International Journal of Bifurcation and Chaos, Vol. 6, No. 2 (1996) 219–249


# CLARIFYING CHAOS: EXAMPLES AND COUNTEREXAMPLES

RAY BROWN
*Applied Chaos Technology Corporation,*
*P.O. Box 1608, Arlington, VA 22210, USA*

LEON O. CHUA
*Department of Electrical Engineering and Computer Sciences,*
*University of California, Berkeley, CA 94720, USA*

Received June 5, 1995; Revised August 16, 1995

Over the past fifteen years there have been various attempts to define chaos. In an effort to find a universally acceptable definition we began constructing new examples of chaotic systems in the hope that the salient features of chaos could be captured. Our efforts to date have failed and the examples we have constructed seem to suggest that no such definition exists. However, these examples have proved to be valuable in spite of our inability to hone a universal definition of chaos from them. Consequently, we present this list of examples and their significance. Some interesting conclusions that we can draw from them are: It is possible to construct simple closed form solutions of chaotic one-dimensional maps; sensitive dependence on initial conditions, the most widely used definition of chaos, has many counterexamples; there are invertible chaotic dynamical systems defined by simple differential equations that do not have horseshoes; three important properties that are thought to characterize chaos, continuous power spectral density, exponentially sensitive dependence on initial conditions, and exponential loss of information (Chaitin's concept of algorithmic complexity), are independent.

Chaos seems to be tied to our notion of rates of divergence of orbits or degradation of information such as is found in systems with positive Lyapunov exponents. The reliance on rates seems to open the door to a pandora's box of rates, both higher and lower than exponential. The intuitive notion of pseudo-randomness, a practical feature of chaos, is present in examples that do not have positive Lyapunov exponents. And in general, nonlinear polynomial rates of degradation of information are also quite "unpredictable".

We conclude that it appears that for any given definition of chaos, there may always be some "clearly" chaotic systems which do not fall under that definition, thus making chaos a cousin to Gödel's undecidability.

## 1. Introduction

This paper originated in our attempt to answer several fundamental questions about chaos by constructing examples or counterexamples.

The questions we sought to answer are[1]

**Is it possible to write down in closed form (without the use of integrals and derivatives) and in terms of elementary algebraic and transcendental functions (trigonometric and exponential functions), the solutions of some chaotic dynamical systems?**

[1]Throughout this paper the term *chaos* is used in the sense of at least one of the typical features widely adopted as chaotic in the literature.

The answer is yes and we present a large class of examples of noninvertible dynamical systems whose time series can be written down in closed form in terms of elementary functions. Further, the construction can be extrapolated to produce examples in any number of dimensions. Non-invertibility seems to be the price one pays to be able to carry out these constructions. This fact suggests the interesting possibility that for any chaotic dynamical system there may be a projection onto a noninvertible chaotic dynamical system for which a closed form solution does exist.

### Does sensitive dependence on initial conditions characterize chaos?

We ask this question because sensitive dependence on initial conditions is widely thought of as a synonym for chaos. But what is meant by this notion varies between researchers. In order to bring some discipline to the use of this notion we start from the original definition and construct counterexamples to show that the answer is no, even when the orbits separate exponentially and the system is bounded. Further, there are systems for which all solutions are almost periodic which have sensitive dependence on initial conditions if the original definition is used. Also, if boundedness is relaxed, there are simple linear non-hyperbolic dynamical systems with sensitive dependence on initial conditions.

### What is the relationship of zero autocorrelation (ZA), sensitive dependence on initial conditions (SD), and exponential loss of information (EL) in dynamical systems?

ZA is thought of as a property of highly random processes such as stationary white noise, hence it is a legitimate candidate for defining chaos. Generally, we consider chaos to be associated with a decaying autocorrelation function, see Houlton and May in Mullin [1994]. However, we only use the most extreme case, i.e., ZA, in our examples, since this represents the most difficult constraint to work with. We present seven examples that illustrate the independence of these three properties. Most importantly ZA does not imply either SD or EL and neither EL nor SD implies ZA.

### Is there an algebraic, as opposed to geometric, method of identifying chaotic dynamical systems?

We ask this question because of the serious difficulties encountered in proving that a system is chaotic by using an existing theory. It would clearly be simpler to determine chaos from the algebraic forms of equations, if it were possible. However, what do we look for? Our examples answer the most basic aspects of this question by constructing examples of maps from one- and two-sided shifts, thus assuring that they are chaotic.

### Is there an algebraic means of relating chaos to forced oscillators?

Since forced oscillators are a common form in which we encounter chaos, is there some way to see from the form of the equations when a forced oscillator produces chaos? This question originates with Hirsch [1985], p. 192. In response to this question, we show how to relate shifts to twist-and-flip maps. The point in showing that twist-and-flip maps can represent shifts is that twist-and-flip maps are the simplest closed form Poincaré maps derivable from forced oscillators and so provide a very good starting point for answering this question. As shown in Brown & Chua [1993] most of the well-known chaotic maps studied, such as the Hénon map, can be derived from a set of induced nonlinear maps (such as twist maps) by means of composition and that these composition maps have a suggestive relationship to the shift map. These examples are intended to suggest that there may be a theory in which maps may be proven to be chaotic by factoring them into simple well-known components rather than going through the elaborate process of applying the Smale–Birkhoff theorem or related theorems that prove the existence of subshifts of finite type.

### How does a random process differ from a chaotic process and what is the relationship between chaos and pseudo-random processes?

Chaotic processes and random processes are distinct. Chaotic processes can be defined by algorithms, random processes cannot. We consider pseudo-random processes to be chaotic processes and present some history of the evolution of the concept of pseudo-random number generators to draw out this relationship. Once this is done, the relationship between these concepts suggests ways of broadening the concept of chaos to include divergence of orbits more extreme than exponential. Along the way we conclude that cryptography is the study of chaotic maps restricted to periodic

orbits. In a practical vein, we show how to construct pseudo-random number generators and stochastic processes from chaotic systems.

**Should chaotic systems be considered to be bounded?**

Many researchers insist that only bounded systems be considered when studying chaos. However, the measure preserving Hénon map is unbounded, [Devaney, 1989]. There are many other examples. Hence, in defining chaos, no restrictions as to boundedness is reasonable.

## 1.1. *Existing examples and definitions*

Due to the widespread efforts of many scientists to construct a suitable definition of chaos we include some comments about existing examples and definitions that have shaped our thinking.

There are many interesting examples that have been set forth by various authors in the past, especially the example of Grebogi *et al.* [1984] of a non-chaotic strange attractor. We do not attempt to include any existing examples. The examples presented here have been developed by the authors, and hopefully we have not inadvertently reproduced anyone's example for which they should receive credit.

There are numerous definitions of chaos that are commonly in use. For example, a dynamical system is chaotic when:

1. It has a Smale horseshoe [Ozorio de Almeida, 1988].
2. It has positive Kolmogorov entropy [Schuster, 1988].
3. It has positive topological entropy [Katok, 1980].
4. It has a positive Lyapunov exponent [Gulick, 1992].
5. Its sequences have positive algorithmic complexity [Ford, 1986].
6. It has a dense set of periodic orbits, is topologically transitive, and has sensitive dependence on initial conditions [Devaney, 1989].
7. It has sensitive dependence on initial conditions and is topologically transitive [Wiggins, 1992].
8. The power spectral density of related time-series has a component which is absolutely continuous with respect to Lesbegue measure [Bergé *et al.*, 1984].
9. A statistically oriented definition of Shil'nikov [1994].

There are more definitions and there are definitions specialized for non-invertible systems (i.e., negative Schwartzian derivatives) for one-dimensional maps. We have not organized these examples around existing definitions of chaos; there are two concepts that we believe are at the heart of our understanding of chaos: (1) rapid loss of correlation between the future and past; (2) rapid loss of information over time. Each of the above definitions entails these concepts in some fashion. All definitions involving entropy are definitions about information. Positive entropy may be thought of as negative information or a loss of information. Sensitive dependence on initial conditions is meant to suggest that small errors in initial conditions (information) extrapolate to large errors very fast, thus information is lost rapidly. Algorithmic complexity is also inseparable from information. The presence of positive Lyapunov exponents is equivalent to some form of exponential loss of decimals during round off errors. In all cases we intend to convey the notion of a rapid loss of information in some form that prevents the future from being predicted from the past.

But what does "rapid" loss of information mean in the context of chaos. Most researchers would say that exponential rates or losses is what we mean by rapid. However, over arbitrarily long, but fixed time scales, polynomial loss of information can exceed exponential, even though exponential losses will eventually exceed polynomial losses in the long run. Hence, there is an element of philosophy that comes with the concept of chaos and we discuss these philosophical issues in the last section.

This paper has evolved from our efforts to obtain a universally acceptable definition of chaos by answering specific questions with examples. In general, we began with a shift in some coordinate system and then used the shift to construct functions of the shift and maps which were compositions involving the shift. Doing this assured us that the resulting examples would usually be chaotic by any definition. This approach does have some hazards. It is possible to construct a nonchaotic map by composing a nonchaotic map with a chaotic map. This at first seems unreasonable. However, in Brown & Chua [1993] we presented a theoretical frame work to build up chaotic maps from nonchaotic components. Reversing this process leads us to the formation of nonchaotic maps from the composition of a chaotic map and a nonchaotic map. This is not new and the seeds for this idea comes from

Devaney [1984] and Devogelaere [1959]. In particular, it is possible to obtain chaotic maps from the composition of periodic maps. If such a composition is written as

$$\Phi = P \circ Q$$

where $P^2 = Q^2 = I$, where $I$ is the identity map, then $P \circ \Phi = Q$, which is not chaotic. This is the situation for many non-integrable Hamiltonian systems having chaotic solutions.

It seems unlikely that we can construct invertible examples of chaotic maps having closed form solutions in terms of elementary functions. In particular, such maps often have both periodic and chaotic solutions, hence a single "formula" for its solution would have to reflect this fact, and this looks impossible.

### Specifying an Ideal Definition

An ideal definition has at least the following features:

1. It is simple.
2. It is easy to apply by researchers in the field.
3. It contains the essential features of the phenomenon defined.
4. It is possible to derive all important and widely recognized features of the phenomenon from the definition.
5. It includes all widely recognized examples of the phenomenon.
6. It excludes anything that is widely recognized as not being an example of the phenomenon.

All definitions of chaos suffer from some defects but the most serious is that the various definitions cannot be derived from each other. Our examples suggest that this is to be expected and that a definition meeting the above criteria is likely out of reach.

## 2. Closed Form Solutions of Chaotic Equations

It is often thought that chaotic dynamical systems is a subset of those systems for which it is not possible to express a closed form solution (no integrals and derivatives) in terms of elementary functions (polynomials and elementary exponential and trigonometric functions). This requirement excludes any function that explicitly involves addition modulo 1, such as the well known example $T(x) = 2x \bmod(1)$ and its two-dimensional analog, the cat map.

In this section we show that it is possible to construct such examples. However, all examples are noninvertible,[2] such as the logistic map and the tent map.

The general construction we give is based on the duplication formula that exists for elliptic curves [Silverman, 1986]. A sketchy note by Ulam and Von Neumann, [1947] in the *Bulletin of the American Mathematical Society*, November 1947, page 1120, may allude to the possibility of the following exposition.

### 2.1. *One-dimensional maps*

Given that the one-sided shift

$$S_a(x) = 2x \bmod(a)$$

where $a > 0$, is a universal paradigm for chaos in any definition, any function of the iterates of this map must be considered as a *candidate* for a chaotic function, so long as it is not constant almost everywhere. We now consider some examples of such functions. We will need the following formulas:

$$(Kx)\bmod(Ka) = K(x \bmod(a))$$

where $a$, $K > 0$. If $f$ is a periodic function on the real line with period $K$ we have

$$f(Kx) = f(Kx \bmod(K)) = f(K(x \bmod(1)))$$

The significance of these formulas is that

$$f(K2^n x_0) = f(K(2^n x_0 \bmod(1)))$$

and so sequences of the form $y_n = f(K2^n x_0)$ are good candidates for solutions of chaotic finite difference equations.

**Example 1.** Closed form solution of the logistic equation for $\lambda = 4$ in terms of elementary functions.

The logistic equation is

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

We only consider the case where $\lambda = 4$.

Consider the sequence

$$y_n = \sin^2(2^n C \bmod(1)\pi) = \sin^2(2^n C\pi)$$

[2]This observation is generic in electronic circuits since locally-passive [Chua, 1980] electronic circuits where all elements are described by invertible (strictly monotone increasing) characteristics cannot oscillate, let alone become chaotic.

where $0 \leq C \leq 1$ is an arbitrary constant determined by the initial conditions. We have

$$y_{n+1} = 4\sin^2(2^n C\pi)\cos^2(2^n C\pi)$$

or

$$y_{n+1} = 4y_n(1 - y_n)$$

and hence $y_n$ is the closed form solution of the logistic equation with parameter value equal to 4. This example shows that the solution of the logistic equation is a function of the one-sided shift, but is not topologically conjugate to a shift.

Let us generalize this example:

**Example 2.** The elliptic logistic equation.

Now consider the sequence

$$y_n = \mathrm{sn}^2(2^n C\,K)$$

where sn is the Jacobi elliptic sine and $K$ is half the period of $\mathrm{sn}(t,\, k)$ and the parameter $k$ is such that $0 < k < 1$. A direct computation gives the finite difference equation

$$y_{n+1} = 4y_n(1 - y_n)\left(\frac{1 - k^2 y_n}{(1 - k^2 y_n^2)^2}\right)$$

which can be called the elliptic logistic equation.

If we consider $2x \bmod(a)$ to be a pseudo-random number generator, then the iterates of the logistic and elliptic logistic finite difference equations are pseudo-random samples of the functions $\sin^2(t)$ and $\mathrm{sn}^2(t,\, k)$ over one period, i.e., half the period of sin and sn.

**Example 3.** Random samples of the tangent, cotangent, and cosine functions.

Now consider the sequence

$$y_n = \tan(2^n C)$$

This sequence satisfies the finite difference equation

$$y_{n+1} = \frac{2y_n}{1 - y_n^2}$$

which must be considered a chaotic dynamical system.

The sequence

$$y_n = \cot(2^n C)$$

leads to the equation

$$y_{n+1} = \left(y_n - \frac{1}{y_n}\right)\Big/ 2$$

and the sequence

$$y_n = \cos(2^n C)$$

leads to the equation

$$y_{n+1} = 2y_n^2 - 1$$

Multiple angle formulas also work. Let

$$y_n = \cos(3^n C)$$

then

$$y_{n+1} = 4y_n^3 - 3y_n$$

and in general if

$$y_n = \cos(k^n C)$$

then

$$y_{n+1} = P(y_n)$$

where $P(x)$ is a polynomial.

All of these finite difference equations provide a pseudo-random number generator for obtaining a pseudo-random sample of random variables having the distributions arctan, arccot, and arccos, respectively. The distribution of arctan is similar to a Gaussian distribution and can be used as such in some cases.

The preceding examples all arise from the duplication formula for the trigonometric functions and the elliptic sine. A duplication formula exists for any elliptic function and so we may obtain a closed form solution for any chaotic equation arising from the use of a duplication formula as above. To illustrate this, we use the elliptic function of Weierstrass, $y^2 = 4x^3 - g_2 x - g_3$, [Silverman, 1986] in our next three examples.

**Example 4.** Equation from the Weierstrass function, $g_2 = 0.0$, $g_3 = 1.0$.

For this case, we have the duplication formula from Abramowitz & Stegun [1965], p. 654:

$$\wp(2x) = \frac{\wp(x)(\wp(x)^3 + 2)}{(4\wp(x)^3 - 1)}$$

Taking

$$y_n = \wp(2^n C)$$

we get the finite difference equation

$$y_{n+1} = \frac{y_n(y_n^3 + 2)}{4y_n^3 - 1}$$

**Example 5.** An equation from the Weierstrass function, $g_2 = 1.0$, $g_3 = 0.0$.

For this case, we have the duplication formula from Abramowitz & Stegun [1965], p. 659:

$$\wp(2x) = \frac{(\wp(x)^2 + 1/4)^2}{\wp(x)(4\wp(x)^2 - 1)}$$

Taking

$$y_n = \wp(2^n C)$$

as before we get the finite difference equation

$$y_{n+1} = \frac{(y_n^2 + 1/4)^2}{y_n(4y_n^2 - 1)}$$

Note that these maps are not on the unit interval. By using the derivative of $\wp(x)$ we obtain the following phase plane mapping.

**Example 6.** An equation from the Weierstrass function, in the phase plane for $g_2 = 0.0$, $g_3 = 1.0$.

For this case, we have the duplication formula from Abramowitz & Stegun [1965] for $\wp'(x)$, p. 659:

$$\wp'(2x) = \frac{2\wp(x)^6 - 10\wp^3(x) - 1}{\wp'(x)^3}$$

Taking $y_n$ as before and $x_n$ as

$$x_n = \wp'(2^n C)$$

we get the finite difference equation

$$\begin{pmatrix} y_{n+1} \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} \dfrac{(y_n^2 + 1/4)^2}{y_n(4y_n^2 - 1)} \\ \dfrac{2y_n^6 - 10y_n^3 - 1}{x_n^3} \end{pmatrix}$$

**Example 7.** General Duplication Formulas.

If

$$f(2x) = F(f(x))$$

and

$$y_n = f(2^n C)$$

then

$$y_{n+1} = F(y_n)$$

If

$$f(3x) = F(f(x))$$

and

$$y_n = f(3^n C)$$

then

$$y_{n+1} = F(y_n)$$

In general, if

$$f(kx) = F(f(x))$$

and

$$y_n = f(k^n C)$$

then

$$y_{n+1} = F(y_n)$$

The process can be reversed in appropriate cases. Given

$$y_{n+1} = F(y_n)$$

we assume that the solution is given by

$$y_n = f(k^n C)$$

and conclude that

$$f(kx) = F(f(x)) \tag{1}$$

If the solution of the finite difference equation actually does arise from a multiple angle formula and $F$, $f$ have all derivatives then a McLauren's series for $f$ is obtainable. In particular $f(0)$ is a fixed point of $F$, and $kf'(0) = F'(f(0))f'(0)$ means that either $f'(0) = 0$ or is an arbitrary constant and that $F'(f(0)) = k$ is a consistency condition that must be satisfied for such a solution to exist. All other derivatives are obtained by differentiating Eq. 1 and evaluating it at $x = 0$.

**Example 8. The Tent Map.**

Let

$$y_n = \frac{\arccos(\cos(2^n C))}{\pi}$$

where the arccosine is the invertible function that maps the interval $[-1, 1]$ to the interval $[0, \pi]$. A direct computation shows that this sequence satisfies the equation

$$y_{n+1} = 1 - |2y_n - 1|.$$

## 2.2. *Noninvertible maps of the plane*

The following map is a variation of an example that appears in a classical text on advanced calculus, see R. C. Buck, [1956], p. 274.

**Example 9.** A chaotic map in the plane.

Consider the mapping of the plane given by

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \dfrac{x^2 - y^2}{r} \\ \dfrac{2xy}{r} \end{pmatrix}$$

where $r = \sqrt{x^2 + y^2}$.

The closed-form solution for initial conditions $(x_0, y_0)$ is

$$T^n\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} r_0 \cos(2^n\theta_0) \\ r_0 \sin(2^n\theta_0) \end{pmatrix}$$

where $r_0 = \sqrt{x_0 + y_0}$ and $\tan(\theta_0) = y_0/x_0$.

This map is chaotic since it is a function of the one-sided shift. To see this change to polar coordinates where $T$ is given by

$$T\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} r \\ 2\theta \bmod(2\pi) \end{pmatrix}$$

**Example 10**

Consider the mapping of the plane given by

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos(r\theta) \\ r \sin(r\theta) \end{pmatrix}$$

where $r = \sqrt{x^2 + y^2}$ and $\theta = \arctan(y/x)$. The orbits of this map are functions of the orbits of a one-sided shift of the form $r_0^n\theta_0 \bmod(2\pi)$. In particular

$$T^n\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos(r^n\theta) \\ r \sin(r^n\theta) \end{pmatrix}$$

For $r_0 \geq 2$ these mappings have all of the chaotic properties of the mapping of the complex unit circle $z \to z^2$ [Beardon, 1991].

## 3. Sensitive Dependence on Initial Conditions

In this section we examine the role of sensitive dependence on initial conditions in defining chaos. We begin with a definition which was originated by Guckenheimer.

**Definition.** The mapping T : S → S is said to have sensitive dependence on initial conditions if there exist a number $\tau > 0$ such that for all $x \in$ S and any neighborhood, $U$ of $x$, there exist a $y \in U$ and $n > 0$ such that $d(\mathrm{T}^n(x), \mathrm{T}^n(y)) > \tau$ [Devaney, 1988].

In this definition $d(x, y)$ represents the distance between the points $x$ and $y$. This definition says that there is some constant (call it a separation constant) $\tau$ such that for any point $x \in S$ and any neighborhood $U$ of $x$, there is some point $y \in U$ (usually we think of $y$ being very close to $x$) that will eventually move away from $x$ (not necessarily permanently) by an amount $\tau$.

**Example 11.** A linear system.

Consider the linear system

$$\dot{x} = x$$

whose solution is

$$x(t) = x_0 \exp(t)$$

For $t = \ln(2)$, it defines the linear one-dimensional map $\mathrm{T} : \mathbf{R} \to \mathbf{R}$ given by:

$$\mathrm{T}(x) = 2x$$

Let $x \in \mathbf{R}$ and choose $\tau = 1$. Given $\varepsilon > 0$, choose any $y$ with $\|x - y\| < \varepsilon$. Then the distance between the $n$th iterates of $x$ and $y$ is $2^n\varepsilon$. By choosing $2^n\varepsilon = \tau$ we can solve for the integer $n$ needed to verify that the definition is satisfied, and thus conclude that T has sensitive dependence on initial conditions. Note that the forward time autocorrelation of $x(t)$ is 1.

**Example 12.** A bounded system in the plane.

Consider the simple twist, T, [Brown & Chua, 1991] with a fixed point at the origin defined on the plane where a disk of radius $\varepsilon$ about the origin has been removed. T is the time-one map determined by the differential equations:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} -ry \\ rx \end{pmatrix}$$

where $r = \sqrt{x^2 + y^2}$. We repeat that we require that $x^2 + y^2 > \varepsilon^2$ for some fixed but arbitrary parameter, $\varepsilon$, to meet the criteria that a disk of radius $\varepsilon$ has been removed from the plane.

T has sensitive dependence on initial conditions since neighboring solutions of these equations rotate

about the origin at different angular velocities and so must eventually separate by an amount $\tau = 2\varepsilon$, the diameter of the disk that has been removed from the plane. In this example, every orbit is confined to an invariant circle with the center at the origin. On each invariant circle T is a rotation, which is ergodic [Walters, 1982] but T is not ergodic on its domain of definition. On each invariant circle every orbit forms a sequence which is almost periodic. The time series for each solution is almost periodic as well. There are no hyperbolic fixed points for T, or positive Lyapunov exponents. In short, T has no chaotic properties by any other definition.

**Example 13.** A simple system on the torus.

Consider the torus in three dimensions centered at the origin, where $R$ is the radius from the origin to the center of the torus, and the radius of a vertical cross section is 1:

$$\begin{pmatrix} x(\theta, \phi) \\ y(\theta, \phi) \\ z(\theta, \phi) \end{pmatrix} = \begin{pmatrix} (R - \cos(\theta))\cos(\phi) \\ (R - \cos(\theta))\sin(\phi) \\ \sin(\theta) \end{pmatrix}$$

Measuring the angle $\theta$ from the origin, we define a rotation around each cross section circle located at the angle $\theta$ as follows: If $x$ is a point on the cross section circle located at angle $\theta$, rotate $x$ around this circle by an angle $\theta$. Hence, the rotation in the direction of $\phi$ depends on the angle $\theta$. For any point on the torus, and any neighborhood about this point, there is a point nearby on a different cross section circle that rotates at a different angular rate and hence the two points must eventually separate by a distance of $\tau = 2$, twice the radius of the cross section of the torus. Each orbit is almost periodic with no two orbits having the same almost period. This system cannot have any chaos for the same reason as the preceding system.

**Example 14.** Linear without exponential loss of information.

The map

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x \\ x + y \end{pmatrix}$$

is linear and has sensitive dependence on initial conditions since for any point $(x, y)$ there is a point within epsilon from which it separates. This is because after $n$ iterations of this map on the difference of the two points (because our map is linear their difference is $(\varepsilon, \delta)$) we have

$$\begin{pmatrix} \varepsilon \\ \varepsilon + n\delta \end{pmatrix}$$

The length of this vector determines the distance the two points have separated after $n$-iterations which is at least the distance $n\delta$. Since this map preserves the integer lattice it may be considered as a map on the torus and as such is equivalent to Example 13.

The conclusion of these examples is that sensitive dependence on initial conditions is not sufficient to define chaos, even with the exponential loss of information.

**Example 9.** Revisited.

If in Example 9 of the previous section we apply the formal definition of sensitive dependence on initial conditions we see that if we choose our reference point as the fixed point $(0, 0)$, than there does not exist any $\tau$ for which the definition can work. The problem is that the definition is formulated in terms of a metric. What is true of this example is that although the terms of the sequence

$$d(T^n(0, 0), T^n(x, y))$$

are constant, the angular coordinate of $T^n(x, y)$ is uncorrelated. Given any metric on a compact manifold we can construct a similar example.

## 4. Relationship of Three Properties of Complex Dynamical Systems

In this section we present examples which show the extent to which three important properties of complex dynamical systems are independent. These properties are: Zero autocorrelation (ZA), sensitive dependence on initial conditions (SD), and exponential loss of information (EL). One definition of chaos requires the power spectral density (Fourier transform of the average autocorrelation function) to have a component that is absolutely continuous with respect to the Lesbegue measure [Bergé *et al.*, 1984]. This definition arises from the analogy of chaos to random processes (the relationship between autocorrelation and continuous spectrum in random processes can be noted in such well-known

references as [Doob, 1953], and [Papoulis, 1984]). In particular, if the autocorrelation, $R(\tau)$ of a process decays fast enough so that

$$\int_{-\infty}^{\infty} |R(\tau)| d\tau < \infty$$

then the Fourier transform of $R(\tau)$ is uniformly continuous by a well known theorem, see [Katznelson, 1976], p. 121, Theorem 1.2. For stationary random processes, which is the type that measure preserving dynamical systems on compact sets define, ergodic theory applies and we know that the coordinate functions (time series) of a strong mixing transformation have an autocorrelation which converges to 0 at infinity, see Walters [1982], p. 45, Theorem 1.23(iii)(3). When the dynamical system is a flow we can define the power spectrum and by our previous comments, it must have a continuous component. Weak mixing transformations can have continuous power spectra with the autocorrelation converging to 0 in an average sense [Cornfeld *et al.*, 1982], p. 29, Theorem 2(iii). If we do not have a flow, then Theorem 3 of this reference suggests another definition for chaos in terms of the Fourier coefficients. In short, the theory of stochastic processes and, a related cousin, ergodic theory, suggest that the autocorrelation or its Fourier transform can be used to define chaos in analogy with general random processes.

Our examples omit the computation of transforms and consider only the average autocorrelation function in forward time. For unbounded examples we use the limiting correlation coefficient. It should be noted that some authors use the convolution in their definition of autocorrelation in order to get a direct simple relation to the Fourier transform through the Wiener–Khintchine theorem, See Hsu [1984], p. 121, Eq. (7.28). Nothing we have said depends on doing this.

The use of the linear correlation coefficient as a measure of chaos is widespread in practice. For example, see the article of Houlton and May, from Mullin [1994], p. 155. In our examples we use Eq. (7.3) from this reference to compute the autocorrelation. If the autocorrelation coefficient is 0 for positive time and non-zero for $t = 0$ for signals having a mean of 0, the analogy with white noise still holds. (For signals without zero mean value we subtract the mean and then apply the definition).

Some chaotic systems have an exponentially decaying autocorrelation, see Houlton and May from Mullin [1994], Fig. 7.2, however we only consider the extreme case of zero autocorrelation in these examples.

For completeness, we must cover seven cases. They are: (1) ZA, without SD and EL; (2) SD, without ZA and EL; (3) EL without SD and ZA; (4) ZA and SD without EL; (5) ZA and EL without SD; (6) SD and EL without ZA; (7) ZA, SD, EL combined.

Case 7 is Example 1; Case 2 is Example 12; Case 6 is Example 11. The remaining four cases are treated in the following sections.

### 4.1. *Autocorrelation and chaos*

**Example 15.** ZA without SD or EL (Case 1)

Let our system be given by the following time-varying linear differential equations:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} -2yt \\ 2xt \end{pmatrix}$$

The general solution is given by:

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} x_0 \cos(t^2) - y_0 \sin(t^2) \\ y_0 \cos(t^2) + x_0 \sin(t^2) \end{pmatrix}$$

In this example, the average autocorrelation is 0 everywhere except at 0, but there is no sensitive dependence on initial conditions due to the fact that the equations are linear and bounded. The absence of EL follows directly from the form of the time series equations.

**Example 16.** ZA and SD without EL (Case 4)

Let our system be given by the equations:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} -2ryt \\ 2rxt \end{pmatrix}$$

where $r = \sqrt{x^2 + y^2}$ and we delete a disk about the origin of radius $\varepsilon$. The general solution is given by the equations:

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} x_0 \cos(u) - y_0 \sin(u) \\ y_0 \cos(u) + x_0 \sin(u) \end{pmatrix}$$

where

$$u = r_0(t^2 - 1) \quad r_0 = \sqrt{x_0^2 + y_0^2}$$

In this example, the average autocorrelation is 0 everywhere except at 0 as in the previous example, and there is sensitive dependence on initial conditions due to the presence of $r_0$ as a factor in the argument of the sine and cosine.

However, there is no exponential loss of information about the angular frequency $r_0$ since the argument of the sine and cosine is equivalent to

$$r_0(t^2 - 1)\text{mod}(2\pi)$$

Dividing by $2\pi$ and sampling this map at the equally spaced time intervals $t = n$ gives, for large $n$, the approximate values

$$r_0(n^2)\text{mod}(1)$$

The analysis of Ford [1986] shows that the loss of binary bits of information in this equation is of the order 2 $\ln(n)$ and hence is not comparable with the loss of information in the sequence defined by the one-sided shift,

$$(2^n)r_0 \bmod(1)$$

which is of the order of $n$.

## 4.2. *Example 17: ZA, SD and EL without a horseshoe.*

We present this example before giving the Case 5 example since that example is derived from Example 17. The system of ODEs that define this example are:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} -ryz \\ rxz \\ z \end{pmatrix}$$

where $r = \sqrt{x^2 + y^2}$.

The general solution of these equations is given by

$$\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} x_0 \cos(u) - y_0 \sin(u) \\ y_0 \cos(u) + x_0 \sin(u) \\ z_0 \exp(t) \end{pmatrix}$$

where

$$u = r_0 z_0(\exp(t) - 1) \quad r_0 = \sqrt{x_0^2 + y_0^2}$$

The time $t$ map, $\Phi_t$, is given by:

$$\Phi \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos(u) & -\sin(u) & 0 \\ \sin(u) & \cos(u) & 0 \\ 0 & 0 & \exp(t) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

where

$$u = rz(\exp(t) - 1) \qquad r = \sqrt{x^2 + y^2}$$

### 4.2.1. *Connection to the one-sided shift*

By choosing $t = \ln(2)$ in $\Phi_t$ we obtain the mapping

$$\Phi_{\ln(2)} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos(rz) & -\sin(rz) & 0 \\ \sin(rz) & \cos(rz) & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Iterating this mapping $n$ times gives the map:

$$\Phi^n_{\ln(2)} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$= \begin{pmatrix} \cos(rz(2^n - 1)) & -\sin(rz(2^n - 1)) & 0 \\ \sin(rz(2^n - 1)) & \cos(rz(2^n - 1)) & 0 \\ 0 & 0 & 2^{n+1} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Choosing the initial condition $(1, 0, 2\pi\theta)$ with $\theta \in [0, 1]$ we have,

$$\Phi^n_{\ln(2)} \begin{pmatrix} 1 \\ 0 \\ 2\pi\theta \end{pmatrix} = \begin{pmatrix} \cos((2^n - 1)2\pi\theta) \\ \sin((2^n - 1)2\pi\theta) \\ 2^{n+1}2\pi\theta \end{pmatrix}$$

The first two components of the right-hand side vector are comparable to the components, in vector form, of the $n$th iterate of the mapping of the complex circle given by

$$S(z) = z^2$$

where we use $z = a + bi$ with $\|z\| = 1$. If we choose as a starting point on the unit circle the point $\exp(2\pi\theta i)$ the $n$th iterate of $S$ is

$$\exp(2^n(2\pi\theta)i) = \cos(2^n(2\pi\theta)) + i\sin(2^n(2\pi\theta))$$

To see the connection to a one-sided shift note that the arguments in the sine and cosine can be replaced by

$$2^n(2\pi\theta)\text{mod}(2\pi)$$

which can also be expressed as

$$2^n\theta \bmod(1)$$

so that the $n$th iterate of $S$ is just the sine and cosine evaluated at the pseudo-random number given by $2^n\theta \bmod(1)$.

#### 4.2.2. *Autocorrelation in forward time*

We return to the general solution of our ODEs which is given by:

$$\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} x_0 \cos(u) - y_0 \sin(u) \\ y_0 \cos(u) + x_0 \sin(u) \\ z_0 \exp(t) \end{pmatrix}$$

where

$$u = r_0 z_0(\exp(t) - 1) \quad \text{and} \quad r_0 = \sqrt{x_0^2 + y_0^2}$$

Using the average autocorrelation for a bounded signal to compute the forward time average autocorrelation for $x(t)$ and $y(t)$ confirms that their autocorrelation functions are the same as white noise.

We conclude from the foregoing that this example has the two most important and universally recognized features of chaotic dynamical systems and so must be considered chaotic. However, the solution of the ODE can be written down in terms of elementary functions.

#### 4.2.3. *Sensitive dependence on initial conditions*

We use the time-one map for this comment. The third component of this map is exponential and hence must have sensitive dependence on initial conditions by Example 11.

#### 4.2.4. *Horseshoes*

We use the time $t$ maps, $\Phi_t$, for these remarks. By direct examination of $\Phi_t$, $t \neq 0$, we see that the only fixed point is the origin, which is not hyperbolic. There are no periodic points since $z(t) \to \infty$ as $t \to \infty$. Hence, there are no horseshoes.

#### 4.2.5. *Example 18: EL and ZA without SD (Case 5)*

If we formulate the preceding example as a nonautonomous, two-dimensional ODE we concede the time-one map, $\Phi_t$, but obtain another interesting example. As a two-dimensional system we have the ODEs:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} -ry \exp(t) \\ rx \exp(t) \end{pmatrix}$$

where $r = \sqrt{x^2 + y^2}$ as before.

The general solution of these equations is given by

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} x_0 \cos(u) - y_0 \sin(u) \\ y_0 \cos(u) + x_0 \sin(u) \end{pmatrix}$$

where

$$u = r_0(\exp(t) - 1) \qquad r_0 = \sqrt{x_0^2 + y_0^2}$$

The same autocorrelation can be obtained for $x(t)$, $y(t)$ as before. Also the link to the one-sided shift is still visible by evaluating the time series at the equally-spaced times $t_n = n \ln(2)$. In this case we have the sequences,

$$\begin{pmatrix} x(t_n) \\ y(t_n) \end{pmatrix} = \begin{pmatrix} x_0 \cos(r_0(2^n - 1)) - y_0 \sin(r_0(2^n - 1)) \\ y_0 \cos(r_0(2^n - 1)) + x_0 \sin(r_0(2^n - 1)) \end{pmatrix}$$

which have an exponential loss of information about the angular frequency, $r_0$.

Sensitive dependence on initial conditions is lost since the fixed point $(0, 0)$ is a solution to our equations so there is no constant $\tau$ such that some solution starting near $(0, 0)$ eventually diverges by the amount $\tau$ at some later time. In fact, all solutions stay a fixed distance from $(0, 0)$ for all time.

In general, given

$$\cos(f(x_0, t))$$

where $\partial f(x_0, t)/\partial t$ is eventually increasing and

$$-\infty < \int_0^\infty \cos(f(x_0, t))dt < \infty$$

then

$$\lim_{N \to \infty} \frac{1}{N} \int_0^N \cos(f(x_0, t + s)) \cos(f(x_0, t))dt = 0 \quad \text{for} \quad s > 0$$

The proof follows from the addition formula for the cosine. $\cos(t^2)$ satisfies these conditions since it gives the *Fresnel* integral

$$\int_0^\infty \cos(t^2)dt = 0.5\sqrt{0.5\pi}$$

### 4.3. *Example without autocorrelation*

To construct an example to cover Case 5 we must first present an interesting example due to Boyd & Chua on which it is based. In Boyd & Chua

[1985] an example of a dynamical system that has no autocorrelation is given. A significant fact about this system is that it has no mean value. Following their analysis we construct a new example of a system having exponential loss of information and having no autocorrelation and which does not have sensitive dependence on initial conditions, Case 5.

We first note that the basis of their example is the function

$$x(t) = R\cos(\ln(t+1)+\theta)$$

where $R$ and $\theta$ depend on the initial conditions. The mean value of this function, as they observe, does not exist since

$$\int_0^N \frac{x(t)dt}{N}$$

is periodic as a function of $N$. We note that $x(t)$ is a solution to the classical Cauchy–Legendre differential equation:

$$(at+b)^2\ddot{x} + (at+b)\dot{x} + x = 0$$

with $a = b = 1$. The general Cauchy/Legendre equation is solved by the substitution $at+b = \exp(z)$ which reduces it to a linear equation having constant coefficients.

The example of Boyd & Chua is given by the equations:

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} R\cos(\ln(1+tx_3(0))+\theta) \\ R\sin(\ln(1+tx_3(0))+\theta) \\ x_3(0)(1+tx_3(0))^{-1} \end{pmatrix}$$

where $R = \sqrt{x_1(0)^2 + x_2(0)^2}$, $x_1(0) = R\cos(\theta)$, $x_2(0) = R\sin(\theta)$ and defines a circuit having no average power.

By modifying these equations to include a twisting action we construct an example which has exponential loss of information, EL, has no autocorrelation or average value and hence does not have ZA, and does not have sensitive dependence on initial conditions (SD):

**Example 19** (Case 3).

Consider the equations:

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} R\cos(\cos(R(\exp(t)-1))\ln(1+t)+\theta) \\ R\sin(\cos(R(\exp(t)-1))\ln(1+t)+\theta) \end{pmatrix}$$

It is routine to derive an autonomous three-dimensional system from these two functions and so we omit the derivation. The autonomous system is given by:

$$\begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{pmatrix} = \begin{pmatrix} -x_2\dot{u}(x_3) \\ x_1\dot{u}(x_3) \\ 1 \end{pmatrix}$$

where $u = \cos(R(\exp(x_3)-1))\log(x_3+1)$. The factor $v = R\exp(x_3)$ in $u$ assures exponential loss of information in the initial condition $R$. By taking the cosine of $v$ we get a bounded factor that is losing information at an exponential rate. We then multiply this by $\log(t+1)$ to get the effect of the dilating timescale in the argument of the sine and cosine. The dilating timescale is the key feature in the Boyd & Chua example that assures that it cannot have an average value, and hence cannot have ZA.

This example illustrates another possibility for chaos, a dynamical system that wanders randomly in time having no average value, and hence cannot be analyzed in a statistical sense.

## 5. An Algebraic Method of Recognizing Chaos: Functions of Shifts

In this section we provide three examples that may suggest an algebraic means of identifying chaos. To construct these examples we derive systems which are functions of the one and two-sided shifts. The rationale for doing this is to develop an intuitive ability to recognize when maps are related to shift maps.

Two examples have already been encountered. They are Examples 9 and 10 given in Sec. 2.2. The third example is Example 20:

**Example 20.** Transformations of the cat map.

Consider

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \dfrac{2x - y(1-x^2)}{1-x^2-2xy} \\ \dfrac{x+y}{1-xy} \end{pmatrix}$$

$T$ is not invertible but is locally invertible and has a hyperbolic fixed point at the origin with intersecting stable and unstable manifolds. These features

follow directly from the map

$$M\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

used to construct $T$. The map $M$ can be restricted to the torus by considering its components with addition mod 1 and it is an Anosov diffeomorphism there. An advantage that $T$ has over $M$ restricted to the torus is that many scientists working with nonlinear phenomena cannot follow the abstract arguments needed to see the relationship of $M$ to the two-sided shift. A disadvantage of $T$ is that it is not globally invertible, whereas $M$ is invertible.

## 6. Representations of the Bilateral and Unilateral Shifts as Twist-and-Flip Maps

In the previous section we gave three examples of sequences and maps that were formed from the one-sided shift. In this section we will show how to represent both the one and two-sided shifts using a composition of maps that arise naturally from forced oscillators and autonomous three-dimensional dynamical systems.

In Brown & Chua [1991a, 1991b] we showed that Poincaré maps in square-wave forced oscillators such as that of Duffing could be factored into two components one of which was a twist map. In Brown & Chua [1993] we showed that the second component in this factorization could vary from something as simple as a 180-degree rotation to very complex maps. In particular, the simple twist-and-flip map was found to be a closed form Poincaré map and it was shown that it was representative of Poincaré maps for a broad class of systems. Generalizations of this paradigm led to other interesting closed form Poincaré maps such as the twist-and-translate map, referred to as the twist-and-shift in Brown & Chua [1991b]. In Brown & Chua [1993] we showed that Poincaré maps for a broad class of driven oscillators could be factored into combinations of twists, translates, dilations, and contractions to name a few, and that each of these driven oscillators could be realized as an electronic circuit. We showed that the Hénon map was an example of this factorization. Earlier Brown [1992] showed that this paradigm could be extended to include Poincaré maps for the generalized Chua circuit. The twist map also occurs as the fundamental object of analysis in KAM theory. Thus we see that maps such as the twist-and-flip map are ubiquitous in dynamics, especially Hamiltonian dynamics.

Definitions of chaos which involve the shift map, such as is the case when we use the Smale horseshoe in the definition of chaos, pose a theoretical and practical obstacle in their use since it is not obvious how these maps relate to general Poincaré maps. (For example, prior to Brown and Chua [1993], the Hénon map was an "heuristic" example of what a Poincaré map might look like. In our 1993 article, we show how to obtain the Hénon map as an actual Poincaré map of an electronic circuit). Consequently, in many cases difficult theorems, such as the Smale–Birkhoff theorem, which are inaccessible to the average working nonlinear scientist, must be invoked to prove the presence of chaos in a dynamical system. An alternative we seek is a direct *algebraic* relationship between shift maps and twist maps which will provide a link by which direct comparisons of Poincaré maps with shift maps can be made.

In this section we will present some representations of the shift as maps that frequently occur in the Poincaré maps of commonly occurring dynamical systems. In particular we will show that the shift map occurs as a twist-and-dilation map and that the shift can have a component that is a sigmoid map.

### 6.1. *Representation of the bilateral shift as a twist-and-dilation map in the plane*

In Sec. 3.1 of Brown & Chua [1993] we suggested the following construction but did not present it explicitly.

**Example 21.** Bilateral shift as a twist-and-dilate map.

We define two maps $T_1$, $T_2$ as follows:

$$T_1\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} r \\ \theta \end{pmatrix}$$

and

$$T_2\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} r \\ \theta \end{pmatrix}$$

We require $-\pi \leq \theta < \pi$, in place of the usual $0 \leq \theta < 2\pi$ in these equations. The composition $T_2 \circ T_1$

is

$$(T_2 \circ T_1)\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} r \\ \theta \end{pmatrix}$$

The composition considered on the torus is the familiar cat map and is known to be a bilateral shift.

In rectangular coordinates the map

$$T_1\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1 + \arctan(y/x)}{r}\begin{pmatrix} x \\ y \end{pmatrix}$$

is a radial dilation or contraction of the vector $(x, y)$. We take the arctangent to have values in the interval $[-\pi/2, \pi/2]$.

The map $T_2$ is the simple twist:

$$T_2\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(r) & -\sin(r) \\ \sin(r) & \cos(r) \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

where $r = \sqrt{x^2 + y^2}$. The composition $T_2 \circ T_1$ is not invertible in the plane but is nearly so in a practical sense. In particular, the points on the curve $\arctan(y/x) = -r$ are all mapped to the origin. But this is a set of measure zero and is unlikely to be encountered in practice. The non-invertibility occurs because we have permitted the arctangent to have negative values. The orbits of this map resemble the samples of a vector random variable with a Gaussian distribution. This map shows how a shift could occur as a Poincaré map in a driven oscillator.

This example also shows that if we remove a set of measure 0, that a forced oscillator does not have to have *any* elliptic island chains such as is seen in the familiar KAM theory. We conjecture that by perturbing $T_1$ so that it is invertible we will force the appearance of elliptic regions on sets of positive measure that do not exist for $T_2 \circ T_1$.

In general, KAM island chains can arise from maps of the form $T_2 \circ T_1$ where $T_2$ is a simple twist and $T_1$ is any number of other possible maps, [Brown & Chua, 1993]. Example 21 shows that the structure and appearance of the island chains, when they occur, can be attributed solely to the nature of the map $T_1$. If $T_1$ is a rotation that preserves the integral curves of $T_2$ then the entire plane is elliptic. In this case, all solutions are almost periodic, and there are no island chains or chaos. If $T_1$ is as in Example 21, then there are no islands chains either. There are some almost periodic orbits as there are in the shift, but in general, chaos prevails exactly as it does in the bilateral shift. In between these two extremes arise the island chains in which the map $T_2 \circ T_1$ is sometimes elliptic, some times hyperbolic, and sometimes parabolic, [Arnold & Avez, 1989].

**Example 22.** Unilateral shift as a twist-and-dilate map.

We define two maps $T_1$, $T_2$ as follows:

$$T_1\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} \theta \\ \theta \end{pmatrix}$$

and

$$T_2\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} r \\ \theta \end{pmatrix}$$

$T_2$ is invertible but $T_1$ is not. In rectangular coordinates we have

$$T_1\begin{pmatrix} x \\ y \end{pmatrix} = \frac{\arctan(y/x)}{r}\begin{pmatrix} x \\ y \end{pmatrix}$$

and $T_2$ is, as in the previous example, the simple twist:

$$T_2\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(r) & -\sin(r) \\ \sin(r) & \cos(r) \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

The composition is a noninvertible map on the plane which is a one-sided shift. Using the constructions in Brown & Chua [1993], Secs. 2.0 and 3.2, we can make this map the Poincaré map of an electronic circuit.

**Example 23.** Unilateral shift as a sigmoid map.

This example is found in Brown & Chua [1993] but we repeat it here:

Let

$$f(x) = x - 0.5(1 - \tanh(0.5\beta(1 - x))$$

and define on the unit interval the map

$$T(x) = f(2x)$$

As $\beta \to \infty$ this map converges pointwise to the unilateral shift, except at a finite number of points. This example shows that the occurrence of a term of the form $f(2x)$ in a forcing term of a differential equation can have the effect of a shift, and thus create chaos. An extreme example of this is the modified Chua equation [Brown, 1992].

## 7. Bifurcations that Form Elliptic Regions

In this section we present an interesting example that is an aside from the main line of exposition. It is a bifurcation process in which we bifurcate from chaos without any elliptic regions as in the bilateral shift example above, to the formation of elliptic regions. This is done by changing a parameter. This example is motivated by the measure-preserving and orientation-preserving Hénon map.

**Example 24.** Elliptic bifurcations.

The map

$$T_1\begin{pmatrix}x\\y\end{pmatrix} = \left(b + \arctan\left(\frac{y}{x}\right)\right)\begin{pmatrix}\cos(r)\\ \sin(r)\end{pmatrix}$$

is a radial dilation or contraction of the vector $(x, y)$ similar to the unilateral shift example. We take the values of the arctangent in the interval $[-\pi/2, \pi/2]$ and take $b > \pi/2$ so that the map is locally invertible since the determinant $(b+\arctan(y/x))$ is never zero for values of $b > \pi/2$. For $T_2$ we take the twist:

$$T_2\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}\cos(f(r)) & -\sin(f(r))\\ \sin(f(r)) & \cos(f(r))\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$$

where $r = \sqrt{x^2+y^2}$ and $f(u) = 1 - au^2$. The map $T_2$ is a twist modeled after the nonlinear fundamental map in the Hénon map [Brown & Chua, 1993]. The composition $T_2 \circ T_1$ depends on the parameter $a$. For $a \approx 0.5$ the orbits resemble the bilateral shift map. As $a$ is decreased to 0, an elliptic region forms and then for $a = 0.0$ all orbits are elliptic.

## 8. Chaos, Randomness, and Pseudo-Random Number Generators

The subjective aspects of our thinking about chaos today are not so different from the thinking about the practical generation of random numbers that occurred beginning in the 1940s with the advent of electronic computers. A review of that dialogue reveals that the search for a practical random number generator was also a search for explicit algorithms that would generate what might be called *uniform* chaos. Because of this suggestive historical connection, we briefly review some comments from that era.

### 8.1. *Some history about pseudo-random processes*

In the late 1930s and on to about 1964, the *pseudo-random number generator problem* was to find a simple, practical, and reproducible source of uniformly distributed numbers on the interval $[0, 1]$ for use in computer simulations of complex problems involving random processes. This problem (in today's terminology) was to determine a satisfactory chaotic dynamical system which mapped $[0, 1]$ into itself and which passed all reasonable tests of *uniform* randomness. As we know, variations on the one-sided shift were settled on as the best candidate at that time and the basic concept of the shift is still the basis of many pseudo-random number generator algorithms today. For example, see Lidl & Niederreiter [1986] and Berlekamp [1984], under cyclic codes. The intense focus on the determination of the best generator of *uniform* deviates was due to the fact that once this mapping was constructed it could be used to obtain all other probability distributions. In a parallel sense, the one- and two-sided shifts are the generators of *uniform* chaos, and many other forms of chaos may be generated by considering functions of the shifts. To follow this analogy further, if $f$ is any invertible mapping of the unit interval, and $S$ is the one-sided shift on the unit interval, then the mapping $f \circ S$ is a chaotic mapping of the interval as well. In the event $f$ is the arctan, then the composition with the shift produces, approximately, a Gaussianly-distributed pseudo-random variable. This analogy goes further. A major cornerstone of the argument about the uniformity of pseudo-random numbers was the statistical tests of randomness. On this point Kendall & Babington-Smith [1938] set forth four tests of randomness [Kendall & Babington-Smith, 1938], p. 154. These tests amounted to the first attempts to define a practical test of when a *deterministic* dynamical system was uniformly chaotic, at least from the statistician's point of view. The use of statistical tests were instrumental in bringing discipline to the notion of pseudo-random numbers and on the subjectivity of this topic Hull & Dobell [1962], p. 240, write:

> "*In this way one may avoid becoming involved in any philosophical arguments about the meaning of randomness, arguments which,*

*according to Kendall and Babington-Smith 'are of an abstract and metaphysical character bordering at times on the theological'."*

The debate over the definition of chaos over the past 15 years could be similarly described in some quarters. In any case, the notion that a system is chaotic when it can be shown to pass some test that the one-sided shift passes is the basis of most of the definitions of chaos today.

By 1964 the debate had settled down and in that year the National Bureau of Standards published the monumental work *Handbook of Mathematical Functions*, edited by Milton Abramowitz and Irene Stegun [Abramowitz & Stegun, 1964]. In the section on probability functions (Chap. 26) they state, with regard to the generation of random numbers:

"*... the use of random numbers in electronic computers has resulted in a need for random numbers to be generated in a completely deterministic way.*"

p. 949.

This remarkable comment alludes to the need to have a deterministic dynamical system (to use today's terminology) which can generate what appears to be uniform random deviates. The algorithm (dynamical system) selected, after much searching and analysis, to provide uniform pseudo-random numbers was iteration of the map $X_{n+1} = (aX_n + b)\mathrm{mod}(T)$, a one-sided shift.[3] The numbers $a$ and $b$ are to be chosen so that

"*(1) the resulting sequence possess the desired statistical properties of random numbers, (2) the period of the sequence is as long as possible, (3) the speed of generation is fast.*"

ibid, p. 949

The specific statistical test recommended as a guide for choosing the numbers $a$ and $b$ is to make the average autocorrelation of the sequence *near zero*, ibid, p. 949.

Of particular practical importance in choosing this dynamical system as a random number generator is that the operations needed for generation of each deviate *only require a shift operation plus two additions*, ibid, p. 950. The authors note (as is well known) that from this dynamical system (our terminology), all other *random variables* can be generated.

In R. W. Hamming's classical text *Numerical Methods for Scientists and Engineers* [Hamming, 1962], we find an interesting passage describing why a sequence of integers generated by the one-sided shift is considered random:

"*... the numbers seem to be in a chaotic order.*"

p. 138

The algorithm he is referring to is $x_n = \rho x_{n-1}\mathrm{mod}(2^k)$, a one-sided shift for integers between $[0, 2^k]$, the same as Abramowitz and Stegun with $b = 0$. Dividing this algorithm by $2^k$ we get the standard one-sided shift $\rho x \,\mathrm{mod}(1)$. The number $2^k$ is chosen for computers having a word length of $k$. The exact value of the number $\rho$ is a practical constraint needed to assure the longest possible sequence is generated on a binary computer of finite word length. For an imaginary computer of infinite word length which can be thought of as the limit of computers of finite word length $k$, we can take $\rho = 2$ and divide all numbers by $2^k$ and let $k \to \infty$ to get the dynamical system $x_{n+1} = 2x_n \,\mathrm{mod}(1)$, the standard one-sided shift on two symbols so often encountered in nonlinear dynamics. More recent discussions of pseudo-random numbers can be found in Knuth [1981]. In general, the simple shift $2x \,\mathrm{mod}(1)$ is not "random" enough for the generation of pseudo-random numbers. In contrast, a suitable pseudo-random number algorithm is given by

$$x_{n+1} = 16807x_n \,\mathrm{mod}(1)$$

when restricted to the subset

$$\frac{1}{m}, \frac{2}{m}, \frac{3}{m}, \ldots, \frac{m-1}{m}$$

where $m = 2147483647 = 2^{31} - 1$, which is a large prime number, see Park & Miller [1988]. If we make this map invertible by embedding it into a

[3] An example of this map characterizes the global dynamics of an idealized multivibrator electronic circuit during the chaotic regime when synchronization is lost [Tang *et al.*, 1983], [Small *et al.*, 1995], as observed in the rolling of a television picture when operating "out of sync".

two-dimensional map as shown in Brown & Chua [1993], pp. 1246–1249, then this orbit lies very near the unstable manifold. The Lyapunov exponent in this case is approximately log(16807). By choice of $m$ the orbit defines a very long pseudo-random sequence. In this case we have made chaos and pseudo-randomness coincide exactly and through this example we see that pseudo-randomness is chaos with a very Lyapunov exponent for which a periodic orbit can be found that is the word length of a given computer and on which the mapping satisfies randomness criteria.

Regardless of the course of the debate about what is pseudo-random, the essential practical need for generating deterministic random deviates on a computer remains intact and any conclusion from this dialogue will have some implications for chaos theory. Perhaps the conclusion about chaos and pseudo-random number generators will be: Given any definition of chaos or pseudo-randomness, there will always be some "clearly chaotic/ pseudo-random" dynamical system whose chaos/ pseudo-randomness cannot be established by that definition.

## 8.2. *Algorithm versus initial conditions*

In this section we ask "Does chaos refer to the complexity of the time series determined by the initial conditions of a dynamical system or does it refer to the algorithm or operations that define the dynamical system?" Clearly, a given dynamical system is no more than its set of orbits (time series) determined by initial conditions. However, it is equally plausible that simple initial conditions give rise to complex orbits by the action of complex dynamical systems. A badly needed example of this, which however does not yet exist, would be to show that there is a point with rational coordinates on the unstable manifold of the Hénon map having rational parameters.

The initial condition view is supported by the error analysis of the one-sided shift, see Ford [1986], pp. 20–21, which reveals that to obtain $k$ bits of accuracy after $n$ iterations of an initial condition requires that $n + k$ bits of information be present in the initial condition at the start. This analysis leads Ford to the view that chaos is algorithmic complexity, i.e., it is the incompressibility of information provided by a dynamical system that defines chaos. This view is consistent with the view that it is the rate at which information is lost in the initial conditions (as in exponential divergence of solutions of an ODE) that determines chaos and it includes the notion of the unpredictability of the future from the past. In algorithmic complexity the emphasis is apparently on the algorithm. But the uncertainty in the analysis of Ford comes from the initial condition on which the algorithm operates. Clearly, the very simple shift algorithm (dynamical system) is not the source of complexity in shift dynamical system, but rather it is the initial conditions. With periodic initial conditions we get periodic orbits, and with almost periodic initial conditions we get almost periodic orbits. In every case the orbit is equivalent to the initial condition. This seems to place the entire source of chaos on the initial conditions. But how do we define complex initial conditions? One way is to use the coin toss experiment, 1 for heads 0 for tails. But this is an algorithm! In fact any attempt to define a complex initial condition leads to an algorithm. Thus we are brought back to the algorithmic complexity view of chaos. Once here, we begin our analysis by reducing the algorithm (at least in thought) to binary bits, which can be considered as an initial condition for a machine to operate on. But what machine? A standardized, very simple machine having simple operations such as a Turing machine. Given the standardized machine, we then examine its operation on the initial conditions which were obtained from the algorithm and *its* initial conditions. Now it seems that the Turing machine, or dynamical system, is primary. But where do we get the interesting initial conditions on which the machine will operate? Our machine, much like the shift operator, is so simple that it cannot be the source of any complexity, so the complexity must come from the initial conditions. Thus the initial conditions are the primary source of complexity. But to get interesting initial conditions requires action on our part, thus some form of dynamics are needed, more complex than our simple machine. It appears that we are stuck in a unfriendly circle when trying to find the primary object which we can call chaotic.

It appears from the example of the Turing machine and the shift that we tend toward making the initial condition primary. This means binary bit sequences are primary. Binary sequences have

analogies with real-valued functions. There are periodic sequences corresponding to periodic functions and there are almost periodic sequences corresponding to almost periodic functions. There are random sequences corresponding to a coin toss, and there are sequences in between random and almost periodic, such as the binary representation of the decimal part of $\pi$ and e, which, according to algorithmic complexity theory [Ford, 1986] are not random. (An implication of this view is, for example, that the number $\pi$ should not be used as a seed for generating pseudo-random numbers on a computer of infinite word length.)

To obscure further which is the primary concept, we offer the following example: Consider the binary number, defined by a function $f : \mathbf{N} \to \{0, 1\}$ where,

$$f(n) = 1 \quad \text{if } n \text{ is prime and 0 otherwise.}$$

Is this a chaotic binary number, or is it the orbit of a dynamical system? Surely it must be both but it is not conceptually generated by an operation or dynamical system in the way that the coin-toss sequence is generated. However, it can be mathematically represented as a dynamical system since the prime numbers form an increasing sequence. If we define $F(n) = n$th prime, then we can define the dynamical system on the range of $F$ by $\Phi(F(n)) = F(n+1)$, which is the shift.

We conclude that the primary source of chaos may be philosophical, not mathematical.

## 8.3. *Rates of information loss*

Since two key features of chaos, exponential degradation of information or separation of orbits, and 0 correlation between the past and the future, have motivated most dynamical systems thinking about what is chaos, it is significant to ask if these two properties of a system are separate, or does one always imply the other? As we saw, the answer is that they are separate. There are dynamical systems in which information is lost at an exponential rate, and which have no autocorrelation and there are non-trivial functions with zero autocorrelations, which arise from dynamical systems which do not have an exponential degradation of information, or even sensitive dependence on initial conditions. In this section we take a closer look at the role of exponential degradation of information in chaos.

We return to the one-sided shift $x_{n+1} = 2x_n \bmod(1)$, which has the general solution $x_n = 2^n x_0 \bmod(1)$. The difference between this sequence and the sequence $x_n = nx_0 \bmod(1)$ is that information in the latter sequence degrades at the rate $\ln(n)$. But it *does* degrade. Further, the sequence generated by $x_n = 2^n x_0 \bmod(1)$ is a subsequence of $x_n = nx_0 \bmod(1)$. But the first equation agrees with our intuitive notion of chaos and the second does not, even though, for each $x_0$, the second sequence contains the first sequence as a subsequence. This fact demonstrates that our intuition about chaos is connected to the *rate* at which information is degraded and not the fact that it is degraded. But the rate at which information is degraded is also a function of the sampling rate. If we sample the function $y = \sin(t)$ at equally spaced time intervals $1, 2, 3, \ldots$ we get the sequence $a_n = \sin(n)$, which is not chaotic. But if we sample the same function at times $t_n = 2^n$ we get the sequence $a_n = \sin(2^n)$. This is chaotic as seen in Example 1 since the effect of the argument $2^n$ in a periodic function of period $2\pi$ is the same as using $2^n \bmod(2\pi)$ which is the same as using the function $2^n(1/2\pi) \bmod(1)$, and which is the sequence of iterates of the dynamical system $x_{n+1} = 2x_n \bmod(1)$ with $x_0 = 1/(2\pi)$.

Any function on a bounded interval can be sampled by use of a random number generator and we can obtain, thereby, a random sample of the function. This is how we generate normally distributed random variables, for example. Thus, given any non-constant function, $f$, on $[0, 1]$, the function $g(n) = f(2^n x_0 \bmod(1))$ is a pseudo-random or chaotic, sequence of the values of $f$, depending on the value of $x_0$.

If sampling at a rate of $n$ is not chaotic but sampling at a rate of $2^n$ is chaotic, then what about the sequence $g(n) = f(n^2 x_0 \bmod(1))$? A result of H. Weyl provided the basis for an analysis of the sequence $y_n = n^2\pi \bmod(1)$ which was seriously considered as a source of pseudo-random numbers for computer applications in the 1950s [Golenko, 1959]. The key theorem of Weyl [1916] (under broad constraints) can be used to show that the sequence $(n^2/\pi) \bmod(1)$ is uniformly distributed in $[0, 1]$. Combined with the fact that this sequence is also independent, $(n^2/\pi)\bmod(1)$ is very suggestive of chaos. Although this sequence may not be entirely acceptable as a generator of *uniformly distributed* pseudo-random numbers, it does require serious

consideration as a chaotic sequence. For a map to be chaotic it should not be necessary that it pass all tests of randomness since this is impractical under the best of circumstances, see Rand [1955], Wold [1948].

This leads to the question "what is the rate of separation of orbits or degradation of information at which we believe that chaos begins?" Since it is customary to express information in terms of $\log_2$, a rate of degradation of $n$ corresponds to the exponential function $\exp(t)$, and a rate $\log_2(n)$ corresponds to $t$. This suggests that we use the jump between polynomial rates and exponential rates as our philosophical criteria. Or when expressed as logarithms, the jump between the log function and the first linear polynomial. This gap leaves a lot of room for shades of complexity to be defined. For example, the digits of $\pi$ appear to be at least as random as those of a pseudo-random number generator used by computers[4] but are not considered chaotic from the algorithmic complexity view [Ford, 1986]. The same situation occurs for any constant that is computed by using a finite algorithm, even though the decimal expansion of many significant constants are too chaotic to predict.

Further, there is a lot of room above exponential rates of divergence for "darker" shades of complexity to be defined. In other words, we can question the sacredness of exponential separation of orbits as the starting point of chaos. For example, what about separation at the rate of $\Gamma(n)$, which is factorial separation. For example, consider the map $T(z) = a^z$, $a > 1.0$. Is this a higher order of chaos, than exponential? So where does chaos begin? At present, there is no precise answer to this question.

Our skeptical approach can even be extended to question whether the Smale horseshoe is sacred. After all, the Smale–Birkhoff theorem says that a diffeomorphism $\Phi$, has a horseshoe when some iterate of $\Phi$, say $\Phi^{100000000}$ is topologically conjugate to a shift. There are realistic examples where this is the case. In fact, given any integer N, there is a chaotic system for which the Smale–Birkhoff iterate is greater than N. (Simple twist-and-flip systems in which the amplitude of the square-wave voltage is small provides such a set of examples.) This is analogous to asking if the one-sided shift

$$T(x) = (1.00000000001)x \bmod(1)$$

is chaos, since it is a root of a shift, i.e. $T^{[1/\log_2(1.00000000001]+1]}$, where $[x]$ denotes the integer part of $x$, is slightly more random than the shift $2x \bmod(1)$. But $T$, as defined by the above formula, is nearly as predictable as a linear function, since the powers of 1.00000000001 increase very slowly.

So what is Chaos? Is it but a subjective notion illustrating the limitations of the human intellect, or is it an intrinsic property of nature such as the sequence of prime numbers? We cannot answer this question.

## 8.4. *Linear versus nonlinear chaos*

In this section we question whether linear systems is a realistic line of demarcation for chaotic processes. We have noted in Sec. 3 that there are unbounded linear systems which have EL and SD. We could just take these examples as an exception and keep linear systems out of the picture. However, there are other examples of interest.

Consider the sequence $1, 4, 9, 16, 25 \ldots$. We can guess that the $n$th term by inspection is $n^2$. We can do this for more complex sequences such as $1, -1/4, 1/64, -1/2304, \ldots$, whose $n$th term is given by the formula:

$$\frac{(-1)^n}{(n!)^2(2^{2n})}$$

and is the $n$th term of the power series expansion of the *Bessel Function the First Kind of Order Zero.* A sequence for which we can write down an equation for the $n$th term is somewhat simpler than a sequence for which we cannot. If we cannot, then is not this system suggestive of chaos? The algorithmic complexity view would argue for this situation. For example, we can write down an infinite series that generates the most important physical constants such as $\pi$ and e. An interesting question whose answer could shed further light on what is chaos is this: Is it always possible to write down a general formula for the $n$th term of the power series expansion for a solution of a *linear* ODE? The answer is not known, and, at present, Bessel's equation has solutions for which a general formula for

[4]The work of G. Chudnovsky and P. Chudnovsky confirms that the first billion digits of $\pi$ are as random as a pseudo-random number generator used on a computer [Schroeder, 1991].

the $n$th term of the power series expansion is not known.

As a side issue, we can even ask if forced nonlinear oscillators are a starting point for chaos in the following sense. The $n$th term of the power series solution for the non-chaotic equation

$$\ddot{y} + y^3 = 0$$

cannot be expressed in closed form. In fact, it involves the initial condition. Thus, does an initial condition of positive algorithmic complexity result in a power series of positive algorithmic complexity?

Another complication is that it can be proven that every dynamical system can be expressed as a *linear* operator on a locally compact topological vector space. The essential construction is found in Doob [1953], Chap. 10, Sec. 1. This construction must be combined with the metric defined in Robinson [1995], p. 37, and some additional observations to get this result. This construction cannot be carried out so as to assure the space has a norm, and so there is no representation in terms of a normed linear space. However, the term *linear* does not *formally* exclude chaos.

## 8.5. *Chaos and cryptography*

It could be said that the study of "chaos" in discrete dynamical systems such as finite algebras is the subject of cryptography. This analogy is persuasive in many ways. Crypto systems are invertible dynamical systems whose orbits are used to generate a sequence of numbers to be used as codes for letters. The seeds are the initial conditions and parameters. The sensitivity to initial conditions is a key specification of a good crypto system. While not stated in those exact terms, crypto theorists do describe the requirements of a good cipher as implying sensitive dependence on initial conditions. For example, W. E. Madryga's crypto algorithm design objectives included the following requirement:

> "*A 1-bit change of the key should produce a radical change in the ciphertext using the same plaintext, and a 1-bit change of the plaintext should produce a radical change in the ciphertext using the same key. This is called the avalanche effect.*
>
> [Schneier, 1994], p. 245

This design objective can clearly be met by a wide range of chaotic dynamical systems. As a further connection between chaos and cryptography we note that the operations used in many important crypto schemes such as DES are a composition of involutions. In Brown & Chua [1993], we show that many measure preserving chaotic maps are compositions of involutions, and Devaney [1976] and Devogelaere [1958] show that such systems, called reversible, are characteristic of Hamiltonian systems. Specifically, the twist-and-flip dynamical systems bear a strong resemblance to rotor machines [Denning, 1982].

Hyperbolic maps on discrete spaces all have positive Lyapunov exponents. In one dimension such maps may be called discrete exponential maps. They make good ciphers because it is hard to compute discrete logarithms. Approximations are of little value precisely because hyperbolic maps on discrete spaces have sensitive dependence on initial conditions and low autocorrelations. Although all discrete space dynamical systems are periodic, they can be related to continuous space dynamical systems which are not, but which have subsets on which they are periodic. If these subsets come close to hyperbolic regions of the continuous systems, then they benefit from chaos. It is possible to see encryption schemes as mappings of continuous space dynamical systems with the discrete domain of the encryption scheme being a subset on which the continuous dynamical system is periodic. For example, the pseudo-random number generator defined by

$$x_{n+1} = 2x_n \bmod(p)$$

where $p$ is a large prime number can be realized as the mapping

$$x_{n+1} = 2x_n \bmod(1)$$

restricted to the periodic subset of numbers $1/p \ldots (p-1)/p$ which is a subset of the interval $[0, 1]$. This embedding of encryption schemes in continuous dynamical systems and vice versa suggest a good source of encryption techniques and interesting dynamical systems. In this regard, the encryption scheme of Blum–Micali

$$x_{n+1} = a^{x_n} \bmod(p)$$

defines a chaotic dynamical system with "gamma-chaos", i.e. chaos that is beyond exponential. The

encryption scheme of Blum–Blum–Schub [Schneier, 1994] defines a dynamical system that makes an interesting connection between number theory and chaos.

In general, the relations that may be drawn between cryptography and chaos seem beneficial to both disciplines.

## 8.6. *Random versus chaotic*

Numerous researchers have asked the question:

**How is a random process different from a chaotic process?**

To fix the discussion we provide some review of random processes and the concept of randomness.

### 8.6.1. *Random variables in mathematics*

In probability theory we define a random variable as a measurable function. This is misleading. For one thing, the order in which measurable functions take on their values is significant. For example, if $f(x)$ is a measurable function on the interval $[0, 1]$, then $f(1-x)$ is a different measurable function. However, within the context of probability theory they are the same random variable. This is because they have the same distribution function. In general, if $\mathbf{T}(x)$ is any measure preserving mapping of $[0, 1]$, then $f(x)$ and $f(\mathbf{T}(x))$ are the same random variable, but they are not the same measurable function.

The first thing that is important about a function in probability theory is its range, not its domain. Second is the frequency with which a function takes on the values in its range. The two functions $f(x)$ and $f(1-x)$ have the same range, and they take on the values in their range with the same frequency. Thus we do not distinguish them in the subject of probability theory.

We may make all of this rigorous from a mathematical point of view if we consider a random variable to be the set of all functions having the same distribution function. One way to do this is to define a random variable, $f$, as the set

$$\{g | \exists \mathbf{T} \ni g(x) = f(\mathbf{T}(x))\}$$

where $\mathbf{T}$ is a measure preserving map of the unit interval.

Now that we have formally defined the concept of random variable, we ask "Where does the randomness come in?" For example, the function $f(x) = x$ is a random variable by definition but this function is not the first thing that comes to mind when thinking about random variables. The problem is that the formal definition of random variable has side stepped the real question of when does a function qualify as random?

### 8.6.2. *Random functions and sequences*

G. Chaitin and I. Kolmogorov independently addressed the question of what is randomness. Their approach is roughly as follows: First we must select a mathematical object for which a concept of randomness could conceivably apply. The most reasonable choice is a function, and the simplest functions to work with are mappings from the integers into some other set such as the real numbers. Every such map defines a sequence so we ask what does it mean for a sequence to be random. Appealing to the history of probability theory to guide us we look for examples. What comes to mind is the *gedanken experiment* which defines a sequence by a coin toss, 1 for heads 0 for tails. This is a mapping from the integers onto the set $\{0, 1\}$. This example will satisfy everyone's idea of randomness and thus if we can *formally* characterize what makes this sequence random we will be on the way to a general formal definition.

The source of randomness in this example is the *ideal* coin toss. What makes this *process* random is that we *imagine* that no matter how many 1's or 0's we have tossed, the value of the next toss is still in doubt. A mathematical way of saying this is that there is no way to write down a formula that will give the value of the next toss in terms of the previous tosses, no matter how many tosses we have made. Since any sequence of 0's and 1's represents a binary number, a coin toss number is a number for which no algorithm can be written down. In particular, no recursive algorithm exists that will generate the number. It can be proven that most numbers on the interval $[0, 1]$ are of this form [P. Martin-Löf, 1966]. That is, if we think of selecting a number "at random" between $[0, 1]$, it will, with probability one, be a number that cannot be written down using an algorithm.

Chaitin and Kolmogrov analyzed finite sequences and derived a definition of what it means for such a sequence to be random. The coin toss

example is an infinite sequence. So we back up for a moment to look at the finite case, the basic idea is the same as the infinite case. We now ask: Given a finite sequence of 1's and 0's when is it random? Clearly, if we can write a finite formula for the $n$th term of the sequence it seems that it would not qualify as random since subsequent values of the sequence are determined from past values. However, if the effort to do this is considerable, then maybe it is random after all. Chaitin and Kolmogrov quantified this idea using the concept of a computer program. If the shortest computer program needed to generate the sequence consumes far fewer bits than those needed to simply print the number, then the sequence must not be random. However, if these two values are close, then the sequence is very random; if they are equal, it is completely random.

Kolmogrov, Alekseev, Solomonoff, and Chaiten have all settled on the name algorithmic complexity theory to describe this line of thinking, the term "random" being relegated to a informal role. P. Martin-Löf [1966] established that the concept of complexity as set forth by Kolmogorov does pass all conceivable tests for randomness. Kolmogrov defined complexity for finite sequences. Chaiten and Alekseev independently extended this notion to infinite sequences (as a historical note we add that Chaiten's work on all aspects of complexity proceeded independently from the Russian school.) We now say that a number for which no finite algorithm can be found which will generate it has positive algorithmic complexity. Whenever we talk about an algorithm it only makes sense to discuss a *finite* algorithm and this is the term used in algorithmic complexity theory.

Through this line of thought we see that the notion of randomness can be made mathematically rigorous. Specifically, using this framework we can say that the numbers which are "random" are those whose binary representation cannot be described in anyone's lifetime or the lifetime of the universe. This is completely consistent with the notion of randomness found in the coin toss gedanken experiment. Now we must link this to dynamical systems in order to draw out the relationship to chaos.

In order to place this entire discussion in the framework of dynamics, one needs a space such as $[0,1]$, and a mapping on this space with which we can describe our situation in terms of the orbit of a map. The way that it has been chosen to do this so that the map itself does not get into the way of our thought process is to use the shift map on the set of binary sequences. This amounts to placing our entire attention on the initial conditions of the system and ignoring the dynamics altogether since the shift map is a very trivial device that makes the framework of the discussion a dynamical system. Thus, in reality, the binary representation of numbers is where all of the complexity lies, at least as it has been formulated thus far.

How do we relate this to chaos? As a first point we note that the shift operator on $[0,1]$ is a chaotic map. If we start with an initial condition of positive algorithmic complexity, the orbit is unpredictable, i.e., there is no recursive algorithm that can describe the orbit since there is none to describe the initial condition. Thus a chaotic dynamical system can generate an orbit that is just like one generated by a coin toss. But, given the same initial condition, we get the same orbit, unlike a coin toss. But there is a muddle in the midst of all of this. The initial condition of the shift corresponds to an infinite coin toss that has already been made and in our analogy we are comparing it to a coin toss we are about to make. The initial condition of the coin toss is determined by whether the coin starts as a head or a tail. This is not comparable to the shift initial condition which is the completed infinite coin toss sequence. The problem is we cannot compare a dynamical system defined by an algorithm with a system (perhaps purely philosophical) for which there is no algorithmic definition. The tenuous connection between these two "systems" is the unspecifiable "chaotic" initial condition of the shift which can nevertheless be proven mathematically to exist. Therefore, given an initial condition for the shift, we can theoretically generate a specific *realization* of a coin toss, but not the coin toss system itself. Or so it seems. We recall that an initial condition of positive algorithmic complexity cannot itself be generated by an algorithm and is thus not realizable. So all we can do is to imagine that all of this can be done.

At this point we are forced into a digression about non-chaotic systems. Every dynamical system, linear or nonlinear, can start with an initial condition of positive algorithmic complexity. Regardless, we would still consider the linear system to be predictable even when starting from a mathematically random initial condition. This is because

a bounded linear map does not separate two orbits that start close together and thus the algorithmic complexity of the initial conditions never gets into the act. This is true even in non-chaotic nonlinear systems such as the twist map in the plane for which all solutions are almost periodic. By ordinary computations we can show that for bounded non-chaotic systems having positive algorithmic complexity in the initial conditions, that the complexity is not relevant because two orbits starting close together stay close together. In linear unbounded systems the algorithmic complexity does eventually enter in, but only because the error enters into the orbit not because the initial conditions were very complex, just that they were unknown. Thus, for some systems, if we measure the initial conditions well enough (even though they may have positive algorithmic complexity) we can predict the future forever. The better we measure, the better we can predict. For other systems such as the shift, no matter how well we measure the initial conditions if the initial conditions have positive algorithmic complexity we cannot predict the future any better than we can for a coin toss. Thus a chaotic dynamical system perpetuates into the future the difficulties posed by initial conditions of positive algorithmic complexity. But for initial conditions that can be described by an algorithm, even the shift is predictable, virtually by definition. The extent to which chaotic processes are "random" is in the initial conditions and how the process advances those initial conditions. If we could avoid initial conditions with positive algorithmic complexity, we would be able to avoid chaos. But mathematical theory tells us that such initial conditions are the rule, not the exception. We pose an intriguing general problem in this regard: Given the Hénon map having rational values for the parameters $a$, $b$, is there a point on the unstable manifold having rational coordinates which is not a hetero- or homoclinic point?

So randomness is a property of a system that cannot be described by an algorithm. As to whether such systems exist is a open question of philosophy, not mathematics or science. The closest we can come to randomness in dynamics is the dynamical system determined by the shift on the interval $[0, 1]$ when the initial condition has positive algorithmic complexity. This is an example of chaos, but not necessarily the universal example. However, no such initial condition is describable in mathematical terms. And if it were, we would still obtain only a single realization of a coin toss and to obtain another realization we must get another initial condition. If we could get to this point we could finally get some help from mathematical theory. Because, if we have one initial condition of positive algorithmic complexity, we have a countable set which is the orbit of the shift starting with the given initial condition. This orbit is dense in the interval $[0, 1]$ and so from a practical point of view is as good as having as many initial conditions as we can use.

But the problem of how to get at least one initial condition remains and is, in essence, the problem of making a random selection. There is no mathematical formulation of "making a random selection", there is only a formulation of how to determine, once a selection of some sequence is made, if the selected sequence is random.

It is interesting that Kolmogorov only coined the term complexity to apply to finite sequences even though the infinite notion was an obvious extension of the finite concept. Alekseev and Chaiten independently took the concept a step further to include infinite sequences. In effect, they took the step to talk about a class of numbers that can never be written down in any period of time and that the only way to convey information about such sequences is to present the sequence itself, since there exists no shorter or more abbreviated method of talking about it by definition, but of course, this is impossible.

So if we are to be mathematically rigorous about randomness we are forced into the realm of algorithmic complexity. To do this within the framework of dynamics leads us to the shift on the set of bi-infinite sequences of symbols where the number of symbols is infinite. To imagine randomness in this setting we simply imagine that the shift acts on a symbol sequence of positive algorithmic complexity that we, in fact, cannot write down in any finite interval of time. If we could, the sequence would have complexity 0 and would not be random.

To summarize in probabilistic terms: The shift on bi-infinite sequences with an initial condition having positive algorithmic complexity (which in fact can never be specified) is the mathematical idealization of a discrete stationary stochastic process of mutually independent random variables. But how do we obtain a realization of this process. We must still make a selection of an initial condition.

### 8.6.3. *Random selections*

Traditionally in probability theory when we talk about a random variable, $f$, we think of making a random selection of a value of a specific realization of the random variable $f$. However, as soon as we entertain this thought we have departed from the realm of mathematics and this reveals the fundamental problem: How do we make rigorous the thought or act of making a random selection?

To make a random selection requires some procedure. But of course procedures are not random. They involve a specific set of orderly reproducible steps, or an *algorithm.* For example the procedure to select lottery numbers using ping pong balls put into motion by air currents is not random. But it is sufficiently complex that no one has yet discovered a way to systematically guess the lotto numbers even on a slightly regular basis.

Another approach would be to use some process as a standard random process. Philosophically one could argue that no process is random, by definition of process. All processes must obey the laws of physics, known or not yet known. In this regard, some have argued that the emissions of photons by excited hydrogen atoms may qualify as random events, but even this can be disputed philosophically since there is no reproducible experiment that can *prove* that this process is random.

Even if there exists "physical" random events whose randomness cannot be proven, we may still ask if it is possible to provide a rigorous mathematical formulation of a random selection?

Given the preceding discussion, we are tempted to conclude that making a random selection has always been an intuitive notion that must defy a precise description since any formal description would imply some form of order that would contradict our notion of random selection. In particular, any algorithm for making a random selection must contradict the non-computability of random sequences. There is no perfect way out of this problem. But there is a practical way out. To make a random selection we use chaotic orbits of chaotic dynamical systems. This is not new, it is just stated in a new way. Whenever we use a random number generator this is what we are doing. In fact, we may reformulate the pseudo-random number generation problem in terms of chaos as follows:

**How do we find large order periodic orbits of dynamical systems which pass a practical set of tests for randomness?**

As noted in Sec. 8.5, one approach is to start with chaotic dynamical systems having large Lyapunov exponents such as the modular mapping $x \to 7^5 x \bmod(1)$ restricted to the periodic orbit of elements of the form $k/(2^{31}-1)$. By recasting this as a problem of finite fields, we can conclude that this orbit is a practical random number source. However, the problem of finding large periodic orbits of chaotic maps does admit a more general approach which has not yet been explored.

So the theoretical problem of making a random selection is exchanged for the theoretical problem of determining the "pseudo-random" periodic orbits of dynamical systems, most likely chaotic dynamical systems. In this regard we have finite field theory, algebraic coding theory, encryption theory, and algorithmic complexity theory to aid in this analysis.

For example, using the notion of complexity of Chaitin and Kolmogrov we can decide how compressible a given orbit is and settle on a threshold of compressibility to be used to select periodic orbits for random number generators. The ideas for this approach are implicit in Chaitin [1975].

### 8.6.4. *Some examples*

In previous sections we have mentioned several examples of chaotic maps that could be considered for the generation of pseudo-random samples of distributions of random variables for use in computer simulations. In this section we have two simple examples of chaotic maps which may be used to write quick computer routines to produce two familiar pseudo-random processes.

**Example 25.** Two-dimensional Gaussian distribution.

$$T_2\begin{pmatrix} x \\ y \end{pmatrix} = \frac{r+2xy}{r^2}\begin{pmatrix} \dfrac{x^2-y^2}{r^2} \\ \dfrac{2xy}{r^2} \end{pmatrix}$$

$T_2$ is made by the use of invertible fundamental maps [Brown & Chua, 1993] in polar coordinates and followed by a change to rectangular coordinates. $T_2$ is a perturbation of $T_1$ which is

unbounded. The iterates of $T_2$ resemble the samples of a two-dimensional pseudo-random variable having a Gaussian distribution.

Using the examples of Sec. 2 we can construct a pseudo-random walk which must be considered as chaos:

**Example 26.** Random walks.

Consider the finite difference equation:

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} 2x_n^2 - 1 \\ y_n + x_n \end{pmatrix}$$

We know that $x_n = \cos(2^n C)$ and so $y_n$ is the sum of $n$ terms of $x_n$. Other "random walks" can be similarly obtained.

Anyone implementing these maps will note certain differences in the orbits from standard stochastic methods that use pseudo-random number generators. The differences arise from the size of the Lyapunov exponent of the underlying map. The level of pseudo-randomness can be increased by using a higher iterate of these maps, thus increasing their Lyapunov exponent. In particular, in place of the sequence

$$y_n = \cos(2^n \pi C)$$

that gives rise to the first component of Example 26, we could use five iterates of the map determined by the sequence:

$$y_n = \cos(7^n \pi C)$$

to get a single sample. The map determined by the preceding sequence is $y_{n+1} = 64y_n^7 - 112y_n^5 + 56y_n^3 - 7y_n$. Taking five iterates of this map as the first component in Example 26 will be the same as using the pseudo-random map $x \to 16807x \bmod(1)$. By starting at a point of the form $1/2147483647$ we obtain a series of random numbers generated by the map

$$x \to 16807 \bmod(2147483647),$$

the map discussed by Park & Miller [1988].

## 8.7. *Chaos and stochastic processes*

In this final section we show how to link stochastic processes and chaotic processes. This idea is not new as seen by the paper of Sakai & Tokumaru [1988]. In that work it was shown that a simple chaotic map could be considered as a first-order auto-regressive process. The work of Chua, Yao & Yang [1990] takes this idea a step further and constructs a switched capacitor circuit for a map that includes $2x \bmod(1)$. This circuit can be used as a practical noise generator. Both works can be advanced by using $16807x \bmod(1)$ in place of $2x \bmod(1)$. The higher Lyapunov exponent provides greater "randomness". (In general, it is well known that every stationary stochastic process can be viewed as a measure-theoretic dynamical system [Doob, 1953].) In this section we carry this idea forward to general moving average processes by addressing the following problem posed by Dr. Douglas Lake, Office of Naval Research:

**Is it possible to use a chaotic dynamical system to model a given stochastic process on a computer?**

We first note that if we model a stochastic process on a computer, the entire model is built up from a random number generator, which is deterministic. Typical choices are modulo arithmetic schemes. These are deterministic and discrete dynamical systems. The simple generator $x_{n+1} = 16807\, x_n \bmod(2147483647)$ is a good example. As noted in Sec. 8.1 the orbits of this map can be considered as periodic orbits of the closely related map $x_{n+1} = 16807\, x_n \bmod(1)$ which has a periodic orbit of numbers of the form $k/2147483647$ and gives the same pseudo-random sequences as the modulo arithmetic scheme. The mod(1) map is a chaotic dynamical system on $[0, 1]$ with Lyapunov exponent 16807, and is a shift with the entropy $\log(16807)$. Hence any formula or algorithm built up from this map is chaotic as well. Hence, the answer is: Every stochastic process modelled on a computer is a chaotic dynamical system. But this is not an answer to the spirit of Dr. Lake's question. The question could also be posed, "*Is there a simple chaotic dynamical system that does not use transcendental functions or the standard random number generators that can be used in a practical sense to construct a model of any stochastic process on a computer?*" (For example, are any of the familiar chaotic maps useful in this regard?) This is the question we address. Of particular importance in addressing this question is that the selected

dynamical system model should be much faster than the conventional approach.

As a first step in answering this question we pose a simpler version of the problem:

**Given a stationary (wide sense) stochastic process, $\psi$, having a autocorrelation function $R(n)$, is there a chaotic dynamical system that can be used to closely approximate the realizations of $\psi$ on a computer?**

The answer to the simpler question is yes. The answer to the first question is not yet known. We give a construction which answers the simpler question.

## 8.8. *Review of moving average processes*

From Doob [1953], pp. 498–499 we recall that a moving average is a process of the form

$$x_n = \sum_{k=-\infty}^{\infty} c_k \xi_{k+n}$$

where $\xi_k$ are orthogonal random variables, and $c_k$ are constants such that

$$\sum_{k=-\infty}^{\infty} |c_k| < \infty$$

The constants $c_k$ can be represented as the Fourier coefficients of a periodic function, $c(\lambda)$:

$$c(\lambda) = \sum_{k=-\infty}^{\infty} c_k \exp(2\pi i k \lambda)$$

The correlation, $R$, of the process $x_n$ is given by

$$R(n) = \int_{-1/2}^{1/2} |c(\lambda)|^2 \exp(2\pi i n \lambda) d\lambda,$$

the Fourier coefficients of $|c(\lambda)|^2$. Given $R$ in advance we can always construct $c(\lambda)$ and thus $x_n$.

In this discussion we will always assume that the processes we are working with are such that $c_k = 0$ for $k < 0$, and so our process may be written as

$$x_n = \sum_{k=0}^{\infty} c_k \xi_k$$

## 8.9. *Construction of mappings for $\xi_k$, $c_k$, and $x_n$*

One way to construct a dynamical system to generate realizations of $x_n$ is to generate dynamical systems for each component needed for $x_n$. In general, Fourier coefficients are not obtained by iteration. However, in special cases they can be obtained by the iteration of a dynamical system. We will show how to do this for a simple case, and then show how this simple case can be extended to obtain a general result.

### 8.9.1. *Simple dynamical system for $c_k$*

The simplest dynamical system that can be used to get Fourier coefficients is:

$$c_{k+1} = r c_k$$

where $0 < |r| < 1$. This finite difference equation corresponds to the map

$$f(x) = rx$$

The general solution is $c_k = r^k c_0$. Since $|r| < 1$, the associated Fourier series converges uniformly. By a direct computation, the correlation coefficients of the associated process are given by

$$R(n) = \frac{r^n c_0}{1 - r^2} = \frac{c_n}{1 - r^2}$$

### 8.9.2. *A dynamical system for $\xi_k$*

Referring to Example 3 of Sec. 2 we are able to construct an orthogonal set of random variables. Let $\eta$ be a uniform random variable on $[0, 1]$. Then,

$$\xi_k = \cos(2^k \pi \eta)$$

form an orthogonal set of random variables. Given a sample of $\xi_0$ we can generate samples of the entire sequence from the relation

$$\xi_{k+1} = 2\xi_k^2 - 1$$

This amounts to using $2x \bmod(1)$, the shift, as a pseudo-random number generator. While this has some limitations, it still has considerable practical value. In particular, when we compose the shift with cosine, the recurrence relation for samples of

$\xi_k$ is a chaotic dynamical system and can be relied upon to give uncorrelated samples of the entire sequence.

### 8.9.3. *A dynamical system for $x_n$*

We obtain a dynamical system for $x_n$ by obtaining a dynamical system for the sequence of partial sums that define $x_n$. In particular, let

$$S_N = \sum_{k=0}^{N} c_k \xi_k$$

then

$$S_{N+1} = S_N + c_{N+1}\xi_{N+1}$$

As $N \to \infty$ $S_N \to x_n$.

## 8.10. *Putting it all together*

To obtain the final dynamical system we will use to generate $x_n$ we need three coordinates for the three parts of the system. The first coordinate generates $c_k$, hence we use the map $x \to rx$. The second coordinate generates samples of $\xi_k$ and is given by $y \to 2y^2 - 1$, the third coordinate generates the sequence of partial sums $S_N$ and is given by $z \to z + xy$. The final assembled map is

$$\mathbf{T}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} rx \\ 2y^2 - 1 \\ z + xy \end{pmatrix}$$

When $\mathbf{T}$ is iterated, the $z$-component converges to a sample of $x_n$.

## 8.11. *The dynamics of* $\mathbf{T}$

By a direct computation the eigenvalues of $D\mathbf{T}$ are, $r$, $-2$, 1, for the stable, unstable, and center manifolds at a fixed point. A fixed point is given by $(0.0, 0.5, z)$. $\mathbf{T}$ is locally invertible but not invertible globally. Hence $\mathbf{T}$ is not a diffeomorphism and thus does not properly have the said manifolds. But it does have analogies of these manifolds that behave the same as if it were a diffeomorphism. A test for chaos is possible but the most direct proof is that the second coordinate is chaotic, thus making the entire map chaotic. There is another problem with $\mathbf{T}$ in this regard: The fixed points are not isolated. All of this limits the application of the theories of differential geometry (stable manifold theorem, Smale–Birkhoff theorem, etc), but not the fact that it is chaotic. Here are the data:

$$D\mathbf{T} = \begin{pmatrix} r & 0 & 0 \\ 0 & 4y & 0 \\ y & x & 1 \end{pmatrix}$$

The eigenvectors are the coordinate vectors when $|r| < 1$. There is a hyperbolic fixed point for every $z$ at $(0, -0.5, z)$ with stable and unstable directions as the $x$-axis and $y$-axis respectively. The unstable direction does not have a "manifold" (no proof) but does define a possibly non-smooth set. The dynamics of $\mathbf{T}$ at any fixed point are the same as if the unstable set were a manifold with "virtual" homoclinic points. This can be seen by noting that as the first coordinate converges to 0, the second coordinate oscillates in a manner similar to an unstable manifold having transverse homoclinic points. Thus, the first two coordinates of this map define a two-dimensional, differentiable, noninvertible map that is quite an interesting example in itself.

For given $r$, $|r| < 1$, the first coordinate always converges to 0 while the last converges to the sample of the stochastic process. Hence the first and last coordinates have attractor characteristics. The second coordinate remains chaotic.

## 8.12. *Generalizations for arbitrary correlations*

In this section we address the general problem of constructing a dynamical system that can be used to model a moving average process having correlation $R(n)$. We divide the problem into two parts, one theoretical, the other practical. First the theoretical problem.

### 8.12.1. *Theoretical dynamical systems*

Since $R(n)$ determines $c(\lambda)$, we need only show how to approximate arbitrary $c_k$ for a finite set of $k$'s with a dynamical system. The theoretical technique to do this follows directly from the theory of finite difference equations. In particular, we may ask "given a finite sequence of $k$ numbers, is there a finite difference equation whose solution has these numbers as the first $k$ values in its solution?" The answer is yes in general. However, the finite difference equation may be of a very high order. This in turn makes the dynamical system of high

dimension. To obtain a low-dimensional mapping we simply express the coefficients $c_k$ in a formula for the $n$th term and take the first difference. Using this along with a coordinate to express the time dependence that may exist results in a four-dimensional system. This would, presumably, be the lowest dimensional general system that could be found. But this is all theory, implementation is harder.

### 8.12.2. *Practical dynamical systems*

A more practical approach is to see how far we can go by generalizing the simple example we have already developed. This is done by assuming the $i$th term of the sequence of coefficients in the moving average can be expressed as:

$$c_i = \sum_0^j \alpha_j r_j^i$$

where $0 < |r_i| < 1$. The $\alpha$s are determined by the first $k$ $c_i$s so that if there are $k$ terms given, we can obtain them exactly by solving $k$ equations in $k$ unknowns. The remaining terms are obtained by iteration. Of course this is a form of extrapolation using a dynamical system and hence is an approximation. However, it can be very useful in the case where the sequence $c_k$s are monotonically decreasing. Still, we do not have a low order system unless we are satisfied with the number $k$ being 1, 2 or 3 for example. We have an example of this approach for $k = 4$:

**Example.**

Let $r_1$, $r_2$ be in $[-1, 1]$ and assume that

$$\begin{aligned} c_0 &= \alpha_1 + \alpha_2 \\ c_1 &= \alpha_1 r_1 + \alpha_2 r_2 \\ c_2 &= \alpha_1 r_1^2 + \alpha_2 r_2^2 \\ c_3 &= \alpha_1 r_1^3 + \alpha_2 r_2^3 \end{aligned}$$

which are 4 equations in 4 unknowns which can be solved for $\alpha_i$, $r_i$ in terms of $c_i$. The general term is

$$c_k = \alpha_1 r_1^k + \alpha_2 r_2^k$$

and represents an extrapolation if the first 4 $c_i$s are known. Given that the first few terms are the most important in a correlation function which must go to 0 and $n \to \infty$, this extrapolation could be harmless in many cases. We caution that this method does require that the $c_k$s are solutions to a *linear* finite difference equation which is not generally true of Fourier coefficients. The convenience of this approach is that $R(n)$ can be directly computed and must solve the same linear finite difference equation as the $c_i$'s.

In the case where there exist a general expression for the $n$th Fourier coefficient of the function $c(\lambda)$ a third order equation can be obtained, but $R(n)$ may not be easy to compute. The approach is illustrated in the following example:

**Example.**

Let $c_k = 1/k$, then $k = 1/c_k$ and a finite difference equation for the reciprocal of the $c_k$ is possible. Specifically,

$$(x_k + 1)x_{k+1} - x_k = 0$$

or

$$x_{k+1} = \frac{x_k}{x_k + 1}$$

The $\mathbf{T}$ corresponding to this set of $c_k$ is given by

$$\mathbf{T}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x/(1+x) \\ 2y^2 - 1 \\ z + y/x \end{pmatrix}$$

To obtain the sequence $c_k = 1/k^2$ we choose

$$\mathbf{T}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x/(1+x) \\ 2y^2 - 1 \\ z + y/x^2 \end{pmatrix}$$

and for an arbitrary function of $k$, say $f(k)$ we use

$$\mathbf{T}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x/(1+x) \\ 2y^2 - 1 \\ z + yf(1/x) \end{pmatrix}$$

where the initial condition must have $x_0 = 1$. As noted at the end of the last section we may improve the randomness of the second coordinate by using five iterates of the map

$$y \to 64y^7 - 112y^5 + 56y^3 - 7y$$

in place of $y \to 2y^2 - 1$.

The analysis of $\mathbf{T}$ is the same as before. One last example that gives a means of obtaining

$R(n)$ in a compact form for a special case of $c(\lambda)$ is given by Bessel's function:

**Example.**

Let $c_k = (0.5)^k/k!$ then

$$R(n) = \sum_0^\infty \frac{(0.5)^k(0.5)^{k+n}}{k!(n+k)!} = \frac{J_n(i)}{i^n}$$

where $J_n(i)$ is Bessel's function of order $n$ evaluated at the complex argument $i = \sqrt{-1}$.

### 8.13. *Defining chaos revisited*

Given that we require that a chaotic dynamical system be one that is definable by a finite algorithm we pose the following question:

Let $\Phi$ be a finite algorithm (dynamical system) and let $z_0$ be an initial condition of positive algorithmic complexity. Then $\Phi^n(z_0)$ is a sequence of numbers of positive algorithmic complexity. Likewise, for any other initial condition $z$

$$f(n) = \frac{|\Phi^n(z_0) - \Phi^n(z)|}{|\Phi^n(z_0)| + |\Phi^n(z)|}$$

is a sequence of numbers of positive algorithmic complexity. For every positive real number $a$ we define

$$g_a(n) = 1 \text{ if } f(n) > a \quad \text{and} \quad 0 \text{ if } f(n) \leq a$$

For what algorithms does $g_a$ converge for every $a < |z - z_0|$?

## 9. Summary and Conclusion

Our conclusion based on these examples seems simple: At present *chaos* is a philosophical term, not a rigorous mathematical term. It may refer to the present day limitations of human thought or it may describe the "randomness" of the sequence of prime numbers. Moreover, chaos may be undecidable in the sense of Gödel in that no matter what definition is given for chaos, there is some example of chaos which cannot be proven to be chaotic from the definition.

What can be said is that there are simple chaotic systems whose time series can be written down in closed form; also, we can conclude that sensitive dependence on initial conditions needs something more in the definition if it is to be a synonym for chaos. Various features of chaos, **ZA, SD, EL** are independent features of complex systems and hopefully can be derivable from a good definition of chaos.

Lastly, we believe that the development of an *algebraic* theory of chaos is possible and that by understanding how shifts algebraically occur in oscillators may lead us to a new, more easily applied, proof of the existence of chaos.

## Acknowledgments

We would like to thank Ms. Becky Popp of the Naval Research Laboratory for extensive conversations about the definition of chaos and for her many valuable suggestions that led to a complete reorganization of this manuscript and the rewriting of many sections. We would also like to thank Jim Heagy, Lou Pecora, and Tom Carroll of the Naval Research Laboratory for lively dialogues concerning chaos and random number generators. We also thank Mingzhou Ding of Florida Atlantic University for numerous lively conversations about the construction of our examples. This work was supported in part by ONR contract N00014-95-C-0153.

## References

Abramowitz, M. & Stegun, I., eds. [1964] *Handbook of Mathematical Functions* (Dover, New York).

Alekseev, V. & Yakobson, M. [1981] "Symbolic dynamics and hyperbolic dynamical systems," *Physics Rep.* **75**, 287.

de Almeida, A. [1988] *Hamiltonian Systems: Chaos and Quantization* (Cambridge University Press, Cambridge).

Arnold, V. & Avez, A. [1989] *Ergodic Problems of Classical Mechanics* (Addison-Wesley, New York).

Beardon, A. [1991] *Iteration of Rational Functions* (Springer-Verlag, New York).

Bergé, P., Pomeau, Y. & Vidal, C. [1984] *Order within Chaos* (John-Wiley & Sons, New York).

Boyd, S. & Chua, L. [1985] "Dynamical system state need not have spectrum," *IEEE Trans. Circuits and Syst.* **CAS-32**(9), 968–969.

Brown, R. & Chua, L. [1991a] "Horseshoes in the twist-and-flip map," *Int. J. of Bifurcation and Chaos* **1**(1), 235–252.

Brown, R. & Chua, L. [1991b] "Generalizing the twist-and-flip paradigm," *Int. J. of Bifurcation and Chaos* **1**(2), 385–416.

Brown, R. [1992] "Generalizations of the Chua equations," *Int. J. of Bifurcation and Chaos* **2**(4), 889–909.

Buck, R. [1956] *Advanced Calculus* (McGraw-Hill, New York).

Chaitin, G. [1975] "Randomness and mathematical proof," *Scientific American* **232**(5).

Chaitin, G. [1994] "Randomness and complexity in pure mathematics," *Int. J. Bifurcation and Chaos* **4**(1).

Cornfeld, I., Fomin, S. & Sinai, Y. [1982] *Ergodic Theory* (Springer-Verlag, Berlin).

Chua, L. O., Yao, Y. & Yang, Q. [1990] "Generating randomness from chaos and constructing chaos with desired randomness," *Int. J. of Circuit Theory and Applications* **18**(3), 215–240.

Chua, L. O. [1980] "Dynamic nonlinear networks: State of the art," *IEEE Trans. Circuits and Systems* **27**(11), 1059–1087.

Denning, D. [1982] *Cryptography and Data Security* (Addison-Wesley, Reading, Massachusetts).

Devaney, R. [1976] "Reversible diffeomorphisms and flows," *Trans. Amer. Math. Soc.*, **218**, 89–113.

Devaney, R. [1989] *An Introduction to Chaotic Dynamical Systems* (Addison-Wesley, New York).

Devogelaere, R. [1958] "On the structure of symmetric periodic solutions of conservative systems with applications," *Contributions to the Theory of Nonlinear Oscillations* **IV**, 53–54.

Doob, J. L. [1953] *Stochastic Processes* (John Wiley & Sons, New York).

Epstein, R. & Carnielli, W. [1989] *Computability, Computable Functions, Logic and the Foundations of Mathematics* (Wadsworth & Cole, Pacific Grove, California).

Ford, J. [1986] "Chaos: Solving the unsolvable, predicting the unpredictable," from *Chaotic Dynamics and Fractals* (Academic Press, New York).

Grebogi, C., Ott, E., Pelikam, S. & Yorke, J. [1984] "Strange attractors that are not chaotic," *Physica* **D13**, 261.

Guckenheimer, J. & Holmes, P. [1983] *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Springer-Verlag, New York).

Gulick, D. [1992] *Encounters with Chaos* (McGraw-Hill, New York).

Hamming, R. [1962] *Numerical Methods for Scientist and Engineers* (McGraw-Hill, New York).

Hirsch, M. [1985] "The chaos of dynamical systems" from *Chaos, Fractals, and Dynamics* (Marcel Dekker, New York).

Hsu, H. P. [1984] *Applied Fourier Analysis* (Harcourt Brace, New York).

Hull, T. & Dobell, A. [1962] "Random number generators," *SIAM* **4**, 230–254.

Katok, A. [1980] "Lyapunov exponents, entropy and periodic points for diffeomorphisms," *Publ. Math. IHES* **51**, 137–174.

Katznelson, Y. [1976] *An Introduction to Harmonic Analysis* (Dover Publications, New York).

Knuth, D. [1981] *The Art of Computer Programming* (Addison-Wesley).

Lewis, R. & Papadimitriou, C. [1981] *Elements of the Theory of Computation* (Prentice-Hall, Englewood Cliffs, N.J.).

Lorenz, E. [1993] *The Essence of Chaos* (University of Washington Press, Seattle).

Martin-Löf, P. [1966] "The definition of random sequences," *Information and Control* **9**, 602–619.

Moon, F. C. [1987] *Chaotic Vibrations* (John Wiley & Sons, New York).

Mullin, T. [1994] *The Nature of Chaos* (Oxford Univ. Press).

Nemytskii, V. & Stepanov, V. [1960] *Qualitative Theory of Differential Equations* (Dover, New York).

Orenstein, D. [1989] "Ergodic theory, randomness, and chaos," *Science* **243**, 182–186.

Papoulis, A. [1984] *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York).

Park, S. & Miller, K. [1988] "Random number generators: Good ones are hard to find," *Commun. ACM* **31**(10), 1192–1201.

Rand Corporation [1955] *A Million Random Digits with 100,000 Normal Deviates* (The Free Press, Glencoe, Illinois).

Robinson, C. [1995] *Dynamical Systems* (CRC Press, Boca Raton).

Sakai, H. & Tokumara, H. [1980] "Autocorrelations of certain chaos," *IEEE Trans. Acoustics, Speech, and Signal Processing* **ASSP-28**(5), 588–590.

Schneier, B. [1994] *Applied Cryptography* (John Wiley, New York).

Schroeder, M. [1991] *Fractals, Chaos, and Power Laws* (Freeman, New York).

Schuster, H. [1988] *Deterministic Chaos* (VCH, Weinheim, Germany).

Shil'nikov, L. [1994] "Chua's circuit: Rigorous results and future problems," *Int. J. of Bifurcation and Chaos* **4**(3), 489–519.

Silverman, J. [1986] *The Arithmetic of Elliptic Curves* (Springer-Verlag, Berlin).

Smell, A. J., Dumbell, K. D. & Smith, P. D. [1995] "Chaos in the Tang–Mees–Chua model of threshold synchronization," *Int. J. Bifurcation and Chaos* **5**(1), 175–187.

Strogatz, S. [1994] *Nonlinear Dynamics and Chaos* (Addison-Wesley, New York).

Tang, Y. S., Mees, A. I. & Chua, L. O. [1983], "Synchronization and chaos," *IEEE Trans. on Circuits and Systems* **30**, 620–626.

Thompson, J. & Stewart, H. [1986] *Nonlinear Dynamics and Chaos* (John Wiley & Sons, New York).

Ulam, S. & Von Neumann, J. [1947] "On combination of stochastic and deterministic processes; Preliminary report," *Bulletin of the AMS*, p. 1120.

Walters, P. [1982] *Introduction to Ergodic Theory* (Springer-Verlag, New York).

Wiggins, S. [1990] *Introduction to Applied Nonlinear Dynamical Systems and Chaos* (Springer-Verlag, New York).

Wiggins, S. [1992] *Chaotic Transport in Dynamical Systems* (Springer-Verlag, New York).

Wold, H. [1948] "Random normal deviates," *Tracts for Computers* (Cambridge University Press, Cambridge).