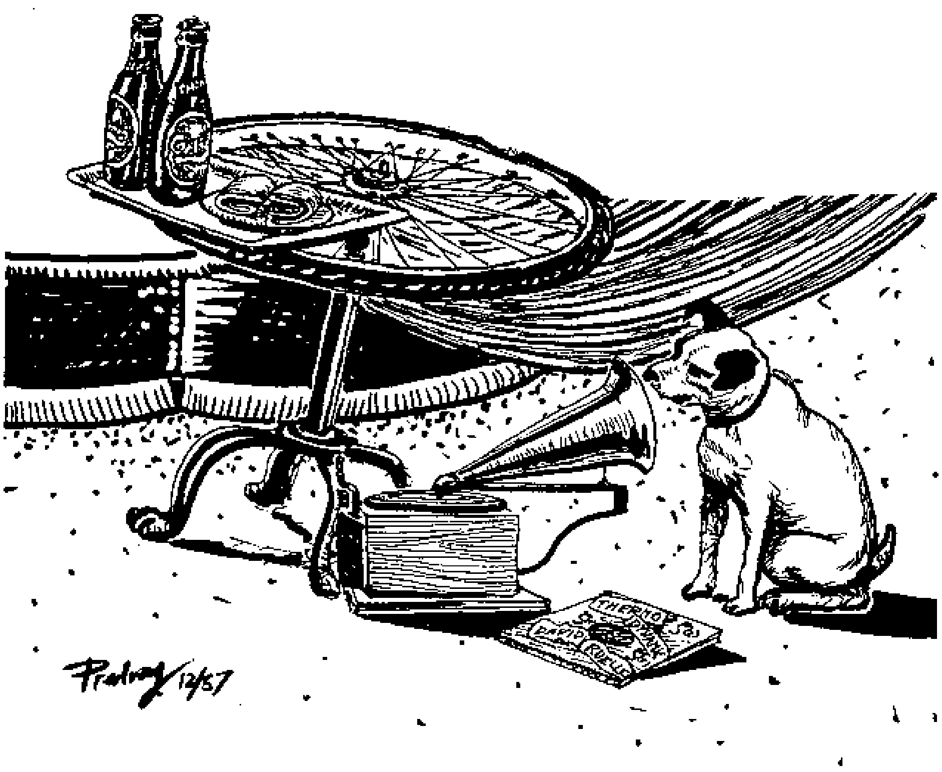
Classical and Quantum Chaos
Part I: Deterministic Chaos



Predrag Cvitanović – Roberto Artuso – Per Dahlqvist – Ronnie Mainieri – Gregor Tanner – Gábor Vattay – Niall Whelan – Andreas Wirzba

version 10, Aug 26 2003 printed August 27, 2003
`www.nbi.dk/ChaosBook/ChaosBook` comments to: predrag@nbi.dk

Contents

Volume I: Classical chaos

Volume II: Quantum chaos

Volume www: Appendices on `www.nbi.dk/ChaosBook`

Contributors

No man but a blockhead ever wrote except for money
Samuel Johnson

This book is a result of collaborative labors of many people over a span of several decades. Coauthors of a chapter or a section are indicated in the byline to the chapter/section title. If you are referring to a specific coauthored section rather than the entire book, cite it as (for example):

> C. Chandre, F.K. Diakonos and P. Schmelcher, section "Discrete cyclist relaxation method", in P. Cvitanović, R. Artuso, R. Mainieri, G. Tanner and G. Vattay, *Chaos: Classical and Quantum* (Niels Bohr Institute, Copenhagen 2003); `www.nbi.dk/ChaosBook`.

Chapters without a byline are written by Predrag Cvitanović. Friends whose contributions and ideas were invaluable to us but have not contributed written text to this book, are listed in the acknowledgements.

Roberto Artuso

Ronnie Mainieri

Gábor Vattay

Gregor Tanner

Acknowledgements

> I feel I never want to write another book. What's the good! I can eke living on stories and little articles, that don't cost a tithe of the output a book costs. Why write novels any more!
>
> D.H. Lawrence

This book owes its existence to the Niels Bohr Institute's and Nordita's hospitable and nurturing environment, and the private, national and cross-national foundations that have supported the collaborators' research over a span of several decades. P.C. thanks M.J. Feigenbaum of Rockefeller University; D. Ruelle of I.H.E.S., Bures-sur-Yvette; I. Procaccia of the Weizmann Institute; P. Hemmer of University of Trondheim; The Max-Planck Institut für Mathematik, Bonn; J. Lowenstein of New York University; Edificio Celi, Milano; and Fundaçaõ de Faca, Porto Seguro, for the hospitality during various stages of this work, and the Carlsberg Foundation and Glen P. Robinson for support.

The authors gratefully acknowledge collaborations and/or stimulating discussions with E. Aurell, V. Baladi, B. Brenner, A. de Carvalho, D.J. Driebe, B. Eckhardt, M.J. Feigenbaum, J. Frøjland, P. Gaspar, P. Gaspard, J. Guckenheimer, G.H. Gunaratne, P. Grassberger, H. Gutowitz, M. Gutzwiller, K.T. Hansen, P.J. Holmes, T. Janssen, R. Klages, Y. Lan, B. Lauritzen, J. Milnor, M. Nordahl, I. Procaccia, J.M. Robbins, P.E. Rosenqvist, D. Ruelle, G. Russberg, M. Sieber, D. Sullivan, N. Søndergaard, T. Tél, C. Tresser, and D. Wintgen.

We thank Dorte Glass for typing parts of the manuscript; B. Lautrup and D. Viswanath for comments and corrections to the preliminary versions of this text; M.V. Berry for the quotation on page 603; H. Fogedby for the quotation on page 239; J. Greensite for the quotation on page 6; Ya.B. Pesin for the remarks quoted on page 612; E.A. Spiegel for quotations on page 1 and page 693.

Fritz Haake's heartfelt lament on page 206 was uttered at the end of the first conference presentation of cycle expansions, in 1988. G.P. Morriss advice to students as how to read the introduction to this book, page 4, was offerred during a 2002 graduate course in Dresden. Kerson Huang's interview of C.N. Yang quoted on page 111 is available on `www.nbi.dk/ChaosBook/extras`.

Who is the 3-legged dog reappearing throughout the book? Long ago, when we were innocent and knew not Borel measurable α to Ω sets, P. Cvitanović asked V. Baladi a question about dynamical zeta functions, who then asked J.-P. Eckmann, who then asked D. Ruelle. The answer was transmitted back: "The master says: 'It is holomorphic in a strip' ". Hence His Master's Voice logo, and the 3-legged dog is us, still eager to fetch the bone. The answer has made it to the book, though not precisely in His Master's words. As a matter of fact, the answer *is* the book. We are still chewing on it.

Profound thanks to all the unsung heroes - students and colleagues, too numerous to list here, who have supported this project over many years in many ways, by surviving pilot courses based on this book, by providing invaluable insights, by teaching us, by inspiring us.

Chapter 1

Overture

> If I have seen less far than other men it is because I have stood behind giants.
> Edoardo Specchio

Rereading classic theoretical physics textbooks leaves a sense that there are holes large enough to steam a Eurostar train through them. Here we learn about harmonic oscillators and Keplerian ellipses - but where is the chapter on chaotic oscillators, the tumbling Hyperion? We have just quantized hydrogen, where is the chapter on the classical 3-body problem and its implications for quantization of helium? We have learned that an instanton is a solution of field-theoretic equations of motion, but shouldn't a strongly nonlinear field theory have turbulent solutions? How are we to think about systems where things fall apart; the center cannot hold; every trajectory is unstable?

This chapter offers a quick survey of the main topics covered in the book. We start out by making promises - we will right wrongs, no longer shall you suffer the slings and arrows of outrageous Science of Perplexity. We relegate a historical overview of the development of chaotic dynamics to appendix A, and head straight to the starting line: A pinball game is used to motivate and illustrate most of the concepts to be developed in this book.

Throughout the book


indicates that the section requires a hearty stomach and is probably best skipped on first reading


fast track points you where to skip to

 tells you where to go for more depth on a particular topic

 indicates an exercise that might clarify a point in the text

 indicates that a figure is still missing - you are urged to fetch it

This is a textbook, not a research monograph, and you should be able to follow the thread of the argument without constant excursions to sources. Hence there are no literature references in the text proper, all learned remarks and bibliographical pointers are relegated to the "Commentary" section at the end of each chapter.

1.1 Why this book?

> It seems sometimes that through a preoccupation with science, we acquire a firmer hold over the vicissitudes of life and meet them with greater calm, but in reality we have done no more than to find a way to escape from our sorrows.
>
> Hermann Minkowski in a letter to David Hilbert

The problem has been with us since Newton's first frustrating (and unsuccessful) crack at the 3-body problem, lunar dynamics. Nature is rich in systems governed by simple deterministic laws whose asymptotic dynamics are complex beyond belief, systems which are locally unstable (almost) everywhere but globally recurrent. How do we describe their long term dynamics?

The answer turns out to be that we have to evaluate a determinant, take a logarithm. It would hardly merit a learned treatise, were it not for the fact that this determinant that we are to compute is fashioned out of infinitely many infinitely small pieces. The feel is of statistical mechanics, and that is how the problem was solved; in 1960's the pieces were counted, and in 1970's they were weighted and assembled together in a fashion that in beauty and in depth ranks along with thermodynamics, partition functions and path integrals amongst the crown jewels of theoretical physics.

Then something happened that might be without parallel; this is an area of science where the advent of cheap computation had actually subtracted from our collective understanding. The computer pictures and numerical plots of fractal science of 1980's have overshadowed the deep insights of the 1970's, and these pictures have since migrated into textbooks. Fractal science posits that certain quantities (Lyapunov exponents, generalized dimensions, . . .) can be estimated on a computer. While some of the numbers so obtained are indeed mathematically sensible characterizations of fractals, they are in no sense observable and measurable on the length and time scales dominated by chaotic dynamics.

Even though the experimental evidence for the fractal geometry of nature is circumstantial, in studies of probabilistically assembled fractal aggregates we know of nothing better than contemplating such quantities. In deterministic systems we can do *much* better. Chaotic dynamics is generated by interplay of locally unstable motions, and interweaving of their global stable and unstable manifolds. These features are robust and accessible in systems as noisy as slices of rat brains. Poincaré, the first to understand deterministic chaos, already said as much (modulo rat brains). Once the topology of chaotic dynamics is understood, a powerful theory yields the macroscopically measurable consequences of chaotic dynamics, such as atomic spectra, transport coefficients, gas pressures.

That is what we will focus on in this book. This book is a self-contained graduate textbook on classical and quantum chaos. We teach you how to evaluate a determinant, take a logarithm, stuff like that. Should take 100 pages or so. Well, we fail - so far we have not found a way to traverse this material in less than a semester, or 200-300 page subset of this text. Nothing to be done about that.

1.2 Chaos ahead

Things fall apart; the centre cannot hold
W.B. Yeats: *The Second Coming*

Study of chaotic dynamical systems is no recent fashion. It did not start with the widespread use of the personal computer. Chaotic systems have been studied for over 200 years. During this time many have contributed, and the field followed no single line of development; rather one sees many interwoven strands of progress.

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 "non-generic", have gotten us very far. The success has lulled us into a habit of expecting simple solutions to simple equations - an expectation tempered 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 and statistical investigations. Nevertheless, a beautiful theory of deterministic chaos, of predictive quality comparable to that of the traditional perturbation expansions for nearly integrable systems, already exists.

In the traditional approach the integrable motions are used as zeroth-order approximations to physical systems, and weak nonlinearities are then accounted for perturbatively. For strongly nonlinear, non-integrable systems such expansions fail completely; the asymptotic time phase space exhibits amazingly rich structure which is not at all apparent in the integrable approximations. However, hidden in this apparent chaos is a rigid skeleton, a tree of *cycles* (periodic orbits) of increasing lengths and self-similar structure. The insight of the modern dynamical systems theory is that the zeroth-order approximations to the harshly chaotic dynamics should be very different from those for the nearly integrable systems: a good starting approximation here is the linear stretching and folding of a baker's map, rather than the periodic motion of a harmonic oscillator.

So, what is chaos, and what is to be done about it? To get some feeling for how and why unstable cycles come about, we start by playing a game of pinball. The reminder of the chapter is a quick tour through the material covered in this book. Do not worry if you do not understand every detail at the first reading – the intention is to give you a feeling for the main themes of the book, details will be filled out later. If you want to get a particular point clarified right now, ☞ on the margin points at the appropriate section.

1.3 The future as in a mirror

> All you need to know about chaos is contained in the introduction of the [Cvitanović *et al "Chaos: Classical and Quantum"*] book. However, in order to understand the introduction you will first have to read the rest of the book.
> Gary Morriss

That deterministic dynamics leads to chaos is no surprise to anyone who has tried pool, billiards or snooker – the game is about beating chaos – so we start our story about what chaos is, and what to do about it, with a game of *pinball*. This might seem a trifle, but the game of pinball is to chaotic dynamics what a pendulum is to integrable systems: thinking clearly about what "chaos" in a game of pinball is will help us tackle more difficult problems, such as computing diffusion constants in deterministic gases, or computing the helium spectrum.

We all have an intuitive feeling for what a ball does as it bounces among the pinball machine's disks, and only high-school level Euclidean geometry

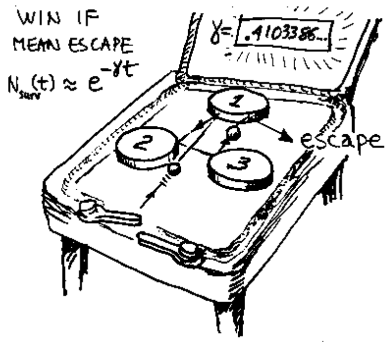


Figure 1.1: A physicist's bare bones game of pinball.

is needed to describe its trajectory. A physicist's pinball game is the game of pinball stripped to its bare essentials: three equidistantly placed reflecting disks in a plane, fig. 1.1. A physicist's pinball is free, frictionless, point-like, spin-less, perfectly elastic, and noiseless. 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.

At the beginning of 18th century Baron Gottfried Wilhelm Leibniz was confident that given the initial conditions one knew all what a deterministic system would do far into the future. He wrote [1.1], anticipating by century and a half the oft quoted Laplace's "Given for one instant an intelligence which could comprehend all the forces by which nature is animated...":

> That everything is brought forth through an established destiny is just as certain as that three times three is nine. [. . .] If, for example, one sphere meets another sphere in free space and if their sizes and their paths and directions before collision are known, we can then foretell and calculate how they will rebound and what course they will take after the impact. Very simple laws are followed which also apply, no matter how many spheres are taken or whether objects are taken other than spheres. From this one sees then that everything proceeds mathematically – that is, infallibly – in the whole wide world, so that if someone could have a sufficient insight into the inner parts of things, and in addition had remembrance and intelligence enough to consider all the circumstances and to take them into account, he would be a prophet and would see the future in the present as in a mirror.

Leibniz chose to illustrate his faith in determinism precisely with the type of physical system that we shall use here as a paradigm of "chaos". His claim is wrong in a deep and subtle way: a state of a physical system can *never* be specified to infinite precision, there is no way to take all the circumstances into account, and a single trajectory cannot be tracked, only a ball of nearby initial points makes physical sense.

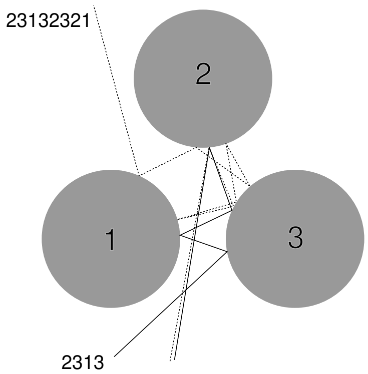


Figure 1.2: Sensitivity to initial conditions: two pinballs that start out very close to each other separate exponentially with time.

1.3.1 What is "chaos"?

> I accept chaos. I am not sure that it accepts me.
> Bob Dylan, *Bringing It All Back Home*

A deterministic system is a system whose present state is *in principle* fully determined by its initial conditions, in contra-distinction to a stochastic system, for which the initial conditions determine the present state only partially, due to noise, or other external circumstances beyond our control. For a stochastic system, the present state reflects the past initial conditions plus the particular realization of the noise encountered along the way.

A deterministic system with sufficiently complicated dynamics can fool us into regarding it as a stochastic one; disentangling the deterministic from the stochastic is the main challenge in many real-life settings, from stock market to palpitations of chicken hearts. So, what is "chaos"?

In a game of pinball, any two trajectories that start out very close to each other separate exponentially with time, and in a finite (and in practice, a very small) number of bounces their separation $\delta\mathbf{x}(t)$ attains the magnitude of L, the characteristic linear extent of the whole system, fig. 1.2. This property of *sensitivity to initial conditions* can be quantified as

$$|\delta\mathbf{x}(t)| \approx e^{\lambda t}|\delta\mathbf{x}(0)|$$

where λ, the mean rate of separation of trajectories of the system, is called the *Lyapunov exponent.* For any finite accuracy $|\delta\mathbf{x}(0)| = \delta x$ of the initial data, the dynamics is predictable only up to a finite *Lyapunov time*

☞ sect. 8.3

$$T_{\mathrm{Lyap}} \approx -\frac{1}{\lambda}\ln|\delta x/L| \,, \tag{1.1}$$

Figure 1.3: Dynamics of a *chaotic* dynamical system is (a) everywhere locally unstable (positive Lyapunov exponent) and (b) globally mixing (positive entropy). (A. Johansen)

despite the deterministic and, for baron Leibniz, infallible simple laws that rule the pinball motion.

A positive Lyapunov exponent does not in itself lead to chaos. One could try to play 1- or 2-disk pinball game, but it would not be much of a game; trajectories would only separate, never to meet again. What is also needed is *mixing*, the coming together again and again of trajectories. While locally the nearby trajectories separate, the interesting dynamics is confined to a globally finite region of the phase space and thus of necessity the separated trajectories are folded back and can re-approach each other arbitrarily closely, infinitely many times. In the case at hand there are 2^n topologically distinct n bounce trajectories that originate from a given disk. More generally, the number of distinct trajectories with n bounces can be quantified as

$$N(n) \approx e^{hn}$$

☞ sect. 10.1

where the *topological entropy* h ($h = \ln 2$ in the case at hand) is the growth rate of the number of topologically distinct trajectories.

☞ sect. 17.1

The appellation "chaos" is a confusing misnomer, as in deterministic dynamics there is no chaos in the everyday sense of the word; everything proceeds mathematically – that is, as baron Leibniz would have it, infallibly. 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. The word "chaos" has in this context taken on a narrow technical meaning. If a deterministic system is locally unstable (positive Lyapunov exponent) and globally mixing (positive entropy) - fig. 1.3 - it is said to be *chaotic*.

While mathematically correct, the definition of chaos as "positive Lyapunov + positive entropy" is useless in practice, as a measurement of these quantities is intrinsically asymptotic and beyond reach for systems observed in nature. More powerful is Poincaré's vision of chaos as the interplay of local instability (unstable periodic orbits) and global mixing (intertwining of their stable and unstable manifolds). In a chaotic system any open ball of initial conditions, no matter how small, will in finite time overlap with any other finite region and in this sense spread over the extent of the entire asymptotically accessible phase space. Once this is grasped, the focus of theory shifts from attempting precise prediction of individual trajectories (which is impossible) to description of the geometry of the space of possible outcomes, and evaluation of averages over this space. How this is accomplished is what this book is about.

A definition of “turbulence” is harder to come by. Intuitively, the word refers to irregular behavior of an infinite-dimensional dynamical system described by deterministic equations of motion - say, a bucket of boiling water described by the Navier-Stokes equations. But in practice the word “turbulence” tends to refer to messy dynamics which we understand poorly. As soon as a phenomenon is understood better, it is reclaimed and renamed: “a route to chaos”, “spatiotemporal chaos”, and so on.

☞ appendix B

In this book we shall develop a theory of chaotic dynamics for low dimensional attractor visualized as a succession of nearly periodic but unstable motions. In the same spirit, we shall think of turbulence in spatially extended systems in terms of recurrent spatiotemporal patterns. Pictorially, dynamics drives a given spatially extended system through a repertoire of unstable patterns; as we watch a turbulent system evolve, every so often we catch a glimpse of a familiar pattern:

$\Longrightarrow$ other swirls $\Longrightarrow$

For any finite spatial resolution, the system follows approximately for a finite time a pattern belonging to a finite alphabet of admissible patterns, and the long term dynamics can be thought of as a walk through the space of such patterns. Recasting this image into mathematics is what this book is about.

1.3.2 When does “chaos” matter?

> Whether ’tis nobler in the mind to suffer
> The slings and arrows of outrageous fortune,
> Or to take arms against a sea of troubles,
> And by opposing end them?
> W. Shakespeare, *Hamlet*

When should we be mindful of chaos? The solar system is “chaotic”, yet we have no trouble keeping track of the annual motions of planets. The rule of thumb is this; if the Lyapunov time (1.1) (the time by which a phase space region initially comparable in size to the observational accuracy extends across the entire accessible phase space) is significantly shorter than the observational time, you need to master the theory that will be developed here. That is why the main successes of the theory are in statistical mechanics, quantum mechanics, and questions of long term stability in celestial mechanics.

As in science popularizations too much has been made of the impact of the "chaos theory", a number of caveats are already needed at this point.

At present the theory is in practice applicable only to systems with a low intrinsic *dimension* – the minimum number of degrees of freedom necessary to capture its essential dynamics. If the system is very turbulent (description of its long time dynamics requires a space of high intrinsic dimension) we are out of luck. Hence insights that the theory offers to elucidation of problems of fully developed turbulence, quantum field theory of strong interactions and early cosmology have been modest at best. Even that is a caveat with qualifications. There are applications – such as spatially extended systems and statistical mechanics applications – where the few important degrees of freedom can be isolated and studied profitably by methods to be described here.

☞ sect. 2.4.1
☞ chapter 20

The theory has had limited practical success applied to the very noisy systems so important in life sciences and in economics. Even though we are often interested in phenomena taking place on time scales much longer than the intrinsic time scale (neuronal interburst intervals, cardiac pulse, etc.), disentangling "chaotic" motions from the environmental noise has been very hard.

1.4 A game of pinball

Formulas hamper the understanding.
S. Smale

We are now going to get down to the brasstacks. But first, a disclaimer: If you understand most of the rest of this chapter on the first reading, you either do not need this book, or you are delusional. If you do not understand it, is not because the people who wrote it are so much smarter than you: the most one can hope for at this stage is to give you a flavor of what lies ahead. If a statement in this chapter mystifies/intrigues, fast forward to a section indicated by ☞ on the margin, read only the parts that you feel you need. Of course, we think that you need to learn ALL of it, or otherwise we would not have written it up in the first place.

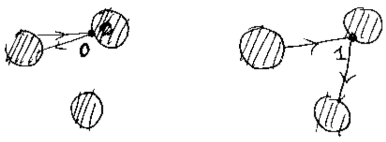


Figure 1.4: Binary labeling of the 3-disk pinball trajectories; a bounce in which the trajectory returns to the preceding disk is labeled 0, and a bounce which results in continuation to the third disk is labeled 1.

Confronted with a potentially chaotic dynamical system, we analyze it through a sequence of three distinct stages; diagnose, count, measure. I. First we determine the intrinsic *dimension* of the system – the minimum number of degrees of freedom necessary to capture its essential dynamics. If the system is very turbulent (description of its long time dynamics requires a space of high intrinsic dimension) we are, at present, out of luck. We know only how to deal with the transitional regime between regular motions and a few chaotic degrees of freedom. That is still something; even an infinite-dimensional system such as a burning flame front can turn out to have a very few chaotic degrees of freedom. In this regime the chaotic dynamics is restricted to a space of low dimension, the number of relevant parameters is small, and we can proceed to step II; we *count* and *classify* all possible topologically distinct trajectories of the system into a hierarchy whose successive layers require increased precision and patience on the part of the observer. This we shall do in sects. 1.4 and 1.4.1. If successful, we can proceed with step III of sect. 1.5.1: investigate the *weights* of the different pieces of the system.

☞ sect. 2.4.1
☞ chapter 9
☞ chapter 10

We commence our analysis of the pinball game with steps I, II: diagnose, count. We shall return to step III – measure – in sect. 1.5.1.

☞ chapter 15

With the game of pinball we are in luck – it is a low dimensional system, free motion in a plane. The motion of a point particle is such that after a collision with one disk it either continues to another disk or it escapes. If we label the three disks by 1, 2 and 3, we can associate every trajectory with an *itinerary*, a sequence of labels which indicates the order in which the disks are visited; for example, the two trajectories in fig. 1.2 have itineraries _2313_, _23132321_ respectively. The itinerary will be finite for a scattering trajectory, coming in from infinity and escaping after a finite number of collisions, infinite for a trapped trajectory, and infinitely repeating for a periodic orbit. Parenthetically, in this subject the words "orbit" and "trajectory" refer to one and the same thing.

✎ 1.1
page 32

Such labeling is the simplest example of *symbolic dynamics*. As the particle cannot collide two times in succession with the same disk, any two consecutive symbols must differ. This is an example of *pruning*, a rule that forbids certain subsequences of symbols. Deriving pruning rules is in general a difficult problem, but with the game of pinball we are lucky - there are no further pruning rules.

The choice of symbols is in no sense unique. For example, as at each bounce we can either proceed to the next disk or return to the previous disk, the above 3-letter alphabet can be replaced by a binary $\{0, 1\}$ alphabet, fig. 1.4. A clever choice of an alphabet will incorporate important features of the dynamics, such as its symmetries.

☞ sect. 9.7

Suppose you wanted to play a good game of pinball, that is, get the pinball to bounce as many times as you possibly can – what would be a winning strategy? The simplest thing would be to try to aim the pinball so it bounces many times between a pair of disks – if you managed to shoot it so it starts out in the periodic orbit bouncing along the line connecting two disk centers, it would stay there forever. Your game would be just as good if you managed to get it to keep bouncing between the three disks forever, or place it on any periodic orbit. The only rub is that any such orbit is *unstable*, so you have to aim very accurately in order to stay close to it for a while. So it is pretty clear that if one is interested in playing well, unstable periodic orbits are important – they form the *skeleton* onto which all trajectories trapped for long times cling.

☞ sect. 30.2

1.4.1 Partitioning with periodic orbits

A trajectory is periodic if it returns to its starting position and momentum. We shall refer to the set of periodic points that belong to a given periodic orbit as a *cycle*.

Short periodic orbits are easily drawn and enumerated - some examples are drawn in fig. 1.5 - but it is rather hard to perceive the systematics of orbits from their shapes. In mechanics a trajectory is fully and uniquely specified by its position and momentum at a given instant, and no two distinct phase space trajectories can intersect. Their projections on arbitrary subspaces, however, can and do intersect, in rather unilluminating ways. In the pinball example the problem is that we are looking at the projections of a 4-dimensional phase space trajectories onto a 2-dimensional subspace, the space coordinates. A clearer picture of the dynamics is obtained by constructing a phase space Poincaré section.

The position of the ball is described by a pair of numbers (the spatial coordinates on the plane) and its velocity by another pair of numbers (the components of the velocity vector). As far as baron Leibniz is concerned, this is a complete description.

Suppose that the pinball has just bounced off disk 1. Depending on its position and outgoing angle, it could proceed to either disk 2 or 3. Not much happens in between the bounces – the ball just travels at constant velocity along a straight line – so we can reduce the four-dimensional flow to a two-

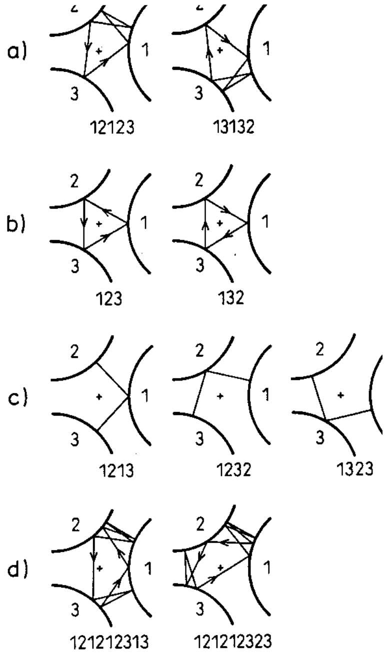


Figure 1.5: Some examples of 3-disk cycles: (a) $\overline{12123}$ and $\overline{13132}$ are mapped into each other by the flip across 1 axis. Similarly (b) $\overline{123}$ and $\overline{132}$ are related by flips, and (c) $\overline{1213}$, $\overline{1232}$ and $\overline{1323}$ by rotations. (d) The cycles $\overline{121212313}$ and $\overline{121212323}$ are related only by time reversal. These symmetries are discussed in more detail in chapter 19. (from ref. [1.2])

dimensional map f that takes the coordinates of the pinball from one disk edge to another disk edge. Let us state this more precisely: the trajectory just after the moment of impact is defined by marking s_n, the arc-length position of the nth bounce along the billiard wall, and $p_n = p \sin \phi_n$ the momentum component parallel to the billiard wall at the point of impact, fig. 1.6. Such section of a flow is called a *Poincaré section*, and the particular choice of coordinates (due to Birkhoff) is particularly smart, as it conserves the phase-space volume. In terms of the Poincaré section, the dynamics is reduced to the *return map* $P : (s_n, p_n) \mapsto (s_{n+1}, p_{n+1})$ from the boundary of a disk to the boundary of the next disk. The explicit form of this map is easily written down, but it is of no importance right now.

☞ sect. 5

Next, we mark in the Poincaré section those initial conditions 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 without further ado. We label the two strips $\mathcal{M}_0$, $\mathcal{M}_1$. Embedded within them there are four strips $\mathcal{M}_{00}$, $\mathcal{M}_{10}$, $\mathcal{M}_{01}$, $\mathcal{M}_{11}$ of initial conditions that survive for two bounces, and so forth, see figs. 1.7 and 1.8. Provided that the disks are sufficiently separated, after n bounces the survivors are divided into 2^n distinct strips: the $\mathcal{M}_i$th strip consists of all points with itinerary $i = s_1 s_2 s_3 \ldots s_n$, $s = \{0, 1\}$. The unstable cycles as a skeleton of chaos are almost visible here: each such patch contains a periodic point $\overline{s_1 s_2 s_3 \ldots s_n}$ with the basic block infinitely repeated. Periodic points are skeletal in the sense that as we look further and further, the strips shrink but the periodic points stay put forever.

We see now why it pays to have a symbolic dynamics; it provides a nav-

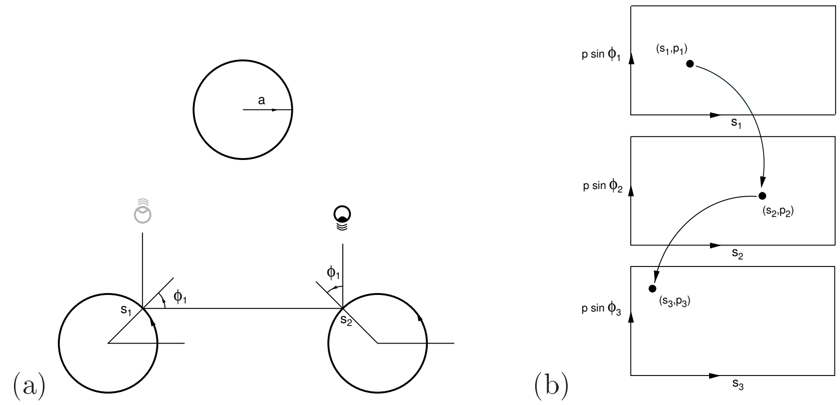
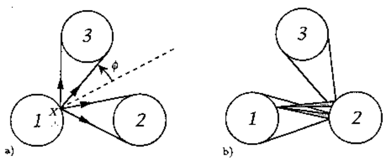


Figure 1.6: (a) The Poincaré section coordinates for the 3-disk game of pinball. (b) Collision sequence $(s_1, p_1) \mapsto (s_2, p_2) \mapsto (s_3, p_3)$ from the boundary of a disk to the boundary of the next disk presented in the Poincaré section coordinates.

Figure 1.7: (a) A trajectory starting out from disk 1 can either hit another disk or escape. (b) Hitting two disks in a sequence requires a much sharper aim. The cones of initial conditions that hit more and more consecutive disks are nested within each other, as in fig. 1.8.



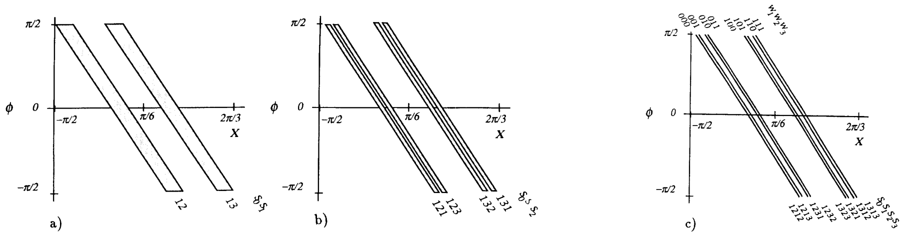


Figure 1.8: Ternary labeled regions of the 3-disk game of pinball phase space Poincaré section which correspond to trajectories that originate on disk 1 and remain confined for (a) one bounce, (b) two bounces, (c) three bounces. The Poincaré sections for trajectories originating on the other two disks are obtained by the appropriate relabeling of the strips (K.T. Hansen [1.3]).

igation chart through chaotic phase space. There exists a unique trajectory for every admissible infinite length itinerary, and a unique itinerary labels every trapped trajectory. For example, the only trajectory labeled by $\overline{12}$ is the 2-cycle bouncing along the line connecting the centers of disks 1 and 2; any other trajectory starting out as $12\ldots$ either eventually escapes or hits the 3rd disk.

1.4.2 Escape rate

☞ remark 7.1

What is a good physical quantity to compute for the game of pinball? Such system, for which almost any trajectory eventually leaves a finite region (the pinball table) never to return, is said to be open, or a *repeller*. The repeller *escape rate* is an eminently measurable quantity. An example of such measurement would be an unstable molecular or nuclear state which can be well approximated by a classical potential with possibility of escape in certain directions. In an experiment many projectiles are injected into such a non-confining potential and their mean escape rate is measured, as in fig. 1.1. The numerical experiment might consist of injecting the pinball between the disks in some random direction and asking how many times the pinball bounces on the average before it escapes the region between the disks.

1.2
page 32

For a theorist a good game of pinball consists in predicting accurately the asymptotic lifetime (or the escape rate) of the pinball. We now show how the periodic orbit theory accomplishes this for us. Each step will be so simple that you can follow even at the cursory pace of this overview, and still the result is surprisingly elegant.

Consider fig. 1.8 again. In each bounce the initial conditions get thinned out, yielding twice as many thin strips as at the previous bounce. The total area that remains at a given time is the sum of the areas of the strips, so that the fraction of survivors after n bounces, or the *survival probability* is given by

$$\begin{aligned}
\hat{\Gamma}_1 &= \frac{|\mathcal{M}_0|}{|\mathcal{M}|} + \frac{|\mathcal{M}_1|}{|\mathcal{M}|}\,, \qquad \hat{\Gamma}_2 = \frac{|\mathcal{M}_{00}|}{|\mathcal{M}|} + \frac{|\mathcal{M}_{10}|}{|\mathcal{M}|} + \frac{|\mathcal{M}_{01}|}{|\mathcal{M}|} + \frac{|\mathcal{M}_{11}|}{|\mathcal{M}|}\,, \\
\hat{\Gamma}_n &= \frac{1}{|\mathcal{M}|} \sum_i^{(n)} |\mathcal{M}_i|\,, \qquad (1.2)
\end{aligned}$$

where i is a label of the ith strip, $|\mathcal{M}|$ is the initial area, and $|\mathcal{M}_i|$ is the area of the ith strip of survivors. $i = 01, 10, 11, \ldots$ is a label, not a binary number. Since at each bounce one routinely loses about the same fraction of trajectories, one expects the sum (1.2) to fall off exponentially with n and tend to the limit

$$\hat{\Gamma}_{n+1}/\hat{\Gamma}_n = e^{-\gamma_n} \to e^{-\gamma}. \tag{1.3}$$

The quantity γ is called the *escape rate* from the repeller.

1.5 Chaos for cyclists

We shall now show that the escape rate γ can be extracted from a highly convergent *exact* expansion by reformulating the sum (1.2) in terms of unstable periodic orbits.

If, when asked what the 3-disk escape rate is for disk radius 1, center-center separation 6, velocity 1, you answer that the continuous time escape rate is roughly $\gamma = 0.4103384077693464893384613078192\ldots$, you do not need this book. If you have no clue, hang on.

1.5.1 Size of a partition

Not only do the periodic points keep track of locations and the ordering of the strips, but, as we shall now show, they also determine their size.

As a trajectory evolves, it carries along and distorts its infinitesimal neighborhood. Let

$$x(t) = f^t(x_0)$$

denote the trajectory of an initial point $x_0 = x(0)$. To linear order, the evolution of the distance to a neighboring trajectory $x_i(t) + \delta x_i(t)$ is given by the Jacobian matrix

$$\delta x_i(t) = \sum_{j=1}^{d} \mathbf{J}^t(x_0)_{ij} \delta x_{0j}\,, \qquad \mathbf{J}^t(x_0)_{ij} = \frac{\partial x_i(t)}{\partial x_{0j}}\,.$$

A trajectory of a pinball moving on a flat surface is specified by two position coordinates and the direction of motion, so in this case $d = 3$. Evaluation of a cycle Jacobian matrix is a longish exercise - here we just state the result. ☞ sect. 5.3 The Jacobian matrix describes the deformation of an infinitesimal neighborhood of $x(t)$ as it goes with the flow; its the eigenvectors and eigenvalues give the directions and the corresponding rates of its expansion or contraction. The trajectories that start out in an infinitesimal neighborhood are separated along the unstable directions (those whose eigenvalues are less than unity in magnitude), approach each other along the stable directions (those whose eigenvalues exceed unity in magnitude), and maintain their distance along the marginal directions (those whose eigenvalues equal unity in magnitude). In our game of pinball the beam of neighboring trajectories is defocused along the unstable eigendirection of the Jacobian matrix $\mathbf{J}$.

As the heights of the strips in fig. 1.8 are effectively constant, we can concentrate on their thickness. If the height is $\approx L$, then the area of the ith strip is $\mathcal{M}_i \approx L l_i$ for a strip of width l_i.

Each strip i in fig. 1.8 contains a periodic point x_i. The finer the intervals, the smaller is the variation in flow across them, and the contribution from the strip of width l_i is well approximated by the contraction around the periodic point x_i within the interval,

$$l_i = a_i/|\Lambda_i| \,, \tag{1.4}$$

where Λ_i is the unstable eigenvalue of the Jacobian matrix $\mathbf{J}^t(x_i)$ evaluated at the ith periodic point for $t = T_p$, the full period (due to the low dimensionality, the Jacobian can have at most one unstable eigenvalue). Note that it is the magnitude of this eigenvalue which is important and we can disregard its sign. The prefactors a_i reflect the overall size of the system and the particular distribution of starting values of x. As the asymptotic trajectories are strongly mixed by bouncing chaotically around the repeller, we expect them to be insensitive to smooth variations in the initial distribution.

☞ sect. 7.3

To proceed with the derivation we need the *hyperbolicity* assumption: for large n the prefactors $a_i \approx O(1)$ are overwhelmed by the exponential growth of Λ_i, so we neglect them. If the hyperbolicity assumption is justified, we can replace $|\mathcal{M}_i| \approx L l_i$ in (1.2) by $1/|\Lambda_i|$ and consider the sum

☞ sect. 11.1.1

$$\Gamma_n = \sum_i^{(n)} 1/|\Lambda_i| \,,$$

where the sum goes over all periodic points of period n. We now define a generating function for sums over all periodic orbits of all lengths:

$$\Gamma(z) = \sum_{n=1}^{\infty} \Gamma_n z^n \,. \tag{1.5}$$

Recall that for large n the nth level sum (1.2) tends to the limit $\Gamma_n \to e^{-n\gamma}$, so the escape rate γ is determined by the smallest $z = e^{\gamma}$ for which (1.5) diverges:

$$\Gamma(z) \approx \sum_{n=1}^{\infty} (ze^{-\gamma})^n = \frac{ze^{-\gamma}}{1 - ze^{-\gamma}} \,. \tag{1.6}$$

This is the property of $\Gamma(z)$ which motivated its definition. We now devise an alternate expression for (1.5) in terms of periodic orbits to make explicit the connection between the escape rate and the periodic orbits:

$$\begin{aligned}
\Gamma(z) &= \sum_{n=1}^{\infty} z^n \sum_{i}^{(n)} |\Lambda_i|^{-1} \\
&= \frac{z}{|\Lambda_0|} + \frac{z}{|\Lambda_1|} + \frac{z^2}{|\Lambda_{00}|} + \frac{z^2}{|\Lambda_{01}|} + \frac{z^2}{|\Lambda_{10}|} + \frac{z^2}{|\Lambda_{11}|} \\
&\quad + \frac{z^3}{|\Lambda_{000}|} + \frac{z^3}{|\Lambda_{001}|} + \frac{z^3}{|\Lambda_{010}|} + \frac{z^3}{|\Lambda_{100}|} + \ldots
\end{aligned} \tag{1.7}$$

For sufficiently small z this sum is convergent. The escape rate γ is now given by the leading pole of (1.6), rather than a numerical extrapolation of a sequence of γ_n extracted from (1.3). As any finite truncation $n < n_{trunc}$ of (1.7) is a polynomial in z, convergent for any z, finding this pole requires that we know something about Γ_n for any n, and that might be a tall order. ☞ sect. 11.4

We could now proceed to estimate the location of the leading singularity of $\Gamma(z)$ from finite truncations of (1.7) by methods such as Padé approximants. However, as we shall now show, it pays to first perform a simple resummation that converts this divergence into a *zero* of a related function.

1.5.2 Dynamical zeta function

If a trajectory retraces a *prime* cycle r times, its expanding eigenvalue is Λ_p^r. A prime cycle p is a single traversal of the orbit; its label is a non-repeating symbol string of n_p symbols. 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. By the chain rule for derivatives the stability of a cycle is the same everywhere along the orbit, so each prime cycle of length n_p contributes n_p terms to the sum (1.7). Hence (1.7) can be rewritten as ✎ 10.5 page 195 ☞ sect. 4.4

$$\Gamma(z) = \sum_p n_p \sum_{r=1}^{\infty} \left(\frac{z^{n_p}}{|\Lambda_p|} \right)^r = \sum_p \frac{n_p t_p}{1 - t_p} \,, \qquad t_p = \frac{z^{n_p}}{|\Lambda_p|} \tag{1.8}$$

where the index p runs through all distinct *prime* cycles. Note that we have resumed the contribution of the cycle p to all times, so truncating the summation up to given p is *not* a finite time $n \le n_p$ approximation, but an asymptotic, *infinite* time estimate based by approximating stabilities of all cycles by a finite number of the shortest cycles and their repeats. The $n_p z^{n_p}$ factors in (1.8) suggest rewriting the sum as a derivative

$$\Gamma(z) = -z\frac{d}{dz}\sum_p \ln(1-t_p)\,.$$

Hence $\Gamma(z)$ is a logarithmic derivative of the infinite product

$$1/\zeta(z) = \prod_p (1-t_p)\,, \qquad t_p = \frac{z^{n_p}}{|\Lambda_p|}\,. \tag{1.9}$$

This function is called the *dynamical zeta function*, in analogy to the Riemann zeta function, which motivates the choice of "zeta" in its definition as $1/\zeta(z)$. This is the prototype formula of the periodic orbit theory. The zero of $1/\zeta(z)$ is a pole of $\Gamma(z)$, and the problem of estimating the asymptotic escape rates from finite n sums such as (1.2) is now reduced to a study of the zeros of the dynamical zeta function (1.9). The escape rate is related by (1.6) to a divergence of $\Gamma(z)$, and $\Gamma(z)$ diverges whenever $1/\zeta(z)$ has a zero.

☞ sect. 16.1
☞ sect. 12.4

Easy, you say: "Zeros of (1.9) can be read off the formula, a zero $z_p = |\Lambda_p|^{1/n_p}$ for each term in the product. What's the problem?" Dead wrong!

1.5.3 Cycle expansions

How are formulas such as (1.9) used? We start by computing the lengths and eigenvalues of the shortest cycles. This usually requires some numerical work, such as the Newton's method searches for periodic solutions; we shall assume that the numerics is under control, and that *all* short cycles up to given length have been found. In our pinball example this can be done by elementary geometrical optics. It is very important not to miss any short cycles, as the calculation is as accurate as the shortest cycle dropped – including cycles longer than the shortest omitted does not improve the accuracy (unless exponentially many more cycles are included). The result of such numerics is a table of the shortest cycles, their periods and their stabilities.

☞ chapter 14

☞ sect. 25.3

Now expand the infinite product (1.9), grouping together the terms of the same total symbol string length

$$
\begin{aligned}
1/\zeta &= (1-t_0)(1-t_1)(1-t_{10})(1-t_{100})\cdots \\
&= 1 - t_0 - t_1 - [t_{10} - t_1 t_0] - [(t_{100} - t_{10} t_0) + (t_{101} - t_{10} t_1)] \\
&\quad -[(t_{1000} - t_0 t_{100}) + (t_{1110} - t_1 t_{110}) \\
&\quad +(t_{1001} - t_1 t_{001} - t_{101} t_0 + t_{10} t_0 t_1)] - \ldots \qquad (1.10)
\end{aligned}
$$

The virtue of the expansion is that the sum of all terms of the same total length n (grouped in brackets above) is a number that is exponentially smaller than a typical term in the sum, for geometrical reasons we explain in the next section.

☞ chapter 15

☞ sect. 15.1

The calculation is now straightforward. We substitute a finite set of the eigenvalues and lengths of the shortest prime cycles into the cycle expansion (1.10), and obtain a polynomial approximation to $1/\zeta$. We then vary z in (1.9) and determine the escape rate γ by finding the smallest $z = e^{\gamma}$ for which (1.10) vanishes.

1.5.4 Shadowing

When you actually start computing this escape rate, you will find out that the convergence is very impressive: only three input numbers (the two fixed points $\overline{0}$, $\overline{1}$ and the 2-cycle $\overline{10}$) already yield the pinball escape rate to 3-4 significant digits! We have omitted an infinity of unstable cycles; so why does approximating the dynamics by a finite number of the shortest cycle eigenvalues work so well?

☞ sect. 15.1.3

The convergence of cycle expansions of dynamical zeta functions is a consequence of the smoothness and analyticity of the underlying flow. Intuitively, one can understand the convergence in terms of the geometrical picture sketched in fig. 1.9; the key observation is that the long orbits are *shadowed* by sequences of shorter orbits.

A typical term in (1.10) is a difference of a long cycle $\{ab\}$ minus its shadowing approximation by shorter cycles $\{a\}$ and $\{b\}$

$$
t_{ab} - t_a t_b = t_{ab}(1 - t_a t_b / t_{ab}) = t_{ab} \left(1 - \left| \frac{\Lambda_{ab}}{\Lambda_a \Lambda_b} \right| \right) , \qquad (1.11)
$$

where a and b are symbol sequences of the two shorter cycles. If all orbits are weighted equally ($t_p = z^{n_p}$), such combinations cancel exactly; if orbits

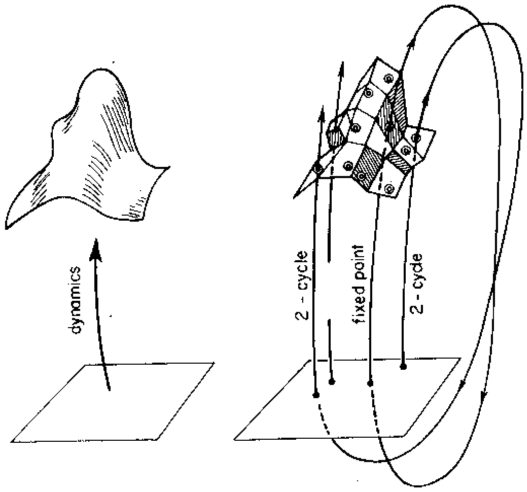


Figure 1.9: Approximation to (a) a smooth dynamics by (b) the skeleton of periodic points, together with their linearized neighborhoods. Indicated are segments of two 1-cycles and a 2-cycle that alternates between the neighborhoods of the two 1-cycles, shadowing first one of the two 1-cycles, and then the other.

of similar symbolic dynamics have similar weights, the weights in such combinations almost cancel.

This can be understood in the context of the pinball game as follows. Consider orbits $\overline{0}$, $\overline{1}$ and $\overline{01}$. The first corresponds to bouncing between any two disks while the second corresponds to bouncing successively around all three, tracing out an equilateral triangle. The cycle $\overline{01}$ starts at one disk, say disk 2. It then bounces from disk 3 back to disk 2 then bounces from disk 1 back to disk 2 and so on, so its itinerary is $\overline{2321}$. In terms of the bounce types shown in fig. 1.4, the trajectory is alternating between 0 and 1. The incoming and outgoing angles when it executes these bounces are very close to the corresponding angles for 0 and 1 cycles. Also the distances traversed between bounces are similar so that the 2-cycle expanding eigenvalue Λ_{01} is close in magnitude to the product of the 1-cycle eigenvalues $\Lambda_0\Lambda_1$.

To understand this on a more general level, try to visualize the partition of a chaotic dynamical system's phase space in terms of cycle neighborhoods 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, fig. 1.9.

The orbits that follow the same symbolic dynamics, such as $\{ab\}$ and a "pseudo orbit" $\{a\}\{b\}$, lie close to each other in the phase space; long shadowing pairs have to start out exponentially close to beat the exponential growth in separation with time. If the weights associated with the orbits are multiplicative along the flow (for example, by the chain rule for products of derivatives) and the flow is smooth, the term in parenthesis in (1.11) falls off exponentially with the cycle length, and therefore the curvature expansions are expected to be highly convergent.

☞ chapter 13

1.6 Evolution

The above derivation of the dynamical zeta function formula for the escape rate has one shortcoming; it estimates the fraction of survivors as a function of the number of pinball bounces, but the physically interesting quantity is the escape rate measured in units of continuous time. For continuous time flows, the escape rate (1.2) is generalized as follows. Define a finite phase space region $\mathcal{M}$ such that a trajectory that exits $\mathcal{M}$ never reenters. For example, any pinball that falls of the edge of a pinball table in fig. 1.1 is gone forever. Start with a uniform distribution of initial points. The fraction of initial x whose trajectories remain within $\mathcal{M}$ at time t is expected to decay exponentially

$$\Gamma(t) = \frac{\int_{\mathcal{M}} dx dy \, \delta(y - f^t(x))}{\int_{\mathcal{M}} dx} \;\to\; e^{-\gamma t} .$$

The integral over x starts a trajectory at every $x \in \mathcal{M}$. The integral over y tests whether this trajectory is still in $\mathcal{M}$ at time t. The kernel of this integral

$$\mathcal{L}^t(y, x) = \delta\big(y - f^t(x)\big) \tag{1.12}$$

is the Dirac delta function, as for a deterministic flow the initial point x maps into a unique point y at time t. For discrete time, $f^n(x)$ is the nth iterate of the map f. For continuous flows, $f^t(x)$ is the trajectory of the initial point x, and it is appropriate to express the finite time kernel $\mathcal{L}^t$ in terms of a generator of infinitesimal time translations

$$\mathcal{L}^t = e^{t\mathcal{A}} ,$$

☞ sect. 7.4.1
☞ chapter 26

very much in the way the quantum evolution is generated by the Hamiltonian H, the generator of infinitesimal time quantum transformations.

As the kernel $\mathcal{L}$ is the key to everything that follows, we shall give it a name, and refer to it and its generalizations as the *evolution operator* for a d-dimensional map or a d-dimensional flow.

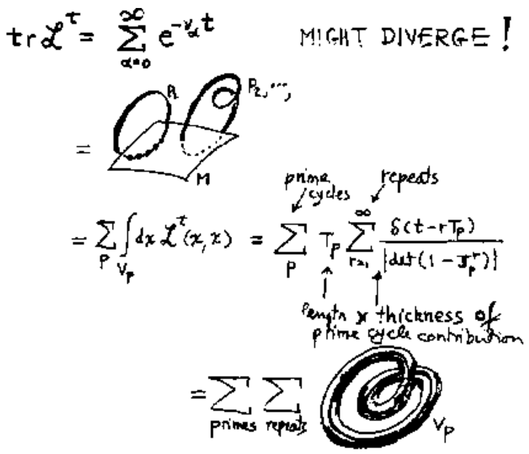


Figure 1.10: The trace of an evolution operator is concentrated in tubes around prime cycles, of length T_p and thickness $1/|\Lambda_p|^r$ for rth repeat of the prime cycle p.

The number of periodic points increases exponentially with the cycle length (in case at hand, as 2^n). As we have already seen, this exponential proliferation of cycles is not as dangerous as it might seem; as a matter of fact, all our computations will be carried out in the $n \to \infty$ limit. Though a quick look at chaotic dynamics might reveal it to be complex beyond belief, it is still generated by a simple deterministic law, and with some luck and insight, our labeling of possible motions will reflect this simplicity. If the rule that gets us from one level of the classification hierarchy to the next does not depend strongly on the level, the resulting hierarchy is approximately self-similar. We now turn such approximate self-similarity to our advantage, by turning it into an operation, the action of the evolution operator, whose iteration encodes the self-similarity.

1.6.1 Trace formula

Recasting dynamics in terms of evolution operators changes everything. So far our formulation has been heuristic, but in the evolution operator formalism the escape rate and any other dynamical average are given by exact formulas, extracted from the spectra of evolution operators. The key tools are the *trace formulas* and the *spectral determinants*.

The trace of an operator is given by the sum of its eigenvalues. The explicit expression (1.12) for $\mathcal{L}^t(x,y)$ enables us to evaluate the trace. Identify y with x and integrate x over the whole phase space. The result is an expression for $\mathrm{tr}\,\mathcal{L}^t$ as a sum over neighborhoods of prime cycles p and their repetitions

☞ sect. 11.3

$$\mathrm{tr}\,\mathcal{L}^t = \sum_p T_p \sum_{r=1}^{\infty} \frac{\delta(t - rT_p)}{\left|\det\left(\mathbf{1} - \mathbf{J}_p^r\right)\right|} \,. \qquad (1.13)$$

This formula has a simple geometrical interpretation sketched in fig. 1.10. After the rth return to a Poincaré section, the initial tube $\mathcal{M}_p$ has been stretched out along the expanding eigendirections, with the overlap with the initial volume given by $1/\left|\det\left(\mathbf{1} - \mathbf{J}_p^r\right)\right| \to 1/|\Lambda_p|$.

The "spiky" sum (1.13) is disquieting in the way reminiscent of the Poisson resummation formulas of Fourier analysis; the left-hand side is the smooth eigenvalue sum $\mathrm{tr}\, e^{\mathcal{A}} = \sum e^{s_\alpha t}$, while the right-hand side equals zero everywhere except for the set $t = rT_p$. A Laplace transform smooths the sum over Dirac delta functions in cycle periods and yields the *trace formula* for the eigenspectrum $s_0, s_1, \cdots$ of the classical evolution operator:

$$\begin{aligned}\int_{0_+}^{\infty} dt\, e^{-st}\, \mathrm{tr}\,\mathcal{L}^t &= \mathrm{tr}\,\frac{1}{s - \mathcal{A}} = \sum_{\alpha=0}^{\infty} \frac{1}{s - s_\alpha} \\ &= \sum_p T_p \sum_{r=1}^{\infty} \frac{e^{r(\beta \cdot A_p - sT_p)}}{\left|\det\left(\mathbf{1} - \mathbf{J}_p^r\right)\right|} \,. \end{aligned} \qquad (1.14)$$

The beauty of the trace formulas lies in the fact that everything on the right-hand-side – prime cycles p, their periods T_p and the stability eigenvalues of $\mathbf{J}_p$ – is an invariant property of the flow, independent of any coordinate choice. ☞ sect. 11.1

1.6.2 Spectral determinant

The eigenvalues of a linear operator are given by the zeros of the appropriate determinant. One way to evaluate determinants is to expand them in terms of traces, using the identities

✎ 4.1 page 81

$$\begin{aligned}\ln\det(s - \mathcal{A}) &= \mathrm{tr}\,\ln(s - \mathcal{A}) \\ \frac{d}{ds}\ln\det(s - \mathcal{A}) &= \mathrm{tr}\,\frac{1}{s - \mathcal{A}} \,,\end{aligned}$$

and integrating over s. In this way the *spectral determinant* of an evolution operator becomes related to the traces that we have just computed: ☞ chapter 12

$$\det(s - \mathcal{A}) = \exp\left(-\sum_p \sum_{r=1}^{\infty} \frac{1}{r} \frac{e^{-sT_p r}}{\left|\det\left(\mathbf{1} - \mathbf{J}_p^r\right)\right|}\right) \,. \qquad (1.15)$$

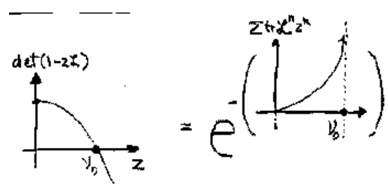


Figure 1.11: Spectral determinant is preferable to the trace as it vanishes smoothly at the leading eigenvalue, while the trace formula diverges.

The s integration leads here to replacement $T_p \to T_p/rT_p$ in the periodic orbit expansion (1.14).

The motivation for recasting the eigenvalue problem in this form is sketched in fig. 1.11; exponentiation improves analyticity and trades in a divergence of the trace sum for a zero of the spectral determinant. The computation of the zeros of $\det(s - \mathcal{A})$ proceeds very much like the computations of sect. 1.5.3.

☞ sect. 12.5.1

1.7 From chaos to statistical mechanics

While the above replacement of dynamics of individual trajectories by evolution operators which propagate densities might feel like just another bit of mathematical voodoo, actually something very radical has taken place. Consider a chaotic flow, such as stirring of red and white paint by some deterministic machine. *If* we were able to track individual trajectories, the fluid would forever remain a striated combination of pure white and pure red; there would be no pink. What is more, if we reversed stirring, we would return back to the perfect white/red separation. However, we know that this cannot be true – in a very few turns of the stirring stick the thickness of the layers goes from centimeters to Ångströms, and the result is irreversibly pink.

Understanding the distinction between evolution of individual trajectories and the evolution of the densities of trajectories is key to understanding statistical mechanics – this is the conceptual basis of the second law of thermodynamics, and the origin of irreversibility of the arrow of time for deterministic systems with time-reversible equations of motion: reversibility is attainable for distributions whose measure in the space of density functions goes exponentially to zero with time.

By going to a description in terms of the asymptotic time evolution operators we give up tracking individual trajectories for long times, but instead gain a very effective description of the asymptotic trajectory densities. This will enable us, for example, to give exact formulas for transport coefficients such as the diffusion constants without *any* probabilistic assumptions (such as the *stosszahlansatz* of Boltzmann).

☞ chapter 20

A century ago it seemed reasonable to assume that statistical mechanics applies only to systems with very many degrees of freedom. More recent is the realization that much of statistical mechanics follows from chaotic dynamics, and already at the level of a few degrees of freedom the evolution of densities is irreversible. Furthermore, the theory that we shall develop here generalizes notions of "measure" and "averaging" to systems far from equilibrium, and transports us into regions hitherto inaccessible with the tools of the equilibrium statistical mechanics.

The results of the equilibrium statistical mechanics do help us, however, to understand the ways in which the simple-minded periodic orbit theory falters. A non-hyperbolicity of the dynamics manifests itself in power-law correlations and even "phase transitions".

☞ chapter 18

1.8 Guide to literature

> But the power of instruction is seldom of much efficacy, except in those happy dispositions where it is almost superfluous.
>
> Gibbon

This text aims to bridge the gap between the physics and mathematics dynamical systems literature. The intended audience is the dream graduate student, with a theoretical bent. As a complementary presentation we recommend Gaspard's monograph [1.4] which covers much of the same ground in a highly readable and scholarly manner.

As far as the prerequisites are concerned - this book is not an introduction to nonlinear dynamics. Nonlinear science requires a one semester basic course (advanced undergraduate or first year graduate). A good start is the textbook by Strogatz [1.5], an introduction to flows, fixed points, manifolds, bifurcations. It is probably the most accessible introduction to nonlinear dynamics - it starts out with differential equations, and its broadly chosen examples and many exercises make it favorite with students. It is not strong on chaos. There the textbook of Alligood, Sauer and Yorke [1.6] is preferable: an elegant introduction to maps, chaos, period doubling, symbolic dynamics, fractals, dimensions - a good companion to this book. An introduction more comfortable to physicists is the textbook by Ott [1.7], with baker's map used to illustrate many key techniques in analysis of chaotic systems. It is perhaps harder than the above two as the first book on nonlinear dynamics.

An introductory course should give students skills in qualitative and numerical analysis of dynamical systems for short times (trajectories, fixed points, bifurcations) and familiarize them with Cantor sets and symbolic dynamics for chaotic dynamics. With this, and a graduate level exposure to statistical mechanics, partial differential equations and quantum mechanics, the stage is set for any of the one-semester advanced courses based on this book. The courses taught so far start out with the introductory chapters on qualitative dynamics, symbolic dynamics and flows, and then continue in different directions:

Deterministic chaos. Chaotic averaging, evolution operators, trace formulas, zeta functions, cycle expansions, Lyapunov exponents, billiards, transport coefficients, thermodynamic formalism, period doubling, renormalization operators.

Spatiotemporal dynamical systems. Partial differential equations for dissipative systems, weak amplitude expansions, normal forms, symmetries and bifurcations, pseudospectral methods, spatiotemporal chaos.

Quantum chaology. Semiclassical propagators, density of states, trace formulas, semiclassical spectral determinants, billiards, semiclassical helium, diffraction, creeping, tunneling, higher $\hbar$ corrections.

This book concentrates on the periodic orbit theory. The role of unstable periodic orbits was already fully appreciated by Poincaré [1.8, 1.9], who noted that hidden in the apparent chaos is a rigid skeleton, a tree of *cycles* (periodic orbits) of increasing lengths and self-similar structure, and suggested that the cycles should be the key to chaotic dynamics. Periodic orbits have been at core of much of the mathematical work on the theory of the classical and quantum dynamical systems ever since. We refer the reader to the reprint selection [1.10] for an overview of some of that literature.

If you find this book not rigorous enough, you should turn to the mathematics literature. The most extensive reference is the treatise by Katok and Hasselblatt [1.11], an impressive compendium of modern dynamical systems theory. The fundamental papers in this field, all still valuable reading, are Smale [1.12], Bowen [1.13] and Sinai [1.14]. Sinai's paper is prescient and offers a vision and a program that ties together dynamical systems and statistical mechanics. It is written for readers versed in statistical mechanics. Markov partitions were introduced by Sinai in ref. [1.15]. The classical text (though certainly not an easy read) on the subject of dynamical zeta functions is Ruelle's *Statistical Mechanics, Thermodynamic Formalism* [1.16]. In Ruelle's monograph transfer operator technique (or the "Perron-Frobenius theory") and Smale's theory of hyperbolic flows are applied to zeta functions and correlation functions. The status of the theory from Ruelle's point of view is compactly summarized in his 1995 Pisa lectures [1.17]. Further excellent mathematical references on thermodynamic formalism are Parry and Pollicott's monograph [1.18] with emphasis on the symbolic dynamics aspects of the formalism, and Baladi's clear and compact reviews of the theory dynamical zeta functions [1.19, 1.20].

A graduate level introduction to statistical mechanics from the dynamical point view is given by Dorfman [1.21]; the Gaspard monograph [1.4] covers the same ground in more depth. Driebe monograph [1.22] offers a nice introduction to the problem of irreversibility in dynamics. The role of "chaos" in statistical mechanics is critically dissected by Bricmont in his highly readable essay *"Science of Chaos or Chaos in Science?"* [1.23].

If you were wandering while reading this introduction "what's up with rat brains?", the answer is yes indeed, there is a line of research in study on neuronal dynamics that focuses on possible unstable periodic states, described for example in ref. [1.25, 1.26, 1.27].

Guide to exercises

God can afford to make mistakes. So can Dada!
Dadaist Manifesto

The essence of this subject is incommunicable in print; the only way to develop intuition about chaotic dynamics is by computing, and the reader is urged to try to work through the essential exercises. Some of the solutions provided might be more illuminating than the main text. So as not to fragment the text, the exercises are indicated by text margin boxes such as the one on this margin, and collected at the end of each chapter. The problems that you should do have **underlined titles**. The rest (**smaller type**) are optional. Difficult problems are marked by any number of *** stars. By the end of the course you should have completed at least three projects: (a) compute everything for a one-dimensional repeller, (b) compute escape rate for a 3-disk game of pinball, (c) compute a part of the quantum 3-disk game of pinball, or the helium spectrum, or if you are interested in statistical rather than the quantum mechanics, compute a transport coefficient. The essential steps are:

15.2 page 301

- **Dynamics**
  1. count prime cycles, exercise 1.1
  2. pinball simulator, exercise 5.1, exercise 14.4
  3. pinball stability, exercise 5.5, exercise 14.4
  4. pinball periodic orbits, exercise 14.6, exercise 14.5
  5. helium integrator, exercise 2.9, exercise 14.8
  6. helium periodic orbits, exercise 29.4, exercise 14.9

- **Averaging, numerical**

  1. pinball escape rate, exercise 12.12
  2. Lyapunov exponent, exercise 17.2

- **Averaging, periodic orbits**

  1. cycle expansions, exercise 15.1, exercise 15.2
  2. pinball escape rate, exercise 15.4, exercise 15.5
  3. cycle expansions for averages, exercise 15.1, exercise 16.3
  4. cycle expansions for diffusion, exercise 20.1
  5. desymmetrization exercise 19.1
  6. semiclassical quantization exercise 28.3
  7. ortho-, para-helium, lowest eigenenergies exercise 29.7

Solutions for some of the problems are included appendix O. Often going through a solution is more instructive than reading the corresponding chapter.

Résumé

The goal of this text is an exposition of the best of all possible theories of deterministic chaos, and the strategy is: 1) count, 2) weigh, 3) add up.

In a chaotic system any open ball of initial conditions, no matter how small, will spread over the entire accessible phase space. Hence the theory focuses on description of the geometry of the space of possible outcomes, and evaluation of averages over this space, rather than attempting the impossible, precise prediction of individual trajectories. The dynamics of distributions of trajectories is described in terms of evolution operators. In the evolution operator formalism the dynamical averages are given by exact formulas, extracted from the spectra of evolution operators. The key tools are the *trace formulas* and the *spectral determinants.*

The theory of evaluation of spectra of evolution operators presented here 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 on a strange attractor in the sense that any long orbit can approximately be pieced together from the nearby periodic orbits of finite length. This notion is made precise by approximating orbits by prime cycles, and evaluating associated curvatures. A curvature measures the deviation of a longer cycle from its approximation by shorter cycles; smoothness and the local instability of the flow implies exponential (or faster) fall-off for (almost) all curvatures. Cycle expansions offer then an efficient method for evaluating classical and quantum observables.

The critical step in the derivation of the dynamical zeta function was the hyperbolicity assumption, that is the assumption of exponential shrinkage of all strips of the pinball repeller. By dropping the a_i prefactors in (1.4), we have given up on any possibility of recovering the precise distribution of starting x (which should anyhow be impossible due to the exponential growth of errors), but in exchange we gain an effective description of the asymptotic behavior of the system. The pleasant surprise of cycle expansions (1.9) is that the infinite time behavior of an unstable system is as easy to determine as the short time behavior.

To keep exposition simple we have here illustrated the utility of cycles and their curvatures by a pinball game, but topics covered in this book – unstable flows, Poincaré sections, Smale horseshoes, symbolic dynamics, pruning, discrete symmetries, periodic orbits, averaging over chaotic sets, evolution operators, dynamical zeta functions, spectral determinants, cycle expansions, quantum trace formulas and zeta functions, and so on to the semiclassical quantization of helium – should give the reader some confidence in the general applicability of the theory. The formalism should work for any average over any chaotic set which satisfies two conditions:

1. the weight associated with the observable under consideration 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.

The theory is applicable to evaluation of a broad class of quantities characterizing chaotic systems, such as the escape rates, Lyapunov exponents, transport coefficients and quantum eigenvalues. One of the surprises is that the quantum mechanics of classically chaotic systems is very much like the classical mechanics of chaotic systems; both are described by nearly the same zeta functions and cycle expansions, with the same dependence on the topology of the classical flow.

References

[1.1] G.W. Leibniz, *Von dem Verhängnisse*

[1.2] P. Cvitanović, B. Eckhardt, P.E. Rosenqvist, G. Russberg and P. Scherer, in G. Casati and B. Chirikov, eds., *Quantum Chaos* (Cambridge University Press, Cambridge 1993).

[1.3] K.T. Hansen, *Symbolic Dynamics in Chaotic Systems*, Ph.D. thesis (Univ. of Oslo, 1994);
`www.nbi.dk/CATS/papers/khansen/thesis/thesis.html`

[1.4] P. Gaspard, *Chaos, Scattering and Statistical Mechanics* (Cambridge Univ. Press, Cambridge 1997).

[1.5] S.H. Strogatz, *Nonlinear Dynamics and Chaos* (Addison-Wesley 1994).

[1.6] K.T. Alligood, T.D. Sauer and J.A. Yorke, *Chaos, an Introduction to Dynamical Systems* (Springer, New York 1996)

[1.7] E. Ott, *Chaos in Dynamical Systems* (Cambridge Univ. Press, Cambridge 1993).

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

[1.9] For a very readable exposition of Poincaré's work and the development of the dynamical systems theory see J. Barrow-Green, *Poincaré and the Three Body Problem*, (Amer. Math. Soc., Providence R.I., 1997), and F. Diacu and P. Holmes, *Celestial Encounters, The Origins of Chaos and Stability* (Princeton Univ. Press, Princeton NJ 1996).

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

[1.11] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems* (Cambridge U. Press, Cambridge 1995).

[1.12] S. Smale, *Differentiable Dynamical Systems*, *Bull. Am. Math. Soc.* **73**, 747 (1967).

[1.13] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Springer Lecture Notes in Math. **470** (1975).

[1.14] Ya.G. Sinai, *Gibbs measures in ergodic theory*, *Russ. Math. Surveys* **166**, 21 (1972).

[1.15] Ya.G. Sinai, *"Construction of Markov partitions"*, *Funkts. Analiz i Ego Pril.* **2**, 70 (1968). English translation: *Functional Anal. Appl.* **2**, 245 (1968).

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

[1.17] D. Ruelle, *"Functional determinants related to dynamical systems and the thermodynamic formalism*, preprint IHES/P/95/30 (March 1995).

[1.18] W. Parry and M. Pollicott, *Zeta Functions and the periodic Structure of Hyperbolic Dynamics*, *Astérisque* **187–188** (Société Mathématique de France, Paris 1990).

[1.19] V. Baladi, *"Dynamical zeta functions"*, in B. Branner and P. Hjorth, eds., *Real and Complex Dynamical Systems* (Kluwer, Dordrecht, 1995).

[1.20] V. Baladi, *Positive Transfer Operators and Decay of Correlations* (World Scientific, Singapore 2000).

[1.21] R. Dorfman, *From Molecular Chaos to Dynamical Chaos* (Cambridge Univ. Press, Cambridge 1998).

[1.22] D.J. Driebe, *Fully Chaotic Map and Broken Time Symmetry* (Kluwer, Dordrecht, 1999).

[1.23] J. Bricmont, *"Science of Chaos or Chaos in Science?"*, available on `www.ma.utexas.edu/mp_arc, #96-116.`

[1.24] V.I. Arnold, *Mathematical Methods in Classical Mechanics* (Springer-Verlag, Berlin, 1978).

[1.25] S.J. Schiff, et al. *"Controlling chaos in the brain"*, *Nature* **370**, 615 (1994).

[1.26] F. Moss, "Chaos under control", *Nature* **370**, 615 (1994).

[1.27] J. Glanz,s *Science* **265**, 1174 (1994).

[1.28] Poul Martin Møller, *En dansk Students Eventyr [The Adventures of a Danish Student]* (Copenhagen 1824).