CHAOS FOR CYCLISTS

PREDRAG CVITANOVIĆ [1] [2]
Niels Bohr Institute,
Blegdamsvej 17, DK-2100 Copenhagen Ø

1 INTRODUCTION

Nonlinear physics presents us with a perplexing variety of complicated fractal objects and strange sets. Notable examples include strange attractors for chaotic dynamical systems, regions of high vorticity in fully developed turbulence and fractal growth processes. By now most of us appreciate the fact that the phase space of a such dynamical system is an infinitely interwoven mixture of islands of stability and regions of chaos. Confronted today with a potentially turbulent dynamical system, we analyze it in three distinct steps. First, we diagnose the intrinsic *dimension* of the system - the lower bound on the number of degrees of freedom needed to capture the essential dynamics - by some numerical dimension algorithm[1]. If the system is very turbulent, its intrinsic dimension is high, and we are, at present, out of luck. So far we have a handle only on the transitional regime between regular motions and weak turbulence. In this regime attractors are of low dimension, the number of important parameters is small, and we can proceed to the second step: we classify qualitatively all motions of the system into a hierarchy whose successive layers require increased precision and patience on part of the observer. This classification is called the *symbolic dynamics* of the system: it is a road map which describes its topology. Parenthetically, through this enumeration number theory enters and comes to play a central and often highly non-trivial role in the study of deterministic chaos.

Having determined what the pieces of the system are, we proceed onto the third, quantitative step: the *scaling*, or metric structure of the dynamical system. Possible trajectories are qualitatively of three distinct types: they are either asymptotically unstable (positive Lyapunov exponent), asymptotically marginal (vanishing Lyapunov) or asymptotically stable (negative Lyapunov). The asymptotically stable orbits can be treated by the traditional integrable system methods. The asymptotically marginal orbits can be called the "border of order"; they remain marginal under any rescaling, and they are a domain of the renormalization theory, a topic beyond the scope of this lecture. Here we shall concentrate on the third class of orbits, the asymptotically unstable orbits which build up chaos. The word "chaos" has in this context taken on a narrow technical meaning. When a physicist says that a certain system exhibits "chaos", he means that the system obeys deterministic laws of evolution, but that the outcome is highly sensitive to small uncertainties in the specification of the initial state. In a chaotic system any open ball of initial conditions, no matter how small, will in finite time spread over the extent of the entire asymptotically accessible phase space. Once this is grasped, the focus of theory shifts from attempting precise prediction (which is impossible) to description of the geometry of the space of possible outcomes, and evaluation of averages over this space.

[1] Carlsberg Fellow

[2] to appear in *Noise and Chaos in Nonlinear Dynamical Systems*, F. Moss ed. (Cambridge Univ. Press, Cambridge 1988)

A precise *quantitative* encoding of the metric structure is given by the *scaling functions* (or transfer operators). Their purpose is twofold:

For an experimentalist, they are the theorist's prediction of the motions expected in a given parameter and phase-space range. Given the observed motions, the symbolic dynamics predicts what motions should be seen next, and the scaling functions predict where they should be seen, and what precision is needed for their observation.

For a theorist, the scaling functions are a tool which resolves the fine asymptotic structure of a chaotic dynamical system and proves that the system is indeed chaotic, and not just a regular motion of period exceeding the endurance of an experimentalist. Furthermore - and that is theoretically very sweet - the scalings often tend to *universal* limits. In such cases, the finer the scale, the better the theorists's prediction! So what at first sight appears to be a bewilderingly complex dynamics might turn out to be a manifestation of a rather simple law, and common to many apparently unrelated phenomena.

In retrospect many triumphs of both classical and quantum physics seem a stroke of luck: a few integrable problems, such as the harmonic oscillator and the Kepler problem, though "nongeneric", have gotten us very far. The success has lulled us into a habit of expecting simple solutions to simple equations - an expectation shattered for many by the recently acquired ability to numerically scan the phase space of non-integrable dynamical systems. The initial impression might be that all our analytic tools have failed us, and that the chaotic systems are amenable only to numerical statistical investigations. However, as we will attempt to show here, we maybe already possess a perturbation theory of the deterministic chaos of predictive quality comparable to that of the traditional perturbation expansions for nearly integrable systems. This theory is based on the observation that the motion in dynamical systems of few degrees of freedom is often organized around a few *fundamental* cycles. These short cycles capture the skeletal topology of the motion in the sense that any long orbit can approximately be pieced together from the fundamental cycles. Computations with such systems require techniques reminiscent of statistical mechanics; however, the actual calculations are crisply deterministic throughout. The strategy will be to expand averages over chaotic phase space regions in terms of short unstable periodic orbits, with the small expansion parameter being the non-uniformity of the flow (here referred to as "curvature") across neighborhoods of periodic points. To get some feeling for how and why unstable cycles come about, we start by playing a game of pinball.

2 PINBALL CHAOS

A physicist's pinball[2,3,4,5] game is a pinball reduced to its bare essentials: three disks in a plane:

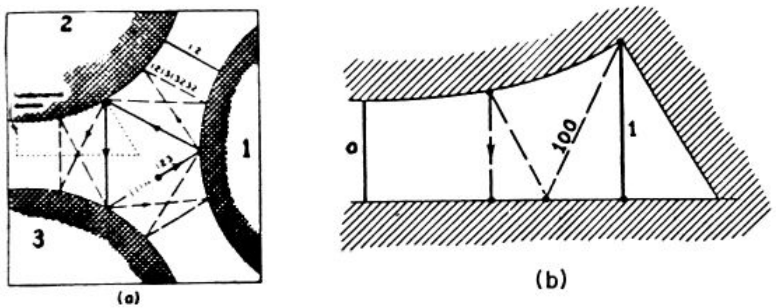

Figure 1. The 3-disk pinball with the disk radius : center separation ratio a:R = 1:2.5. (a) the three disks, with $\overline{12}$, $\overline{123}$ and $\overline{121232313}$ cycles indicated. (b) the fundamental domain, *i.e.* a wedge consisting of a section of a disk, two segments of symmetry axes acting as straight mirror walls, and an escape gap. The above cycles restricted to the fundamental domain are now the two fixed points $\overline{0}$ and $\overline{1}$ and the $\overline{100}$ cycle.

Point-like pinballs are shot at the disks from random starting positions and angles; they spend some time bouncing between the disks and then escape. For a physicist

a good game consists in predicting the asymptotic lifetime (or the escape rate) of a pinball to many significant digits. The unstable cycles as a skeleton of chaos are almost visible here: a good strategy for keeping the ball bouncing as long as possible is to aim it as close as possible to a periodic trajectory. Short periodic trajectories are easily drawn and enumerated - some examples are given in fig. 1 - but it is rather hard to extract the systematics of the orbits from their physical space trajectories.

A clearer picture of the dynamics is obtained by constructing a phase space Poincaré section. We start by exploiting the sixfold symmetry of the disks, and restrict the pinball to bouncing in the fundamental domain (fig. 1b). The whole pinball plane can be retiled by 6 copies of the fundamental domain, but the details are inessential for present considerations. We define our Poincaré section by marking x_i, the position of the ith bounce off the bottom wall (fig. 1b), and $\sin\phi_i$, the momentum component parallel to the bottom wall. $(x, \sin\phi)$ coordinates are a convenient choice, because they are phase-space volume preserving[6].

We next mark the initial conditions $\Omega_{.\epsilon_1}$ which do not escape in one bounce. There are two strips of survivors, as the trajectories originating from one disk can hit either of the other two disks, or escape. We label the two strips with $\epsilon = 0, 1$. There are four strips $\Omega_{.\epsilon_1\epsilon_2}$ that survive four bounces, and so forth. Another way to look at the survivors after two bounces is to plot $\Omega_{\epsilon_1.\epsilon_2}$, the intersection of $\Omega_{.\epsilon_2}$ with the strips $\Omega_{\epsilon_1.}$ obtained by time reversal ($\sin\phi \rightarrow -\sin\phi$). Provided that the disks are sufficiently well separated, what emerges is a complete binary Cantor set with the usual Smale horseshoe[8] foliation and symbolic dynamics.

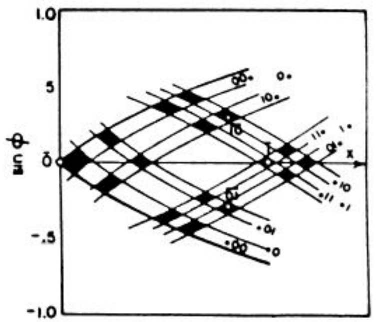

Figure 2. The Poincaré section of the phase space for the fundamental domain pinball, fig. 1b. Indicated are the fixed points $\overline{0}$, $\overline{1}$ and the 2-cycle $\overline{01}$,$\overline{10}$, together with strips which survive 1, 2, ... bounces. Iteration corresponds to bit shift, so for example region ...01.01... maps into ...010.1....

After n iterations the survivors are divided into 2^n distinct neighborhoods: the ith neighborhood consists of all points with itinerary $i = \epsilon_1\epsilon_2\epsilon_3\ldots\epsilon_n$, $\epsilon_i = \{0, 1\}$. Each such patch contains a periodic point $\overline{\epsilon_1\epsilon_2\epsilon_3\ldots\epsilon_n}$ with the basic block infinitely repeated. Periodic points are skeletal in the sense that as we look further and further, they stay put forever, while the finite covers shrink onto them. The periodic points are dense on the asymptotic repeller, and their number increases exponentially with cycle length. As we shall see, this exponential proliferation of cycles is not as dangerous as it might seem.

Before continuing with the pinball as an illustration of the origin and structure of unstable cycles, we turn briefly to the role of cycles in more general settings.

3 CYCLES AS THE SKELETON OF CHAOS

Consider a general three-dimensional flow sketched in fig. 3. To be interesting, the flow should be recurrent; otherwise it is a transient state that we cannot observe for long times. If the flow is recurrent, we can cut it by a Poincaré section; if it is a map of a compact disk domain onto itself, it must have at least one fixed point. Now consider the ways in which the flow can deform the neighborhood of a fixed point. There are essentially two possibilities: the neighborhood can return wrapped around the fixed point (the fixed point is stable or elliptic - see fig. 3a), or squeezed, stretched and folded (the fixed point is unstable or hyperbolic - see fig. 3b).

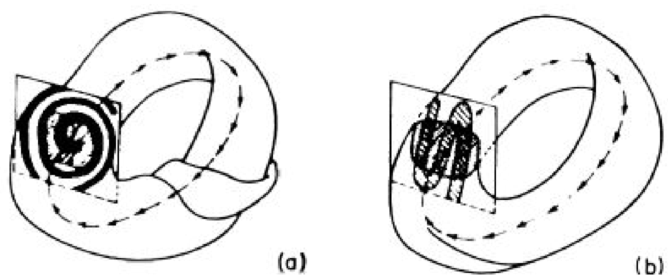

Figure 3. A recurrent flow around (a) an elliptic fixed point, (b) a hyperbolic fixed point.

In the traditional approach we use integrable motions of fig. 3a as zeroth-order approximations to physical systems, and account for weak non-linearities perturbatively. We tend to think of a dynamical system as a smooth system whose evolution can be followed by integrating a set of differential equations. When this is actually followed through to asymptotic times, we discover that the strongly non linear systems show an amazingly rich structure which is not at all apparent in their formulation in terms of differential equations. However, hidden in this apparent chaos is a rigid skeleton, a tree of *cycles* (periodic orbits) of increasing lengths and self-similar structure. The new insight is that the zeroth-order approximations to harshly chaotic dynamics should be very different from those for the nearly integrable systems: a good starting approximation here is the stretching and kneading of a baker's map of fig. 3b, rather than the winding of a harmonic oscillator of fig. 3a.

There are a number of deep reasons[9] for the special role the cycles play in the theory of dynamical systems; we shall motivate them here by appealing to the concept of *invariance*. A deterministic dynamical system can be presented in an infinite number of ways. The theorist's choice of variables and parameters is a matter of aesthetics; the experimentalist's is constrained by instrumental limitations, and so the same physics is presented to us in many obscurely related guises as a phase-space trajectory, as a time-delay plot, as a map such as a Poincaré section of a flow[12]. How are we to efficiently identify the system under its different guises? What is clearly needed is a *complete invariant characterization*, something like specifying a representation of a Lie group by the values of its Casimir operators. Given such characterization, one could extract the invariants from both theory and the experiment:

	EXPERIMENT	THEORY
c_1	1.763...	1.7512638...
c_2	0.32...	0.253437...
c_3	0.02...	0.019253...
...	...	...

and quantify the closeness of the two by computing something like a "distance" in the space of dynamical systems

$$\lim_{n\to\infty} \sum_{i=1}^{n} (c_i^{exp} - c_i^{the})^2$$

We shall argue here that for low dimensional deterministic dynamical systems the cycles (periodic orbits) provide a possibly optimal[13] invariant description of a dynamical system, with the following virtues:

1. cycle symbol sequences are *topological* invariants: they give the spatial layout of a strange set

2. cycle eigenvalues are *metric* invariants: they give the scale of each piece of a strange set

3. cycles are *dense* on the asymptotic non-wandering set

4. cycles are ordered *hierarchically*: short cycles give good approximations to a strange set, longer cycles only refinements. Errors due to neglecting long cycles can be bounded, and typically fall off exponentially with the cutoff cycle length

5. cycles are *structurally robust*: eigenvalues of short cycles vary slowly with smooth parameter changes

6. asymptotic averages (such as generalized dimensions, escape rates, quantum mechanical eigenstates and other "thermodynamic" averages) can be computed from short cycles by means of *cycle expansions*

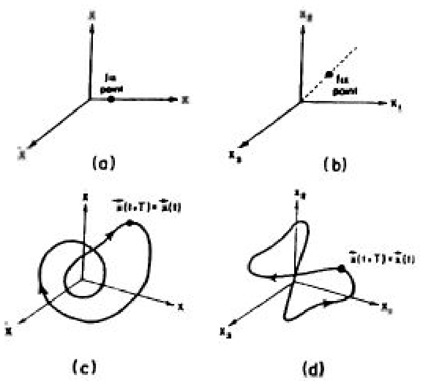

Figure 4. A fixed point and a cycle remain a fixed point and a cycle in any representation of a dynamical systems. Here (a) and (c) phase space is built

from x and its derivatives. (b) and (d) might be the same trajectories in the time-delay coordinates.

Points 1, 2: That the cycle topology and eigenvalues are invariant properties of dynamical systems follows from elementary considerations. If the same dynamics is given by a map f in one set of coordinates, and a map g in the next, then f and g (or any other good representation) are related by a reparametrization and a coordinate transformation $f = h^{-1} \circ g \circ h$. As both f and g are arbitrary representations of the dynamical system, the explicit form of the conjugacy h is of no interest, only the properties invariant under any transformation h are of general import. The most obvious invariant properties are topological; a fixed point must be a fixed point in any representation, a trajectory which exactly returns to the initial point (a cycle) must do so in any representation (fig. 4). Furthermore, a good representation should not mutilate the data; h must be a smooth transformation which maps nearby cycle points of f into nearby cycle points of g. This smoothness guarantees that the cycles are not only topological invariants, but that their linearized neighborhoods are also metrically invariant. In particular, the cycle eigenvalues (eigenvalues of the Jacobians $df^{(n)}(x_k)/dx$ of periodic orbits $f^{(n)}(x_k) = x_k$) are invariant.

Point 3: The cycles are intuitively expected to be dense because on a connected chaotic set a typical trajectory is expected to behave ergodically, and pass infinitely many times arbitrarily close to any point on the set, including the initial point of the trajectory itself. Generically one expects to be able to gently move the initial point in such a way that that the trajectory returns precisely to the initial point:

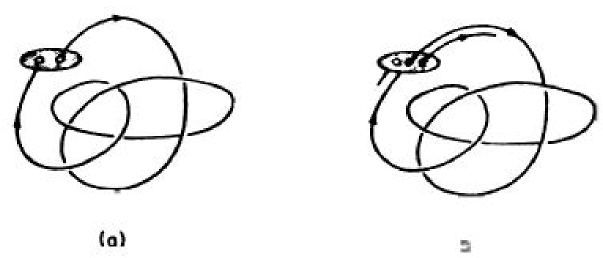

Figure 5. (a) A close recurrence of an unstable trajectory (b) can be exploited to locate a nearby cycle.

This is by no means guaranteed to work, and it must be checked for the particular system at hand. A variety of ergodic (but insufficiently mixing) counter-examples can be constructed - the most familiar being a quasiperiodic motion on a torus.

Point 5: An important virtue of cycles is their ***structural robustness***. Many quantities customarily associated with dynamical systems depend on the notion of "structural stability"[8], *i.e.* robustness of strange sets to small parameter variations. This is certainly not a property of generic dynamical systems, such as the Hénon map[15] $(x, y) \to (1 - ax^2 + y, bx)$. For example, while numerical studies indicate that for $a = 1.4$, $b = 0.3$ the attractor is "strange", parameter variation as minute as changing a to $a = 1.39945219$ destroys this attractor and replaces is with a stable cycle of length 13. Still, the short unstable cycles of length less than 13 are structurally robust in the sense that they are only slightly distorted by such parameter changes, and averages computed using them as a skeleton are insensitive to small deformations of the strange set. In contrast, lack of structural stability wreaks havoc with quantities such as Lyapunov exponents, for which there is no guarantee that they converge in any finite numerical computation.

The theoretical advance that we will concentrate on here is **point 4:** we now know how to control the errors due to neglecting longer cycles. As we shall explain in sect. 5, even though the number of invariants is infinite (unlike, for example, the number of Casimir invariants for a compact Lie group) the dynamics can be well approximated to any finite accuracy by a small finite set of invariants. The origin of this convergence is geometrical, as we shall now show by returning to our game of pinball.

4 PINBALL ESCAPE RATE

Consider fig. 2 again. In each bounce the initial conditions get thinned out, yielding twice as many thin strips as at the previous bounce. The phase-space volume is preserved by the flow, so they are contracted along the stable eigendirections, and ejected along the unstable eigendirections; the total fraction of survivors after n bounces is proportional to

$$\Gamma_n = \sum_i^{(n)} l_i. \tag{1}$$

where i is a binary label of the ith strip, and l_i is the width of the ith strip. One expects the sum (1) to fall off exponentially with n, and tend to a limit

$$\Gamma_n = e^{-n/T_n} \to e^{-n/T} = e^{-n\gamma}. \tag{2}$$

T is the ***asymptotic lifetime*** of a random initial pinball; $\gamma = 1/T$ is the pinball ***escape rate***. We shall now show that this asymptotic escape rate can be extracted from a highly convergent ***exact*** expansion by reformulating the sum (1) in terms of unstable periodic orbits.

Each neighborhood i in fig. 2 contains a periodic point $\bar{i}$. The finer the intervals, the smaller is the variation in flow across them, and the strip width l_i is well approximated by the contraction around the periodic point, $l_i = a_i/|\Lambda_i|$. Here Λ_i is the expanding eigenvalue of the linearized map evaluated on the periodic point i, and a_i is a prefactor defined by

$$a_i = l_i|\Lambda_i|. \tag{3}$$

Now we make the only approximation in our derivation of the ζ function: for large n the prefactors $a_i \approx O(1)$ are overwhelmed by the exponential growth of Λ_i, so we neglect them. This is called the ***hyperbolicity*** assumption. a_i reflect the particular distribution of incoming pinballs; the asymptotic trajectories are strongly mixed by bouncing chaotically between the disks and we expect them to be insensitive to smooth variations in the initial distribution. If the hyperbolicity assumption is justified, we can replace l_i in (1) by $1/\Lambda_i$ and form a formal sum over all periodic orbits of all lengths:

$$\begin{aligned}\Omega(z) &= \sum_{n=1}^{\infty} z^n \sum_i^{(n)} |\Lambda_i|^{-1} = z/|\Lambda_{\overline{0}}| + z/|\Lambda_{\overline{1}}| + z^2/|\Lambda_{\overline{00}}| + z^2/|\Lambda_{\overline{01}}| \\ &\quad + z^2/|\Lambda_{\overline{10}}| + z^2/|\Lambda_{\overline{11}}| + z^3/|\Lambda_{\overline{000}}| + z^3/|\Lambda_{\overline{001}}| + \dots \end{aligned} \tag{4}$$

As for large n the nth level sum (2) tends to the limit $e^{-n\gamma}$, the escape rate γ is determined by the smallest $z = e^{\gamma}$ for which (4) diverges:

$$\Omega(z) \approx \sum_{n=0}^{\infty} (ze^{-\gamma})^n = \frac{1}{1 - ze^{-\gamma}} \tag{5}$$

This observation is the motivation for introducing the sum (4). Rather than attempting to estimate the escape rate from the $n \to \infty$ limit of preasymptotic sums (1), we shall determine γ from the singularities of (4).

If a trajectory retraces itself r times, its expanding eigenvalue is Λ_p^r, where p is a *prime* cycle. A prime cycle is a single traversal of the orbit; its label is a non-repeating symbol string. There is only one prime cycle for each cyclic permutation class. For example, $p = \overline{0011} = \overline{1001} = \overline{1100} = \overline{0110}$ is prime, but $\overline{0101} = \overline{01}$ is not. The stability of a cycle is (by the chain rule) the same everywhere along the orbit, so each prime cycle of length n_p contributes n_p terms to the sum (4). Hence (4) can be rewritten as

$$\Omega(z) = \sum_p n_p \sum_{r=1}^{\infty} \left(z^{n_p}|\Lambda_p^{-1}|\right)^r = \sum_p \frac{n_p z^{n_p}|\Lambda_p^{-1}|}{1 - z^{n_p}|\Lambda_p^{-1}|},$$

where the index p runs through all distinct *prime* cycles. The $n_p z^{n_p}$ factors in the sum suggest rewriting it as a logarithmic derivative $\Omega(z) = -z\frac{d}{dz}\ln\zeta(z)$. The resulting infinite product

$$1/\zeta(z) = \prod_p (1 - t_p), \quad t_p = z^{n_p}|\Lambda_p^{-1}|, \tag{6}$$

is an example of a dynamical ζ function[16]. The name is motivated by the (purely formal) similarity of the infinite product to the Euler product representation of the Riemann ζ function.

The above derivation of the ζ function formula for the escape rate has one shortcoming; it estimates the fraction of survivors as a function of the number of pinball bounces. However, the physically interesting quantity is the mean lifetime; giving the same weight to all paths with the same number of bounces overestimates the contributions of the long trajectories and underestimates the short trajectories (remember, the flight paths between disks are of different lengths). The correct weight[17,18] is obtained by replacing the discrete "topological" time n_p by the actual cycle period T_p in (6):

$$t_p = z^{T_p}|\Lambda_p^{-1}|. \tag{7}$$

or $e^{T_p \gamma}$

Perhaps more surprisingly, (6) also yields ***quantum*** resonances, with the quantum amplitude associated with a given cycle being essentially the square root of the classical weight (we refer the reader to refs. [3,5] for detailed discussions).

Expression (6) is the main result of this section; the problem of estimating the asymptotic escape rates from finite n sums such as (1) is now reduced to a study of the singularities of the ζ function (6). The escape rate is related by (5) to a divergence of $\Omega(z)$, and $\Omega(z)$ diverges whenever $1/\zeta(z)$ or $\zeta(z)$ has a zero. Glancing back, we see that the derivation is very general, and should work for any average over any strange set which satisfies two conditions: 1. the weight associated with a part of the set is multiplicative along the trajectory; 2. the set is organized in such a way that the nearby points in the symbolic dynamics have nearby weights.

We conclude this section by a few general comments on the relation of the finite sum (1) to the dynamical ζ function (6). Not so long ago most physicists were inclined to believe that given a deterministic rule, a sum like (1) could be evaluated to any desired precision. For short finite times this is indeed true: every interval in (1) can be accurately determined, and there is no need for a fancy theory. However, if a dynamical system is unstable, any uncertainty in the initial conditions grows exponentially and attain the size of the system in a finite time. The difficulty with estimating the $n \to \infty$ limit from (1) is at least twofold:

1. due to the exponential growth in number of intervals, and the exponential decrease in attainable accuracy, the maximal attainable cycle length is of order 5 to 20;

2. the preasymptotic sequence T_n in (2) is not unique, because in general the intervals l_i in the sum (1) should be weighted by the probability distribution of initial pinballs.

In contrast, with ζ function (6) the infinite time behavior of an unstable system is as easy to determine as the short time behavior by direct evaluation of (1). The only critical step in the derivation of the ζ function was the hyperbolicity assumption, *i.e.* assumption of exponential shrinkage for all parts of the strange set. By dropping the prefactors (3), we have given up on any possibility of recovering the precise distribution of starting pinballs coordinates (which should anyhow be impossible due to the exponential growth of errors), but in return gained a very effective description of the asymptotic behavior of the system.

Perhaps it is worth emphasizing that the Euler product formula (6) is an *exact* expression for the asymptotically strange set. Our cycle expansions will be dominated by short cycles, but that does not mean that we are using finite covers to approximate the set: by resummation that led to (6) we have already been lifted to the topologically exact $n \to \infty$ strange set. The approximation will consist in approximating the strange set with "nearby" Cantor sets with a finite number of already asymptotically exact scales. Our experience is that computations with the exact cycles expression (6) are both quicker and of better convergence than computations that extrapolate from finite cover estimates such as (1).

5 CYCLE EXPANSIONS AND CURVATURES

How are formulas such as (6) used? We start by computing the lengths and eigenvalues of shortest cycles. In our pinball example this can be done by elementary geometrical optics; in general potentials or maps this requires some numerical integrations and Newton's method searches for periodic solutions. The result is a table like this (the disk radius : center separation ratio is a:R = 1:6):

p	T_p	Λ_p
o	4	9.89897948557
i	4.26794919243	-11.7714551964
io	8.31652948517	-124.094801992
oio	12.3217466162	-1240.54255704
ioi	12.580807741	1449.54507485
oioo	16.3222764744	-12295.706862
ioii	16.8490718592	-17079.0190089
iooi	16.5852429061	14459.97595
ooioo	20.3223300257	-121733.838705
...		

The next step is the key step in our approach[19]: we observe that the expansion of the Euler product (6)

$$1/\zeta = 1 - \sum_f t_f - \sum_n c_n, \tag{8}$$

allows a regrouping of terms into dominant *fundamental* contributions t_f and decreasing *curvature* corrections c_n. For example, if the strange set is labelled by binary symbol sequences, as is the pinball of fig. 3, then the Euler product (6) is given by

$$\begin{aligned} 1/\zeta &= (1-t_0)(1-t_1)(1-t_{10})(1-t_{100})(1-t_{101})(1-t_{1000}) \\ &\quad (1-t_{1001})(1-t_{1011})(1-t_{10000})(1-t_{10001}) \\ &\quad (1-t_{10010})(1-t_{10011})(1-t_{10101})(1-t_{10111})\cdots \end{aligned} \tag{9}$$

The curvature expansion is obtained by multiplying out the Euler product and grouping together the terms of the same total symbol string length:

$$\begin{aligned} 1/\zeta &= 1 - t_0 - t_1 - [t_{10} - t_1 t_0] - [(t_{100} - t_{10} t_0) - (t_{101} - t_{10} t_1)] \\ &\quad -[(t_{1000} - t_0 t_{100}) + (t_{1110} - t_1 t_{110}) \\ &\quad +(t_{1001} - t_1 t_{001} - t_{101} t_0 + t_{10} t_0 t_1)] - \ldots \end{aligned} \tag{10}$$

The fundamental cycles t_0, t_1 have no shorter approximants; they are the "building blocks" of the dynamics in the sense that all longer orbits can be approximately pieced together from them. We call the sum of all terms of same total length n (grouped in brackets above) the nth curvature correction c_n, for geometrical reasons we shall soon try to explain. It is also often possible to pair off individual longer cycles and their shorter approximants (grouped in parenthesis above). If all orbits are weighted equally ($t_p = z^{n_p}$), such combinations cancel exactly; if orbits of similar symbolic dynamics have similar weights, the weights in such combinations will almost cancel.

Given the curvature expansion (10), the calculation is straightforward. We substitute the eigenvalues and lengths of prime cycles (for the example at hand, up to 5 pinball bounces - total of 14 cycles) into the curvature expansion (10), and obtain a polynomial approximation[20] to $1/\zeta$. We then vary the exponent γ in (7) and determine the escape rate γ by finding the smallest γ for which (10) vanishes. The zeros are easily determined by standard numerical methods (such as the Newton algorithm), with accuracy as good as 7 significant digits for the pinball example considered here.

The convergence can be illustrated by listing γ computed from truncations of (10) to different maximal cycle lengths. (disk radius to disk-disk center separation ratio is a:R = 1:6). The first column gives the maximal cycle length used, the second the estimate of the classical escape rate from the full 3-disk cycle expansion, the third from the fundamental domain expansion[5].

n	full 3 disks	fund. dom.
1		0.407693
2	0.43578	0.410280
3	0.40491	0.410336
4	0.40945	0.410338
5	0.41037	0.410338
6	0.41034	

For comparison, a numerical simulation of ref. [3] yields $\gamma = .410\ldots$, and the $n = 2,3$ approximations of ref. [3] yield 0.3102, 0.4508 respectively.

If one has some experience with numerical estimates of dimensions, one realizes that the convergence here is very impressive; only three input numbers (the two fixed points $\bar{0}$, $\bar{1}$ and the 2-cycle $\overline{10}$) already yield the escape rate to 4 significant digits! We have omitted an infinity of unstable cycles; so why does approximating the dynamics by a finite number of cycle eigenvalues work so well?

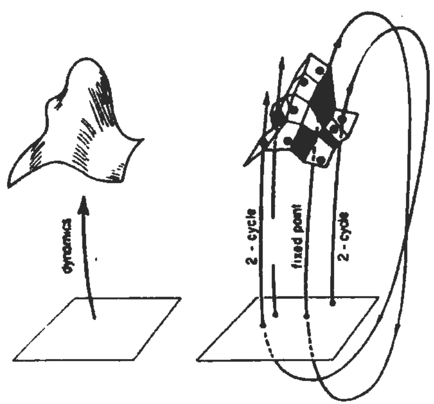

Figure 6. Approximation to (a) a smooth dynamics by (b) the skeleton of periodic points, together with their linearized neighborhoods.

A typical curvature expansion term in (10) is a *difference* of a long cycle $\{ab\}$ minus its shadowing approximation by shorter cycles $\{a\}$ and $\{b\}$:

$$t_{ab} - t_a t_b = t_{ab}(1 - t_a t_b / t_{ab}) = t_{ab}\left(1 - \left|\frac{\Lambda_{ab}}{\Lambda_a \Lambda_b}\right| e^{(T_a + T_b - T_{ab})\gamma}\right) \qquad (11)$$

To understand why this should be small compared to t_{ab}, try to visualise the description of a chaotic dynamical system in terms of cycles as a tessellation of the dynamical system, with smooth flow approximated by its periodic orbit skeleton, each "face" centered on a periodic point, and the scale of the "face" determined by the linearization of the flow around the periodic point (see fig. 6).

The orbits that follow the same symbolic dynamics, such as $\{ab\}$ and a "pseudo orbit" $\{a\}\{b\}$, lie physically close; longer and longer orbits resolve the dynamics with finer and finer resolution in the phase space. If the weights associated with the orbits are multiplicative along the flow (for example, products of derivatives) and the flow is smooth, the term in parenthesis in (11) falls off *exponentially* with the cycle length, and therefore the curvature expansions are expected to be highly convergent.

We have here used the curvature expansions to evaluate pinball escape rates, but the technique is much more general. It is, for example, applicable[17] to a broad class of "thermodynamical" averages such as those used in the extraction of generalized dimensions[21], with the sum (1) generalized to

$$\Gamma_n = \sum_i^{(n)} \frac{p_i^q}{l_i^\tau} , \qquad (12)$$

as well as to the quantum periodic orbits sums[4,5]. The simplest example is evaluation of the topological entropy h: the cycles are counted by setting $\tau = 0$, $t_p = e^{-n_p q(\tau)}$ and determining $h = -q(0)$ for which (8) has a zero. For the value of τ such that $q(\tau) = 0$, (12) is the classical definition of the Hausdorff-Besicovitch dimension $D_H = -\tau$, and so on. Each such application requires determination of the correct cycle weigth t_p - the rest is just machinery.

6 SUMMARY

A motion on a strange attractor can be approximated by shadowing the orbit by a sequence of nearby periodic orbits of finite length. This notion is here made precise by approximating orbits by primitive cycles, and evaluating associated curvatures. A curvature measures the deviation of a longer cycle from its approximation by shorter cycles; the smoothness of the dynamical system implies exponential fall-off for (almost) all curvatures. The technical prerequisite for implementing this shadowing is a good understanding of the symbolic dynamics of the classical dynamical system. The resulting curvature expansions offer an efficient method for evaluating classical and quantum periodic orbit sums; accurate estimates can be obtained by using as input the lengths and eigenvalues of a few prime cycles.

For reasons of clarity we have here illustrated the utility of cycles and their curvatures by the simple pinball example of fig. 1, but detailed investigations of the Hénon-type maps, scaling functions and quantum pinballs[19,17,5] give us some confidence in the general feasibility of the cycle analysis advocated here.

The cycle expansions such as (8) outperform the pedestrian methods such as extrapolations from the finite cover sums (1) for a number of reasons. The cycle expansion is a better averaging procedure than the naive box counting algorithms because the strange attractor is here pieced together in a controlled way from neighborhoods ("space average") rather than explored by a long ergodic trajectory ("time average"). The cycle expansion is (as explained above) co-ordinate and reparametrization invariant - a finite nth level sum (1) is not. Cycles are of finite period but infinite duration, so the cycle eigenvalues are already evaluated in the $n \to \infty$ limit, but for the sum (1) the limit has to be estimated by numerical extrapolations. And, crucially, the higher terms in the cycle expansion (8) are deviations of longer primitive cycles from their approximations by shorter cycles $c_{10} = -t_{10}+t_1t_0$, $c_{1001} = -t_{1001} + t_1t_{001} + t_{101}t_0 - t_{10}t_0t_1, \ldots$, which vanish exactly in piecewise linear approximations and are expected to fall off exponentially for smooth dynamical flows.

Acknowledgment: This lecture is a presentation of results an extended collaboration with R. Artuso, E. Aurell and B. Eckhardt, refs. [17,4,5]. The author is grateful to the Carlsberg Foundation for support.

References

[1] P. Grassberger and I. Procaccia, *Phys. Rev. A* **31**, 1872 (1985)

[2] B. Eckhardt, *J. Phys.* **A 20**, 5971 (1987)

[3] P. Gaspard and S.A. Rice, *J. Chem. Phys.* **90**, 2225 (1989), **90**, 2242 (1989) **90**, 2255 (1989)

[4] P. Cvitanović and B. Eckhardt, preprint (February 1989)

[5] P. Cvitanović, B. Eckhardt, and P. Scherer (in preparation)

[6] G.D. Birkhoff, *Acta Math.* **50**, 359 (1927), reprinted in ref. [7].

[7] R.S. MacKay and J.D. Miess, *Hamiltonian Dynamical Systems* (Adam Hilger, Bristol 1987)

[8] S. Smale, *Bull. Am. Math. Soc.* **73**, 747 (1967)

[9] The important role played by periodic orbits was already noted by H. Poincaré[10], and has been at core of much of the mathematical work on the theory of the dynamical systems ever since. We refer the reader to the classic text of Ruelle[11] and the reprint selection [7] for a summary and references to the mathematical literature.

[10] H. Poincaré, *Les méthodes nouvelles de la méchanique céleste* (Guthier-Villars, Paris 1892-99)

[11] D. Ruelle, *Statistical Mechanics, Thermodynamic Formalism*, (Addison-Wesley, Reading 1978).

[12] The figures of chapter IV. of P. Bergè, Y. Pomeau and C. Vidal, *Order within Chaos*, (Wiley and Sons, New York 1984) provide a good illustration of this non-uniqueness of representations of dynamical systems.

[13] The hope is that this is the maximum invariant information that can be extracted from a dynamical system; however, it is not known whether cycles indeed suffice. For example, it is argued in ref. [14] that this is not the case.

[14] M.J. Feigenbaum, *J. Stat. Phys.* **52**, 527 (1988).

[15] M. Hènon, *Comm. Math. Phys.* **50**, 69 (1976)

[16] See ref. [11], sect 7.23.

[17] R. Artuso, E. Aurell and P. Cvitanović, in preparation.

[18] A lucid introduction to escape from repellers is given in L.P. Kadanoff and C. Tang, *Proc. Natl. Acad. Sci.* **81**, 1276 (1984). The ζ-function formulation is given in D. Ruelle, *J. Stat. Phys.* **44**, 281 (1986); W. Parry and M. Pollicott, *Ann. Math.* **118**, 573 (1983) and discussed in ref.[19].

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

[20] The convergence can be improved further by estimating c_n for large n and summing the asymptotic tails, see ref. [17].

[21] P. Grassberger, *Phys. Lett.* **97A**, 227 (1983); **107A**, 101 (1985); H.G.E. Hentschel and I. Procaccia, *Physica* **8D**, 435 (1983); R. Benzi, G. Paladin, G. Parisi and A. Vulpiani, *J. Phys.* **A17**, 3521 (1984); T.C. Halsey, M.H. Jensen, L.P. Kadanoff, I. Procaccia and B.I. Shraiman, *Phys. Rev.* **A107**, 1141 (1986)