NONLINEAR AND COMPLEX DYNAMICS IN ECONOMICS

by

WILLIAM A. BARNETT
Department of Economics, Campus Box 1208
Washington University
One Brookings Drive
St. Louis, Missouri 63130-4810
U.S.A.

ALFREDO MEDIO
Department of Economics
The University of Venice
Ca’ Foscari, 30123 Venice
ITALY

and

APOSTOLOS SERLETIS*
Department of Economics
The University of Calgary
Calgary, Alberta T2N 1N4
CANADA

*We thank Demitre Serletis for research assistance.

1 Introduction

According to an unsophisticated but perhaps still prevailing view, the output of deterministic dynamical systems can in principle be predicted exactly and - assuming that the model representing the real system is correct - errors in prediction will be of the same order of errors in observation and measurement of the variables. On the contrary - so the story runs - random processes describe systems of irreducible complexity owing to the presence of an indefinitely large number of degrees of freedom, whose behavior can only be predicted in probabilistic terms.

This simplifying view was completely upset by the discovery of chaos, i.e., deterministic systems with stochastic behavior. It is now well known that perfectly deterministic systems (i.e., systems with no stochastic components) of low dimensions (i.e., with a small number of state variables) and with simple nonlinearities (i.e., a single quadratic function) can have stochastic behavior. This means that for chaotic systems, if the measurements that determine their states are only finitely precise - and this must be the case for any concrete, physically meaningful system - the observed outcome may be as random as that of the spinning wheel of a roulette and essentially unpredictable. The discovery that such systems exist and are indeed ubiquitous has brought about a profound re-consideration of the issue of randomness.

It is not difficult to understand why these theoretical findings have captured the imagination of many economists. Since many important topics in economics are typically formalized by means of systems of ordinary differential or difference equations, these findings alone should be sufficient to motivate economists' broad interests in chaos theory. But there exists a question, or rather a group of questions, in economics, usually labelled 'business cycles', for which the field of mathematical research under discussion is eminently important[1].

A scanty observation of the time series of most variables of economic interest, such as the price of an individual commodity or the exchange rate between two currencies, shows the presence of bounded and more or less regular fluctuations, with or without an underlying trend. Even more interestingly, this oscillating behavior seems also to characterize the aggregate activity of industrialized economies, as represented by their main economic indicators. Economists have long been concerned with the explanation of this phenomenon. The literature on

[1]There exist, of course, other areas of research in economics for which chaos theory is, or could be shown to be, very important, i.e., technical progress. We believe, however, that the case of business cycles can best illustrate the role of nonlinear dynamical analysis in general and of chaos theory in particular, especially when we look at it in a historical perspective.

the subject is enormous and the number of different theories equally vast. However, if we restrict ourselves to the 'mathematical' investigation of economic fluctuations, we observe that two basic, mutually competing approaches have dominated this area of research in modern times.

The origin of the first approach - which we shall label 'econometric approach to business cycles' - may be traced back to the seminal works of Eugene Slutsky (1927) and Ragnar Frisch (1933) and was later developed, and given the status of orthodoxy, by the works of the Cowles Commission in the 1940s and 1950s. The fundamental idea of the econometric approach is the distinction between impulse and propagation mechanisms. In the typical version of this approach, serially uncorrelated shocks affect the relevant variables through distributed lags (the propagation mechanism), leading to serially correlated fluctuations in the variables themselves[2]. As Slutsky showed, even simple linear non-oscillatory propagation mechanisms, when excited by random, structureless shocks, can produce output sequences which are qualitatively similar to certain observed macroeconomic cycles.

The ability of the econometric approach to provide an explanation of business cycles was called in question largely on the ground that explaining fluctuations by means of random shocks amounts to a confession of ignorance. An alternative approach - which we shall label 'nonlinear disequilibrium' - was then developed by a school of economists who, somewhat misleadingly, was associated to the name of Keynes. The basic idea of these authors was that instability and fluctuations are essentially due to market failures and consequently they must be primarily explained by deterministic models, i.e., by models where random variables play no essential role. Classical examples of such models can be found in the works of Nicholas Kaldor (1940), John Hicks (1950), and Richard Goodwin (1951).

Mathematically, these models were characterized by the presence of nonlinearity of certain basic functional relationships of the system and lags in its reaction-mechanisms. The typical result was that, under certain configurations of the parameters, the equilibrium of the system can loose its stability, giving rise to a stable periodic solution (a 'limit cycle'), which was taken as an idealized description of self-sustained real fluctuations, with each boom containing the seeds of the following slump and vice versa. The nonlinear disequilibrium approach

[2] For completeness's sake, among the impulse-propagation models of the cycle, one should distinguish between those in which random external events affect economic 'fundamentals' (essentially, tastes and technology), and those in which those events directly change only agents' expectations. The latter case has been extensively studied in recent years in the economic literature under the label 'sunspots'.

to the analysis of business cycles was very popular in the forties and fifties, but its appeal to economists seems to have declined rapidly thereafter and a recent, not hostile textbook of macroeconomics [Olivier Blanchard and Stanley Fischer (1987, pp. 277)] declares it "largely disappeared".

The reasons for the crisis of the Keynesian style of theorizing and the associated nonlinear disequilibrium theories of the cycle are manifold, not all of them perhaps pertaining to scientific reasoning, and a full investigation of this interesting issue is out of the question here. However, there exist two fundamental criticisms, raised against the nonlinear disequilibrium approach mainly by supporters of the rational expectations hypothesis, which are relevant to our discussion and can be briefly summarized. First, in the nonlinear disequilibrium approach, agents' expectations, either explicitly modelled or implicitly derived from the overall structure of the model, are, under most circumstances, incompatible with agents' 'rational' behavior. Second, the nonlinear disequilibrium approach has been 'refuted' by empirical observation, as time series generated by the relevant models do not agree with available data[3].

The first of these criticisms can be best appreciated by making reference to the original formulation of the rational expectations hypothesis. In John Muth's (1961, pp. 316) own words, "the expectations of firms (or, more generally, the subjective probability distribution of outcomes) tend to be distributed, for the same information set, about the prediction of the theory (or the 'objective' probability distributions of outcomes)". If expectations were not rational in the sense defined above – so Muth's argument continues – "there would be opportunities for economists in commodity speculation, running a firm, or selling the information to the present owners"[4]. This argument does have some validity if the outcomes of the dynamical system under consideration can be accurately predicted once the 'true' model is known, i.e., if the outcome is periodic. In this case, if agents are 'rational', fluctuations can be explained only by exogenous random shocks[5].

[3]For example, so the argument runs, a time series generated by a model characterized by a stable limit cycle will have a power spectrum exhibiting a sharp peak corresponding to the dominant frequency, plus perhaps a few minor peaks corresponding to subharmonics. On the other hand, aggregate time series of actual variables would typically have a broad band power spectrum, often with a predominance of low frequencies - see, for example, Clive Granger and Paul Newbold (1977).

[4]*Ibid.*, pp. 330.

[5]Even when the outcome is periodic, however, if the periodicity is long and the time path very complicated, one may question the idea that well-informed real economic agents would actually forecast it correctly: all the more so if the outcome is quasi-periodic, possibly with a large number of incommensurable frequencies.

However, if the theory implies chaotic, unpredictable dynamics of the system, the rational expectations argument loses much of its strength and non-optimizing rules of behavior – such as adaptive reaction mechanisms of the kind assumed by the nonlinear disequilibrium models, or 'bounded rationality' *à la* Simon – might not be as irrational as they may seem at first sight. At any rate, the presence of chaos makes the hypothesis of costless information and the infinitely powerful learning (and calculating) ability of economic agents, implicit in the perfect foresight–rational expectations hypothesis, much harder to accept.

Equally strong reservations can be raised in relation to the second criticism mentioned above. It is well known, for example, that deterministic chaotic systems can generate output qualitatively similar to the actual economic time series. However, none of these broad considerations can be used as a conclusive argument that business fluctuations are actually the output of chaotic deterministic systems. They do, however, strongly suggest that, in order to describe complex dynamics mathematically, one does not necessarily have to make recourse to exogenous, unexplained shocks. The alternative option - the deterministic description of irregular fluctuations - provides economists with new research opportunities undoubtedly worth exploiting.

In the following sections, we shall try to make the basic notions of complexity, chaos, and the other related concepts more precise, having in mind their (actual or potential) applications to economically motivated questions. In so doing, we shall divide the presentation into two broad parts, nicknamed 'deductive' and 'inductive'. The former will deal with the analysis of given dynamical systems, whereas the 'inductive' part will consider the complementary problem of studying economic time series as output of unknown systems.

In particular, we have four tasks before us. First, we divide the 'deductive' part of the paper into two subparts, nicknamed 'geometric' and 'ergodic', and we (thus) discuss two fundamentally different approaches to the study of dynamical systems - the geometric approach (based on the theory of differential/difference equations) and the ergodic approach (based on the axiomatic formulation of probability theory). Second, we discuss the question of predictability in a rigorous manner to provide a very powerful, but abstract way of characterizing chaotic behavior. Third, we introduce specific applications in microeconomics, macroeconomics, and finance, and discuss the policy relevancy of chaos. Finally, we briefly discuss several statistical techniques devised to detect independence, nonlinearity, and chaos in time-series data, and report the evidence of chaotic dynamics in economics and finance.

2 The Geometric Approach

Chaos theory is a very technical subject and a proper understanding of the issues at stake requires that some fundamental concepts and results be discussed in detail.

Generally speaking, in order to generate complex dynamics a deterministic model must have two essential properties: (i) there must be continuous- or discrete-time lags between variables and (ii) there must be some nonlinearity in the functional relationships of the model. In applied disciplines including economics, the first of these features is typically represented by means of systems of differential or difference equations and - even though there exist other mathematical formulations of dynamics which are interesting and economically relevant - in this paper we shall concentrate our attention on them. The geometric (or topological) approach to dynamics, which can be largely identified with the qualitative theory of differential/difference equations, aims at the study of the asymptotic geometric properties of the orbit structure of a dynamical system.

2.1 Continuous and Discrete Dynamical Systems

Typically, a system of ordinary differential equations will be written as[6]

$$\dot{x} = f(x), \qquad x \in \mathbb{R}^n \tag{1}$$

where $f : U \to \mathbb{R}^n$ with U an open subset of $\mathbb{R}^n$ and $\dot{x} \equiv dx/dt$. The vector x denotes the physical (economic) variables to be studied, or some appropriate transformations of them; $t \in \mathbb{R}$ indicates time. In this case, the space $\mathbb{R}^n$ of dependent variables is referred to as *phase space* or *state space*, while $\mathbb{R}^n \times \mathbb{R}$ is called the *space of motions*.

Equation (1) is often referred to as a *vector field*, since a solution of (1) at each point x is a curve in $\mathbb{R}^n$, whose velocity vector is given by $f(x)$. A solution of Equation (1) is a function

$$\phi : I \to \mathbb{R}^n$$

[6]Systems described by Equation (1), in which f does not depend directly on the independent variable t are called *autonomous*. If f does depend on t directly, we shall write

$$\dot{x} = f(x,t), \qquad (x,t) \in \mathbb{R}^n \times \mathbb{R}$$

and $f : U \to \mathbb{R}^n$ with U an open subset of $\mathbb{R}^n \times \mathbb{R}$. Equations of this type are called *non-autonomous*. In economics they are used, for example, to investigate technical progress.

where I is an interval in $\mathbb{R}$ (in economic applications, typically $I = [0, +\infty)$), such that ϕ is differentiable on I, $[\phi(t)] \in U$ for all $t \in I$, and

$$\dot{\phi}(t) = f[\phi(t)], \qquad \forall\, t \in I.$$

The set $\{\phi(t) \mid t \in I\}$ is the *orbit* of ϕ: it is contained in the phase space; the set $\{(t, \phi(t)) \mid t \in I\}$ is the *trajectory* of ϕ: it is contained in the space of motions. However, in applications, the terms 'orbit' and 'trajectory' are often used as synonyms. If we wish to indicate the dependence on initial conditions explicitly, then a solution of Equation (1) passing through the point x_0 at time t_0 is denoted by

$$\phi(t, t_0, x_0),$$

(if t_0 is equal to zero it can be omitted). For a solution $\phi(t, x_0)$ to exist, continuity of f is sufficient. For such a solution to be unique, it is sufficient that f be continuous and differentiable in U.

We can also think of solutions of ordinary differential equations in a slightly different manner, which is now becoming prevalent in dynamical system theory and will be very helpful for understanding some of the concepts discussed in the following sections. Suppose we denote by $\phi_t(x)$ the point in $\mathbb{R}^n$ reached by the system at time t starting from the point x at time 0, under the action of the vector field f of Equation (1). Then the totality of solutions of (1) can be represented by the *one-parameter family* of maps of the phase-space onto itself, $\phi_t : \mathbb{R}^n \to \mathbb{R}^n$, which is called *phase flow* or, for short, *flow* generated by the vector field f, by analogy with fluid flow where we think of the time evolution as a streamline.

If we now take t as a fixed parameter and considering that, for autonomous vector fields, time-translated solutions remain solutions [i.e., if $\phi(t)$ is a solution of Equation (1), $\phi(t + \tau)$ is also a solution for any $\tau \in \mathbb{R}$], the problem may be formulated as

$$x_{t+1} = T(x_t), \qquad x \in \mathbb{R}^n,\ t \in \mathbb{N} \tag{2}$$

where $T = \phi_\tau$ and τ is the fixed value of the parameter t, normalized so that $\tau = 1$.

Thus, a difference equation like (2) can be derived from a differential equation like (1). This need not be that case and many problems in economics as well as in other areas of research give rise directly to discrete-time dynamical systems. In fact, *non invertible* maps such as the celebrated logistic map extensively discussed later in this essay could not be derived from a system of ordinary differential equations.

Equations like (2) are often referred to as *iterated maps* since its orbit is obtained recursively given an initial condition x_t. For example, if we compose T with itself, then we get the second iterate

$$x_{t+2} = T \circ T(x_t) = T^2(x_t)$$

and by induction on n we get the nth iterate,

$$x_{t+n} = T \circ T^{n-1}(x_t) = T^n(x_t)$$

Hence, by the notation $T^n(x)$, we mean T composed with itself $n-1$ times - not the nth derivative of T or the nth power of T[7].

Notice the following difference between the orbits of continuous-time and those of discrete-time systems: the former are continuous *curves* in the state space, whereas the latter are *sequences of points* in space. Also, the fact that a map is a function implies that, starting from any given point in space, there exists only one forward orbit. If the function is non-invertible, however, backward orbits are not defined[8].

The short-run dynamics of individual orbits can usually be described with sufficient accuracy by means of straightforward numerical integration of the differential equations or iteration of the maps. In applications, however, and specifically in economic ones, we are often concerned not with short-term properties of individual solutions, but with the global qualitative properties of bundles of solutions which start from certain practically relevant subsets of initial conditions. Those global properties, however, can only be effectively investigated in relation to trajectories which are somehow recurrent (i.e., broadly speaking, those trajectories which come back again and again to any region of the state space which they once visited).

In what follows, we shall concentrate mainly on that part of the state space of a system which corresponds to recurrent trajectories in the sense just indicated, which will be made more precise below. Even so, a comprehensive analysis of the global behavior of a nonlinear system may not be possible. In this case, the best research strategy is probably a combination of analytical and numerical investigation, the latter playing very much the same role as experiments do in natural sciences.

[7]As an example, if $T(x) = -x^3$, then $T^2(x) = T \circ T(x) = -(-x^3)^3 = x^9$ and $T^3(x) = T \circ T \circ T(x) = T \circ T^2(x) = -(x^9)^3 = -x^{27}$.

[8]A map is invertible if and only if it is one-to-one. For example, tha map $T\colon \mathbb{R} \to \mathbb{R}$ defined by $T(x) = x^2$ is not one-to-one, since $T(1) = 1 = T(-1)$. However, the map $T\colon [0,\infty) \to \mathbb{R}$ defined by $T(x) = x^2$ is one-to-one (and therefore invertible).

2.2 Conservative and Dissipative Systems.

Dynamical systems, whether of a continuous or of a discrete type, can be classified into conservative and dissipative ones. A system is said to be *conservative*, if certain physical properties of the system remain constant in time. Formally, we may say that the flow associated with the system given by Equation (1) preserves volumes if at all points the so-called *Lie derivative* (or *divergence*) is zero, i.e., we have

$$\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} = 0.$$

Analogously, the map given by Equation (2) is said to preserve volumes in the state space if we have at all points

$$\mid \det D_x T(x) \mid = 1$$

where $D_x T(x)$ denotes the matrix of partial derivatives of $T(x)$.

An especially interesting class of conservative systems is formed by Hamiltonian systems. A continuous-time (autonomous) system of ordinary differential equations like (1), $\dot{x} = f(x)$, where $x = (k,q),\ k,q \in \mathbb{R}^n$ is said to be *Hamiltonian* if it is possible to define a continuous function $H(k,q)\colon \mathbb{R}^{2n} \to \mathbb{R}$ - called *Hamiltonian function* - such that

$$\begin{aligned} \dot{k} &= (\partial H/\partial q) \\ \dot{q} &= -(\partial H/\partial k). \end{aligned}$$

If we consider that $(d/dt)H(k,q) = (\partial H/\partial k)\dot{k} + (\partial H/\partial q)\dot{q} = 0$, we can deduce that $H(k,q)$ is constant under the flow.

From the fact that in conservative systems volumes remain constant under the flow (or map), we may deduce that those systems cannot have attracting regions in the phase space, i.e., there can never be asymptotically stable fixed points, or limit cycles, or strange attractors. Since strange attractors (to be defined later) are the main object of our investigation and conservative systems are relatively rare in economic applications, we shall not pursue their general study here[9].

[9] One interesting economic example of a conservative dynamical system is the well-known model of infinite horizon optimal growth, which can be formulated as follows

$$\max_{\dot{k}} \int_0^\infty u(k,\dot{k}) e^{-\rho t} dt$$

with $(k,\dot{k}) \in S \subset \mathbb{R}^{2n}$, and $k(0) = k_0$. In this formulation, u is a concave utility function, k denotes

Unlike conservative ones, *dissipative* dynamical systems, on which most of this essay concentrates, are characterized by contraction of phase space volumes with increasing time. Dissipation can be formally described by the property that divergence is negative, i.e., we have

$$\sum_{i=1}^{n} \frac{\partial f_i(x)}{\partial x_i} < 0$$

or, in the discrete-time case,

$$\mid \det D_x T(x) \mid < 1.$$

Because of dissipation, the dynamics of a system whose phase space is n-dimensional, will eventually be confined to a subset of dimension smaller than n. Thus, in sharp contrast to the situation encountered in conservative systems, dissipation permits one to distinguish between *transient* and *permanent* behavior. For dissipative systems, the latter may be quite simple even when the number of phase space variables is very large.

To better understand this point, think of an n-dimensional system of differential equations characterized by a unique, globally asymptotically stable equilibrium point. Clearly, for such a system, the flow will contract any n-dimensional set of initial conditions to a zero-dimensional final state, a point in $\mathbb{R}^n$. Think also of an n-dimensional ($n \geq 2$) dissipative system characterized by a unique, globally stable limit cycle. Here, too, once the transients have died out, we are left with a one-dimensional orbit, the cycle.

---

the capital stock, $\dot{k}$ is net investment, $\rho \in \mathbb{R}^+$ is the positive discount rate, and the set S is convex and embodies the technological restrictions.

To attack this problem by means of the Pontryagin Maximal Principle, we must first of all introduce an auxiliary vector-valued variable $q \in \mathbb{R}^n$ and define a function

$$H(k,q) \equiv \max_{\dot{k};(k,\dot{k})\in S} \{u(k,\dot{k}) + q\dot{k}\}.$$

where the variables q can be interpreted as prices of investment goods and the Hamiltonian function $H(k,q)$ can be interpreted as the (maximum) current value of national income, evaluated in terms of utility. The necessary (though not sufficient) condition for maximization is that the time evolution of k and q satisfies the following system of differential equations

$$\dot{k} = \frac{\partial H}{\partial q} \qquad \text{and} \qquad \dot{q} = -\frac{\partial H}{\partial k} + \rho q$$

which can be thought of as a Hamiltonian system, plus a (linear) perturbation given by the term ρq.

The asymptotic, permanent regime of a dissipative system is the only observable behavior, in the sense that it is not ephemeral, can be repeated and therefore be 'seen' (i.e., on the screen of a computer), and is often easier to investigate than the overall orbit structure. Even though transients may sometimes last for a very long time and their behavior may be an interesting subject for investigation, for dissipative systems we shall concentrate instead on the long-run behavior of the system, ignoring the transient behavior associated with the start up of the system. That is, we shall consider only the attractor (or attractors, in general) to which trajectories from a range of initial conditions are attracted, to understand the asymptotic properties of a dynamical system. That is, we shall concentrate on the asymptotic properties of a dynamical system, devoting our attention mainly to the attractors of a system, i.e., to the sets of points to which trajectories starting from a range of initial conditions tend as time goes by.

2.3 Invariant and Attracting Sets

To discuss recurrence properties of orbits of a dynamical system, we shall start from the notion of invariant sets. Such sets play an important role in the organization of system orbits in the state space and their investigation is an indispensable first step in the study of the dynamics of a system. Formally, for the discrete dynamical system given by Equation (2), we say that the set $S \subseteq X$ is *invariant* under the action of the map T, if we have:

$$
\begin{aligned}
\phi_t(S) &\subseteq S, \qquad \forall\, t \in \mathbb{R} \\
[\text{respectively, } T^n(S) &\subseteq S, \qquad \forall\, n \in \mathbb{N}]
\end{aligned}
$$

This says specifically that as we apply the map T to any point of S, then we obtain yet another point of S.

When constructing a mathematical model of the time evolution of certain physical, or economic variables, we often wish to impose constraints on the set A of 'reasonable' values of those variables. For example, quantities such as capital stock, consumption or relative prices should remain positive, or at least non-negative for all times; quantities such as the saving ratio or the ratio between factor remunerations and total income should be between zero and one at all times, etc. In other words, we want the 'acceptable' set A to be invariant. The invariance of A is a necessary (although not sufficient) condition for the validity of a dynamical model and in particular of its, implicit or explicit, adjustment mechanisms.

In most cases of practical interest, however, finding invariant sets is not enough. We also wish to locate the region(s) of the state space which ultimately capture all the orbits originating in a certain (not too small) domain. For this purpose, we suggest the following definition.

Definition 1 *A closed invariant set A is said to be an attracting set if for every open set $V \supset A$, there exists an open neighborhood U of A, such that for all $x \in U$ (except perhaps certain subsets of Lebesgue measure zero)*, $T^n(x) \in V$ *for all* $n > N > 0$ *and* $T^n(x) \to A$ *for* $n \to \infty$.

For an attracting set we can also define the *basin* (or *domain*) *of attraction*, as the set of points each of which gives rise to an orbit that is caught by the attracting set. Formally, we can define the basin of attraction as the set $B = \bigcup_{t \leq 0} \phi_t(V)$ [for maps, $B = \bigcup_{n \leq 0} G^n(V)$].

The fact that a set is attracting does not mean that all its parts are attracting too. Therefore, in order to describe the asymptotic regime of a system, we need the stronger concept of *attractor*. A desirable property of an attractor - as distinguished from an attracting set - is *indecomposability,* or *irreducibility*. This property obtains when an orbit, starting from any point on the attractor, as time goes by gets arbitrarily close to any other point. Strangely enough, there is no straightforward and universally adopted definition of attractor, and although the properties of the simpler cases can be easily dealt with, more complicated types of attractors present difficult conceptual problems. In an operational, non-rigorous sense, an attractor is a set on which experimental points generated by a map (or a flow) accumulate for large t[10]. We shall retain this broad, operational definition here, deferring a more sophisticated discussion of the question of attractiveness and observability to the part of this paper concerning the ergodic approach[11].

The simplest type of an attractor is a stable *fixed point*, or, using a terminology more common in economics, a stable *equilibrium*[12]. Ascertaining the existence

[10] See, for example, Eckmann and David Ruelle (1985, pp.623).

[11] The notion of attractiveness is intimately related to that of 'stability' of orbits. Given the vastity of the subject, we cannot deal with it in any detail here and shall refer the reader to the relevant bibliography - for a recent, very clear discussion of this topic, see Paul Glendinning (1994).

[12] In the recent times, economists often use a notion of equilibrium somewhat different from that of mathematicians and physicists, sometimes labelled 'dynamic' or 'sequence equilibrium'. Broadly speaking, the latter implies that certain conditions hold (in a nutshell, 'all markets clear') at all times, while the system evolves in time. As one of these author argued elsewhere [Alfredo Medio (1992, pp. 11-12)], this representation in fact implies the presence of two dynamic

of a fixed/equilibrium point mathematically amounts to finding the solutions of a system of algebraic equations. In the continuous-time case $\dot{x} = f(x)$,the set of equilibria is defined by $E = \{\overline{x} \mid f(\overline{x}) = 0\}$, i.e., the set of values of x such that its rate of change in time is 0. Analogously, in the discrete-time case $x_{t+1} = T(x_t)$, we have $E = \{\overline{x} \mid \overline{x} - T(\overline{x}) = 0\}$, i.e., the values of x which are mapped to themselves by T.

As an example, consider the 'logistic' map

$$\begin{aligned} x_{t+1} &= T_r(x_t) \\ &= r x_t(1 - x_t), \qquad x \in [0,1],\ r \in (0,4]. \end{aligned} \tag{3}$$

To find the fixed points of (3), we put $x_{t+1} = x_t = \bar{x}$ and solve for $\bar{x}$, finding $\bar{x}_1 = 0$ and $\bar{x}_2 = 1 - 1/r$ - see Figure 1.

To get some idea of the importance of fixed points, in Figure 2 we plot the phase diagram of the logistic map for different values of the *tuning* (or *control*) *parameter*, r. Notice that the height of the phase curve hill depends on the value r. For $r < 1$, the only fixed point in the interval $0 \leq x \leq 1$ is $\overline{x} = 0$, but for $r > 1$, there are two fixed points. Using graphical iteration (an algorithmic process of drawing vertical and horizontal segments first to the phase curve and then to the diagonal, $x_{t+1} = x_t$, which reflects it back to the curve), it is easy to show that all trajectories for starting values in the interval $0 \leq x \leq 1$ and for $r < 1$ approach the final value $\overline{x} = 0$. The point $\overline{x} = 0$ is the attractor for those orbits and the interval $0 \leq x \leq 1$ is the basin of attraction for that attractor.

In general, we can examine the dynamical information contained in the derivative of the map at the fixed point, $T'(\overline{x})$. If $|T'(\overline{x})| \neq 1, \overline{x}$ is called *hyperbolic fixed point*. In fact a fixed point $\overline{x}$ is *stable* (or *attracting*) if $|T'(\overline{x})| < 1$, *unstable* (or *repelling*) if $|T'(\overline{x})| > 1$, and *superstable* (or *superattractive*) if $|T'(\overline{x})| = 0$ - superstable in the sense that convergence to the fixed point is very rapid. Fixed points whose derivatives are equal to one in absolute value are called *nonhyperbolic* (or *neutral*) *fixed points*.

Next in the scale of complexity of invariant sets, we consider *stable periodic solutions*, or *limit cycles*. For maps, a point $\overline{x}$ is a periodic point of T with period k, if $T^k(\overline{x}) = \overline{x}$ for $k > 1$ and $T^j(\overline{x}) \neq \overline{x}$ for $0 < j < k$. In other words, $\overline{x}$ is a periodic point of T with period k if it is a fixed point of T^k. In this case we

mechanisms: a short-run, often hidden, dynamics generating a temporary equilibrium ('market clearance'), and a long-run dynamics describing the evolution of equilibria in time, not necessarily converging to a stationary state (fixed point).

say that $\bar{x}$ has period k under T, and the orbit is a sequence of k distinct points $\{\bar{x}, T(\bar{x}), \ldots, T^k(\bar{x})\}$ which, under the iterated action of T, are repeatedly visited by the system, always in the same order. Since all points between $\bar{x}$ and $T^k(\bar{x})$ are also period k points, the resulting sequence is known as a *period k cycle* or alternatively a *k-period cycle*. Notice that k is the *least* period - if $k = 1$, then $\bar{x}$ is a fixed point for T[13].

The third basic type of attractor is called *quasiperiodic***.** If we consider the motion of a dynamical system after all transients have died out, the simplest way of looking at a quasiperiodic attractor is to describe its dynamics as a mechanism consisting of two or more independent periodic motions - see Robert Hilborn (1994, pp. 154-157) for a non-technical discussion. Quasiperiodic orbits can look quite complicated, since the motion never exactly repeats itself (hence, *quasi*), but the motion is not chaotic (as it was wrongly once conjectured). Notice that quasiperiodic dynamics have been found to occur in economically motivated dynamical models - see, for example, Hans-Walter Lorenz (1993), Medio (1992, chapter 12), and Medio and Giorgio Negroni (1996).

Attractors with an orbit structure more complicated than that of periodic or quasiperiodic systems are called *chaotic* or *strange attractors***.** The strangeness of an attractor mostly refers to its geometric characteristic of being a 'fractal' set, whereas chaotic is often referred to a dynamic property, known as 'sensitive dependence on initial conditions', or equivalently, 'divergence of nearby orbits'. Notice that strangeness, as defined by fractal dimension, and chaoticity, as defined by sensitive dependence on initial conditions, are independent properties. Thus, we have chaotic attractors that are not fractal and strange attractors that are not chaotic.

As we shall see, separation of nearby orbits, or, equivalently, amplification of errors is the basic mechanism that makes accurate prediction of the future course of chaotic orbits impossible, except in the short run. On the other hand, as chaotic attractor are bounded objects, the expansion that characterizes their orbits must be accompanied by a 'folding' action that prevents them to escape to infinity. The coupling of 'stretching and folding' of orbits is the distinguishing feature of chaos and it is at the root of both the complexity of its dynamics and the 'strangeness' of its geometry.

In dissipative systems, a chaotic attractor typically arises when the overall con-

[13]For example, the point 1 lies on a 2-cycle for $T(x) = -x^3$, since $T(1) = -1$ and $T(-1) = 1$. Similarly, the point 0 lies on a 3-cycle for $T(x) = -\frac{3}{2}x^2 + \frac{5}{2}x + 1$, since $T(0) = 1$, $T(1) = 2$, and $T(2) = 0$.

traction of volumes, which characterizes those systems, takes place by shrinking in some directions, accompanied by (less rapid) stretching in the others. However, one-dimensional, non-invertible maps that generate chaotic orbits characterized by sensitive dependence on initial conditions - such as, for example, the logistic map - pose a puzzling problem. Strictly speaking, they are not conservative or dissipative: they might indeed be called ‘anti-dissipative’. These maps only have a stretching action and their output remains bounded due to the effect of the (nonmonotonic) nonlinearity. We could think of these maps as limit cases of (dissipative) two-dimensional, invertible maps with very strong contraction in one direction - so strong that, in the limit, only one dimension is left, along which nearby orbits separate.

In what follows, we shall discuss the ‘fractal’ property of chaotic attractor briefly, whereas the ‘sensitive dependence on initial conditions’ property of chaos will be given greater attention here and in the ergodic section of the paper, since this property of chaos is, in our opinion, the most relevant to economics.

2.4 Fractal Dimension

The term ‘fractal’ was coined by Benoit Mandelbrot (1985) and it refers to geometrical objects characterized by non-integral dimensions and ‘self-similarity’. Intuitively, a snowflake can be taken as a natural fractal[14]. The problem of defining measurement criteria finer than the familiar Euclidean dimensions (length, area, volume) in order to quantify the geometric properties of ‘broken’ or ‘porous’ objects was tackled by mathematicians long before the name and properties of fractals became popular. There now exists a rather large number of criteria for measuring qualities that otherwise have no clear definition (such as, for example, the degree of roughness or brokenness of an object), but we shall limit ourselves here to discuss the simplest type concisely.

Let S be a set of points in a space of Euclidean dimension p (think, for example, of the points on the real line generated by the iterations of a one-dimensional map). We now consider certain boxes of side ε (or, equivalently, certain spheres of radius ε), and calculate the minimum number of such cells, $N(\varepsilon)$, necessary to ‘cover’ S. Then, the *fractal dimension* D of the set S will be given by the following limit (assuming it exists)

[14] The term *fractal* comes from the Latin *fractus* which means broken.

$$D \equiv \lim_{\varepsilon \to 0} \frac{\log(N(\varepsilon))}{\log(1/\varepsilon)} \tag{4}$$

The quantity defined in Equation (4) is also called the (Kolmogorov) *capacity dimension*. It is easily seen that, for the most familiar geometrical objects, it provides perfectly intuitive results. For example, if S consists of just one point, $N(\varepsilon) = 1$ and $D = 0$; if it is a segment of unit length, $N(\varepsilon) = 1/\varepsilon$, and $D = 1$; if it is a plane of unit area, $N(\varepsilon) = 1/\varepsilon^2$ and $D = 2$; finally, if S is a cube of unit area, $N(\varepsilon) = 1/\varepsilon^3$ and $D = 3$, etc. That is to say, for 'regular' geometric objects, dimension D does not differ from the usual Euclidean dimension, and, in particular, D is an integer.

The fractal dimension, however, is not always an integer. Let us consider the fractal called *Cantor set* (or *Cantor dust)* - named after the German mathematician George Cantor (1845-1918). To make a Cantor set, start with a line segment of unit length. Remove the middle third and repeat this process without end, each time on twice as many line segments as before. The Cantor set is the set of points that remains, which are infinitely many but their total length is zero. What is the fractal dimension of the Cantor set? By making use of the notion of capacity dimension, we shall have $N(\varepsilon) = 1$ for $\varepsilon = 1$, $N(\varepsilon) = 2$ for $\varepsilon = 1/3$, and, generalizing, $N(\varepsilon) = 2^n$ for $\varepsilon = (1/3)^n$. Taking the limit for $n \to \infty$ (or, equivalently, taking the limit for $\varepsilon \to 0$), we can write

$$D = \lim_{\substack{n \to \infty \\ (\varepsilon \to 0)}} \frac{\log 2^n}{\log 3^n} \simeq 0.63$$

We have thus quantitatively characterized a geometric set that is more complex than the usual Euclidean objects. Indeed the dimension of the Cantor set is a non-integer. We might say that the Cantor dust is an object 'greater' than a point (dimension 0) but 'smaller' than a segment (dimension 1). It can also be verified that the Cantor set is characterized by self-similarity.

Let us consider another fractal, namely the Sierpinski triangle - named after the Polish mathematician Vaclav Sierpinski (1882-1969). To construct a Sierpinski triangle, we start with an equilateral triangle of unit side length. Connect the inner midpoints of the sides with lines and remove the inner triangle of the four equal triangles. Repeat this process to infinity, each time on three times as many triangles as before. At its infinite stage of growth, when the Sierpinski triangle is complete and fully grown, it will consist of an infinite number of triangles with a total area of zero. What is the value of D for the Sierpinski triangle? Since after

n stages we are left with $N(\varepsilon) = 3^n$ triangles of side length $\varepsilon = 1/2^n$, taking the limit for $n \to \infty$ (or, equivalently, $\varepsilon \to 0$) we have

$$D = \lim_{\substack{n\to\infty \\ (\varepsilon\to 0)}} \frac{\log 3^n}{\log 2^n} \simeq 1.584$$

Here, D is smaller than 2, in spite of the fact that the triangle is embedded in two dimensions[15].

The concept of fractal dimension is useful in the geometric analysis of dynamical systems, because it can be conceived of as a measure of the way trajectories fill the phase space under the action of a flow or a map. A non-integer fractal dimension, for example, indicates that trajectories of a system fill up less than an integer subspace of the phase space - see Medio (1992, chapter 7) for a non-rigorous, but intuitive discussion. Also, the concept of fractal dimension is useful in the quantitative analysis of chaotic attractors. For example, the dimension of the attractor of a system [as measured by (4)] can be taken as an index of complexity, as indicated by the essential dimension of the system.

2.5 Lyapunov Exponents

To provide a rigorous characterization, as well as a way of measuring sensitive dependence on initial conditions, we shall now discuss a powerful conceptual tool known as Lyapunov exponents. They provide an extremely useful tool for characterizing the behavior of nonlinear dynamical systems. They measure the (infinitesimal) exponential rate at which nearby orbits are moving apart. A positive Lyapunov exponent is an operational definition of chaotic behavior[16].

Although Lyapunov exponents could be discussed in a rather general framework, we shall deal with the issue in the context of one-dimensional maps, since they are by far the most common type of dynamical system encountered in economic applications of chaos theory. Consider, therefore, the map given by Equation (2), with $T : U \to \mathbb{R}$, U being a subset of $\mathbb{R}$. We want to describe the evolution in time of two orbits originating from two nearby points x_0 and $x_0 + \varepsilon$ (where ε is

[15]Similarly, the value of D for the Sierpinski gasket (a 3-dimensional version of the Sierpinski triangle) is 2 and the value of D for the Koch snowflake (named after the Swedish mathematician Helge von Koch who proposed it in 1904) is 1.261... – see Hans Lauwerier (1991) for a discussion of these and other fractal objects.

[16]Notice, however, that it is possible to have sensitive dependence on initial conditions with orbit divergence less than exponential. In this case, no Lyapunov exponent will be positive.

the difference, assumed to be infinitesimally small, between x_0 and $x_0+\varepsilon$). If we apply the map function T, n times to each point, the difference between the results will be related to ε as follows:

$$d_n = e^{n\lambda(x_0)}\varepsilon$$

where d_n is the difference between the two points after they have been iterated by the map T, n times and $\lambda(x_0)$ is the rate of convergence or divergence.

Taking the logarithm of the above equation and solving for $\lambda(x_0)$ gives

$$\lambda(x_0) = \frac{1}{n}\log|\frac{d_n}{\varepsilon}|.$$

Asymptotically, we shall have[17]

$$\begin{aligned}
\lambda(x_0) &= \lim_{n\to\infty}\frac{1}{n}\log|\frac{d_n}{\varepsilon}| \qquad (5)\\
&= \lim_{n\to\infty}\frac{1}{n}\log|T'(x_{n-1)}...T'(x_1)T'(x_0)|\\
&= \lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}\log|T'(x_j)|
\end{aligned}$$

The quantity $\lambda(x_0)$ is called Lyapunov exponent. Note that the right hand side of (5) is an average along an orbit (a time average) of the logarithm of the derivative[18]. From Equation (5), the interpretation of $\lambda(x_0)$ is straightforward: it is the (local) asymptotic exponential rate of divergence of nearby orbits[19].

As an example, let

$$T_\Lambda(x) = \begin{cases} 2x & \text{for} \quad 0 \le x \le 1/2 \\ 2(1-x) & \text{for} \quad 1/2 \le x \le 1 \end{cases} \qquad (6)$$

be the symmetric 'tent' map. Clearly, $\lambda(x_0)$ is not defined if x_0 is such that $x_j = T_\Lambda^j(x_0) = 1/2$ for some j (because the derivative is not defined). For other points $x_0 \in [0,1]$, $|T_\Lambda'(x_j)| = 2$ for all j, so that $\lambda(x_0) = \log 2$.

[17]Notice that $\frac{d_n}{\varepsilon} = T'(x_{n-1})...T'(x_1)T'(x_0)$.

[18]Notice that, in general, Lyapunov exponents depend on the selected initial conditions. We shall see later under what conditions they may be independent of them.

[19]It is local, since we evaluate the rate of separation in the limit, as $\varepsilon \to 0$. It is asymptotic, since we evaluate it in the limit of indefinitely large number of iterations, as $n \to \infty$ - assuming that the limit exists.

As another example, consider the logistic map, $T_r(x)$, given by Equation (3). Since $T_r'(x_j) = r(1-2x_j)$, the Lyapunov exponent is given by

$$\begin{aligned}\lambda(x_0) &= \lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}\log\left|r(1-2x_j)\right| \\ &= \log r + \lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}\log|1-2x_j|\end{aligned}$$

Clearly, if $x_0 = 0$ or 1, then $\lambda(x_0) = \log r$. For points $x_0 \in (0,1)$ and for $r = 4$, $\lambda(x_0) = \log 2$.

The sign of Lyapunov exponents is especially important to classify different types of dynamical behavior. In particular, the presence of a positive Lyapunov exponent signals that nearby orbits diverge exponentially in the corresponding direction. In its turn, this indicates that observation errors will be amplified by the action of the map. We shall see in what follows that the presence of a positive Lyapunov exponent is intimately related to the lack of predictability of dynamical systems, and thus it is an essential feature of chaotic behavior[20].

2.6 Topological Conjugacy

Before we move on, we shall discuss a fundamental type of equivalence relation between maps, called *topological conjugacy*. It plays an important role in the study of dynamical systems, since it shows that two apparently different systems may actually be dynamically equivalent.

Definition 2 *Let $f : X \to X$ and $g : Y \to Y$ be two maps. A homeomorphism*[21] *$h : X \to Y$ is called a topological conjugacy (or topological equivalence) if $h \circ f = g \circ h$. We also say that f and g are topologically conjugate by h, or that f and g are conjugate.*

The relationship can be described pictorially as a commutative diagram

[20]The calculation of Lyapunov exponents in the general, multidimensional case is more complex and cannot be discussed here in any detail.

[21]A map $h : X \to Y$ is a *homeomorphism* if and only if h is continuous, one-to-one, onto, and has a continuous inverse. In this case, we say that the domain and codomain are homeomorphic to one another.

$$\begin{tikzcd}
x \in X \arrow[r, "f"] \arrow[d, "h"'] & f(x) \in X \arrow[d, "h"] \\
h(x) \in Y \arrow[r, "g"] & h(f(x)) = g(h(x)) \in Y
\end{tikzcd}$$

Broadly speaking, the diagram says that if we start with an element $x \in X$ in the upper left corner and then follow the arrows in either possible direction, we always end up at the same element $h(f(x)) = g(h(x)) \in Y$ in the lower right corner. The homeomorphic property of h and the condition $h \circ f = g \circ h$ guarantee that the dynamics of f on X and that of g on Y are essentially the same - see Clark Robinson (1995) for more details.

As an example, it is easy to show that the logistic map (with $r = 4$), $T_4(x)$, on $[0,1]$ and $T_\Lambda(x)$ on $[0,1]$ are topologically conjugate by $h(x) = \sin^2(\pi x/2)$. Clearly, h transforms the unit interval $[0,1]$ to itself in a one-to-one fashion, it is continuous, and its inverse, h^{-1}, is also continuous. To prove the conjugacy, it suffices to show that $T_4(x) \circ h = h \circ T_\Lambda(x)$.

2.7 Transition to Chaos

In the previous sections, we have provided a classification of attractors and discussed the distinct properties of chaotic attractors. The relevance of these procedures would be greatly enhanced if, in addition, we could describe the qualitative *changes* in the orbit structure of the system which take place when the control parameters are varied. In this way, we would obtain not only a snapshot of chaotic dynamics, but also a description of its emergence. Moreover, if we could provide a rigorous and exhaustive classification of the ways in which complex behavior may appear, transition to chaos could be predicted theoretically, and potentially turbulent mechanisms could be detected in practical applications - and their undesirable effects could be avoided by acting on the relevant parameters.

Unfortunately, the present state of the art does not permit us to define the prerequisites of chaotic behavior with sufficient precision and generality. In order to forecast the appearance of chaos in a dynamical system, we are for the time being left with a limited number of theoretical predictive criteria and a list of certain typical (but by no means exclusive) 'routes to chaos'. Typically, transition to chaos takes place through *bifurcations*. A bifurcation is an essentially nonlinear phenomenon and describes a qualitative change in the orbit structure of a (discrete or continuous-time) dynamical system when one or more parameter is changed. Bifurcation theory is a vast and complex area and we shall consider it here only incidentally.

There exist various types of routes to chaos, generated by so-called codimension one bifurcations (that is, bifurcations depending on a single parameter). In what follows, we shall only (briefly) deal with *period-doubling*, probably the best known route to chaos at least in the economics literature - see, for example, William Baumol and Jess Benhabib (1989). For a discussion of other routes to chaos (such as intermittency and the quasiperiodic route), see Medio (1992, chapter 9).

Period-doubling takes place in both discrete and continuous-time dynamical systems, and can be most simply described by considering the dynamics of the logistic map, $T_r(x)$ given by equation (3), for different values of r. If $r < 1$, the phase curve will lie entirely below the $x_{t+1} = x_t$ line in the positive quadrant - see Figure 3(a) - and $\bar{x} = 0$ is the only fixed point (in fact $\bar{x} = 0$ is an equilibrium for all r). Figures 3(a) and 3(b) give the phase and state space representations of $T_r(x)$ for $r = 0.6$ and $x_0 = 0.01$. Notice that the only fixed point is at $T_r(x) = \bar{x} = 0$.

As r increases beyond 1, $\bar{x} = 0$ loses stability, but a new (positive) fixed point, $\bar{x} = 1 - 1/r$, appears at the intersection of the $x_{t+1} = x_t$ line and the phase curve, as shown in Figure 4(a). In fact, for $r = 2$ the fixed point $\bar{x} = 1 - 1/r$ becomes superstable - since $T_2'(1/2) = 0$. Therefore, for $1 < r < 3$ there are two fixed points: $\bar{x} = 0$, which is unstable, and $\bar{x} = 1 - 1/r$, which is stable. From Figure 4(b) we see that the trajectory approaches some positive unique value (a so-called *single limit point*) between 0 and 1.

As r goes through $r = 3$, a bifurcation called 'flip' occurs and the situation changes. The fixed point $x = 1 - 1/r$ turns into a repeller, since $|T_r'(x)| > 1$, and a stable 2-*cycle* (or an *orbit of period* 2) is born: $x, T_r(x), T_r^2(x) = x$. For example, for $r = 3.2360679775$, there is a superstable orbit of period 2: 0.5, 0.8090169943..., 0.5 - see the state diagram in Figure 5(b).

Let us briefly describe how this happens. For an orbit of period 2 we need to consider the function of $T_r \circ T_r(x)$ - abbreviated $T_r^2(x)$ - and the associated dynamic equation

$$\begin{aligned} T_r^2(x) &= T_r \circ T_r(x) \\ &= r^2 x(1-x)(1-rx(1-x)). \end{aligned} \tag{7}$$

This is again a nonlinear system and its dynamic behavior can be studied as r varies using the same analysis as before. In particular, the fixed points of $T_r^2(x)$ can be found by equating $T_r^2(x)$ with x and solving the resulting 4th order equation. Hence

$$\begin{aligned} x &= T_r^2(x) \\ &= r^2x(1-x)(1-rx(1-x)) \\ &= -r^3x^4+2r^3x^3-(r^2+r^3)x^2+r^2x \end{aligned}$$

whence we can derive the four fixed points, namely:

$$\begin{aligned} \bar{x}_1 &= 0 \\ \bar{x}_2 &= 1-1/r \\ \bar{x}_3 &= \frac{1}{2r}\left(r+1+\sqrt{(r-3)(r+1)}\right) \\ \bar{x}_4 &= \frac{1}{2r}\left(r+1-\sqrt{(r-3)(r+1)}\right) \end{aligned}$$

Clearly, the four fixed points of $T_r^2(x)$ are the two fixed points of $T_r(x)$ and the two elements of the 2-cycle, which have no counterpart in $T_r(x)$ - see the phase diagram in Figure 5(a).

The fixed points of the second-order system (7) are characterized by the derivative of $T_r^2(x)$, $(T_r^2)'(x)$. Since $(T_r^2)'(0)=r^2$ and $(T_r^2)'(1-1/r)=(2-r)^2$, for values of r between 3 and 3.45, each of the fixed points $\bar{x}=0$ and $\bar{x}=1-1/r$ (which are still present) are unstable. The other two fixed points, however, $\bar{x}=\frac{1}{2r}\left(r+1\pm\sqrt{(r-3)(r+1)}\right)$, are both stable, thus implying that each of them locally attracts the dynamics of the second-order system (7).

With respect to Figure 5(a), for r between 3 and 3.45, the trajectories of the first-order system (3) no longer converge to the fixed point $\bar{x}=1-1/r$ (point B), but escape from it and diverge towards the pair of fixed points, $\bar{x}=\frac{1}{2r}\left(r+1\pm\sqrt{(r-3)(r+1)}\right)$ - points D and C, respectively. Any one of them is unstable under the first-order system (3) - since $|T_r'(x)|>1$ at both C and D, so that the trajectories once in any one of these points are initially repelled. Points C and D, however, are stable under the second-order system (7) - since $(T_r^2)'(x)$ at both C and D is less than 1 in absolute value, so that after having moved away from each of C and D in the first step, trajectories come back to each of these points in the second step, thus making the dynamics of system (7) stable with respect to each of C and D. Summarizing, for r between 3 and 3.45, the trajectories of $T_r(x)$ oscillate in the set $\{$C, D$\}$, giving rise to a stable 2-cycle for $T_r(x)$, as it is shown in Figure 5(b). In this case the system is said to undergo a *flip bifurcation* - see John Guckenheimer and Philip Holmes (1983).

If r is increased further, then the two stable fixed points of $T_r^2(x)$ become unstable. In particular, both fixed points of $T_r^2(x)$ will bifurcate at the same r value, leading to an orbit of period 4. In other words

$$\begin{aligned} T_r^2 \circ T_r^2(x) &= T_r^4(x) \\ &= T_r \circ T_r \circ T_r \circ T_r(x) \end{aligned}$$

will have eight fixed points, four of which will be stable. For example, for $r = 3.498561699$ there is a superstable orbit of period 4 : 0.5, 0.874..., 0.383..., and 0.827... - see the phase and state space representations in Figures 6(a) and 6(b).

The same bifurcation scenario will repeat over and over again as r is increased, yielding orbits of period 16, 32, 64, and so on *ad infinitum*. However, the sequence $\{r_\kappa\}$ of values of r at which κ-cycles appear has a finite accumulation point $r_\infty \simeq 3.569946$, involving an infinite number of period doubling bifurcations[22]. The limit set corresponding to r_∞ is a geometric object with a non-integer fractal dimension $\simeq 0.538$ and a Lyapunov exponent equal to zero, and consequently the motion on it is not chaotic in the sense defined above. In fact Mitchell Feigenbaum (1978) discovered that convergence of r to r_∞ is controlled by the universal parameter $\delta \simeq 4.669202$ - known as the *Feigenbaum attractor*. The computation of δ is based on the formula

$$\delta = \lim_{\kappa \to \infty} \left(\frac{r_\kappa - r_{\kappa-1}}{r_{\kappa+1} - r_\kappa} \right)$$

where $(r_\kappa - r_{\kappa-1})$ and $(r_{\kappa+1} - r_\kappa)$ are the distances on the real line between successive flip bifurcations.

Past r_∞, we enter what is usually called the ‘chaotic zone’. For $r_\infty < r < 4$ the model will behave either periodically or aperiodically - in the latter case, the dynamics may be nonchaotic (zero Lyapunov exponent, no sensitive dependence

[22]The values of r for which these transitions from one cycle to another cycle occur, are called *bifurcation points*, the transitions are called *bifurcations*, and the phenomenon is called *period-doubling*.

on initial conditions) or chaotic (positive Lyapunov exponent, sensitive dependence on initial conditions). There is, for example, a tiny interval near $r = 3.83$ (a so-called *window of stability* or *periodicity*) where a stable 3-cycle occurs - see Figures 7(a) and 7(b). Just past $r = 3.83$, the period doubling occurs again, leading to orbits of period 6, 12, 24, and so on, also governed by the Feigenbaum constant. In fact, for r between r_∞ and 4 there is a denumerably infinite number of periodic windows and still an indenumerable number of values of r for which the model behaves aperiodically (chaotically or not). For $r = 4$, we have a completely chaotic orbit, as is illustrated in the state space diagram of Figure 8.

In fact, the different period lengths κ of stable periodic orbits appear in a universal order, with higher-period cycles being associated with higher values of r. In particular, if r_κ is the value of r at which a stable κ-cycle first appears as r is increased, then $r_\kappa > r_q$ if $\kappa \succ q$ (where $\kappa \succ q$ simply means that "κ is listed before q") in the following Sharkovski (1964) ordering (in which we first list the odd numbers except one, then 2 times the odds, 2^2 times the odds, etc., and at the end the powers of 2 in decreasing order - representing the period doubling)

$$\begin{aligned} 3 &\succ 5 \succ \ldots \succ 2\cdot 3 \succ 2\cdot 5 \succ \ldots \succ 2^2\cdot 3 \succ 2^2\cdot 5 \succ \ldots \\ &\succ 2^3\cdot 3 \succ 2^3\cdot 5 \succ \ldots \succ 2^3 \succ 2^2 \succ 2 \succ 1 \end{aligned}$$

This ordering seems strange, but it turns out to be the ordering which expresses which periods imply which other periods. For example, the minimum r value for an orbit of period $\kappa = 2\cdot 3 = 6$ is larger than the minimum r value for an orbit of period $\kappa = 2^2\cdot 3 = 12$, because $6 \succ 12$ in the Sharkovski ordering. One consequence of this ordering is that the existence of a stable κ $(= 3)$-cycle guarantees the existence of any other stable q-cycle for some $r_q < r_\kappa$ - see, for example, Tien-Yien Li and James Yorke (1975).

3 The Ergodic Approach

We have so far been discussing dynamical systems mainly from a geometric (or topological) point of view. This approach, being intuitively appealing and lending itself to suggestive graphical representations, has been tremendously successful in the study of low-dimensional systems, such as, for example, (discrete- and continuous-time) systems with one and perhaps two variables. For higher-dimensional systems, however, the geometric approach has encountered rather formidable obstacles and rigorous results and classifications are few.

Thus, it is sometimes convenient to change perspective and adopt a different approach, based on the axiomatic formulation of probability theory and aimed at the investigation of statistical properties of orbits. This requires the use and understanding of some basic notions and results of set theory and measure theory, which we shall briefly review. We shall see that in many aspects the ergodic theory of dynamical systems parallels the geometric one. Moreover, the ergodic approach is more powerful and effective in dealing with basic questions such as complexity and predictability as well as with the relation between deterministic and stochastic systems.

3.1 Some Elementary Measure Theory

Definition 3 *Let X be a set of points x. A system $\Im$ of subsets of X is called a σ-algebra if*

a. $\emptyset, X \in \Im$*;*

b. $A \in \Im$ *implies* $A^c \in \Im$*; and*

c. $A_n \in \Im, n = 1, 2, \ldots,$ *implies* $\cup A_n \in \Im, \cap A_n \in \Im$.

That is, $\Im$ is a σ-algebra if the null set ($\emptyset$) and X are in $\Im$, the set A and its compliment (A^c) are in $\Im$, and given a sequence $\{A_n\}_{n=1}^{\infty}$ of subsets of X, $A_n \in \Im$, then the union $\cup A_n$ and the intersection $\cap A_n$ are in $\Im$. The space X together with a σ-algebra $\Im$ of its subsets is a *measurable space*, and is denoted by $(X, \Im)$. Since we are dealing here with metric spaces (i.e., with spaces endowed with a distance, such as $\mathbb{R}^n$), among the various σ-algebras, we shall consider the *Borel-σ-algebra*, i.e., the smallest such algebra containing the collections of open (or closed) subsets of X.

Definition 4 *Let $(X, \Im)$ be a measurable space. A real-valued function $\mu = \mu(A), A \in \Im$, taking values in $[0, \infty]$, is a measure if*

a. $\mu(\emptyset) = 0$*;*

b. $\mu(A) \geq 0$ *for all* $A \in \Im$*; and*

c. if $\{A_n\}_{n=1}^{\infty}$ *is a disjoint sequence of* $\Im$*-sets, then* $\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$.

Thus a measure assigns zero to the empty set, is nonnegative, and is countably additive. The triple $(X,\Im,\mu)$ is called a *measure space*. We shall be interested in finite measures (that is, $0 \leq \mu < \infty$) in which case μ can suitably be rescaled so that $\mu(X) = 1$. When $\mu(X) = 1$, $\mu(X)$ is called a *probability measure* and $(X,\Im,\mu)$ is called a *probability space*. In this case X is the sample space (or space) of elementary events, the sets A in $\Im$ are events, and $\mu(A)$ is the probability of the event A.

Two especially important examples of probability measures, which will be used in the sequel, are the *Dirac* and the *Lebesque measures*. The former, also called *Dirac delta* and usually denoted by δ_x is the probability measure that assigns value 1 to all the subsets A of X that contain a given point x, and value zero to all the subsets that do not contain it. Formally, $\delta_x(A) = 1$ if $x \in A$ and $\delta_x(A) = 0$ if $x \notin A$. The *Lebesque measure* on the real line (henceforth denoted by m) is the measure that assigns to each interval its length as its measure[23]. In particular, the Lebesque measure of (a,b), as well as of any of the intervals $(a,b]$, $[a,b]$, or $[a,b)$, is simply its length $|b-a|$.

Definition 5 *Let $(X,\Im,\mu)$ be a probability space. A transformation T of X into X is measurable if, for every $A \in \Im$, $T^{-1}A = \{x : Tx \in A\} \in \Im$.*

Notice that $T^{-1}A$ denotes the pre-image of A – the set of points that are mapped onto A by T in one step.

Definition 6 *A measurable transformation T is said to preserve a measure μ if for every $A \in \Im$, $\mu(T^{-1}A) = \mu(A)$. If T is measure-preserving (with respect to μ), μ is called T-invariant.*

When these concepts are applied to the investigation of dynamical systems, X will typically correspond to the phase space and the elements x of X to states (or positions) of the system. The subsets A (the events) will correspond to certain interesting configurations of orbits of the system in the phase space (such as, for example, fixed points, limit cycles, or strange attractors, basins of attraction, etc.). Finally, the transformation T will correspond to the (flow) map governing the evolution of the state of the system in time[24]. We shall refer to a measure-preserving

[23]Thus, the Lebesgue measure corresponds to the intuitive notion of length (for one-dimensional sets) and volume (for k-dimensional ones). It also provides an intuitive and physically relevant notion of probability.

[24]We can think, for example, of the transformation T as a mechanism for recursively generating a sample $\{x, Tx, \ldots, T^n x\}$ from the domain X from an initial condition $x \in X$. Since, in the present context, there is no essential difference between discrete- and continuous-time dynamical systems, in what follows we shall discuss the issue in terms of maps, i.e., discrete-time systems.

transformation T on the probability space $(X, \Im, \mu)$ as a *dynamical system*, denoted by $(X, \Im, \mu, T)$.

3.2 Ergodicity

As we want to study the statistical properties of orbits generated by measure-preserving transformations, we need to calculate averages over time. In fact, certain basic quantities such as Lyapunov exponents (which, as we have seen above, measure the rate of divergence of nearby orbits) or metric entropy (which, as we shall see below, measures the rate of information production of observations of a system) can be looked at as time averages.

The existence of such averages is guaranteed by the following theorem.

Theorem 1 *(Birkhoff and Khinchin). Let $(X, \Im, \mu)$ be a probability space, T measure preserving and ergodic on X, and f an integrable function. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \widehat{f}(x)$$

exists for μ-almost every $x \in X$[25]*. $\widehat{f}(x)$ is T-invariant, i.e., $\widehat{f}(T(x)) = \widehat{f}(x)$.*

In general, time averages depend on x, meaning that they may be different for orbits originating from different initial states. This happens, for example, when the space X is decomposable (under the action of T), in the sense that there exist, say, two subspaces X_1 and X_2, both invariant with respect to T (i.e., T maps points of X_1 only to X_1 and points of X_2 only to X_2)[26]. It is for this reason that we shall concentrate in a fundamental class of invariant measures that satisfy the requirement of indecomposability, and known as *ergodic measures*. This will ensure that X is (dynamically) indecomposable - a requirement, for it to be called chaotic.

Definition 7 *Given a dynamical system $(X, \Im, \mu, T)$, the measure-preserving transformation T is said to be ergodic (or indecomposable) if $T^{-1}(A) = A$, for some*

[25] 'μ-almost every $x \in X$' means 'all points $x \in X$, except a set of points to which μ assigns zero value'.

[26] The dynamic decomposability of the system - a geometric, or topological fact - is reflected in the existence of T-invariant measures μ that are decomposable in the sense that they can be represented as a weighted average of invariant measures. For example, in the case mentioned above, we can write $\mu = \alpha\mu_1 + (1-\alpha)\mu_2$, where $\alpha \in (0,1)$ and μ_1 and μ_2 may or may not be further decomposed.

$A \in \Im$, implies either $\mu(A) = 1$ or $\mu(A) = 0$. In this case, the T-invariant measure μ is also said to be an ergodic measure for T.

As an example, consider a discrete-time dynamical system characterized by an attracting periodic orbit of period k, $\{x, Tx, \ldots, T^k x = x\}$. In this case, the measure that assigns the value $1/k$ to each point of the orbit is invariant and ergodic.

To discuss ergodic properties of dynamical systems, let $f{:}X \to \mathbb{R}$ be a measurable function, representing a measurement made on the system (such as, for example, the number of times that an orbit generated by the map T and starting from the point x visits a set A when T is iterated n times). As it was argued in the previous section, it is interesting and sometimes necessary to consider the *time average* of f (the average value of f evaluated along the forward trajectory), defined by

$$\widehat{f}(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x).$$

If $\widehat{f}(x)$ exists, it may be thought of as an equilibrium value of f.

Alternatively, we could evaluate the *space* (or *phase*) *average* of f, by considering f as a function of x and multiplying that value by the probability that the system visits the set A. This average is the expectation (or mean value) of $f(x)$ evaluated on the space X

$$\overline{f} = \int_X f(x) d\mu(x).$$

The following result establishes the connection between the time average and the space average.

Theorem 2 *Let $(X, \Im, \mu)$ be a probability space. If the measure-preserving transformation T is ergodic then the limit function $\widehat{f}(x)$ defined in the Birkhoff-Khinchin theorem above is a constant and we have*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \widehat{f}(x) = \int_X f(x) d\mu(x)$$

In words, the theorem states that if T is ergodic, then the time average, $\widehat{f}$ equals the space average $\overline{f}$.

3.3 Lyapunov Exponents Revisited

Clearly, ergodic theory provides an alternative (and often simpler) way of calculating average properties of a dynamical system. In fact, we can use the ideas just discussed, to reformulate the definition of Lyapunov exponents. In particular, if we choose $f(x) = \log|T'(x)|$ and apply the ergodic theorem, then the Lyapunov exponent of a map T can be written as

$$\lambda = \int_X \log\left|T'(x)\right| \mathrm{d}\mu(x)$$

which is independent of the initial condition. The quantity $\log|T'(x)|$ - the natural logarithm of the absolute value of the slope of the curve generated by the map T in the (x_{t+1}, x_t) plane - measures the (exponential) rate at which small discrepancies between trajectories (or small errors) are amplified by the action of the map. In the general case in which that slope varies with x, its different values are weighted by μ. It follows that slopes obtaining over sets of values of x whose μ-measure is zero, do not affect the final results.

All this can be easily illustrated by two simple examples. Let us first compute the Lyapunov exponent for the asymmetric tent map

$$T_{\hat{\Lambda}}(x) = \begin{cases} x/a & \text{for } 0 \leq x \leq a \\ (1-x)/(1-a) & \text{for } a \leq x \leq 1 \end{cases}$$

In this case, it is easy to see that the map $T_{\hat{\Lambda}}(x)$ preserves the Lebesque measure whose density function is $\rho(x) = 1$, implying that $\mu(x) = \int_x \rho(x)dx = \int_x dx$ and $d\mu(x) = dx$. Consequently,

$$\begin{aligned}
\lambda &= \int_0^1 \log\left|T'_{\hat{\Lambda}}(x)\right| d\mu(x) \\
&= \int_0^a \log\left(\frac{1}{a}\right) dx + \int_a^1 \log\left(\frac{1}{1-a}\right) dx \\
&= a\log(\frac{1}{a}) + (1-a)\log(\frac{1}{1-a})
\end{aligned}$$

Clearly, for $a = 1/2$, we are in the case of the symmetric tent map, $T_\Lambda(x)$, and $\lambda = \log 2$.

As another example, let us compute the Lyapunov exponent for the logistic map (for $r = 4$), $T_4(x) = 4x(1-x)$, using its conjugacy with the symmetric tent map, $T_\Lambda(x)$, $h(x) = \sin^2(\pi x/2)$. Since the tent map preserves Lebesque measure, the conjugacy also induces an invariant measure $\mu(x)$ for the logistic map, with density function $\rho(x) = \pi^{-1}[x(1-x)]^{-1/2}$. Consequently, the Lyapunov exponent for the logistic map (for $r = 4$) is given by[27]

$$\begin{aligned} \lambda &= \int_0^1 \log\left|T_4'(x)\right| \frac{\mathrm{d}x}{\pi[x(1-x)]^{1/2}} \\ &= \int_0^1 \frac{\log|4-8x|}{\pi[x(1-x)]^{1/2}} dx \\ &= \log 2. \end{aligned}$$

4 Predictability, Entropy

The rather formidable apparatus described above will allow us to discuss the question of predictability in a rigorous manner. In so doing, however,we must first remove a possible source of confusion. In particular, given that the ergodic approach analyzes dynamical systems by means of probabilistic methods, one might immediately point out that since the outcomes of deterministic dynamical systems are not random events, measure and probability theories are not the appropriate tools of analysis. *Prima facie,* this seems to be a convincing argument – if the system is deterministic, we know the equations of motion, and we can measure its state with infinite precision, then there is nothing left to discuss.

However, infinite precision of observation is a purely mathematical expression, and it has no physical counterpart. When dynamical system theory is applied to real systems, a distinction must be made between *states* of a system, i.e., points in a state space, and *observable states*, i.e., subsets (or cells) of the state space, whose (non-zero) size reflects our limited power of observation[28]. This will be consistent with the fact that in real systems perfect foresight only makes sense when it is interpreted as an asymptotic state of affairs which is approached as economic agents accumulate information and learn about the position of the system. Much of what follows concerns the conditions under which, given precise knowledge of the equations of the system (i.e., given a deterministic system),

[27] Recall that we are using the invariant measure $\mu(x) = \int_X \rho(x)dx = \int_X \frac{1}{\pi[x(1-x)]^{1/2}}dx$, which implies that $d\mu(x) = \frac{1}{\pi[x(1-x)]^{1/2}}dx$.

[28] As will become apparent in the discussion that follows, this distinction is not important for systems whose orbit structure is simple, such as, for example, systems characterized by a stable fixed point or a stable limit cycle. For these systems, that is, the (unrealistic) assumption of infinite precision of observation is a convenient simplification.

but an imprecise, albeit arbitrarily accurate, observation of its state, prediction is possible.

We can now characterize the concept of predictability in a rigorous fashion, in terms of a quantity called entropy. Let $(p_1, ..., p_N)$ be a finite probability distribution, i.e., $p_i \geq 0$ for all i and $p_1 + ... + p_N = 1$, for the occurrence of events $A_1, ..., A_N$. The entropy of this distribution is

$$H = -\sum_{i=1}^{N} p_i \log(p_i),$$

with $0 \log 0 = 0$. H measures the degree of indeterminacy (uncertainty) of an event. It attains its largest value $(\log N)$ for $p_1 = ... = p_N = 1/N$, meaning that the distribution has maximal indeterminacy, and its minimum value (zero) when one of the p's is equal to one, the others being zero[29].

We shall now apply this entropy idea to a description of the state space behavior of a dynamical system, $(X, \Im, \mu, T)$. Let a single trajectory run for a long time to map out an attractor, and let $P = (P_1, \ldots, P_N)$ be a finite μ-measurable partition of X[30]. The entropy of P will be equal to

$$H(P) = -\sum_{i=1}^{N} \mu(P_i) \log \mu(P_i),$$

where $\mu(P_i)$ measures the probability of finding the system in the 'cell' P_i.

However, when dealing with a dynamical system, we are not interested in the entropy of a partition of the state space (i.e., the information in a single experiment), but in the entropy of the system (i.e., the rate at which replications of the experiment produce information). To make this idea more precise, for each P_i, we write $T^{-k}P_i$ for the set of points that led to P_i in k steps. We then denote by $T^{-k}P$ the partition $(T^{-k}P_1, ..., T^{-k}P_N)$, which is deduced from P by time evolution. Finally, we define the 'super-partition'[31]

[29] For example, in a game of dice, the maximum entropy of a throw (the maximum uncertainty about its outcome) obtains when each of the six facets of a die has the same probability $(1/6)$ of turning up. An unfair player can reduce the uncertainty of the outcome by 'loading' the dice and thereby increasing the probability of one or more of the six faces (and correspondingly decreasing the probability of the others).

[30] A partition can also be viewed as a function $P : X \to \{P_1, ..., P_N\}$ such that, for each point of the state space $x \in X$, $P(x)$ is the element of the partition, the cell of X, in which x is contained.

[31] Given two partitions P_1 and P_2, the operation $P_1 \bigvee P_2$ consists of all the possible intersections of the elements of P_1 and P_2 and it is called a 'span'.

$$\bigvee_{i=1}^{n} T^{-i}P \equiv P \vee T^{-1}P \vee \cdots \vee T^{-n}P$$

which is generated by P in a time interval of length n. The entropy of the super-partition, $\bigvee_{i=1}^{n} T^{-i}P$, namely $H(\bigvee_{i=1}^{n} T^{-i}P)$, can be calculated analogously, summing over all the cells of $\bigvee_{i=1}^{n} T^{-i}P$. A moment's reflection will suggest that, whereas an element of the original partition P corresponds to a (approximately observed) state of a dynamical system, an element of the super-partition, $\bigvee_{i=1}^{n} T^{-i}P$, corresponds to a *sequence* of n states.

If we now divide $H(\bigvee_{i=1}^{n} T^{-i}P)$ by the number of observations n, we obtain the average amount of information contained in - the average amount of uncertainty about - the 'super-experiment' consisting in the repeated observation of the system along a typical orbit. If we increase the number of observations indefinitely, we obtain[32]

$$h(T,P) = \lim_{n\to\infty} \frac{1}{n} H(\bigvee_{i=1}^{n} T^{-i}P),$$

which is the entropy of the transformation T with respect to the partition P[33].

From the definitions above, the link between entropy and predictability should be clear. Zero entropy means that, if we observe the state of a dynamical system long enough, although with finite precision (and we know the "true" law of motion), then there is no uncertainty left about the future. On the contrary, positive entropy means that, no matter how long we observe the system, additional observations are informative, i.e., the future is unpredictable.

To investigate this point a little further, $h(T,P)$ can be looked at as the limit of a fraction - the numerator is the entropy of a 'super-partition' obtained by iterating T, and the denominator is the number of iterations. Loosely speaking, if when the number of iterations increases, $H(\bigvee_{i=1}^{n} T^{-i}P)$ remains bounded, the limit will be zero; if it grows linearly with n the limit will be a finite, positive value; if it grows more than linearly, the limit will be unbounded. To interpret this result, consider that each cell of the partition $\bigvee_{i=1}^{n} T^{-i}P$ corresponds to a sequence of length n of cells of P (i.e., to an orbit of length n of the system, observed with P-precision). Considering the definition of entropy, one can verify

[32]See Patrick Billingsley (1965, pp. 81-82) or Ricardo Mañé (1987, pp. 216) for a proof that this limit exists.

[33]In the literature, we also find the expression $h(\mu,P)$ where the system is identified by the invariant measure. In the present context, the two exressions are entirely equivalent.

that $H(\bigvee_{i=1}^{n} T^{-i}P)$ will increase with n linearly according to whether, increasing the number of observations, the number of possible sequences will increase *exponentially*. From this point of view, it is easy to understand why 'simple' systems (i.e., those characterized by attractors which are fixed points or periodic orbits) have zero entropy. Transients apart, for those systems the possible sequences of states are limited and their number does not increases at all with the number of observations. Complex systems are precisely those for which the number of possible sequences of states grows exponentially with the number of observations. For finite-dimensional, deterministic systems characterized by bounded attractors, entropy is bounded above by the sum of the positive Lyapunov exponents and is therefore finite.[34].

So far we have been talking about entropy relative to a specific partition. The entropy of a transformation T, or equivalently the entropy of the T-invariant measure μ, is

$$h(T) = \mathrm{Sup}_P \; h(T,P)$$

where the supremum is taken over all finite partitions. The quantity $h(T)$ is also known as *K*(olmogorov)-*S*(inai), or *metric entropy*. Unless we indicate differently, by entropy we mean K-S entropy. Actual computation of K-S entropy, $h(T)$, directly from its definition looks a rather desperate project. Fortunately, a result from Kolmogorov and Sinai guarantees that, under conditions often verified in specific problems, the entropy of a system $h(T)$ can be obtained from the computation of its entropy relative to a given partition, $h(T,P)$. Formally, we have the

[34] The entropy of a system with respect to a given partition can be given an alternative, very illuminating formulation by making use of the auxiliary concept of *conditional entropy* of (partition) $\mathcal{A}$ given (partition) $\mathcal{B}$, defined by

$$H(\mathcal{A}|\mathcal{B}) = -\sum_{A,B} \mu(A \cap B) \log \; \mu(A|B)$$

where A, B denote elements of the partitions $\mathcal{A}$ and $\mathcal{B}$, respectively. Intuitively, conditional entropy can be viewed as the average amount of uncertainty of the experiment $\mathcal{A}$ when the outcome of the experiment $\mathcal{B}$ is known. It can be shown [see Billingsley (1965, pp. 81-82)] that

$$\lim_{n\to\infty} \frac{1}{n} H(\bigvee_{i=1}^{n} T^{-i}P) = \lim_{n\to\infty} H(P \left| \bigvee_{i=1}^{n} T^{-i}P \right.)$$

This equation provides another useful interpretation of $h(T,P)$: it is *t*he amount of uncertainty of - the amount of information contained in - an observation of the system in the partitioned state space, conditional upon the (finite-precision) knowledge of its state in the infinitely remote past.

following[35]

Theorem 3 *(Kolmogorov-Sinai). Let $(X,\Im,\mu)$ be a probability space; T a transformation preserving μ and P a partition of $(X,\Im,\mu)$ with finite entropy. If $\bigvee_{i=1}^{\infty} T^{-i}P = \Im$ mod 0, then $h(T) = h(T,P)$. In this case, P is called a* generator.

As an example, consider the symmetric tent map given by equation (?) and the partition of the interval consisting of the two sub-intervals located, respectively to the left and to the right of the $1/2$ point. Thus, we have a partition $P = \{P_1, P_2\}$ of $[0,1]$, where $P_1 = \{0 < x < 1/2\}$ and $P_2 = \{1/2 < x < 1\}$. Then the atoms of $T^{-1}P_1$ are the two subintervals $\{0 < x < 1/4\}$ and $\{3/4 < x < 1\}$ and the atoms of $T^{-1}P_2$ are the two subintervals $\{1/4 < x < 1/2\}$ and $\{1/2 < x < 3/4\}$. Hence, taking all possible intersections of subintervals, the span $\{T^{-1}P \bigvee P\}$ consists of the four subintervals $\{0 < x < 1/4\}$, $\{1/4 < x < 1/2\}$, $\{1/2 < x < 3/4\}$, and $\{3/4 < x < 1\}$. Repeating the same operation m times the span $\{\bigvee_{i=0}^{m-1} T^{-i}P\ \}$ is formed by 2^m subintervals of equal length 2^{-m}, defined by $\{x : (i-1)/2^m < x < i/2^m\}$, $1 \leq i \leq 2^m$. Moreover, considering that (in the case of the tent map) the span $\bigvee_{i=0}^{\infty} T^{-i}P$ contains any open subinterval of $[0,1]$ and therefore, if we use the Borel σ-algebra, the selected partition is a generator, we can apply the Kolmogorov-Sinai Theorem and have $h(T) = h(T,P)$. Finally, taking into account the fact that the tent map preserves the Lebesgue measure m (which, we recall, assigns to each measurable set a value equal to its length), we conclude that the K-S entropy of the tent map is

$$
\begin{aligned}
h(T) &= \lim_{m\to\infty} \frac{1}{m} H\left(\bigvee_{i=0}^{m-1} T^{-i}P\right) \\
&= \lim_{m\to\infty} \frac{1}{m}\left[-2^m(2^{-m}\log(2^{-m}))\right] \\
&= \log 2.
\end{aligned}
$$

Before concluding this section, we would like to notice that entropy is closely linked with another type of statistical invariants, the Lyapunov exponents. It can be shown that in general we have the following inequality

$$h(T) \leq \sum_{i:\lambda_i>0} \lambda_i,$$

[35]For discussion and proof of the K-S theorem, see Billinglsey (1965, pp. 84-85) or Mañé (1987, pp. 218-22).

where λ denotes a Lyapunov exponent. For observable chaotic systems, strict equality holds[36]. As we have seen before, the equality indeed holds for the tent map.

The close relation between entropy and Lyapunov exponents is not surprising. We have already observed that entropy crucially depends on the rate at which the number of new possible sequences of 'coarsed-grained' states of the system grows as the number of observations increases. But this rate is strictly related to the rate of divergence of nearby orbits, which is measured by the Lyapunov exponents. Thus, the presence of one positive Lyapunov exponent on the attractor signals positive entropy and unpredictability of the system.

5 Isomorphism

In the discussion of dynamical systems from a geometric point of view, we have encountered the notion of topological equivalence. Analogously, there exists a fundamental type of equivalence relation between measure-preserving transformations, called *isomorphism*, which plays a very important role in ergodic theory and which we shall use in the sequel.

Definition 8 *Two transformations T and $\widehat{T}$ acting, respectively, on the state spaces X and $\widehat{X}$, and preserving, respectively, the measures μ and $\widehat{\mu}$, are isomorphic if a one-to-one and invertible map θ exists such that (excluding perhaps certain sets of measure zero)*

a. $\widehat{T} \circ \theta = \theta \circ T$*; and*

b. the map θ preserves the probability structure, i.e., if I and $\widehat{I}$ are, respectively, measurable subsets of X and $\widehat{X}$, then $\mu(I) = \widehat{\mu} \circ \theta(I)$ and $\widehat{\mu}(\widehat{I}) = m \circ \theta^{-1}(\widehat{I})$.

Certain properties such as ergodicity and entropy are invariant under isomorphism. Consequently, isomorphic transformations have the same entropy[37].

As an example, we will show that the logistic map (with $r = 4$), $T_4(x)$, and the tent map, $T_\Lambda(x)$, are isomorphic, and therefore have the same entropy. By making use of the topological conjugacy, $\theta(x) = \sin^2(\pi x/2)$, we have already shown that $T_4(x)$ and $T_\Lambda(x)$ are topologically conjugate, that is $\theta \circ T_4(x) = T_\Lambda(x) \circ$

[36] For technical details, see Donald Ornstein and Benjamin Weiss (1991, pp. 78-85) or Ruelle (1989, pp. 71-77).

[37] The reverse is true only for a certain class of transformations called Bernoulli.

θ. In general, however, topological conjugacy need not imply measure-theoretic isomorphism. So we still have to prove that $T_4(x)$ preserves a measure ρ such that, for almost all subintervals I of $[0,1]$, $\rho(I) = m(\theta(I))$, where m is the Lebesgue measure which, as we already know, is $T_\Lambda(x)$-invariant.

If these measures are absolutely continuous, the last equation is equivalent to $\rho(dx) = (d\theta/dx)dx$, whence, making use of the definition of $\theta(x)$ and considering that $\cos(\theta) = [1-\sin^2(\theta)]^{1/2}$, we obtain

$$\rho(dx) = \frac{dx}{\pi[x(1-x)]^{1/2}}.$$

Now, considering that the counter-image of each interval I under $T_\Lambda(x)$ consists of two subintervals whose length is half the length of I, it is easily verified that $T_\Lambda(x)$ preserves the Lebesgue measure. If we also consider that $T_4^{-1}(x) = \theta^{-1} \circ T_\Lambda^{-1}(x) \circ \theta$, we can establish that $T_4(x)$ preserves $m\theta$ and therefore it preserves ρ. Since isomorphism preserves entropy, we can conclude that the logistic map $T_4(x)$, has entropy equal to $\log 2 > 0$ and its outcome is therefore unpredictable[38].

The implications for economics of the results just obtained are puzzling. For example, consider the case in which models of optimal growth give rise to dynamic, logistic-type equations with chaotic parameter. The sequences thus generated are optimal in the sense that they solve a problem of intertemporal maximization of rational agents, in an economy satisfying the requirements of competitive equilibrium at each point of time. In the absence of (exogenous) random disturbances, along optimal trajectories agents' expectations are supposed to be always fulfilled. While the latter assumption may be acceptable when the dynamics of the system are simple (i.e., convergence to a steady state or to a periodic orbit), it makes little sense if the dynamics are chaotic.

6 Chaos in Dynamic Economic Models

Chaos represents a radical change of perspective on business cycles. Business cycles receive an endogenous explanation and are traced back to the strong nonlinear deterministic structure that can pervade the economic system. This is different from the (currently dominant) exogenous approach to economic fluctuations, based on the assumption that economic equilibria are determinate and intrinsically stable, so that in the absence of continuing exogenous shocks the economy tends

[38]Notice that, in this case, the metric entropy and the unique Lyapunov exponent are equal.

towards a steady state, but because of stochastic shocks a stationary pattern of fluctuations is observed.

Richard Goodwin was one of the first (back in the 1950's and 1960's) to understand the relevance of chaos theory for economics - see Goodwin (1982), for a collection of the relevant papers. Recently, however, there has been a revival of interest in dynamical systems theory, and there is a group of economists who look at economic fluctuations as deterministic phenomena, endogenously created by market forces, and aggregator (utility and production) functions. They agree with Goodwin that chaos theory has great implications for both theory and policy. For example, chaos could help unify different approaches to structural macroeconomics. As Jean-Michel Grandmont (1985) has shown for different parameter values even the most classical of economic models can produce stable solutions (characterizing classical economics) or more complex solutions, such as cycles or even chaos (characterizing much of Keynesian economics).

In what follows, we shall briefly review some representative theoretical microeconomic and macroeconomic models that predict cycles and chaos as outcomes of reasonable economic hypotheses. Our purpose is not to provide a complete survey of all existing dynamic economic models that predict chaos. The reader that is interested in a more exhaustive survey should also consult William Brock (1988), Michele Boldrin and Michael Woodford (1992), Kazuo Nishimura and Gerhard Sorger (1996), and Pietro Reichlin (1997).

6.1 Rational Choice and Chaos

Benhabib and Richard Day (1981), using a standard micro-framework, showed that rational choice can lead to erratic behavior when preferences depend on past experience. Following Benhabib and Day (1981), consider the (logarithmic representation of the) Cobb-Douglas utility function

$$u(x,y;\alpha) = \alpha \log x + (1-\alpha) \log y$$

with $0 < \alpha < 1$. Maximizing subject to (the usual budget constraint)

$$p_x x + p_y y = I \tag{8}$$

yields the Marshallian demand functions

$$x = \alpha \frac{I}{p_x} \qquad \text{and} \qquad y = (1-\alpha)\frac{I}{p_y} \tag{9}$$

Assuming, however, that preferences depend on past experience, as in Benhabib and Day (1981), according to a function

$$\alpha_t = r x_{t-1} y_{t-1} \tag{10}$$

where r is an 'experience dependence' parameter, then the demand for x and y is described by a first-order difference equation in x and y, respectively. For example, by substituting (10) into (9) and exploiting the budget constraint (8), the demand for x is obtained (under the assumption of constant prices) as

$$x_t = \frac{rI}{p_x p_y} x_{t-1}(I - p_x x_{t-1}) \tag{11}$$

Clearly, Equation (11) describes a one-humped curve like the logistic map (4). In fact, for $p_x = p_y = I = 1$, Equation (11) reduces to Equation (4). Therefore, the specification of experience dependent preferences generates chaotic behavior for appropriate values of the experience dependence parameter, r.

6.2 Descriptive Growth Theory and Chaos

Following Day (1982), we consider the descriptive one-sector model due to Robert Solow (1956). Under the assumptions that aggregate saving equals gross investment and that the capital stock exists for exactly one period, this system can be written as a first-order system in discrete time as

$$(1+\nu)k_{t+1} = s f(k_t) \tag{12}$$

where k is capital per worker, f a neoclassical production function, and the two parameters $\nu > -1$ and $s \in [0,1]$ represent, respectively, the rates of population growth and saving. Under the usual convexity assumptions, the phaseline of Equation (12) is an increasing concave function through the origin, with two fixed points. The trivial steady state at 0 is asymptotically unstable while the other (positive) fixed point is globally stable, attracting orbits that start at any initial value $k_0 > 0$.

Day (1982) extended the above neoclassical one-sector model of capital accumulation by introducing a pollution effect that reduces productivity as in the following (Cobb-Douglas type) production function

$$f(k_t) = B k_t^{\varphi} (\varsigma - k_t)^{\gamma} \tag{13}$$

where $k_t \leq \varsigma =$ constant (acting as a saturation level of capital per worker) and $(\varsigma - k_t)^\gamma$ reflects the effect of pollution on per capita output. In particular, when k increases, pollution also increases and less output can be produced with a given stock of capital than in the standard model. With (13), the neoclassical model (12) becomes

$$(1+\nu)k_{t+1} = sBk_t^\varphi(\varsigma - k_t)^\gamma \tag{14}$$

which for $B = \gamma = \varsigma = 1$ reduces to

$$k_{t+1} = rk_t(1-k_t) \tag{15}$$

where $r = sB/(1+\nu)$. Equation (15) is formally identical with the logistic map (4). Hence, all properties of the logistic map apply here as well. Moreover, the general five-parameter map (14) is also chaotic for appropriate values of the parameters - see Day (1982) or Lorenz (1993) for details.

6.3 Optimal Growth Theory with Money and Chaos

In this section we consider one version of the neoclassical growth model - Miquel Sidrauski's (1967) optimal growth model with money. It is assumed that the economy is composed of a large number of identical infinitely lived households, each maximizing (at time t) a lifetime utility function of the form

$$\sum_{t=0}^{\infty} \beta^t u(c_t, m_t)$$

where c and m are consumption and real money balances per capita. Ignoring capital accumulation, production, and interest-bearing public debt, the representative household's budget constraint for period t is assumed to be

$$P_t(m_t + c_t) = P_t y + H_t + P_{t-1}m_{t-1}$$

where y is a constant endowment and H_t is per capita lump-sum government transfers, assumed to be equal to μM_{t-1} (where $\mu > 0$ is the constant rate of money growth). Assuming additive instantaneous utility, the equilibrium fixed points for the system are obtained by solving the following first-order difference equation [see Costas Azariadis (1993, section 26.3) for more details]

$$m_{t+1}u_c(y, m_{t+1}) = \frac{1+\mu}{\beta}\left[u_c(y, m_t) - u_m(y, m_t)\right] m_t \tag{16}$$

If we drop the separability assumption and instead consider

$$u(c,m) = \frac{\left(c^{1/2}m^{1/2}\right)^{1-\sigma}}{1-\sigma}, \qquad \sigma > 0, \sigma \neq 1$$

where σ is the reciprocal of the intertemporal elasticity of substitution between current and future values of the aggregate commodity $(cm)^{1/2}$, then Equation (16) simplifies to

$$x_{t+1}^{\alpha} = \frac{1+\mu}{\beta} x_t^{\alpha}(1-x_t) \tag{17}$$

where $x_t = y/m_t$ and $\alpha = (\sigma - 3)/2$, assumed to be positive. Equation (17) has a unique positive steady state

$$\bar{x} = 1 - \frac{\beta}{1+\mu} \tag{18}$$

Substituting (18) into (17) to eliminate $(1+\mu)/\beta$, we obtain

$$x_{t+1} = x_t \left(\frac{1-x_t}{1-\bar{x}}\right)^{1/\alpha}$$

which for $\alpha = 1$ reduces to the logistic map - see Kiminori Matsuyama (1991) or Azariadis (1993, section 26.4) for more details regarding the dynamic behavior of this system.

6.4 Policy Relevancy of Chaos

As it has just been shown chaos can be produced, for some parameter settings, from even many of the most classical economic models - including models in which there is continuous market clearing, rational expectations, overlapping generations, perfect competition, no externalities, and no forms of market failure. The issue has been whether or not the parameter settings that can produce chaos are economically 'reasonable'. With large enough nonlinear, dynamic models to be viewed as possible approximations to reality, there are no currently available conclusions regarding the plausibility of the subset of the parameter set that can support chaos.

But there is also the question about whether or not we should care. In positive economics, there is good reason to care. Understanding the behavior of an economy that is chaotic is not possible with a model that is not chaotic, since chaotic solution paths have many properties that cannot be produced from non-chaotic solutions. But on the normative side, the usefulness of chaos is much less clear. Grandmont's (1985) model, for example, produces Pareto optimal chaotic solution paths. The fact that the solutions are chaotic does not alone provide any justification for government intervention, and indeed any such intervention could produce a stable, but Pareto inferior solution. In fact, James Bullard and Alison Butler (1993) have argued that the existing theoretical results on chaos have no policy relevance, since in chaotic models the justification for intervention always can be identified with a form of market failure entered into the structure of the model, and hence the chaos is an independent and policy-irrelevant feature of those models.

There is an exemption, however. Woodford (1989) has argued that chaos might produce increased Pareto-sensitivity to market failure. If that is the case, then there is an interaction between chaos and the policy implications of market failure, with small market failures producing increased Pareto loss, when the economy also is chaotic. This could be an important result and could result in high policy relevancy for chaos, but at present Woodford's speculation remains only a supposition, and has not been confirmed in theory or practice. Hence, at present, the policy relevance of chaos must remain in doubt.

7 Efficient Markets and Chaos

Recently the efficient markets hypothesis and the notions connected with it have provided the basis for a great deal of research in financial economics. A voluminous literature has developed supporting this hypothesis. Briefly stated, the hypothesis claims that asset prices are rationally related to economic realities and always incorporate all the information available to the market. This implies the absence of exploitable excess profit opportunities. However, despite the widespread allegiance to the notion of market efficiency, a number of studies have suggested that certain asset prices are not rationally related to economic realities. For example, Laurence Summers (1986) argues that market valuations differ substantially and persistently from rational valuations and that existing evidence (based on common techniques) does not establish that financial markets are efficient.

Motivated by these considerations, in this section we provide a review of the literature with respect to the efficient markets hypothesis, discuss some of the more recent testing methodologies, and finally consider the intersection between the efficient markets theory and chaos theory.

7.1 The Random Walk Hypothesis

Standard asset pricing models typically imply the *martingale model*, according to which tomorrow's price is expected to be the same as today's price. Symbolically, a stochastic process x_t follows a martingale if

$$E_t(x_{t+1}|\Omega_t) = x_t \tag{19}$$

where Ω_t is the time t information set - assumed to include x_t. Equation (19) says that if x_t follows a martingale the best forecast of x_{t+1} that could be constructed based on current information Ω_t would just equal x_t. Alternatively, the martingale model implies that $(x_{t+1} - x_t)$ is a *fair game* (a game which is neither in your favor nor your opponent's)[39]

$$E_t[(x_{t+1} - x_t)|\Omega_t] = 0. \tag{20}$$

Clearly, x_t is a martingale if and only if $(x_{t+1} - x_t)$ is a fair game. It is for this reason that fair games are sometimes called *martingale differences*[40].

The fair game model (20) says that increments in value (changes in price adjusted for dividends) are unpredictable, conditional on the information set Ω_t. In this sense, information Ω_t is fully reflected in prices and hence useless in predicting rates of return. The hypothesis that prices fully reflect available information has come to be known as the *efficient markets hypothesis*. In fact Eugene Fama (1970) defined three types of (informational) capital market efficiency (not to be confused with allocational or Pareto-efficiency), each of which is based on a different notion of exactly what type of information is understood to be relevant. In particular, markets are weak-form, semistrong-form, and strong-form efficient if the information set includes past prices and returns alone, all public information, and any information public as well as private, respectively. Clearly, strong-form efficiency implies semistrong-form efficiency, which in turn implies weak-form efficiency, but the reverse implications do not follow, since a market easily could

[39] A stochastic process z_t is a fair game if z_t has the property $E_t(z_{t+1}|\Omega_t) = 0$.

[40] The martingale process is a special case of the more general submartingale process. In particular, x_t is a *submartingale* if it has the property $E_t(x_{t+1}|\Omega_t) > x_t$. Note that the submartingale is also a fair game where x_{t+1} is expected to be greater than x_t. In terms of the $(x_{t+1} - x_t)$ process the submartingale model implies that $E_t[(x_{t+1} - x_t)|\Omega_t] > 0$. Stephen LeRoy (1989, pp. 1593-4) also offers an example in which $E_t[(x_{t+1} - x_t)|\Omega_t] < 0$, in which case x_t will be a *supermartingale*.

be weak-form efficient but not semistrong-form efficient or semistrong-form efficient but not strong-form efficient.

The martingale model given by (19) can be written equivalently as

$$x_{t+1} = x_t + \varepsilon_t$$

where ε_t is the martingale difference. When written in this form the martingale looks identical to the *random walk model* - the forerunner of the theory of efficient capital markets. The martingale, however, is less restrictive than the random walk. In particular, the martingale difference requires only independence of the conditional expectation of price changes from the available information, as risk neutrality implies, whereas the (more restrictive) random walk model requires this and also independence involving the higher conditional moments (i.e., variance, skewness, and kurtosis) of the probability distribution of price changes. By not requiring probabilistic independence between successive price changes, the martingale difference model is entirely consistent with the fact that price changes, although uncorrelated, tend not to be independent over time but to have clusters of volatility and tranquility (i.e., dependence in the higher conditional moments) - a phenomenon originally noted for stock market prices by Mandelbrot (1963) and Fama (1965).

7.2 Tests of the Random Walk Hypothesis

The random walk and martingale hypotheses imply a unit root in the level of the price or logarithm of the price series - notice that a unit root is a necessary but not sufficient condition for the random walk and martingale models to hold. Hence, these models can be tested using recent advances in the theory of integrated regressors. The literature on unit root testing is vast and, in what follows, we shall only briefly illustrate some of the issues that have arisen in the broader search for unit roots in financial asset prices.

Charles Nelson and Charles Plosser (1982), using the augmented Dickey-Fuller (ADF) unit root testing procedure [see David Dickey and Wayne Fuller (1981)] test the null hypothesis of *difference-stationarity* against the *trend-stationarity* alternative. In particular, in the context of financial asset prices, one would estimate the following regression

$$\Delta y_t = \alpha_0 + \alpha_1 y_{t-1} + \sum_{j=1}^{\ell} c_j \Delta y_{t-j} + \varepsilon_t$$

where y denotes the logarithm of the series. The null hypothesis of a single unit root is rejected if α_1 is negative and significantly different from zero. A trend variable should not be included, since the presence of a trend in financial asset prices is a clear violation of market efficiency, whether or not the asset price has a unit root. The optimal lag length, ℓ, can be chosen using data-dependent methods, that have desirable statistical properties when applied to unit root tests. Based on such ADF unit root tests, Nelson and Plosser (1982) argue that most macroeconomic and financial time series have a unit root.

Pierre Perron (1989), however, argues that most time series [and in particular those used by Nelson and Plosser (1982)] are trend stationary if one allows for a one-time change in the intercept or in the slope (or both) of the trend function. The postulate is that certain 'big shocks' do not represent a realization of the underlying data generation mechanism of the series under consideration and that the null should be tested against the trend-stationary alternative by allowing, under both the null and the alternative hypotheses, for the presence of a one-time break (at a known point in time) in the intercept or in the slope (or both) of the trend function[41]. Hence, whether the unit root model is rejected or not depends on how big shocks are treated. If they are treated like any other shock, then ADF unit root testing procedures are appropriate and the unit root null hypothesis cannot (in general) be rejected. If, however, they are treated differently, then Perron-type procedures are appropriate and the null hypothesis of a unit root will most likely be rejected.

It is also important to note that in the tests that we discussed so far the unit root is the null hypothesis to be tested and that the way in which classical hypothesis testing is carried out ensures that the null hypothesis is accepted unless there is strong evidence against it. In fact, Denis Kwiatkowski, Peter Phillips, Peter Schmidt, and Yongcheol Shin (1992) argue that such unit root tests fail to reject a unit root because they have low power against relevant alternatives and they propose tests (known as KPSS tests) of the hypothesis of stationarity against the alternative of a unit root. They argue that such tests should complement unit root tests and that by testing both the unit root hypothesis and the stationarity hypothesis, one can distinguish series that appear to be stationary, series that appear

[41]Perron's (1989) assumption that the break point is uncorrelated with the data has been criticized, on the basis that problems associated with 'pre-testing' are applicable to his methodology and that the structural break should instead be treated as being correlated with the data. More recently, a number of studies treat the selection of the break point as the outcome of an estimation procedure and transform Perron's (1989) conditional (on structural change at a known point in time) unit root test into an unconditional unit root test.

to be integrated, and series that are not very informative about whether or not they are stationary or have a unit root.

Finally, given that integration tests are sensitive to the class of models considered (and may be misleading because of misspecification), *fractionally*-integrated representations, which nest the unit-root phenomenon in a more general model, have also been used - see Richard Baillie (1996) for a survey. Fractional integration is a popular way to parameterize long-memory processes. If such processes are estimated with the usual autoregressive-moving average model, without considering fractional orders of integration, the estimated autoregressive process can exhibit spuriously high persistence close to a unit root. Since financial asset prices might depart from their means with long memory, one could condition the unit root tests on the alternative of a fractional integrated process, rather than the usual alternative of the series being stationary. In this case, if we fail to reject an autoregressive unit root, we know it is not a spurious finding due to neglect of the relevant alternative of fractional integration and long memory.

Despite the fact that the random walk and martingale hypotheses are contained in the null hypothesis of a unit root, unit root tests are not predictability tests. They are designed to reveal whether a series is difference-stationary or trend stationary and as such they are tests of the permanent/temporary nature of shocks. More recently a series of papers including those by James Poterba and Summers (1988), and Andrew Lo and Craig MacKinlay (1988) have argued that the efficient markets theory can be tested by comparing the relative variability of returns over different horizons using the variance ratio methodology of John Cochrane (1988). They have shown that asset prices are mean reverting over long investment horizons - that is, a given price change tends to be reversed over the next several years by a predictable change in the opposite direction. Similar results have been obtained by Fama and Kenneth French (1988), using an alternative but closely related test based on predictability of multiperiod returns. Of course, mean-reverting behavior in asset prices is consistent with transitory deviations from equilibrium which are both large and persistent, and implies positive autocorrelation in returns over short horizons and negative autocorrelation over longer horizons.

Predictability of financial asset returns is a broad and very active research topic and a complete survey of the vast literature is beyond the scope of the present paper. We shall notice, however, that a general consensus has emerged that asset returns are predictable. As John Campbell, Lo, and MacKinlay (1997, pp. 80) put it “[r]ecent econometric advances and empirical evidence seem to suggest that financial asset returns are predictable to some degree. Thirty years ago this would have been tantamount to an outright rejection of market efficiency. However, modern financial economics teaches us that other, perfectly rational, factors may account for such predictability. The fine structure of securities markets and frictions in the trading process can generate predictability. Time-varying expected returns due to changing business conditions can generate predictability. A certain degree of predictability may be necessary to reward investors for bearing certain dynamic risks".

7.3 Random Walk versus Chaos

Most of the empirical tests that we discussed in the previous subsection are designed to detect 'linear' structure in financial data - that is, linear predictability is the focus. However, as Campbell, Lo, and MacKinlay (1997, pp. 467) argue "many aspects of economic behavior may not be linear. Experimental evidence and casual introspection suggest that investors' attitudes towards risk and expected return are nonlinear. The terms of many financial contracts such as options and other derivative securities are nonlinear. And the strategic interactions among market participants, the process by which information is incorporated into security prices, and the dynamics of economy-wide fluctuations are all inherently nonlinear. Therefore, a natural frontier for financial econometrics is the modeling of nonlinear phenomena".

It is for this reason that interest in deterministic nonlinear chaotic processes has in the recent past experienced a tremendous rate of development. Besides its obvious intellectual appeal, chaos is interesting because of its ability to generate output that mimics the output of stochastic systems thereby offering an alternative explanation for the behavior of asset prices. In fact, the possible existence of chaos could be exploitable and even invaluable. If, for example, chaos can be shown to exist in asset prices, the implication would be that profitable, nonlinearity-based trading rules exist (at least in the short run and provided the actual generating mechanism is known). Prediction, however, over long periods is all but impossible, due to the sensitive dependence on initial conditions property of chaos.

Clearly then, an important area for potentially productive research is to test for chaos and (in the event that it exists) to identify the nonlinear deterministic system that generates it. We turn to such tests in the following section.

8 Tests of Nonlinearity and Chaos

Although the exciting recent advances in deterministic nonlinear dynamical systems theory have had immediate implications for the 'hard' sciences, the impact on economics and finance has been less dramatic for at least two reasons. First, unlike most hard scientists, economists are generally not specific about functional form when modeling economic phenomena as deterministic nonlinear dynamical systems. Thus they rarely have theoretical reasons for expecting to find one form of nonlinearity rather than another. Second, economists mostly use non-experimental data, rendering it almost impossible to recover the deterministic dynamical system governing economic phenomena, even if such a system exists and is low-dimensional.

Despite these caveats, the mathematics of deterministic nonlinear dynamical systems has motivated several univariate statistical tests for independence, nonlinearity, and chaos, to which we now turn.

8.1 The Correlation Dimension Test

The concept and measurement of fractal dimension are not only necessary to understand the finer geometrical nature of strange attractors, but they are also fundamental tools for providing quantitative analyses of such attractors. Unfortunately, however, fractal dimension [as defined by Equation (4)] cannot be computed easily in practice, and convergence of the limit may not be guaranteed. To remedy this, Peter Grassberger and Itamar Procaccia (1983) suggested the concept of *correlation dimension* (or *correlation exponent*) which is, at the moment, prevailing in applications. The basic idea is that of replacing the box-counting algorithm, necessary to compute $N(\varepsilon)$ in Equation (4), with the measurement of correlations between points of a long time series on the attractor. Hence, the correlation dimension (unlike the fractal dimension) is a probabilistic, not a metric, dimension.

To briefly discuss the correlation dimension test for chaos, let us start with the 1-dimensional series, $\{x_t\}_{t=1}^n$, which can be embedded into a series of m-dimensional vectors $X_t = (x_t, x_{t-1}, ..., x_{t-m+1})'$ giving the series $\{x_t\}_{t=m}^n$. The selected value of m is called the *embedding dimension* and each X_t is known as an *m-history* of the series $\{x_t\}_{t=1}^n$. This converts the series of scalars into a slightly shorter series of (m-dimensional) vectors with overlapping entries[42]. In particu-

[42]In creating the embedding one could also take a forward orientation, $X_t = (x_t, x_{t+1}, ..., x_{t+m-1})'$, without affecting the results.

lar, from the sample size n, $N = n - m + 1$ m-histories can be made[43]. Assuming that the true, but unknown, system which generated $\{x_t\}_{t=1}^{n}$ is ϑ-dimensional and provided that $m \geq 2\vartheta + 1$, then the N m-histories recreate the dynamics of the data generation process and can be used to analyze the dynamics of the system.

The correlation dimension test is based on the *correlation function* (or *correlation integral*), $C(N,m,\varepsilon)$, which for a given embedding dimension m is given by:

$$C(N,m,\varepsilon) = \frac{1}{N(N-1)} \sum_{m \leq t \neq s \leq N} H\left(\varepsilon - \|X_t - X_s\|\right)$$

where ε is a sufficiently small number, $H(z)$ is the Heavside function, which maps positive arguments into 1, and nonpositive arguments into 0, i.e.,

$$H(z) = \begin{cases} 1 & \text{if } z > 0 \\ 0 & \text{otherwise,} \end{cases}$$

and $\|.\|$ denotes the distance induced by the selected norm[44]. In other words, the correlation integral is the number of pairs (t,s) such that each corresponding component of X_t and X_s are near to each other, nearness being measured in terms of distance being less than ε. Intuitively, $C(N,m,\varepsilon)$ measures the probability that the distance between any two m-histories is less than ε. If $C(N,m,\varepsilon)$ is large (which means close to 1) for a very small ε, then the data is very well correlated.

To move from the correlation function to the correlation dimension, one proceeds by looking to see how $C(N,m,\varepsilon)$ changes as ε changes. One expects $C(N,m,\varepsilon)$ to increase with ε (since increasing ε increases the number of neighbouring points that get included in the correlation integral). In fact, Grassberger and Procaccia (1983) have shown that for small values of ε, $C(N,m,\varepsilon)$ grows exponentially at the rate of D_c

$$C(N,m,\varepsilon) = \eta e^{D_c}$$

[43]For example, the series $\{x_1, \ldots, x_6\}$ would give the following four overlapping 3-histories: $X_3 = (x_1,x_2,x_3)'$, $X_4 = (x_2,x_3,x_4)'$, $X_5 = (x_3,x_4,x_5)'$, and $X_6 = (x_4,x_5,x_6)'$.

[44]Brock (1986, Theorem 2.4) shows that the correlation integral is independent of the choice of norm. The type most often used is the maximum norm (which is also more convenient for computer applications): $\|X_t - X_s\| = \max_{k\in[0,m-1]}\{|x_{t+k} - x_{s+k}|\}$, where$\|.\|$ is Euclidean distance. Using this norm the correlation integral may be written as $C(N,m,\varepsilon) = \frac{1}{N(N-1)} \sum_{m\leq t\neq s\leq N} \prod_{k=0}^{m-1} H\left(\varepsilon - |X_{t+k} - X_{s+k}|\right)$ since $H\left(\varepsilon - \|X_t - X_s\|\right) = \prod_{k=0}^{m-1} H\left(\varepsilon - |X_{t+k} - X_{s+k}|\right)$, i.e., if any $|x_{t+k} - x_{s+k}| \geq \varepsilon$ then $H(.) = 0$.

where η is some constant and D_c is the above mentioned correlation dimension.

If the increase in $C(N,m,\varepsilon)$ is slow as ε is increased, then most data points have to be near to each other, and the data is well correlated. If, however, the increase is fast, then the data are rather uncorrelated. Hence, the higher the correlation dimension [and the faster the increase in $C(N,m,\varepsilon)$ as ε is increased], the less correlated the data is and the system is regarded stochastic. On the other hand, the lower the correlation dimension [and the slower the increase in $C(N,m,\varepsilon)$ as ε is increased], the more correlated the data is and the system is regarded as essentially deterministic, even if fairly complicated.

The correlation dimension can be defined as

$$D_c = \lim_{\varepsilon \to 0} \frac{d \log C(N,m,\varepsilon)}{d \log \varepsilon}$$

that is, by the slope of the regression of $\log C(N,m,\varepsilon)$ versus $\log \varepsilon$ for small values of ε. As a practical matter one investigates the estimated value of D_c as m is increased. If as m increases D_c continues to rise, then the system is stochastic. If, however, the data are generated by a deterministic process (consistent with chaotic behavior), then D_c reaches a finite saturation limit beyond some relatively small m[45]. The correlation dimension can therefore be used to distinguish true stochastic processes from deterministic chaos (which may be low-dimensional or high-dimensional)[46].

While the correlation dimension measure is therefore potentially very useful in testing for chaos, the sampling properties of the correlation dimension are, however, unknown. As William Barnett, Ronald Gallant, Melvin Hinich, Jochen Jungeilges, Daniel Kaplan, and Mark Jensen (1995, pp. 306) put it "[i]f the only source of stochasticity is noise in the data, and if that noise is slight, then it is possible to filter the noise out of the data and use the correlation dimension test deterministically. However, if the economic structure that generated the data contains a stochastic disturbance within its equations, the correlation dimension is stochastic and its derived distribution is important in producing reliable inference".

[45]Since D_c can be used to characterize both chaos and stochastic dynamics (i.e., D_c is a finite number in the case of chaos and equal to infinity in the case of an independent and identically distributed stochastic process), one often finds in the literature expressions like 'deterministic chaos' (meaning simply chaos) and 'stochastic chaos' (meaning standard stochastic dynamics). This terminology, however, is confusing in contexts other than that of the correlation dimension analysis and we shall not use it in this paper.

[46]It is to be noted that Grassberger and Procaccia (1983) have shown that $D_c \leq D$, i.e., D_c is a lower bound for D - see also Medio (1992) for other measures of fractal dimension and their relation to D_c and D.

Moreover, if the correlation dimension is very large as in the case of high-dimensional chaos, it will be very difficult to estimate it without an enormous amount of data. In this regard, Ruelle (1990) argues that a chaotic series can only be distinguished if it has a correlation dimension well below $2\log_{10} N$, where N is the size of the data set, suggesting that with economic time series the correlation dimension can only distinguish low-dimensional chaos from high-dimensional stochastic processes - see also Grassberger and Procaccia (1983) for more details[47].

8.2 The BDS Test

To deal with the problems of using the correlation dimension test, Brock, Davis Dechert, Blake LeBaron, and José Scheinkman (1996) devised a new statistical test which is known as the BDS test. The BDS tests the null hypothesis of whiteness (independent and identically distributed observations) against an unspecified alternative using a nonparametric technique.

The BDS test is based on the Grassberger and Procaccia (1983) correlation integral as the test statistic. In particular, under the null hypothesis of whiteness, the BDS statistic is

$$W(N,m,\varepsilon) = \sqrt{N}\frac{C(N,m,\varepsilon) - C(N,1,\varepsilon)^m}{\widehat{\sigma}(N,m,\varepsilon)}$$

where $\widehat{\sigma}(N,m,\varepsilon)$ is an estimate of the asymptotic standard deviation of $C(N,m,\varepsilon) - C(N,1,\varepsilon)^m$ - the formula for $\widehat{\sigma}(N,m,\varepsilon)$ can be found in Brock et al. (1996). The BDS statistic is asymptotically standard normal under the whiteness null hypothesis - see Brock et al. (1996) for details.

The intuition behind the BDS statistic is as follows. $C(N,m,\varepsilon)$ is an estimate of the probability that the distance between any two m-histories, X_t and X_s of the series $\{x_t\}$ is less than ε. If $\{x_t\}$ were independent then for $t \neq s$ the probability of this joint event equals the product of the individual probabilities. Moreover, if $\{x_t\}$ were also identically distributed then all of the m probabilities under the product sign are the same. The BDS statistic therefore tests the null hypothesis that $C(N,m,\varepsilon) = C(N,1,\varepsilon)^m$ - the null hypothesis of whiteness[48].

Since the asymptotic distribution of the BDS test statistic is known under the null hypothesis of whiteness, the BDS test provides a direct (formal) statistical test

[47] Therefore, to detect an attractor with $D_c = 2$ we need at least 460 data points, with $D_c = 3$ at least $10,000$ data points, and with $D_c = 4$ at least $210,000$ data points.

[48] Note that whiteness implies that $C(N,m,\varepsilon) = C(N,1,\varepsilon)^m$ but the converse is not true.

for whiteness against general dependence, which includes both nonwhite linear and nonwhite nonlinear dependence. Hence, the BDS test does not provide a direct test for nonlinearity or for chaos, since the sampling distribution of the test statistic is not known (either in finite samples or asymptotically) under the null hypothesis of nonlinearity, linearity, or chaos. It is, however, possible to use the BDS test to produce indirect evidence about nonlinear dependence [whether chaotic (i.e., nonlinear deterministic) or stochastic], which is necessary but not sufficient for chaos - see Barnett et al. (1997) and Barnett and Melvin Hinich (1992) for a discussion of these issues.

8.3 The Hinich Bispectrum Test

Hinich (1982) argues that the bispectrum in the frequency domain is easier to interpret than the multiplicity of third order moments $\{C_{xxx}(r,s) : s \leq r, r = 0,1,2,...\}$ in the time domain. For frequencies ω_1 and ω_2 in the principal domain given by

$$\Omega = \{(\omega_1, \omega_2) : 0 < \omega_1 < 0.5, \omega_2 < \omega_1, 2\omega_1 + \omega_2 < 1\},$$

the bispectrum, $B_{xxx}(\omega_1, \omega_2)$, is defined by

$$B_{xxx}(\omega_1, \omega_2) = \sum_{r=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} C_{xxx}(r,s) \exp[-i2\pi(\omega_1 r + \omega_2 s)].$$

The bispectrum is the double Fourier transformation of the third order moments function and is the third order polyspectrum. The regular power spectrum is the second order polyspectrum and is a function of only one frequency.

The skewness function $\Gamma(\omega_1, \omega_2)$ is defined in terms of the bispectrum as follows

$$\Gamma^2(\omega_1, \omega_2) = \frac{|B_{xxx}(\omega_1, \omega_2)|^2}{S_{xx}(\omega_1) S_{xx}(\omega_2) S_{xx}(\omega_1 + \omega_2)}, \tag{21}$$

where $S_{xx}(\omega)$ is the (ordinary power) spectrum of $x(t)$ at frequency ω. Since the bispectrum is complex valued, the absolute value (vertical lines) in Equation (21) designates modulus. David Brillinger (1965) proves that the skewness function $\Gamma(\omega_1, \omega_2)$ is constant over all frequencies $(\omega_1, \omega_2) \in \Omega$ if $\{x(t)\}$ is linear; while $\Gamma(\omega_1, \omega_2)$ is flat at zero over all frequencies if $\{x(t)\}$ is Gaussian. Linearity and Gaussianity can be tested using a sample estimator of the skewness function. But observe that those flatness conditions are necessary but not sufficient for general linearity and Gaussianity, respectively. On the other hand, flatness of the skewness function is necessary and sufficient for third order nonlinear dependence. The Hinich (1982) 'linearity test' tests the null hypothesis that the skewness function is flat, and hence is a test of lack of third order nonlinear dependence. For details of the test, see Hinich (1982).

8.4 The NEGM Test

As it was argued earlier, the distinctive feature of chaotic systems is sensitive dependence on initial conditions - that is, exponential divergence of trajectories with similar initial conditions. The most important tool for diagnosing the presence of sensitive dependence on initial conditions (and thereby of chaoticity) is provided by the dominant Lyapunov exponent, λ. This exponent measures average exponential divergence or convergence between trajectories that differ only in having an 'infinitesimally small' difference in their initial conditions and remains well-defined for noisy systems. A bounded system with a positive Lyapunov exponent is one operational definition of chaotic behavior.

One early method for calculating the dominant Lyapunov exponent is that proposed by Alan Wolf, Jack Swift, Harry Swinney, and John Vastano (1985). This method, however, requires long data series and is sensitive to dynamic noise, so inflated estimates of the dominant Lyapunov exponent are obtained. Recently, Douglas Nychka, Stephen Ellner, Ronald Gallant, and Daniel McCaffrey (1992) have proposed a regression method, involving the use of neural network models, to test for positivity of the dominant Lyapunov exponent. The Nychka et al. (1992), hereafter NEGM, Lyapunov exponent estimator is a regression (or Jacobian) method, unlike the Wolf et al. (1985) direct method which [as Brock and Chera Sayers (1988) have found] requires long data series and is sensitive to dynamic noise.

Assume that the data $\{x_t\}$ are real-valued and are generated by a nonlinear autoregressive model of the form

$$x_t = f(x_{t-L}, x_{t-2L}, \ldots, x_{t-mL}) + e_t \tag{22}$$

for $1 \leq t \leq N$, where L is the time-delay parameter and m is the length of the autoregression. Here f is a smooth unknown function, and $\{e_t\}$ is a sequence of independent random variables with zero mean and unknown constant variance. The Nychka et al. (1992) approach to estimation of the maximum Lyapunov exponent involves producing a state-space representation of (22)

$$X_t = F(X_{t-L}) + E_t, \quad F : \mathbb{R}^m \to \mathbb{R}^m$$

where $X_t = (x_t, x_{t-L}, \ldots, x_{t-mL+L})'$, $F(X_{t-L}) = (f(x_{t-L}, \ldots, x_{t-mL}), x_{t-L}, \ldots, x_{t-mL+L})'$, and $E_t = (e_t, 0, \ldots, 0)'$, and using a Jacobian-based method to estimate λ through the intermediate step of estimating the individual Jacobian matrices

$$J_t = \frac{\partial F(X_t)}{\partial X'}.$$

After using several nonparametric methods, McCaffrey et al. (1992) recommend using either thin plate splines or neural nets to estimate J_t. Estimation based on neural nets involves the use of the a neural net with q units in the hidden layer

$$f(X_{t-L}, \theta) = \beta_0 + \sum_{j=1}^{q} \beta_j \psi(\gamma_{0j} + \sum_{i=1}^{m} \gamma_{ij} x_{t-iL})$$

where ψ is a known (hidden) nonlinear 'activation function' [usually the logistic distribution function $\psi(u) = 1/(1+\exp(-u))$]. The parameter vector θ is then fit to the data by nonlinear least squares. That is, one computes the estimate $\widehat{\theta}$ to minimize the sum of squares $S(\theta) = \sum_{t=1}^{N} [x_t - f(X_{t-1}, \theta)]^2$, and uses $\widehat{F}(X_t) = (f(x_{t-L}, \ldots, x_{t-mL}, \widehat{\theta}), x_{t-L}, \ldots, x_{t-mL+L})'$ to approximate $F(X_t)$.

As appropriate values of $L, m,$ and q, are unknown, Nychka et al. (1992) recommend selecting that value of the triple (L, m, q) that minimizes the Bayesian Information Criterion (BIC) - see Gideon Schwartz (1978). As shown by Gallant and Halbert White (1992), we can use $\widehat{J_t} = \partial \widehat{F}(X_t)/\partial X'$ as a nonparametric estimator of J_t when (L, m, q) are selected to minimize BIC. The estimate of the dominant Lyapunov exponent then is

$$\widehat{\lambda} = \frac{1}{2N} \log |\widehat{v}_1(N)|$$

where $\widehat{v}_1(N)$ is the largest eigenvalue of the matrix $\widehat{T}_N' \widehat{T}_N$ and where $\widehat{T}_N = \widehat{J}_N \widehat{J}_{N-1}, \ldots, \widehat{J}_1$.

8.5 The White Test

In White's (1989) test, the time series is fitted by a single hidden-layer feed-forward neural network, which is used to determine whether any nonlinear structure remains in the residuals of an autoregressive (AR) process fitted to the same time series. The null hypothesis for the test is 'linearity in the mean' relative to an information set. A process that is linear in the mean has a conditional mean function that is a linear function of the elements of the information set, which usually contains lagged observations on the process[49].

The rationale for White's test can be summarized as follows: under the null hypothesis of linearity in the mean, the residuals obtained by applying a linear filter to the process should not be correlated with any measurable function of the history of the process. White's test uses a fitted neural net to produce the measurable function of the process's history and an AR process as the linear filter. White's method then tests the hypothesis that the fitted function does not correlate with the residuals of the AR process. The resulting test statistic has an asymptotic χ^2 distribution under the null of linearity in the mean[50].

8.6 The Kaplan Test

We begin our discussion of Daniel Kaplan's (1994) test by reviewing its origins in the chaos literature, although the test is currently being used as a test of linear stochastic process against general nonlinearity, whether or not noisy or chaotic. In the case of chaos, a time series plot of the output of a chaotic system may be very difficult to distinguish visually from a stochastic process. However, plots of the solution paths in phase space (x_{t+1} plotted against x_t and lagged values of x_t) often reveal deterministic structure that was not evident in a plot of x_t versus t - see, for example, Figure 9. A test based upon continuity in phase space has been proposed by Kaplan (1994).

Briefly, he used the fact that deterministic solution paths, unlike stochastic processes, have the following property: points that are nearby are also nearby under their image in phase space. Using this fact, he has produced a test statistic, which has a strictly positive lower bound for a stochastic process, but not for a deterministic solution path. By computing the test statistic from an adequately large number of linear processes that plausibly might have produced the data, the approach can be used to test for linearity against the alternative of noisy nonlinear dynamics. The procedure involves producing linear stochastic process surrogates

[49]For a formal definition of linearity in the mean, see Tae-Hwy Lee, White, and Granger (1993, section 1). Note that a process that is not linear in the mean is said to exhibit 'neglected nonlinearity'. Also, a process that is linear is also linear in the mean, but the converse need not be true.

[50]See Lee, White, and Granger (1993, section 2) for a presentation of the test statistic's formula and computation method.

for the data and determining whether the surrogates or a noisy continuous nonlinear dynamical solution path better describe the data. Linearity is rejected, if the value of the test statistic from the surrogates is never small enough relative to the value of the statistic computed from the data[51].

9 Evidence on Nonlinearity and Chaos

There have been a great deal of studies over the past few years testing for nonlinearity or chaos on economic and financial data. Thus we devote a good deal of space to this empirical work. In this section we present a discussion of the empirical evidence on economic and financial data, look at the controversies that have arisen about the available results, address one important question regarding the power of some of the best known tests for nonlinearity or chaos against various alternatives, and raise the issue of whether dynamical systems theory is practical in economics.

9.1 Evidence on Economic Data

In Table 1 we list 7 studies that have used various economic time series to test for nonlinearity or chaos. In Columns 2 to 5 we present the data set; the number of observations; the testing procedure used; and the results obtained. Clearly, there is a broad consensus of support for the proposition that the data generating processes are characterized by a pattern of nonlinear dependence, but there is no consensus at all on whether there is chaos in economic time series. For example, Brock and Sayers (1988), Murray Frank and Thanasis Stengos (1988), and Frank, Ramazan Gencay, and Stengos (1988) find no evidence of chaos in U.S., Canadian, and international, respectively, macroeconomic time series.

On the other hand, Barnett and Ping Chen (1988), claimed successful detection of chaos in the (demand-side) U.S. Divisia monetary aggregates. Their conclusion was further confirmed by Gregory DeCoster and Douglas Mitchell (1991, 1994). This published claim of successful detection of chaos has generated considerable controversy, as in James Ramsey, Sayers, and Philip Rothman (1990) and Ramsey and Rothman (1994), who by re-examining the data utilized in Barnett and Chen (1988) show that there is no evidence for the presence of chaos. In fact, they raised similar questions regarding virtually all of the other published tests of chaos.

[51] See Kaplan (1994) or Barnett et al. (1997) for more details about Kaplan's procedure.

Further results relevant to this controversy have recently been provided by Apostolos Serletis (1995). Building on Barnett and Chen (1988), Serletis (1995) contrasts the random walk behavior of the velocity of money to chaotic dynamics, motivated by the notion that velocity follows a deterministic, dynamic, and nonlinear process which generates output that mimics the output of stochastic systems. In doing so, he tests for chaos using the Lyapunov exponent estimator of Nychka et al. (1992) and reports evidence of chaos in the Divisia L velocity series.

Although from a theoretical point of view, it would be extremely interesting to obtain empirical verification that macroeconomic series have actually been generated by deterministic chaotic systems, it is fair to say that those series are not the most suitable ones for the calculation of chaos indicators. This is for at least two reasons. First of all, the series are rather short with regard to the calculations to be performed, since they are usually recorded at best only monthly; secondly, they have probably been contaminated by a substantial dose of noise (this is particularly true for aggregate time series like GNP). We should not be surprised, therefore, that exercises of this kind have not yet led to particularly encouraging results.

9.2 Evidence on Financial Data

As can be seen from Table 2 (where we summarize the evidence in the same fashion as in Table 1), there is already a substantial literature testing for nonlinear dynamics on financial data, using various inference methods - for other unpublished work on testing nonlinearity and chaos on financial data, see Abhay Abhyankar, Laurence Copeland, and Woon Wong (1997, table 1). In fact, the analysis of financial time series has led to results which are as a whole more interesting and more reliable than those of macroeconomic series. This is probably due to the much larger number of data available and their superior quality (measurement in most cases is more precise, at least when we do not have to make recourse to broad aggregation).

Scheinkman and LeBaron (1989) studied United States weekly returns on the Center for Research in Security Prices (CRSP) value-weighted index, employing the BDS statistic, and found rather strong evidence of nonlinearity and some evidence of chaos[52]. Some very similar results have been obtained by Frank and Stengos (1989), investigating daily prices (from the mid 1970's to the mid 1980's)

[52]In order to verify the presence of a nonlinear structure in the data, they also suggested employing the so-called 'shuffling diagnostic'. This procedure involves studying the residuals obtained by adapting an autoregressive model to a series and then reshuffling these residuals. If the residuals

for gold and silver, using the correlation dimension and the Kolmogorov entropy. Their estimate of the correlation dimension was between 6 and 7 for the original series and much greater and non-converging for the reshuffled data.

More recently, Serletis and Periklis Gogas (1997) test for chaos in seven East European black market exchange rates, using the Kees Koedijk and Clements Kool (1992) monthly data (from January 1955 through May 1990). In doing so, they use three inference methods, the BDS test, the NEGM test, as well as the Lyapunov exponent estimator of Gencay and Dechert (1992). They find some consistency in inference across methods, and conclude, based on the NEGM test, that there is evidence consistent with a chaotic nonlinear generation process in two out of the seven series - the Russian ruble and East German mark. Altogether, these and similar results seem to suggest that financial series provide a more promising field of research for the methods in question.

A notable feature of the literature just summarized is that most researchers, in order to find sufficient observations to implement the tests, use data periods measured in years. The longer the data period, however, the less plausible is the assumption that the underlying data generation process has remained stationary, thereby making the results difficult to interpret. In fact, different conclusions have been reached by researchers using high-frequency data over short periods. For example, Abhyankar, Copeland, and Wong (1995) examine the behavior of the U.K. Financial Times Stock Exchange 100 (FTSE 100) index, over the first six months of 1993 (using 1-, 5-, 15-, 30-, and 60-minute returns). Using the Hinich (1982) bispectral linearity test, the BDS test, and the NEGM test, they find evidence of nonlinearity, but no evidence of chaos.

More recently, Abhyankar, Copeland, and Wong (1997) test for nonlinear dependence and chaos in real-time returns on the world's four most important stock-market indices - the FTSE 100, the Standard & Poor 500 (S&P 500) index, the Deutscher Aktienindex (DAX), and the Nikkei 225 Stock Average. Using the BDS and the NEGM tests, and 15-second, 1-minute, and 5-minute returns (from September 1 to November 30, 1991), they reject the hypothesis of independence in favor of a nonlinear structure for all data series, but find no evidence of low-dimensional chaotic processes.

---

are totally random (i.e., if the series under scrutiny is not characterized by chaos), the dimension of the residuals and that of the shuffled residuals should be approximately equal. On the contrary, if the residuals are chaotic and have some structure, then the reshuffling must reduce or eliminate the structure and consequently increase the correlation dimension. The correlation dimension of their reshuffled residuals always appeared to be much greater than that of the original residuals, which was interpreted as being consistent with chaos.

Of course, there is other work, using high-frequency data over short periods, that finds order in the apparent chaos of financial markets. For example, the article by Shoaleh Ghashghaie, Wolfgang Breymann, Joachim Peinke, Peter Talkner, and Yadolah Dodge (1996) analyzes all worldwide 1,472,241 bid-ask quotes on U.S. dollar-German mark exchange rates between October 1, 1992 and September 30, 1993. It applies physical principles and provides a mathematical explanation of how one trading pattern led into and then influenced another. As the authors conclude, "...we have reason to believe that the qualitative picture of turbulence that has developed during the past 70 years will help our understanding of the apparently remote field of financial markets".

9.3 Controversies

As discussed in the previous two subsections, there is little agreement about the existence of chaos or even of nonlinearity in economic and financial data, and some economists continue to insist that linearity remains a good assumption for such data, despite the fact that theory provides very little support for that assumption. It should be noted, however, that the available tests search for evidence of nonlinearity or chaos in data without restricting the boundary of the system that could have produced that nonlinearity or chaos. Hence these tests should reject linearity, even if the structure of the economy is linear, but the economy is subject to shocks from a surrounding nonlinear or chaotic physical environment, as through nonlinear climatological or weather dynamics. Under such circumstances, linearity would seem an unlikely inference[53].

Since the available tests are not structural and hence have no ability to identify the source of detected chaos, the alternative hypothesis of the available tests is that no natural deterministic explanation exists for the observed economic fluctuations anywhere in the universe. In other words, the alternative hypothesis is that economic fluctuations are produced by supernatural shocks or by inherent randomness in the sense of quantum physics. Considering the implausibility of the alternative hypothesis, one would think that findings of chaos in such nonparametric tests would produce little controversy, while any claims to the contrary would be subjected to careful examination. Yet, in fact the opposite seems to be the case.

We argued earlier that the controversies might stem from the high noise level

[53]In other words, not only is there no reason in economic theory to expect linearity within the structure of the economy, but there is even less reason to expect to find linearity in nature, which produces shocks to the system.

that exists in most aggregated economic time series and the relatively low sample sizes that are available with economic data. However, it also appears that the controversies are produced by the nature of the tests themselves, rather than by the nature of the hypothesis, since linearity is a very strong null hypothesis, and hence should be easy to reject with any test and any economic or financial time series on which an adequate sample size is available. In particular, there may be very little robustness of such tests across variations in sample size, test method, and data aggregation method. That possibility was the subject of Barnett et al. (1995), who used five of the most widely used tests for nonlinearity or chaos with various monetary aggregate data series of various sample sizes and acquired results that differed substantially across tests and over sample sizes, as well as over the statistical index number formulas used to aggregate over the same component data.

9.4 Single Blind Controlled Competition

It is possible that none of the tests for chaos and nonlinear dynamics that we have discussed completely dominates the others, since some tests may have higher power against certain alternatives than other tests, without any of the tests necessarily having higher power against all alternatives. If this is the case, each of the tests may have its own comparative advantages, and there may even be a gain from using more than one of the tests in a sequence designed to narrow down the alternatives.

To explore this possibility, Barnett with the assistance of Jensen designed and ran a single blind controlled experiment, in which they produced simulated data from various processes having linear, nonlinear chaotic, or nonlinear nonchaotic signal. They transmitted each simulated data set by e-mail to experts in running each of the statistical tests that were entered into the competition. The e-mailed data included no identification of the generating process, so those individuals who ran the tests had no way of knowing the nature of the data generating process, other than the sample size, and there were two sample sizes: a 'small sample' size of 380 and a 'large sample' size of 2000 observations.

In fact five generating models were used to produce samples of the small and large size. The models were a fully deterministic, chaotic Feigenbaum recursion (Model I), a generalized autoregressive conditional heteroskedasticity (GARCH) process (Model II), a nonlinear moving average process (Model III), an autoregressive conditional heteroskedasticity (ARCH) process (Model IV), and an autoregressive moving average (ARMA) process (Model V). Details of the parameter settings and noise generation method can be found in Barnett et al. (1996). The tests entered into this competition were Hinich's bispectrum test, the BDS test, White's test, Kaplan's test, and the NEGM test of chaos.

The results of the competition are available in Barnett et al. (1997) and are summarized in Table 3. They provide the most systematic available comparison of tests of nonlinearity and indeed do suggest differing powers of each test against certain alternative hypotheses. In comparing the results of the tests, however, one factor seemed to be especially important: subtle differences existed in the definition of the null hypothesis, with some of the tests being tests of the null of linearity, defined in three different manners in the derivation of the test's properties, and one test being a test of the null of chaos. Hence there were four null hypotheses that had to be considered to be able to compare each test's power relative to each test's own definition of the null.

Since the tests do not all have the same null hypothesis, differences among them are not due solely to differences in power against alternatives. Hence one could consider using some of them sequentially in an attempt to narrow down the inference on the nature of the process. For example, the Hinich test and the White test could be used initially to find out whether the process lacks third order nonlinear dependence and is linear in the mean. If either test rejects its null, one could try to narrow down the nature of the nonlinearity further by running the NEGM test to see if there is evidence of chaos. Alternatively, if the Hinich and White tests both lead to acceptance of the null, one could run the BDS or Kaplan test to see if the process appears to be fully linear. If the data leads to rejection of full linearity but acceptance of linearity in the mean, then the data may exhibit stochastic volatility of the ARCH or GARCH type.

In short, the available tests provide useful information, and such comparisons of other tests could help further to narrow down alternatives. But ultimately we are left with the problem of isolating the nature of detected nonlinearity or chaos to be within the structure of the economy. This final challenge remains unsolved, especially in the case of chaos.

9.5 Testability of Chaos within the Economy

Recently there has been considerable criticism of the existing research on chaos, as for example in Granger's (1994) review of Benhabib's (1992) book. However, it is unwise to take a strong opinion (either pro or con) in that area of research. Contrary to popular opinion within the profession, there have been no published tests of chaos 'within the structure of the economic system', and there is very little chance that any such tests will be available in this field for a very long time. Such tests are simply beyond the state of the art.

All of the published tests of chaos in economic data test for evidence of chaos in the data. If chaos is found, the test has no way of determining whether or not the source of the chaos is from within the structure of the economy or perhaps is from within the chaotic weather systems that surround the planet. Considering the fact that chaos is clearly evident in many natural phenomena, and considering the fact that natural phenomena introduce shocks into the economy, the observation of chaotic behavior in some economic variables should be no surprise, but should give us no reason to believe that the economic system is chaotic, or is not chaotic.

To determine whether the source of chaos in economic data is from within the economic system, a model of the economy must be constructed. The null hypothesis that then must be tested is the hypothesis that the parameters of the model are within the subset of the parameter space that supports the chaotic bifurcation regime of the dynamic system. Currently, however, we do not have the mathematical tools to find and characterize that subset, when more than three parameters exist. Hence, with any usable model of any economy, the set that defines the null hypothesis cannot be located - and no one can test a null hypothesis that cannot be located and defined.

Since we cannot test the hypothesis, we may instead wish to consider whether or not chaos is plausible on philosophical ground. On that basis, the question would be whether the economy should be viewed as having evolved naturally, as in the natural sciences, or was the product of intentional human design by economic 'engineers'. Systems intentionally designed (by engineers) to be stable are stable and not chaotic, if designed optimally. Nature, however, was not designed by human beings, and is chaotic - the weather, for example, will never converge to a steady state. Which view is more appropriate to understanding the dynamics of actual economies is not clear.

10 Conclusion

We have reviewed a great deal of high quality research on nonlinear and complex dynamics and evidence concerning chaotic nonlinear dynamics in economic and financial time series. There are many reasons for this interest. Chaos, for example, represents a radical change of perspective on business cycles. Business cycles receive an endogenous explanation and are traced back to the strong nonlinear deterministic structure that can pervade the economic system. This is different from the (currently dominant) exogenous approach to economic fluctuations, based on the assumption that economic equilibria are determinate and intrinsically stable, so that in the absence of continuing exogenous shocks the economy tends towards a steady state, but because of stochastic shocks a stationary pattern of fluctuations is observed.

Chaos could also help unify different approaches to structural macroeconomics. As Grandmont (1985) has shown, for different parameter values even the most classical of economic models can produce stable solutions (characterizing classical economics) or more complex solutions, such as cycles or even chaos (characterizing much of Keynesian economics). Finally, if forecasting is a goal, the possible existence of chaos could be exploitable and even invaluable. If, for example, chaos can be shown to exist in asset prices, the implication would be that profitable, nonlinearity-based trading rules exist (at least in the short run and provided the actual generating mechanism is known). Prediction, however, over long periods is all but impossible, due to the 'sensitive dependence on initial conditions' property of chaos.

However, as we argued in the previous section, we do not have the slightest idea of whether or not the economy exhibits chaotic nonlinear dynamics (and hence we are not justified in excluding the possibility). Until the difficult problems of testing for chaos 'within the structure of the economic system' are solved, the best that we can do is to test for chaos in economic data, without being able to isolate its source. But even that objective has proven to be difficult. While there have been many published tests for chaotic nonlinear dynamics, little agreement exists among economists about the correct conclusions.

11 References

ABARBANEL, HENRY D. I., BROWN, REGGIE, SIDOROWICH, JOHN J., AND TSIMRING, LEV SH. "The Analysis of Observed Chaotic Data in Physical Systems", *Reviews of Modern Physics*, Oct. 1993, 65 (4), pp. 1331-1392.

ABHYANKAR, ABHAY H. COPELAND, LAURENCE S., AND WONG WOON. "Nonlinear Dynamics in Real-Time Equity Market Indices: Evidence From the UK", *Econ. J.*, July 1995, 105 (431), pp. 864-880.

ABHYANKAR, ABHAY H. COPELAND, LAURENCE S., AND WONG WOON. "Uncovering Nonlinear Structure in Real-Time Stock-Market Indexes: The S&P

500, the DAX, the Nikkei 225, and the FTSE-100", *J. Bus. Econ. Statist.*, Jan. 1997, 15 (1), pp. 1-14.

AZARIADIS, COSTAS. *Intertemporal Macroeconomics*. Blackwell Publishers, 1993.

BAILLIE, RICHARD T. "Long Memory Processes and Fractional Integration in Econometrics", *J. Econometrics*, July 1996, 73 (1), pp. 5-59.

BARNETT, WILLIAM A. AND CHEN, PING. "The Aggregation-Theoretic Monetary Aggregates are Chaotic and Have Strange Attractors: An Econometric Application of Mathematical Chaos", in *Dynamic Econometric Modeling*. Eds.: WILLIAM A. BARNETT, ERNST R. BERNDT, AND HALBERT WHITE. Cambridge: Cambridge U. Press, 1988, pp. 199-246.

BARNETT, WILLIAM A., GALLANT, A. RONALD, HINICH, MELVIN J., JUNGEILGES, JOCHEN, KAPLAN, DANIEL, AND JENSEN, MARK J. "A Single-Blind Controlled Competition Among Tests for Nonlinearity and Chaos", *J. Econometrics*, 1997 (forthcoming).

BARNETT, WILLIAM A., GALLANT, A. RONALD, HINICH, MELVIN J., JUNGEILGES, JOCHEN, KAPLAN, DANIEL, AND JENSEN, MARK J. "Robustness of Nonlinearity and Chaos Tests to Measurement Error, Inference Method, and Sample Size", *J. of Econ. Behav. Organ.*, July 1995, 27 (2), pp. 301-320.

BARNETT, WILLIAM A., GALLANT, A. RONALD, HINICH, MELVIN J., JUNGEILGES, JOCHEN, KAPLAN, DANIEL, AND JENSEN, MARK J. "An Experimental Design to Compare Tests of Nonlinearity and Chaos", in *Nonlinear Dynamics in Economics*. Eds.: WILLIAM A. BARNETT, ALAN KIRMAN, AND MARK SALMON. Cambridge: Cambridge U. Press, 1996.

BARNETT, WILLIAM A. AND HINICH, MELVIN J. "Empirical Chaotic Dynamics in Economics", *Annals of Operations Research*, 1992, 37 (1-4), pp. 1-15.

BAUMOL, WILLIAM AND BENHABIB, JESS. "Chaos: Significance, Mechanism, and Economic Applications", *J. Econ. Perspectives*, Winter 1989, 3 (1), pp. 77-106.

BENHABIB, JESS. *Cycles and Chaos in Economic Equilibrium*. Princeton, NJ: Princeton U. Press., 1992.

BENHABIB, JESS AND DAY, RICHARD, H. "Rational Choice and Erratic Behavior", *Review of Economic Studies*, July 1981, 48 (3), 459-471.

BILLINGSLEY, PATRICK., 1965, *Ergodic Theory and Information*, Wiley

BLANCHARD, OLIVIER AND FISCHER, STANLEY. *Lectures on Macroeconomics*. MIT Press, 1989.

BOLDRIN, MICHELE. "Path of Optimal Accumulation in Two-sector Models", in *Economic Complexity: Chaos, Sunspots, Bubbles, and Nonlinearity*. Eds.: WILLIAM A. BARNETT, JOHN GEWEKE, AND KARL SHELL. Cambridge: Cambrigde U. Press, 1989, pp.

BOLDRIN, MICHELE, AND WOODFORD, MICHAEL. "Equilibrium Models Displaying Endogenous Fluctuations and Chaos: a Survey", *J. Mon. Econ*., Mar. 1990, 25 (2), pp. 189-223

BRILLINGER, DAVID R. "An Introduction to the Polyspectrum", *Annals of Mathematical Statistics*, 1965, 36, pp. 1351-1374.

BROCK, WILLIAM A. "Nonlinearity and Complex Dynamics in Economics and Finance", in *The Economy as an Evolving Complex System.* Eds.: PHILIP W. ANDERSON, KENNETH J. ARROW, AND DAVID PINES. Addison-Wesley Publishing Company, 1988, pp. 77-97.

BROCK, WIILIAM A., DECHERT, W. DAVIS, LEBARON, BLAKE, AND SCHEINKMAN, JOSÉ A. "A Test for Independence Based on the Correlation Dimension", *Econ. Reviews*, Aug. 1996, 15 (3), pp. 197-235.

BROCK, WILLIAM A. AND SAYERS, CHERA. "Is the Business Cycle Characterized by Deterministic Chaos?", *J. Monet. Econ*. July 1988, 22 (1), pp. 71-90.

BULLARD, JAMES AND BUTLER, ALISON. "Nonlinearity and Chaos in Economic Models: Implications for Policy Decisions", *Econ. J*., July 1993, 103 (419), pp. 849-867.

CAMPBELL, JOHN Y., LO, ANDREW W., AND MACKINLAY, A. CRAIG. *The Econometrics of Financial Markets.* Princeton, NJ: Princeton U. Press, 1997.

COCHRANE, JOHN H. "How Big is the Random Walk in GNP?", *J. Polit. Econ*., Oct. 1988, 96 (5), pp. 893-920.

DAY, RICHARD H. "Irregular Growth Cycles", *Amer. Econ. Rev.*, 1982, 72 (3), pp. 406-414.

DECOSTER, GREGORY P. AND MITCHELL, DOUGLAS W. "Nonlinear Monetary Dynamics", *J. Bus. Econ. Statist*., 1991, 9, pp. 455-462.

DECOSTER, GREGORY P. AND MITCHELL, DOUGLAS W. “Reply”, *J. Bus. Econ. Statist*., 1994, 12, pp. 136-137.

DICKEY, DAVID A. AND FULLER, WAYNE A. “Likelihood Ratio Tests for Autoregressive Time Series With a Unit Root”, *Econometrica*, July 1981, 49 (4), pp. 1057-1072.

ECKMANN, J. AND RUELLE, DAVID. “Ergodic Theory of Chaos and Strange Attractors”, *Reviews of Modern Physics*, July 1985, 57 (3), pp. 617-656.

FAMA, EUGENE F. “Efficient Capital Markets: A Review of Theory and Empirical Work”, *J. Finance*, May 1970, 25 (2), pp. 383-417.

FAMA, EUGENE F. “The Behavior of Stock Market Prices”, *J. Business*, January 1965, 38 (1), pp. 34-105.

FAMA, EUGENE F. AND FRENCH, KENNETH R.“Permanent and Temporary Components of Stock Prices”, *J. Polit. Econ.*, Apr. 1988, 96 (2), pp. 246-273.

FEIGENBAUM, MITCHELL J. “Quantitative Universality for a Class of Nonlinear Transformations”, *J. Stat. Phys.*, 1978, 19, pp. 25-52.

FRANK, MURRAY, GENCAY, RAMAZAN, AND STENGOS, THANASIS.“International Chaos”, *Europ. Econ. Rev.*, 1988, 32, pp. 1569-1584.

FRANK, MURRAY AND STENGOS, THANASIS.“Some Evidence Concerning Macroeconomic Chaos”, *J. Monet. Econ.*, 1988, 22, pp. 423-438.

FRANK, MURRAY AND STENGOS, THANASIS.“Measuring the Strangeness of Gold and Silver Rates of Return”, *Rev. Econ. Stud.*, 1989, 56, pp. 553-567.

FRISCH, RAGNAR. “Propagation Problems and Impulse Problems in Dynamic Economics”, in *Economic Essays in Honour of Gustav Cassel*. Eds.: London: Allen and Unwin 1933, pp. 171-205.

GALLANT, RONALD AND WHITE, HALBERT. “On Learning the Derivatives of an Unknown Mapping With Multilayer Feedforward Networks”, *Neural Networks*, 1992, 5, pp. 129-138.

GENCAY, RAMAZAN AND DECHERT, W. DAVIS. “An Algorithm for the $n-$Lyapunov Exponents of an $n-$Dimensional Unknown Dynamical System”, *Physica D*, Oct. 1992, 59 (1-3), pp. 142-157.

GHASHGHAIE, SHOALEH, BREYMANN, WOLFGANG, PEINKE, JOACHIM, TALKNER, PETER, AND DODGE, YADOLAH. "Turbulent Cascades in Foreign Exchange Markets", *Nature*, June 1996, 381 (6585), pp. 767-770.

GLENDINNING, PAUL A. *Stability, Instability, and Chaos*. Cambridge: Cambridge U. Press, 1994.

GOODWIN, RICHARD M. *Essays in Economic Dynamics*. London: Macmilllan Press Ltd., 1982.

GOODWIN, RICHARD M. "The Non-linear Accelerator and the Persistence of Business Cycles", *Econometrica*, Jan. 1951, 19 (1), pp. 1-17.

GRANDMONT, JEAN-MICHEL. "On Endogenous Competitive Business Cycles", *Econometrica*, Sept. 1985, 53 (5), pp. 995-1045.

GRANGER, CLIVE W. J. AND NEWBOLD, PAUL. *Forecasting Economic Time Series*. New York: Academic Press 1977.

GRANGER, CLIVE W. J. "Is Chaotic Economic Theory Relevant for Economics? A Review Article of: Jess Benhabib: Cycles and Chaos in Economic Equilibrium", *J. of Intern. and Comparative Econ.*, 1994, 3, pp. 139-145.

GRASSBERGER, PETER AND PROCACCIA, ITAMAR. "Characterization of Strange Attractors", *Physical Review Letters*, Jan. 1983, 50 (5), pp. 346-349.

GUCKENHEIMER, JOHN, AND HOLMES, PHILIP. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields.* New York: Springer-Verlag, 1983.

HICKS, JOHN R. *A Contribution to the Theory of the Trade Cycle*. Oxford: Oxford U. Press, 1950.

HILBORN, ROBERT C. *Chaos and Nonlinear Dynamics: An Introduction for Scientists and Engineers*. New York: Oxford U. Press, 1994.

HINICH, MELVIN J. "Testing for Gaussianity and Linearity of a Stationary Time Series", *J. of Time Series Analysis*, 1982, 3, pp. 169-176.

HSIEH, DAVID A. "Chaos and Nonlinear Dynamics: Applications to Financial Markets", *J. Finance*, Dec. 1991, 46 (5), pp. 1839-1877.

KALDOR, NICHOLAS. "A Model of the Trade Cycle", *Econ. J*, March 1940, 50 (197), pp. 78-92.

KAPLAN, DANIEL T. "Exceptional Events as Evidence for Determinism", *Phys-*

ica D, May 1994, 73 (1-2), pp. 38-48.

KWIATKOWSKI, DENIS, PHILLIPS, PETER C.B., SCHMIDT, PETER, AND SHIN, YONGCHEOL. "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root", *J. Econometrics*, Oct.-Dec. 1992, 54 (1-3), pp. 159-178.

KOEDIJK, KEES G., AND KOOL, CLEMENTS, J. M. "Tail Estimates of East European Exchange Rates", *J. Bus. Econ. Statist.*, 1992, 10, pp. 83-96.

LAUWERIER, HANS. *Fractals: Endlessly Repeated Geometrical Figures*. Princeton, NJ: Princeton U. Press, 1991.

LEE, TAE-HWY, WHITE, HALBERT, AND GRANGER, CLIVE W. J. "Testing for Neglected Nonlinearities in Time Series Models", *J. Econometrics*, Apr. 1993, 56 (3), pp. 269-290.

LEROY, STEPHEN F. "Efficient Capital Markets and Martingales", *J. Econ. Lit*. Dec. 1989, 27 (4), pp. 1583-1621.

LI, TIEN-YIEN AND YORKE, JAMES A. "Period Three Implies Chaos", *Amer. Math. Monthly*, Dec. 1975, 82 (10), pp. 985-992.

LO, ANDREW W. AND MACKINLAY, A. CRAIG. "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test", *Rev. Fin. Stud.* 1988, 1, pp. 41-66.

LORENZ, HANS-WALTER. *Nonlinear Dynamical Economics and Chaotic Motion*, 2nd edition. Berlin: Springer Verlag, 1993.

MANDELBROT, BENOIT B. "The Variation of Certain Speculative Stock Prices", *J. Business*, Oct. 1963, 36 (4), pp. 394-419.

MANDELBROT, BENOIT B. "Self-Affine Fractals and Fractal Dimension", *Physica Scripta*, 1985, 32, 257-260.

MATSUYAMA, KIMINORI. "Endogenous Price Fluctuations in an Optimizing Model of a Monetary Economy", *Econometrica*, Nov. 1991, 59 (6), pp. 1617-1631.

MAÑÉ, RICARDO. *Ergodic Theory and Differentiable Dynamics*. Berlin: Springer-Verlag, 1987.

MCCAFFREY, DANIEL, ELLNER, STEPHEN, GALLANT, RONALD, AND

NYCHKA, DOUGLAS. “Estimating the Lyapunov Exponent of a Chaotic System with Nonparametric Regression”, *J. of the American Statistical Association*, Sept. 1992, 87 (419), pp. 682-695.

MEDIO, ALFREDO *Chaotic Dynamics. Theory and Applications to Economics*. Cambridge: Cambridge U. Press, 1992.

MEDIO, ALFREDO, AND NEGRONI, GIORGIO. “Chaotic Dynamics in Overlapping Generations Models with Production”, in *Nonlinear Dynamics and Economics*. Eds.: WILLIAM A. BARNETT, ALAN KIRMAN, AND MARK SALMON. Cambridge: Cambridge U. Press, 1996. pp. 3-44.

MUTH, JOHN F. “Rational Expectations and the Theory of Price Movements”, *Econometrica*, Nov. 1961, 29 (6), pp. 315-335.

NELSON, CHARLES R. AND PLOSSER, CHARLES I. “Trends and Random Walks in Macroeconomic Time Series”, *J. Monet. Econ.*, 1982, 10, pp. 139-162.

NISHIMURA, KAZUO, AND SORGER, GERHARD. “Optimal Cycles and Chaos: A Survey”, *Studies in Nonlinear Dynamics and Econometrics,* April 1996, 1 (1), pp. 11-28.

NYCHKA, DOUGLAS, ELLNER, STEPHEN, GALLANT, RONALD, AND MCCAFFREY, DANIEL. “Finding Chaos in Noisy Systems”, *J. of the Royal Statistical Society B* 1992, 54 (2), pp. 399-426.

ORNSTEIN, DONALD S. AND WEISS, B. “Statistical Properties of Chaotic Systems”, *Bulletin* (New Series) *of the American Mathematical Society*, Jan. 1991, 24 (1), pp. 11-116.

OTT, EDWARD. “Strange Attractors and Chaotic Motions of Dynamical Systems”, *Reviews of Modern Physics*, Oct. 1981, 53 (4), pp. 655-671.

PERRON, PIERRE. “The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis”, *Econometrica*, Nov. 1989, 57 (6), pp. 1361-1401.

POTERBA, JAMES M. AND SUMMERS, LAURENCE H. “Mean Reversion in Stock Prices: Evidence and Implications”, *J. Fin. Econ.*, 1988, 22, pp. 27-59.

RAMSEY, JAMES B. AND ROTHMAN, PHILIP. Comment on “Nonlinear Monetary Dynamics” by DeCoster and Mitchell, *J. Bus. Econ. Statist.*, 1994, 12, pp. 135-136.

RAMSEY, JAMES B., SAYERS, CHERA L., AND ROTHMAN, PHILIP. “The

Statistical Properties of Dimension Calculations Using Small Data Sets: Some Economic Applications", *Int. Econ. Rev*., Nov. 1990, 31 (4), pp. 991-1020.

REICHLIN, PIETRO. "Equilibrium Cycles in an Overlapping Generations Economy with Production", *J. Econ. Theory*, Oct. 1986, 40 (1), pp. 89-102.

REICHLIN, PIETRO. "Endogenous Cycles in Competitive Models: An Overview", *Studies in Nonolinear Dynamics and Econometrics*, Jan. 1997, 1 (4), pp. 175-185.

ROBINSON, CLARK. *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos*. CRC Press Inc., 1995.

RUELLE, DAVID. *Chaotic Evolution and Strange Attractors*. Cambridge: Cambridge U. Press, 1989.

RUELLE, DAVID. "Deterministic Chaos: The Science and the Fiction", *Proc. R. Soc. London A*, Feb. 1990, 427 (1873), pp. 241-248.

SCHEINKMAN, JOSÉ A. "Nonlinearities in Economic Dynamics", *Econ. J.*, March 1990, 100, pp. 33-48.

SCHEINKMAN, JOSÉ A. AND LEBARON, BLAKE "Nonlinear Dynamics and Stock Returns", *J. Business*, July 1989, 62 (3), pp. 311-337.

SCHWARTZ, GIDEON. "Estimating the Dimension of a Model", *The Annals of Statistics*, 1978, 6, pp. 461-464.

SERLETIS, APOSTOLOS. "Random Walks, Breaking Trend Functions, and the Chaotic Structure of the Velocity of Money", *J. Bus. Econ. Statist*., Oct. 1995, 13 (4), pp. 453-458.

SERLETIS, APOSTOLOS, AND GOGAS, PERIKLIS. "Chaos in East European Black-Market Exchange Rates", *Ricerche Economiche*, 1997 (forthcoming).

SHARKOVSKI, A.N. "Coexistence of Cycles of a Continuous Map of a Line Into Itself", *Ukrain. Math. Zeitschrift*, 1964, 16, pp. 61-71.

SIDRAUSKI, MIGUEL. "Rational Choice and Patterns of Growth in a Monetary Economy", *Amer. Econ. Rev.*, June 1967, 57 (2), pp. 534-544.

SLUTSKY, EUGENE. "The Summation of Random Causes as the Source of Cyclical Processes", *Econometrica*, 1927, 5, pp. 105-146.

SOLOW, ROBERT M. "A Contribution to the Theory of Economic Growth", *Quart. J. Econ.*, Feb. 1956, 70 (1), pp. 65-94.

SUGIHARA, GEORGE, AND MAY, ROBERT M. “Nonlinear Forecasting as a Way of Distinguishing Chaos from Measurement Error in Time Series”, *Nature*, Apr. 1990, 344 (6268), pp. 734-741.

SUMMERS, LAURENCE H. “Does the Stock Market Rationally Reflect Fundamental Values”, *J. Finance*, July 1986, 41 (3), pp. 591-601.

WHITE, HALBERT. “Some Asymptotic Results for Learning in Single Hidden-Layer Feedforward Network Models”, *J. of the American Statistical Association*, Dec. 1989, 84 (408), pp. 1003-1013.

WOLF, ALAN, SWIFT, JACK B., SWINNEY, HARRY L., AND VASTANO, JOHN A. “Determining Lyapunov Exponents from a Time Series”, *Physica* 16*D*, 1985, 16, pp. 285-317.

WOODFORD, MICHAEL. “Imperfect Financial Intermediation and Complex Dynamics”, in *Economic Complexity: Chaos, Sunspots, Bubbles, and Nonlinearity*. Eds.: WILLIAM A. BARNETT, JOHN GEWEKE, AND KARL SHELL. Cambridge: Cambridge U. Press 1989, pp. 309-338.

WOODFORD, MICHAEL. “Stationary Sunspot Equilibria in a Finance Constrained Economy”, *J. Econ. Theory*, Oct. 1986, 40 (1), pp. 128-137.

Figure 1:

Figure 2:

Figure 3:

Figure 4:

Figure 5:

Figure 6:

Figure 7:

Figure 8:

Figure 9:

TABLE 1.

SUMMARY OF PUBLISHED RESULTS OF NONLINEARITY AND CHAOS TESTING ON ECONOMIC DATA

Study	Data	N	Tests	Results
Barnett & Chen (1988)	Various weekly simple-sum and Divisia monetary aggregates	807	Correlation dimension	Evidenc and cha aggrega
DeCoster & Mitchell (1990)	Various weekly simple-sum and Divisia monetary aggregates	469-842	Correlation dimension	Evidenc and cha aggrega
Ramsey, Sayers, & Rothman (1990)	Barnett & Chen's (1988) Divisia M2 series and Scheinkman & LeBaron's (1989) CRSP value weighted daily stock returns	807-5,200	Correlation dimension	No evid
Serletis (1995)	Various monthly simple-sum and Divisia velocity series	396	NEGM	Evidenc velocity
Brock & Sayers (1988)	Several U.S. macroeconomic time series, using (mostly) quarterly data from the late 40's to the mid 80's	134	Correlation dimension	Nonline evidenc chaos
Frank & Stengos (1988)	Several Canadian macroeconomic time series, using (mostly) quarterly data since 1947	147	Correlation dimension	Nonline evidenc
Frank, Gencay, & Stengos (1988)	Real GNP series for Germany, Italy, Japan, and the U.K., using quarterly data since 1960	87-108	Correlation dimension	Nonline evidenc

TABLE 2.

SUMMARY OF PUBLISHED RESULTS OF NONLINEARITY
AND CHAOS TESTING ON FINANCIAL DATA

Study	Data	N	Tests	R
Serletis & Gogas (1997)	Seven East European black-market exchange rates	438	a. BDS b. NEGM c. Gencay & Dechert	a. b c.
Abhyankar, Copeland, and Wong (1997)	Real-time returns on four stock-market indices	2,268 - 97,185	a. BDS b. NEGM	a. b.
Abhyankar, Copeland, and Wong (1997)	FTSE 100	60,000	a. Bispectral linearity test b. BDS c. NEGM	a. b. c.
Hsieh (1991)	Weekly S&P 500 and CRSP value weighted returns	1,297 - 2,017	BDS	N
Frank & Stengos (1989)	Gold and silver rates of return	2,900 - 3,100	a. Correlation dimension b. Kolmogorov entropy	a. b.
Hinich & Patterson (1989)	Dow Jones industrial average	750	Bispectral Gaussianity and linearity tests	N n
Scheinkman & LeBaron (1989)	Daily CRSP value weighted returns	5,200	BDS	E
Brockett, Hinich & Patterson (1988)	10 Common U.S. stocks and $-yen spot and forward exchange rates	400	Bispectral Gaussianity and linearity tests	N n

TABLE 3.

RESULTS OF A SINGLE-BLIND CONTROLLED COMPETITION AMONG TESTS FOR NONLINEARITY AND CHAOS

		Small Sample		Large Sample	
Test	Null hypothesis	Successes	Failures	Successes	Failures
Hinich	Lack of 3rd order nonlinear dependence	3	2	3 plus ambiguous in 1 case	1 plus ambiguous in 1 case
BDS	Linear process	2	Ambiguous in 3 cases	5	0
NEGM	Chaos	5	0	5	0
White	Linearity in mean	4	1	4	1
Kaplan	Linear process	5	0	5	0

Source: Barnett, Gallant, Hinich, Jungeilges, Kaplan, and Jensen (1997, tables 1-4, 6-7, and 9-10).