Baire's Theorem, a fundamental theorem of analysis.* Because Baire's Theorem is not generally included in undergraduate coursework, we include it here with proof. The theorem appears in various forms--the form below is that most suited to our purposes.

Baire's Theorem. Suppose S is a complete metric space. Then the countable intersection of dense open sets in S is dense in S.

PROOF. Let $\{D_i : i = 1,2,3,\ldots\}$ be a family of dense open sets in S. Let $D = \bigcap_{i=1}^{\infty} D_i$. We show that any ball b in S intersects D.

Since D_1 is dense in S, there is a point p_1 in $b \cap D_1$. Now, because finite intersections of open sets are open, $b \cap D_1$ is open. Hence, we can find a positive $\varepsilon_1 < 1$, such that the closure of the ball $b_1 = b(p_1,\varepsilon_1)$ is entirely contained in $b \cap D_1$. Similarly, we can find a new point p_2 and ball $b_2 = b(p_2, \varepsilon_2)$, with $\varepsilon_2 < \min(\frac{1}{2}, \varepsilon_1)$, such that the closure of b_2 is entirely contained in $b_1 \cap D_2$. It is clear that we may continue this process to obtain a sequence (p_n), such that

$$(1)\quad p_n \in \overline{b_n} = \overline{b(p_n,\varepsilon_n)} \quad \text{where } \varepsilon_n < \frac{1}{n}, \text{ and}$$

$$(2)\quad \overline{b_n} \subset b_{n-1} \cap D_n \quad \text{and } \overline{b_1} \subset b \cap D_1.$$

It follows that (p_n) is a Cauchy sequence in S. Since S is complete, (p_n) must converge to a point p in S. By (2), p must be contained in $b \cap D$. ■

---

* Kitai's proof is essentially a generalization of the argument employed by Rolewicz, and does not make use of Baire's Theorem.

We will now take up the proof of theorem 3.3.1, beginning with the following lemma.

Lemma 3.4. Let T be a continuous linear operator T on a separable Hilbert space H. The set of all universal vectors for T (in H) is a countable intersection of open sets in H.

PROOF. Let $X = \{x_i : i = 1,2, \dots \}$ be dense in H. Let $D = \{v \in H: \mathcal{O}_T(v)$ is dense in $H\}$. Let $W_{i,n} = \{y \in H: \| T^k y - x_i \| < \frac{1}{n}$ for some $k = 0, 1, 2, \dots \}$. We shall show that each $W_{i,n}$ is open in H, and that $D = \bigcap_{i,n} W_{i,n}$.

First, note that $W_{i,n}$ is the union over k of the inverse images of the set $b(x_i, \frac{1}{n})$ under T^k. These inverse images must be open sets, since each ball $b(x_i, \frac{1}{n})$ is open and T^k is continuous for all k. Hence each $W_{i,n}$ is a union of open sets, and therefore open.

Next, let $d \in D$ be arbitrary. Since $\mathcal{O}_T(d)$ is dense in H, given any x_i in X and positive integer n there exists a positive integer k such that $\| T^k d - x_i \| < \frac{1}{n}$. Hence $d \in W_{i,n}$ for all positive integers *i* and *n*. It follows that $d \in \bigcap_{i,n} W_{i,n}$ and hence $D \subset \bigcap_{i,n} W_{i,n}$. Now, let $w \in \bigcap_{i,n} W_{i,n}$ be arbitrary, and let b be any ball in H. Since X is dense in H, find x_j in X such that $x_j \in b$. Choose a positive integer N such that $b(x_j, \frac{1}{N}) \subset b$. Since $w \in \bigcap_{i,n} W_{i,n}$, in particular we must have $w \in W_{j,N}$. Thus there is a positive integer k such that $\| T^k w - x_j \| < \frac{1}{N}$. Hence $T^k w \in b$. Thus $\mathcal{O}_T(w)$ intersects b. It follows that $\mathcal{O}_T(w)$ is dense in H, and hence $\bigcap_{i,n} W_{i,n} \subset D$. Thus $D = \bigcap_{i,n} W_{i,n}$. ■

With these two results established, we now present the proof of Theorem 3.3. Again, we follow Shapiro and Gethner.

Theorem 3.3 (restated). Suppose T is a continuous linear operator on a separable Hilbert space H. Suppose there exists a dense subset D of H and a right inverse S for T (TS = identity for H) such that $\|T^n x\| \to 0$ and $\|S^n x\| \to 0$ for all $x \in D$. Then H has universal vectors for T.

PROOF. We know by Lemma 3.4 above that the collection of vectors with dense orbit is a countable intersection of open sets $W_{i,n}$ in H (see proof of Lemma 3.4). Thus if we show each set $W_{i,n}$ is dense, Baire's theorem will guarantee the existence of vectors with dense orbit.*

Fix $W = W_{i,n}$. Let $\varepsilon = \frac{1}{n}$ and $x = x_i$ for convenience. Let $z \in H$ and $\delta > 0$ be arbitrary. We must establish the existence of a $y \in W$ such that $y \in b(z, \delta)$. Since D is dense in H, choose x_0 and z_0 in D such that $\| z - z_0 \| < \frac{\delta}{2}$ and $\| x - x_0 \| < \frac{\varepsilon}{2}$. Since $\|T^n x_0\| \to 0$ and $\|S^n x_0\| \to 0$ for all $x_0 \in D$, we may choose an N large enough that $\|T^N z_0\| < \frac{\varepsilon}{2}$ and $\| S^N x_0 \| < \frac{\delta}{2}$. Let $y = S^N x_0 + z_0$. Then

$$\| y - z \| \leq \| y - z_0 \| + \| z_0 - z \| = \| S^N x_0 \| + \| z_0 - z \| < \frac{\delta}{2} + \frac{\delta}{2} = \delta,$$

so $y \in b(z, \delta)$ as desired. We need only verify that $y \in W$ to complete the proof. Since TS is the identity on H, $T^N S^N$ is also the identity on H, so

$$\begin{aligned} \| T^N y - x \| &= \| T^N(S^N x_0 + z_0) - x \| = \| T^N S^N x_0 - x + T^N z_0 \| \\ &\leq \| x_0 - x \| + \| T^N z_0 \| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Thus $y \in W$, and W is dense. Since W was arbitrary, each $W_{i,n}$ is dense. Baire's Theorem yields the result. ■

---

* H is complete by definition, so Baire's Theorem applies.

We should remark that the existence of a single universal vector implies the existence of a dense collection of universal vectors. Suppose x is a universal vector for an operator T so that $\mathcal{O}_T(x)$ is dense. It follows easily from the definition of $\mathcal{O}_T(x)$ that for all n, $T^n x$ is also universal. Thus $\mathcal{O}_T(x)$ is actually a dense collection of universal vectors. As a final result, we make use of this fact, noting that the existence of a universal vector establishes sensitive dependence initial conditions.

Theorem 3.5. Let H be a Hilbert space, and let $T: H \to H$ be linear. If T has a universal vector d, then T has sensitive dependence on initial conditions.

PROOF. Let $d \in H$ have dense orbit. Fix $\delta > 0$. Let $x \in H$ and $\varepsilon > 0$ be arbitrary. By the density of $\mathcal{O}_T(d)$, there is an N such that $T^N d \in b(0,\varepsilon)$. Note that $y = x + T^N d$ is in $b(x,\varepsilon)$. By the remark above $\mathcal{O}_T(T^N d)$ is dense, hence unbounded. Choose n such that $\|T^n T^N d\| > \delta$. Then,

$$\| T^n y - T^n x \| = \| T^n(x + T^N d) - T^n x \| = \| T^n x + T^n T^N d - T^n x \| = \| T^n T^N d \| > \delta. \blacksquare$$

Hence the condition of sensitive dependence for 2B follows immediately from the results of this subsection. We included the argument for sensitive dependence found in Part 2 above because it does not rely on the Baire's Theorem and the other results of Part 3, and as such may be more accessible to a reader having less mathematical experience. Both proofs of sensitive dependence are due to Joel Shapiro.

References

1. R. L. Devaney, *An introduction to chaotic dynamical systems*, Addison-Wesley, New York, 1989.
2. R. G. Douglas, *Banach algebra techniques in operator theory*, Academic Press,

New York, 1972.

3. R. M. Gethner and J. H. Shapiro, *Universal vectors for operators on spaces of holomorphic functions,* Proc. Amer. Math. Soc. **100** (1987), 281-288.
4. C. Kitai, *Invariant closed sets for linear operators*, Thesis, University of Toronto, 1982.
5. S. Rolewicz, *On the orbits of elements*, Studia Math. **33** (1969), 17-22.