Macroeconomic Dynamics, **2**, 1998, 505–532. Printed in the United States of America.

MD SURVEY

NONLINEAR DYNAMICS AND CHAOS PART I: A GEOMETRICAL APPROACH

ALFREDO MEDIO
University "Ca' Foscari" of Venice

This paper is the first part of a two-part survey reviewing some basic concepts and methods of the modern theory of dynamical systems. The survey is introduced by a preliminary discussion of the relevance of nonlinear dynamics and chaos for economics. We then discuss the dynamic behavior of nonlinear systems of difference and differential equations such as those commonly employed in the analysis of economically motivated models. Part I of the survey focuses on the geometrical properties of orbits. In particular, we discuss the notion of attractor and the different types of attractors generated by discrete- and continuous-time dynamical systems, such as fixed and periodic points, limit cycles, quasiperiodic and chaotic attractors. The notions of (noninteger) fractal dimension and Lyapunov characteristic exponent also are explained, as well as the main routes to chaos.

Keywords: Nonlinear Dynamics, Attractor, Cycles, Chaos, Fractal Dimension, Lyapunov Exponents

1. INTRODUCTION

According to an unsophisticated, yet rather common, view, the output of deterministic dynamical systems, in principle, can be predicted exactly and, assuming that the model representing the real system is correct, errors in prediction will be of the same order as errors in observation and measurement of the variables. On the contrary, so the argument runs, random processes describe systems of irreducible complexity owing to the presence of an indefinitely large number of degrees of freedom, whose behavior can be predicted only 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 random components) of low dimensions (i.e., with a small number of state variables) and with simple nonlinearities (e.g., a single quadratic function) can have stochastic behavior. This means that, for chaotic systems, if the measurements that determine their states are only

The author gratefully acknowledges the financial support of the Italian Ministry of the University (MURST) and the National Council of Research (CNR). Address Correspondence to: Professor Alfredo Medio, Department of Economics, University "Ca' Foscari" of Venice, 873 S. Giobbe, Venice 30121, Italy; e-mail: medio@unive.it.

 1365-1005/98 $9.50

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 brought about a profound reconsideration of the issue of determinism and randomness; the analytical and numerical investigation of their dynamics led to the production of a vast literature and to a deeper understanding of difficult conceptual problems in a number of different scientific disciplines.

Because many important topics in economics typically are formalized by means of dynamical systems (mostly 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 labeled "business cycles," for which the field of mathematical research under discussion is eminently important. The literature on 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 few basic, mutually competing approaches have dominated this area of research in modern times.

The first approach, which we label "exogenous cycle theory" (EXCT), essentially explains fluctuations as the result of random perturbations. Historically, its origin may be traced back to the seminal works of Slutzky (1927), Yule (1927), and Frisch (1993) and was developed later, and given the status of orthodoxy, by the works of the Cowles Commission in the 1940's and 1950's. The fundamental idea of this approach is the distinction between impulse and propagation mechanisms. In its typical version, serially uncorrelated shocks affect the relevant variables through distributed lags (the propagation mechanism), leading to serially correlated fluctuations in the variables themselves.[1] As Slutzky showed, even simple linear nonoscillatory propagation mechanisms, when excited by random, structureless shocks, can produce output sequences that are qualitatively similar to certain observed macroeconomic cycles.

The ability of the EXCT approach to provide an explanation of business cycles was called into question largely on the grounds that explaining fluctuations by means of random shocks amounts to a confession of ignorance. An alternative approach, which we label "endogenous disequilibrium cycle theory" (ENDCT), then was developed by a school of economists that, 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 explained primarily by deterministic models, i.e., by models in which random variables play no essential role. Classical examples of such models can be found in the works of Kaldor (1940), Hicks (1950), and 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 lose 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 ENDCT approach 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 [Blanchard and Fischer 1989, p. 277)] declares that it has "largely disappeared."

The reasons for the crisis of the Keynesian style of theorizing and the associated ENDCT 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 exists a fundamental criticism, raised against the ENDCT approach mainly by supporters of the rational expectations hypothesis [see Muth (1961)], which are relevant to our discussion and can be summarized briefly as follows. In the ENDCT models, agents' expectations, either explicitly modeled, or implicitly derived from the overall structure of the model, are, under most circumstances, incompatible with agents' rational behavior. In particular, in those models, agents do not make an efficient use of their information, including the knowledge of the models themselves. Consequently, they make systematic errors and/or do not exploit fully the outstanding profit opportunities.

A third line of inquiry, which could be labeled "equilibrium cycle theory," belongs to the tradition of general competitive equilibrium (GCE) that dominated contemporary economic theory. In its Walrasian formulation, GCE was conceived to explain the interdependence of economic variables at a given moment, rather than their evolution through time. However, its more elegant and sophisticated modern versions, originating in the works of Arrow and Debreu, have a built-in dynamic characterization, as intertemporal equilibrium defines certain time paths of the relevant variables, which can be called "equilibrium dynamics." Given the constraints imposed on equilibrium dynamics by the assumptions usually adopted by the prevailing theory (concavity of utility and production functions, constant returns to scale, intertemporally independent tastes and technology, agents' perfect foresight in the absence of exogenous perturbations, etc.), equilibrium dynamics may not be the most promising breeding ground for instability and dynamic complexity. However, in the past 15 or 20 years, a growing amount of research has been focused on characterizations of "disequilibrium dynamics"—mainly in the form of optimal-growth or overlapping-generations models—capable of generating oscillatory or more complex behavior of economic variables.[2] There is now a general consensus that irregular dynamics is by no means incompatible with perfect-foresight competitive intertemporal equilibrium, although there is still disagreement on whether that type of behavior is likely in real economies.

Having thus briefly recalled some of the main themes of economists' long-lasting debate on business-cycle theory, we may now wonder what the impact on that debate of recent advances in dynamical system theory has been so far.

Some of the effects of the theoretical developments that concern us here are of a general nature and consist of a deeper and more sophisticated understanding of nonlinear dynamics. This, we believe, will have a positive and lasting influence on economic dynamics in general and the analysis of economic fluctuations in particular, quite independently of the specific views and tenets of the various schools of thought. Apart from these broad considerations, however, a specific result, i.e., the discovery (or, more accurately, the rediscovery) of stochastic behavior of purely deterministic systems, may have a more direct and nonneutral effect on the current disputes among proponents of different theories of economic fluctuations, and reopen issues that were thought to have been resolved definitively.

On the one hand, the perfect-foresight hypothesis of standard equilibrium theory clashes with the results showing that the outcome of the dynamics may well be chaotic, and therefore essentially unpredictable in the sense that was sketched before and is discussed more rigorously later. In the presence of chaos, the assumptions of costless information and infinitely powerful monitoring and calculating ability of economic agents, implicit in the perfect-foresight hypothesis, are very hard to swallow.

On the other hand, and for exactly the same reason, the criticism of "irrationality" leveled against disequilibrium dynamical models such as those suggested by the ENDCT needs to be reconsidered. The criticism is perfectly valid if the outcomes of the dynamical system under consideration can be predicted accurately once the true model is known, e.g., if the outcome is periodic. However, if the theory implies chaotic (i.e., unpredictable) dynamics of the system, the perfect-foresight hypothesis must be dropped and replaced by more convincing expectation-generating mechanisms. In this context, adaptive mechanisms of the kind assumed by the ENDCT models, or other nonoptimizing rules of behavior, might turn out to be the most rational strategy available.

Chaos theory has not only affected the theoretical, deductive aspects of the investigation of business cycles, but it also has suggested new avenues for the empirically oriented, inductive side of that investigation. It is now well understood that purely deterministic dynamical systems can generate output mimicking true stochastic processes as accurately as we may wish. Although this is by no means a conclusive argument that business fluctuations are actually the output of chaotic deterministic systems, the result in question strongly suggests that, to describe complex dynamics mathematically, one does not necessarily have to make recourse to exogenous, unexplained shocks. The alternative option—the deterministic, or partially deterministic[3] description of irregular fluctuations—provides economists with new research opportunities undoubtedly worth exploiting.

Most questions raised among economists by the development of the modern theory of nonlinear dynamical systems are still open and actively debated in the literature. However, those questions are far from trivial mathematically and they cannot be discussed effectively in a nontechnical fashion, lest the meaning and the novelty of the results be lost. To maximize their fruitfulness for economics, the new concepts and methods of nonlinear dynamics should become part of the standard toolbox of economic theorists. This will permit economists to pursue a greater degree of independence from mathematicians and physical scientists, creatively adapting mathematical tools to their models rather than vice versa. Moreover, only a first-hand study of those methods and a sustained effort in applying them with originality to concrete economic problems will allow us to distinguish between what is of purely mathematical significance and what is also physically, i.e., economically, relevant. This process of assimilation of ideas and techniques occurred in the past with remarkable success for other areas of mathematics, such as calculus or matrix algebra, and, hopefully, it will take place for nonlinear dynamics as well.

We would like to contribute to this endeavor by reviewing some basic topics of the modern theory of dynamical systems. Because the subject matter is vast and complicated and the literature enormous, we have to be very selective. The discussion focuses on the concept of attractor and the properties of the different types of attractors occurring in economically motivated models. The presentation is divided into two parts, which are labeled "geometrical" and "ergodic" to reflect two fundamental theoretical approaches to the study of dynamical systems.

2. CONTINUOUS- AND DISCRETE-TIME SYSTEMS: FLOWS AND MAPS

In applied disciplines including economics, dynamical problems often are formulated as systems of differential or difference equations and, even though there exist other mathematical formulations of dynamics that are interesting and economically relevant, in this paper we concentrate on the former. The geometrical (or topological) approach to dynamics, which can be identified largely with the qualitative theory of differential/difference equations, studies the properties of the orbits of a dynamical system, looked at as geometrical structures in the state space.

Typically, a system of ordinary differential equations (ODE's) is written as[4]

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

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

Equation (1) also is called *vector field*, because its solution is a curve in the state space, whose velocity vector at each point x is given by $f(x)$. A solution of (1) is often written as a function $x(t)$, where $x : I \to \mathbf{R}^n$ and I is an interval in $\mathbf{R}$. In economic applications, where one wants to study the behavior of variables from here to eternity, typically $I = [0, +\infty)$.

We can also think of solutions of ODE's in a slightly different manner, which is becoming prevalent in dynamical system theory and will be very helpful for understanding some of the concepts discussed in the following sections. If we denote by $\phi(t, x) \equiv \phi_t(x)$ the state in $\mathbf{R}^n$ reached by the system at time t starting from x at t_0, then the totality of solutions of (1) can be represented by a map[5] $\phi : \mathbf{R} \times U \to \mathbf{R}^n$ or, equivalently, $\phi_t : U \to \mathbf{R}^n$, that satisfies (1) in the sense that

$$\frac{\partial}{\partial t}\phi(t, x) = f[\phi(t, x)]$$

for all t and x for which the solution exists. The map ϕ_t is called the *flow* (or the *flow map*) generated by f. The term flow often is used broadly to refer to continuous-time dynamical systems, as distinguished from discrete-time maps. A sufficient condition for a solution $\phi(t, x)$ through a point $x \in U$ to exist is that f be continuous. For such a solution to be unique, it is sufficient that f be continuous and differentiable in U.

The set of points $\{\phi(t, x) \mid t \in I\}$ defines an *orbit* of ϕ. It is a solution curve in the state space, parameterized by time. In this case, time typically appears in the form of arrows indicating the time direction of motion. The set $\{(t, \phi(t)) \mid t \in I\}$ defines a *trajectory* of ϕ and it is contained in the space of motions. However, in applications, the terms orbit and trajectory often are used as synonyms.

Remark 1. Notice the following properties of flows: (i) $\phi(0, x) = x$; (ii) time-translated solutions remain solutions; i.e., if $t, s \in \mathbf{R}$ are two different instants in time, we have $\phi(t + s, x) = \phi(s, \phi(t, x)) = \phi(t, \phi(s, x)) = \phi(s + t, x)$.

If time t is allowed to take only uniformly distributed, discrete values, separated by a fixed interval τ, from a (continuous-time) flow $\phi_t(x)$, we can derive a (discrete-time) map

$$x_{n+1} = G(x_n), \qquad \mathbf{(2)}$$

where $G = \phi_\tau$. Certain properties of continuous-time dynamical systems are preserved by this transformation and can be studied by considering the discrete-time systems derived from them. As we shall see later, for example, the Lyapunov characteristic exponents of a system like (1) are known to be the same as those of the associated discrete-time system defined by the so-called time-one-map $G = \phi_1$.

Remark 2. There exists another way of deriving a discrete-time map from a continuous-time dynamical system, called a Poincaré map, which describes the sequence of positions of a system generated by the intersections of an orbit in continuous time and a given space with a lower dimension, called surface of section. Clearly, in this case the time intervals between different pairs of states of the systems need not be equal. Poincaré maps are a powerful method of investigation of dynamical systems but they are seldom used in economics. [An exception is provided by Medio (1992, pp. 210–213).]

Remark 3. Whereas orbits of ODE's are continuous curves, orbits of maps are discrete sets of points. This has a number of important consequences, the most important of which can be appreciated intuitively. If the solution of a system of ODE's is unique, two solution curves cannot intersect one another in the state space. It follows that, for continuous-time dynamical systems of dimension one and two, the orbit structure must be constrained drastically. In the former, simpler case, we can only have fixed points and orbits leading to (or away from) them; in the two-dimensional case, nothing more complex than periodic orbits can occur. The restriction does not apply to maps, however, where orbits, so to speak, can "jump around." Consequently, even simple, one-dimensional nonlinear maps can generate very complicated orbits.

Of course, whenever the problem at hand allows it, discrete dynamical systems may be studied in their own right. More generally, we often wish to study problems resulting in types of maps that cannot be generated by systems of ODE's, in particular *noninvertible maps*.[6] Among other things, lack of invertibility in discrete-time systems raises a question of interpretation that is absent in continuous-time formulations. To see this, let us again consider system (2) and, for simplicity's sake, let $|\tau| = 1$. The fact that G is a function implies that, starting from any given point in space, there exists only one sequence $S = \{x^0, x^1, x^2, \ldots\}$ generated by applying G repeatedly to an initial point x^0. Whether S is interpreted as an orbit going forward or backward in time entirely depends on the time parameter t: The former alternative holds if we put $\tau = +1$, and therefore $t \in \mathbf{Z}^+$ (the set of nonnegative integers), the latter if $\tau = -1$ and therefore $-t \in \mathbf{Z}^+$. In the backward-moving case, (2) will be conveniently rewritten as

$$x_{t-1} = G(x_t) \tag{3}$$

or, equivalently, as

$$x_t = G(x_{t+1}). \tag{4}$$

If G is invertible, its well-defined inverse G^{-1} also will generate a unique sequence emanating from the same initial point and going in the time direction opposite to that of S. In this case therefore, system (2) unambiguously identifies forward and backward dynamics. However, if the map G is noninvertible (i.e., it is *many-to-one*) and $t \in \mathbf{Z}^+$, there will be a unique forward-moving orbit, but infinitely many backward-moving ones. Vice versa, if $-t \in \mathbf{Z}^+$, there will be a unique backward-moving orbit, but infinitely many forward-moving ones.

Interestingly, there exist economically motivated dynamical systems, in particular, some of those belonging to the vast family of overlapping-generations models, which result in backward-moving maps like (4). Mathematically, the situation is clear enough but, when applied to economic problems, it gives rise to delicate questions of interpretation. Real time flows only forward and real economic agents worry about the future not the past. The relevant question here is: Given a dynamical equation like (4) generating backward-moving orbits, what can we learn about the forward dynamics implicit in the economic specifications of the model? Some of the answers to this question can be found by applying a rather subtle mathematical technique known as natural extension, which relates the dynamics of a variable x in the original space to appropriately defined dynamics in a larger space whose elements are infinite sequences of x generated by G. We cannot discuss the question in any detail here but refer the interested reader to the specialized literature.[7]

For a system of ODE's like (1), a general solution $\phi(t, x)$ seldom can be written in a closed form, i.e., as a combination of known functions (powers, exponentials, logarithms, sines, cosines, etc.) or in the form of converging power series. Unfortunately, closed-form solutions are available only for special cases: linear systems with constant coefficients; one-dimensional ODE's (i.e., those for which $n = 1$); a small number of *very special* nonlinear differential equations of order greater than one. The generality of nonlinear systems that are studied in applications therefore escapes full analytical investigation; that is, an exact mathematical description of solution orbits cannot be found. Analogous difficulties arise when dynamical systems are represented by means of nonlinear maps. In this case, too, explicit solutions are generally available only for linear systems. In fact, even one-dimensional nonlinear maps usually cannot be given an explicit solution, although, for certain classes of such maps we can obtain a fairly complete qualitative analysis of their dynamics by means of a geometrical argument.

This is a very unsatisfactory situation since one-dimensional systems of ODE's, as well as linear systems of any dimensions in both continuous and discrete time, are not very interesting because their behavior is morphologically rather limited and they cannot be used effectively to represent cyclical or complex dynamics.[8]

Therefore, if we want to study interesting dynamical problems described by nonlinear differential or difference equations, we must change our orientation and adapt our goals to the available means. The short-run dynamics of individual orbits usually can 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 often are concerned not with short-term properties of individual solutions, but with the global qualitative properties of bundles of solutions that start from certain practically relevant subsets of initial conditions. Those properties can be investigated effectively by studying the asymptotic behavior of orbits and concentrating the analysis on regions of the state space that are *persistent* in the weak sense that trajectories never leave them, or in the stronger sense that trajectories are attracted to them. In what follows, we try to make these broad considerations more precise, providing criteria for a classification of basic types of dynamic behavior.

3. INVARIANT SETS, ATTRACTING SETS, AND ATTRACTORS

To discuss the persistence properties of orbits of a dynamical system, we shall start from the notion of invariant set. Formally, for the dynamical system (1) (respectively, (2)], we say that the set $S \subseteq U$ is invariant under the action of the flow ϕ_t or the map G, if we have

$$\phi_t(S) \subseteq S \qquad \forall t \in \mathbf{R}$$

$$[\text{respectively, } G^n(S) \subseteq S \qquad \forall n \in \mathbf{Z}].$$

If the abovementioned property holds for $\forall t \geq 0$ (respectively, $\forall n \geq 0$), S is called forward invariant. Notice that, if G is not invertible, G^n is defined only for $n \in \mathbf{Z}^+$. When constructing a mathematical model of the time evolution of certain physical, or economic, variables, we often wish to impose constraints on the acceptable values of those variables. For example, quantities such as capital stock, consumption, or relative prices should remain positive, or at least nonnegative 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. Clearly, the invariance of the set A of the acceptable values is a necessary (although not sufficient) condition for the validity of a dynamical model.

Investigation of invariant sets is an indispensable first step in the study of the dynamics of a system. Anybody who has performed numerical simulations of dynamical systems knows how important it is to locate regions of the state space such that the variables never escape from them. Invariant sets together with the orbits leading to or away from them play an essential role in the organization of the state space. The sets of points approaching an invariant set as $t \to \pm\infty$ sometimes are called *stable and unstable manifolds*, respectively. For nonhyperbolic sets, there also exist associated invariant sets called *center manifolds*, in which orbits may be converging or diverging. Center manifolds play an important role in the analysis of bifurcations [see e.g., Palis and Takens (1993)].

Remark 4. The notion of *hyperbolicity* is complex and cannot be discussed here in detail. Broadly speaking, for a dynamical system defined by a diffeomorphism $G : X \to X$, we say that a fixed point $p \in X$ of G is hyperbolic if the Jacobian matrix evaluated at p has no eigenvalue of modulus one. Hyperbolic fixed points have the property that small perturbations of the controlling map do not change the local dynamics qualitatively. The notion of hyperbolicity and many related results can be generalized so as to define hyperbolic structures for invariant sets different from fixed points [see Palis and Takens (1993, pp. 154–168)].

In most cases of practical interest, finding invariant sets is not enough. We also wish to locate the regions of the state space that ultimately attract all of the orbits originating in a certain (not too small) domain. The notion of attractiveness is intimately related to that of stability of orbits. Given the vastness of the subject (a few dozen different definitions of stability can be found in the literature), we deal with it only to the extent required by our main theme and suggest the following definition, based on Milnor (1985, p. 183) with minor adjustments.

DEFINITION 1. *A closed subset A of the state space X, invariant under a map G, is said to be Lyapunov stable if it has arbitrarily small neighborhoods V with* $G(V) \subset V$. *Moreover, A is asymptotically stable if it is Lyapunov stable and it has an open basin of attraction* $B(A)$. $B(A)$ *is the set of all points* $x \in X$ *such that* $G^n(x) \rightarrow A$ *for* $n \rightarrow \infty$. *In this case, if we choose V such that its closure* $cl(V) \subset B(A)$ *then* $A = \cap_{n>0} G^n(V)$.

This definition can be adapted easily to a flow ϕ_t. An asymptotically stable set also is called an *attracting set*. Notice that the location of the basin of attraction of a set seldom can be defined exactly and the structure of the basin of attraction (*b* of *a*) of complex (chaotic) sets also can have an extremely complicated structure.

The fact that a set is attracting does not mean that all of its parts are attracting too. Therefore, to describe the asymptotic regime of a system, we need the stronger concept of *attractor*. Besides being attractive, it also would be desirable for an attractor to be *indecomposable* or *irreducible*, in the sense that the asymptotic properties of different orbits originating in its basin of attraction $B(A)$ are qualitatively the same, independently of the initial conditions. Topologically, we say that an invariant set A is indecomposable if it is not the union of invariant subsets. In Part II, we discuss a deeper, probabilistic notion of indecomposability.

Strangely enough, there is no straightforward and universally adopted definition of attractor and, although the properties of the simpler cases can be dealt with easily, more complicated situations present difficult conceptual problems. For example, there exist well-known cases of dynamical systems that do not possess any set that can be defined as "an indecomposable attracting set," even though their orbits asymptotically converge toward a unique compact set [see Milnor (1985, p. 178)]. For the time being, we retain the operational, nonrigorous notion of an attractor as *a set on which experimental points generated by a flow or a map accumulate for large t*. The question of attractiveness is revisited in Part II from a measure–theoretic point of view.

Figure 1 shows a (chaotic) attractor with its basin of attraction.

The simplest type of attractor is a *stable fixed point*, or, using a terminology more common in economics, a stable *equilibrium*.

Remark 5. Economists often use a notion of equilibrium rather different from that of mathematicians and physicists, sometimes labeled 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. This representation implies the presence of two dynamic mechanisms: (i) short-run, fast dynamics generating a temporary equilibrium (market clearance), and (ii) long-run, slow dynamics describing the evolution of equilibria in time, not necessarily converging to a stationary state (fixed point). Thus, economic models of dynamic equilibrium implicitly postulate the hardly innocuous assumptions that the short-run dynamics are stable and the short-run adjustments do not interfere with the long-run evolution of the system.

https://doi.org/10.1017/S1365100598009079 Published online by Cambridge University Press

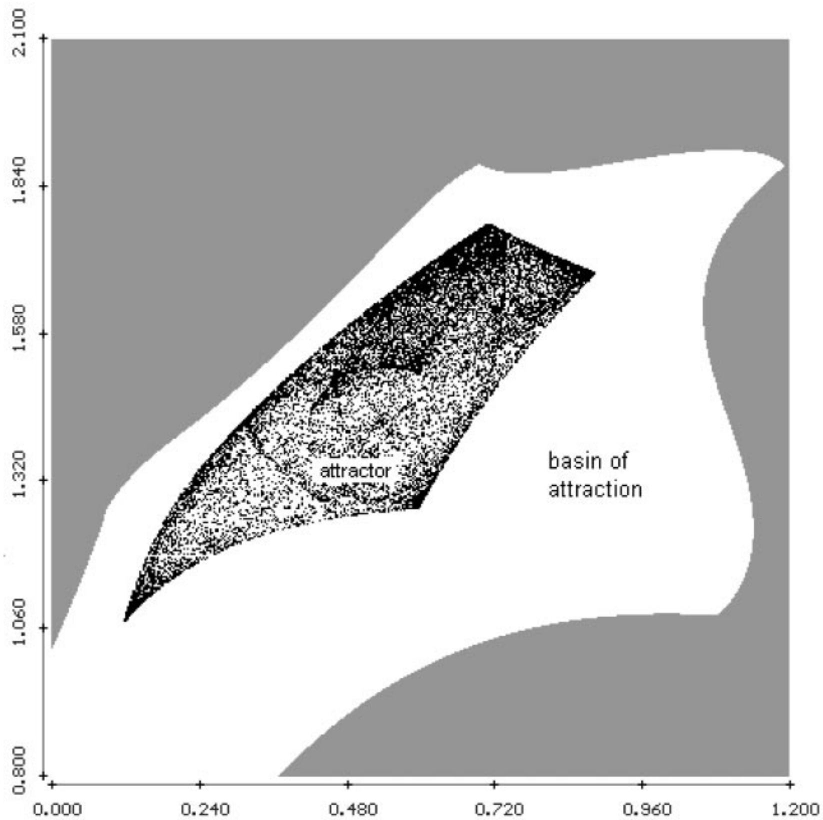


FIGURE 1. Chaotic attractor with its basin of attraction.

Mathematically, to prove the existence of a fixed/equilibrium point, we must solve a system of algebraic equations. In the continuous-time case (1),

$$\dot{x} = f(x),$$

the set of equilibria is defined by $E = \{\bar{x} \mid f(\bar{x}) = 0\}$, i.e., the set of values of x such that its rate of change in time is nil. Analogously, in the discrete-time case, $x_{n+1} = G(x_n)$, we have $E = \{\bar{x} \mid \bar{x} - G(\bar{x}) = 0\}$, i.e., the values of x that are mapped to themselves by G. Because the functions f and G generally are nonlinear, there are no ready-made methods to find the equilibrium solutions exactly, although geometrical and numerical techniques often give us all the qualitative information we need. Notice that linear systems typically have either one or no solution, whereas nonlinear systems typically have either no solutions, or a finite number of them. It follows that only nonlinear systems may generate the economically interesting phenomenon of (finite) multiple equilibria. In both the continuous- and the discrete-time cases, local stability of fixed points can be ascertained by well-known linearization techniques, whereas global stability sometimes can be established by the equally well-known second, or direct, method of Lyapunov.

Next in the scale of complexity of attractors, we consider *stable periodic solutions*, or *limit cycles*.

For a system of ODE's, we say that its solution is periodic if, for some positive τ, $\phi_\tau(x) = x$. The periodic orbit then is defined as the set $\Gamma = \{\phi_t(x) \mid t \in [0, \tau)\}$. The orbit Γ is stable if there exists a neighborhood U of it such that $\phi_t(x) \to \Gamma$ as $t \to \infty$. The *smallest* possible $\tau > 0$ for which this property holds is called the *period* of the orbit and

$$\omega/2\pi = 1/\tau$$

is its *frequency*, measured as number of cycles per unit of time. (Naturally, ω is the frequency measured in radians per unit of time.)

For maps, an n-periodic point $\bar{x}$ is a value of x such that $G^n(\bar{x}) = \bar{x}$. The orbit in this case is a sequence of n distinct points $\{\bar{x}, G(\bar{x}), \ldots, G^{n-1}(\bar{x})\}$ which, under the iterated action of G, are visited repeatedly by the system, always in the same order. Notice the following interesting facts:

1. Periodic solutions only exist for systems of ODE's of dimension ≥ 2. Linear systems of ODE's (of dimension ≥ 2) only have periodic solutions for very special parameter configurations. One-dimensional, discrete-time systems defined by a map G cannot have periodic solutions if G is a monotonically increasing function.
2. For a discrete-time dynamical system described by a map G, a periodic solution of period n corresponds to n fixed points of the map G^n. Thus, in principle, such periodic solutions always can be studied as fixed points. Needless to say, the greater the period n, the more complicated the calculations to find the corresponding fixed points.

To ascertain stability of limit cycles is not an easy matter except in the simpler cases. Earlier contributions to the nonlinear theory of business cycles typically described the economy by means of two-dimensional systems of ODE's. In this simple case, the existence of one or more limit cycles can be established by means of the Poincaré–Bendixson (PB) theorem.[9] There also exist analytical methods to establish existence and stability of cycles generated through a Hopf bifurcation (defined below). Because this paper is only marginally concerned with cycles as such, we do not expand this matter any further.

For reasons already discussed above, the stability of periodic orbits in discrete-time systems can, *in principle*, be analyzed by the same methods applied to fixed points. We are saying "in principle" because, for high-order periodic points, the necessary calculations may be very complicated.

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 oscillators (i.e., subsystems whose dynamics are periodic) with frequencies $\omega_1, \ldots, \omega_n$, such that two or more of the n frequencies are incommensurable (i.e., their ratios are irrational numbers). Clearly, in this case no real number τ can be found such that the system returns exactly to the same point from which it started. More precisely, we have the following:

DEFINITION 2. *A function $h : \mathbf{R} \to \mathbf{R}^n$ is called quasiperiodic if it can be written in the form $h(t) = H(\omega_1 t, \dots, \omega_n t)$, where H is periodic of period 2π in each of its arguments, and two or more of the n positive numbers ω_i (the frequencies) are incommensurable.*

In continuous time, the simplest case of quasiperiodic dynamics is given by a couple of oscillators, described by the linear system of ODE's, with state space $\mathbf{R}^4$

$$\begin{cases} \dot{x}_1 = -\omega_1 x_2 \\ \dot{x}_2 = \omega_1 x_1 \end{cases} \qquad \begin{cases} \dot{y}_1 = -\omega_2 y_2 \\ \dot{y}_2 = \omega_2 y_1. \end{cases} \tag{5}$$

In terms of polar coordinates, the dynamics of (5) can be described by the simpler system

$$\begin{cases} \dot{\theta}_1 = \omega_1, \\ \dot{\theta}_2 = \omega_2, \end{cases} \tag{6}$$

where $\dot{\theta}_i$ indicates the angular velocity and the state space is now $T^2 = S^1 \times S^1$, where S^1 denotes a circle (the space T^n is called an n-dimensional torus). There are two possibilities: (1) ω_1/ω_2 is a rational number; i.e., it can be expressed as a ratio of two integers p/q. In this case, there is a continuum of periodic orbits of period q. (2) ω_1/ω_2 is an irrational number. In this case, starting from any initial point in T^2, the orbit wanders on the torus, getting arbitrarily near any other point in it without ever returning to exactly the initial position.

Analogously, in discrete time, the simplest quasiperiodic map—the fixed rotation of the circle—can be written as

$$z_{n+1} = c z_n, \tag{7}$$

where $z \in S^1$ is a point on the unit circle, $c = e^{i2\pi\alpha}$, where $i = \sqrt{-1}$ and α is irrational.

Quasiperiodic orbits can look quite complicated, especially if there exist many incommensurable frequencies. As we see later, it was even conjectured (wrongly) that quasiperiodicity is the typical route to chaos.

Quasiperiodic dynamics have been found to occur in economically motivated dynamical models [see, e.g., Reichlin (1986); Woodford (1986); Medio (1992, Ch. 12), Venditti (1996)], and we have more to say about them later.

Attractors with an orbit structure more complicated than that of periodic or quasiperiodic systems are called *chaotic* or *strange* attractors. Although terminology is not yet uniform, the term chaotic often is referred to a dynamic property, i.e., the sensitive dependence on initial conditions (SDIC) or, equivalently, the divergence of nearby orbits, whereas the strangeness of an attractor mostly refers to its geometric characteristic of being a fractal set.

Remark 6. Notice that chaoticity, as defined by sensitive dependence on initial conditions, and strangeness, as defined by a fractal dimension (more precisely, by a noninteger FD), are independent properties. Thus, we have chaotic attractors that are not fractal and strange (fractal) attractors that are not chaotic. A well-known example of the second case is the so-called Feigenbaum attractor; an equally well-known example of the first case is the attractor generated by the so-called tent map (with slope = 2). For continuous-time dynamical systems, however, it is conjectured that, in general, strange attractors are also chaotic.

Chaotic dynamics typically occur when the overall contraction of volumes, which characterizes dissipative dynamical systems, takes place by shrinking in some directions, accompanied by less rapid stretching in the others. (However, as we see later, we can have one-dimensional, discrete-time maps for which there is no shrinking at all, even though, because of nonlinearity, the orbits of the system are confined to a bounded region of the space.) This fact has profound consequences because it implies that there may be an unstable motion *within the attractor*. An important consequence of this instability is that pairs of orbits that originate

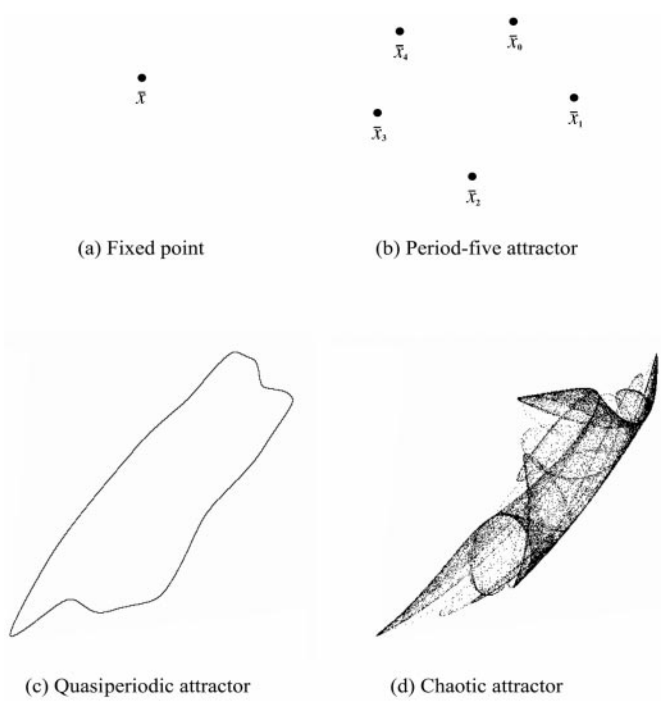


Figure 2. Different types of attractors in discrete time.

https://doi.org/10.1017/S1365100598009079 Published online by Cambridge University Press

from points arbitrarily near one another on the attractor become (exponentially) separated as time goes by. Thus, arbitrarily small errors in the measurement of initial conditions are magnified by the action of the flow or the map. As we show in Part II, 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, because chaotic attractors are bounded objects, the expansion that characterize their orbits must be accompanied by a folding action that prevents them from escaping to infinity. The combination of stretching and folding of orbits is a distinguishing feature of chaos and it is at the root of both the complexity of its dynamics and the strangeness of its geometry.

We discuss the fractal property of chaotic attractors briefly in the following section, whereas SDIC is given greater attention both here and in the ergodic part of the paper, because this property of chaos is, in our opinion, the most relevant to economics.

Figures 2 and 3 show the main different types of attractors, in discrete and continuous time, respectively.

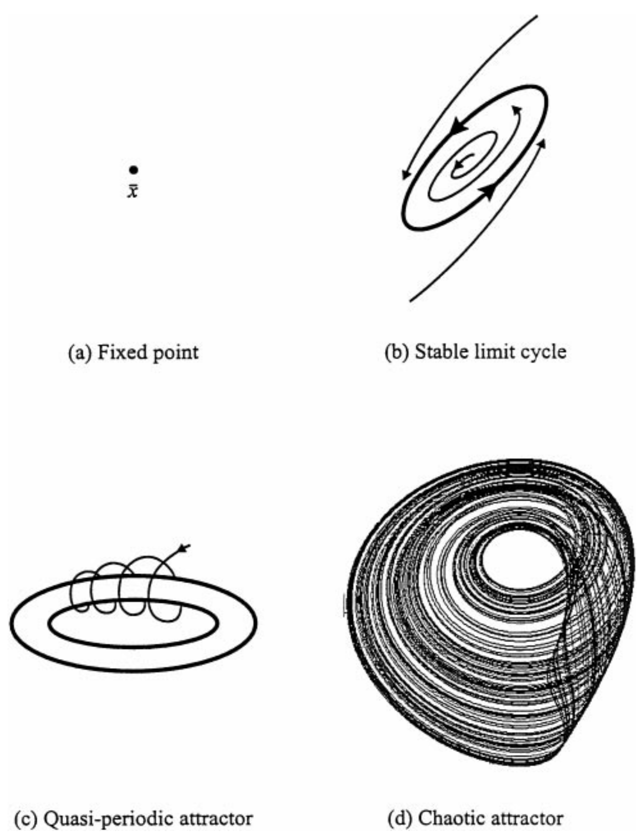


FIGURE 3. Different types of attractors in continuous time.

https://doi.org/10.1017/S1365100598009079 Published online by Cambridge University Press

4. FRACTAL DIMENSION

The term fractal was coined by Benoit Mandelbrot and it refers to geometrical objects characterized by self-similarity, i.e., objects having the same structure on all scales. Each part of a fractal object thus can be viewed as a reduced scale of the whole. Intuitively, a snowflake can be taken as a natural fractal. When applied to geometrical analysis of dynamical systems, the concept of FD can be conceived of as a measure of the way in which orbits fill the state space under the action of a flow or a map.

A simple, but precise, mathematical example can be given as follows: Consider a segment of unit length; divide it into three equal subsegments and remove the intermediate one. You will be left with two segments of length one-third. Divide each of them into three segments of length one-ninth and remove the (two) intermediate ones. If this procedure is iterated n times and we let n go to infinity, we obtain a set S called a Cantor set after the famous mathematician. Let us then pose the following question: What is the dimension of S? Were n finite, S would be a collection of segments and its Euclidean dimension would clearly be one. On the other hand, were S a finite collection of points, its dimension would be zero, but in the limit for $n \to \infty$, S is an infinite collection of points and therefore the answer is not obvious. To deal with this problem, the traditional Euclidean notion of dimension is insufficient and we need the more sophisticated criterion of FD. There exist a rather large family of such dimensions, but we limit ourselves here to the simplest example.

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 in the construction of the Cantor set or, more generally, 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, $N(\epsilon)$, necessary to cover S.

Then, the FD D of the set S will be given by the following limit (assuming it exists):

$$D \equiv \lim_{\epsilon \to 0} \frac{\log[N(\epsilon)]}{\log(1/\epsilon)}. \qquad \textbf{(8)}$$

The quantity defined in (8) is 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, $N(\epsilon) = 1$, and $D = 0$; if it is a segment of unit length, $N(\epsilon) = 1/\epsilon$, and $D = 1$; finally, if S is a plane of unit area, $N(\epsilon), = 1/\epsilon^2$ and $D = 2$, etc. That is to say, for regular geometrical objects, the dimension D does not differ from the usual Euclidean dimension and, in particular, D is an integer. By making use of the notion of capacity dimension, we can determine the dimension of the Cantor set constructed above. To do this, let us try to evaluate the limit (8), proceeding step by step. Consider first a (one-dimensional) box of side ϵ. Clearly, we have $N(\epsilon) = 1$ for $\epsilon = 1$, $N(\epsilon) = 2$ for

$\epsilon = 1/3$, and, generalizing, $N(\epsilon) = 2^n$ for $\epsilon = (1/3)^n$. Taking the limit for $n \to \infty$ (or, equivalently, taking the limit for $\epsilon \to 0$), we can write

$$D = \lim_{\substack{n \to \infty \\ (\epsilon \to 0)}} \frac{\log 2^n}{\log 3^n} \approx 0.63. \tag{9}$$

We thus have provided a quantitative characterization of a geometric set that is more complex than the usual Euclidean objects. Indeed, we have just found that the dimension of S is a noninteger. We might say that S is an object dimensionally greater than a point but smaller than a segment. It also can be verified that the set S is characterized by self-similarity.

The concept and measurement of FD are not only necessary to understand the finer geometrical nature of strange attractors, but they are also very useful tools for providing quantitative analyses of such attractors, both in theoretical and applied investigations of dynamical systems. A well-known example of an application of the notion of FD to the analysis of time series is the so-called Brock-Dechert-Scheinkmann (BDS) test [see Brock et al. (1987)].

5. LYAPUNOV CHARACTERISTIC EXPONENTS

To provide a rigorous characterization, as well as a way of measuring SDIC, we discuss a powerful conceptual tool known as Lyapunov characteristic exponents (LCE's). LCE's are discussed here only in relation to maps because, as we mentioned before, for the present purpose the analysis of continuous-time systems can be reduced to the discrete-time case. We start with a simple, one-dimensional setting and then generalize it. Consider the map

$$x_{n+1} = T(x_n) \tag{10}$$

with $T : U \to \mathbf{R}$, U being an open subset of $\mathbf{R}$. We now want to describe the evolution in time of two orbits originating from two nearby points x_0 and $\hat{x}_0$. At the nth iteration, we have

$$\hat{x}_n - x_n = T^n(\hat{x}_0) - T^n(x_0) = G(\hat{x}_0, x_0). \tag{11}$$

If we now take x_0 as a constant, expand G in a Taylor series around $\hat{x}_0 = x_0$, and retain only the first-order term, we have

$$\hat{x}_n - x_n \simeq \left.\frac{dT^n}{dx}\right|_{\hat{x}_0 = x_0} (\hat{x}_0 - x_0) = [T'(x_0)T'(x_1)\cdots T'(x_{n-1})](\hat{x}_0 - x_0), \tag{12}$$

where, provided that the derivative is different from zero, the approximation can be made arbitrarily accurate by taking $(\hat{x}_0 - x_0)$ sufficiently small. Asymptotically,

https://doi.org/10.1017/S1365100598009079 Published online by Cambridge University Press

we then have

$$\lim_{n\to\infty} |\hat{x}_n - x_n| \simeq e^{n\lambda(x_0)}|\hat{x}_0 - x_0|. \quad \textbf{(13)}$$

The quantity

$$\lambda(x_0) = \lim_{n\to\infty} \frac{1}{n} \log|T'(x_0)T'(x_1)\cdots T'(x_{n-1})| = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \log|T'(x_j)| \quad \textbf{(14)}$$

is called the LCE. From (13), the interpretation of $\lambda(x_0)$ is straightforward: It is the (local) asymptotic exponential rate of divergence of nearby orbits.[10]

In the multidimensional case, the calculation of LCE's can be done in a similar, albeit more complex, manner. Generally, we have

$$x_{n+1} = T(x_n),$$

where $x_n = (x_n^1, \ldots, x_n^m) \in \mathbf{R}^m$ and $\boldsymbol{T}$ is a vector of functions $T_i : \mathbf{R}^m \to \mathbf{R}$ $(i = 1, \ldots, m)$. Consider now a point x_0 in the state space and a nearby point $\hat{x}_0$. The rate of change of their distance, as measured, for example, by the Euclidean norm $\|\cdot\|$, will evolve under the action of the map T according to the ratio

$$\|T^n(\hat{x}_0) - T^n(x_0)\| / \|\hat{x}_0 - x_0\|. \quad \textbf{(15)}$$

Because we want to study the time evolution of nearby orbits, in the limit as $\hat{x}_0 \to x_0$, (15) can be expressed as

$$\frac{\left\| \prod_{i=0}^{n-1} DT(x_i)w \right\|}{\|w\|} = \frac{\|DT^n(x_0)w\|}{\|w\|}, \quad \textbf{(16)}$$

where $DT(x)$ is $(m \times m)$ matrix with typical element $[\partial T_i / \partial x_j]$ and w is a vector in the *tangent space* at x_0. We now can define the LCE's of the vector w (which generally depend on x_0) as follows:

$$\lambda(x_0, w) = \lim_{n\to\infty} \frac{1}{n} \log \frac{\|DT^n(x_0)w\|}{\|w\|} \quad \textbf{(17)}$$

There generally will be m such exponents, some of which may coincide. From the formula (17), we notice that, in general, LCE's depend not only on w but also

https://doi.org/10.1017/S1365100598009079 Published online by Cambridge University Press

on x_0. In Part II, we see under what conditions LCE's may become independent of initial conditions. A thorough discussion of the analytical and computational problems related to LCE's is provided by Benettin et al. (1980).

LCE's provide an extremely useful tool for characterizing the behavior of nonlinear dynamical systems. Because they are invariant with respect to smooth, invertible changes of coordinates, it is sometimes possible to evaluate the LCE's of a map by calculating those of a different and simpler one. The signs of LCE's are especially important to classify different types of dynamical behavior. The typical situations can be summarized thusly:

1. Asymptotically stable, fixed points of an m-dimensional system, both in discrete and continuous time, are associated with m negative LCE's.
2. The m-dimensional continuous-time systems with nonchaotic attractors different from fixed points have $m - n$ negative LCE's ($n < m$), and n zero LCE's with $n = 1$ for a limit cycle or $n = k$ for quasiperiodic dynamics on a T^k torus.
3. Chaotic attractors, both in discrete and continuous time, are associated with the presence of at least a positive LCE, which 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. As we see in Part II, the presence of one or more positive LCE is intimately related to the lack of predictability of dynamical systems, which is an essential feature of chaotic behavior.

6. TRANSITION TO CHAOS

In the preceding sections, we provide a classification of attractors and discuss some distinctive properties of chaotic attractors. The relevance of these findings would be greatly enhanced if, in addition, we could describe the *qualitative changes in the orbit structure* of the system that take place when certain control parameters are varied. In this way, we would obtain not only a snapshot of chaotic dynamics, but also, so to speak, a stroboscopic 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. 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—such as the appearance or disappearance of a fixed point or of another invariant set—when one or more parameter is changed. Bifurcation theory is a vast and complex area of investigation and we consider it here only concisely.

There exist three canonical types of routes to chaos generated by so-called codimension-one bifurcations, i.e., bifurcations depending on a single parameter and we briefly deal with them in turn.

6.1. Period Doubling

This is probably the best known route to chaos, at least in economic literature. It takes place in both discrete- and continuous-time systems and can be described most simply by considering its occurrence in the celebrated logistic map

$$x_{t+1} = f(x_t) = rx_t(1 - x_t) \qquad x \in [0, 1], \quad r \in (1, 4]. \tag{18}$$

For values of r slightly greater than 1, map (18) has two nonnegative equilibria, $\bar{x}_1 = 0$ and $\bar{x}_2 = 1 - 1/r$. The former is unstable and the latter is stable. As r goes through $r_1 = 3$, a bifurcation called a flip occurs and the situation changes: $\bar{x}_2$ becomes unstable as well and a stable two-cycle is born. It is easy to verify that the two points of the cycle of the map f are also stable equilibria of the map $G(x) \equiv f^2$, i.e., the second iterate of f. As we further increase r, the initially stable two-cycle loses stability at $r_2 = 1 + \sqrt{6}$ when a second flip bifurcation occurs and a stable four-cycle is created. This scenario is repeated over and over again and leads to an infinite sequence of flip bifurcations and *period doublings*. The sequence $\{r_k\}$ of values of r at which k-cycles occur has a finite accumulation point $r_\infty \approx 3.569446$, involving an infinity of periodic orbits, all unstable. The limit set corresponding to r_∞ is known as the Feigenbaum attractor. It is a geometric object with a noninteger FD ≈ 0.538 and an *LCE equal to zero* and, consequently, the motion on it is not chaotic in the sense defined above.

Remark 7. Notice that the Feigenbaum attractor, although Lyapunov stable, is not asymptotically stable and therefore is not an attracting set as defined above [cf. Milnor (1985, pp. 182–183)].

Past r_∞, we enter what usually is called the chaotic zone (CZ). This phrase means that, in CZ, there exist (infinitely many) values of r for which orbits are nonperiodic but not chaotic (i.e., the LCE $\lambda = 0$) or even chaotic (i.e., $\lambda > 0$ and orbits therefore are characterized by sensitive dependence on initial conditions). This does not mean, however, that only aperiodic or chaotic orbits exist in CZ. Although the chaotic set $\{r \mid \lambda > 0\}$ is infinite and it is believed to have positive Lebesgue measure (i.e., choosing r at random, we have a nonzero chance of getting a chaotic value), it nowhere forms an interval. Thus, starting from a chaotic value of r, ever so slight changes in this parameter may destroy the chaotic dynamics and bring about a periodic solution with $\lambda < 0$. On the contrary, in the CZ there exist infinitely many periodic windows, i.e., intervals of values of r to which there correspond stable periodic orbits. Some of these windows—for example the period-3 window—are particularly evident and can be located easily by simple inspection of a bifurcation diagram (see Figure 4). As we see later while revisiting

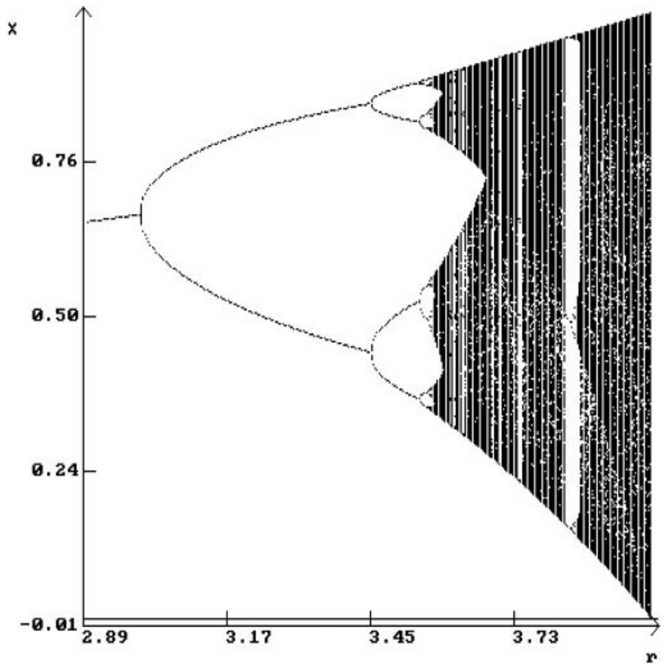


FIGURE 4. Period-doubling route to chaos.

the question of LCE's, the case $r = 4$ is special because, for that value, we can determine the unique LCE exactly and thereby establish the chaotic nature of orbits with certainty.

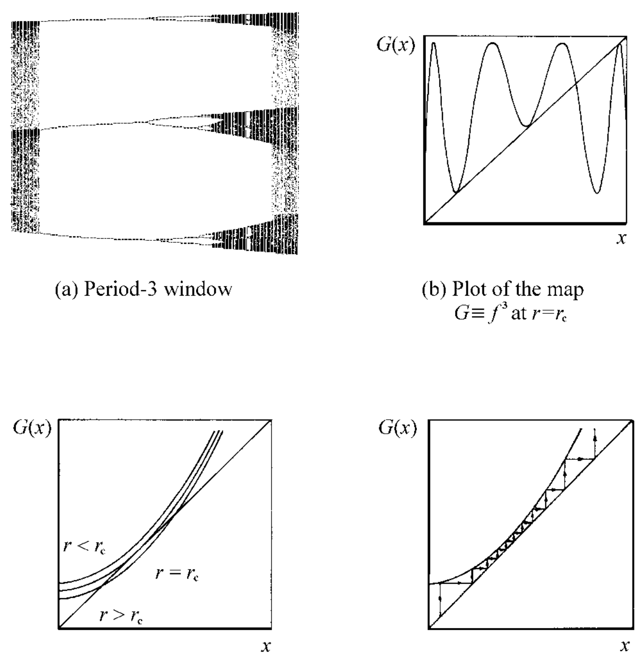
The period-doubling process is ubiquitous and can be found in a discrete-time map of any dimensions as well as in continuous-time systems of differential equations of dimension equal to or greater than 3. In continuous time, the opening step of a period-doubling sequence takes place typically through a Hopf bifurcation.[11] Notice that the sequence is not always complete, i.e., it need not lead to a chaotic zone. That is, for certain maps when we increase the controlling parameter monotonically, we may have the curious phenomenon of a partial period doubling followed by a period halving that eventually leads back to a stable fixed point.

6.2. Intermittency and Crisis

Broadly speaking, *intermittency* is a phenomenon characterized by alternation of simple, quiet dynamics and bursts of wild oscillations. There is strong experimental evidence of intermittency in physical sciences and it is also present in the output generated by economically motivated dynamical models. In true experiments as well as in numerical simulations, intermittency is a sign of impending chaos. There exist different types of intermittency, but here we only mention the simplest one, which can be discussed by means of the familiar logistic map (18). Consider the situation for values of the parameter r in the vicinity of $r_c = 1 + 2\sqrt{2} \approx 3.828427$, i.e., near the left boundary of the period-3 window. For $r < r_c$, the map f has two unstable fixed points, namely 0 and $1 - 1/r$. As r *increases* through the critical value r_c, the map f acquires a stable and an unstable period-3 cycle. The three points of each cycle are stable (respectively, unstable) fixed points of the map $G(x) \equiv f^3(x)$. If we reverse the procedure and *decrease* r, at $r = r_c$ three pairs of (stable and unstable) fixed points of G will coalesce and disappear. The bifurcation occurring at r_c is known as *saddle node*, or *tangent bifurcation*, and is associated with the sudden, catastrophic disappearance, or appearance, of equilibria.

For $r > r_c$, the asymptotic dynamics generated by f are a period-3 orbit, whereas, for values of r slightly smaller than r_c, we have a regular, almost periodic, motion, which is interrupted from time to time by bursts of apparently chaotic behavior. Although the overall motion is aperiodic, most iterates of the map are concentrated in the neighborhoods of the three points of the (disappeared) period-3 orbit, and the average duration of regular dynamics is a continuous (inverse) function of the distance $|r - r_c|$.

This rather curious behavior can be better understood by inspecting Figure 5.



(a) Period-3 window

(b) Plot of the map $G \equiv f^3$ at $r = r_c$

(c) Expanded view of (b) near a tangency point for $r \gtreqless r_c$

(d) Slow motion "through the channel" for $r < r_c$

FIGURE 5. Intermittency.

https://doi.org/10.1017/S1365100598009079 Published online by Cambridge University Press

For r slightly smaller than r_c, the motion of the system slows down in the vicinity of the locus where a stable fixed point was, as if it were still looking for it. After a certain number of apparently regular iterations (the greater the number the nearer r is to r_c), the system leaves the neighborhood and wanders away in an irregular fashion until it is reinjected in the channel between the curve of the map and the bisector, and so on and so forth.

Inspection of Figures 4 and 5a also reveals another interesting case of qualitative change in the orbit structure. If we continue to increase the parameter r past r_c, we have a series of period-doubling bifurcations leading to orbits of period $3 \cdot 2^i$ ($i = 2, 3, \ldots$), and eventually to the so-called periodic chaos (or noisy periodicity), i.e., chaotic behavior restricted to three narrow bands, each of them visited periodically. Past a certain critical value of r, the situation changes again discontinuously and the three bands suddenly merge and broaden to form a single chaotic region. This phenomenon is sometimes called (interior) *crisis*.

6.3. Quasiperiodic Route to Chaos

The idea that quasiperiodicity is the fundamental intermediate step in the route to chaos, is a long-standing one. Over 50 years ago, the Russian physicist Landau suggested that turbulence in time,[12] or chaos, is a regime approached by a dynamical system through an infinite sequence of Hopf bifurcations that takes place when a certain parameter is changed. Thus, the dynamics of the system would be periodic after the first bifurcation, and quasiperiodic after successive bifurcations, with an ever-increasing degree of quasiperiodicity, leading in the limit to turbulent, chaotic behavior.

The conjecture that Landau's scenario is the only, or even the most likely, route to chaos was rejected on the basis of two basic results, namely:

- From a theoretical point of view, Ruelle and Takens (1971) and Newhouse et al. (1978) proved that systems with four-torus (or even three-torus) attractors are unlikely to be observed, because they are easily perturbed to chaos.
- Lorenz's (1963) work on turbulence showed that complexity (in the sense of a large dimension of the system) is not a necessary condition for chaos to occur, and that low-dimensional systems are perfectly capable of producing a chaotic output.

Thus, although experimental results seem to suggest the possibility of a direct transition from quasiperiodicity to chaos, mathematically this is still an open question. Some of the analytical aspects of this problem have been investigated by studying the circle map, which, in turn, is related to the so-called Neimark (or Neimark-Saker) bifurcation of discrete-time dynamical systems. This bifurcation and the ensuing periodic/quasiperiodic/chaotic scenario is interesting to economists because it has been shown to occur in economically motivated models, such as the celebrated overlapping-generations model. We deal with this issue in greater detail in Part II.

6.4. Homoclinic Bifurcations

All the bifurcations we have considered so far are "local" in the sense that their occurrence can be ascertained by investigating changes taking place around a certain fixed point (or a cycle). In particular, those bifurcations can be related to the fact that certain eigenvalues go through zero (continuous time) or through unity in modulus (discrete time). There also exist, however, bifurcations that are "global" in the sense that they are generated by changes in the structure of the system occurring far from a fixed point (or a cycle) and leave the local orbit structure unchanged. These bifurcations are much more difficult to ascertain analytically and we shall discuss them here only very briefly.

In certain cases, global bifurcations may open a route to chaos. A well-known example is the so-called *homoclinic bifurcation* that occurs when, changing a certain controlling parameter, the stable and unstable manifolds of a saddle point intersect. (An analogous event is the *heteroclinic bifurcation*, which is related to intersections of stable and unstable manifolds of two different saddle points.) Homoclinic bifurcations may lead to the creation (or annihilation) of chaotic attractors. An interesting phenomenon related to homoclinic bifurcations is the so-called *homoclinic tangency* which, as the name suggests, takes place when stable and unstable manifolds become tangent. This phenomenon was shown to be generic for a broad class of dynamical systems and may give rise to chaotic attractors or to the coexistence of infinitely many periodic attractors. In the latter case, strictly speaking, one could not define the dynamics as chaotic. However, because the basin of attraction of each periodic attractor must be very small, in the presence of perturbations such as those introduced by numerical simulations, the orbits of a system exhibiting homoclinic tangency may be practically indistinguishable from those of truly chaotic systems.[13] Homoclinic bifurcations as defined before belong to a broad class of global bifurcations including various types of saddle collisions, occurring when the stable or the unstable manifold of an invariant set of a saddle type collides with another invariant set such as a limit cycle. An interesting example of saddle collision, nicknamed "blue sky catastrophe," is depicted by Abraham and Shaw [(1988, p. 143); cf. also Medio (1992, pp. 169–173)]. Global bifurcations of the saddle collision type and, in particular, homoclinic and heteroclinic tangencies, have been found in economically motivated models [see, e.g., Brock and Hommes (1997, pp. 1059–1095)].

6.5. Hysteresis

An interesting nonlinear phenomenon related to (local or global) bifurcations leading to catastrophic appearance or disappearance of invariant sets such as fixed points, periodic orbits, or chaotic attractors is the so-called *hysteresis*. Broadly

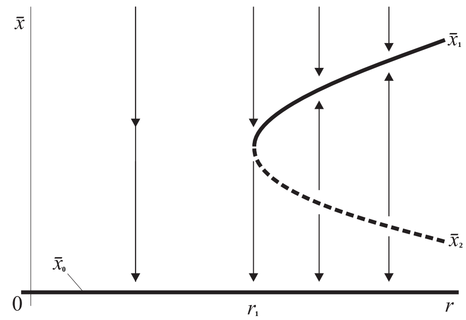


Figure 6. Hysteresis.

speaking, we say that hysteresis occurs when the effects on the dynamics of a system of changes of a parameter in one direction are not reversed when the parameter is changed back to the original value. A simple example is illustrated by Figure 6.

On the ordinate of Figure 6, we measure *the equilibrium values* of the state variable, x, whereas on the abscissa we measure the controlling parameter r. We assume that, for all positive initial conditions, the following properties hold: (i) for $0 < r < r_1$, the system is characterized by a unique, stable equilibrium $\bar{x}_0 = 0$; and (ii) for $r > r_1$, there are three equilibria $\bar{x}_0 = 0$ which is still stable, $\bar{x}_1$ which is stable, and $\bar{x}_2$ which is unstable; (iii) at $r = r_1$, the two nontrivial equilibria coalesce and a saddle-node bifurcation occurs. Suppose now that we fix $r > r_1$. For initial conditions inside or above the parabolic curve of equilibria, the system will converge toward the equilibrium $\bar{x}_1$. Suppose now we fix the parameter r at progressively lower values. Clearly, for $r < r_1$, the system, starting from the same initial conditions as before, will converge to zero. If we now revert the procedure and increase r past r_1, keeping the initial conditions near zero, the system will continue to converge to the trivial equilibrium. If we wish to move back to $\bar{x}_1$, increasing the parameter is not enough, we need to give the system a push and move it again inside the parabola.

The phenomenon of hysteresis is well known in economic literature and it suggests interesting considerations of economic policy [for a recent application to a problem of labor economics, see Brunello and Medio (1996)].

7. TOPOLOGICAL EQUIVALENCE

When confronted with a problem P_1 too hard to solve, we sometimes can overcome the difficulty by defining another, easier problem P_2, solving it and then establishing some equivalence relation between P_1 and P_2 so that the results established for the latter can be extended to the former. This strategy is applied to the study of dynamical systems by means of the concept of topological equivalence, which can be defined as follows:

DEFINITION 3. *Two maps $G_1 : X \to X$ and $G_2 : Y \to Y$ are topological equivalent if there exists a homeomorphism $h : X \to Y$ such that the following diagram commutes*

$$\begin{array}{ccc} X & \xrightarrow{G_1} & X \\ h \downarrow & & \downarrow h; \\ Y & \xrightarrow[G_2]{} & Y \end{array}$$

i.e., we have $h \circ G_1 = G_2 \circ h$.

Under topological equivalence, h maps orbits generated by G_1 to orbits generated by G_2, preserving the order of points as well as most basic properties of the dynamics although some details may be lost.[14] An analogous definition is available for continuous-time dynamical systems with similar implications. A celebrated and widely exploited case of topological equivalence—first studied by Ulam and von Neumann in 1947—relates the logistic map with the parameter r equal to 4, and the so-called tent map which, being semilinear, is much easier to analyze. We make use of this result (and some additional equivalence properties of those maps) in Part II.

NOTES

1. 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. In recent years, the latter case has been studied extensively in the economic literature under the label "sunspots."

2. This is no place for a comprehensive survey of the literature on these and similar variations of the basic equilibrium dynamics model, but we quote some of the most significant contributions: on the role of impatience, see Benhabib and Nishimura (1979), Deneckere and Pelikan (1986), and Boldrin and Montrucchio (1986); on overlapping-generations models, Benhabib and Day (1982), Grandmont (1985), and Reichlin (1986); on imperfectly competitive markets, Woodford (1986); on intertemporally dependent utility, Ryder and Heal (1973).

3. The phrase "partially deterministic" denotes here a dynamical system with an identifiable determinist core plus a stochastic perturbation.

4. Systems described by equation (1), in which f does not depend directly on t, are called *autonomous*. If f does depend on t directly, we can write

$$\dot{x} = f(x, t), \quad (x, t) \in \mathbf{R}^n \times \mathbf{R} \tag{$\mathbf{1'}$}$$

and $f : U \to \mathbf{R}^n$ with U an open subset of $\mathbf{R}^n \times \mathbf{R}$. Equations of type (1′) are called *nonautonomous*. In economics they are used, for example, to investigate technical progress.

5. Broadly speaking, a *map is a function* whose iterates are used to describe the dynamics of a system in a certain space. The term *mapping* also is used.

6. For a map to be conceived as a *flow map*, it must be *orientation-preserving*. For a map $\phi_t(x) : U \to \mathbf{R}^n$ defined above, the latter property holds if and only if

$$\det[D_x \phi_t(x)] > 0, \qquad \forall x \in U,$$

where $D_x \phi_t = [\partial \phi_t^j / \partial x_i]$ $(i, j = 1, \ldots, n)$ denotes the matrix of partial derivatives.

7. See, for example, Cornfeld et al. (1994). For a recent application to OLG models, see Medio (1998).

8. However, linear theory is important in the analysis of nonlinear systems because it can be employed to investigate qualitatively their *local* behavior, e.g., their behavior in an arbitrarily small neighborhood of a single point or of a periodic orbit. This is particularly important in stability analysis and in the study of (local) bifurcations.

9. The PB theorem states that if the orbit of a continuous-time, two-dimensional dynamical system enters and never leaves a closed and bounded region C, and there are no fixed points in C, then there exists at least one periodic orbit (a *limit cycle*) in C. In general, if only one such orbit exists, it is asymptotically stable. [Cf. Glendinning (1994, pp. 132–137).]

10. It is local because we evaluate the rate of separation in the limit for $\hat{x}_0 \to x_0$. It is asymptotic because we evaluate it in the limit of indefinitely large number of iterations—assuming that the limit exists.

11. The Hopf bifurcation takes place when a stable fixed point of a continuous-time dynamical system loses its stability as the real part of a pair of complex conjugate eigenvalues of the Jacobian matrix evaluated at the fixed point, goes through zero with nonzero velocity, and becomes positive. That leads to the creation of a (stable or unstable) limit cycle. For a more thorough discussion of the Hopf bifurcation, see, for example, Glendinning (1994, pp. 224–244); for an application to an economically motivated model, see Invernizzi and Medio (1991). An example of a period-doubling route to chaos starting from a Hopf bifurcation can be found in Medio (1991; 1992, Ch. 13).

12. We specify "turbulence in time," to distinguish it from "full turbulence" that takes place in both time and space, which we do not discuss in the present survey.

13. For a rigorous, detailed discussion of homoclinic tangency and related phenomena, see Palis and Takens (1983).

14. For example, topological equivalence does not distinguish between nodes and foci. Equivalence is stricter when h is a C^k diffeomorphism with $k > 0$. In this case, the term C^k conjugacy is used.

REFERENCES

Abraham, R.H. & C.D. Shaw (1988) *Dynamics, the Geometry of Behavior*, Pts. I and II. Santa Cruz: Aerial.

Benettin, G., L. Galgani, A. Giorgilli & J.M. Strelcyn (1980) Lyapunov characteristic exponents for smooth dynamical systems: A method for computing them all. *Meccanica* 15, 21–30.

Benhabib, J. & R.H. Day (1982) A characterization of erratic dynamics in the overlapping generations models. *Journal of Economic Dynamics and Control* 4, 37–55.

Benhabib, J. & K. Nishimura (1979) The Hopf bifurcation and the existence and stability of closed orbits in multisector models of optimal economic growth. *Journal of Economic Theory* 21, 417–444.

Blanchard, O.J. & S. Fischer (1989) *Lectures on Macroeconomics*. Cambridge, MA: MIT Press.

Boldrin, M. & L. Montrucchio (1986) On the indeterminacy of capital accumulation paths. *Journal of Economic Theory* 40, 26–39.

Brock, W.A., W.D. Dechert & J. Scheinkmann (1987) A Test for Independence Based on the Correlation Dimension. SSRI working paper 8702, University of Wisconsin.

Brock, W.A. & C.H. Hommes (1997) A rational route to randomness. *Econometrica* 65, 1059–1095.

Brunello, G. & A. Medio (1996) A Job Competition Model of Workplace Training and Education. Nota di Lavoro 96.16. University of Venice.

Cornfeld, I.P., S.V. Fomin & Ya.G. Sinai (1994) *Ergodic Theory*. Berlin: Springer-Verlag.

Day, R.H. (1982) Irregular growth cycles. *American Economic Review* 72, 406–414.

Deneckere, R. & S. Pelikan (1986) Competitive chaos. *Journal of Economic Theory* 40, 13–25.

Frisch, R. (1933) Propagation problems and impulse problems in dynamic economics. In *Economic Essays in Honour of Gustav Cassel*, pp. 171–205. London: Allen and Unwin.

Glendinning, P. (1994) *Stability, Instability and Chaos: An Introduction to the Theory of Nonlinear Differential Equations*. Cambridge, England: Cambridge University Press.

Goodwin, R.M. (1951) The nonlinear accelerator and the persistence of business cycles. *Econometrica* 18, 1–17.

Grandmont, J.M. (1985) On endogenous competetive business cycles. *Econometrica* 53, 995–1045.

Hicks, J.R. (1950) *A Contribution to the Theory of the Trade Cycle*. Oxford: Oxford University Press.

Kaldor, N. (1940) A model of the trade cycle. *Economic Journal* 50, 78–92.

Invernizzi S. & A. Medio (1991) On lags and chaos in economic dynamic models. *Journal of Mathematical Economics* 29, 521–550.

Lorenz, E.N. (1963) Deterministic non-periodic flows. *Journal of Atmospheric Science* 20, 131–140.

Medio, A. (1991) Continuous-time models of chaotic dynamics in economics. *Journal of Economic Behavior and Organization* 16, 115–151.

Medio, A. (1992) *Chaotic Dynamics. Theory and Applications to Economics*. Cambridge, England: Cambridge University Press.

Medio, A. (1998) The problem of backward dynamics in economics. Nota di Lavoro 98.20. University of Venice.

Milnor, J. (1985) On the concept of attractor. *Communications in Mathematical Physics* 99, 177–195.

Muth, J. (1961) Rational expectations and the theory of the price movements. *Econometrica* 29, 315–335.

Newhouse, S.E., D. Ruelle & F. Takens (1978). Occurrence of strange axiom A attractors near quasiperiodic flows on T^m, $m \geq 3$. *Communications of Mathematical Physics* 64, 35–40.

Palis, J. & F. Takens (1993). *Hyperbolicity & Sensitive Chaotic Dynamics at Homoclinic Bifurcations*. Cambridge, England: Cambridge University Press.

Reichlin, P. (1986) Equilibrium cycles in an overlapping generations economy with production. *Journal of Economic Theory* 40, 89–102.

Ruelle, D. & F. Takens (1971) On the nature of turbulence. *Communications of Mathematical Physics* 20, 167–192; 23, 343–344.

Ryder, H.E. & G.M. Heal (1973) Optimal growth with intertemporary dependent preferences. *Review of Economic Studies* 1, 1–31.

Slutzky, E. (1927) The summation of random causes as the source of cyclical processes. *Econometrica* 5, 105–146.

Venditti, A. (1996) Hopf bifurcation and quasi-periodic dynamics in discrete multisector growth models. *Research in Economics/Ricerche Economiche* 50, 267–291.

Yule, G.U. (1927) On a method of investigating periodicities in disturbed series. *Philosophical Transactions of the Royal Society of London, Series A: Mathematical and Physical Sciences* 226, 267–298.

Ulam, S.M. & J. von Neumann (1947) On combination of stochastic and deterministic processes. *Bulletin of the American Mathematical Society* 53, 1120–1132.

Woodford, M. (1986) Stationary sunspot equilibria in a finance constrained economy. *Journal of Economic Theory* 40, 128–137.