Unpredictability and Undecidability in Dynamical Systems

Cristopher Moore

Department of Physics, Cornell University, Ithaca, New York 14853

(Received 18 August 1989)

We show that motion with as few as three degrees of freedom (for instance, a particle moving in a three-dimensional potential) can be equivalent to a Turing machine, and so be capable of universal computation. Such systems possess a type of unpredictability qualitatively stronger than that which has been previously discussed in the study of low-dimensional chaos: Even if the initial conditions are known *exactly*, virtually any question about their long-term dynamics is undecidable.

PACS numbers: 05.45.+b, 02.50.+s, 05.40.+j

Traditionally, physicists studied integrable systems where a formula could be found for all time describing a system's future state. When we widened our scope in order to study the so-called "chaotic" systems, we were forced to relax our definition of what constitutes a solution to a problem, since no such formula exists. Instead, we content ourselves with measuring and describing the various statistical properties of a system, its scaling behavior, and so on: We can do this because the individual trajectories are essentially random.

In this paper we introduce a class of dynamical systems, with as few as two or three degrees of freedom, in which even these modest goals are impossible: rather than being merely random, the dynamics is highly *complex*. We will show that these systems have a stronger kind of unpredictability than a typical chaotic system, and that as a consequence almost nothing can be said about their long-term behavior.

Complex behavior of this kind has already been discussed in various distributed, many-degree-of-freedom systems such as cellular automata (CA's),[1] partial differential equations (PDE's),[2] neural networks,[3] and hard-sphere gases;[4] what is new about our examples is that they are embedded in a smooth, finite-dimensional dynamics. In particular, we will show how in principle to construct a three-dimensional potential in which the motion of a single particle can correspond to universal computation (i.e., a Turing machine).

First, we discuss in what sense a canonical chaotic system is actually quite simple. Many physical systems contain within their dynamics a "horseshoe map" of the plane[5] like that shown in Fig. 1, in which some area in phase space is stretched, folded, and mapped back onto itself. These include[6,7] damped pendula, beam dynamics in accelerators, spin-orbit coupling, and, in fact, any chaotic Hamiltonian system or area-preserving map such as the standard map, as well as generic two-dimensional maps such as the Hénon map. In these systems, the horseshoe is generally thought of as the generator of chaos.

FIG. 1. The Smale horseshoe map.

This is because if we label points in the Cantor set with two-sided infinite sequences

$$a=(a_i)=\ldots a_{-2}a_{-1}.a_0a_1a_2\ldots$$

as shown in Fig. 2, the horseshoe is equivalent to the *shift map* σ: $a_i \rightarrow a_{i+1}$. For instance, the fixed point in the lower-left-hand corner has the following sequence $(0)=\ldots000.000\ldots$. [We denote by (w) the periodic continuation of some sequence w.]

This map is then considered chaotic because errors grow exponentially as faraway discrepancies in the point's symbol sequence are shifted towards the origin. This also implies that, to accurately predict the system t steps in the future, we need roughly t digits of the initial

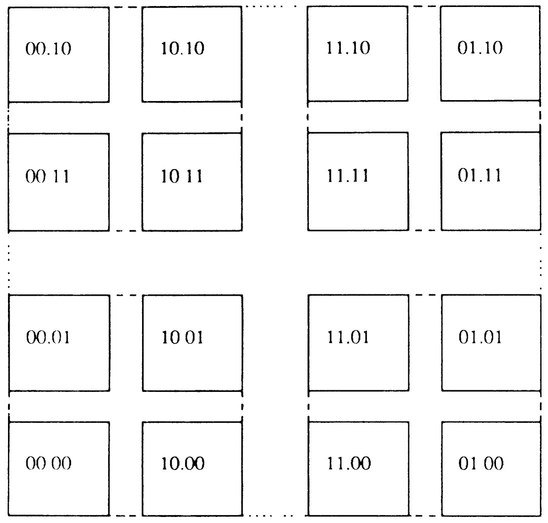

FIG. 2. The invariant Cantor set contained in the horseshoe. Points are given an "address" consisting of a two-sided infinite sequence, in such a way that the map corresponds to shifting the sequence. The left and right halves dictate the x and y coordinates, respectively.

conditions; thus the amount of information needed for prediction grows with time. This is the usual type of "unpredictability" associated with chaotic systems.

However, what we actually do with this information is trivial: If you want the ith digit t steps in the future, just look up the $(i+t)$th digit. In addition, long-term dynamical questions about the system are easy to answer: For instance, the periodic points can be enumerated exactly, since they correspond simply to periodic sequences (0), (1), (01), (10), etc. This simple symbolic description makes it possible to calculate escape rates, Lyapunov exponents, etc., with exponential convergence.[8,9] Thus we can say that this system is *chaotic*, but not *complex*.

How can we invent a more complex action on sequences while retaining the smoothness of the dynamics in the plane? Consider[10] the following *generalized shift* (GS) map:

$$\Phi\colon\ a \rightarrow \sigma^{F(a)}(a \oplus G(a))\,.$$

Here F is a map from a to the integers, and G is a map from a to finite sequences. The notation reads: First replace a finite number of cells in a with the sequence $G(a)$. Then shift the sequence left or right by the amount $F(a)$. Furthermore, we require that F and G depend on a finite number of cells in a; we will call this area of a the *domain of dependence* (DOD). We will take examples with a small DOD, for instance the three cells centered around the origin; unless the DOD is kept as small as possible, things get out of hand.

For example, consider the following GS, which we will call Φ:

$a_{-1}.a_0a_{+1}$	F	G
0.00	-1	0.11
0.01	$+1$	1.01
0.10	$+1$	1.11
0.11	-1	0.00
1.00	$+1$	0.01
1.01	-1	0.10
1.10	$+1$	0.11
1.11	-1	0.01

To illustrate the dynamics, we show that the sequence $(0)1.1(0)=\ldots0001.1000\ldots$ is fixed. The value on the DOD is 1.10, so the table tells us to change it to $G=0.11$, and then shift left since $F=+1$; but then we have the original sequence again.

Then it turns out[10] that *any generalized shift is equivalent to a piecewise linear map of the plane*, in which shifting left and right correspond to stretching the y and z directions, respectively. This particular one is equivalent to the map shown in Fig. 3, and is 1-1 and onto: The eight blocks correspond to the $2^3=8$ possible states of the DOD. To construct the map, we simply use the same labeling of points with sequences that we did in the horseshoe. (In this figure we have closed the gaps of the Cantor set, and just shown the blocks—the discontinuities between blocks can be smoothed in a C^∞ way.[10])

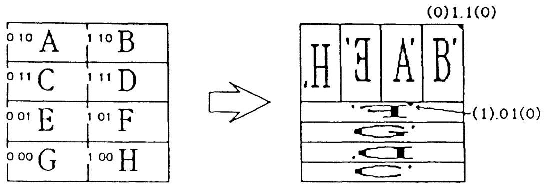

FIG. 3. The map of the plane equivalent to the generalized shift example given in the text. The two fixed points are shown. The map is constructed as follows: Consider the block A, where the DOD is 0.10. The function G tells us to replace this with 1.11; each cell replaced corresponds to a reflection, and we end up upside down on D. Then since $F=+1$, we apply the horseshoe map once to shift left (if $F=-1$, we use the inverse) and arrive at the image A' shown.

We pause here to note the difference between these maps and CA's. In a CA, a rule is applied everywhere at once; in a GS, the action is purely local, and outside the DOD is just a shift. This turns out to be essential to the smoothness of the dynamics, as it corresponds to differentiability; as maps of the Cantor set, CA's are continuous but nowhere differentiable.

In Fig. 4, we simulate this map on a random initial condition. The evolution of the sequence is shown as successive rows of a space-time diagram, with black and white pixels for 0 and 1, respectively. Instead of shifting the sequence left or right by F, for purposes of diagramming we shift the DOD by $-F$; in other words, instead of shifting the entire sequence we shift the "decimal point" the other way, leaving most of the sequence unchanged. Thus the area outside the DOD is stationary and appears as vertical stripes.

These simulations show highly complex behavior, including intermittent periodicity in the neighborhood of various periodic points. These show up in the diagrams as propagating structures reminiscent of "gliders" in CA's:[11] For instance, the fixed point discussed above appears to move to the right each step, since the sequence around it is shifted to the left. The spectrum of periodic points is very irregular; the first few are

x	T	S_T
(0)1.1(0)	1 (fixed)	$+1$
(1).01(0)	1 (fixed)	-1
(0)10.101(0)	7	-1
(0)10.100(1)	7	-1
(10)0.1000(0)	16	2
$(10^n)0.10^{n+2}(10^n)$	$15+n$	$n+1$ (for all $n \geq 1$)

where T is the period and S_T is the total amount of shifting left or right during the course of the orbit. Some of these are visible in Fig. 4.

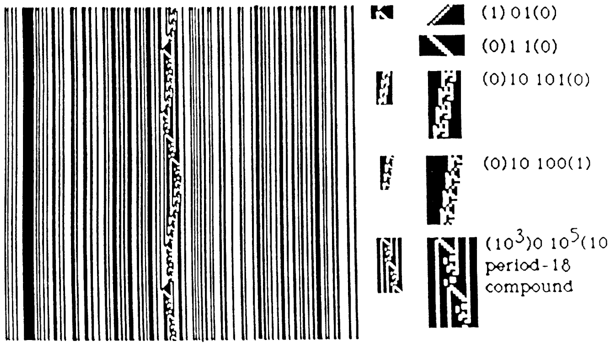

FIG. 4. Simulations of the example. The successive sequences are plotted as rows, with time going downward. Black and white correspond to 0 and 1, respectively. Several periodic points are "magnified" and shown on the side.

When we simulate the equivalent map of the plane, shown in Fig. 3, orbits appear to fill up the entire square; thus this particular example appears to be ergodic. However, the divergence of close initial conditions is slower than exponential. This corresponds to the irregular motion left and right of the DOD, as shown in the simulation; since shifting corresponds to doubling one direction or the other, the distance between nearby points grows as $d \approx 2^{S_t}$, where

$$S_t = \sum_{i=0}^{t} F(\Phi^i(x))$$

is the total amount of shifting after t steps. Since S_t grows slower than linearly with t, the map is not hyperbolic and the divergence is subexponential. In this example the behavior turns out numerically to be diffusive ($S_t \propto t^{1/2}$); examples exist[10] where S_t grows logarithmically with t, maing d grow as a power law. In general, properties like these are highly irregular and difficult to classify, as we will now see.

How can such maps be understood? To address this, we connect them with *Turing machines*.[3] These are idealized computers; they consist of a box with a finite number of internal states, and an infinite "tape" on which sequences of symbols may be written. The box is allowed to read only the one symbol at its present position on the tape, and on the basis of that and its internal state, it may (1) change its internal state, (2) change the tape symbol there, and (3) move one space left or right.

Now it turns out that Turing machines are *computationally universal* in the following sense: For any program written in Pascal, C++, or whatever, there is a Turing machine which will perform it using the tape as its registers and memory. Thus Turing machines can perform any finite computation: produce the digits of π on the tape, test a number written on the tape for primality, etc.

It is customary for the Turing machine to have some special states, designated as "halt" states, to announce that the computation is complete. Then a basic question a dynamicist might ask is, will a given Turing machine, with a given initial state, ever halt? This is obviously a hard question: For instance, we could construct a Turing machine which searches for counterexamples to Fermat's last theorem, and halts only if it finds one. Then to answer this so-called *halting problem* one would have to prove or disprove Fermat's last theorem.

In fact, Turing proved that this question is *undecidable*; i.e., there is no general algorithm to answer it. This can be generalized to other questions: Let H be the set of sequences on which the Turing machine will eventually halt. Then Rice's theorem[12] shows that virtually any question about H is undecidable, such as whether H is nonzero, infinite, dense, etc. Considered as a dynamical system, these are all questions about a basin of attraction: Thus, the long-time behavior of a Turing machine is completely unpredictable.

Note that this is a completely different kind of unpredictability from the chaotic behavior of the horseshoe. The horseshoe is unpredictable because our initial description will have small errors, and these errors grow until our prediction is completely off. Turing machines are unpredictable *even if the initial conditions are known exactly.*

Now we show the correspondence between Turing machines and GS's. We use a simple coding trick to absorb the "box" directly onto the tape, as follows: If s is the internal state and (t_i) the tape, write

$$(s, \ldots, t_{-2}t_{-1}t_0t_1, \ldots) \rightarrow \ldots t_{-2}t_{-1}.st_0t_1 \ldots$$

to obtain a single sequence. Then the action of the Turing machine is easily reproduced by a GS with F and G reading s and t_0, modifying them appropriately, and then shifting left or right.

How does the undecidability of the halting problem carry over? Being in a halt state corresponds to specifying one or more cells of the sequence, which in turn corresponds to being in a particular block A of the Cantor set; so asking if the Turing machine ever halts for a given initial state becomes asking if a given initial point x will ever fall into A. But we have the result that this question is undecidable, even if x is known exactly (for instance, if its coordinates in the Cantor set are rational). This means that the basin of attraction of A is not a *recursive* set, i.e., that there is no algorithm to test whether or not a point is in it. Using various coding tricks and Rice's theorem,[10] we can extend this property to other dynamically generated sets, for instance, the set of periodic points: This, too, is nonrecursive; i.e., it is undecidable whether a given point x is periodic.

This undecidability gives these sets a much more complex kind of structure than a fractal or multifractal; at every time scale, and correspondingly at every scale of magnification, qualitatively new behavior shows up. No nice scaling behavior ever takes over, unlike the situation previously discussed with regard to basins of attraction.[13] This corresponds to the fact that there is no way to

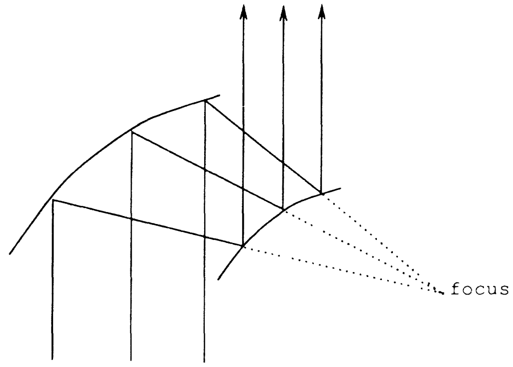

FIG. 5. Expanding or contracting trajectories in a billiard. The two parabolic arcs share the same focus.

"jump ahead" as there is with the horseshoe; if after 10^8 steps x has not yet fallen into A, it still might. The best one can do is to simulate the system and see what happens. This is the "computational irreducibility" referred to by Wolfram.[1]

Thus even the simplest long-term properties of the motion are undecidable: For instance, we can calculate the average amount of shifting S_t for any finite t, but its long-time behavior is uncomputable. This means that even the question of whether or not the map is chaotic (i.e., $S_t \rightarrow \infty$ as $t \rightarrow \infty$) is undecidable.

Can this type of dynamics be found in physical systems? In fact, it can, even in the motion of a single particle in a three-dimensional potential. The easiest way to do this is as a billiard. Let particles pass upward at unit speed through a plane surface of section, defining a point in the Cantor set. Then it is a simple matter to reflect the particle around so it approaches again from below; to carry out expansion or contraction along a given axis, we employ parabolic "mirrors" as shown in Fig. 5. We emphasize that only a finite number of such mirrors are needed—one set for each block of the Cantor set. Thus a billiard can be devised with any generalized shift map as its Poincaré map. We can then "soften" these billiard walls to make a smooth potential for the particle to move in.

These systems give us a "toy model" in which complexity in finite-dimensional systems can be discussed in a precise manner. Although the systems shown here are somewhat contrived, it is plausible that a similar level of complexity obtains in, say, the three-body problem. (It seems that with only one degree of freedom much less complexity is possible.[14])

In conclusion, we have exhibited a kind of motion that is unpredictable in a qualitatively stronger way than what is usually referred to as "chaos," even though the divergence of nearby initial conditions is typically less than exponential. Virtually any question about its long-term behavior is undecidable. Long-term average quantities like escape rates, Lyapunov exponents, or the measure of a basin of attraction are impossible to compute. No nice scaling behavior, in time or phase space, ever takes over. Spectra of periodic points and basins of attraction are completely irregular. Even the question of whether or not it is chaotic is impossible to answer.

Thanks to Philip Homes, Richard Durrett, Juris Hartmanis, Predrag Cvitanović, Charles Bennett, Robert MacKay, and Jim Sethna for conversations, ideas, and support. I also thank Mats Nordahl and David Umberger for careful readings of (and comments on) the manuscript. This work was supported by the NSF. I thank the Niels Bohr Institute-NORDITA in Copenhagen and the Nonlinear Systems Laboratory of the University of Warwick for their hospitality.

[1] S. Wolfram, Commun. Math. Phys. **96**, 15 (1984); Phys. Rev. Lett. **54**, 735 (1984).

[2] S. Omohundro, Physica (Amsterdam) **10D**, 128–134 (1984).

[3] Marvin Minsky, *Computation: Finite and Infinite Machines* (Prentice-Hall, Englewood Cliffs, NJ, 1967).

[4] E. Fredkin and T. Toffoli, Int. J. Theor. Phys. **21**, 219–253 (1982).

[5] S. Smale, in *Differential and Combinatorial Topology,* edited by S. S. Cairns (Princeton Univ. Press, Princeton, NJ, 1963), pp. 63–80.

[6] John Guckenheimer and Philip Holmes, *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields* (Springer-Verlag, Berlin, 1983).

[7] *Hamiltonian Dynamical Systems,* edited by R. S. MacKay and J. D. Meiss (Hilger, London, 1987).

[8] P. Cvitanović, Phys. Rev. Lett. **61**, 2729 (1988).

[9] R. Artuso, E. Aurell, and P. Cvitanović (to be published).

[10] C. Moore (to be published).

[11] Martin Gardner, *Wheels, Life, and Other Mathematical Amusements* (Freeman, San Francisco, 1983).

[12] H. Rogers, Jr., *Theory of Recursive Functions and Effective Computability* (McGraw-Hill, New York, 1967).

[13] C. Crebogi, S. W. McDonald, E. Ott, and J. A. Yorke, Phys. Lett. **99A**, 415 (1983).

[14] C. Moore (to be published).