

8. S. P. Havera and F. C. Bellrose, *Wetlands* 4, 19 (1985).
9. F. H. Nichols, J. E. Cloern, S. N. Luoma, D. H. Peterson, *Science* 231, 567 (1986).
10. F. C. White, J. R. Hairston, W. N. Musser, H. F. Perkins, J. F. Reed, *J. Soil Water Conserv.* 36, 172 (1981).
11. P. Crosson, *Conservation Tillage and Conventional Tillage: A Comparative Assessment* (Soil Conservation Society of America, Ankeny, IA 1981); R. E. Sijka, G. W. Langdale, D. L. Karlen, *Adv. Agron.* 37, 155 (1984). Conservation tillage reduces soil loss and surface runoff of nitrogen, but it is not certain if leaching and contamination of ground water by nitrogen is also reduced [T. J. Logan, J. W. Davidson, J. L. Baker, M. R. Overcash, Eds., *Effects of Conservation Tillage on Groundwater Quality: Nitrates and Pesticides* (Lewis, Chelsea, MI, 1987)].
12. S. Postel, in *The State of the World*, L. Brown, Ed. (World Watch Institute, Washington, DC, 1986), chap. 3.
13. M. R. Gebhardt, T. C. Daniel, E. E. Schweizer, R. R. Allmarus, *Science* 230, 625 (1985).
14. P. F. Hendrix et al., *BioScience* 36, 374 (1987).
15. R. L. Blevins, M. S. Smith, G. W. Thomas, in *No-Tillage Agriculture: Principles and Practices*, R. E. Phillips and S. H. Phillips, Eds. (Van Nostrand Reinhold, New York, 1984); J. W. Doran, *Soil Sci. Soc. Am. J.* 44, 765 (1980).
16. G. J. House, B. R. Stinner, D. A. Crossley, E. P. Odum, *J. Appl. Ecol.* 21, 991 (1984); B. R. Stinner, D. A. Crossley, E. P. Odum, R. L. Todd, *Ecology* 65, 354 (1984).
17. E. P. Odum and M. G. Turner, *The Georgia Landscape: A Changing Resource* (Institute of Ecology, University of Georgia, Athens, 1988).
18. ———, in *Trends in Landscape Ecology*, I. S. Zonneveld and R. T. T. Forman, Eds. (Springer-Verlag, New York, in press).
19. W. C. Johnson and D. M. Sharpe, *For. Sci.* 22, 307 (1976).
20. M. G. Turner and C. L. Ruscher, *Landscape Ecol.* 1, 241 (1988).
21. M. G. Turner, *Environ. Manage.* 11, 237 (1987).
22. G. W. Johnson, *The Wasted Land* (Univ. of North Carolina Press, Chapel Hill, 1938).
23. J. S. Olson, in *Productivity of World Ecosystems* (National Academy of Sciences, Washington, DC, 1975); R. H. Waring and W. H. Schlesinger, *Forest Ecosystems, Concepts and Management* (Academic Press, New York, 1985).
24. See reviews by H. T. Odum, H. E. Daly, R. Costanza, and others [in *Integration of Economy and Ecology*, A. M. Jansson, Ed. (Univ. of Stockholm Press, Stockholm, Sweden, 1984)].
25. M. G. Turner, E. P. Odum, R. Costanza, T. P. Springer, *Environ. Manage.* 12, 237 (1988).
26. K. Parker, E. P. Odum, J. L. Cooley, *Natural Ecosystems of the Sunbelt* (Institute of Ecology, University of Georgia, Athens, 1981).
27. D. F. Spencer, S. B. Alpert, H. H. Gilman, *Science* 232, 609 (1986).
28. E. P. Odum, *BioScience* 34, 558 (1984).
29. H. W. Odum, *Southern Regions of the United States* (Univ. of North Carolina Press, Chapel Hill, 1936).
30. E. P. Odum, *J. Soil Water Conserv.* 42, 412 (1987).
31. I am indebted to M. G. Turner, a research associate on the Georgia Project, which was supported by the W. K. Kellogg Foundation. Other members of the task force were E. O. Box, J. L. Cooley, F. B. Golley, J. E. Kundell, J. A. Little, V. Meentemeyer, D. Morrison, H. R. Pulliam, and R. Westmacot. I also acknowledge work of D. A. Crossley, D. C. Colman, P. F. Hendrix, B. R. Stinner, G. J. House, and P. M. Groffman of the agroecosystem team, supported by grants from the National Science Foundation.

# Ergodic Theory, Randomness, and “Chaos”

D. S. ORNSTEIN

**Ergodic theory is the theory of the long-term statistical behavior of dynamical systems. The baker’s transformation is an object of ergodic theory that provides a paradigm for the possibility of deterministic chaos. It can now be shown that this connection is more than an analogy and that at some level of abstraction a large number of systems governed by Newton’s laws are the same as the baker’s transformation. Going to this level of abstraction helps to organize the possible kinds of random behavior. The theory also gives new concrete results. For example, one can show that the same process could be produced by a mechanism governed by Newton’s laws or by a mechanism governed by coin tossing. It also gives a statistical analog of structural stability.**

**E**RGODIC THEORY AROSE OUT OF AN ATTEMPT TO UNDERSTAND the long-term statistical or probabilistic behavior of dynamical systems such as the motions of a billiard ball or the motions of the earth’s atmosphere. The theory focused on certain mathematical objects called abstract dynamical systems or measure-preserving flows. The idea here is to abstract out the statistical properties and ignore other properties of the dynamical system. Thus two systems are considered the same when viewed as an abstract system (we call these isomorphic) if, after we ignore sets (or events) of probability zero, there is a one-to-one correspondence between the points in their phase spaces (see below) so that corresponding sets have the same probability and evolve in the same

way (in other words, maintain the correspondences for all time) (1). If we preserved the topology instead of probabilities, we would be studying the qualitative theory of ordinary differential equations initiated by Poincaré.

Abstract dynamical systems are natural objects from the mathematical point of view, and they arise in many different contexts (even in areas as far afield as number theory); elucidating their structure is considered an important mathematical problem. Much of the work in ergodic theory has had little to do with its initial motivation, but recently certain problems, some of which had been unsolved for more than a decade, have been solved, and a group of results has been obtained that does relate to concrete systems such as the billiard system. These results, which I will refer to as isomorphism theory, center around a better understanding of a certain abstract dynamical system called the baker’s transformation. This is a map of the unit square onto itself (Fig. 1). We first stretch the square, doubling its width and halving its height. We next stack the right half of the elongated rectangle above the left half (the shape is again a square). These two steps give the baker’s transformation. If looked at properly the baker’s transformation can serve as the mathematical model for coin tossing (this is easy to see), and in this sense is completely random while it is deterministic in the sense that every point moves in a definite way.

The baker’s transformation is often used as a paradigm for explaining the possibility of deterministic chaos (2, 3), that is, systems that evolve according to Newton’s laws but nevertheless appear to be random. In recent years people have become increasingly aware of the ubiquity of this phenomenon.

The author is professor of mathematics at Stanford University, Stanford, CA 94305.

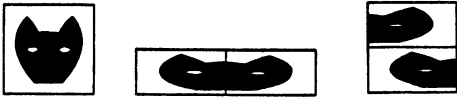


Fig. 1. The baker's transformation.

New results follow from isomorphism theory:

1) At the level of abstraction of isomorphism there is a unique (4, 5) system that is the most random possible. I will refer to it as the Bernoulli flow or  $B_t$ .

Even though "most random" is not well defined, I will explain some precise results that will make the above statement clear. It is important, however, to identify isomorphic systems and, unless we do this, "most random" has no meaning. Because there is no agreed definition of "chaotic" in the literature on chaos, result 1 gives us something definitive to aim at when studying specific systems.

2) If we look at  $B_t$ , in other words, a system isomorphic to  $B_t$ , at regular time intervals (for example, every hour on the hour), then this new system is isomorphic to the baker's transformation.

3) There are systems that obey Newton's laws that are isomorphic to  $B_t$ . A billiard system with one or several convex obstacles is an example (without the obstacle it is not chaotic). Other examples of  $B_t$  are geodesic flow on a manifold of negative curvature and the Lorenz attractor. Dissipative examples of  $B_t$  are provided by Smale's axiom A attractors. Isomorphism theory provides the only method so far for completely determining the statistical structure of specific chaotic systems, in other words, determining them up to isomorphism.

I conjecture that "most" chaotic systems that arise naturally are abstractly the same as  $B_t$ . This would mean that much of the diversity that we see arises from different ways of looking at the same abstract system.

Isomorphism in general can be very hard to visualize—very different systems can be isomorphic—but this is the price we have to pay for the generality of our results. On the other hand, in special cases the theory produces isomorphisms that preserve some of the geometry as well as the statistical structure. I will explain these results in terms of two examples, both of which are isomorphic to  $B_t$ .

The first is the motion of a billiard ball on a rectangular table with a convex obstacle. This example exhibits what people call deterministic chaos. The second example is the random jumping of a point on the same table, the jumps being determined by flipping a biased coin. This is generally thought of as truly "random."

In terms of these examples we have the following results:

4) If we perturb the obstacle slightly in our billiard system, we get essentially the same collection of infinite trajectories. More precisely, the system and its perturbation are isomorphic and corresponding points are close, with high probability. This means that we cannot tell if we are looking at our original system through a "viewer" (6) that distorts slightly or at our perturbed system (through a viewer that does not distort at all), even if we look for all time.

This result is surprising because effects of perturbing the obstacle are cumulative whereas the distortions produced by a viewer are not (the viewer does not interfere with the orbit, it only misreads it slightly).

5) If we design the random system carefully (still using a biased coin), then this random system and the billiard system will have essentially the same trajectories even in the long run (the short-term approximation of trajectories is not hard). One can even reproduce the random system exactly by observing the billiard system through a specially constructed viewer of finite precision that does not distort very much. Note that the viewer, as before, is not random, and thus

our random system could be produced by Newton's laws or by coin flipping.

This article has a dual purpose. I would like to explain some results in ergodic theory by describing what they say in a simple concrete context. I also hope that these results will shed some light on the nature of chaos. For a more complete technical version of this article, see (7), which is based largely on (8).

## The Phase Space of a Dynamical System

Consider the example of a billiard ball moving on a rectangular table with a convex obstacle. The ball moves in a straight line at unit speed and bounces off the obstacle and the edges with the angle of reflection equal to the angle of incidence.

The configuration of the system is completely determined by the position and velocity of the ball. The phase space is the set of all configurations of the system. Each point (configuration) in the phase space moves along a well-defined trajectory or orbit in the phase space.

Phase space also has a probability structure: every reasonable (in other words, measurable) set in phase space is assigned a probability—the probability of finding the system in that set. The probability structure has the property that, as a set evolves in time, its probability does not change.

## Isomorphism and Abstract Dynamical Systems

**Definition:** Two systems are isomorphic if (after we ignore sets of probability zero) there is a one-to-one correspondence between the points of their phase spaces, where corresponding sets have the same probability and orbits correspond to orbits in a time-preserving manner.

Isomorphism means "the same" relative to the structure we are interested in (here, the statistical dynamical structure) and abstracts out everything but this structure. We could think of an abstract system as an equivalence class under isomorphism or, more directly:

**Definition:** An abstract system is a set of points that play the role of the phase space, in which each reasonable (Lebesgue measurable) subset has a probability (that does not change as the subset moves) and a rule that tells us where a point will move in  $t$  units of time (9). Isomorphic systems can be thought of as being the same abstract system (10).

An application of isomorphism theory is that all our billiard systems with obstacles are isomorphic to each other and to  $B_t$  but not to the one with no obstacles.

The classification of chaotic dynamical systems by their abstract system ran into difficulties because the abstract system could not be determined in any specific case. However, isomorphism theory gives criteria for determining if a specific dynamical system is isomorphic to  $B_t$ . This gives, essentially, the only cases of chaotic systems for which the abstract dynamical system associated with a specific system can be determined exactly.

## Stability

Individual orbits of chaotic systems can be extremely unstable and nonreproducible. On the other hand, for many chaotic systems, we can prove that the system as a whole is stable. We will describe our result in the case of billiard balls with convex obstacles.

**Theorem 1.** Pick a small  $1 \gg \alpha > 0$ . Suppose we change the position or curvature of the obstacle by a small amount (how small

depends on  $\alpha$ ). Then there is (after we ignore sets of probability zero) a one-to-one correspondence between points of the phase space so that, with probability  $1 - \alpha$ , corresponding points are closer than  $\alpha$ , corresponding sets have the same probability, and orbits of the original system correspond to orbits of the changed system in a time-preserving manner. [To be completely accurate we may need a small rescaling of time,  $t \rightarrow ct$ ,  $|1 - c| < \alpha$ , in one of the systems (5).]

Let me rephrase this in a more picturesque way. We can think of the above correspondence as being implemented by looking at our system through a viewer. In other words, when we look at the ball through the viewer, we see a slightly distorted position and velocity, the one that corresponds (via the above correspondence) to the position and velocity we are looking at. The viewer thus distorts each orbit by misreading it slightly. Our result says that the perturbed system is identical to the original system seen through our viewer, which distorts only slightly. (We get the same collection of orbits, and all joint probabilities are the same.)

The interest in this result lies in the fact that changes introduced by changing the obstacles are cumulative, their effect builds up as time goes on, whereas the changes introduced by the viewer clearly do not build up. The viewer does not interfere with the orbit, it only misreads it slightly.

Theorem 1 is delicate. For example, no matter how closely we approximate the obstacle by a polygon (a large change in curvature), any orbit of the original system and any orbit of the changed system will be  $D$  apart on average, where  $D$  is the average distance between points in the phase space (even if we rescale time). This means that, if we cannot distinguish points that are too close, we can change the system by an amount that is too small to see but still be able to observe the effect by looking at the system for a long time.

Our results on stability are analogous to the structural stability results of Anosov and Smale. Both kinds of stability hold for smooth perturbations of axiom A systems. (These systems include most of the examples that can be analyzed rigorously.) The main differences are as follows:

- 1) Our correspondence between phase spaces (the one in theorem 1) preserves probability whereas that of Anosov and Smale preserves the topology (homeomorphism). Note that a homeomorphism can take a set of probability 1 to a set of probability zero.
- 2) The structural stability correspondence moves all points by a small amount. Our correspondence moves most (except for a set of small probability) points by only a small amount.
- 3) Our correspondence takes orbits to orbits with a constant rescaling of time ( $t \rightarrow ct$ ,  $c \approx 1$ ). In structural stability the rescaling may differ from point to point (so corresponding sets may not evolve in the same way).

## Random Systems and Their Phase Space

Following Kolmogorov, I will now introduce the analog of phase space for random processes. This will cast random processes and Newtonian systems in the same mathematical framework and allow for rigorous comparison.

Let us do this for coin tossing. We describe the outcome of each coin tossing experiment (the experiment goes on for all time) as a doubly infinite sequence of heads (H's) and tails (T's). The phase space is the collection of these sequences. Each subset is assigned a probability. For example, the set of all sequences that are H at time 3 and T at time 5 gets probability  $1/4$ . The passage of time shifts each sequence to the left (what used to be time 1 is now time 0).

**Definition:** The dynamical system or transformation I have just described is called the Bernoulli shift  $B_{(v_2, v_2)}$  and may be identified

with the baker's transformation. If instead of flipping a coin, we spin a roulette wheel with three slots of probability  $p_1, p_2, p_3$ , we would get the Bernoulli shift  $B_{(p_1, p_2, p_3)}$ .

We identify  $B_{(v_2, v_2)}$  with the baker's transformation as follows. We let H,T correspond to 0,1 and split the two-sided infinite sequences into two one-sided infinite sequences such as ... 1000.1101 ~ 0.0001 ... , 0.1101 ... . These one-sided sequences give the dyadic expansion of the  $x$  and  $y$  coordinates of a point in the square. The reader can check that under the correspondence I have just given between sequences of H,T and points in the square, coin tossing probabilities correspond to ordinary area and that the shift on sequences H,T corresponds to the baker's transformation defined in the introduction.

The random process described in the introduction, called a semi-Markov process, gives rise to a dynamical system in the same way. The process can be defined as follows. We stay at one of a finite number of points  $x_i$  on the billiard table for time  $t_i$  and then jump to one of a pair of points according to a flip of a biased coin (the pair of points depends on  $x_i$ ). It can be shown that this new dynamical system is isomorphic to  $B_t$ .

We can think of the dynamical system that we get from a random process as an abstract model for the minimal mechanism capable of producing that process. If we start with a dynamical system and make a measurement (which can be modeled as a function on the phase space), then we can think of the outcome of the measurement at various times as a sample path of a stationary process.

The phase space model for a random process means, in some sense, that every random process could be produced by a deterministic mechanism. These mechanisms, however, are not governed by Newton's laws. A much harder problem, mathematically, is to show that the same system can arise both from Newton's laws and from processes based on coin flipping.

## Random Versus Deterministic Systems

I will give an example of how the randomness of systems isomorphic to  $B_t$  manifests itself in a more concrete way by comparing a system governed by Newton's laws, our billiard system, with a random system based on coin flipping, our semi-Markov system.

**Theorem 2.** There exists, in theory, a nonrandom viewer that is as reliable as we want, and the viewer has the property that the billiard system with a convex obstacle seen through this viewer is identical (all joint probabilities are the same) to some semi-Markov system. Furthermore, we can still completely reconstruct the orbit we are looking at from the random orbit that we see.

Technically the viewer is a measurable function from the phase space to the finite collection of points on the table. It is reliable in the sense that for most points in the phase space (position and velocity pairs) the position of the ball we are looking at is close to the position of the special point that we see. This viewer models finite precision measurements.

In the billiard system, we can get as much information as we want about the state at time zero by making arbitrarily fine observations at time zero. The analogy of this with the viewer is that we make regular observations of fixed accuracy and we get as much information as we want about the state at time zero by making the sequence of observations go longer into the past and into the future.

Theorem 2 says that the semi-Markov system can be thought of as being produced by Newton's laws. Theorem 2 holds for all systems isomorphic to  $B_t$ . A sense in which  $B_t$  is "the most random" is that theorem 2 only holds for systems isomorphic to  $B_t$ .

A common belief is that the appearance of randomness comes



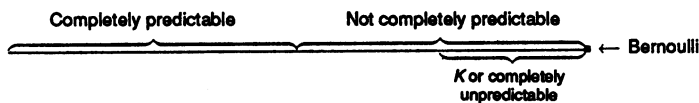


Fig. 2. The kinds of chaotic behavior.

Table 1. Predictability of finite precision observations at discrete time intervals (every hour on the hour).

Class of flow	Holds for	Knowledge of the past tells us
Completely predictable	All observations	Everything about the future
Not completely predictable	Some observations	Not everything about the future
$K$	All observations	Arbitrarily little about fixed time in distant future
Bernoulli	All observations	Vanishingly little about the long-term future

from some microscopic random phenomenon, such as thermal agitation, which is magnified because of sensitivity to initial conditions. Theorem 2 says that we can get random-looking behavior without assuming any underlying randomness. However, we can still ask about the effect of adding some small underlying randomness. It can be shown that in certain cases we cannot distinguish between randomness in the system and randomness in the device through which we view the system. The system that results from adding a small amount of randomness could be reproduced exactly (all joint probabilities are the same) if we were to look at the original system through a viewer that distorts randomly but not very much (with high probability). Such a viewer does not interfere with the process and its effect is not cumulative, whereas the effects of the original perturbation do accumulate, making this result somewhat surprising. [An exact statement and some cases where this holds are given in (7).]

## Bernoulli Flow $B_t$ , the Most Random Flow Possible

**Theorem 3.** There is an abstract system that we call  $B_t$  with the property that if we only look at it on the hour, then it is isomorphic to the Bernoulli shift,  $B_{(1/2, 1/2)}$ . If we were to look at  $B_t$  on the minute instead of on the hour, we would get another Bernoulli shift; in fact, we get all Bernoulli shifts in this way. Any flow for which we can get a Bernoulli shift by discretizing time is isomorphic to  $B_t$  after an appropriate rescaling of time (5). Thus  $B_t$  is “most random” in the sense that independent processes are the most random discrete time processes.

A more definitive justification for calling  $B_t$  “most random” comes from examining the predictability of measurements on a system isomorphic to  $B_t$ . Recall that a measurement gives rise to a stationary process. We will say that measurements (with the same range) on different systems give rise to the same process if all joint probabilities are the same.

Before discussing  $B_t$ , I should point out that there is a class of systems that are the least random—where all measurements (even those of finite precision) are predictable in the sense that, if made at regular intervals of time, the past determines the future. We call these systems completely predictable.

Isomorphism theory shows:

**Theorem 4.** If a system has any observation that is not predictable, then the set of processes arising from observations on the system includes all the processes that can arise from observations on

$B_t$ . Furthermore, if a system is not isomorphic to  $B_t$ , then it gives rise to some measurement that is more predictable (in a sense discussed in the next section) than any measurement on  $B_t$ .

The following theorem, which is related to theorem 4, shows that  $B_t$  is present, in some sense, in all chaotic systems and that it has a unique place among these systems.

**Theorem 5.** Any system that is not completely predictable has  $B_t$  [with a possible rescaling (5, 11) of time] as a factor. Furthermore, the only factors of  $B_t$  are  $B_t$  (with a rescaling of time that slows the speed of  $B_t$ ).

System  $a$  is a factor of system  $b$  if there is a many-to-one correspondence from the phase space of  $b$  to that of  $a$ , where corresponding sets have the same probability and evolve in the same way. In other words, we can make  $b$  isomorphic to  $a$  by lumping points together and treating these lumps as single points. Lumps must, of course, go into lumps under time evolution.

One of the most important features of isomorphism theory is that it allows one to show that specific systems are isomorphic to  $B_t$ . We have already encountered several such examples. A result of Pesin says that, when the phase space is three-dimensional, then essentially the only possibilities are completely predictable or  $B_t$ .

**Theorem 6.** Suppose we have a Hamiltonian system with a three-dimensional phase space, after fixing the energy. The phase space will then break up into possibly one but at most a countable number of invariant sets with the property that on each of these the system is either completely predictable or [after rescaling time (5)] isomorphic to  $B_t$  or isomorphic to a minor modification of  $B_t$  (12).

No one has been able to rigorously analyze the double pendulum (a pendulum with a joint in the middle) but Pesin’s theorem says that it is either the most random possible or completely predictable. Computer studies make it highly unlikely that it is completely predictable.

## The Varieties of Chaotic Behavior

At one extreme we have a single flow  $B_t$ , whereas at the other extreme we have the class of flows that are completely predictable (the billiard system with no obstacles is an example). The flows that are not completely predictable are generally considered chaotic. Not being completely predictable is essentially the same as having sensitivity to initial conditions, positive Lyapunov exponents, or positive entropy.

A flow that is not completely predictable may have some predictable and some nonpredictable measurements. There is a subclass that is especially chaotic, the so-called  $K$  flows, where no measurement is predictable.

It was once hoped that the general picture would be fairly simple, at least in discrete time. It had been conjectured that the only discrete-time  $K$  systems were the Bernoulli shifts and that every discrete system was the product of a  $K$  system and a completely predictable system (the Pinsker conjecture). The latter would mean that every stationary process could be realized as a function of the joint output of a completely predictable system and a  $K$  system (the two systems not interacting).

**Theorem 7.** There exist systems that are not the product of a completely predictable system and a completely nonpredictable system.

**Theorem 8.** There are  $K$  transformations that are not Bernoulli shifts and  $K$  flows that are not  $B_t$  (Fig. 2).

There are many theorems to the effect that the  $K$  class is extremely complicated. In particular, the fact that so many different systems are isomorphic to  $B_t$  is not due to a lack of possible alternatives. These examples also show that there are chaotic processes that cannot be

closely simulated by more general semi-Markov processes (for example, if we jump to more than two points and replace coin flipping with a more complicated independent process). Thus, not all randomness can be understood in terms of coin flipping or independent processes (7).

We can also characterize the different degrees of randomness by the predictability of observation made on the systems (Table 1). I will not state these results precisely but refer the reader to (7) for the precise statement; Table 1 should give the reader a rough idea.

## Dissipative Systems

The main example I used to illustrate our results was billiards, a conservative system. These results, however, apply to a large class of dissipative systems, those that satisfy axiom A. In this situation there is an invariant set in the phase space, the attractor. The attractor has an invariant probability structure and is exactly the kind of dynamical system I have been describing (it is even isomorphic to  $B_1$ ). The difference is that the probability of our system being on the attractor is zero. However, with probability 1, every trajectory of our system is asymptotic to a generic trajectory on the attractor, and thus the dynamics of the attractor still governs the observed long-term behavior of the system.

I conjecture that the above situation is typical of dissipative systems. In particular, I conjecture that long-term statistical averages exist for dissipative as well as conservative systems.

## Historical Overview of Isomorphism Theory for Chaotic Systems

Isomorphism theory for chaotic systems was initiated in 1958 when Kolmogorov and Sinai (13) introduced the concept of entropy to ergodic theory and used it to solve a long-standing problem by showing that not all Bernoulli shifts were isomorphic. [They showed that the entropy of  $B(p_1, p_2, p_3)$  was  $p_1 \log p_1 + p_2 \log p_2 + p_3 \log p_3$ , and so forth, and that shifts of different entropy could not be isomorphic (13).] Sinai showed that a Bernoulli shift was a factor of anything not completely predictable (in discrete time) (14).

On the concrete side, Adler and Weiss proved an isomorphism theorem for automorphisms of the 2-torus (15), and Sinai (16) and Anosov (17) showed that a large class of systems (including billiard balls with obstacles) were completely nonpredictable. Anosov also showed that these systems were stable in a topological rather than a statistical sense (structural stability) (17).

The first of the group of results discussed in this article came in 1970 when Ornstein showed that Bernoulli shifts of the same entropy were isomorphic (18) by methods different from those of Sinai and those of Adler and Weiss. By a still different set of ideas Ornstein showed that the completely unpredictable class contained more than the Bernoulli shifts (19) and that not every transformation was the direct product of a completely unpredictable and a completely predictable transformation (20).

With the method used to prove the isomorphism theorem for Bernoulli shifts, Ornstein showed that Bernoulli shifts could arise in the context of real (continuous) time, that there was a unique Bernoulli flow (which strings together all of the Bernoulli shifts), that the only factors of  $B_t$  are  $B_t$ , and that  $B_t$  is a factor of any system that is not completely predictable (21).

The connection with concrete systems was made when Weiss and Ornstein showed that the geodesic flow on a surface of negative curvature is isomorphic to  $B_t$  (22). Since then it has been shown that

a large class of specific flows were isomorphic to  $B_t$ . This was done by using the work of Sinai, Anosov, Pesin, and others (23) to check a criterion (8) that makes the isomorphism theorem work.

In many cases the isomorphisms produced by the theory do not move points very much and we get theorem 1, a statistical version of Anosov's results on structural stability, and theorem 2 (7).

There is a series of results (24), due mainly to Feldman, Rudolph, Thouvenot, Weiss, and Ornstein, that is a continuation of the abstract results mentioned above and that contains some of the deepest results in this subject. Unfortunately I do not have the space to describe them.

There is also an area of finitary codes that are a continuation of the result of Adler and Weiss but are less directly related to the results of this article. The ones most closely related are (25).

---

### REFERENCES AND NOTES

1. I would emphasize that we are discussing statistical properties and ignoring events of probability zero, even in concrete examples. For example, imbedded Smale horseshoes (a topological version of the baker's transformation) typically have probability zero (when a probability structure exists) and so lie outside the scope of ergodic theory.
2. I. Ekeland, *Mathematics and the Unexpected* (Univ. of Chicago Press, Chicago, 1988) (Ekeland calls the baker's transformation Arnold's cat).
3. I. Prigogine, *From Being to Becoming* (Freeman, San Francisco, 1980), appendix A and p. 187.
4. Technically,  $B_t$  and  $B_t$  after a rescaling of time (for example, doubling the speed of the flow) are not isomorphic. To be absolutely precise we need to say that  $B_t$  is unique up to rescaling of time.
5. When I rescale in this article, the speed change will not be different at different points, as it can be in structural stability theory, because the latter change could affect the way that sets evolve and so have a drastic effect on the qualitative behavior of the system.
6. Technically the viewer is just a measurable function; in this case it is the correspondence that gives our isomorphism.
7. D. S. Ornstein and B. Weiss, *Am. Math. Soc. Bull.*, in press.
8. D. S. Ornstein, *Ergodic Theory, Randomness, and Dynamical Systems* (Yale Mathematical Monograph 5, Yale Univ. Press, New Haven, 1974).
9. Abstract systems can retain a lot of structure. For example, every surface of negative curvature carries an abstract system, the horocycle flow, from which the surface can be completely reconstructed.
10. The information that we have lost in this abstraction is contained in a function on the phase space, in the case of the billiard system the function that identifies a point in the abstract phase space with a particular position and velocity.
11. The rescalings of time are those rescalings that make the entropy of  $B_t$  less than or equal to the entropy of the first system.
12. The modification of  $B_t$  is the product of  $B_t$  and a rotation of the circle. Such a system can be thought of as taking a system isomorphic to  $B_t$  together with a watch. A configuration of this system consists of a configuration of the old system and the time of day. Pesin's theorem is more general, holding for any smooth flow on a three-dimensional compact manifold that preserves a smooth measure.
13. Ja. G. Sinai, *Dokl. Akad. Nauk SSSR* **124**, 768 (1959); A. N. Kolmogorov, *ibid.* **119**, 861 (1958); *ibid.* **124**, 754 (1959).
14. Ja. G. Sinai, *ibid.* **147**, 797 (1962) [*Sov. Math. Dokl.* **3**, 1725 (1962)].
15. R. Adler and B. Weiss, *Am. Math. Soc. Mem.* **98** (1970).
16. Ja. G. Sinai, *Dokl. Akad. Nauk SSSR* **136**, 549 (1961); *Usp. Mat. Nauk* **27**, 137 (1972).
17. D. V. Anosov, *Proc. Steklov Inst. No. 90* (1967).
18. D. S. Ornstein, *Adv. Math.* **4**, 337 (1970). The following books are devoted mainly to an exposition of this theorem: P. Shields, *The Theory of Bernoulli Shifts* (Univ. of Chicago Press, Chicago, 1973); M. Smorodinsky, *Ergodic Theory, Entropy* (Springer Lecture Notes 214, Springer, New York, 1970).
19. D. S. Ornstein, *Adv. Math.* **10**, 49 (1973). The  $K$  class is shown to be large in D. S. Ornstein and P. Shields, *ibid.*, p. 63.
20. D. S. Ornstein, *ibid.*, p. 89; *ibid.*, p. 103.
21. ———, *ibid.* **5**, 349 (1970); in *Contributions to Ergodic Theory and Probability* (Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1970), pp. 178–218; *Adv. Math.* **10**, 124 (1973).
22. B. Weiss and D. S. Ornstein, *Isr. J. Math.* **14**, 184 (1973).
23. Billiard balls are treated in G. Gallavotti and D. S. Ornstein, *Commun. Math. Phys.* **38**, 83 (1974). Axiom A and the Lorenz attractor are treated in M. Ratner, *Isr. J. Math.* **17**, 380 (1974), and L. Bunimovic, *Sov. Math. Izv.* **8**, 219 (1974). Pesin's theorem can be found in Ya. Pesin, *Russ. Math. Surv.* **32** (no. 4), 55 (1977).
24. Entropy theory is developed for continuous time (without reducing to discrete time) in J. Feldman, *Isr. J. Math.* **36**, 321 (1980). The finite fiber factors of a Bernoulli shift are classified in D. Rudolph, *J. Anal. Math.* **34**, 36 (1978). A theory parallel to isomorphism theory for nonconstant speed changes is developed in J. Feldman, *Isr. J. Math.* **24**, 16 (1976); D. Ornstein, D. Rudolph, B. Weiss, *Am. Math. Soc. Mem.* **37** (no. 262) (1982); D. Rudolph, *ibid.* **54** (no. 323) (1985). Isomorphism theory for amenable group actions is developed in D. Ornstein and B. Weiss, *J. Anal. Math.* **48**, 1 (1987). Isomorphism theory is developed for

products where one factor is Bernoulli in J. P. Thouvenot, *Isr. J. Math.* **21**, 177 (1975); *ibid.* **27**, 150 (1977).

25. M. Keane and M. Smorodinsky, *Ann. Math.* **109** (no. 2), 397 (1976); R. Adler and B. Marcus, *Am. Math. Soc. Mem.* **20** (no. 219) (1979).

26. Most of this article is based on a joint paper with Weiss (7). I thank I. Vardi for comments and discussions and for extensive help in organizing and revising this manuscript and the manuscript of (7). Supported in part by National Science Foundation grant DMS86-05098.

# Peptide and Protein Synthesis by Segment Synthesis-Condensation

EMIL THOMAS KAISER, HISAKAZU MIHARA, GENEVIEVE A. LAFORET,  
JEFFERY W. KELLY, LEE WALTERS, MARK A. FINDEIS, TOMIKAZU SASAKI

The chemical synthesis of biologically active peptides and polypeptides can be achieved by using a convergent strategy of condensing protected peptide segments to form the desired molecule. An oxime support increases the ease with which intermediate protected peptides can be synthesized and makes this approach useful for the synthesis of peptides in which secondary structural elements have been redesigned. The extension of these methods to large peptides and proteins, for which folding of secondary structures into functional tertiary structures is critical, is discussed. Models of apolipoproteins, the homeo domain from the developmental protein encoded by the *Antennapedia* gene of *Drosophila*, a part of the Cro repressor, and the enzyme ribonuclease T<sub>1</sub> and a structural analog have been synthesized with this method.

IN THIS ARTICLE WE DESCRIBE EFFORTS IN OUR LABORATORY to prepare models for biologically active peptides and polypeptides in which we have focused our attention on secondary structural elements such as  $\alpha$  helices and have, to a first approximation, been able to neglect tertiary structure (1, 2). We then proceed to the question of whether the principles that we have developed for the design of secondary structural units in such molecules can be extended to the replacement of naturally occurring secondary structural elements by redesigned units in proteins where folding to form tertiary structures is crucial.

## How to Construct Model Proteins

One of the most powerful techniques for the construction of proteins has been the cloning of genes followed by the expression of the corresponding naturally occurring proteins in suitable host systems (3). To obtain modified protein structures, the techniques of site-directed or cassette mutagenesis have been used to modify the gene structures. Alternatively, with the considerable progress that has been made in DNA synthesis, both natural and modified

The authors are at the Laboratory of Bioorganic Chemistry and Biochemistry, Rockefeller University, New York, NY 10021. This article is dedicated to the memory of Professor E. Thomas Kaiser.

proteins can be prepared through the expression of corresponding synthetic genes.

In view of the power of the molecular biological methods, it is reasonable to ask whether chemical synthesis of proteins remains a viable alternative. The Merrifield solid-phase procedure (4), in which the carboxyl terminus of the growing peptide chain is covalently anchored to a solid support, revolutionized peptide synthesis. When used in combination with purification techniques such as high-performance liquid chromatography (HPLC), the preparation of peptides 30 or 40 amino acids in length has become almost routine. However, when one proceeds to the stepwise synthesis of longer peptides and of small proteins in the range of 50 to perhaps 150 amino acids in length, the problems that arise in the purification and characterization of materials often become formidable. Although small proteins up to about 150 amino acids in length have been synthesized with the stepwise solid-phase method (5), the purification of such materials has been difficult and the homogeneity of the resultant peptides has been difficult to establish. Furthermore, if a family of mutants or structural analogs of such small proteins are to be prepared by the solid-phase method, most of the molecule must be resynthesized for each mutant, unless the mutation is near the amino terminus.

Nevertheless, the appeal of chemical methods for the preparation of small proteins remains high because such methods have flexibility that would be difficult or impossible to achieve by the molecular biological approaches. In particular, as we have already demonstrated for the opioid peptide  $\beta$ -endorphin, through chemical synthesis it is possible to introduce nonpeptidic structural regions, to replace a right-handed helix by a left-handed helix, and to use unnatural amino acids (6). Also, isotopically labeled amino acids useful for spectroscopic studies can be introduced at specific locations in the peptide or protein molecule by chemical methodology (7).

The classical approach to preparing peptides of 30 to 40 residues in length was to synthesize well-characterized protected peptide segments through solution-phase methods and then to couple the resultant segments in solution to make the desired product. Such a segment-condensation approach has been successfully applied to the synthesis of enzymes (8). Solution syntheses of the peptide segments allowed the purity of the growing peptides to be monitored at each stage of the synthesis, but the synthetic efforts were enormously tedious, limiting the number of peptides that could be constructed in a reasonable length of time. Furthermore, there was no assurance that the assembly of the protected peptide segments would proceed