Classical and Quantum Chaos:

Part III: Material available on www.nbi.dk/ChaosBook

version 9.3.9 **Feb 15 2003** printed February 16, 2003
www.nbi.dk/ChaosBook comments to: predrag@nbi.dk

Appendix A

A brief history of chaos

Laws of attribution

1. **Arnold's Law**: everything that is discovered is named after someone else (including Arnold's law)
2. **Berry's Law**: sometimes, the sequence of antecedents seems endless. So, nothing is discovered for the first time.
3. **Whiteheads's Law**: Everything of importance has been said before by someone who did not discover it.

M.V. Berry

A.1 Chaos is born

(R. Mainieri)

Classical mechanics has not stood still since Newton. The formalism that we use today was developed by Euler and Lagrange. By the end of the 1800's the three problems that would lead to the notion of chaotic dynamics were already known: the three-body problem, the ergodic hypothesis, and nonlinear oscillators.

A.1.1 Three-body problem

Trying to predict the motion of the Moon has preoccupied astronomers since antiquity. Accurate understanding of its motion was important for determining the longitude of ships while traversing open seas. Kepler's Rudolphine tables had

been a great improvement over previous tables, and Kepler was justly proud of his achievements. He wrote in the introduction to the announcement of Kepler's third law, *Harmonice Mundi* (Linz, 1619) in a style that would not fly with the contemporary *Physical Review Letters* editors:

> What I prophesied two-and-twenty years ago, as soon as I discovered the five solids among the heavenly orbits – what I firmly believed long before I had seen Ptolemy's *Harmonics* – what I had promised my friends in the title of this book, which I named before I was sure of my discovery – what sixteen years ago, I urged as the thing to be sought – that for which I joined Tycho Brahé, for which I settled in Prague, for which I have devoted the best part of my life to astronomical contemplations, at length I have brought to light, and recognized its truth beyond my most sanguine expectations. It is not eighteen months since I got the first glimpse of light, three months since the dawn, very few days since the unveiled sun, most admirable to gaze upon, burst upon me. Nothing holds me; I will indulge my sacred fury; I will triumph over mankind by the honest confession that I have stolen the golden vases of the Egyptians to build up a tabernacle for my God far away from the confines of Egypt. If you forgive me, I rejoice; if you are angry, I can bear it; the die is cast, the book is written, to be read either now or in posterity, I care not which; it may well wait a century for a reader, as God has waited six thousand years for an observer.

Bernoulli used Newton's work on mechanics to derive the elliptic orbits of Kepler and set an example of how equations of motion could be solved by integrating. But the motion of the Moon is not well approximated by an ellipse with the Earth at a focus; at least the effects of the Sun have to be taken into account if one wants to reproduce the data the classical Greeks already possessed. To do that one has to consider the motion of three bodies: the Moon, the Earth, and the Sun. When the planets are replaced by point particles of arbitrary masses, the problem to be solved is known as the three-body problem. The three-body problem was also a model to another concern in astronomy. In the Newtonian model of the solar system it is possible for one of the planets to go from an elliptic orbit around the Sun to an orbit that escaped its domain or that plunged right into it. Knowing if any of the planets would do so became the problem of the stability of the solar system. A planet would not meet this terrible end if solar system consisted of two celestial bodies, but whether such fate could befall in the three-body case remained unclear.

After many failed attempts to solve the three-body problem, natural philosophers started to suspect that it was impossible to integrate. The usual technique for integrating problems was to find the conserved quantities, quantities that do not change with time and allow one to relate the momenta and positions different times. The first sign on the impossibility of integrating the three-body problem came from a result of Burns that showed that there were no conserved quantities that were polynomial in the momenta and positions. Burns' result did not

preclude the possibility of more complicated conserved quantities. This problem was settled by Poincaré and Sundman in two very different ways.

In an attempt to promote the journal *Acta Mathematica*, Mittag-Leffler got the permission of the King Oscar II of Sweden and Norway to establish a mathematical competition. Several questions were posed (although the king would have preferred only one), and the prize of 2500 kroner would go to the best submission. One of the questions was formulated by Weierstrass:

> Given a system of arbitrary mass points that attract each other according to Newton's laws, under the assumption that no two points ever collide, try to find a representation of the coordinates of each point as a series in a variable that is some known function of time and for all of whose values the series converges uniformly.
>
> This problem, whose solution would considerably extend our understanding of the solar system, ...

Poincaré's submission won the prize. He showed that conserved quantities that were analytic in the momenta and positions could not exist. To show that he introduced methods that were very geometrical in spirit: the importance of phase flow, the role of periodic orbits and their cross sections, the homoclinic points.

The interesting thing about Poincaré's work was that it did not solve the problem posed. He did not find a function that would give the coordinates as a function of time for all times. He did not show that it was impossible either, but rather that it could not be done with the Bernoulli technique of finding a conserved quantity and trying to integrate. Integration would seem unlikely from Poincaré's prize-winning memoir, but it was accomplished by the Finnish-born Swedish mathematician Sundman. Sundman showed that to integrate the three-body problem one had to confront the two-body collisions. He did that by making them go away through a trick known as regularization of the collision manifold. The trick is not to expand the coordinates as a function of time t, but rather as a function of $\sqrt[3]{t}$. To solve the problem for all times he used a conformal map into a strip. This allowed Sundman to obtain a series expansion for the coordinates valid for all times, solving the problem that was proposed by Weirstrass in the King Oscar II's competition.

The Sundman's series are not used today to compute the trajectories of any three-body system. That is more simply accomplished by numerical methods or through series that, although divergent, produce better numerical results. The conformal map and the collision regularization mean that the series are effectively in the variable $1 - e^{-\sqrt[3]{t}}$. Quite rapidly this gets exponentially close to one, the radius of convergence of the series. Many terms, more terms than any one has ever wanted to compute, are needed to achieve numerical convergence. Though Sundman's work deserves better credit than it gets, it did not live up to

Weirstrass's expectations, and the series solution did not "considerably extend our understanding of the solar system." The work that followed from Poincaré did.

A.1.2 Ergodic hypothesis

The second problem that played a key role in development of chaotic dynamics was the ergodic hypothesis of Boltzmann. Maxwell and Boltzmann had combined the mechanics of Newton with notions of probability in order to create statistical mechanics, deriving thermodynamics from the equations of mechanics. To evaluate the heat capacity of even a simple system, Boltzmann had to make a great simplifying assumption of ergodicity: that the dynamical system would visit every part of the phase space allowed by conservations law equally often. This hypothesis was extended to other averages used in statistical mechanics and was called the ergodic hypothesis. It was reformulated by Poincaré to say that a trajectory comes as close as desired to any phase space point.

Proving the ergodic hypothesis turned out to be very difficult. By the end of our own century it has only been shown true for a few systems and wrong for quite a few others. Early on, as a mathematical necessity, the proof of the hypothesis was broken down into two parts. First one would show that the mechanical system was ergodic (it would go near any point) and then one would show that it would go near each point equally often and regularly so that the computed averages made mathematical sense. Koopman took the first step in proving the ergodic hypothesis when he noticed that it was possible to reformulate it using the recently developed methods of Hilbert spaces. This was an important step that showed that it was possible to take a finite-dimensional nonlinear problem and reformulate it as a infinite-dimensional linear problem. This does not make the problem easier, but it does allow one to use a different set of mathematical tools on the problem. Shortly after Koopman started lecturing on his method, von Neumann proved a version of the ergodic hypothesis, giving it the status of a theorem. He proved that if the mechanical system was ergodic, then the computed averages would make sense. Soon afterwards Birkhoff published a much stronger version of the theorem.

A.1.3 Nonlinear oscillators

The third problem that was very influential in the development of the theory of chaotic dynamical systems was the work on the nonlinear oscillators. The problem is to construct mechanical models that would aid our understanding of physical systems. Lord Rayleigh came to the problem through his interest in understanding how musical instruments generate sound. In the first approximation one can construct a model of a musical instrument as a linear oscillator. But real

instruments do not produce a simple tone forever as the linear oscillator does, so Lord Rayleigh modified this simple model by adding friction and more realistic models for the spring. By a clever use of negative friction he created two basic models for the musical instruments. These models have more than a pure tone and decay with time when not stroked. In his book *The Theory of Sound* Lord Rayleigh introduced a series of methods that would prove quite general, such as the notion of a limit cycle, a periodic motion a system goes to regardless of the initial conditions.

A.2 Chaos grows up

(R. Mainieri)

The theorems of von Neumann and Birkhoff on the ergodic hypothesis were published in 1912 and 1913. This line of enquiry developed in two directions. One direction took an abstract approach and considered dynamical systems as transformations of measurable spaces into themselves. Could we classify these transformations in a meaningful way? This lead Kolmogorov to the introduction of the concept of entropy for dynamical systems. With entropy as a dynamical invariant it became possible to classify a set of abstract dynamical systems known as the Bernoulli systems. The other line that developed from the ergodic hypothesis was in trying to find mechanical systems that are ergodic. An ergodic system could not have stable orbits, as these would break ergodicity. So in 1898 Hadamard published a paper with a playful title of "... billiards ...," where he showed that the motion of balls on surfaces of constant negative curvature is everywhere unstable. This dynamical system was to prove very useful and it was taken up by Birkhoff. Morse in 1923 showed that it was possible to enumerate the orbits of a ball on a surface of constant negative curvature. He did this by introducing a symbolic code to each orbit and showed that the number of possible codes grew exponentially with the length of the code. With contributions by Artin, Hedlund, and Hopf it was eventually proven that the motion of a ball on a surface of constant negative curvature was ergodic. The importance of this result escaped most physicists, one exception being Krylov, who understood that a physical billiard was a dynamical system on a surface of negative curvature, but with the curvature concentrated along the lines of collision. Sinai, who was the first to show that a physical billiard can be ergodic, knew Krylov's work well.

The work of Lord Rayleigh also received vigorous development. It prompted many experiments and some theoretical development by van der Pol, Duffing, and Hayashi. They found other systems in which the nonlinear oscillator played a role and classified the possible motions of these systems. This concreteness of experiments, and the possibility of analysis was too much of temptation for Mary Lucy Cartwright and J.E. Littlewood, who set out to prove that many of

the structures conjectured by the experimentalists and theoretical physicists did indeed follow from the equations of motion. Birkhoff had found a "remarkable curve" in a two dimensional map; it appeared to be non-differentiable and it would be nice to see if a smooth flow could generate such a curve. The work of Cartwright and Littlewood lead to the work of Levinson, which in turn provided the basis for the horseshoe construction of Smale.

In Russia, Lyapunov paralleled the methods of Poincaré and initiated the strong Russian dynamical systems school. Andronov carried on with the study of nonlinear oscillators and in 1937 introduced together with Pontryagin the notion of coarse systems. They were formalizing the understanding garnered from the study of nonlinear oscillators, the understanding that many of the details on how these oscillators work do not affect the overall picture of the phase space: there will still be limit cycles if one changes the dissipation or spring force function by a little bit. And changing the system a little bit has the great advantage of eliminating exceptional cases in the mathematical analysis. Coarse systems were the concept that caught Smale's attention and enticed him to study dynamical systems.

A.3 Chaos with us

(R. Mainieri)

In the fall of 1961 Steven Smale was invited to Kiev where he met Arnold, Anosov, Sinai, and Novikov. He lectured there, and spent a lot of time with Anosov. He suggested a series of conjectures, most of which Anosov proved within a year. It was Anosov who showed that there are dynamical systems for which all points (as opposed to a non–wandering set) admit the hyperbolic structure, and it was in honor of this result that Smale named these systems Axiom-A. In Kiev Smale found a receptive audience that had been thinking about these problems. Smale's result catalyzed their thoughts and initiated a chain of developments that persisted into the 1970's.

Smale collected his results and their development in the 1967 review article on dynamical systems, entitled "Differentiable dynamical systems". There are many great ideas in this paper: the global foliation of invariant sets of the map into disjoint stable and unstable parts; the existence of a horseshoe and enumeration and ordering of all its orbits; the use of zeta functions to study dynamical systems. The emphasis of the paper is on the global properties of the dynamical system, on how to understand the topology of the orbits. Smale's account takes you from a local differential equation (in the form of vector fields) to the global topological description in terms of horseshoes.

The path traversed from ergodicity to entropy is a little more confusing. The

general character of entropy was understood by Weiner, who seemed to have spoken to Shannon. In 1948 Shannon published his results on information theory, where he discusses the entropy of the shift transformation. Kolmogorov went far beyond and suggested a definition of the metric entropy of an area preserving transformation in order to classify Bernoulli shifts. The suggestion was taken by his student Sinai and the results published in 1959. In 1960 Rohlin connected these results to measure-theoretical notions of entropy. The next step was published in 1965 by Adler and Palis, and also Adler, Konheim, McAndrew; these papers showed that one could define the notion of topological entropy and use it as an invariant to classify continuous maps. In 1967 Anosov and Sinai applied the notion of entropy to the study of dynamical systems. It was in the context of studying the entropy associated to a dynamical system that Sinai introduced Markov partitions in 1968.

Markov partitions allow one to relate dynamical systems and statistical mechanics; this has been a very fruitful relationship. It adds measure notions to the topological framework laid down in Smale's paper. Markov partitions divide the phase space of the dynamical system into nice little boxes that map into each other. Each box is labeled by a code and the dynamics on the phase space maps the codes around, inducing a symbolic dynamics. From the number of boxes needed to cover all the space, Sinai was able to define the notion of entropy of a dynamical system. In 1970 Bowen came up independently with the same ideas, although there was presumably some flow of information back and forth before these papers got published. Bowen also introduced the important concept of shadowing of chaotic orbits. We do not know whether at this point the relations with statistical mechanics were clear to every one. They became explicit in the work of Ruelle. Ruelle understood that the topology of the orbits could be specified by a symbolic code, and that one could associate an "energy" to each orbit. The energies could be formally combined in a "partition function" to generate the invariant measure of the system.

After Smale, Sinai, Bowen, and Ruelle had laid the foundations of the statistical mechanics approach to chaotic systems, research turned to studying particular cases. The simplest case to consider is one-dimensional maps. The topology of the orbits for parabola-like maps was worked out in 1973 by Metropolis, Stein, and Stein. The more general one-dimensional case was worked out in 1976 by Milnor and Thurston in a widely circulated preprint, whose extended version eventually got published in 1988.

A lecture of Smale and the results of Metropolis, Stein, and Stein inspired Feigenbaum to study simple maps. This lead him to the discovery of the universality in quadratic maps and the application of ideas from field-theory to dynamical systems. Feigenbaum's work was the culmination in the study of one-dimensional systems; a complete analysis of a nontrivial transition to chaos. Feigenbaum introduced many new ideas into the field: the use of the renormalization group which lead him to introduce functional equations in the study of dynamical sys-

tems, the scaling function which completed the link between dynamical systems and statistical mechanics, and the use of presentation functions as the dynamics of scaling functions.

The work in more than one dimension progressed very slowly and is still far from completed. The first result in trying to understand the topology of the orbits in two dimensions (the equivalent of Metropolis, Stein, and Stein, or Milnor and Thurston's work) was obtained by Thurston. Around 1975 Thurston was giving lectures "On the geometry and dynamics of diffeomorphisms of surfaces". Thurston's techniques exposed in that lecture have not been applied in physics, but much of the classification that Thurston developed can be obtained from the notion of a "pruning front" developed independently by Cvitanović.

Once one develops an understanding for the topology of the orbits of a dynamical system, one needs to be able to compute its properties. Ruelle had already generalized the zeta function introduced by Artin and Mazur so that it could be used to compute the average value of observables. The difficulty with Ruelle's zeta function is that it does not converge very well. Starting out from Smale's observation that a chaotic dynamical system is dense with a set of periodic orbits, Cvitanović used these orbits as a skeleton on which to evaluate the averages of observables, and organized such calculations in terms of rapidly converging cycle expansions. This convergence is attained by using the shorter orbits used as a basis for shadowing the longer orbits.

This account is far from complete, but we hope that it will help get a sense of perspective on the field. It is not a fad and it will not die anytime soon.

Remark A.1 Notion of global foliations. For each paper cited in dynamical systems literature, there are many results that went into its development. As an example, take the notion of global foliations that we attribute to Smale. As far as we can trace the idea, it goes back to René Thom; local foliations were already used by Hadamard. Smale attended a seminar of Thom in 1958 or 1959. In that seminar Thom was explaining his notion of transversality. One of Thom's disciples introduced Smale to Brazilian mathematician Peixoto. Peixoto (who had learned the results of the Andronov-Pontryagin school from Lefschetz) was the closest Smale had ever come until then to the Andronov-Pontryagin school. It was from Peixoto that Smale learned about structural stability, a notion that got him enthusiastic about dynamical systems, as it blended well with his topological background. It was from discussions with Peixoto that Smale got the problems in dynamical systems that lead him to his 1960 paper on Morse inequalities. The next year Smale published his result on the hyperbolic structure of the nonwandering set. Smale was not the first to consider a hyperbolic point, Poincaré had already done that; but Smale was the first to introduce a global hyperbolic structure. By 1960 Smale was already lecturing on the horseshoe as a structurally stable dynamical system with an infinity of periodic points and promoting his global viewpoint.

(R. Mainieri)

Remark A.2 Levels of ergodicity. In the mid 1970's A. Katok and Ya.B. Pesin tried to use geometry to establish positive Lyapunov exponents. A. Katok and J.-M. Strelcyn carried out the program and developed a theory of general dynamical systems with singularities. They studied uniformly hyperbolic systems (as strong as Anosov's), but with sets of singularities. Under iterations a dense set of points hits the singularities. Even more important are the points that never hit the singularity set. In order to establish some control over how they approach the set, one looks at trajectories that apporach the set by some given ϵ^n, or faster.

Ya.G. Sinai, L. Bunimovich and Chernov studied the geometry of billiards in a very detailed way. A. Katok and Ya.B. Pesin's idea was much more robust. Look at the discontinuity set (geometry of it matters not at all), take an ϵ neighborhood around it. Given that the Lebesgue measure is ϵ^α and the stability grows not faster than $(\text{distance})^n$, A. Katok and J.-M. Strelcyn prove that the Lyapunov exponent is non-zero.

In mid 1980's Ya.B. Pesin studied the dissipative case. Now the problem has no invariant Lebesgue measure. Assuming uniform hyperbolicity, with singularities, and tying together Lebesgue measure and discontinutities, and given that the stability grows not faster than $(\text{distance})^n$, Ya.B. Pesin proved that the Lyapunov exponent is non-zero, and that SBR measure exists. He also proved that the Lorenz, Lozi and Byelikh attractors satisfy these conditions.

In the the systems were uniformly hyperbolic, all trouble was in differentials. For the Hénon attractor, already the differentials are nonhyperbolic. The points do not separate uniformly, but the analogue of the singularity set can be obtained by excizing the regions that do not separate. Hence there are 3 levels of ergodic systems:

1. Anosov flow
2. Anosov flow + singularity set
   - the Hamiltonian systems: general case A. Katok and J.-M. Strelcyn, billiards Ya.G. Sinai and L. Bunimovich.
   - the dissipative case: Ya.B. Pesin
3. Hénon
   - The first proof was given by M. Benedicks and L. Carleson [?].
   - A more readable proof is given in M. Benedicks and L.-S. Young [?]

(based on Ya.B. Pesin's comments)

A.3.1 Periodic orbit theory

The history of the periodic orbit theory is rich and curious, and the recent advances are to equal degree inspired by a century of separate development of three disparate subjects; 1. *classical chaotic dynamics*, initiated by Poincaré and put on its modern footing by Smale, Ruelle, and many others; 2. *quantum theory* initiated by Bohr, with the modern "chaotic" formulation by Gutzwiller; and 3. analytic *number theory* initiated by Riemann and formulated as a spectral problem by Selberg. Following totally different lines of reasoning and driven by

very different motivations, the three separate roads all arrive at formally nearly identical *trace formulas*, *zeta functions* and *spectral determinants*.

That these topics should be related is far from obvious. Connection between dynamics and number theory arises from Selberg's observation that description of geodesic motion and wave mechanics on spaces of constant negative curvature is essentially a number-theoretic problem. *A posteriori*, one can say that zeta functions arise in both classical and quantum mechanics because in both the dynamical evolution can be described by the action of linear evolution (or transfer) operators on infinite-dimensional vector spaces. The spectra of these operators are given by the zeros of appropriate determinants. One way to evaluate determinants is to expand them in terms of traces, $\log \det = \mathrm{tr} \log$, and in this way the spectrum of an evolution operator becames related to its traces, that is, periodic orbits. A perhaps deeper way of restating this is to observe that the trace formulas perform the same service in all of the above problems; they relate the spectrum of lengths (local dynamics) to the spectrum of eigenvalues (global averages), and for nonlinear geometries they play a role analogous to that the Fourier transform plays for the circle.

A.4 Death of the Old Quantum Theory

> In 1913 Otto Stern and Max Theodor Felix von Laue went up for a walk up the Uetliberg. On the top they sat down and talked about physics. In particular they talked about the new atom model of Bohr. There and then they made the "Uetli Schwur": If that crazy model of Bohr turned out to be right, then they would leave physics. It did and they didn't.
>
> A. Pais, *Inward Bound: of Matter and Forces in the Physical World*

In an afternoon of May 1991 Dieter Wintgen is sitting in his office at the Niels Bohr Institute beaming with the unparalleled glee of a boy who has just committed a major mischief. The starting words of the manuscript he has just penned are

> The failure of the Copenhagen School to obtain a reasonable ...

34 years old at the time, Dieter was a scruffy kind of guy, always in sandals and holed out jeans, a left winger and a mountain climber, working around the clock with his students Gregor and Klaus to complete the work that Bohr himself would have loved to see done back in 1916: a "planetary" calculation of the helium spectrum.

Never mind that the "Copenhagen School" refers not to the old quantum theory, but to something else. The old quantum theory was no theory at all; it was a set of rules bringing some order to a set of phenomena which defied logic of classical theory. The electrons were supposed to describe planetary orbits around the nucleus; their wave aspects were yet to be discovered. The foundations seemed obscure, but Bohr's answer for the once-ionized helium to hydrogen ratio was correct to five significant figures and hard to ignore. The old quantum theory marched on, until by 1924 it reached an impasse: the helium spectrum and the Zeeman effect were its death knell.

Since the late 1890's it had been known that the helium spectrum consists of the orthohelium and parahelium lines. In 1915 Bohr suggested that the two kinds of helium lines might be associated with two distinct shapes of orbits (a suggestion that turned out to be wrong). In 1916 he got Kramers to work on the problem, and wrote to Rutherford: "I have used all my spare time in the last months to make a serious attempt to solve the problem of ordinary helium spectrum ...I think really that at last I have a clue to the problem." To other colleagues he wrote that "the theory was worked out in the fall of 1916" and of having obtained a "partial agreement with the measurements." Nevertheless, the Bohr-Sommerfeld theory, while by and large successful for hydrogen, was a disaster for neutral helium. Heroic efforts of the young generation, including Kramers and Heisenberg, were of no avail.

For a while Heisenberg thought that he had the ionization potential for helium, which he had obtained by a simple perturbative scheme. He wrote enthusiastic letters to Sommerfeld and was drawn into a collaboration with Max Born to compute the spectrum of helium using Born's systematic perturbative scheme. In first approximation, they reproduced the earlier calculations. The next level of corrections turned out to be larger than the computed effect. The concluding paragraph of Max Born's classic "Vorlesungen über Atommechanik" from 1925 sums it up in a somber tone:

> (...) the systematic application of the principles of the quantum theory (...) gives results in agreement with experiment only in those cases where the motion of a single electron is considered; it fails even in the treatment of the motion of the two electrons in the helium atom.
>
> This is not surprising, for the principles used are not really consistent. (...) A complete systematic transformation of the classical mechanics into a discontinuous mechanics is the goal towards which the quantum theory strives.

That year Heisenberg suffered a bout of hay fever, and the old quantum theory was dead. In 1926 he gave the first quantitative explanation of the helium spectrum. He used wave mechanics, electron spin and the Pauli exclusion principle, none of which belonged to the old quantum theory, and planetary orbits of electrons were cast away for nearly half a century.

Why did Pauli and Heisenberg fail with the helium atom? It was not the fault of the old quantum mechanics, but rather it reflected their lack of understanding of the subtleties of classical mechanics. Today we know what they missed in 1913-24: the role of conjugate points (topological indices) along classical trajectories was not accounted for, and they had no idea of the importance of periodic orbits in nonintegrable systems.

Since then the calculation for helium using the methods of the old quantum mechanics has been fixed. Leopold and Percival added the topological indices in 1980, and in 1991 Wintgen and collaborators orbits. Dieter had good reasons to gloat; while the rest of us were preparing to sharpen our pencils and supercomputers in order to approach the dreaded 3-body problem, they just went ahead and did it. What it took - and much else - is described in this book. One is also free to ponder what quantum theory would look like today if all this was worked out in 1917.

Remark A.3 Sources. This tale, aside from a few personal recollections, is in large part lifted from Abraham Pais' accounts of the demise of the old quantum theory [?, ?], as well as Jammer's account [1]. The helium spectrum is taken up in chapter 25. In August 1994 Dieter Wintgen died in a climbing accident in the Swiss Alps.

Appendix B

Infinite-dimensional flows

Flows described by partial differential equations are considered infinite dimensional because if one writes them down as a set of ordinary differential equations (ODE) then one needs an infinity of the ordinary kind to represent the dynamics of one equation of the partial kind (PDE). Even though the phase space is infinite dimensional, for many systems of physical interest the global attractor is finite dimensional. We illustrate how this works with a concrete example, the Kuramoto-Sivashinsky system.

B.0.1 Partial differential equations

First, a few words about partial differential equations in general. Many of the partial differential equations of mathematical physics can be written in the quasi-linear form

$$\partial_t u = \mathbf{A} u + N(u)\,, \tag{B.1}$$

where u is a function (possibly a vector function) of the coordinate x and time t, $\mathbf{A}$ is a linear operator, usually containing the Laplacian and a few other derivatives of u, and $N(u)$ is the nonlinear part of the equation (terms like $u\partial_x u$ in (2.14)).

Not all equations are stated in the form (B.1), but they can easily be so transformed, just as the ordinary differential equations can be rewritten as first-order systems. We will illustrate the method with a variant of the D'Alambert's wave equation describing a plucked string:

$$\partial_{tt} y = \left(c + \frac{1}{2} (\partial_x y)^2 \right) \partial_{xx} y\,. \tag{B.2}$$

Were the term $\partial_x y$ small, this equation would be just the ordinary wave equation. To rewrite the equation in the first order form (B.1), we need a field $u = (y, w)$ that is two-dimensional,

$$\partial_t \begin{bmatrix} y \\ w \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ c\partial_{xx} & 0 \end{bmatrix} \begin{bmatrix} y \\ w \end{bmatrix} + \begin{bmatrix} 0 \\ \partial_{xx} y(\partial_x y)^2/2 \end{bmatrix} . \tag{B.3}$$

The [2×2] matrix is the linear operator $\mathbf{A}$ and the vector on the far right is the nonlinear function $N(u)$. Unlike ordinary functions, differentiations are part of the function. The nonlinear part can also be expressed as a function on the infinite set of numbers that represent the field, as exemplified by the Kuramoto-Sivashinsky system (2.14).

☞ chapter 4.2

The usual technique for solving the linear part is to use Fourier methods. Just as in the ordinary differential equation case, one can integrate the linear part of

$$\partial_t u = \mathbf{A} u \tag{B.4}$$

to obtain

$$u(x,t) = e^{t\mathbf{A}} u(x,0) . \tag{B.5}$$

If u is expressed as Fourier series $\sum_k a_k \exp(ikx)$, as we will do for the Kuramoto-Shivashinsky system, then we can determine the action of $e^{t\mathbf{A}}$ on $u(x,0)$. This can be done because differentiations in $\mathbf{A}$ act rather simply on the exponentials. For example,

$$e^{t\partial_x} u(x,0) = e^{t\partial_x} \sum_k a_k e^{ikx} = \sum_k a_k \frac{(it)^k}{k!} e^{ikx} . \tag{B.6}$$

Depending on the behavior of the linear part, one distinguishes three classes of partial differential equations: *diffusion, wave, and potential.* The classification relies on the solution by a Fourier series, as in (B.5). In mathematical literature these equations are also called *parabolic, hyperbolic* and *elliptic.* If the nonlinear part $N(u)$ is as big as the linear part, the classification is not a good indication of behavior, and one can encounter features of one class of equations while studying the others.

In *diffusion-type* equations the modes of high frequency tend to become smooth, and all initial conditions tend to an attractor, called the *inertial manifold.* The Kuramoto-Sivashinsky system studied below is of this type. The solution being

attracted to the inertial manifold does not mean that the amplitudes of all but a finite number of modes go to zero (alas were we so lucky), but that there is a finite set of modes that could be used to describe any solution of the inertial manifold. The only catch is that there is no simple way to discover what these inertial manifold modes might be.

In *wave-like* equations the high frequency modes do not die out and the solutions tend to be distributions. The equations can be solved by variations on the WKB idea: the wave-like equations can be approximated by the trajectories of the wave fronts.

☞ chapter 23

2.10 on p. 50

Elliptic equations have no time dependence and do not represent dynamical systems.

Appendix C

Stability of Hamiltonian flows

C.1 Symplectic invariance

(M.J. Feigenbaum and P. Cvitanović)

The symplectic structure of Hamilton's equations buys us much more than the incompressibility, or the phase space volume conservation alluded to in sect. 5.1. We assume you are at home with Hamiltonian formalism. If you would like to see the Hamilton's equations derived, Hamilton's original line of reasoning is retraced in sect. 23.2.1. The evolution equations for any p, q dependent quantity $Q = Q(p, q)$ are given by (7.33).

☞ sect. 23.2.1

In terms of the Poisson brackets, the time evolution equation for $Q = Q(p, q)$ is given by (7.35). We now recast the symplectic condition (5.8) in a form convenient for using the symplectic constraints on $\mathbf{J}$. Writing $x(t) = x' = [p', q']$ and the Jacobian matrix and its inverse

$$\mathbf{J} = \begin{pmatrix} \frac{\partial q'}{\partial q} & \frac{\partial q'}{\partial p} \\ \frac{\partial p'}{\partial q} & \frac{\partial p'}{\partial p} \end{pmatrix}, \qquad \mathbf{J}^{-1} = \begin{pmatrix} \frac{\partial q}{\partial q'} & \frac{\partial q}{\partial p'} \\ \frac{\partial p}{\partial q'} & \frac{\partial p}{\partial p'} \end{pmatrix}, \tag{C.1}$$

we can spell out the symplectic invariance condition (5.8):

$$\begin{aligned} \frac{\partial q'_k}{\partial q_i}\frac{\partial p'_k}{\partial q_j} - \frac{\partial p'_k}{\partial q_i}\frac{\partial q'_k}{\partial q_j} &= 0 \\ \frac{\partial q'_k}{\partial p_i}\frac{\partial p'_k}{\partial p_j} - \frac{\partial p'_k}{\partial p_i}\frac{\partial q'_k}{\partial p_j} &= 0 \\ \frac{\partial q'_k}{\partial q_i}\frac{\partial p'_k}{\partial p_j} - \frac{\partial p'_k}{\partial q_i}\frac{\partial q'_k}{\partial p_j} &= \delta_{ij}\,. \end{aligned} \tag{C.2}$$

From (5.9) we obtain

$$\frac{\partial q_i}{\partial q'_j} = \frac{\partial p'_j}{\partial p_i}\,, \quad \frac{\partial p_i}{\partial p'_j} = \frac{\partial q'_j}{\partial q_i}\,, \quad \frac{\partial q_i}{\partial p'_j} = -\frac{\partial q'_j}{\partial p_i}\,, \quad \frac{\partial p_i}{\partial q'_j} = -\frac{\partial p'_j}{\partial q_i}\,. \tag{C.3}$$

Taken together, (C.3) and (C.2) imply that the flow conserves the $[p,q]$ Poisson brackets

$$\begin{aligned} [q_i, q_j] &= \frac{\partial q_i}{\partial p'_k}\frac{\partial q_j}{\partial q'_k} - \frac{\partial q_j}{\partial p'_k}\frac{\partial q_i}{\partial q'_k} = 0 \\ [p_i, p_j] &= 0\,, \qquad [p_i, q_j] = \delta_{ij}\,, \end{aligned} \tag{C.4}$$

that is, the transformations induced by a Hamiltonian flow are *canonical*, preserving the form of the equations of motion. The first two relations are symmetric under i, j interchange and yield $D(D-1)/2$ constraints each; the last relation yields D^2 constraints. Hence only $(2D)^2 - 2D(D-1)/2 - D^2 = 2D^2 + D$ elements of $\mathbf{J}$ are linearly independent, as it behooves group elements of the symplectic group $Sp(2D)$.

We have now succeeded in making the full set of constraints explicit - as we shall see in appendix D, this will enable us to implement dynamics in such a way that the symplectic invariance will be automatically preserved.

Consider the symplectic product of two infinitesimal vectors

$$\begin{aligned} (\delta x, \delta\hat{x}) &= \delta x^T \omega \delta\hat{x} = \delta p_i \delta\hat{q}_i - \delta q_i \delta\hat{p}_i \\ &= \sum_{i=1}^{D} \{\text{oriented area in the } (p_i, q_i) \text{ plane}\}\,. \end{aligned} \tag{C.5}$$

Time t later we have

$$(\delta x', \delta\hat{x}') = \delta x^T \mathbf{J}^T \omega \mathbf{J} \delta\hat{x} = \delta x^T \omega \delta\hat{x}\,.$$

This has the following geometrical meaning. We imagine there is a reference phase space point. We then define two other points infinitesimally close so that the vectors δx and $\delta\hat{x}$ describe their displacements relative to the reference point. Under the dynamics, the three points are mapped to three new points which are still infinitesimally close to one another. The meaning of the above expression is that the symplectic area spanned by the three final points is the same as that spanned by the inital points. The integral (Stokes theorem) version of this infinitesimal area invariance states that for Hamiltonian flows the D oriented

areas $\mathcal{V}_i$ bounded by D loops $\Omega\mathcal{V}_i$, one per each (p_i, q_i) plane, are separately conserved:

$$\int_{\mathcal{V}} dp \wedge dq = \oint_{\Omega\mathcal{V}} p \cdot dq = \text{invariant}\,. \tag{C.6}$$

C.2 Monodromy matrix for Hamiltonian flows

(G. Tanner)

It is not the Jacobian matrix of the flow, but the *monodromy* matrix, which enters the trace formula. This matrix gives the time dependence of a displacement perpendicular to the flow on the energy manifold. Indeed, we discover some trivial parts in the Jacobian matrix $\mathbf{J}$. An initial displacement in the direction of the flow $x = \omega\nabla H(x)$ transfers according to $\delta x(t) = x_l(t)\delta t$ with δt time independent. The projection of any displacement on δx on $\nabla H(x)$ is constant, i.e. $\nabla H(x(t))\delta x(t) = \delta E$. We get the equations of motion for the monodromy matrix directly choosing a suitable local coordinate system on the orbit $x(t)$ in form of the (non singular) transformation $\mathbf{U}(x(t))$:

$$\tilde{\mathbf{J}}(x(t)) = \mathbf{U}^{-1}(x(t))\,\mathbf{J}(x(t))\,\mathbf{U}(x(0)) \tag{C.7}$$

These lead to

$$\begin{aligned} \dot{\tilde{\mathbf{J}}} &= \tilde{\mathbf{L}}\tilde{\mathbf{J}} \\ \text{with} \qquad \tilde{\mathbf{L}} &= \mathbf{U}^{-1}(\mathbf{L}\mathbf{U} - \dot{\mathbf{U}}) \end{aligned} \tag{C.8}$$

Note that the properties a) – c) are only fulfilled for $\tilde{\mathbf{J}}$ and $\tilde{\mathbf{L}}$, if $\mathbf{U}$ itself is symplectic.
Choosing $x_E = \nabla H(t)/|\nabla H(t)|^2$ and x_t as local coordinates uncovers the two trivial eigenvalues 1 of the transformed matrix in (C.7) at any time t. Setting $\mathbf{U} = (x_t^T, x_E^T, x_1^T, \ldots, x_{2d-2}^T)$ gives

$$\tilde{\mathbf{J}} = \begin{pmatrix} 1 & * & * & \ldots & * \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & * & & & \\ \vdots & \vdots & & \mathbf{m} & \\ 0 & * & & & \end{pmatrix}; \qquad \tilde{\mathbf{L}} = \begin{pmatrix} 0 & * & * & \ldots & * \\ 0 & 0 & 0 & \ldots & 0 \\ 0 & * & & & \\ \vdots & \vdots & & 1 & \\ 0 & * & & & \end{pmatrix}, \tag{C.9}$$

The matrix $\mathbf{m}$ is now the monodromy matrix and the equation of motion are given by

$$\dot{\mathbf{m}} = \mathrm{l}\,\mathbf{m}. \tag{C.10}$$

The vectors $x_1, \ldots, x_{2d-2}$ must span the space perpendicular to the flow on the energy manifold.

For a system with two degrees of freedom, the matrix $\mathbf{U}(\mathbf{t})$ can be written down explicitly, i.e.

$$\mathbf{U}(t) = (x_t, x_1, x_E, x_2) = \begin{pmatrix} \dot{x} & -\dot{y} & -\dot{u}/q^2 & -\dot{v}/q^2 \\ \dot{y} & \dot{x} & -\dot{v}/q^2 & \dot{u}/q^2 \\ \dot{u} & \dot{v} & \dot{x}/q^2 & -\dot{y}/q^2 \\ \dot{v} & -\dot{u} & \dot{y}/q^2 & \dot{x}/q^2 \end{pmatrix} \tag{C.11}$$

with $x^T = (x, y; u, v)$ and $q = |\nabla H| = |\dot{x}|$. The matrix $\mathbf{U}$ is non singular and symplectic at every phase space point x (except the equilibrium points $\dot{x} = 0$). The matrix elements for l are given (C.13). One distinguishes 4 classes of eigenvalues of $\mathbf{m}$.

- *stable* or *elliptic*, if $\Lambda = e^{\pm i\pi\nu}$ and $\nu \in]0, 1[$.
- *marginal*, if $\Lambda = \pm 1$.
- *hyperbolic*, *inverse hyperbolic*, if $\Lambda = e^{\pm\lambda}$, $\Lambda = -e^{\pm\lambda}$; $\lambda > 0$ is called the Lyapunov exponent of the periodic orbit.
- *loxodromic*, if $\Lambda = e^{\pm u \pm i\Psi}$ with u and Ψ real. This is the most general case possible only in systems with 3 or more degree of freedoms.

For 2 degrees of freedom, i.e. $\mathbf{m}$ is a (2×2) matrix, the eigenvalues are determined by

$$\lambda = \frac{\mathrm{Tr}(\mathbf{m}) \pm \sqrt{\mathrm{Tr}(\mathbf{m})^2 - 4}}{2}, \tag{C.12}$$

i.e. $\mathrm{Tr}(\mathbf{m}) = 2$ separates stable and unstable behavior.

The l matrix elements for the local transformation (C.11) are

$$\begin{aligned} \tilde{\mathrm{l}}_{11} &= \frac{1}{q}[(h_x^2 - h_y^2 - h_u^2 + h_v^2)(h_{xu} - h_{yv}) + 2(h_x h_y - h_u h_v)(h_{xv} + h_{yu}) \\ &\quad -(h_x h_u + h_y h_v)(h_{xx} + h_{yy} - h_{uu} - h_{vv})] \\ \tilde{\mathrm{l}}_{12} &= \frac{1}{q^2}[(h_x^2 + h_v^2)(h_{yy} + h_{uu}) + (h_y^2 + h_u^2)(h_{xx} + h_{vv}) \\ &\quad -2(h_x h_u + h_y h_v)(h_{xu} + h_{yv}) - 2(h_x h_y - h_u h_v)(h_{xy} - h_{uv})] \\ \tilde{\mathrm{l}}_{21} &= -(h_x^2 + h_y^2)(h_{uu} + h_{vv}) - (h_u^2 + h_v^2)(h_{xx} + h_{yy}) \\ &\quad +2(h_x h_u - h_y h_v)(h_{xu} - h_{yv}) + 2(h_x h_v + h_y h_u)(h_{xv} + h_{yu}) \\ \tilde{\mathrm{l}}_{22} &= -\tilde{\mathrm{l}}_{11}, \end{aligned} \tag{C.13}$$

with h_i, h_{ij} is the derivative of the Hamiltonian H with respect to the phase space coordinates and $q = |\nabla H|^2$.

Remark C.1 The sign convention of the Poisson bracket. The Poisson bracket is antisymmetric in its arguments and there is a freedom to define it with the opposite sign convention. When such freedoms exist in physics, it is certain that both conventions are in use and this is no exception. In several texts you will see the right hand side of (7.34) defined as $[B, A]$ so that (7.35) is $\frac{dQ}{dt} = [Q, H]$. As long as one is consistent, there should be no problem.

Remark C.2 The sign convention of ω. The overall sign of ω, the symplectic invariant in (5.4), is set by the convention that the Hamilton's principal function (for energy conserving flows) is given by $R(q, q', t) = \int_q^{q'} p_i dq_i - Et$. With this sign convention the action along a classical path is minimal, and the kinetic energy of a free particle is positive.

Appendix D

Implementing evolution

The large body of accrued wisdom on the subject of flows called fluid dynamics is about physicaly real flows of media with (typically) continuous densities. On the other hand, the flows in state spaces of complicated systems are not only abstract, but may lead to fractal or other forms of complication in the densities. The latter frequently require more abstract tools that we develop in this text. To sharpen our intuition about those, it might be helpful to outline the more tangible fluid dynamical vision.

D.1 Material invariants

> I'm a material girl in a material world.
> Madonna, *Material Girl*
>
> (E.A. Spiegel and P. Cvitanović)

We consider first the simplest property of a flow called a *material invariant* by fluid dynamicists, who have no choice but to ponder such matters. A material invariant $I(x)$ is a property attached to each point x that is preserved by the flow, $I(x) = I(f^t(x))$; for example, at this point the fluid is green. As $I(x) = I(f^t(x))$ is invariant, its total time derivative vanishes, $\dot{I}(x) = 0$. Written in terms of partial derivatives this is the *conservation equation* for the material invariant

$$\partial_t I + v_i \partial_i I = 0\,. \tag{D.1}$$

Let the *density* of representative points be $\rho(x,t)$. To each representative point x we assign the value $I(x)$ of some physical property of a system. The manner in which the flow redistributes $I(x)$ is governed by a partial differential

equation whose form is relatively simple because the representative points are neither created nor destroyed. This conservation property is expressed in the integral statement

$$\partial_t \int_{\mathcal{V}} dx\, \rho I = -\int_{\partial\mathcal{V}} d\sigma\, n_i v_i \rho I\,, \tag{D.2}$$

where $\mathcal{V}$ is an arbitrary volume in $\mathcal{M}$, $\partial\mathcal{V}$ is its surface and $\mathbf{n}$ is its outward normal. Repeated indices are summed over throughout and ∂_t is the partial derivative with respect to time.

We may use the divergence theorem to turn the surface integral into a volume integral and obtain

$$\int_{\mathcal{V}} \left[\partial_t(\rho I) + \partial_i(v_i \rho I)\right] dx = 0\,,$$

where ∂_i is the partial derivative operator with respect to x_i. Since the integration is over an arbitrary volume, we conclude that

$$\partial_t(\rho I) + \partial_i(\rho I v_i) = 0\,. \tag{D.3}$$

The choice $I \equiv 1$ yields the *continuity equation* for the density:

$$\partial_t \rho + \partial_i(\rho v_i) = 0\,. \tag{D.4}$$

If the density itself is a material invariant, combining (D.1) and (D.4) we conclude that $\partial_i v_i = 0$ and $\mathbf{J}^t(x_0) = 1$. An example of such incompressible flow is the Hamiltonian flow of sect. 5.1.1. For incompressible flows the continuity equation (D.4) becomes a statement of conservation of the phase-space volume (see sect. 5.1.1), or the Liouville theorem

$$\partial_t \rho + v_i \partial_i \rho = 0\,. \tag{D.5}$$

D.2 Koopmania

The way in which time evolution acts on densities may be rephrased in the language of functional analysis, by introducing the *Koopman operator*, whose action

on a phase space function $a(x)$ is to replace it by its downstream value time t later, $a(x) \to a(x(t))$ evaluated at the trajectory point $x(t)$:

$$\mathcal{K}^t a(x) \;=\; a(f^t(x))\,. \tag{D.6}$$

Observable $a(x)$ has no explicit time dependence; all time dependence is carried in its evaluation at $x(t)$ rather than at $x = x(0)$.

Suppose we are starting with an initial density of representative points $\rho(x)$: then the average value of $a(x)$ evolves as

$$\langle a\rangle(t) \;=\; \frac{1}{|\rho_{\mathcal{M}}|}\int_{\mathcal{M}} dx\; a(f^t(x))\rho(x) \;=\; \frac{1}{|\rho_{\mathcal{M}}|}\int_{\mathcal{M}} dx\;\left[\mathcal{K}^t a(x)\right]\rho(x)\,.$$

An alternative point of view (analogous to the shift from the Heisenberg to the Schrödinger picture in quantum mechanics) is to push dynamical effects into the density. In contrast to the Koopman operator which advances the trajectory by time t, the Perron-Frobenius operator (7.10) depends on the trajectory point time t in the past, so the Perron-Frobenius operator is the adjoint of the Koopman operator

?? on p. ??

$$\int_{\mathcal{M}} dx\;\left[\mathcal{K}^t a(x)\right]\;\rho(x) = \int_{\mathcal{M}} dx\, a(x)\;\left[\mathcal{L}^t \rho(x)\right]\;. \tag{D.7}$$

Checking this is an easy change of variables exercise. For finite dimensional deterministic invertible flows the Koopman operator (D.6) is simply the inverse of the Perron-Frobenius operator (7.6), so in what follows we shall not distinguish the two. However, for infinite dimensional flows contracting forward in time and for stochastic flows such inverses do not exist, and there you need to be more careful.

sect. 2.4.2

The family of Koopman's operators $\left\{\mathcal{K}^t\right\}_{t\in\mathbb{R}_+}$ forms a semigroup parametrized by time

(a) $\mathcal{K}^0 = I$

(b) $\mathcal{K}^t\mathcal{K}^{t'} = \mathcal{K}^{t+t'} \qquad t, t' \geq 0 \qquad$ (semigroup property) ,

with the *generator* of the semigroup, the generator of infinitesimal time translations defined by

$$\mathcal{A} \;=\; \lim_{t\to 0^+} \frac{1}{t}\left(\mathcal{K}^t - I\right)\,.$$

(If the flow is finite-dimensional and invertible, $\mathcal{A}$ is a generator of a group). The explicit form of $\mathcal{A}$ follows from expanding dynamical evolution up to first order, as in (2.4):

$$\mathcal{A}a(x) \;=\; \lim_{t\to 0^+} \frac{1}{t}\left(a(f^t(x)) - a(x)\right) \;=\; v_i(x)\partial_i a(x)\,. \tag{D.8}$$

Of course, that is nothing but the definition of the time derivative, so the equation of motion for $a(x)$ is

$$\left(\frac{d}{dt} - \mathcal{A}\right) a(x) \;=\; 0\,. \tag{D.9}$$

☞ appendix D.3

The finite time Koopman operator (D.6) can be formally expressed by exponentiating the time evolution generator $\mathcal{A}$ as

$$\mathcal{K}^t = e^{t\mathcal{A}}\,. \tag{D.10}$$

7.10 on p. 118

The generator $\mathcal{A}$ looks very much like the generator of translations. Indeed, for a constant velocity field dynamical evolution is nothing but a translation by time $\times$ velocity:

7.11 on p. 118

$$e^{tv\frac{\partial}{\partial x}} a(x) = a(x+tv)\,. \tag{D.11}$$

As we will not need to implement a computational formula for general $e^{t\mathcal{A}}$ in what follows, we relegate making sense of such operators to appendix D.3. Here we limit ourselves to a brief remark about the notion of "spectrum" of a linear operator.

☞ appendix D.3

The Koopman operator $\mathcal{K}$ acts multiplicatively in time, so it is reasonable to suppose that there exist constants $M > 0$, $\beta \geq 0$ such that $||\mathcal{K}^t|| \leq Me^{t\beta}$ for all $t \geq 0$. What does that mean? The operator norm is define in the same spirit in which we defined the matrix norms in sect. K.2: We are assuming that no value of $\mathcal{K}^t\rho(x)$ grows faster than exponentially for any choice of function $\rho(x)$, so that the fastest possible growth can be bounded by $e^{t\beta}$, a reasonable expectation in the light of the simplest example studied so far, the exact escape rate (7.14). If that is so, multiplying $\mathcal{K}^t$ by $e^{-t\beta}$ we construct a new operator $e^{-t\beta}\mathcal{K}^t = e^{t(\mathcal{A}-\beta)}$ which decays exponentially for large t, $||e^{t(\mathcal{A}-\beta)}|| \leq M$. We say that $e^{-t\beta}\mathcal{K}^t$ is an element of a *bounded* semigroup with generator $\mathcal{A} - \beta I$. Given this bound, it follows by the Laplace transform

$$\int_0^\infty dt\, e^{-st}\mathcal{K}^t = \frac{1}{s-\mathcal{A}}\,, \qquad \text{Re}\, s > \beta\,, \tag{D.12}$$

☞ sect. K.2

that the *resolvent* operator $(s-\mathcal{A})^{-1}$ is bounded

$$\left|\left|\frac{1}{s-\mathcal{A}}\right|\right| \leq \int_0^\infty dt\, e^{-st} M e^{t\beta} = \frac{M}{s-\beta}\,.$$

If one is interested in the spectrum of $\mathcal{K}$, as we will be, the resolvent operator is a natural object to study. The main lesson of this brief aside is that for the continuous time flows the Laplace transform is the tool that brings down the generator in (7.29) into the resolvent form (7.31) and enables us to study its spectrum.

D.3 Implementing evolution

(R. Artuso and P. Cvitanović)

We now come back to the semigroup of operators $\mathcal{K}^t$. We have introduced the generator of the semigroup (7.26) as

$$\mathcal{A} \;=\; \frac{d}{dt}\mathcal{K}^t\Big|_{t=0}\,.$$

If we now take the derivative at arbitrary times we get

$$\begin{aligned}\left(\frac{d}{dt}\mathcal{K}^t\psi\right)(x) &= \lim_{\eta\to 0}\frac{\psi(f^{t+\eta}(x)) - \psi(f^t(x))}{\eta} \\ &= v_i(f^t(x))\,\frac{\partial}{\partial\tilde{x}_i}\psi(\tilde{x})\Big|_{\tilde{x}=f^t(x)} \\ &= \left(\mathcal{K}^t\mathcal{A}\psi\right)(x)\end{aligned}$$

which can be formally integrated like an ordinary differential equation yielding

7.10 on p. 118

$$\mathcal{K}^t \;=\; e^{t\mathcal{A}}\,. \tag{D.13}$$

This guarantees that the Laplace transform manipulations in sect. 7.4 are correct. Though the formal expression of the semigroup (D.13) is quite simple one has to take care in implementing its action. If we express the exponential through the power series

$$\mathcal{K}^t \;=\; \sum_{k=0}^{\infty} \frac{t^k}{k!}\mathcal{A}^k\,, \tag{D.14}$$

we encounter the problem that the infinitesimal generator (7.26) contains non-commuting pieces, that is, there are i, j combinations for which the commutator does not satisfy

$$\left[\frac{\partial}{\partial x_i}, v_j(x)\right] = 0\,.$$

To derive a more useful representation, we follow the strategy used for finite-dimensional matrix operators in sects. 4.2 and 4.3 and use the semigroup property to write

$$\mathcal{K}^t \;=\; \prod_{m=1}^{t/\delta\tau} \mathcal{K}^{\delta\tau}$$

as the starting point for a discretized approximation to the continuous time dynamics, with time step $\delta\tau$. Omitting terms from the second order onwards in the expansion of $\mathcal{K}^{\delta\tau}$ yields an error of order $O(\delta\tau^2)$. This might be acceptable if the time step $\delta\tau$ is sufficiently small. In practice we write the Euler product

$$\mathcal{K}^t \;=\; \prod_{m=1}^{t/\delta\tau} \left(1 + \delta\tau\mathcal{A}_{(m)}\right) \;+\; O(\delta\tau^2) \tag{D.15}$$

where

$$\left(\mathcal{A}_{(m)}\psi\right)(x) \;=\; v_i(f^{m\,\delta\tau}(x))\,\frac{\partial\psi}{\partial\tilde{x}_i}\bigg|_{\tilde{x}=f^{m\delta\tau}(x)}$$

As far as the x dependence is concerned, $e^{\delta\tau\mathcal{A}_i}$ acts as

$$e^{\delta\tau\mathcal{A}_i}\begin{Bmatrix} x_1 \\ \cdot \\ x_i \\ x_d \end{Bmatrix} \to \begin{Bmatrix} x_1 \\ \cdot \\ x_i + \delta\tau v_i(x) \\ x_d \end{Bmatrix}. \tag{D.16}$$

2.6 on p. 49

We see that the product form (D.15) of the operator is nothing else but a prescription for finite time step integration of the equations of motion - in this case the simplest Euler type integrator which advances the trajectory by $\delta\tau\times$velocity at each time step.

D.3.1 A symplectic integrator

The procedure we described above is only a starting point for more sophisticated approximations. As an example on how to get a sharper bound on the error term consider the Hamiltonian flow $\mathcal{A} = \mathcal{B} + \mathcal{C}$, $\mathcal{B} = p_i\frac{\partial}{\partial q_i}$, $\mathcal{C} = -\partial_i V(q)\frac{\partial}{\partial p_i}$. Clearly the potential and the kinetic parts do not commute. We make sense of the formal solution (D.15) by spliting it into infinitesimal steps and keeping terms up to $\delta\tau^2$ in

D.2 on p. 599

$$\mathcal{K}^{\delta\tau} = \hat{\mathcal{K}}^{\delta\tau} + \frac{1}{24}(\delta\tau)^3[\mathcal{B}+2\mathcal{C},[\mathcal{B},\mathcal{C}]] + \cdots\,, \tag{D.17}$$

where

$$\hat{\mathcal{K}}^{\delta\tau} = e^{\frac{1}{2}\delta\tau\mathcal{B}}e^{\delta\tau\mathcal{C}}e^{\frac{1}{2}\delta\tau\mathcal{B}}\,. \tag{D.18}$$

The approximate infinitesimal Liouville operator $\hat{\mathcal{K}}^{\delta\tau}$ is of the form that now generates evolution as a sequence of mappings induced by (7.30), a free flight by $\frac{1}{2}\delta\tau\mathcal{B}$, scattering by $\delta\tau\partial V(q')$, followed again by $\frac{1}{2}\delta\tau\mathcal{B}$ free flight:

$$\begin{aligned}
e^{\frac{1}{2}\delta\tau\mathcal{B}}\begin{Bmatrix} q \\ p \end{Bmatrix} &\to \begin{Bmatrix} q' \\ p' \end{Bmatrix} = \begin{Bmatrix} q - \frac{\delta\tau}{2}p \\ p \end{Bmatrix} \\
e^{\delta\tau\mathcal{C}}\begin{Bmatrix} q' \\ p' \end{Bmatrix} &\to \begin{Bmatrix} q'' \\ p'' \end{Bmatrix} = \begin{Bmatrix} q' \\ p' + \delta\tau\partial V(q') \end{Bmatrix} \\
e^{\frac{1}{2}\delta\tau\mathcal{B}}\begin{Bmatrix} q'' \\ p'' \end{Bmatrix} &\to \begin{Bmatrix} q''' \\ p''' \end{Bmatrix} = \begin{Bmatrix} q' - \frac{\delta\tau}{2}p'' \\ p'' \end{Bmatrix}
\end{aligned} \tag{D.19}$$

Collecting the terms we obtain an integration rule for this type of symplectic flow which is better than the straight Euler integration (D.16) as it is accurate up to order $\delta\tau^2$:

$$\begin{aligned}
q_{n+1} &= q_n - \delta\tau\, p_n - \frac{(\delta\tau)^2}{2}\partial V\left(q_n - \delta\tau p_n/2\right) \\
p_{n+1} &= p_n + \delta\tau\partial V\left(q_n - \delta\tau p_n/2\right)
\end{aligned} \tag{D.20}$$

The Jacobian matrix of one integration step is given by

$$\mathbf{J} = \begin{pmatrix} 1 & -\delta\tau/2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \delta\tau\partial V(q') & 1 \end{pmatrix}\begin{pmatrix} 1 & -\delta\tau/2 \\ 0 & 1 \end{pmatrix}. \tag{D.21}$$

Note that the billiard flow (5.22) is an example of such symplectic integrator. In that case the free flight is interupted by instantaneous wall reflections, and can be integrated out.

Commentary

Remark D.1 Koopman operators. The "Heisenberg picture" in dynamical system theory has been introduced by Koopman refs. [2, 3], see also ref. [1]. Inspired by the contemporary advances in quantum mechanics, Koopman [2] observed in 1931 that $\mathcal{K}^t$ is unitary on $L^2(\mu)$ Hilbert spaces. The Liouville/Koopman operator is the classical analogue of the quantum evolution operator — the kernel of $\mathcal{L}^t(y,x)$ introduced in (7.16) (see also sect. 8.2) is the analogue of the Green's function. The relation between the spectrum of the Koopman operator and classical ergodicity was formalized by von Neumann [3]. We shall not use Hilbert spaces here and the operators that we shall study *will not* be unitary. For a discussion of the relation between the Perron-Frobenius operators and the Koopman operators for finite dimensional deterministic invertible flows, infinite dimensional contracting flows, and stochastic flows, see Lasota-Mackey [1] and Gaspard [4].

Remark D.2 Symplectic integration. The reviews [6] and [7] offer a good starting point for exploring the symplectic integrators literature.

References

[D.1] A. Lasota and M.C. Mackey, *Chaos, Fractals and Noise* (Springer, New York 1994).

[D.2] B.O. Koopman, *Proc. Nat. Acad. Sci. USA* **17**, 315 (1931).

[D.3] J. von Neumann, *Ann. Math.* **33**, 587 (1932).

[D.4] B.A. Shadwick, J.C. Bowman, and P.J. Morrison, *Exactly Conservative Integrators*, `chao-dyn/9507012`, Submitted to *SIAM J. Sci. Comput.*

[D.5] D.J.D. Earn, *Symplectic integration without roundoff error*, `astro-ph/9408024`.

[D.6] P.J. Channell and C. Scovel, *Nonlinearity* **3**, 231 (1990).

[D.7] J.M. Sanz-Serna and M.P. Calvo, *Numerical Hamiltonian problems* (Chapman and Hall, London, 1994).

Exercises

Exercise D.1 **Symplectic volume preservation.** *Check that the sequence of mappings (D.19) is volume preserving,* $\det \hat{U} = 1$.

Exercise D.2 **Noncommutativity.** *Check that the commutators in (D.17) are not vanishing by showing that*

$$[\mathcal{B},\mathcal{C}] = -p\left(V''\frac{\partial}{\partial p} - V'\frac{\partial}{\partial q}\right) .$$

Exercise D.3 **Symplectic leapfrog integrator.** *Implement (D.20) for 2-dimensional Hamiltonian flows; compare it with Runge-Kutta integrator by integrating trajectories in some (chaotic) Hamiltonian flow.*

Appendix E

Symbolic dynamics techniques

The kneading theory for unimodal mappings is developed in sect. E.1. Pruning theory for Bernoulli shifts (an exercise mostly of formal interest) is discussed in sect. **??**. The prime factorization for dynamical itineraries of sect. E.2 illustrates the sense in which prime cycles are "prime" - the product structure of zeta functions is a consequence of the unique factorization property of symbol sequences.

E.1 Topological zeta functions for infinite subshifts

(Per Dahlqvist)

The Markov graph methods outlined in chapter 9 are well suited for symbolic dynamics of finite subshift type. A sequence of well defined rules leads to the answer, the topological zeta function, which turns out to be a polynomial. For infinite subshifts one would have to go through an infinite sequence of graph constructions and it is of course very difficult to make any asymptotic statements about the outcome. Luckily, for some simple systems the goal can be reached by much simpler means. This is the case for unimodal maps.

We will restrict our attention to the topological zeta function for unimodal maps with one external parameter $f_A(x) = Ag(x)$. As usual, symbolic dynamics is introduced by mapping a time series $\ldots x_{i-1}x_ix_{i+1}\ldots$ onto a sequence of symbols $\ldots s_{i-1}s_is_{i+1}\ldots$ where

$$\begin{aligned} s_i &= 0 \quad x_i < x_c \\ s_i &= C \quad x_i = x_c \\ s_i &= 1 \quad x_i > x_c \end{aligned} \tag{E.1}$$

$I(C)$	$\zeta_{top}^{-1}(z)/(1-z)$
$1C$	
$101C$	
$1011101C$	
$H^\infty(1)$	$\prod_{n=0}^{\infty}(1-z^{2^n})$
$10111C$	
$1011111C$	
101^∞	$(1-2z^2)/(1+z)$
$10111111C$	
$101111C$	
$1011C$	
$101101C$	
$10C$	$(1-z-z^2)$
$10010C$	
$100101C$	

$I(C)$	$\zeta_{top}^{-1}(z)/(1-z)$
$1001C$	
$100111C$	
$10011C$	
$100110C$	
$100C$	
$100010C$	
$10001C$	
$100011C$	
$1000C$	
$100001C$	
$10000C$	
$100000C$	
10^∞	$(1-2z)/(1-z)$

Table E.1: All ordered kneading sequences up to length seven, as well as some longer kneading sequences. Harmonic extension $H^\infty(1)$ is defined below.

and x_c is the critical point of the map (that is maximum of g). In addition to the usual binary alphabet we have added a symbol C for the critical point. The kneading sequence K_Λ is the itinerary of the critical point. The crucial observation is that no periodic orbit can have a topological coordinate (see sect. E.1.1) beyond that of the kneading sequence. The kneading sequence thus inserts a border in the list of periodic orbits (ordered according to maximal topological coordinate), cycles up to this limit are allowed, all beyond are pruned. All unimodal maps (obeying some further constraints) with the same kneading sequence thus have the same set of periodic orbitsand the same topological zeta function. The topological coordinate of the kneading sequence increases with increasing Λ.

The kneading sequence can be of one of three types

1. It maps to the critical point again, after n iterations. If so, we adopt the convention to terminate the kneading sequence with a C, and refer to the kneading sequence as finite.
2. Preperiodic, that is it is infinite but with a periodic tail.
3. Aperiodic.

As an archetype unimodal map we will choose the *tent map*

$$x \mapsto f(x) = \begin{cases} \Lambda x & x \in [0,1/2] \\ \Lambda(1-x) & x \in (1/2,1] \end{cases} , \tag{E.2}$$

where the parameter $\Lambda \in (1,2]$. The topological entropy is $h = \log\Lambda$. This follows from the fact any trajectory of the map is bounded, the escape rate is

strictly zero, and so the dynamical zeta function

$$1/\zeta(z) = \prod_p \left(1 - \frac{z^{n_p}}{|\Lambda_p|}\right) = \prod_p \left(1 - \left(\frac{z}{\Lambda}\right)^{n_p}\right) = 1/\zeta_{\rm top}(z/\Lambda)$$

has its leading zero at $z = 1$.

The set of periodic points of the tent map is countable. A consequence of this fact is that the set of parameter values for which the kneading sequence is periodic or preperiodic are countable and thus of measure zero and consequently *the kneading sequence is aperiodic for almost all* Λ. For general unimodal maps the corresponding statement is that the kneading sequence is aperiodic for almost all topological entropies.

For a given periodic kneading sequence of period n, $\underline{K_\Lambda} = PC = s_1 s_2 \dots s_{n-1} C$ there is a simple expansion for the topological zeta function. Then the expanded zeta function is a polynomial of degree n

$$1/\zeta_{\rm top}(z) = \prod_p (1 - z_p^n) = (1-z)\sum_{i=0}^{n-1} a_i z^i , \qquad a_i = \prod_{j=1}^{i} (-1)^{s_j} \tag{E.3}$$

and $a_0 = 1$.

Aperiodic and preperiodic kneading sequences are accounted for by simply replacing n by ∞.

Example. Consider as an example the kneading sequence $K_\Lambda = 10C$. From (E.3) we get the topological zeta function $1/\zeta_{top}(z) = (1-z)(1-z-z^2)$, see table E.1. This can also be realized by redefining the alphabet. The only forbidden subsequence is 100. All allowed periodic orbits, except $\overline{0}$, can can be built from a alphabet with letters $\underline{10}$ and $\underline{1}$. We write this alphabet as $\{\underline{10},\underline{1};\overline{0}\}$, yielding the topological zeta function $1/\zeta_{top}(z) = (1-z)(1-z-z^2)$. The leading zero is the inverse golden mean $z_0 = (\sqrt{5}-1)/2$.

Example. As another example we consider the preperiodic kneading sequence $K_\Lambda = 101^\infty$. From (E.3) we get the topological zeta function $1/\zeta_{top}(z) = (1-z)(1-2z^2)/(1+z)$, see table E.1. This can again be realized by redefining the alphabet. There are now an infinite number of forbidden subsequences, namely $101^{2n}0$ where $n \geq 0$. These pruning rules are respected by the alphabet $\{\underline{01^{2n+1}};\overline{1},\overline{0}\}$, yielding the topological zeta function above. The pole in the zeta function $\zeta_{top}^{-1}(z)$ is a consequence of the infinite alphabet.

An important consequence of (E.3) is that the sequence $\{a_i\}$ has a periodic tail if and only if the kneading sequence has one (however, their period may differ by a factor of two). We know already that the kneading sequence is aperiodic for almost all Λ.

The analytic structure of the function represented by the infinite series $\sum a_i z_i$ with unity as radius of convergence, depends on whether the tail of $\{a_i\}$ is periodic or not. If the period of the tail is N we can write

$$1/\zeta_{top}(z) = p(z) + q(z)(1 + z^N + z^{2N} \ldots) = p(z) + \frac{q(z)}{1 - z^N} \quad ,$$

for some polynomials $p(z)$ and $q(z)$. The result is a set of poles spread out along the unit circle. This applies to the preperiodic case. An aperiodic sequence of coefficients would formally correspond to infinite N and it is natural to assume that the singularities will fill the unit circle. There is indeed a theorem ensuring that this is the case [**?**], provided the a_i's can only take on a finite number of values. The unit circle becomes a *natural boundary*, already apparent in a finite polynomial approximations to the topological zeta function, as in fig. 10.2. A function with a natural boundary lacks an analytic continuation outside it.

To conclude: The topological zeta function $1/\zeta_{top}$ for unimodal maps has the unit circle as a natural boundary for almost all topological entropies and for the tent map (E.2), for almost all Λ.

Let us now focus on the relation between the analytic structure of the topological zeta function and the number of periodic orbits, or rather (10.6), the number N_n of fixed points of $f^n(x)$. The trace formula is (see sect. 10.4)

$$N_n = \mathrm{tr}\, T^n = \frac{1}{2\pi i} \oint_{\gamma_r} dz\, z^{-n} \frac{d}{dz} \log \zeta_{top}^{-1}$$

where γ_r is a (circular) contour encircling the origin $z = 0$ in clockwise direction. Residue calculus turns this into a sum over zeros z_0 and poles z_p of ζ_{top}^{-1}

$$N_n = \sum_{z_0 : r < |z_0| < R} z_0^{-n} - \sum_{z_p : r < |z_p| < R} z_0^{-n} + \frac{1}{2\pi i} \oint_{\gamma_R} dz\, z^{-n} \frac{d}{dz} \log \zeta_{top}^{-1}$$

and a contribution from a large circle γ_R. For meromorphic topological zeta functions one may let $R \to \infty$ with vanishing contribution from γ_R, and N_n will be a sum of exponentials.

The leading zero is associated with the topological entropy, as discussed in chapter 9.

We have also seen that for preperiodic kneading there will be poles on the unit circle.

To appreciate the role of natural boundaries we will consider a (very) special example. Cascades of period doublings is a central concept for the description of unimodal maps. This motivates a close study of the function

$$\Xi(z) = \prod_{n=0}^{\infty} (1 - z^{2^n}) \quad . \tag{E.4}$$

This function will appear again when we derive (E.3).

The expansion of $\Xi(z)$ begins as $\Xi(z) = 1 - z - z^2 + z^3 - z^4 + z^5 \ldots$. The radius of convergence is obviously unity. The simple rule governing the expansion will effectively prohibit any periodicity among the coefficients making the unit circle a natural boundary.

It is easy to see that $\Xi(z) = 0$ if $z = \exp(2\pi m/2^n)$ for any integer m and n. (Strictly speaking we mean that $\Xi(z) \to 0$ when $z \to \exp(2\pi m/2^n)$ from inside). Consequently, zeros are dense on the unit circle. One can also show that singular points are dense on the unit circle, for instance $|\Xi(z)| \to \infty$ when $z \to \exp(2\pi m/3^n)$ for any integer m and n.

As an example, the topological zeta function at the accumulation point of the first Feigenbaum cascade is $\zeta_{top}^{-1}(z) = (1 - z)\Xi(z)$. Then $N_n = 2^{l+1}$ if $n = 2^l$, otherwise $N_n = 0$. The growth rate in the number of cycles is anything but exponential. It is clear that N_n cannot be a sum of exponentials, the contour γ_R cannot be pushed away to infinity, R is restricted to $R \leq 1$ and N_n is entirely determined by $\int_{\gamma_R}$ which picks up its contribution from the natural boundary.

We have so far studied the analytic structure for some special cases and we know that the unit circle is a natural boundary for almost all Λ. But how does it look out there in the complex plane for some typical parameter values? To explore that we will imagine a journey from the origin $z = 0$ out towards the unit circle. While traveling we let the parameter Λ change slowly. The trip will have a distinct science fiction flavor. The first zero we encounter is the one connected to the topological entropy. Obviously it moves smoothly and slowly. When we move outward to the unit circle we encounter zeros in increasing densities. The closer to the unit circle they are, the wilder and stranger they move. They move from and back to the horizon, where they are created and destroyed through bizarre bifurcations. For some special values of the parameter the unit circle suddenly gets transparent and and we get (infinitely) short glimpses of another world beyond the horizon.

We end this section by deriving eqs (E.5) and (E.6). The impenetrable prose is hopefully explained by the accompanying tables.

We know one thing from chapter 9, namely for that finite kneading sequence of length n the topological polynomial is of degree n. The graph contains a node which is connected to itself only via the symbol 0. This implies that a factor $(1-z)$ may be factored out and $\zeta_{top}(z) = (1-z)\sum_{i=0}^{n-1} a_i z^i$. The problem is to find the coefficients a_i.

periodic orbits	finite kneading sequences
$\overline{P1} = A^\infty(P)$	
	PC
$\overline{P0}$	
	$P0PC$
$\overline{P0P1}$	
	$P0P1P0PC$
$\downarrow$	$\downarrow$
$H^\infty(P)$	$H^\infty(P)$

Table E.2: Relation between periodic orbits and finite kneading sequences in a harmonic cascade. The string P is assumed to contain an odd number of 1's.

The ordered list of (finite) kneading sequences table E.1 and the ordered list of periodic orbits (on maximal form) are intimately related. In table E.2 we indicate how they are nested during a period doubling cascade. Every finite kneading sequence PC is bracketed by two periodic orbits, $\overline{P1}$ and $\overline{P0}$. We have $\overline{P1} < PC < \overline{P0}$ if P contains an odd number of 1's, and $\overline{P0} < PC < \overline{P1}$ otherwise. From now on we will assume that P contains an odd number of 1's. The other case can be worked out in complete analogy. The first and second harmonic of PC are displayed in table E.2. The periodic orbit $\overline{P1}$ (and the corresponding infinite kneading sequence) is sometimes referred to as the antiharmonic extension of PC (denoted $A^\infty(P)$) and the accumulation point of the cascade is called the harmonic extension of PC [?] (denoted $H^\infty(P)$).

A central result is the fact that a period doubling cascade of PC is not interfered by any other sequence. Another way to express this is that a kneading sequence PC and its harmonic are adjacent in the list of kneading sequences to any order.

$I(C)$			$\zeta_{top}^{-1}(z)/(1-z)$
P_1	=	$100C$	$1-z-z^2-z^3$
$H^\infty(P_1)$	=	$10001001100\ldots$	$1-z-z^2-z^3-z^4+z^5+z^6+z^7-z^8\ldots$
P'	=	$10001C$	$1-z-z^2-z^3-z^4+z^5$
$A^\infty(P_2)$	=	$1000110001\ldots$	$1-z-z^2-z^3-z^4+z^5-z^6-z^7-z^8\ldots$
P_2	=	$1000C$	$1-z-z^2-z^3-z^4$

Table E.3: Example of a step in the iterative construction of the list of kneading sequences PC.

Table E.3 illustrates another central result in the combinatorics of kneading sequences. We suppose that P_1C and P_2C are neighbors in the list of order 5

(meaning that the shortest finite kneading sequence $P'C$ between P_1C and P_2C is longer than 5.) The important result is that P' (of length $n'=6$) has to coincide with the first $n'-1$ letters of both $H^\infty(P_1)$ and $A^\infty(P_2)$. This is exemplified in the left column of table E.3. This fact makes it possible to generate the list of kneading sequences in an iterative way.

The zeta function at the accumulation point $H^\infty(P_1)$ is

$$\zeta_{P_1}^{-1}(z)\Xi(z^{n_1}) \quad , \tag{E.5}$$

and just before $A^\infty(P_2)$

$$\zeta_{P_2}^{-1}(z)/(1-z^{n_2}) \quad . \tag{E.6}$$

A short calculation shows that this is exactly what one would obtain by applying (E.3) to the antiharmonic and harmonic extensions directly, provided that it applies to $\zeta_{P_1}^{-1}(z)$ and $\zeta_{P_2}^{-1}(z)$. This is the key observation.

Recall now the product representation of the zeta function $\zeta^{-1} = \prod_p (1-z^{n_p})$. We will now make use of the fact that the zeta function associated with $P'C$ is a polynomial of order n'. There is no periodic orbit of length shorter than $n'+1$ between $H^\infty(P_1)$ and $A^\infty(P_2)$. It thus follows that the coefficients of this polynomial coincides with those of (E.5) and (E.6), see Table E.3. We can thus conclude that our rule can be applied directly to $P'C$.

This can be used as an induction step in proving that the rule can be applied to every finite and infinite kneading sequences.

Remark E.1 How to prove things. The explicit relation between the kneading sequence and the coefficients of the topological zeta function is not commonly seen in the literature. The result can proven by combining some theorems of Milnor and Thurston [?]. That approach is hardly instructive in the present context. Our derivation was inspired by Metropolis, Stein and Stein classical paper [?]. For further detail, consult [?].

E.1.1 Periodic orbits of unimodal maps

A *periodic point* (or a *cycle* point) x_i belonging to a cycle of period n is a real solution of

$$f^n(x_i) = f(f(\ldots f(x_i)\ldots)) = x_i\,, \quad i = 0,1,2,\ldots,n-1 \tag{E.7}$$

The nth iterate of a unimodal map crosses the diagonal at most 2^n times. Similarly, the backward and the forward Smale horseshoes intersect at most 2^n times, and therefore there will be 2^n or fewer periodic points of length n. A cycle of length n corresponds to an infinite repetition of a length n symbol string, customarily indicated by a line over the string:

$$S = (s_1 s_2 s_3 \dots s_n)^\infty = \overline{s_1 s_2 s_3 \dots s_n}\,.$$

If $\overline{s_1 s_2 \dots s_n}$ is the symbol string associated with x_0, its cyclic permutation $\overline{s_k s_{k+1} \dots s_n s_1 \dots s_{k-1}}$ corresponds to the point x_{k-1} in the same cycle. A cycle p is called *prime* if its itinerary S cannot be written as a repetition of a shorter block S'.

Each cycle yields n rational values of γ. It follows from (**??**) that if the repeating string $s_1, s_2, \dots s_n$ contains an odd number "1"s, the string of well ordered symbols $w_1 w_2 \dots w_n$ has to be of the double length before it repeats itself. The value γ is a geometrical sum which we can write as the finite sum

$$\gamma(\overline{s_1 s_2 \dots s_n}) = \frac{2^{2n}}{2^{2n}-1} \sum_{t=1}^{2n} w_t / 2^t$$

Using this we can calculate the $\hat\gamma(S)$ for all short cycles. For orbits up to length 5 this is done in table **??**.

Here we give explicit formulas for the topological coordinate of a periodic point, given its itinerary. For the purpose of what follows it is convenient to compactify the itineraries by replacing the binary alphabet $s_i = \{0, 1\}$ by the infinite alphabet

$$\{a_1, a_2, a_3, a_4, \cdots; \overline{0}\} = \{1, 10, 100, 1000, \dots; \overline{0}\}\,. \tag{E.8}$$

In this notation the itinerary $S = a_i a_j a_k a_l \cdots$ and the corresponding topological coordinate (9.13) are related by $\gamma(S) = .1^i 0^j 1^k 0^l \cdots$. For example:

$$\begin{array}{lllllll} S & = & 111011101001000\dots & = & a_1 a_1 a_2 a_1 a_1 a_2 a_3 a_4 \dots \\ \gamma(S) & = & .101101001110000\dots & = & .1^1 0^1 1^2 0^1 1^1 0^2 1^3 0^4 \dots \end{array}$$

Cycle points whose itineraries start with $w_1 = w_2 = \ldots = w_i = 0$, $w_{i+1} = 1$ remain on the left branch of the tent map for i iterations, and satisfy $\gamma(0 \dots 0 S) = \gamma(S)/2^i$.

A *periodic point* (or a *cycle* point) x_i belonging to a cycle of period n is a real solution of

$$f^n(x_i) = f(f(\dots f(x_i)\dots)) = x_i\,, \quad i = 0, 1, 2, \dots, n-1\,. \tag{E.9}$$

The nth iterate of a unimodal map has at most 2^n monotone segments, and therefore there will be 2^n or fewer periodic points of length n. A periodic orbit of length n corresponds to an infinite repetition of a length n symbol string, customarily indicated by a line over the string:

$$S = (s_1 s_2 s_3 \dots s_n)^\infty = \overline{s_1 s_2 s_3 \dots s_n}\,.$$

As all itineraries are infinite, we shall adopt convention that a finite string itinerary $S = s_1 s_2 s_3 \dots s_n$ stands for infinite repetition of a finite block, and routinely omit the overline. If $\overline{s_1 s_2 \dots s_n}$ is the symbol string associated with x_0, its cyclic permutation $\overline{s_k s_{k+1} \dots s_n s_1 \dots s_{k-1}}$ corresponds to the point x_{k-1} in the same cycle. A periodic orbit p is called *prime* if its itinerary S cannot be written as a repetition of a shorter block S'.

Periodic points correspond to rational values of γ, but we have to distinguish *even* and *odd* cycles. The even (odd) cycles contain even (odd) number of a_i in the repeating block, with periodic points given by

$$\gamma(a_i a_j \cdots a_k a_\ell) = \begin{cases} \frac{2^n}{2^n-1}.1^i 0^j \cdots 1^k & \text{even} \\ \frac{1}{2^n+1}\left(1 + 2^n \times .1^i 0^j \cdots 1^\ell\right) & \text{odd} \end{cases}\,, \tag{E.10}$$

where $n = i + j + \cdots + k + \ell$ is the cycle period. The maximal value cycle point is given by the cyclic permutation of S with the largest a_i as the first symbol, followed by the smallest available a_j as the next symbol, and so on. For example:

$$\begin{array}{llllllll} \hat\gamma(1) & = & \gamma(a_1) & = & .10101\dots & = & .\overline{10} & = & 2/3 \\ \hat\gamma(10) & = & \gamma(a_2) & = & .1^2 0^2 \dots & = & .\overline{1100} & = & 4/5 \\ \hat\gamma(100) & = & \gamma(a_3) & = & .1^3 0^3 \dots & = & .\overline{111000} & = & 8/9 \\ \hat\gamma(101) & = & \gamma(a_2 a_1) & = & .1^2 0^1 \dots & = & .\overline{110} & = & 6/7 \end{array}$$

An example of a cycle where only the third symbol determines the maximal value cycle point is

$$\hat\gamma(1101110) = \gamma(a_2 a_1 a_2 a_1 a_1) = .\overline{11011010010010} = 100/129\,.$$

Maximal values of all cycles up to length 5 are given in table!?

E.2 Prime factorization for dynamical itineraries

The Möbius function is not only a number-theoretic function, but can be used to manipulate ordered sets of noncommuting objects such as symbol strings. Let $\mathcal{P} = \{p_1, p_2, p_3, \cdots\}$ be an ordered set of *prime* strings, and

$$\mathcal{N} = \{n\} = \left\{ p_1^{k_1} p_2^{k_2} p_3^{k_3} \cdots p_j^{k_j} \right\} ,$$

$j \in \mathbb{N}$, $k_i \in \mathbb{Z}_+$, be the set of all strings n obtained by the ordered concatenation of the "primes" p_i. By construction, every string n has a unique prime factorization. We say that a string has a divisor d if it contains d as a substring, and define the string division n/d as n with the substring d deleted. Now we can do things like this: defining $t_n := t_{p_1}^{k_1} t_{p_2}^{k_2} \cdots t_{p_j}^{k_j}$ we can write the inverse dynamical zeta function (15.2) as

$$\prod_p (1 - t_p) = \sum_n \mu(n) t_n ,$$

and, if we care (we do in the case of the Riemann zeta function), the dynamical zeta function as .

$$\prod_p \frac{1}{1 - t_p} = \sum_n t_n \tag{E.11}$$

A striking aspect of this formula is its resemblance to the factorization of natural numbers into primes: the relation of the cycle expansion (E.11) to the product over prime cycles is analogous to the Riemann zeta (exercise c) represented as a sum over natural numbers vs. its Euler product representation.

We now implement this factorization explicitly by decomposing recursively binary strings into ordered concatenations of prime strings. There are 2 strings of length 1, both prime: $p_1 = 0, p_2 = 1$. There are 4 strings of length 2: 00, 01, 11, 10. The first three are ordered concatenations of primes: $00 = p_1^2$, $01 = p_1 p_2$, $11 = p_2^2$; by ordered concatenations we mean that $p_1 p_2$ is legal, but $p_2 p_1$ is not. The remaining string is the only prime of length 2, $p_3 = 10$. Proceeding by discarding the strings which are concatenations of shorter primes $p_1^{k_1} p_2^{k_2} \cdots p_j^{k_j}$, with primes lexically ordered, we generate the standard list of primes, in agreement with table 9.2: 0, 1, 10, 101, 100, 1000, 1001, 1011, 10000, 10001, 10010, 10011, 10110, 10111, 100000, 100001, 100010, 100011, 100110, 100111, 101100, 101110, 101111, ... This factorization is illustrated in table E.4.

factors	string	factors	string	factors	string	factors	string
p_1	0	p_1^4	0000	p_1^5	00000	$p_1^2 p_5$	00101
p_2	1	$p_1^3 p_2$	0001	$p_1^4 p_2$	00001	$p_1 p_2 p_5$	01101
p_1^2	00	$p_1^2 p_2^2$	0011	$p_1^3 p_2^2$	00011	$p_2^2 p_5$	11101
$p_1 p_2$	01	$p_1 p_2^3$	0111	$p_1^2 p_2^3$	00111	$p_3 p_5$	10101
p_2^2	11	p_2^4	1111	$p_1 p_2^4$	01111	$p_1 p_6$	01000
p_3	10	$p_1^2 p_3$	0010	p_2^5	11111	$p_2 p_6$	11000
p_1^3	000	$p_1 p_2 p_3$	0110	$p_1^3 p_3$	00010	$p_1 p_7$	01001
$p_1^2 p_2$	001	$p_2^2 p_3$	1110	$p_1^2 p_2 p_3$	00110	$p_2 p_7$	11001
$p_1 p_2^2$	011	p_3^2	1010	$p_1 p_2^2 p_3$	01110	$p_1 p_8$	01011
p_2^3	111	$p_1 p_4$	0100	$p_2^3 p_3$	11110	$p_2 p_8$	11011
$p_1 p_3$	010	$p_2 p_4$	1100	$p_1 p_3^2$	01010	p_9	10000
$p_2 p_3$	110	$p_1 p_5$	0101	$p_2 p_3^2$	11010	p_{10}	10001
p_4	100	$p_2 p_5$	1101	$p_1^2 p_4$	00100	p_{11}	10010
p_5	101	p_6	1000	$p_1 p_2 p_4$	01100	p_{12}	10011
		p_7	1001	$p_2^2 p_4$	11100	p_{13}	10110
		p_8	1011	$p_3 p_4$	10100	p_{14}	10111

Table E.4: Factorization of all periodic points strings up to length 5 into ordered concatenations $p_1^{k_1} p_2^{k_2} \cdots p_n^{k_n}$ of prime strings $p_1 = 0, p_2 = 1, p_3 = 10, p_4 = 100, \ldots, p_{14} = 10111$.

E.2.1 Prime factorization for spectral determinants

Following sect. E.2, the spectral determinant cycle expansions is obtained by expanding F as a multinomial in prime cycle weights t_p

$$F = \prod_p \sum_{k=0}^{\infty} C_{p^k} t_p^k = \sum_{k_1 k_2 k_3 \cdots = 0}^{\infty} \tau_{p_1^{k_1} p_2^{k_2} p_3^{k_3} \cdots} \tag{E.12}$$

where the sum goes over all pseudocycles. In the above we have defined

$$\tau_{p_1^{k_1} p_2^{k_2} p_3^{k_3} \cdots} = \prod_{i=1}^{\infty} C_{p_i^{k_i}} t_{p_i}^{k_i} . \tag{E.13}$$

c on p. 214

A striking aspect of the spectral determinant cycle expansion is its resemblance to the factorization of natural numbers into primes: as we already noted in sect. E.2, the relation of the cycle expansion (E.12) to the product formula (12.9) is analogous to the Riemann zeta represented as a sum over natural numbers vs. its Euler product representation.

This is somewhat unexpected, as the cycle weights (for example, the stability eigenvalues (??)) factorize exactly with respect to r repetitions of a prime cycle,

$t_{pp\dots p} = t_p^r$, but only approximately (*shadowing*) with respect to subdividing a string into prime substrings, $t_{p_1p_2} \approx t_{p_1}t_{p_2}$.

The coefficients C_{p^k} have a simple form only in 1-d, given by the Euler formula (13.26). In higher dimensions C_{p^k} can be evaluated by expanding (12.9), $F(z) = \prod_p F_p$, where

$$F_p = 1 - \left(\sum_{r=1}^{\infty} \frac{t_p^r}{r d_{p,r}}\right) + \frac{1}{2}\left(\sum_{r=1}^{\infty} \frac{t_p^r}{r d_{p,r}}\right)^2 - \dots .$$

Expanding and recollecting terms, and suppressing the p cycle label for the moment, we obtain

$$\begin{aligned} F_p &= \sum_{r=1}^{\infty} C_k t^k, \quad C_k = (-)^k c_k / D_k, \\ D_k &= \prod_{r=1}^{k} d_r = \prod_{a=1}^{d}\prod_{r=1}^{k}(1-u_a^r) \end{aligned} \tag{E.14}$$

where evaluation of c_k requires a certain amount of not too luminous algebra:

$$\begin{aligned} c_0 &= 1 \\ c_1 &= 1 \\ c_2 &= \frac{1}{2}\left(\frac{d_2}{d_1} - d_1\right) = \frac{1}{2}\left(\prod_{a=1}^{d}(1+u_a) - \prod_{a=1}^{d}(1-u_a)\right) \\ c_3 &= \frac{1}{3!}\left(\frac{d_2 d_3}{d_1^2} + 2d_1 d_2 - 3d_3\right) \\ &= \frac{1}{6}\left(\prod_{a=1}^{d}(1+2u_a+2u_a^2+u_a^3)\right. \\ &\quad \left. +2\prod_{a=1}^{d}(1-u_a-u_a^2+u_a^3) - 3\prod_{a=1}^{d}(1-u_a^3)\right) \end{aligned}$$

etc.. For example, for a general 2-dimensional map we have

$$F_p = 1 - \frac{1}{D_1}t + \frac{u_1+u_2}{D_2}t^2 - \frac{u_1u_2(1+u_1)(1+u_2)+u_1^3+u_2^3}{D_3}t^3 + \dots . \tag{E.15}$$

We discuss the convergence of such cycle expansions in sect. J.4.

With $\tau_{p_1^{k_1}p_2^{k_2}\dots p_n^{k_n}}$ defined as above, the prime factorization of symbol strings is unique in the sense that *each symbol string can be written as a unique concatenation of prime strings*, up to a convention on ordering of primes. This factorization is a nontrivial example of the utility of generalized Möbius inversion, sect. E.2.

How is the factorization of sect. E.2 used in practice? Suppose we have computed (or perhaps even measured in an experiment) all prime cycles up to length n, that is we have a list of t_p's and the corresponding Jacobian matrix eigenvalues $\Lambda_{p,1}, \Lambda_{p,2}, \dots \Lambda_{p,d}$. A cycle expansion of the Selberg product is obtained by generating all strings in order of increasing length j allowed by the symbolic dynamics and constructing the multinomial

$$F = \sum_n \tau_n \tag{E.16}$$

where $n = s_1 s_2 \cdots s_j$, s_i range over the alphabet, in the present case $\{0,1\}$. Factorizing every string $n = s_1 s_2 \cdots s_j = p_1^{k_1} p_2^{k_2} \cdots p_j^{k_j}$ as in table E.4, and substituting $\tau_{p_1^{k_1}p_2^{k_2}\dots}$ we obtain a multinomial approximation to F. For example, $\tau_{001001010101} = \tau_{\underline{001}\,\underline{001}\,\underline{01}\,\underline{01}\,\underline{01}} = \tau_{001^2}\tau_{01^3}$, and τ_{01^3}, τ_{001^2} are known functions of the corresponding cycle eigenvalues. The zeros of F can now be easily determined by standard numerical methods. The fact that as far as the symbolic dynamics is concerned, the cycle expansion of a Selberg product is simply an average over all symbolic strings makes Selberg products rather pretty.

To be more explicit, we illustrate the above by expressing binary strings as concatenations of prime factors. We start by computing N_n, the number of terms in the expansion (E.12) of the total cycle length n. Setting $C_{p^k}t_p^k = z^{n_p k}$ in (E.12), we obtain

$$\sum_{n=0}^{\infty} N_n z^n = \prod_p \sum_{k=0}^{\infty} z^{n_p k} = \frac{1}{\prod_p (1 - z^{n_p})} \ .$$

So the generating function for the number of terms in the Selberg product is the topological zeta function. For the complete binary dynamics we have $N_n = 2^n$ contributing terms of length n:

$$\zeta_{top} = \frac{1}{\prod_p (1 - z^{n_p})} = \frac{1}{1-2z} = \sum_{n=0}^{\infty} 2^n z^n$$

Hence the number of distinct terms in the expansion (E.12) is the same as the number of binary strings, and conversely, the set of binary strings of length n suffices to label all terms of the total cycle length n in the expansion (E.12).

Appendix F

Counting itineraries

F.1 Counting curvatures

One consequence of the finitness of topological polynomials is that the contributions to curvatures at every order are even in number, half with positive and half with negative sign. For instance, for complete binary labelling (15.5),

$$\begin{aligned} c_4 &= -t_{0001} - t_{0011} - t_{0111} - t_0 t_{01} t_1 \\ &\quad + t_0 t_{001} + t_0 t_{011} + t_{001} t_1 + t_{011} t_1 \,. \end{aligned} \tag{F.1}$$

We see that 2^3 terms contribute to c_4, and exactly half of them appear with a negative sign - hence if all binary strings are admissible, this term vanishes in the counting expression.

F.2 on p. 617

Such counting rules arise from the identity

$$\prod_p (1 + t_p) = \prod_p \frac{1 - {t_p}^2}{1 - t_p} \,. \tag{F.2}$$

Substituting $t_p = z^{n_p}$ and using (10.15) we obtain for unrestricted symbol dynamics with N letters

$$\prod_p^{\infty} (1 + z^{n_p}) = \frac{1 - Nz^2}{1 - Nz} = 1 + Nz + \sum_{k=2}^{\infty} z^k \left(N^k - N^{k-1} \right)$$

The z^n coefficient in the above expansion is the number of terms contributing to c_n curvature, so we find that for a complete symbolic dynamics of N symbols

and $n > 1$, the number of terms contributing to c_n is $(N-1)N^{k-1}$ (of which half carry a minus sign).

F.4 on p. 618

We find that for complete symbolic dynamics of N symbols and $n > 1$, the number of terms contributing to c_n is $(N-1)N^{n-1}$. So, superficially, not much is gained by going from periodic orbits trace sums which get N^n contributions of n to the curvature expansions with $N^n(1-1/N)$. However, the point is not the number of the terms, but the cancellations between them.

Exercises

Exercise F.1 Lefschetz zeta function. *Elucidate the relation between the topological zeta function and the Lefschetz zeta function. This should help make sense of sect. ??.*

Exercise F.2 Counting the 3-disk pinball counterterms. *Verify that the number of terms in the 3-disk pinball curvature expansion (15.31) is given by*

$$\begin{aligned}\prod_p (1+t_p) &= \frac{1-3z^4-2z^6}{1-3z^2-2z^3} = 1+3z^2+2z^3+\frac{z^4(6+12z+2z^2)}{1-3z^2-2z^3} \\ &= 1+3z^2+2z^3+6z^4+12z^5+20z^6+48z^7+84z^8+184z^9+\end{aligned} \tag{F.3}$$

This means that, for example, c_6 has a total of 20 terms, in agreement with the explicit 3-disk cycle expansion (15.32).

Exercise F.3 Cycle expansion denominators****.** *Prove that the denominator of c_k is indeed D_k, as asserted (E.14).*

Exercise F.4 Counting subsets of cycles. *The techniques developed above can be generalized to counting subsets of cycles. Consider the simplest example of a dynamical system with a complete binary tree, a repeller map (9.10) with two straight branches, which we label 0 and 1. Every cycle weight for such map factorizes, with a factor t_0 for each 0, and factor t_1 for each 1 in its symbol string. The transition matrix traces (10.5) collapse to $tr(T^k) = (t_0+t_1)^k$, and $1/\zeta$ is simply*

$$\prod_p (1-t_p) = 1-t_0-t_1 \tag{F.4}$$

Substituting into the identity

$$\prod_p (1+t_p) = \prod_p \frac{1-{t_p}^2}{1-t_p}$$

we obtain

$$\begin{aligned}\prod_p (1+t_p) &= \frac{1-t_0^2-t_1^2}{1-t_0-t_1} = 1+t_0+t_1+\frac{2t_0t_1}{1-t_0-t_1} \\ &= 1+t_0+t_1+\sum_{n=2}^{\infty}\sum_{k=1}^{n-1} 2\binom{n-2}{k-1} t_0^k t_1^{n-k}.\end{aligned} \tag{F.5}$$

Hence for $n \geq 2$ the number of terms in the expansion ?! with k 0's and $n-k$ 1's in their symbol sequences is $2\binom{n-2}{k-1}$. This is the degeneracy of distinct cycle eigenvalues in fig.?!; for systems with non-uniform hyperbolicity this degeneracy is lifted (see fig. ?!).

In order to count the number of prime cycles in each such subset we denote with $M_{n,k}$ $(n = 1, 2, \ldots ; k = \{0, 1\}$ for $n = 1$; $k = 1, \ldots, n-1$ for $n \geq 2)$ the number of prime n-cycles whose labels contain k zeros, use binomial string counting and Möbius inversion and obtain

$$\begin{aligned} M_{1,0} &= M_{1,1} = 1 \\ nM_{n,k} &= \sum_{m \left| \frac{n}{k} \right.} \mu(m) \binom{n/m}{k/m}, \quad n \geq 2\,, k = 1, \ldots, n-1 \end{aligned}$$

where the sum is over all m which divide both n and k.

Appendix G

Finding cycles

(C. Chandre)

G.1 Newton-Raphson method

G.1.1 Contraction rate

Consider a d-dimensional map $x' = f(x)$ with an unstable fixed point x_*. The Newton-Raphson algorithm is obtained by iterating the following map

$$x' = g(x) = x - (J(x) - 1)^{-1} (f(x) - x)\,.$$

The linearization of g near x_* leads to

$$x_* + \epsilon' = x_* + \epsilon - (J(x_*) - \mathbf{1})^{-1} (f(x_*) + J(x_*)\epsilon - x_* - \epsilon) + O\left(\|\epsilon\|^2\right),$$

where $\epsilon = x - x_*$. Therefore,

$$x' - x_* = O\left((x - x_*)^2\right).$$

After n steps and if the initial guess x_0 is close to x_*, the error decreases super-exponentially

$$g^n(x_0) - x_* = O\left((x_0 - x_*)^{2^n}\right).$$

G.1.2 Computation of the inverse

The Newton-Raphson method for finding n-cycles of d-dimensional mappings using the multi-shooting method reduces to the following equation

$$\begin{pmatrix} \mathbf{1} & & & -Df(x_n) \\ -Df(x_1) & \mathbf{1} & & \\ & \cdots & \mathbf{1} & \\ & & -Df(x_{n-1}) & \mathbf{1} \end{pmatrix} \begin{pmatrix} \delta_1 \\ \delta_2 \\ \cdots \\ \delta_n \end{pmatrix} = - \begin{pmatrix} F_1 \\ F_2 \\ \cdots \\ F_n \end{pmatrix}, \quad \text{(G.1)}$$

where $Df(x)$ is the $[d \times d]$ Jacobian matrix of the map evaluated at the point x, and $\delta_m = x'_m - x_m$ and $F_m = x_m - f(x_{m-1})$ are d-dimensional vectors. By some starightforward algebra, the vectors δ_m are expressed as functions of the vectors F_m:

$$\delta_m = -\sum_{k=1}^{m} \beta_{k,m-1} F_k - \beta_{1,m-1} \left(\mathbf{1} - \beta_{1,n}\right)^{-1} \left(\sum_{k=1}^{n} \beta_{k,n} F_k \right), \quad \text{(G.2)}$$

for $m = 1, \ldots, n$, where $\beta_{k,m} = Df(x_m)Df(x_{m-1}) \cdots Df(x_k)$ for $k < m$ and $\beta_{k,m} = \mathbf{1}$ for $k \geq m$. Therefore, finding n-cycles by a Newton-Raphson method with multiple shooting requires the inversing of a $[d \times d]$ matrix $\mathbf{1} - Df(x_n)Df(x_{n-1}) \cdots Df(x_1)$.

G.2 Hybrid Newton-Raphson / relaxation method

Consider a d-dimensional map $x' = f(x)$ with an unstable fixed point x_*. The transformed map is the following one:

$$x' = g(x) = x + \gamma C(f(x) - x),$$

where $\gamma > 0$ and C is a $d \times d$ invertible constant matrix. We notice that x_* is also a fixed point of g. Consider the stability matrix at the fixed point x_*

$$A_g = \left. \frac{dg}{dx} \right|_{x=x_*} = 1 + \gamma C(A_f - 1).$$

The matrix C is constructed such that the eigenvalues of A_g are of modulus less than one. Assume that A_f is diagonalizable: In the basis of diagonalization, the matrix writes:

$$\tilde{A}_g = 1 + \gamma \tilde{C}(\tilde{A}_f - 1),$$

where $\tilde{A}_f$ is diagonal with elements μ_i. We restrict the set of matrices $\tilde{C}$ to diagonal matrices with $\tilde{C}_{ii} = \epsilon_i$ where $\epsilon_i = \pm 1$. Thus $\tilde{A}_g$ is diagonal with eigenvalues

$\gamma_i = 1 + \gamma\epsilon_i(\mu_i - 1)$. The choice of γ and ϵ_i is such that $|\gamma_i| < 1$. It is easy to see that if $\mathrm{Re}(\mu_i) < 1$ one has to choose $\epsilon_i = 1$, and if $\mathrm{Re}(\mu_i) > 1$, $\epsilon_i = -1$. If λ is chosen such that

$$0 < \gamma < \min_{i=1,\ldots,d} \frac{2|\mathrm{Re}(\mu_i) - 1|}{|\mu_i - 1|^2},$$

all the eigenvalues of A_g have modulus less that one. The contraction rate at the fixed point for the map g is then $\max_i |1 + \gamma\epsilon_i(\mu_i - 1)|$. We notice that if $\mathrm{Re}(\mu_i) = 1$, it is not possible to stabilize x_* by the set of matrices γC.
From the construction of C, we see that 2^d choices of matrices are possible. For example, for two-dimensional systems, these matrices are

$$C \in \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\}.$$

For 2-dimensional dissipative maps, the eigenvalues satisfy $\mathrm{Re}(\mu_1)\mathrm{Re}(\mu_2) \leq \det Df < 1$. The case $(\mathrm{Re}(\mu_1) > 1, \mathrm{Re}(\mu_2) > 1)$ which is stabilized by $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ has to be discarded. The minimal set is reduced to three matrices.

Appendix H

Applications

> Man who says it cannot be done should not interrupt man doing it.
> Sayings of Vattay Gábor

In this appendix we show that the multidimensional Lyapunov exponents and relaxation exponents (dynamo rates) of vector fields can be expressed in terms of leading eigenvalues of appropriate evolution operators.

H.1 Evolution operator for Lyapunov exponents

Lyapunov exponents were introduced and computed for 1-d maps in sect. **??**. For higher-dimensional flows only the Jacobian matrices are multiplicative, not individual eigenvalues, and the construction of the evolution operator for evaluation of the Lyapunov spectra requires the extension of evolution equations to the flow in the tangent space. We now develop the requisite theory.

Here we construct a multiplicative evolution operator (H.4) whose spectral determinant (H.8) yields the leading Lyapunov exponent of a d-dimensional flow (and is entire for Axiom A flows).

The key idea is to extending the dynamical system by the tangent space of the flow, suggested by the standard numerical methods for evaluation of Lyapunov exponents: start at x_0 with an initial infinitesimal tangent space vector $\eta(0) \in TU_x$, and let the flow transport it along the trajectory $x(t) = f^t(x_0)$.

The dynamics in the $(x, \eta) \in U \times TU_x$ space is governed by the system of

equations of variations [21]:

$$\dot{x} = \mathbf{v}(x)\,, \quad \dot{\eta} = \mathbf{D}\mathbf{v}(x)\eta\,.$$

Here $\mathbf{D}\mathbf{v}(x)$ is the derivative matrix of the flow. We write the solution as

$$x(t) = f^t(x_0)\,, \quad \eta(t) = \mathbf{J}^t(x_0)\cdot\eta_0\,, \tag{H.1}$$

with the tangent space vector η transported by the stability matrix $\mathbf{J}^t(x_0) = \partial x(t)/\partial x_0$.

As explained in sect. 4.1, the growth rate of this vector is multiplicative along the trajectory and can be represented as $\eta(t) = |\eta(t)|/|\eta(0)|\mathbf{u}(t)$ where $\mathbf{u}(t)$ is a "unit" vector in some norm $||.||$. For asymptotic times and for almost every initial $(x_0, \eta(0))$, this factor converges to the leading eigenvalue of the linearized stability matrix of the flow.

We implement this multiplicative evaluation of stability eigenvalues by adjoining the d-dimensional transverse tangent space $\eta \in TU_x$; $\eta(x)\mathbf{v}(x) = 0$ to the $(d+1)$-dimensional dynamical evolution space $x \in U \subset \mathbb{R}^{d+1}$. In order to determine the length of the vector η we introduce a homogeneous differentiable scalar function $g(\eta) = ||\eta||$. It has the property $g(\Lambda\eta) = |\Lambda|g(\eta)$ for any Λ. An example is the projection of a vector to its dth component

$$g\begin{pmatrix}\eta_1\\ \eta_2\\ \cdots\\ \eta_d\end{pmatrix} = |\eta_d|\,.$$

Any vector $\eta \in TU_x$ can now be represented by the product $\eta = \Lambda\mathbf{u}$, where $\mathbf{u}$ is a "unit" vector in the sense that its norm is $||\mathbf{u}|| = 1$, and the factor

$$\Lambda^t(x_0, \mathbf{u}_0) = g(\eta(t)) = g(\mathbf{J}^t(x_0)\cdot\mathbf{u}_0) \tag{H.2}$$

is the multiplicative "stretching" factor.

Unlike the leading eigenvalue of the Jacobian the stretching factor is multiplicative along the trajectory:

$$\Lambda^{t'+t}(x_0, \mathbf{u}_0) = \Lambda^{t'}(x(t), \mathbf{u}(t))\,\Lambda^t(x_0, \mathbf{u}_0).$$

H.1 on p. 635

The $\mathbf{u}$ evolution constrained to $ET_{g,x}$, the space of unit transverse tangent vectors, is given by rescaling of (H.1):

$$\mathbf{u}' = R^t(x, \mathbf{u}) = \frac{1}{\Lambda^t(x, \mathbf{u})}\mathbf{J}^t(x)\cdot\mathbf{u}\,. \tag{H.3}$$

Eqs. (H.1), (H.2) and (H.3) enable us to define a *multiplicative* evolution operator on the extended space $U \times ET_{g,x}$

$$\mathcal{L}^t(x', \mathbf{u}'; x, \mathbf{u}) = \delta\big(x' - f^t(x)\big)\,\frac{\delta\big(\mathbf{u}' - R^t(x, \mathbf{u})\big)}{|\Lambda^t(x, \mathbf{u})|^{\beta-1}}\,, \tag{H.4}$$

where β is a variable.

To evaluate the expectation value of $\log|\Lambda^t(x, \mathbf{u})|$ which is the Lyapunov exponent we again have to take the proper derivative of the leading eigenvalue of (H.4). In order to derive the trace formula for the operator (H.4) we need to evaluate $\mathrm{Tr}\,\mathcal{L}^t = \int dx d\mathbf{u}\,\mathcal{L}^t(\mathbf{u}, x; \mathbf{u}, x)$. The $\int dx$ integral yields a weighted sum over prime periodic orbits p and their repetitions r:

$$\begin{aligned}\mathrm{Tr}\,\mathcal{L}^t &= \sum_p T_p \sum_{r=1}^{\infty} \frac{\delta(t - rT_p)}{|\det(1 - \mathbf{J}_p^r)|}\Delta_{p,r},\\ \Delta_{p,r} &= \int_g d\mathbf{u}\,\frac{\delta\big(\mathbf{u} - R^{T_p r}(x_p, \mathbf{u})\big)}{|\Lambda^{T_p r}(x_p, \mathbf{u})|^{\beta-1}}\,,\end{aligned} \tag{H.5}$$

where $\mathbf{J}_p$ is the prime cycle p transverse stability matrix. As we shall see below, $\Delta_{p,r}$ is intrinsic to cycle p, and independent of any particular cycle point x_p.

We note next that if the trajectory $f^t(x)$ is periodic with period T, the tangent space contains d periodic solutions

$$\mathbf{e}_i(x(T+t)) = \mathbf{e}_i(x(t))\,, \quad i = 1, ..., d,$$

corresponding to the d unit eigenvectors $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_d\}$ of the transverse stability matrix, with "stretching" factors (H.2) given by its eigenvalues

$$\mathbf{J}_p(x)\cdot\mathbf{e}_i(x) = \Lambda_{p,i}\,\mathbf{e}_i(x)\,, \quad i = 1, ..., d. \qquad \text{(no summation on } i)$$

The $\int d\mathbf{u}$ integral in (H.5) picks up contributions from these periodic solutions. In order to compute the stability of the ith eigendirection solution, it is convenient to

expand the variation around the eigenvector $\mathbf{e}_i$ in the stability matrix eigenbasis $\delta\mathbf{u} = \sum \delta u_\ell \, \mathbf{e}_\ell$. The variation of the map (H.3) at a complete period $t = T$ is then given by

$$\begin{aligned} \delta R^T(\mathbf{e}_i) &= \frac{\mathbf{J}\cdot\delta\mathbf{u}}{g(\mathbf{J}\cdot\mathbf{e}_i)} - \frac{\mathbf{J}\cdot\mathbf{e}_i}{g(\mathbf{J}\cdot\mathbf{e}_i)^2}\left(\frac{\partial g(\mathbf{e}_i)}{\partial\mathbf{u}}\cdot\mathbf{J}\cdot\delta\mathbf{u}\right) \\ &= \sum_{k\neq i}\frac{\Lambda_{p,k}}{\Lambda_{p,i}}\left(\mathbf{e}_k - \mathbf{e}_i\frac{\partial g(\mathbf{e}_i)}{\partial u_k}\right)\delta u_k \,. \end{aligned} \tag{H.6}$$

The δu_i component does not contribute to this sum since $g(\mathbf{e}_i + du_i\mathbf{e}_i) = 1 + du_i$ implies $\partial g(\mathbf{e}_i)/\partial u_i = 1$. Indeed, infinitesimal variations $\delta\mathbf{u}$ must satisfy

$$g(\mathbf{u}+\delta\mathbf{u}) = g(\mathbf{u}) = 1 \quad\Longrightarrow\quad \sum_{\ell=1}^{d}\delta u_\ell\frac{\partial g(\mathbf{u})}{\partial u_\ell} = 0\,,$$

so the allowed variations are of form

$$\delta\mathbf{u} = \sum_{k\neq i}\left(\mathbf{e}_k - \mathbf{e}_i\frac{\partial g(\mathbf{e}_i)}{\partial u_k}\right)c_k\,, \quad |c_k| \ll 1\,,$$

and in the neighborhood of the $\mathbf{e}_i$ eigenvector the $\int d\mathbf{u}$ integral can be expressed as

$$\int_g d\mathbf{u} = \int\prod_{k\neq i} dc_k\,.$$

Inserting these variations into the $\int d\mathbf{u}$ integral we obtain

$$\begin{aligned} \int_g d\mathbf{u} \qquad & \delta\big(\mathbf{e}_i + \delta\mathbf{u} - R^T(\mathbf{e}_i) - \delta R^T(\mathbf{e}_i) + \ldots\big) \\ = \qquad & \int\prod_{k\neq i} dc_k\,\delta((1-\Lambda_k/\Lambda_i)c_k + \ldots) \\ & = \prod_{k\neq i}\frac{1}{|1-\Lambda_k/\Lambda_i|}\,, \end{aligned}$$

and the $\int d\mathbf{u}$ trace (H.5) becomes

$$\Delta_{p,r} = \sum_{i=1}^{d}\frac{1}{|\,\Lambda_{p,i}^r\,|^{\beta-1}}\prod_{k\neq i}\frac{1}{|\,1-\Lambda_{p,k}^r/\Lambda_{p,i}^r\,|}\,. \tag{H.7}$$

The corresponding spectral determinant is obtained by observing that the Laplace transform of the trace (11.19) is a logarithmic derivative $\mathrm{Tr}\,\mathcal{L}(s) = -\frac{d}{ds}\log F(s)$ of the spectral determinant:

$$F(\beta,s) = \exp\left(-\sum_{p,r}\frac{e^{sT_pr}}{r\,|\det(1-\mathbf{J}_p^r)\,|}\Delta_{p,r}(\beta)\right). \tag{H.8}$$

This determinant is the central result of this section. Its zeros correspond to the eigenvalues of the evolution operator (H.4), and can be evaluated by the cycle expansion methods.

The leading zero of (H.8) is called "pressure" (or free energy)

$$P(\beta) = s_0(\beta). \tag{H.9}$$

The average Lyapunov exponent is then given by the first derivative of the pressure at $\beta = 1$:

$$\overline{\lambda} = P'(1). \tag{H.10}$$

The simplest application of (H.8) is to 2-dimensional hyperbolic Hamiltonian maps. The stability eigenvalues are related by $\Lambda_1 = 1/\Lambda_2 = \Lambda$, and the spectral determinant is given by

$$\begin{aligned} F(\beta,z) &= \exp\left(-\sum_{p,r}\frac{z^{rn_p}}{r\,|\,\Lambda_p^r\,|}\frac{1}{(1-1/\Lambda_p^r)^2}\Delta_{p,r}(\beta)\right) \\ \Delta_{p,r}(\beta) &= \frac{|\,\Lambda_p^r\,|^{1-\beta}}{1-1/\Lambda_p^{2r}} + \frac{|\,\Lambda_p^r\,|^{\beta-3}}{1-1/\Lambda_p^{2r}}\,. \end{aligned} \tag{H.11}$$

The dynamics (H.3) can be restricted to a u unit eigenvector neighborhood corresponding to the largest eigenvalue of the Jacobi matrix. On this neighborhood the largest eigenvalue of the Jacobi matrix is the only fixed point, and the spectral determinant obtained by keeping only the largest term the $\Delta_{p,r}$ sum in (H.7) is also entire.

In case of maps it is practical to introduce the logarithm of the leading zero and to call it "pressure"

$$P(\beta) = \log z_0(\beta). \tag{H.12}$$

The average of the Lyapunov exponent of the map is then given by the first derivative of the pressure at $\beta = 1$:

$$\overline{\lambda} = P'(1). \tag{H.13}$$

By factorizing the determinant (H.11) into products of zeta functions we can conclude that the leading zero of the (H.4) can also be recovered from the leading zeta function

$$1/\zeta_0(\beta, z) = \exp\left(-\sum_{p,r} \frac{z^{rn_p}}{r|\Lambda_p^r|^\beta}\right). \tag{H.14}$$

This zeta function plays a key role in thermodynamic applications as we will will see in Chapter 17.

H.2 Advection of vector fields by chaotic flows

Fluid motions can move embedded vector fields around. An example is the magnetic field of the Sun which is "frozen" in the fluid motion. A passively evolving vector field $\mathbf{V}$ is governed by an equation of the form

$$\partial_t \mathbf{V} + \mathbf{u}\cdot\nabla\mathbf{V} - \mathbf{V}\cdot\nabla\mathbf{u} = 0, \tag{H.15}$$

where $\mathbf{u}(x,t)$ represents the velocity field of the fluid. The strength of the vector field can grow or decay during its time evolution. The amplification of the vector field in such a process is called the "dynamo effect". In a strongly chaotic fluid motion we can characterize the asymptotic behavior of the field with an exponent

$$\mathbf{V}(x,t) \sim \mathbf{V}(x)e^{\nu t}, \tag{H.16}$$

where ν is called the fast dynamo rate. The goal of this section is to show that periodic orbit theory can be developed for such a highly non-trivial system as well.

We can write the solution of (H.15) formally, as shown by Cauchy. Let $\mathbf{x}(t,\mathbf{a})$ be the position of the fluid particle that was at the point $\mathbf{a}$ at $t = 0$. Then the field evolves according to

$$\mathbf{V}(\mathbf{x},t) = \mathbf{J}(\mathbf{a},t)\mathbf{V}(\mathbf{a},0) \quad , \tag{H.17}$$

where $\mathbf{J}(\mathbf{a},t) = \partial(\mathbf{x})/\partial(\mathbf{a})$ is the Jacobian matrix of the transformation that moves the fluid into itself $\mathbf{x} = \mathbf{x}(\mathbf{a},t)$.

We write $\mathbf{x} = f^t(\mathbf{a})$, where f^t is the flow that maps the initial positions of the fluid particles into their positions at time t. Its inverse, $\mathbf{a} = f^{-t}(\mathbf{x})$, maps particles at time t and position $\mathbf{x}$ back to their initial positions. Then we can write (H.17)

$$V_i(\mathbf{x},t) = \sum_j \int d^3\mathbf{a}\ \mathcal{L}^t_{ij}(\mathbf{x},\mathbf{a})V_j(\mathbf{a},0) \quad , \tag{H.18}$$

with

$$\mathcal{L}^t_{ij}(\mathbf{x},\mathbf{a}) = \delta(\mathbf{a} - f^{-t}(\mathbf{x}))\frac{\partial x_i}{\partial a_j} \quad . \tag{H.19}$$

For large times, the effect of $\mathcal{L}^t$ is dominated by its leading eigenvalue, $e^{\nu_0 t}$ with $Re(\nu_0) > Re(\nu_i)$, $i = 1,2,3,....$ In this way the transfer operator furnishes the fast dynamo rate, $\nu := \nu_0$.

The trace of the transfer operator is the sum over all periodic orbit contributions, with each cycle weighted by its intrinsic stability

$$\mathbf{Tr}\mathcal{L}^t = \sum_p T_p \sum_{r=1}^{\infty} \frac{\mathrm{tr}\,\mathbf{J}_p^r}{\left|\det\left(\mathbf{1} - \mathbf{J}_p^{-r}\right)\right|}\delta(t - rT_p). \tag{H.20}$$

We can construct the corresponding spectral determinant as usual

$$F(s) = \exp\left[-\sum_p \sum_{r=1}^{\infty} \frac{1}{r}\frac{\mathrm{tr}\,\mathbf{J}_p^r}{\left|\det\left(\mathbf{1} - \mathbf{J}_p^{-r}\right)\right|}e^{srT_p}\right] \quad . \tag{H.21}$$

Note that in this formuli we have omitted a term arising from the Jacobian transformation along the orbit which would give $1 + \mathrm{tr}\ \mathbf{J}_p^r$ in the numerator rather than just the trace of $\mathbf{J}_p^r$. Since the extra term corresponds to advection along the orbit, and this does not evolve the magnetic field, we have chosen to ignore it. It is also interesting to note that the negative powers of the Jacobian occur in the denominator, since we have f^{-t} in (H.19).

In order to simplify $F(s)$, we factor the denominator cycle stability determinants into products of expanding and contracting eigenvalues. For a 3-dimensional fluid flow with cycles possessing one expanding eigenvalue Λ_p (with $|\Lambda_p| > 1$), and

one contracting eigenvalue λ_p (with $|\lambda_p| < 1$) the determinant may be expanded as follows:

$$\left|\det\left(\mathbf{1}-\mathbf{J}_p^{-r}\right)\right|^{-1} = |(1-\Lambda_p^{-r})(1-\lambda_p^{-r})|^{-1} = |\lambda_p|^r \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \Lambda_p^{-jr}\lambda_p^{kr} \quad . \text{(H.22)}$$

With this decomposition we can rewrite the exponent in (H.21) as

$$\sum_p \sum_{r=1}^{\infty} \frac{1}{r} \frac{(\lambda_p^r+\Lambda_p^r)e^{srT_p}}{\left|\det\left(\mathbf{1}-\mathbf{J}_p^{-r}\right)\right|} = \sum_p \sum_{j,k=0}^{\infty}\sum_{r=1}^{\infty} \frac{1}{r}\left(|\lambda_p|\Lambda_p^{-j}\lambda_p^k e^{sT_p}\right)^r (\lambda_p^r+\Lambda_p^r) \quad , \text{(H.23)}$$

which has the form of the expansion of a logarithm:

$$\sum_p \sum_{j,k} \left[\log\left(1-e^{sT_p}|\lambda_p|\Lambda_p^{1-j}\lambda_p^k\right) + \log\left(1-e^{sT_p}|\lambda_p|\Lambda_p^{-j}\lambda_p^{1+k}\right)\right] \quad . \text{(H.24)}$$

The spectral determinant is therefore of the form,

$$F(s) = F_e(s)F_c(s) \quad , \text{(H.25)}$$

where

$$F_e(s) = \prod_p \prod_{j,k=0}^{\infty} \left(1-t_p^{(jk)}\Lambda_p\right) \quad , \text{(H.26)}$$

$$F_c(s) = \prod_p \prod_{j,k=0}^{\infty} \left(1-t_p^{(jk)}\lambda_p\right) , \text{(H.27)}$$

with

$$t_p^{(jk)} = e^{sT_p}|\lambda_p|\frac{\lambda_p^k}{\Lambda_p^j} \quad . \text{(H.28)}$$

The two factors present in $F(s)$ correspond to the expanding and contracting exponents. (Had we not neglected a term in (H.21), there would be a third factor corresponding to the translation.)

For 2-d Hamiltonian volume preserving systems, $\lambda = 1/\Lambda$ and (H.26) reduces to

$$F_e(s) = \prod_p \prod_{k=0}^{\infty} \left(1-\frac{t_p}{\Lambda_p^{k-1}}\right)^{k+1} \quad , \quad t_p = \frac{e^{sT_p}}{|\,\Lambda_p\,|} \quad . \text{(H.29)}$$

With $\sigma_p = \Lambda_p/|\Lambda_p|$, the Hamiltonian zeta function (the $j = k = 0$ part of the product (H.27)) is given by

$$1/\zeta_{dyn}(s) = \prod_p \left(1-\sigma_p e^{sT_p}\right) . \text{(H.30)}$$

This is a curious formula — the zeta function depends only on the return times, not on the eigenvalues of the cycles. Furthermore, the identity,

$$\frac{\Lambda+1/\Lambda}{|(1-\Lambda)(1-1/\Lambda)|} = \sigma + \frac{2}{|(1-\Lambda)(1-1/\Lambda)|} \quad ,$$

when substituted into (H.25), leads to a relation between the vector and scalar advection spectral determinants:

$$F_{dyn}(s) = F_0^2(s)/\zeta_{dyn}(s) . \text{(H.31)}$$

The spectral determinants in this equation are entire for hyperbolic (axiom A) systems, since both of them correspond to multiplicative operators.

In the case of a flow governed by a map, we can adapt the formulas (H.29) and (H.30) for the dynamo determinants by simply making the substitution

$$z^{n_p} = e^{sT_p} \quad , \text{(H.32)}$$

where n_p is the integer order of the cycle. Then we find the spectral determinant $F_e(z)$ given by equation (H.29) but with

$$t_p = \frac{z^{n_p}}{|\Lambda_p|} \text{(H.33)}$$

for the weights, and

$$1/\zeta_{dyn}(z) = \Pi_p \left(1-\sigma_p z^{n_p}\right) \text{(H.34)}$$

for the zeta-function

For *maps* with finite Markov partition the inverse zeta function (H.34) reduces to a polynomial for z since curvature terms in the cycle expansion vanish. For example, for maps with complete binary partition, and with the fixed point stabilities of opposite signs, the cycle expansion reduces to

$$1/\zeta_{dyn}(s) = 1. \tag{H.35}$$

For such *maps* the dynamo spectral determinant is simply the square of the scalar advection spectral determinant, and therefore all its zeros are double. In other words, for flows governed by such discrete maps, the fast dynamo rate equals the scalar advection rate.

In contrast, for three-dimensional *flows*, the dynamo effect is distinct from the scalar advection. For example, for flows with finite symbolic dynamical grammars, (H.31) implies that the dynamo zeta function is a ratio of two entire determinants:

$$1/\zeta_{dyn}(s) = F_{dyn}(s)/F_0^2(s)\,. \tag{H.36}$$

This relation implies that for *flows* the zeta function has double poles at the zeros of the scalar advection spectral determinant, with zeros of the dynamo spectral determinant no longer coinciding with the zeros of the scalar advection spectral determinant; Usually the leading zero of the dynamo spectral determinant is larger than the scalar advection rate, and the rate of decay of the magnetic field is no longer governed by the scalar advection.

H.2 on p. 635

Commentary

Remark H.1 Dynamo zeta. The dynamo zeta (H.34) has been introduced by Aurell and Gilbert [3] and reviewed in ref. [5]. Our exposition follows ref. [4].

References

[H.1] J. H. Hannay and A. M. Ozorio de Almeida, *J. Phys.* **A 17**, 3429, (1984).

[H.2] Ya.B. Pesin, *Uspekhi Mat. Nauk* **32,** 55 (1977), [*Russian Math. Surveys* **32,** 55 (1977)].

[H.3] E. Aurell and A. Gilbert, *Geophys. & Astrophys. Fluid Dynamics* (1993).

[H.4] N.J. Balmforth, P. Cvitanović, G.R. Ierley, E.A. Spiegel and G. Vattay, *Stochastic Processes in Astrophysics*, *Annals of New York Academy of Sciences* **706**, 148 (1993).

[H.5] S. Childress and A.D. Gilbert *Stretch, Twist, Fold: The Fast Dynamo*, *Lecture Notes in Physics* **37** (Springer Verlag, Berlin 1995).

[H.6] J. Balatoni and A. Renyi, *Publi. Math. Inst. Hung. Acad. Sci.* **1**, 9 (1956); (english translation **1**, 588 (Akademia Budapest, 1976)).

[H.7] R. Benzi, G. Paladin, G. Parisi and A. Vulpiani, *J. Phys.* **A17**, 3521 (1984).

[H.8] Even though the thermodynamic formalism is of older vintage (we refer the reader to ref. [16] for a comprehensive overview), we adhere here to the notational conventions of ref. [9] which are more current in the physics literature.

[H.9] T.C. Halsey, M.H. Jensen, L.P. Kadanoff, I. Procaccia and B.I. Shraiman, *Phys. Rev.* **A107**, 1141 (1986).

[H.10] 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, em J. Phys. **A17**, 3521 (1984).

[H.11] C. Grebogi, E. Ott and J.A. Yorke, *Phys. Rev.* **A 36**, 3522 (1987).

[H.12] C. Grebogi, E. Ott and J. Yorke, *Physica* **D 7**, 181 (1983).

[H.13] C. Grebogi, E. Ott and J.A. Yorke, *Phys. Rev.* **A36**, 3522 (1987).

[H.14] C. Grebogi, E. Ott and J. Yorke, *Phys. Rev. A* **37,** 1711 (1988).

[H.15] E. Ott, C. Grebogi and J.A. Yorke, *Phys. Lett. A* **135**, 343 (1989).

[H.16] P. Grassberger and I. Procaccia, *Phys. Rev. A* **31**, 1872 (1985).

[H.17] C. Shannon, *Bell System Technical Journal*, **27**, 379 (1948).

[H.18] H. Fujisaka, *Progr. Theor. Phys* **70**, 1264 (1983).

[H.19] A. Politi, R. Badii and P. Grassberger, *J. Phys. A* **15**, L763 (1988); P. Grassberger, R. Badii and A. Politi, *J. Stat. Phys.* **51**, 135 (1988).

[H.20] M. Barnsley, *Fractals Everywhere*, (Academic Press, New York 1988).

[H.21] J. Balatoni and A. Renyi, *Publi. Math. Inst. Hung. Acad. Sci.* **1**, 9 (1956); (english translation **1**, 588 (Akademia Budapest, 1976)).

[H.22] J.D. Crawford and J.R. Cary, *Physica* **D6**, 223 (1983)

[H.23] P. Collet and S. Isola, *Commun.Math.Phys.* **139**, 551 (1991)

[H.24] F. Christiansen, S. Isola, G. Paladin and H.H. Rugh, *J. Phys.* **A 23**, L1301 (1990).

[H.25] A.S. Pikovsky, unpublished.

[H.26] C. Beck, "Higher correlation functions of chaotic dynamical systems - a graph theoretical approach", *Nonlinearity* **4**, 1131 (1991); to be published.

[H.27] The 4-point correlation function is given in ref. [26].

[H.28] G. Hackenbroich and F. von Oppen, "Semiclassical theory of transport in antidot lattices", *Z. Phys.* **B 97**, 157 (1995).

[H.29] M.J. Feigenbaum, *J. Stat. Phys.* **21**, 669 (1979); reprinted in ref. [4].

[H.30] P. Szépfalusy, T. Tél, A. Csordás and Z. Kovács, *Phys. Rev.* **A 36**, 3525 (1987).

Exercises

Exercise H.1 **Stretching factor.** *Prove the multiplicative property of the stretching factor (H.2). Why should we extend the phase space with the tangent space?*

Exercise H.2 **Dynamo rate.** *Suppose that the fluid dynamics is highly dissipative and can be well approximated by the piecewise linear map*

$$f(x) = \begin{cases} 1+ax & \text{if} \quad x<0, \\ 1-bx & \text{if} \quad x>0, \end{cases} \tag{H.37}$$

on an appropriate surface of section ($a, b > 2$). Suppose also that the return time is constant T_a for $x < 0$ and T_b for $x > 0$. Show that the dynamo zeta is

$$1/\zeta_{dyn}(s) = 1 - e^{sT_a} + e^{sT_b}. \tag{H.38}$$

Show also that the escape rate is the leading zero of

$$1/\zeta_0(s) = 1 - e^{sT_a}/a - e^{sT_b}/b. \tag{H.39}$$

Calculate the dynamo and the escape rates analytically if $b = a^2$ and $T_b = 2T_a$. Do the calculation for the case when you reverse the signs of the slopes of the map. What is the difference?

Appendix I

Discrete symmetries

I.1 Preliminaries and definitions

(P. Cvitanović and A. Wirzba)

In the following we will define what we mean by the concepts *group*, *representation*, *symmetry of a dynamical system*, and *invariance*.

Group axioms. First, we define a group in abstract terms: A group G is a set of elements $g_1, g_2, g_3, \ldots$ for which a law of *composition* or *group multiplication* is given such that the product $g_2 \circ g_1$ (which we will also just abbreviate as $g_2 g_1$) of any two elements satisfies the following conditions:

1. If $g_1, g_2 \in G$, then $g_2 \circ g_1 \in G$.
2. The group multiplication is associative: $g_3 \circ (g_2 \circ g_1) = (g_3 \circ g_2) \circ g_1$.
3. The group G contains an element e called *identity* such that $g \circ e = e \circ g = g$ for every element $g \in G$.
4. For every element $g \in G$, there exists an unique element $h \in G$ such that $h \circ g = g \circ h = e$. The element h is called *inverse* of g, and is denoted by $h = g^{-1}$.

A *finite* group is a group with a finite number of elements

$$G = \{e, g_2, \ldots, g_{|G|}\} ,$$

where $|G|$, the number of elements, will be referred to as *order* of the group.

Matrix group on vector space. We will now apply these abstract group definitions to the set of $[d \times d]$-dimensional non-singular matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots$ acting in a d-dimensional vector space $V \in \mathbb{R}^d$, that is the product of matrices $\mathbf{A}$ and $\mathbf{B}$ gives the single matrix $\mathbf{C}$, such that

$$\mathbf{C}v = \mathbf{B}(\mathbf{A}v) \in V, \qquad \forall v \in V. \tag{I.1}$$

The identity of the group is the unit matrix $\mathbf{1}$ which leaves all vectors in V unchanged. Every matrix in the group has a unique inverse.

Linear representation of a group. Let us now map the abstract group G *homeomorphically* on a group of matrices $\mathsf{D}(G)$ in the vector space V, that isin such a way that the group properties, especially the group multiplication, are preserved:

1. Any $g \in G$ is mapped to a matrix $\mathbf{D}(g) \in \mathsf{D}(G)$.
2. The group product $g_2 \circ g_1 \in G$ is mapped onto the matrix product $\mathbf{D}(g_2 \circ g_1) = \mathbf{D}(g_2)\mathbf{D}(g_1)$.
3. The associativity is preserved: $\mathbf{D}(g_3 \circ (g_2 \circ g_1)) = \mathbf{D}(g_3)(\mathbf{D}(g_2)\mathbf{D}(g_1)) = (\mathbf{D}(g_3)(\mathbf{D}(g_2))\mathbf{D}(g_1)$.
4. The identity element $e \in G$ is mapped onto the unit matrix $\mathsf{D}(e) = \mathbf{1}$ and the inverse element $g^{-1} \in G$ is mapped onto the inverse matrix $\mathbf{D}(g^{-1}) = [\mathbf{D}(g)]^{-1} \equiv \mathbf{D^{-1}}(g)$.

We call the so defined matrix group $\mathsf{D}(G)$ a linear or matrix *representation* of the group G in the *representation space* V. Note that the matrix operation on a vector is by definition linear. We use the specification *linear* in order to discriminate the matrix representations from other operator representations that do not have to be linear, in general. Throughout this appendix we only consider linear representations.

If the dimensionality of V is d, we say the representation is an *d-dimensional representation* or has the *degree* d. The matrices $\mathbf{D}(g) \in \mathsf{D}(G)$ are non-singular $[d \times d]$ matrices, which we will also just abbreviate as $\mathbf{g}$, that is$x' = \mathbf{g}x$ corresponds to the normal matrix operation $x'_i = \sum_{j=1}^{d} (\mathbf{g})_{ij} x_j = \sum_{j=1}^{d} g_{ij} x_j$.

Character of a representation. The character of $\chi_\alpha(g)$ of an d-dimensional representation $\mathbf{D}(g)$ of the group element g of a discrete group G is defined as trace

$$\chi_\alpha(g) = \sum_{i=1}^{n} \mathbf{D}_{ii}(g) \equiv \mathrm{tr}\, \mathbf{D}(g)\,.$$

Note especially that $\chi(e) = n$, since $\mathbf{D}_{ij}(e) = \delta_{ij}$ for $1 \le i, j \le n$.

Faithful representations. If the homomorphismus mapping G on $\mathsf{D}(G)$ becomes an isomorphism, the representation is said to be *faithful*. In this case the order of the group of matrices $D(G)$ is equal to the order $|G|$ of the group. In general, however, there will be several elements of G that will be mapped on the unit matrix $\mathbf{D}(e) = \mathbf{1}$. This property can be used to define a subgroup $H \subset G$ of the group G consisting of all elements $h \in G$ that are mapped to the unit matrix of a given representation. Then the cosidered representation is a faithful representation of the *factor group* G/H.

Equivalent representations. From this remarks it should be clear that the representation of a group is by no means unique. If the basis in the d-dimensional vector space V is changed, the matrices $\mathbf{D}(g)$ have to be replaced by their transformations $\mathbf{D}'(g)$. In this case, however, the new matrices $\mathbf{D}'(g)$ and the old matrices $\mathbf{D}(g)$ are related by an equivalence transformation through a non-singular matrix $\mathbf{C}$

$$\mathbf{D}'(g) = \mathbf{C}\,\mathbf{D}(g)\,\mathbf{C}^{-1}\,.$$

Thus, the group of matrices $\mathbf{D}'(g)$ form an equivalent representation $\mathsf{D}'(G)$ to the representation $\mathsf{D}(G)$ of the group G. The equivalent representations have the same structure, although the matrices look different. Because of the cylic nature of the trace and because equivalent representations have the same dimension, the character of equivalent representations is the same

$$\chi(g) = \sum_{i=1}^{n} \mathbf{D}'_{ii}(g) = \mathrm{tr}\,\left(\mathbf{D}'(g)\right) = \mathrm{tr}\,\left(\mathbf{C}\mathbf{D}(g)\mathbf{C}^{-1}\right)\,.$$

Regular representation of a group. The *regular* representation of a group is a special representation that is defined as follows: If we define the elements of a finite group as $g_1, g_2, \ldots, g_{|G|}$, the multiplying from the left by any element g_ν permutes the $g_1, g_2, \ldots, g_{|G|}$ among themselves. We can represent the element g_ν by the permutations of the $|G|$ "coordinates" $g_1, g_2, \ldots, g_{|G|}$. Thus for $i, j = 1, \ldots, |G|$, we define the *regular representation*

$$\mathbf{D}_{ij}(g_\nu) = \begin{cases} \delta_{jl_i} & \text{if } g_\nu g_i = g_{l_i} \text{ with } l_i = 1, \ldots, |G|\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

In this regular representation the diagonal elements of all matrices are zero except for the element g_{ν_0} with $g_{\nu_0} g_i = g_i$, that isfor the identity element e. So in the regular representation the character is given by

$$\chi(g) = \begin{cases} 1 & \text{for} \quad g = e\,, \\ 0 & \text{for} \quad g \neq e\,. \end{cases}$$

Passive and active coordinate transformations. We have to discriminate between *active* and *passive* coordinate transformations. An active (coordinate) transformation corresponds to an non-singular $d \times d$ matrix that actively shifts/changes the vector $x \in \mathcal{M}$

$$x \to \mathbf{T}x.$$

The corresponding passive coordinate transformation changes the coordinate system with respect to which the vector $f(x) \in \mathcal{M}$ is measured. Thus it is given by

$$f(x) \to \mathbf{T}^{-1} f(x) = f(\mathbf{T}^{-1} x).$$

Note that the combination of an passive and active coordinate transformation results to the identity

$$f(x) = \mathbf{T}^{-1} f(\mathbf{T}x)\,.$$

On the other hand, the evolution operator $\mathcal{L}(x,y)$ satisfies the following identity

$$\mathcal{L}(x,y) = \left| \det \left(\frac{\partial \mathbf{T}x}{\partial x} \right) \right| \mathcal{L}(\mathbf{T}x, \mathbf{T}y) = |\det \mathbf{T}|\, \mathcal{L}(\mathbf{T}x, \mathbf{T}y).$$

Note the appearance of $\det \mathbf{T}$ instead of $\det \mathbf{T}^{-1}$ and therefore the contravariant transformation property of $\mathcal{L}(x,y)$ in correspondence to maps $f(x)$. If the coordinate transformation $\mathbf{T}$ belongs to the linear non-singular representation of a *discrete* (that isfinite) symmetry group G, then $|\det \mathbf{T}| = 1$, since for any element g of a finite group G, where exists a number m such that

$$g^n \equiv \underbrace{g \circ g \circ \ldots \circ g}_{m \text{ times}} = e.$$

Thus $\mathbf{T}$ corresponds to the m-th root of $\mathbf{1}$ and the modulus of its determinant is unity.

Symmetry of dynamical system. A dynamical system $(\mathcal{M}, f)$ is invariant under a discrete symmetry group $G = \{e, g_2, \ldots, g_{|G|}\}$, if the map $f : \mathcal{M} \to \mathcal{M}$ (or the continous flow f^t) from the d-dimensional manifold $\mathcal{M}$ into itself (with d finite) is invariant

$$f(\mathbf{g}x) = \mathbf{g} f(x)$$

for *any* coordinate $x \in \mathcal{M}$ and *any* finite non-singular linear representation (that isa non-singular $d \times d$ matrix) $\mathbf{g}$ of any element $g \in G$. So a *symmetry for a dynamical system* $(\mathcal{M}, f)$ has to satisfy the two conditions

1) $\mathbf{g}x \in \mathcal{M} \quad \forall x \in \mathcal{M}$ and $\forall g \in G$,
2) $[\mathbf{D}(g), f] = 0 \quad \forall f : \mathcal{M} \to \mathcal{M}$ and $\forall g \in G$.

Group integration. Note the following laws

$$\frac{1}{|G|} \sum_{g \in G} = 1$$

and therefore

$$\frac{1}{|G|} \sum_{g \in G} \mathbf{D}(g_i) = \mathbf{D}(g_{i_0}), \quad i_0 \text{ fixed}\,.$$

However,

$$\frac{1}{|G|} \sum_{g \in G} \mathbf{D}(g) = \mathbf{0},$$

where $\mathbf{0}$ is the zero matrix of same dimension as the representations $\mathbf{D}(g) \in D(G)$. In particular,

$$\frac{1}{|G|} \sum_{g \in G} \chi_\alpha(g) = \frac{1}{|G|} \sum_{g \in G} \sum_{i=1}^{d_\alpha} D(g)_{ii} = 0.$$

Furthermore, if we consider all non-equilavent irreducible representations of a group G, then the quantities $D^{(\alpha)}_{ij}(g)$ for fixed α, i and j

Orthonormality of characters. But what can we say about

$$\frac{1}{|G|}\sum_{g\in G}\chi_\alpha(hg)\chi_\alpha(g^{-1}k^{-1}) \quad \text{with } h,k\in G \text{ fixed ?}$$

Note the following relation

$$\delta_{ab}\delta_{cd}=\frac{1}{n}\delta_{ad}\delta_{cb}+\left(\delta_{ab}\delta_{cd}-\frac{1}{n}\delta_{ad}\delta_{cb}\right).$$

Projection operators. The projection operator onto the α irreducible subspace of dimension d_α is given by

$$\mathbf{P}_\alpha=\frac{d_\alpha}{|G|}\sum_{g\in G}\chi_\alpha(g)\mathbf{g}^{-1}.$$

Note that $\mathbf{P}_\alpha$ is a $[d\times d]$-dimensional matrix as the representation $\mathbf{g}$.

Irreducible subspaces of the evolution operator.

$$\mathcal{L}=\sum_\alpha \mathrm{tr}\,\mathcal{L}_\alpha$$

with

$$\mathcal{L}_\alpha(y,x)=\frac{d_\alpha}{|G|}\sum_{g\in G}\chi_\alpha(g)\mathcal{L}(\mathbf{g}^{-1}y,x),$$

where the prefactor d_α reflects the fact that a d_α-dimensional representation occurs d_α times.

I.2 C_{4v} factorization

If an N-disk arrangement has C_N symmetry, and the disk visitation sequence is given by disk labels $\{\epsilon_1\epsilon_2\epsilon_3\ldots\}$, only the relative increments $\rho_i=\epsilon_{i+1}-\epsilon_i \bmod N$ matter. Symmetries under reflections across axes increase the group to C_{Nv} and add relations between symbols: $\{\epsilon_i\}$ and $\{N-\epsilon_i\}$ differ only by a reflection.

As a consequence of this reflection increments become decrements until the next reflection and vice versa. Consider four equal disks placed on the vertices of a square (fig. I.1a). The symmetry group consists of the identity $\mathbf{e}$, the two reflections σ_x, σ_y across x, y axes, the two diagonal reflections σ_{13}, σ_{24}, and the three rotations C_4, C_2 and C_4^3 by angles $\pi/2$, π and $3\pi/2$. We start by exploiting the C_4 subgroup symmetry in order to replace the absolute labels $\epsilon_i\in\{1,2,3,4\}$ by relative increments $\rho_i\in\{1,2,3\}$. By reflection across diagonals, an increment by 3 is equivalent to an increment by 1 and a reflection; this new symbol will be called $\underline{1}$. Our convention will be to first perform the increment and then to change the orientation due to the reflection. As an example, consider the fundamental domain cycle 112. Taking the disk $1\to$ disk 2 segment as the starting segment, this symbol string is mapped into the disk visitation sequence $1_{+1}2_{+1}3_{+2}1\ldots=\overline{123}$, where the subscript indicates the increments (or decrements) between neighboring symbols; the period of the cycle $\overline{112}$ is thus 3 in both the fundamental domain and the full space. Similarly, the cycle $\overline{1\underline{1}2}$ will be mapped into $1_{+1}2_{-1}1_{-2}3_{-1}2_{+1}3_{+2}1=\overline{121323}$ (note that the fundamental domain symbol $\underline{1}$ corresponds to a flip in orientation after the second and fifth symbols); this time the period in the full space is twice that of the fundamental domain. In particular, the fundamental domain fixed points correspond to the following 4-disk cycles:

4-disk		reduced
12	$\leftrightarrow$	$\underline{1}$
1234	$\leftrightarrow$	1
13	$\leftrightarrow$	2

Conversions for all periodic orbits of reduced symbol period less than 5 are listed in table I.1.

This symbolic dynamics is closely related to the group-theoretic structure of the dynamics: the global 4-disk trajectory can be generated by mapping the fundamental domain trajectories onto the full 4-disk space by the accumulated product of the C_{4v} group elements $g_1=C$, $g_2=C^2$, $g_{\underline{1}}=\sigma_{diag}C=\sigma_{axis}$, where C is a rotation by $\pi/2$. In the $\overline{1\underline{1}2}$ example worked out above, this yields $g_{\underline{1}12}=g_2g_1g_{\underline{1}}=C^2C\sigma_{axis}=\sigma_{diag}$, listed in the last column of table I.1. Our convention is to multiply group elements in the reverse order with respect to the symbol sequence. We need these group elements for our next step, the dynamical zeta function factorizations.

The C_{4v} group has four one-dimensional representations, either symmetric (A_1) or antisymmetric (A_2) under both types of reflections, or symmetric under one and antisymmetric under the other (B_1, B_2), and a degenerate pair of two-dimensional representations E. Substituting the C_{4v} characters

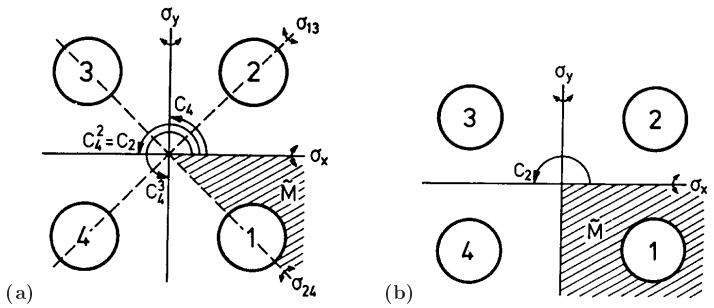


Figure I.1: (a) The symmetries of four disks on a square. (b) The symmetries of four disks on a rectangle. The fundamental domains are indicated by the shaded wedges.

$\bar{p}$	p	$\mathbf{h}_{\bar{p}}$
0	1 2	σ_x
1	1 2 3 4	C_4
2	1 3	C_2, σ_{13}
01	12 14	σ_{24}
02	12 43	σ_y
12	124 134 23	C_4^3
001	121 232 343 414	C_4
002	121 343	C_2
011	121 434	σ_y
012	121 323	σ_{13}
021	124 324	σ_{13}
022	124 213	σ_x
112	123	e
122	124 231 342 413	C_4

$\bar{p}$	p	$\mathbf{h}_{\bar{p}}$
0001	1212 1414	σ_{24}
0002	1212 4343	σ_y
0011	1212 3434	C_2
0012	1212 4141 3434 2323	C_4^3
0021 (a)	1213 4142 3431 2324	C_4^3
0022	1213	e
0102 (a)	1214 2321 3432 4143	C_4
0111	1214 3234	σ_{13}
0112 (b)	1214 2123	σ_x
0121 (b)	1213 2124	σ_x
0122	1213 1413	σ_{24}
0211	1243 2134	σ_x
0212	1243 1423	σ_{24}
0221	1242 1424	σ_{24}
0222	1242 4313	σ_y
1112	1234 2341 3412 4123	C_4
1122	1231 3413	C_2
1222	1242 4131 3424 2313	C_4^3

Table I.1: C_{4v} correspondence between the ternary fundamental domain prime cycles $\bar{p}$ and the full 4-disk {1,2,3,4} labelled cycles p, together with the C_{4v} transformation that maps the end point of the $\bar{p}$ cycle into an irreducible segment of the p cycle. For typographical convenience, the symbol $\underline{1}$ of sect. I.2 has been replaced by 0, so that the ternary alphabet is $\{0,1,2\}$. The degeneracy of the p cycle is $m_p = 8n_{\bar{p}}/n_p$. Orbit $\overline{2}$ is the sole boundary orbit, invariant both under a rotation by π and a reflection across a diagonal. The two pairs of cycles marked by (a) and (b) are related by time reversal, but cannot be mapped into each other by C_{4v} transformations.

C_{4v}	A_1	A_2	B_1	B_2	E
e	1	1	1	1	2
C_2	1	1	1	1	-2
C_4, C_4^3	1	1	-1	-1	0
σ_{axes}	1	-1	1	-1	0
σ_{diag}	1	-1	-1	1	0

into (19.15) we obtain:

$h_{\bar{p}}$			A_1	A_2	B_1	B_2	E
e:	$(1-t_{\bar{p}})^8$	=	$(1-t_{\bar{p}})$	$(1-t_{\bar{p}})$	$(1-t_{\bar{p}})$	$(1-t_{\bar{p}})$	$(1-t_{\bar{p}})^4$
C_2:	$(1-t_{\bar{p}}^2)^4$	=	$(1-t_{\bar{p}})$	$(1-t_{\bar{p}})$	$(1-t_{\bar{p}})$	$(1-t_{\bar{p}})$	$(1+t_{\bar{p}})^4$
C_4, C_4^3:	$(1-t_{\bar{p}}^4)^2$	=	$(1-t_{\bar{p}})$	$(1-t_{\bar{p}})$	$(1+t_{\bar{p}})$	$(1+t_{\bar{p}})$	$(1+t_{\bar{p}}^2)^2$
σ_{axes}:	$(1-t_{\bar{p}}^2)^4$	=	$(1-t_{\bar{p}})$	$(1+t_{\bar{p}})$	$(1-t_{\bar{p}})$	$(1+t_{\bar{p}})$	$(1-t_{\bar{p}}^2)^2$
σ_{diag}:	$(1-t_{\bar{p}}^2)^4$	=	$(1-t_{\bar{p}})$	$(1+t_{\bar{p}})$	$(1+t_{\bar{p}})$	$(1-t_{\bar{p}})$	$(1-t_{\bar{p}}^2)^2$

The possible irreducible segment group elements $\mathbf{h}_{\bar{p}}$ are listed in the first column; σ_{axes} denotes a reflection across either the x-axis or the y-axis, and σ_{diag} denotes a reflection across a diagonal (see fig. I.1a). In addition, degenerate pairs of boundary orbits can run along the symmetry lines in the full space, with the fundamental domain group theory weights $\mathbf{h}_p = (C_2 + \sigma_x)/2$ (axes) and $\mathbf{h}_p = (C_2 + \sigma_{13})/2$ (diagonals) respectively:

			A_1	A_2	B_1	B_2	E
axes:	$(1-t_{\bar{p}}^2)^2$	=	$(1-t_{\bar{p}})(1-0t_{\bar{p}})(1-t_{\bar{p}})(1-0t_{\bar{p}})(1+t_{\bar{p}})^2$				
diagonals:	$(1-t_{\bar{p}}^2)^2$	=	$(1-t_{\bar{p}})(1-0t_{\bar{p}})(1-0t_{\bar{p}})(1-t_{\bar{p}})(1+t_{\bar{p}})^2$				(I.2)

(we have assumed that $t_{\bar{p}}$ does not change sign under reflections across symmetry axes). For the 4-disk arrangement considered here only the diagonal orbits $\overline{13}$, $\overline{24}$ occur; they correspond to the $\overline{2}$ fixed point in the fundamental domain.

The A_1 subspace in C_{4v} cycle expansion is given by

$$\begin{aligned}
1/\zeta_{A_1} &= (1-t_0)(1-t_1)(1-t_2)(1-t_{01})(1-t_{02})(1-t_{12}) \\
&\quad (1-t_{001})(1-t_{002})(1-t_{011})(1-t_{012})(1-t_{021})(1-t_{022})(1-t_{112}) \\
&\quad (1-t_{122})(1-t_{0001})(1-t_{0002})(1-t_{0011})(1-t_{0012})(1-t_{0021})\dots \\
&= 1-t_0-t_1-t_2-(t_{01}-t_0t_1)-(t_{02}-t_0t_2)-(t_{12}-t_1t_2) \\
&\quad -(t_{001}-t_0t_{01})-(t_{002}-t_0t_{02})-(t_{011}-t_1t_{01}) \\
&\quad -(t_{022}-t_2t_{02})-(t_{112}-t_1t_{12})-(t_{122}-t_2t_{12}) \\
&\quad -(t_{012}+t_{021}+t_0t_1t_2-t_0t_{12}-t_1t_{02}-t_2t_{01})\dots \qquad \text{(I.3)}
\end{aligned}$$

(for typographical convenience, $\underline{1}$ is replaced by 0 in the remainder of this section). For one-dimensional representations, the characters can be read off the symbol strings: $\chi_{A_2}(\mathbf{h}_{\tilde{\mathbf{p}}}) = (-1)^{n_0}$, $\chi_{B_1}(\mathbf{h}_{\tilde{\mathbf{p}}}) = (-1)^{n_1}$, $\chi_{B_2}(\mathbf{h}_{\tilde{\mathbf{p}}}) = (-1)^{n_0+n_1}$, where n_0 and n_1 are the number of times symbols 0, 1 appear in the $\tilde{p}$ symbol string. For B_2 all t_p with an odd total number of 0's and 1's change sign:

$$\begin{aligned}
1/\zeta_{B_2} &= (1+t_0)(1+t_1)(1-t_2)(1-t_{01})(1+t_{02})(1+t_{12}) \\
&\quad (1+t_{001})(1-t_{002})(1+t_{011})(1-t_{012})(1-t_{021})(1+t_{022})(1-t_{112}) \\
&\quad (1+t_{122})(1-t_{0001})(1+t_{0002})(1-t_{0011})(1+t_{0012})(1+t_{0021})\dots \\
&= 1+t_0+t_1-t_2-(t_{01}-t_0t_1)+(t_{02}-t_0t_2)+(t_{12}-t_1t_2) \\
&\quad +(t_{001}-t_0t_{01})-(t_{002}-t_0t_{02})+(t_{011}-t_1t_{01}) \\
&\quad +(t_{022}-t_2t_{02})-(t_{112}-t_1t_{12})+(t_{122}-t_2t_{12}) \\
&\quad -(t_{012}+t_{021}+t_0t_1t_2-t_0t_{12}-t_1t_{02}-t_2t_{01})\dots
\end{aligned} \tag{I.4}$$

The form of the remaining cycle expansions depends crucially on the special role played by the boundary orbits: by (I.2) the orbit t_2 does not contribute to A_2 and B_1,

$$\begin{aligned}
1/\zeta_{A_2} &= (1+t_0)(1-t_1)(1+t_{01})(1+t_{02})(1-t_{12}) \\
&\quad (1-t_{001})(1-t_{002})(1+t_{011})(1+t_{012})(1+t_{021})(1+t_{022})(1-t_{112}) \\
&\quad (1-t_{122})(1+t_{0001})(1+t_{0002})(1-t_{0011})(1-t_{0012})(1-t_{0021})\dots \\
&= 1+t_0-t_1+(t_{01}-t_0t_1)+t_{02}-t_{12} \\
&\quad -(t_{001}-t_0t_{01})-(t_{002}-t_0t_{02})+(t_{011}-t_1t_{01}) \\
&\quad +t_{022}-t_{122}-(t_{112}-t_1t_{12})+(t_{012}+t_{021}-t_0t_{12}-t_1t_{02})\dots
\end{aligned} \tag{I.5}$$

and

$$\begin{aligned}
1/\zeta_{B_1} &= (1-t_0)(1+t_1)(1+t_{01})(1-t_{02})(1+t_{12}) \\
&\quad (1+t_{001})(1-t_{002})(1-t_{011})(1+t_{012})(1+t_{021})(1-t_{022})(1-t_{112}) \\
&\quad (1+t_{122})(1+t_{0001})(1-t_{0002})(1-t_{0011})(1+t_{0012})(1+t_{0021})\dots \\
&= 1-t_0+t_1+(t_{01}-t_0t_1)-t_{02}+t_{12} \\
&\quad +(t_{001}-t_0t_{01})-(t_{002}-t_0t_{02})-(t_{011}-t_1t_{01}) \\
&\quad -t_{022}+t_{122}-(t_{112}-t_1t_{12})+(t_{012}+t_{021}-t_0t_{12}-t_1t_{02})\dots
\end{aligned} \tag{I.6}$$

In the above we have assumed that t_2 does not change sign under C_{4v} reflections. For the mixed-symmetry subspace E the curvature expansion is given by

$$\begin{aligned}
1/\zeta_E &= 1+t_2+(-{t_0}^2+{t_1}^2)+(2t_{002}-t_2{t_0}^2-2t_{112}+t_2{t_1}^2) \\
&\quad +(2t_{0011}-2t_{0022}+2t_2t_{002}-{t_{01}}^2-{t_{02}}^2+2t_{1122}-2t_2t_{112} \\
&\quad +{t_{12}}^2-{t_0}^2{t_1}^2)+(2t_{00002}-2t_{00112}+2t_2t_{0011}-2t_{00121}-2t_{00211} \\
&\quad +2t_{00222}-2t_2t_{0022}+2t_{01012}+2t_{01021}-2t_{01102}-t_2{t_{01}}^2+2t_{02022} \\
&\quad -t_2{t_{02}}^2+2t_{11112}-2t_{11222}+2t_2t_{1122}-2t_{12122}+t_2{t_{12}}^2-t_2{t_0}^2{t_1}^2 \\
&\quad +2t_{002}(-{t_0}^2+{t_1}^2)-2t_{112}(-{t_0}^2+{t_1}^2))
\end{aligned} \tag{I.7}$$

A quick test of the $\zeta = \zeta_{A_1}\zeta_{A_2}\zeta_{B_1}\zeta_{B_2}\zeta_E^2$ factorization is afforded by the topological polynomial; substituting $t_p = z^{n_p}$ into the expansion yields

$$1/\zeta_{A_1} = 1-3z\ , \quad 1/\zeta_{A_2} = 1/\zeta_{B_1} = 1\ , \quad 1/\zeta_{B_2} = 1/\zeta_E = 1+z\ ,$$

in agreement with (10.42). 15.9 on p. 292

Remark I.1 Labelling conventions While there is a variety of labelling conventions [15, 13] for the reduced C_{4v} dynamics, we prefer the one introduced here because of its close relation to the group-theoretic structure of the dynamics: the global 4-disk trajectory can be generated by mapping the fundamental domain trajectories onto the full 4-disk space by the accumulated product of the C_{4v} group elements.

I.3 C_{2v} factorization

An arrangement of four identical disks on the vertices of a rectangle has C_{2v} symmetry (fig. I.1b). C_{2v} consists of $\{e, \sigma_x, \sigma_y, C_2\}$, *i.e.*, the reflections across the symmetry axes and a rotation by π.

This system affords a rather easy visualization of the conversion of a 4-disk dynamics into a fundamental domain symbolic dynamics. An orbit leaving the fundamental domain through one of the axis may be folded back by a reflection on that axis; with these symmetry operations $g_0 = \sigma_x$ and $g_1 = \sigma_y$ we associate labels 1 and 0, respectively. Orbits going to the diagonally opposed disk cross the boundaries of the fundamental domain twice; the product of these two reflections is just $C_2 = \sigma_x\sigma_y$, to which we assign the label 2. For example, a ternary string 0 0 1 0 2 0 1 ... is converted into 12143123..., and the associated group-theory weight is given by $\dots g_1g_0g_2g_0g_1g_0g_0$.

Short ternary cycles and the corresponding 4-disk cycles are listed in table I.2. Note that already at length three there is a pair of cycles ($012 = 143$ and $021 = 142$) related by time reversal, but *not* by any C_{2v} symmetries.

$\bar{p}$	p	g
0	14	σ_y
1	12	σ_x
2	13	C_2
01	14 32	C_2
02	14 23	σ_x
12	12 43	σ_y
001	141 232	σ_x
002	141 323	C_2
011	143 412	σ_y
012	143	e
021	142	e
022	142 413	σ_y
112	121 343	C_2
122	124 213	σ_x

$\bar{p}$	p	g
0001	1414 3232	C_2
0002	1414 2323	σ_x
0011	1412	e
0012	1412 4143	σ_y
0021	1413 4142	σ_y
0022	1413	e
0102	1432 4123	σ_y
0111	1434 3212	C_2
0112	1434 2343	σ_x
0121	1431 2342	σ_x
0122	1431 3213	C_2
0211	1421 2312	σ_x
0212	1421 3243	C_2
0221	1424 3242	C_2
0222	1424 2313	σ_x
1112	1212 4343	σ_y
1122	1213	e
1222	1242 4313	σ_y

Table I.2: C_{2v} correspondence between the ternary $\{0,1,2\}$ fundamental domain prime cycles $\bar{p}$ and the full 4-disk {1,2,3,4} cycles p, together with the C_{2v} transformation that maps the end point of the $\bar{p}$ cycle into an irreducible segment of the p cycle. The degeneracy of the p cycle is $m_p = 4n_{\bar{p}}/n_p$. Note that the 012 and 021 cycles are related by time reversal, but cannot be mapped into each other by C_{2v} transformations. The full space orbit listed here is generated from the symmetry reduced code by the rules given in sect. I.3, starting from disk 1.

The above is the complete description of the symbolic dynamics for 4 sufficiently separated equal disks placed at corners of a rectangle. However, if the fundamental domain requires further partitioning, the ternary description is insufficient. For example, in the stadium billiard fundamental domain one has to distinguish between bounces off the straight and the curved sections of the billiard wall; in that case five symbols suffice for constructing the covering symbolic dynamics.

The group C_{2v} has four one-dimensional representations, distinguished by their behavior under axis reflections. The A_1 representation is symmetric with respect to both reflections; the A_2 representation is antisymmetric with respect to both. The B_1 and B_2 representations are symmetric under one and antisymmetric under the other reflection. The character table is

C_{2v}	A_1	A_2	B_1	B_2
e	1	1	1	1
C_2	1	1	-1	-1
σ_x	1	-1	1	-1
σ_y	1	-1	-1	1

Substituted into the factorized determinant (19.14), the contributions of pe-

riodic orbits split as follows

$$\begin{array}{rrcccc} g_{\bar{p}} & & & A_1 & A_2 & B_1 & B_2 \\ e\colon & (1-t_{\bar{p}})^4 & = & (1-t_{\bar{p}}) & (1-t_{\bar{p}}) & (1-t_{\bar{p}}) & (1-t_{\bar{p}}) \\ C_2\colon & (1-t_{\bar{p}}^2)^2 & = & (1-t_{\bar{p}}) & (1-t_{\bar{p}}) & (1-t_{\bar{p}}) & (1-t_{\bar{p}}) \\ \sigma_x\colon & (1-t_{\bar{p}}^2)^2 & = & (1-t_{\bar{p}}) & (1+t_{\bar{p}}) & (1-t_{\bar{p}}) & (1+t_{\bar{p}}) \\ \sigma_y\colon & (1-t_{\bar{p}}^2)^2 & = & (1-t_{\bar{p}}) & (1+t_{\bar{p}}) & (1+t_{\bar{p}}) & (1-t_{\bar{p}}) \end{array}$$

Cycle expansions follow by substituting cycles and their group theory factors from table I.2. For A_1 all characters are $+1$, and the corresponding cycle expansion is given in (I.3). Similarly, the totally antisymmetric subspace factorization A_2 is given by (I.4), the B_2 factorization of C_{4v}. For B_1 all t_p with an odd total number of 0's and 2's change sign:

$$\begin{aligned} 1/\zeta_{B_1} &= (1+t_0)(1-t_1)(1+t_2)(1+t_{01})(1-t_{02})(1+t_{12}) \\ &\quad (1-t_{001})(1+t_{002})(1+t_{011})(1-t_{012})(1-t_{021})(1+t_{022})(1+t_{112}) \\ &\quad (1-t_{122})(1+t_{0001})(1-t_{0002})(1-t_{0011})(1+t_{0012})(1+t_{0021})\dots \\ &= 1+t_0-t_1+t_2+(t_{01}-t_0t_1)-(t_{02}-t_0t_2)+(t_{12}-t_1t_2) \\ &\quad -(t_{001}-t_0t_{01})+(t_{002}-t_0t_{02})+(t_{011}-t_1t_{01}) \\ &\quad +(t_{022}-t_2t_{02})+(t_{112}-t_1t_{12})-(t_{122}-t_2t_{12}) \\ &\quad -(t_{012}+t_{021}+t_0t_1t_2-t_0t_{12}-t_1t_{02}-t_2t_{01})\dots \end{aligned} \tag{I.8}$$

For B_2 all t_p with an odd total number of 1's and 2's change sign:

$$\begin{aligned} 1/\zeta_{B_2} &= (1-t_0)(1+t_1)(1+t_2)(1+t_{01})(1+t_{02})(1-t_{12}) \\ &\quad (1+t_{001})(1+t_{002})(1-t_{011})(1-t_{012})(1-t_{021})(1-t_{022})(1+t_{112}) \\ &\quad (1+t_{122})(1+t_{0001})(1+t_{0002})(1-t_{0011})(1-t_{0012})(1-t_{0021})\dots \\ &= 1-t_0+t_1+t_2+(t_{01}-t_0t_1)+(t_{02}-t_0t_2)-(t_{12}-t_1t_2) \\ &\quad +(t_{001}-t_0t_{01})+(t_{002}-t_0t_{02})-(t_{011}-t_1t_{01}) \\ &\quad -(t_{022}-t_2t_{02})+(t_{112}-t_1t_{12})+(t_{122}-t_2t_{12}) \\ &\quad -(t_{012}+t_{021}+t_0t_1t_2-t_0t_{12}-t_1t_{02}-t_2t_{01})\dots \end{aligned} \tag{I.9}$$

Note that all of the above cycle expansions group long orbits together with their pseudoorbit shadows, so that the shadowing arguments for convergence still apply.

The topological polynomial factorizes as

$$\frac{1}{\zeta_{A_1}} = 1-3z \quad , \quad \frac{1}{\zeta_{A_2}} = \frac{1}{\zeta_{B_1}} = \frac{1}{\zeta_{B_2}} = 1+z,$$

consistent with the 4-disk factorization (10.42).

Remark I.2 C_{2v} symmetry C_{2v} is the symmetry of several systems studied in the literature, such as the stadium billiard [?], and the 2-dimensional anisotropic Kepler potential [24].

I.4 Hénon map symmetries

We note here a few simple symmetries of the Hénon map (3.13). For $b \neq 0$ the Hénon map is reversible: the backward iteration of (3.14) is given by

$$x_{n-1} = -\frac{1}{b}(1 - ax_n^2 - x_{n+1}) \,. \tag{I.10}$$

Hence the time reversal amounts to $b \to 1/b$, $a \to a/b^2$ symmetry in the parameter plane, together with $x \to -x/b$ in the coordinate plane, and there is no need to explore the (a, b) parameter plane outside the strip $b \in \{-1, 1\}$. For $b = -1$ the map is orientation and area preserving (see (17.1) below),

$$x_{n-1} = 1 - ax_n^2 - x_{n+1} \,, \tag{I.11}$$

the backward and the forward iteration are the same, and the non–wandering set is symmetric across the $x_{n+1} = x_n$ diagonal. This is one of the simplest models of a Poincaré return map for a Hamiltonian flow. For the orientation reversing $b = 1$ case we have

$$x_{n-1} = 1 - ax_n^2 + x_{n+1} \,, \tag{I.12}$$

and the non–wandering set is symmetric across the $x_{n+1} = -x_n$ diagonal.

I.5 Symmetries of the symbol square

• advanced section •

Depending on the type of dynamical system, the symbol square might have a variety of symmetries. Under the time reversal

$$\cdots s_{-2}s_{-1}s_0.s_1s_2s_3\cdots \to \cdots s_3s_2s_1.s_0s_{-1}s_{-2}\cdots$$

the points in the symbol square for an orientation preserving map are symmetric across the diagonal $\gamma = \delta$, and for the orientation reversing case they are

symmetric with respect to the $\gamma = 1 - \delta$ diagonal. Consequently the periodic orbits appear either in dual pairs $p = s_1s_2s_3\ldots s_n$, $\overline{p} = s_ns_{n-1}s_{n-2}\ldots s_1$, or are self-dual under time reversal, $S_p = S_{\overline{p}}$. For the orientation preserving case a self-dual cycle of odd period has at least one point on the symmetry diagonal. In particular, all fixed points lie on the symmetry diagonal. Determination of such symmetry lines can be of considerable practical utility, as it reduces some of the periodic orbit searches to 1-dimensional searches.

Remark I.3 Symmetries of the symbol square. For a more detailed discussion of the symbolic dynamics symmetries, see refs. [4, ?].

Appendix J

Convergence of spectral determinants

J.1 Curvature expansions: geometric picture

If you has some experience with numerical estimates of fractal dimensions, you will note that the numerical convergence of cycle expansions for systems such as the 3-disk game of pinball, table 15.2, is very impressive; only three input numbers (the two fixed points $\overline{0}$, $\overline{1}$ 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 cycles work so well?

Looking at the cycle expansions simply as sums of unrelated contributions is not specially encouraging: the cycle expansion (15.2) is not absolutely convergent in the sense of Dirichlet series of sect. 15.5, so what one makes of it depends on the way the terms are arranged.

The simplest estimate of the error introduced by approximating smooth flow by periodic orbits is to think of the approximation as a tessalation of a smooth curve by piecewise linear tiles, fig. 1.9.

J.1.1 Tessalation of a smooth flow by cycles

One of the early high accuracy computations of π was due to Euler. Euler computed the circumference of the circee of unit radius by inscribing into it a regular polygon with N sides; the error of such computation is proportional to $1-\cos(2\pi/N) \propto N^{-2}$. In a periodic orbit tessalation of a smooth flow, we cover the phase space by e^{hn} tiles at the nth level of resolution, where h is the

topological entropy, the growth rate of the number of tiles. Hence we expect the error in approximating a smooth flow by e^{hn} linear segments to be exponentially small, of order $N^{-2} \propto e^{-2hn}$.

J.1.2 Shadowing and convergence of curvature expansions

We have shown in chapter 10 that if the symbolic dynamics is defined by a finite grammar, a finite number of cycles, let us say the first k terms in the cycle expansion are necessary to correctly count the pieces of the Cantor set generated by the dynamical system.

They are composed of products of non–intersecting loops on the Markov graph, see (10.13). We refer to this set of non–intersecting loops as the *fundamental* cycles of the strange set. It is only after these terms have been included that the cycle expansion is expected to converge smoothly, that is only for $n > k$ are the curvatures c_n in (9.2??) a measure of the variation of the quality of a linearized covering of the dynamical Cantor set by the length n cycles, and expected to fall off rapidly with n.

The rate of fall-off of the cycle expansion coefficients can be estimated by observing that for subshifts of finite type the contributions from longer orbits in curvature expansions such as (15.5) can always be grouped into shadowing combinations of pseudo-cycles. For example, a cycle with itinerary $\overline{ab}= s_1 s_2 \cdots s_n$ will appear in combination of form

$$1/\zeta = 1 - \cdots - (t_{ab} - t_a t_b) - \cdots ,$$

with $\overline{ab}$ shadowed by cycle $\overline{a}$ followed by cycle $\overline{b}$, where $a = s_1 s_2 \cdots s_m$, $b = s_{m+1} \cdots s_{n-1} s_n$, and s_k labels the Markov partition $\mathcal{M}_{s_k}$ (9.4) that the trajectory traverses at the kth return. If the two trajectories coincide in the first m symbols, at the mth return to a Poincaré section they can land anywhere in the phase space $\mathcal{M}$

$$\left| f^{T_a}(x_a) - f^{T_{a\cdots}}(x_{a\cdots}) \right| \approx 1 ,$$

where we have assumed that the $\mathcal{M}$ is compact, and that the maximal possible separation across $\mathcal{M}$ is $O(1)$. Here x_a is a point on the $\overline{a}$ cycle of period T_a, and $x_{a\cdots}$ is a nearby point whose trajectory tracks the cycle $\overline{a}$ for the first m Poincaré section returns completed at the time $T_{a\cdots}$. An estimate of the maximal separation of the initial points of the two neighboring trajectories is achieved by Taylor expanding around $x_{a\cdots} = x_{\overline{a}} + \delta x_{a\cdots}$

$$f^{T_a}(x_{\overline{a}}) - f^{T_{a\cdots}}(x_{a\cdots}) \approx \frac{\partial f^{T_a}(x_{\overline{a}})}{\partial x} \cdot \delta x_{a\cdots} = \mathbf{J}_a \cdot \delta x_{a\cdots} ,$$

hence the hyperbolicity of the flow forces the initial points of neighboring trajectories that track each other for at least m consecutive symbols to lie exponentially close

$$|\delta x_{a\cdots}| \propto \frac{1}{|\Lambda_a|} .$$

Similarly, for any observable (8.1) integrated along the two nearby trajectories

$$A^{T_{a\cdots}}(x_{a\cdots}) \approx A^{T_a}(x_{\overline{a}}) + \left.\frac{\partial A^{T_a}}{\partial x}\right|_{x=x_{\overline{a}}} \cdot \delta x_{a\cdots} ,$$

so

$$\left| A^{T_{a\cdots}}(x_{a\cdots}) - A^{T_a}(x_{\overline{a}}) \right| \propto \frac{T_a \mathrm{Const}}{|\Lambda_a|} ,$$

As the time of return is itself an integral along the trajectory, return times of nearby trajectories are exponentially close

$$|T_{a\cdots} - T_a| \propto \frac{T_a \mathrm{Const}}{|\Lambda_a|} ,$$

and so are the trajectory stabilities

$$\left| A^{T_{a\cdots}}(x_{a\cdots}) - A^{T_a}(x_{\overline{a}}) \right| \propto \frac{T_a \mathrm{Const}}{|\Lambda_a|} ,$$

Substituting t_{ab} one finds

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

Since with increasing m segments of $\overline{ab}$ come closer to $\overline{a}$, the differences in action and the ratio of the eigenvalues converge exponentially with the eigenvalue of the orbit $\overline{a}$,

$$T_a + T_b - T_{ab} \approx \mathrm{Const} \times \Lambda_a^{-j} , \quad |\Lambda_a \Lambda_b / \Lambda_{ab}| \approx \exp(-\mathrm{Const}/\Lambda_{ab})$$

n	$t_{ab} - t_a t_b$	$T_{ab} - (T_a + T_b)$	$\log \left\lvert \frac{\Lambda_a \Lambda_b}{\Lambda_{ab}} \right\rvert$	$ab - a \cdot b$
2	-5.23465150784$\times 10^{4}$	4.85802927371$\times 10^{2}$	-6.3$\times 10^{2}$	01-0·1
3	-7.96028600139$\times 10^{6}$	5.21713101432$\times 10^{3}$	-9.8$\times 10^{3}$	001-0·01
4	-1.03326529874$\times 10^{7}$	5.29858199419$\times 10^{4}$	-1.3$\times 10^{3}$	0001-0·001
5	-1.27481522016$\times 10^{9}$	5.35513574697$\times 10^{5}$	-1.6$\times 10^{4}$	00001-0·0001
6	-1.52544704823$\times 10^{11}$	5.40999882625$\times 10^{6}$	-1.8$\times 10^{5}$	000001-0·00001
2	-5.23465150784$\times 10^{4}$	4.85802927371$\times 10^{2}$	-6.3$\times 10^{2}$	01-0·1
3	5.30414752996$\times 10^{6}$	-3.67093656690$\times 10^{3}$	7.7$\times 10^{3}$	011-01·1
4	-5.40934261680$\times 10^{8}$	3.14925761316$\times 10^{4}$	-9.2$\times 10^{4}$	0111-011·1
5	4.99129508833$\times 10^{10}$	-2.67292822795$\times 10^{5}$	1.0$\times 10^{4}$	01111-0111·1
6	-4.39246000586$\times 10^{12}$	2.27087116266$\times 10^{6}$	-1.0$\times 10^{5}$	011111-01111·1

Table J.1: Demonstration of shadowing in curvature combinations of cycle weights of form $t_{ab} - t_a t_b$, the 3-disk fundamental domain cycles at $R : d = 6$, table 14.3. The ratio $\Lambda_a \Lambda_b / \Lambda_{ab}$ is approaching unity exponentially fast.

Expanding the exponentials one thus finds that this term in the cycle expansion is of the order of

$$t_{a^j b} - t_a t_{a^{j-1} b} \approx \text{Const} \times t_{a^j b} \Lambda_a^{-j} \,. \tag{J.1}$$

Even though the number of terms in a cycle expansion grows exponentially, the shadowing cancellations improve the convergence by an exponential factor compared to trace formulas, and extend the radius of convergence of the periodic orbit sums. Table J.1 shows some examples of such compensations between long cycles and their pseudo-cycle shadows.

It is crucial that the curvature expansion is grouped (and truncated) by topologically related cycles and pseudo-cycles; truncations that ignore topology, such as inclusion of all cycles with $T_p < T_{max}$, will contain orbits unmatched by shadowed orbits, and exhibit a mediocre convergence compared with the curvature expansions.

Note that the existence of a pole at $z = 1/c$ implies that the cycle expansions have a finite radius of convergence, and that analytic continuations will be required for extraction of the non-leading zeros of $1/\zeta$. Preferably, one should work with cycle expansions of Selberg products, as discussed in sect. 15.1.3.

J.1.3 No shadowing, poorer convergence

Conversely, if the dynamics is not of a finite subshift type, there is no finite topological polynomial, there are no "curvature" corrections, and the convergence of the cycle expansions will be poor.

J.2 On importance of pruning

If the grammar is not finite and there is no finite topological polynomial, there will be no "curvature" expansions, and the convergence will be poor. That is the generic case, and one strategy for dealing with it is to find a good sequence of approximate but finite grammars; for each approximate grammar cycle expansions yield exponentially accurate eigenvalues, with successive approximate grammars converging toward the desired infinite grammar system.

When the dynamical system's symbolic dynamics does not have a finite grammar, and we are not able to arrange its cycle expansion into curvature combinations (15.5), the series is truncated as in sect. 15.4, by including all pseudo-cycles such that $|\Lambda_{p_1} \cdots \Lambda_{p_k}| \leq |\Lambda_P|$, where P is the most unstable prime cycle included into truncation. The truncation error should then be of order $O(e^{hT_P} T_P / |\Lambda_P|)$, with h the topological entropy, and e^{hT_P} roughly the number of pseudo-cycles of stability $\approx |\Lambda_P|$. In this case the cycle averaging formulas do not converge significantly better than the approximations such as the trace formula (16.18).

Numerical results (see for example the plots of the accuracy of the cycle expansion truncations for the Hénon map in ref. [2]) indicate that the truncation error of most averages tracks closely the fluctuations due to the irregular growth in the number of cycles. It is not known whether one can exploit the sum rules such as the mass flow conservation (16.11) to improve the accuracy of dynamical averaging.

J.3 Ma-the-matical caveats

> "Lo duca e io per quel cammino ascoso intrammo a ritornar nel chiaro monde; e sanza cura aver d'alcun riposa salimmo sù, el primo e io secondo, tanto ch'i' vidi de le cose belle che porta 'l ciel, per un perutgio tondo."
>
> Dante

The periodic orbit theory is learned in stages. At first glance, it seems totally impenetrable. After basic exercises are gone through, it seems totally trivial; all that seems to be at stake are elementary manipulations with traces, determinants, derivatives. But if start thinking about you will get a more and more uncomfortable feeling that from the mathematical point of view, this is a perilous enterprise indeed. In chapter 13 we shall explain which parts of this enterprise are really solid; here you give a fortaste of what objections a mathematician might rise.

Birkhoff's 1931 ergodic theorem states that the time average (8.4) exists almost everywhere, and, if the flow is ergodic, it implies that $\langle a(x)\rangle = \langle a\rangle$ is a constant for almost all x. The problem is that the above cycle averaging formulas implicitly rely on ergodic hypothesis: they are strictly correct only if the dynamical system is locally hyperbolic and globally mixing. If one takes a β derivative of both sides of (**??**)

$$\rho_\beta(y)e^{ts(\beta)} = \int_{\mathcal{M}} dx\, \delta(y - f^t(x))e^{\beta\cdot A^t(x)}\rho_\beta(x)\,,$$

and integrates over y

$$\int_{\mathcal{M}} dy\ \frac{\partial}{\partial\beta}\rho_\beta(y)\Big|_{\beta=0} \ +\ t\,\frac{\partial s}{\partial\beta}\Big|_{\beta=0}\int_{\mathcal{M}} dy\,\rho_0(y) \ = \int_{\mathcal{M}} dx\ A^t(x)\rho_0(x) + \int_{\mathcal{M}} dx\ \frac{\partial}{\partial\beta}\rho_\beta(x)\Big|_{\beta=0}\,,$$

one obtains in the long time limit

$$\frac{\partial s}{\partial\beta}\Big|_{\beta=0} = \int_{\mathcal{M}} dy\,\rho_0(x)\,\langle a(x)\rangle\ . \tag{J.2}$$

This is the expectation value (8.12) only if the time average (8.4) equals the space average (8.9), $\langle a(x)\rangle = \langle a\rangle$, for all x except a subset $x \in \mathcal{M}$ of zero measure; if the phase space is foliated into non-communicating subspaces $\mathcal{M} = \mathcal{M}_1 + \mathcal{M}_2$ of finite measure such that $f^t(\mathcal{M}_1) \cap \mathcal{M}_2 = \emptyset$ for all t, this fails. In other words, we have tacitly assumed metric indecomposability or transitivity. We have also glossed over the nature of the "phase space" $\mathcal{M}$. For example, if the dynamical system is open, such as the 3-disk game of pinball, $\mathcal{M}$ in the expectation value integral (8.18) is a Cantor set, the closure of the union of all periodic orbits.

Alternatively, $\mathcal{M}$ can be considered continuous, but then the measure ρ_0 in (J.2) is highly singular. The beauty of the periodic orbit theory is that instead of using an arbitrary coordinatization of $\mathcal{M}$ it partitions the phase space by the intrinsic topology of the dynamical flow and builds the correct measure from cycle invariants, the stability eigenvalues of periodic orbits.

Were we to restrict the applications of the formalism only to systems which have been rigorously proven to be ergodic, we might as well fold up the shop right now. For example, even for something as simple as the Hénon mapping we do not know whether the asymptotic time attractor is strange or periodic. Physics applications require a more pragmatic attitude. In the cycle expansions approach we construct the invariant set of the given dynamical system as a closure of the

4.2 on p. 80

union of periodic orbits, and investigate how robust are the averages computed on this set. This turns out to depend very much on the observable being averaged over; dynamical averages exhibit "phase transitions"(to be discussed in sect. **??**), and the above cycle averaging formulas apply in the "hyperbolic phase" where the average is dominated by exponentially many exponentially small contributions, but fail in a phase dominated by few marginally stable orbits. Here the noise - always present, no matter how weak - helps us by erasing an infinity of small traps that the deterministic dynamics might fall into.

Still, in spite of all the caveats, periodic orbit theory is a beautiful theory, and the cycle averaging formulas are the most elegant and powerful tool available today for evaluation of dynamical averages for low dimensional chaotic deterministic systems.

J.4 Estimate of the nth cumulant

An immediate consequence of the exponential spacing of the eigenvalues is that the convergence of the Selberg product expansion (E.12) as function of the topological cycle length, $F(z) = \sum_n C_n z^n$, is faster than exponential. Consider a d–dimensional map for which all Jacobian matrix eigenvalues in (**??**) are equal: $u_p = \Lambda_{p,1} = \Lambda_{p,2} = \cdots = \Lambda_{p,d}$. The stability eigenvalues are generally not isotropic; however, to obtain qualitative bounds on the spectrum, we replace all stability eigenvalues with the least expanding one. In this case the p cycle contribution to the product (12.9) reduces to

$$\begin{aligned} F_p(z) &= \prod_{k_1\cdots k_d=0}^{\infty}\left(1 - t_p u_p^{k_1+k_2+\cdots+k_d}\right) \\ &= \prod_{k=0}^{\infty}\left(1 - t_p u_p^k\right)^{m_k}\,; \qquad m_k = \binom{d-1+k}{d-1} = \frac{(k+d-1)!}{k!(d-1)!} \\ &= \prod_{k=0}^{\infty}\sum_{\ell=0}^{m_k}\binom{m_k}{\ell}\left(-u_p^k t_p\right)^\ell \end{aligned} \tag{J.3}$$

In one dimension the expansion can be given in closed form (13.26), and the coefficients C_k in (E.12) are given by

$$\tau_{p^k} = (-1)^k\frac{u_p^{\frac{k(k-1)}{2}}}{\prod_{j=1}^k(1-u_p^j)}t_p^k\ . \tag{J.4}$$

Hence the coefficients in the $F(z) = \sum_n C_n z^n$ expansion of the spectral determinant (15.8) fall off faster than exponentially, as $|C_n| \approx u^{n(n-1)/2}$. In contrast,

the cycle expansions of dynamical zeta functions fall of "only" exponentially; in numerical applications, the difference is dramatic.

In higher dimensions the expansions are not quite as compact. The leading power of u and its coefficient are easily evaluated by use of binomial expansions (J.3) of the $(1+tu^k)^{m_k}$ factors. More precisely, the leading u^n terms in t^k coefficients are of form

$$\begin{aligned}\prod_{k=0}^{\infty}(1+tu^k)^{m_k} &= \ldots + u^{m_1+2m_2+\ldots+jm_j}\, t^{1+m_1+m_2+\ldots+m_j} + \ldots \\ &= \ldots + \left(u^{\frac{m_d}{d+1}}t\right)^{\binom{d+m}{m}} + \ldots \approx \ldots + u^{\frac{\sqrt[d]{d!}}{(d-1)!}n^{\frac{d+1}{d}}}t^n + \ldots\end{aligned}$$

Hence the coefficients in the $F(z)$ expansion fall off faster than exponentially, as $u^{n^{1+1/d}}$. The Selberg products are entire functions in any dimension, provided that the symbolic dynamics is a finite subshift, and all cycle eigenvalues are sufficiently bounded away from 1.

The case of particular interest in many applications are the 2-d Hamiltonian mappings; their symplectic structure implies that $u_p = \Lambda_{p,1} = 1/\Lambda_{p,2}$, and the Selberg product (12.23) In this case the expansion corresponding to (13.26) is given by (13.27) and the coefficients fall off asymptotically as $C_n \approx u^{n^{3/2}}$.

Appendix K

Infinite dimensional operators

(A. Wirzba)

This appendix taken from ref. [1] summarizes the definitions and properties for trace-class and Hilbert-Schmidt matrices, the determinants over infinite dimensional matrices and possible regularization schemes for matrices or operators which are not of trace-class.

K.1 Matrix-valued functions

(P. Cvitanović)

As a preliminary we summarize some of the properties of functions of finite-dimensional matrices.

The derivative of a matrix is a matrix with elements

$$\mathbf{A}'(x) = \frac{d\mathbf{A}(x)}{dx}\,, \qquad A'_{ij}(x) = \frac{d}{dx}A_{ij}(x)\,. \tag{K.1}$$

Derivatives of products of matrices are evaluated by the chain rule

$$\frac{d}{dx}(\mathbf{AB}) = \frac{d\mathbf{A}}{dx}\mathbf{B} + \mathbf{A}\frac{d\mathbf{B}}{dx}\,. \tag{K.2}$$

A matrix and its derivative matrix in general do not commute

$$\frac{d}{dx}\mathbf{A}^2 = \frac{d\mathbf{A}}{dx}\mathbf{A} + \mathbf{A}\frac{d\mathbf{A}}{dx}\,. \tag{K.3}$$

The derivative of the inverse of a matrix, follows from $\frac{d}{dx}(\mathbf{A}\mathbf{A}^{-1}) = 0$:

$$\frac{d}{dx}\mathbf{A}^{-1} = -\frac{1}{\mathbf{A}}\frac{d\mathbf{A}}{dx}\frac{1}{\mathbf{A}}\,. \tag{K.4}$$

A function of a single variable that can be expressed in terms of additions and multiplications generalizes to a matrix-valued function by replacing the variable by the matrix.

In particular, the exponential of a constant matrix can be defined either by its series expansion, or as a limit of an infinite product:

$$e^{\mathbf{A}} = \sum_{k=0}^{\infty}\frac{1}{k!}\mathbf{A}^k\,, \qquad \mathbf{A}^0 = \mathbf{1} \tag{K.5}$$

$$= \lim_{N\to\infty}\left(\mathbf{1}+\frac{1}{N}\mathbf{A}\right)^N \tag{K.6}$$

The first equation follows from the second one by the binomial theorem, so these indeed are equivalent definitions. That the terms of order $O(N^{-2})$ or smaller do not matter follows from the bound

$$\left(1+\frac{x-\epsilon}{N}\right)^N < \left(1+\frac{x+\delta x_N}{N}\right)^N < \left(1+\frac{x+\epsilon}{N}\right)^N\,,$$

where $|\delta x_N| < \epsilon$. If $\lim \delta x_N \to 0$ as $N\to\infty$, the extra terms do not contribute.

Consider now the determinant

$$\det\left(e^{\mathbf{A}}\right) = \lim_{N\to\infty}\left(\det\left(\mathbf{1}+\mathbf{A}/N\right)\right)^N\,.$$

To the leading order in $1/N$

$$\det\left(\mathbf{1}+\mathbf{A}/N\right) = 1+\frac{1}{N}\mathrm{tr}\,\mathbf{A}+O(N^{-2})\,.$$

hence

$$\det e^{\mathbf{A}} = \lim_{N\to\infty}\left(1+\frac{1}{N}\mathrm{tr}\,\mathbf{A}+O(N^{-2})\right)^N = e^{\mathrm{tr}\,\mathbf{A}} \tag{K.7}$$

Due to non-commutativity of matrices, generalization of a function of several variables to a function is not as straightforward. Expression involving several matrices depend on their commutation relations. For example, the commutator expansion

$$e^{t\mathbf{A}}\mathbf{B}e^{-t\mathbf{A}} = \mathbf{B}+t[\mathbf{A},\mathbf{B}]+\frac{t^2}{2}[\mathbf{A},[\mathbf{A},\mathbf{B}]]+\frac{t^3}{3!}[\mathbf{A},[\mathbf{A},[\mathbf{A},\mathbf{B}]]]+\cdots \tag{K.8}$$

sometimes used to establish the equivalence of the Heisenberg and Schrödinger pictures of quantum mechanics follows by recursive evaluation of t derivaties

$$\frac{d}{dt}\left(e^{t\mathbf{A}}\mathbf{B}e^{-t\mathbf{A}}\right) = e^{t\mathbf{A}}[\mathbf{A},\mathbf{B}]e^{-t\mathbf{A}}\,.$$

Manipulations of such ilk yield

$$e^{(\mathbf{A}+\mathbf{B})/N} = e^{\mathbf{A}/N}e^{\mathbf{B}/N}-\frac{1}{2N^2}[\mathbf{A},\mathbf{B}]+O(N^{-3})\,,$$

and the Trotter product formula: if $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{A}=\mathbf{B}+\mathbf{C}$ are matrices, then

$$e^{\mathbf{A}} = \lim_{N\to\infty}\left(e^{\mathbf{B}/N}e^{\mathbf{C}/N}\right)^N \tag{K.9}$$

K.2 Operator norms

(R. Mainieri and P. Cvitanović)

The limit used in the above definition involves matrices - operators in vector spaces - rather than numbers, and its convergence can be checked using tools familiar from calculus. We briefly review those tools here, as throughout the text we will have to consider many different operators and how they converge.

The $n\to\infty$ convergence of partial products

$$\mathbf{E}_n = \prod_{0\le m<n}\left(\mathbf{1}+\frac{t}{m}\mathbf{A}\right)$$

can be verified using the Cauchy criterion, which states that the sequence $\{\mathbf{E}_n\}$ converges if the differences $\|\mathbf{E}_k-\mathbf{E}_j\|\to 0$ as $k,j\to\infty$. To make sense of this we

need to define a sensible norm $\|\cdots\|$. Norm of a matrix is based on the Euclidean norm for a vector: the idea is to assign to a matrix $\mathbf{M}$ a norm that is the largest possible change it can cause to the length of a unit vector $\hat{n}$:

$$\|\mathbf{M}\| = \sup_{\hat{n}} \|\mathbf{M}\hat{n}\|\,, \qquad \|\hat{n}\| = 1\,. \tag{K.10}$$

We say that $\|\cdot\|$ is the operator norm induced by the vector norm $\|\cdot\|$. Constructing a norm for a finite-dimensional matrix is easy, but had $\mathbf{M}$ been an operator in an infinite-dimensional space, we would also have to specify the space $\hat{n}$ belongs to. In the finite-dimensional case, the sum of the absolute values of the components of a vector is also a norm; the induced operator norm for a matrix $\mathbf{M}$ with components M_{ij} in that case can be defined by

$$\|\mathbf{M}\| = \max_i \sum_j |M_{ij}|\,. \tag{K.11}$$

The operator norm (K.11) and the vector norm (K.10) are only rarely distinguished by different notation, a bit of notational laziness that we shall uphold.

Now that we have learned how to make sense out of norms of operators, we can check that

$$\|e^{t\mathbf{A}}\| \le e^{t\|\mathbf{A}\|}\,. \tag{K.12}$$

2.8 on p. 49

As $\|\mathbf{A}\|$ is a number, the norm of $e^{t\mathbf{A}}$ is finite and therefore well defined. In particular, the exponential of a matrix is well defined for all values of t, and the linear differential equation (4.7) has a solution for all times.

K.3 Trace class and Hilbert-Schmidt class

This section is mainly an extract from ref. [9]. Refs. [7, 10, 11, 14] should be consulted for more details and proofs. The trace class and Hilbert-Schmidt property will be defined here for linear, in general non-hermitian operators $\mathbf{A} \in \mathcal{L}(\mathcal{H})$: $\mathcal{H} \to \mathcal{H}$ (where $\mathcal{H}$ is a separable Hilbert space). The transcription to matrix elements (used in the prior chapters) is simply $a_{ij} = \langle\phi_i, \mathbf{A}\phi_j\rangle$ where $\{\phi_n\}$ is an orthonormal basis of $\mathcal{H}$ and $\langle\ ,\ \rangle$ is the inner product in $\mathcal{H}$ (see sect. K.5 where the theory of *von Koch matrices* of ref. [12] is discussed). So, the trace is the generalization of the usual notion of the sum of the diagonal elements of a matrix; but because infinite sums are involved, not all operators will have a trace:

Definition:

(a) An operator $\mathbf{A}$ is called **trace class**, $\mathbf{A} \in \mathcal{J}_1$, if and only if, for every orthonormal basis, $\{\phi_n\}$:

$$\sum_n |\langle\phi_n, \mathbf{A}\phi_n\rangle| < \infty. \tag{K.13}$$

The family of all trace class operators is denoted by $\mathcal{J}_1$.

(b) An operator $\mathbf{A}$ is called **Hilbert-Schmidt**, $\mathbf{A} \in \mathcal{J}_2$, if and only if, for every orthonormal basis, $\{\phi_n\}$:

$$\sum_n \|\mathbf{A}\phi_n\|^2 < \infty. \tag{K.14}$$

The family of all Hilbert-Schmidt operators is denoted by $\mathcal{J}_2$.

Bounded operators are dual to trace class operators. They satisfy the the following condition: $|\langle\psi, B\phi\rangle| \le C\|\psi\|\|\phi\|$ with $C < \infty$ and $\psi, \phi \in \mathcal{H}$. If they have eigenvalues, these are bounded too. The family of bounded operators is denoted by $\mathcal{B}(\mathcal{H})$ with the norm $\|B\| = \sup_{\phi\neq 0} \frac{\|\mathbf{B}\phi\|}{\|\phi\|}$ for $\phi \in \mathcal{H}$. Examples for bounded operators are unitary operators and especially the unit matrix. In fact, every bounded operator can be written as linear combination of four unitary operators.

A bounded operator $\mathbf{C}$ is *compact*, if it is the norm limit of finite rank operators.

An operator $\mathbf{A}$ is called *positive*, $\mathbf{A} \ge 0$, if $\langle\mathbf{A}\phi, \phi\rangle \ge 0\ \ \forall\phi \in \mathcal{H}$. Notice that $\mathbf{A}^\dagger\mathbf{A} \ge 0$. We define $|\mathbf{A}| = \sqrt{\mathbf{A}^\dagger\mathbf{A}}$.

The most important properties of the trace and Hilbert-Schmidt classes are summarized in (see refs. [7, 9]):

(a) $\mathcal{J}_1$ and $\mathcal{J}_2$ are $*$ideals., i.e., they are vector spaces closed under scalar multiplication, sums, adjoints, and multiplication with bounded operators.

(b) $\mathbf{A} \in \mathcal{J}_1$ if and only if $\mathbf{A} = \mathbf{BC}$ with $\mathbf{B}, \mathbf{C} \in \mathcal{J}_2$.

(c) $\mathcal{J}_1 \subset \mathcal{J}_2 \subset$ Compact operators.

(d) For any operator $\mathbf{A}$, we have $\mathbf{A} \in \mathcal{J}_2$ if $\sum_n \|\mathbf{A}\phi_n\|^2 < \infty$ for a single basis. For any operator $\mathbf{A} \ge 0$ we have $\mathbf{A} \in \mathcal{J}_1$ if $\sum_n |\langle\phi_n, \mathbf{A}\phi_n\rangle| < \infty$ for a single basis.

(e) If $\mathbf{A} \in \mathcal{J}_1$, $\mathrm{Tr}(\mathbf{A}) = \sum\langle\phi_n, \mathbf{A}\phi_n\rangle$ is independent of the basis used.

(f) Tr is linear and obeys $\mathrm{Tr}(\mathbf{A}^\dagger) = \overline{\mathrm{Tr}(\mathbf{A})}$; $\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$ if either $\mathbf{A} \in \mathcal{J}_1$ and $\mathbf{B}$ bounded, $\mathbf{A}$ bounded and $\mathbf{B} \in \mathcal{J}_1$ or both $\mathbf{A}, \mathbf{B} \in \mathcal{J}_2$.

(g) $\mathcal{J}_2$ endowed with the inner product $\langle \mathbf{A}, \mathbf{B}\rangle_2 = \mathrm{Tr}(\mathbf{A}^\dagger \mathbf{B})$ is a Hilbert space. If $\|\mathbf{A}\|_2 = [\,\mathrm{Tr}(\mathbf{A}^\dagger \mathbf{A})\,]^{\frac{1}{2}}$, then $\|\mathbf{A}\|_2 \geq \|\mathbf{A}\|$ and $\mathcal{J}_2$ is the $\|\ \|_2$-closure of the *finite* rank operators.

(h) $\mathcal{J}_1$ endowed with the norm $\|\mathbf{A}\|_1 = \mathrm{Tr}(\sqrt{\mathbf{A}^\dagger \mathbf{A}})$ is a Banach space. $\|\mathbf{A}\|_1 \geq \|\mathbf{A}\|_2 \geq \|\mathbf{A}\|$ and $\mathcal{J}_1$ is the $\|\ \|_1$-norm closure of the *finite* rank operators. The dual space of $\mathcal{J}_1$ is $\mathcal{B}(\mathcal{H})$, the family of bounded operators with the duality $\langle \mathbf{B}, \mathbf{A}\rangle = \mathrm{Tr}(\mathbf{B}\mathbf{A})$.

(i) If $\mathbf{A}, \mathbf{B} \in \mathcal{J}_2$, then $\|\mathbf{A}\mathbf{B}\|_1 \leq \|\mathbf{A}\|_2 \|\mathbf{B}\|_2$. If $\mathbf{A} \in \mathcal{J}_2$ and $\mathbf{B} \in \mathcal{B}(\mathcal{H})$, then $\|\mathbf{A}\mathbf{B}\|_2 \leq \|\mathbf{A}\|_2 \|\mathbf{B}\|$. If $\mathbf{A} \in \mathcal{J}_1$ and $\mathbf{B} \in \mathcal{B}(\mathcal{H})$, then $\|\mathbf{A}\mathbf{B}\|_1 \leq \|\mathbf{A}\|_1 \|\mathbf{B}\|$.

Note the most important property for proving that an operator is trace class is the decomposition **(b)** into two Hilbert-Schmidt ones, as the Hilbert-Schmidt property can easily be verified in one single orthonormal basis (see **(d)**). Property **(e)** ensures then that the trace is the same in any basis. Properties **(a)** and **(f)** show that trace class operators behave in complete analogy to finite rank operators. The proof whether a matrix is trace-class (or Hilbert-Schmidt) or not simplifies enormously for diagonal matrices, as then the second part of property **(d)** is directly applicable: just the moduli of the eigenvalues (or – in case of Hilbert-Schmidt – the squares of the eigenvalues) have to be summed up in order to answer that question. A good strategy in checking the trace-class character of a general matrix $\mathbf{A}$ is therefore the decomposition of that matrix into two matrices $\mathbf{B}$ and $\mathbf{C}$ where one, say $\mathbf{C}$, should be chosen to be diagonal and either just barely of Hilbert-Schmidt character leaving enough freedom for its partner $\mathbf{B}$ or of trace-class character such that one only has to show the boundedness for $\mathbf{B}$.

K.4 Determinants of trace class operators

This section is mainly based on refs. [8, 10] which should be consulted for more details and proofs. See also refs. [11, 14].

Pre-definitions (Alternating algebra and Fock spaces):
Given a Hilbert space $\mathcal{H}$, $\otimes^n \mathcal{H}$ is defined as the vector space of multi-linear functionals on $\mathcal{H}$ with $\phi_1 \otimes \cdots \otimes \phi_n \in \otimes^n \mathcal{H}$ in case $\phi_1, \ldots, \phi_n \in \mathcal{H}$. $\bigwedge^n(\mathcal{H})$ is defined as the subspace of $\otimes^n \mathcal{H}$ spanned by the wedge-product

$$\phi_1 \wedge \cdots \wedge \phi_n = \frac{1}{\sqrt{n!}} \sum_{\pi \in \mathcal{P}_n} \epsilon(\pi) [\phi_{\pi(1)} \otimes \cdots \otimes \phi_{\pi(n)}] \tag{K.15}$$

where $\mathcal{P}_n$ is the group of all permutations of n letters and $\epsilon(\pi) = \pm 1$ depending on whether π is an even or odd permutation, respectively. The inner product in

$\bigwedge^n(\mathcal{H})$ is given by

$$(\phi_1 \wedge \cdots \wedge \phi_n, \eta_1 \wedge \cdots \wedge \eta_n) = \det\{(\phi_i, \eta_j)\} \tag{K.16}$$

where $\det\{a_{ij}\} = \sum_{\pi \in \mathcal{P}_n} \epsilon(\pi) a_{1\pi(1)} \cdots a_{n\pi(n)}$. $\bigwedge^n(\mathbf{A})$ is defined as functor (a functor satisfies $\bigwedge^n(\mathbf{A}\mathbf{B}) = \bigwedge^n(\mathbf{A}) \bigwedge^n(\mathbf{B})$) on $\bigwedge^n(\mathcal{H})$ with

$$\bigwedge^n(\mathbf{A})\,(\phi_1 \wedge \cdots \wedge \phi_n) = \mathbf{A}\phi_1 \wedge \cdots \wedge \mathbf{A}\phi_n\ . \tag{K.17}$$

When $n = 0$, $\bigwedge^n(\mathcal{H})$ is defined to be C and $\bigwedge^n(\mathbf{A})$ as $1\colon C \to C$.

Properties: If $\mathbf{A}$ trace class, i.e., $\mathbf{A} \in \mathcal{J}_1$, then for any k, $\bigwedge^k(\mathbf{A})$ is trace class, and for any orthonormal basis $\{\phi_n\}$ the cumulant

$$\mathrm{Tr}\left(\bigwedge^k(\mathbf{A})\right) = \sum_{i_1 < \cdots < i_k} ((\phi_{i_1} \wedge \cdots \wedge \phi_{i_k}), (\mathbf{A}\phi_{i_1} \wedge \cdots \wedge \mathbf{A}\phi_{i_k})) < \infty \tag{K.18}$$

is independent of the basis (with the understanding that $\mathrm{Tr}\bigwedge^0(\mathbf{A}) \equiv 1$).

Definition: Let $\mathbf{A} \in \mathcal{J}_1$, then $\mathbf{det(1+A)}$ is defined as

$$\det(\mathbf{1} + \mathbf{A}) = \sum_{k=0}^{\infty} \mathrm{Tr}\left(\bigwedge^k(\mathbf{A})\right) \tag{K.19}$$

Properties:

Let $\mathbf{A}$ be a linear operator on a separable Hilbert space $\mathcal{H}$ and $\{\phi_j\}_1^\infty$ an orthonormal basis.

(a) $\sum_{k=0}^{\infty} \mathrm{Tr}\left(\bigwedge^k(\mathbf{A})\right)$ converges for each $\mathbf{A} \in \mathcal{J}_1$.

(b) $|\det(\mathbf{1} + \mathbf{A})| \leq \prod_{j=1}^{\infty} (1 + \mu_j(\mathbf{A}))$ where $\mu_j(\mathbf{A})$ are the *singular* values of $\mathbf{A}$, i.e., the eigenvalues of $|\mathbf{A}| = \sqrt{\mathbf{A}^\dagger \mathbf{A}}$.

(c) $|\det(\mathbf{1} + \mathbf{A})| \leq \exp(\|\mathbf{A}\|_1)$.

(d) For any $\mathbf{A}_1, \ldots, \mathbf{A}_n \in \mathcal{J}_1$, $\langle z_1, \ldots, z_n\rangle \mapsto \det\left(\mathbf{1} + \sum_{i=1}^n z_i \mathbf{A}_i\right)$ is an entire analytic function.

(e) If $\mathbf{A}, \mathbf{B} \in \mathcal{J}_1$, then

$$\begin{aligned} \det(\mathbf{1}+\mathbf{A})\det(\mathbf{1}+\mathbf{B}) &= \det\left(\mathbf{1}+\mathbf{A}+\mathbf{B}+\mathbf{AB}\right) \\ &= \det\left((\mathbf{1}+\mathbf{A})(\mathbf{1}+\mathbf{B})\right) \\ &= \det\left((\mathbf{1}+\mathbf{B})(\mathbf{1}+\mathbf{A})\right) \ . \end{aligned} \tag{K.20}$$

If $\mathbf{A} \in \mathcal{J}_1$ and $\mathbf{U}$ unitary, then

$$\det\left(\mathbf{U}^{-1}(\mathbf{1}+\mathbf{A})\mathbf{U}\right) = \det\left(\mathbf{1}+\mathbf{U}^{-1}\mathbf{A}\mathbf{U}\right) = \det(\mathbf{1}+\mathbf{A}) \ . \tag{K.21}$$

(f) If $\mathbf{A} \in \mathcal{J}_1$, then $(\mathbf{1}+\mathbf{A})$ is invertible if and only if $\det(\mathbf{1}+\mathbf{A}) \neq 0$.

(g) If $\lambda \neq 0$ is an n-times degenerate eigenvalue of $\mathbf{A} \in \mathcal{J}_1$, then $\det(\mathbf{1}+z\mathbf{A})$ has a zero of order n at $z = -1/\lambda$.

(h) For any ϵ, there is a $C_\epsilon(\mathbf{A})$, depending on $\mathbf{A} \in \mathcal{J}_1$, so that $|\det(\mathbf{1}+z\mathbf{A})| \leq C_\epsilon(\mathbf{A})\exp(\epsilon|z|)$.

(i) For any $\mathbf{A} \in \mathcal{J}_1$,

$$\det(\mathbf{1}+\mathbf{A}) = \prod_{j=1}^{N(\mathbf{A})} (1+\lambda_j(\mathbf{A})) \tag{K.22}$$

where here and in the following $\{\lambda_j(\mathbf{A})\}_{j=1}^{N(\mathbf{A})}$ are the eigenvalues of $\mathbf{A}$ counted with algebraic multiplicity .

(j) *Lidskii's theorem:* For any $\mathbf{A} \in \mathcal{J}_1$,

$$\mathrm{Tr}(\mathbf{A}) = \sum_{j=1}^{N(\mathbf{A})} \lambda_j(\mathbf{A}) < \infty \ . \tag{K.23}$$

(k) If $\mathbf{A} \in \mathcal{J}_1$, then

$$\begin{aligned} \mathrm{Tr}\left(\bigwedge^k(\mathbf{A})\right) &= \sum_{j=1}^{N\left(\bigwedge^k(\mathbf{A})\right)} \lambda_j\left(\bigwedge^k(\mathbf{A})\right) \\ &= \sum_{1\leq j_1<\cdots<j_k\leq N(\mathbf{A})} \lambda_{j_1}(\mathbf{A})\cdots\lambda_{j_k}(\mathbf{A}) < \infty. \end{aligned}$$

(l) If $\mathbf{A} \in \mathcal{J}_1$, then

$$\det(1+z\mathbf{A}) = \sum_{k=0}^{\infty} z^k \sum_{1\leq j_1<\cdots<j_k\leq N(\mathbf{A})} \lambda_{j_1}(\mathbf{A})\cdots\lambda_{j_k}(\mathbf{A}) < \infty. \tag{K.24}$$

(m) If $\mathbf{A} \in \mathcal{J}_1$, then for $|z|$ small (that is $|z|\max|\lambda_j(\mathbf{A})| < 1$) the series $\sum_{k=1}^{\infty} z^k \mathrm{Tr}\left((-\mathbf{A})^k\right)/k$ converges and

$$\begin{aligned} \det(1+z\mathbf{A}) &= \exp\left(-\sum_{k=1}^{\infty}\frac{z^k}{k}\mathrm{Tr}\left((-\mathbf{A})^k\right)\right) \\ &= \exp\left(\mathrm{Tr}\ln(\mathbf{1}+z\mathbf{A})\right) \ . \end{aligned} \tag{K.25}$$

(n) *The Plemelj-Smithies formula:* Define $\alpha_m(\mathbf{A})$ for $\mathbf{A} \in \mathcal{J}_1$ by

$$\det(\mathbf{1}+z\mathbf{A}) = \sum_{m=0}^{\infty} z^m \frac{\alpha_m(\mathbf{A})}{m!} \ . \tag{K.26}$$

Then $\alpha_m(\mathbf{A})$ is given by the $m \times m$ determinant:

$$\alpha_m(\mathbf{A}) = \begin{vmatrix} \mathrm{Tr}(\mathbf{A}) & m-1 & 0 & \cdots & 0 \\ \mathrm{Tr}(\mathbf{A}^2) & \mathrm{Tr}(\mathbf{A}) & m-2 & \cdots & 0 \\ \mathrm{Tr}(\mathbf{A}^3) & \mathrm{Tr}(\mathbf{A}^2) & \mathrm{Tr}(\mathbf{A}) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ & & & & 1 \\ \mathrm{Tr}(\mathbf{A}^m) & \mathrm{Tr}(\mathbf{A}^{(m-1)}) & \mathrm{Tr}(\mathbf{A}^{(m-2)}) & \cdots & \mathrm{Tr}(\mathbf{A}) \end{vmatrix} \tag{K.27}$$

with the understanding that $\alpha_0(\mathbf{A}) \equiv 1$ and $\alpha_1(\mathbf{A}) \equiv \mathrm{Tr}(\mathbf{A})$. Thus the cumulants $c_m(\mathbf{A}) \equiv \alpha_m(\mathbf{A})/m!$ satisfy the following recursion relation

$$\begin{aligned} c_m(\mathbf{A}) &= \frac{1}{m}\sum_{k=1}^{m}(-1)^{k+1}c_{m-k}(\mathbf{A})\,\mathrm{Tr}(\mathbf{A}^k) \quad \text{for } m \geq 1 \\ c_0(\mathbf{A}) &\equiv 1 \ . \end{aligned} \tag{K.28}$$

Note that in the context of quantum mechanics formula (K.26) is the quantum analog to the curvature expansion of the semiclassical zeta function with $\mathrm{Tr}(\mathbf{A}^m)$ corresponding to the sum of all periodic orbits (prime and also repeated ones) of *total* topological length m, that is let c_m(s.c.) denote the m^{th} curvature term, then the curvature expansion of the semiclassical zeta function is given by the recursion relation

$$\begin{aligned} c_m(\text{s.c.}) &= \frac{1}{m}\sum_{k=1}^{m}(-1)^{k+m+1}c_{m-k}(\text{s.c.}) \sum_{\substack{p;r>0 \\ \text{with } [p]r=k}} [p]\,\frac{t_p(k)^r}{1-\left(\frac{1}{\Lambda_p}\right)^r} \quad \text{for } m \geq 1 \\ c_0(\text{s.c.}) &\equiv 1 \ . \end{aligned} \tag{K.29}$$

In fact, in the cumulant expansion (K.26) as well as in the curvature expansion there are large cancellations involved. Let us order – without lost of generality – the eigenvalues of the operator $\mathbf{A} \in \mathcal{J}_1$ as follows:

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{i-1}| \geq |\lambda_i| \geq |\lambda_{i+1}| \geq \cdots$$

(This is always possible because of $\sum_{i=1}^{N(\mathbf{A})} |\lambda_i| < \infty$.) Then, in the standard (Plemelj-Smithies) cumulant evaluation of the determinant, eq. (K.26), we have enormous cancellations of big numbers, e.g. at the $k^{\rm th}$ cumulant order ($k > 3$), all the intrinsically large 'numbers' λ_1^k, $\lambda_1^{k-1}\lambda_2$, ..., $\lambda_1^{k-2}\lambda_2\lambda_3$, ... and many more have to cancel out exactly until only $\sum_{1\le j_1<\cdots<j_k\le N(\mathbf{A})} \lambda_{j_1}\cdots\lambda_{j_k}$ is finally left over. Algebraically, the fact that there are these large cancellations is of course of no importance. However, if the determinant is calculated numerically, the big cancellations might spoil the result or even the convergence. Now, the curvature expansion of the semiclassical zeta function, as it is known today, *is* the semi-classical approximation to the curvature expansion (unfortunately) in the Plemelj-Smithies form. As the exact quantum mechanical result is approximated semi-classically, the errors introduced in the approximation might lead to big effects as they are done with respect to large quantities which eventually cancel out and not – as it would be of course better – with respect to the small surviving cumulants. Thus it would be very desirable to have a semi-classical analog to the reduced cumulant expansion (K.24) or even to (K.22) directly. It might not be possible to find a direct semi-classical analog for the individual eigenvalues λ_j. Thus the direct construction of the semi-classical equivalent to (K.22) is rather unlikely. However, in order to have a semi-classical "cumulant" summation without large cancellations – see (K.24) – it would be just sufficient to find the semi-classical analog of each complete cumulant (K.24) and not of the single eigenvalues. Whether this will eventually be possible is still an open question.

K.5 Von Koch matrices

Implicitly, many of the above properties are based on the theory of von Koch matrices [11, 12, 13]: An infinite matrix $\mathbf{1}-\mathbf{A} = \|\delta_{jk} - a_{jk}\|_1^\infty$, consisting of complex numbers, is called a matrix with an *absolutely convergent determinant*, if the series $\sum |a_{j_1k_1}a_{j_2k_2}\cdots a_{j_n,k_n}|$ converges, where the sum extends over all pairs of systems of indices $(j_1, j_2, \cdots, j_n)$ and $(k_1, k_2, \cdots, k_n)$ which differ from each other only by a permutation, and $j_1 < j_2 < \cdots j_n$ $(n = 1, 2, \cdots)$. Then the limit

$$\lim_{n\to\infty} \det\|\delta_{jk} - a_{jk}\|_1^n = \det(\mathbf{1}-\mathbf{A})$$

exists and is called the determinant of the matrix $\mathbf{1}-\mathbf{A}$. It can be represented in the form

$$\det(\mathbf{1}-\mathbf{A}) = 1 - \sum_{j=1}^{\infty} a_{jj} + \frac{1}{2!}\sum_{j,k=1}^{\infty} \begin{vmatrix} a_{jj} & a_{jk} \\ a_{kj} & a_{kk} \end{vmatrix} - \frac{1}{3!}\sum_{j,k,m=1}^{\infty} \begin{vmatrix} a_{jj} & a_{jk} & a_{jm} \\ a_{kj} & a_{kk} & a_{km} \\ a_{mj} & a_{mk} & a_{mm} \end{vmatrix} + \cdots ,$$

where the series on the r.h.s. will remain convergent even if the numbers a_{jk} $(j, k = 1, 2, \cdots)$ are replaced by their moduli and if all the terms obtained by expanding the determinants are taken with the plus sign. The matrix $\mathbf{1}-\mathbf{A}$ is called *von Koch matrix*, if both conditions

$$\sum_{j=1}^{\infty} |a_{jj}| \;<\; \infty \,, \tag{K.30}$$

$$\sum_{j,k=1}^{\infty} |a_{jk}|^2 \;<\; \infty \tag{K.31}$$

are fulfilled. Then the following holds (see ref. [11, 13]): (1) Every von Koch matrix has an absolutely convergent determinant. If the elements of a von Koch matrix are functions of some parameter μ ($a_{jk} = a_{jk}(\mu)$, $j, k = 1, 2, \cdots$) and both series in the defining condition converge uniformly in the domain of the parameter μ, then as $n \to \infty$ the determinant $\det\|\delta_{jk} - a_{jk}(\mu)\|_1^n$ tends to the determinant $\det(\mathbf{1} + \mathbf{A}(\mu))$ uniformly with respect to μ, over the domain of μ. (2) If the matrices $\mathbf{1}-\mathbf{A}$ and $\mathbf{1}-\mathbf{B}$ are von Koch matrices, then their product $\mathbf{1}-\mathbf{C} = (\mathbf{1}-\mathbf{A})(\mathbf{1}-\mathbf{B})$ is a von Koch matrix, and

$$\det(\mathbf{1}-\mathbf{C}) = \det(\mathbf{1}-\mathbf{A})\det(\mathbf{1}-\mathbf{B}) \,. \tag{K.32}$$

Note that every trace-class matrix $\mathbf{A} \in \mathcal{J}_1$ is also a von Koch matrix (and that any matrix satisfying condition (K.31) is Hilbert-Schmidt and vice versa). The inverse implication, however, is not true: von Koch matrices are not automatically trace-class. The caveat is that the definition of von Koch matrices is basis-dependent, whereas the trace-class property is basis-*independent*. As the traces involve infinite sums, the basis-independence is not at all trivial. An example for an infinite matrix which is von Koch, but not trace-class is the following:

$$\mathbf{A}_{ij} = \begin{cases} 2/j & \text{for} \quad i-j=-1 \quad \text{and} \quad j \text{ even} \,, \\ 2/i & \text{for} \quad i-j=+1 \quad \text{and} \quad i \text{ even} \,, \\ 0 & \text{else} \,, \end{cases}$$

i.e.,

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1/2 & 0 & 0 & \cdots \\ 0 & 0 & 1/2 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 1/3 & \ddots \\ 0 & 0 & 0 & 0 & 1/3 & 0 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} . \tag{K.33}$$

Obviously, condition (K.30) is fulfilled by definition. Secondly, condition (K.31) is satisfied as $\sum_{n=1}^{\infty} 2/n^2 < \infty$. However, the sum over the moduli of the eigenvalues is just twice the harmonic series $\sum_{n=1}^{\infty} 1/n$ which does not converge. The matrix (K.33) violates the trace-class definition (K.13), as in its eigenbasis the sum over the moduli of its diagonal elements is infinite. Thus the *absolute* convergence is traded for a *conditional* convergence, since the sum over the eigenvalues themselves can be arranged to still be zero, if the eigenvalues with the same modulus are summed first. Absolute convergence is of course essential, if sums have to be rearranged or exchanged. Thus, the trace-class property is indispensable for any controlled unitary transformation of an infinite determinant, as then there will be necessarily a change of basis and in general also a re-ordering of the corresponding traces. Therefore the claim that *a Hilbert-Schmidt operator with a vanishing trace is automatically trace-class* is false. In general, such an operator has to be regularized in addition (see next chapter).

K.6 Regularization

Many interesting operators are not of trace class (although they might be in some $\mathcal{J}_p$ with $p > 1$ - an operator A is in $\mathcal{J}_p$ iff $\mathrm{Tr}|A|^p < \infty$ in any orthonormal basis). In order to compute determinants of such operators, an extension of the cumulant expansion is needed which in fact corresponds to a regularization procedure [8, 10]:
E.g. let $\mathbf{A} \in \mathcal{J}_p$ with $p \leq n$. Define

$$R_n(z\mathbf{A}) = (\mathbf{1} + z\mathbf{A}) \exp\left(\sum_{k=1}^{n-1} \frac{(-z)^k}{k} \mathbf{A}^k\right) - \mathbf{1} \tag{K.34}$$

as the regulated version of the operator $z\mathbf{A}$. Then the regulated operator $R_n(z\mathbf{A})$ is trace class, i.e., $R_n(z\mathbf{A}) \in \mathcal{J}_1$. Define now $\det_n(\mathbf{1} + z\mathbf{A}) = \det(\mathbf{1} + R_n(z\mathbf{A}))$. Then the regulated determinant

$$\det{}_n(\mathbf{1} + z\mathbf{A}) = \prod_{j=1}^{N(z\mathbf{A})} \left[(1 + z\lambda_j(\mathbf{A})) \exp\left(\sum_{k=1}^{n-1} \frac{(-z\lambda_j(\mathbf{A}))^k}{k}\right)\right] < \infty \,. \tag{K.35}$$

exists and is finite. The corresponding Plemelj-Smithies formula now reads [10]:

$$\det{}_n(\mathbf{1} + z\mathbf{A}) = \sum_{m=0}^{\infty} z^m \frac{\alpha_m^{(n)}(\mathbf{A})}{m!} \,. \tag{K.36}$$

with $\alpha_m^{(n)}(\mathbf{A})$ given by the $m \times m$ determinant:

$$\alpha_m^{(n)}(\mathbf{A}) = \begin{vmatrix} \sigma_1^{(n)} & m-1 & 0 & \cdots & 0 \\ \sigma_2^{(n)} & \sigma_1^{(n)} & m-2 & \cdots & 0 \\ \sigma_3^{(n)} & \sigma_2^{(n)} & \sigma_1^{(n)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ & & & & 1 \\ \sigma_m^{(n)} & \sigma_{m-1}^{(n)} & \sigma_{m-2}^{(n)} & \cdots & \sigma_1^{(n)} \end{vmatrix} \tag{K.37}$$

where

$$\sigma_k^{(n)} = \begin{cases} \mathrm{Tr}(\mathbf{A}^k) & k \geq n \\ 0 & k \leq n-1 \end{cases}$$

As Simon [10] says simply, the beauty of (K.37) is that we get $\det_n(\mathbf{1} + \mathbf{A})$ from the standard Plemelj-Smithies formula (K.26) by simply setting $\mathrm{Tr}(\mathbf{A})$, $\mathrm{Tr}(\mathbf{A}^2)$, ..., $\mathrm{Tr}(\mathbf{A}^{n-1})$ to zero.

See also ref. [15] where $\{\lambda_j\}$ are the eigenvalues of an elliptic (pseudo)-differential operator $\mathbf{H}$ of order m on a compact or bounded manifold of dimension d, $0 < \lambda_0 \leq \lambda_1 \leq \cdots$ and $\lambda_k \uparrow +\infty$. and the Fredholm determinant

$$\Delta(\lambda) = \prod_{k=0}^{\infty} \left(1 - \frac{\lambda}{\lambda_k}\right) \tag{K.38}$$

is regulated in the case $\mu \equiv d/m > 1$ as Weierstrass product

$$\Delta(\lambda) = \prod_{k=0}^{\infty} \left[\left(1 - \frac{\lambda}{\lambda_k}\right) \exp\left(\frac{\lambda}{\lambda_k} + \frac{\lambda^2}{2\lambda_k^2} + \cdots + \frac{\lambda^{[\mu]}}{[\mu]\lambda_k^{[\mu]}}\right)\right] \tag{K.39}$$

where $[\mu]$ denotes the integer part of μ. This is, see ref. [15], the unique entire function of order μ having zeros at $\{\lambda_k\}$ and subject to the normalization conditions

$$\ln \Delta(0) = \frac{d}{d\lambda} \ln \Delta(0) = \cdots = \frac{d^{[\mu]}}{d\lambda^{[\mu]}} \ln \Delta(0) = 0 \,. \tag{K.40}$$

Clearly eq. (K.39) is the same as (K.35); one just has to identify $z = -\lambda$, $\mathbf{A} = 1/\mathbf{H}$ and $n - 1 = [\mu]$. An example is the regularization of the spectral determinant

$$\Delta(E) = \det\left[(E - \mathbf{H})\right] \tag{K.41}$$

which – as it stands – would only make sense for a finite dimensional basis (or finite dimensional matrices). In ref. [17] the regulated spectral determinant for the example of the hyperbola billiard in two dimensions (thus $d = 2$, $m = 2$ and hence $\mu = 1$) is given as

$$\Delta(E) = \det\,[(E - \mathbf{H})\Omega(E, \mathbf{H})] \tag{K.42}$$

where

$$\Omega(E, \mathbf{H}) = -\mathbf{H}^{-1} e^{E\mathbf{H}^{-1}} \tag{K.43}$$

such that the spectral determinant in the eigenbasis of $\mathbf{H}$ (with eigenvalues $E_n \neq 0$) reads

$$\Delta(E) = \prod_n \left(1 - \frac{E}{E_n}\right) e^{E/E_n} < \infty \ . \tag{K.44}$$

Note that $\mathbf{H}^{-1}$ is for this example of Hilbert-Schmidt character.

References

[K.1] A. Wirzba, *Quantum Mechanics and Semiclassics of Hyperbolic n-Disk Scattering*, Habilitationsschrift, Technische Universität, Germany, 1997, HAB, chao-dyn/9712015, *Physics Reports* in press.

[K.2] A. Grothendieck, *"La théorie de Fredholm", Bull. Soc. Math. France*, **84**, 319 (1956).

[K.3] A. Grothendieck, *Produits tensoriels topologiques et espaces nucléaires*, Amer. Meth. Soc. **16**, Providence R. I. (1955).

[K.4] C.A. Tracy and H. Widom, CHECK THIS!: *Fredholm Determinants, Differential Equations and Matrix Models*, hep-th/9306042.

[K.5] M.G. Krein, *On the Trace Formula in Perturbation Theory* Mat.. Sborn. (N.S.) **33** (1953) 597-626; *Perturbation Determinants and Formula for Traces of Unitary and Self-adjoint Operators* Sov. Math.-Dokl. **3** (1962) 707-710. M.S. Birman and M.G. Krein, *On the Theory of Wave Operators and Scattering Operators*, Sov. Math.-Dokl. **3** (1962) 740-744.

[K.6] J. Friedel, Nuovo Cim. Suppl. **7** (1958) 287-301.

[K.7] M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Vol. I: *Functional Analysis*, Chap. VI, Academic Press (New York), 1972.

[K.8] M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Vol. IV: *Analysis of Operators*, Chap. XIII.17, Academic Press (New York), 1976.

[K.9] B. Simon, *Quantum Mechanics for Hamiltonians defined as Quadratic Forms*, Princeton Series in Physics, 1971, Appendix.

[K.10] B. Simon, *Notes on Infinite Determinants of Hilbert Space Operators*, Adv. Math. **24** (1977) 244-273.

[K.11] I.C. Gohberg and M.G. Krein, *Introduction to the theory of linear nonselfadjoint operators*, Translations of Mathematical Monographs **18**, Amer. Math. Soc. (1969).

[K.12] H. von Koch, *Sur quelques points de la théorie des déterminants infinis*, Acta. Math. **24** (1900) 89-122; *Sur la convergence des déterminants infinis*, Rend. Circ. Mat. Palermo **28** (1909) 255-266.

[K.13] E. Hille and J.D. Tamarkin, *On the characteristic values of linear integral equations*, Acta Math. **57** (1931) 1-76.

[K.14] T. Kato, *Perturbation Theory of Linear Operators* (Springer, New York, 1966), Chap. X, § 1.3 and § 1.4.

[K.15] A. Voros, *Spectral Functions, Special Functions and the Selberg Zeta Function*, *Comm. Math Phys.* **110**, 439 (1987).

[K.16] A. Voros, *Unstable periodic orbits and semiclassical quantisation*, *J. Phys.* **A 21**, 685 (1988).

[K.17] J.P. Keating and M. Sieber, *Calculation of Spectral Determinants*, preprint 1994.

Appendix L

What reviewers say

L.1 Niels Bohr

"The most important work since that Schrödinger killed the cat."

L.2 Richard P. Feynman

"Great doorstop!"

L.3 Divakar Viswanath

I have read the first six chapters. The introduction chapter is fabulous. The computation of the escape rate for the pin-ball game in the first chapter is absolutely riveting. There can't be a better way to begin. I find the book to be very stimulating.

L.4 Benny Lautrup

I am now reading your book as meticulously. I have lots of little comments, and one big one, which I can tell you immediately. I think that your opening chapter with its many literary references, jokes, and finesses, loses out in the end, because the main message gets obscured. Don't make section quotes unless they are really pertinent (although I love the first one about standing behind giants), but why cite Kierkegaard in Danish (or at all for that matter)?

L.5 Professor Gatto Nero

This book, which I have received unsolicited from the Szczsyrk Oblast Agricultural and Mazuth Office Press appears to be a collage of LaTeX clips from random papers authored by the motley collection of co-authors whose number one usually associates with an experimental high energy *Phys. Rev. Letter*, rather than a book that aspires to be the Landau-Lifshitz of chaos.

Entire rain forests went down so this not inconsiderable tome can be printed and bound. Why these ravings were not left on the Web where they more properly belong is not explained anywhere in the text. If it is an antiBourbaki, as one has in the antimatter world, then why not do the real thing? A Landau-Lifshitz for nonlinear Science, written as it should be done. The nonScience book to end all nonScience books.

Appendix M

Solutions

Chapter 1

Solution 1.1: 3-disk symbolic dynamics. *Some of the cycles are listed in table 9.1 and drawn in fig. 19.3.*

Solution 10.5: 3-disk prime cycle counting. *The formula for arbitrary length cycles is derived in sect. 10.4.*

Solution 1.2: Sensitivity to initial conditions. *To estimate the pinball sensitivity we consider a narrow beam of point particles bouncing between two disks, fig. M.1(a). Or if you find this easier to visualize, think of a narrow ray of light. We assume that the ray of light is focused along the axis between the two points. This is where the least unstable periodic orbit lies, so its stability should give us an upper bound on the number of bounces we can expect to achieve. To estimate the stability we assume that the ray of light has a width $w(t)$ and a "dispersion angle" $\theta(t)$ (we assume both are small), fig. M.1(b). Between bounces the*

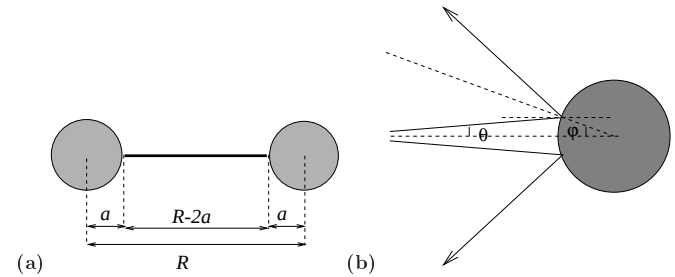


Figure M.1: The 2-disk pinball (a) geometry, (b) defocusing of scattered rays.

dispersion angle stays constant while the width increases as

$$w(t) \approx w(t') + (t - t')\theta$$

At each bounce the width stays constant while the angle increases by

$$\theta_{n+1} = \theta_n + 2\phi \approx \theta_n + w(t)/a.$$

where θ_n denotes the angle after bounce n. Denoting the width of the ray at the nth bounce by w_n then we obtain the pair of coupled equations

$$w_{n+1} = w_n + (R - 2a)\,\theta_n \tag{M.1}$$

$$\theta_n = \theta_{n-1} + \frac{w_n}{a} \tag{M.2}$$

where we ignore corrections of order w_n^2 and θ_n^2. Solving for θ_n we find

$$\theta_n = \theta_0 + \frac{1}{a}\sum_{j=1}^{n} w_n.$$

Assuming $\theta_0 = 0$ then

$$w_{n+1} = w_n + \frac{R - 2a}{a}\sum_{j=1}^{n} w_n$$

Plugging in the values in the question we find the width at each bounce in Ångstrøms grows as 1, 5, 29, 169, 985, etc. To find the asymptotic behavior for a large number of bounces we try an solution of the form $w_n = ax^n$. Substituting this into the equation above and ignoring terms that do not grow exponentially we find solutions

$$w_n \approx a w_n^{asym} = a(3 \pm 2\sqrt{2})^n$$

The solution with the positive sign will clearly dominate. The constant a we cannot determine by this local analysis although it is clearly proportional to w_0. However, the asymptotic solution is a good approximation even for quite a small number of bounces. To find an estimate of a we see that w_n/w_n^{asym} very rapidly converges to 0.146447, thus

$$w_n \approx 0.146447 w_0 (3 + 2\sqrt{2})^n \approx 0.1 \times w_0 \times 5.83^n$$

The outside edges of the ray of light will miss the disk when the width of the ray exceeds 2 cm; this occurs after 11 bounces.

(Adam Prügel-Bennett)

Solution 1.2: Sensitivity to initial conditions, another try. *Adam's estimate is not very good - do you have a better one? The first problem with it is that the instability is very underestimated. As we shall check in exercise 5.5, the exact formula for the 2-cycle stability is $\Lambda = R - 1 + R\sqrt{1 - 2/R}$. For $R = 6$, $a = 1$ this yields $w_n/w_0 \approx (5 + 2\sqrt{6})^n = 9.898979^n$, so if that were the whole story, the pinball would be not likely to make it much beyond 8 bounces.*

The second problem is that local instability overestimates the escape rate from an enclosure; trajectories are reinjected by scatterers. In the 3-disk pinball the particle leaving a disk can be reinjected by hitting either of other 2 disks, hence $w_n/w_0 \approx (9.9/2)^n$. This interplay between local instability and global reinjection will be cast into the exact formula (??) involving "Lyapunov exponent" and "Kolmogorov entropy". In order to relate this estimate to our best continuous time escape rate estimate $\gamma = 0.4103\ldots$ (see table 15.2), we will have to also compute the mean free flight time (15.20). As a crude estimate, we take the shortest disk-to-disk distance, $\langle T \rangle = R - 2 = 4$. The continuous time escape rate result implies that $w_n/w_0 \approx e^{(R-2)\gamma n} = (5.16)^n$, in the same ballpark as the above expansion-reinjection estimate.

(Predrag Cvitanović)

Solution 4.1: Trace-log of a matrix. *1) one method is to first check that this is true for any Hermitian matrix M. Then write an arbitrary complex matrix as sum $M = A + zB$, A, B Hermitian, Taylor expand in z and prove by analytic continuation that the identity applies to arbitrary M.*

(David Mermin)

2) another method: evaluate $\frac{d}{dt}\det\left(e^{t \ln M}\right)$ by definition of derivative in terms of infinitesimals.

(Kasper Juel Eriksen)

3) check appendix K.1

Chapter 2

Solution 2.2: Evolution as a group. *Can you see any other group replacing time? Try $\mathbb{Z}_{17}$ as an example; to mess things up try a non-commutative group.*

(Ronnie Mainieri)

Solution 2.9: Classical collinear helium dynamics. *An example of a solution are A. Prügel-Bennett's programs, available at*
`www.nbi.dk/ChaosBook/extras/`.

Chapter 3

Solution 5.1: A pinball simulator. *An example of a pretty Xwindows pinball is A. Prügel-Bennett's xpinball.c program, available at* `www.nbi.dk/ChaosBook/extras/xpinball.tar.gz`.

Chapter 4

Solution 4.2: How unstable is the Hénon attractor?

1. *Evaluate numerically the Lyapunov exponent by iterating the Hénon map; For $a = 1.4$, $b = 0.3$ the answer should be close to $\lambda = 0.41$. If you have a good estimate and a plot of the convergence of your estimate with t, please send us your results for possibel inclusion into this text.*
2. *Both Lyapunov exponents for $a = 1.39945219$, $b = 0.3$ are negative, roughly $\lambda_1 = -0.2712, \lambda_2 = -0.9328$ (check that these values do respect the constant volume contraction condition (4.36) for the Hénon map). Why? because after a long transient exploration of the Hénon map's non–wandering set, on average after some 11,000 iterates, almost every inital point fall into a stable 13-cycle. You can check its existence by starting at on of its periodic points $(x_p, y_p = -0.2061, -0.3181)$.*

   If you missed the stable 13-cycle (as all students in one of the courses did), you should treat your computer experiments with great deal of scepticism.

As the product of eigenvalues is constant $-b$, you need to evaluate only the expanding eigenvalue. There are many ways to implement this calculation - here are a few:

1. *The most naive way - take the log of distance of two nearby trajectories, iterate until you run out of accuracy. Tray this many times, estimate an average.*
2. *slighly smarter: as above, but keep rescaling the length of the vector connecting neighboring points so it remains small, average over the sum of logs of rescaling factors. You can run this forever.*
3. *keep multiplying the [2×2] Jacobin stability matrix (4.35) until you run out of accuracy. Compute the log of the leading eigenvalue (4.18), tray this many times, estimate an average.*
4. *slighly smarter: as above, but start with an arbitrary inital tangent space vector, keep rescaling the length of the vector connecting neighboring points so it remains small. Can run this forever.*
5. *There is probably no need to use the QR decomposition method or any other such numerical method for this 2-dimensional problem.*

(Yueheng Lan and P. Cvitanović)

Solution 5.4: Billiard exercises. *Korsch and Jodl [?] have a whole book of numerical exercises with billiards, including 3-disks.*

Solution 6.2: Linearization for maps. *(difficulty: medium) The first few terms of the map h that conjugates f to αz*

$$f(z) = h^{-1}(\alpha h(z))\,.$$

are determined many places, for example in ref. [4].

There are conditions on the derivative of f at the origin to assure that the conjugation is always possible. These conditions are formulated in ref. [11], among others.

Chapter 7

Solution 7.1: Integrating over Dirac delta functions.

(b) It does not.

(c) Integrate by parts

$$\begin{aligned} 0 &= \int dx \frac{\partial}{\partial x}\left(g(x)\delta(f(x))\right) \\ &= \int dx\,\left(g'(x)\delta(f(x)) + g(x)f'(x)\delta'(f(x))\right) \end{aligned}$$

Taking $g(x) = 1/f'(x)$ we obtain

$$\int dx\,\delta'(f(x)) = \sum_{x^*\in \mathit{Zero}\,f} \frac{f''(x^*)}{|f'(x^*)|^3}$$

Chapter 12

Solution 12.1: Numerical estimate of the escape rate for a 1-d repeller *The logistic map is defined by $x_{n+1} = Ax_n(1-x_n)$. For $A \le 4$ any point in the unit interval $[0,1]$ will remain in the interval forever. For $A > 4$ almost all points starting in the unit interval will eventually escape towards $-\infty$.*

The rate of escape can be easily measured by numerical experiment. We define the fraction of initial conditions that leave the interval after n iterations to be Γ_n. Fig. M.2 shows a plot of $\log(\Gamma_n)$ versus n, computed by starting with $10\,000\,000$ random initial points. Asymptotically the escape rate falls off exponentially as

$$\Gamma(n) = C\mathrm{e}^{-\gamma n}\,.$$

Fig. M.2 suggests that this formula is very accurate even for relatively small n. We estimate γ by measuring the slope of the curve in fig. M.2. To avoid errors due to rounding and

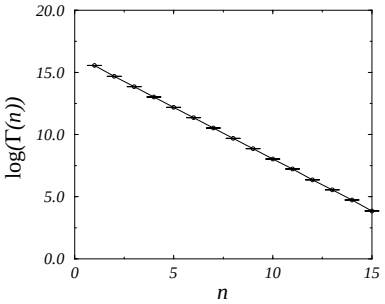


Figure M.2: Plot of $\log(\Gamma(n))$ versus n for the logistic map $x_{n+1} = 6x_n(1-x_n)$. Error bars show estimated errors in the mean assuming a binomial distribution. $10\,000\,000$ random initial starting points were used.

transients only the points $5 \leq n \leq 10$ were used. A linear regression fit yields the escape rate for $A = 6$:

$$\gamma = 0.8315 \pm 0.0001\,,$$

where the error is from statistical fluctuations (there may be systematic errors either due to rounding or because we are not in the true asymptotic regime).

(Adam Prügel-Bennet)

Solution 10.8: Spectrum of the "golden mean" pruned map.

1. *The idea is that with the redefinition $2 = 10$, the alphabet $\{1,2\}$ is unrestricted binary, and due to the piecewise linearity of the map, the stability weights factor in a way similar to (11.10).*
2. *As in (12.9), the spectral determinant for the Perron-Frobenius operator takes form (12.21)*

$$\det(1 - z\mathcal{L}) = \prod_{k=0}^{\infty} \frac{1}{\zeta_k}\,, \qquad \frac{1}{\zeta_k} = \prod_p \left(1 - \frac{z^{n_p}}{|\Lambda_p|\Lambda_p^k}\right)\,.$$

The mapping is piecewise linear, so the form of the topological zeta function worked out in (10.16) already suggests the form of the answer. The alphabet $\{1,2\}$ is unrestricted binary, so the dynamical zeta functions receive contributions only from the two fixed points, with all other cycle contributions cancelled exactly. The $1/\zeta_0$ is the spectral determinant for the transfer operator like the one in (7.13) with the $T_{00} = 0$, and in general

$$\frac{1}{\zeta_k} = \left(1 - \frac{z}{|\Lambda_1|\Lambda_1^k}\right)\left(1 - \frac{z^2}{|\Lambda_2|\Lambda_2^k}\right)\left(1 - \frac{z^3}{|\Lambda_{12}|\Lambda_{12}^k}\right)\cdots$$

$$= 1 - (-1)^k\left(\frac{z}{\Lambda^{k+1}} + \frac{z^2}{\Lambda^{2k+2}}\right)\,. \tag{M.3}$$

The factor $(-1)^k$ arises because both stabilities Λ_1 and Λ_2 include a factor $-\Lambda$ from the right branch of the map. This answer contradicts (10.37). Which answer is the right one?

Solution 12.2: Dynamical zeta functions

1. *Work through section sect. 12.3.2.*
2. *Generalize the transition matrix (9.14) to a transfer operator.*

Solution 12.5: Dynamical zeta functions as ratios of spectral determinants. *Try inserting a factor equal to one in the zeta function and then expanding it. The problem is solved in sect. 12.5.1.*

Solution 15.3: Escape rate for the Ulam map. *The answer is worked out in* Nonlinearity **3**, *325;* **3**, *361 (1990).*

Solution 7.7: Eigenvalues of the skew Ulam tent map Perron-Frobenius operator. *The first few eigenvalues are*

$$e^{s_0} = 1\,, \qquad e^{s_1} = \frac{2}{\Lambda_0} - 1$$
$$e^{s_2} = \frac{1}{4} + \frac{3}{4}\left(\frac{2}{\Lambda_0} - 1\right)^2\,, \qquad e^{s_3} = \frac{1}{2}\left(\frac{2}{\Lambda_0} - 1\right) + \frac{1}{2}\left(\frac{2}{\Lambda_0} - 1\right)^3 \ldots$$

Solution 12.9: Dynamical zeta functions for Hamiltonian maps. *Read sect. 12.5.2.*

Solution 7.5: Invariant measure. *Compare the second map to the construction of Exercise 10.6.*

Solution 13.3: Euler formula. *Let*

$$P = \prod_{k=0}^{\infty}(1 + tu^k) = \sum_{n=0}^{\infty} P_n t^n$$

then

$$P_n = \frac{1}{n!}\left.\frac{\partial^n P}{\partial t^n}\right|_{t=0} = \frac{1}{n!}\sum_{i_n \neq i_{n-1} \neq \cdots \neq i_1} u^{i_n + i_{n-1} + \cdots + i_1}$$
$$= \sum_{i_n > i_{n-1} > \cdots i_1 \geq 0} u^{i_n + i_{n-1} + \cdots + i_1} \tag{M.4}$$

Clearly $P_0 = 1$, *and*

$$P_1 = \sum_{i=0} u^i$$

multiplying both sides by $1 - u$

$$(1-u)P_1 = 1 + u + u^2 + \cdots - (u + u^2 + \cdots) = 1$$

(since, for $|u| < 1$, $\lim_{n\to\infty} u^n = 0$*). Thus* $P_1 = 1/(1-u)$. *Similarly*

$$P_2 = \sum_{i>j\geq 0} u^{i+j}$$

Graphically the allowed values of i *and* j *are*

j

Performing the same trick as for P_1

$$(1-u)P_2 = \sum_{i>j\geq 0} u^{i+j} - \sum_{i>j\geq 0} u^{i+(j+1)}$$

The only terms that survive are those for which $j = i-1$ (that is the top diagonal in the figure) thus

$$(1-u)P_2 = u^{-1}\sum_{i=1}^{\infty} u^{2i}$$

and

$$(1-u)(1-u^2)P_2 = u^{-1}\left(u^2+u^4+\cdots-(u^4+u^6+\cdots)\right) = u$$

Thus

$$P_2 = \frac{u}{(1-u)(1-u^2)}$$

In general

$$(1-u)P_n = \sum_{i_n>i_{n-1}>\cdots i_1\geq 0} u^{i_n+i_{n-1}+\cdots+i_1} - \sum_{i_n>i_{n-1}>\cdots i_1\geq 0} u^{i_n+i_{n-1}+\cdots+(i_1+1)} \tag{M.5}$$

$$= u^{-1}\sum_{i_n>i_{n-1}>\cdots i_2\geq 1} u^{i_n+i_{n-1}+\cdots+2i_2} \tag{M.6}$$

since only the term $i_1 = i_2 - 1$ survives. Repeating this trick

$$(1-u)(1-u^2)P_n = u^{-1-2}\sum_{i_n>i_{n-1}>\cdots i_3\geq 2} u^{i_n+i_{n-1}+\cdots+3i_3}$$

and

$$\prod_{i=1}^{n}(1-u^i)\,P_n = u^{-(1+2+\cdots+n)}u^{n(n-1)} = u^{n(n-1)/2}$$

Thus

$$P_n = \frac{u^{n(n-1)/2}}{\prod_{i=1}^{n}(1-u^i)}.$$

(Adam Prügel-Bennet)

Solution 13.3: Euler formula, 2nd method. *The coefficients Q_k in (13.3) are given explicitly by the* Euler formula

$$Q_k = \frac{1}{1-\Lambda^{-1}}\,\frac{\Lambda^{-1}}{1-\Lambda^{-2}}\,\cdots\,\frac{\Lambda^{-k+1}}{1-\Lambda^{-k}} \quad . \tag{M.7}$$

Such a formula is easily proved by considering the finite order product

$$\mathcal{W}_j(z,\gamma) \;=\; \prod_{l=0}^{j}(1+z\gamma^l) \;=\; \sum_{l=0}^{j+1}\Gamma_l z^l$$

Since we have that

$$(1+z\gamma^{j+1})\mathcal{W}_j(z,\gamma) \;=\; (1+z)\mathcal{W}_j(\gamma z,\gamma)\,,$$

we get the following identity for the coefficients

$$\Gamma_m + \Gamma_{m-1}\gamma^{j+1} \;=\; \Gamma_m\gamma^m + \Gamma_{m-1}\gamma^{m-1} \;\; m = 1,\ldots\,.$$

Starting with $\Gamma_0 = 1$, we recursively get

$$\Gamma_1 \;=\; \frac{1-\gamma^{j+1}}{1-\gamma} \qquad \Gamma_2 \;=\; \frac{(1-\gamma^{j+1})(\gamma-\gamma^{j+1})}{(1-\gamma)(1-\gamma^2)}\ldots\,.$$

the Euler formula (13.4) follows once we take the $j\to\infty$ limit for $|\gamma|<1$.

(Robert Artuso)

Solution 13.3: Euler formula, 3rd method. *First define*

$$f(t,u) := \prod_{k=0}^{\infty}(1+tu^k)\,. \tag{M.8}$$

Note that

$$f(t,u) = (1+t)f(tu,u)\,, \tag{M.9}$$

by factoring out the first term in the product. Now make the ansatz

$$f(t,u) = \sum_{n=0}^{\infty} t^n g_n(u)\,, \tag{M.10}$$

plug it into (M.9), compare the coefficients of t^n and get

$$g_n(u) = u^n g_n(u) + u^{n-1} g_{n-1}(u) \ . \tag{M.11}$$

Of course $g_0(u) = 1$. Therefore by solving the recursion (M.11) and by noting that $\sum_{k=1}^{n-1} k = \frac{n(n-1)}{2}$ one finally arrives at

$$g_n(u) = \frac{u^{\frac{n(n-1)}{2}}}{\prod_{k=1}^{n}(1-u^k)} \ . \tag{M.12}$$

Euler got this formula and he and Jacobi got many nice number theoretical results from it, most prominent the pentagonal number theorem, which says that in the series expansion of $\prod_{k=1}^{\infty}(1-q^k)$ all terms cancel except those which have as an exponent the circumference of a regular pentagon with integer base length.

(Juri Rolf)

Solution 13.4: 2-d product expansion. *Now let us try to apply the same trick as above to the two dimensional situation*

$$h(t,u) := \prod_{k=0}^{\infty}(1+tu^k)^{k+1} \ . \tag{M.13}$$

Write down the first terms and note that similar to (M.9)

$$h(t,u) = f(t,u)h(tu,u) \ , \tag{M.14}$$

where f is the Euler product (M.8). Now make the ansatz

$$h(t,u) = \sum_{n=0}^{\infty} t^n a_n(u) \tag{M.15}$$

and use the series expansion for f in (M.14) to get the recursion

$$a_n(u) = \frac{1}{1-u^n}\sum_{m=0}^{n-1} u^m a_m(u) g_{n-m}(u) \ . \tag{M.16}$$

With this one can at least compute the generalized Euler product effectively, but it would be nice if one could use it for a proof of the general behaviour of the coefficients a_n.

(Juri Rolf)

Chapter ??

Solution ??: Reduction of 3-disk symbolic dynamics. *The answer is given in sect. 19.6.*

Solution ??: Unimodal map symbolic dynamics. *Hint: write down an arbitrary binary number such as $\gamma = .1101001101000\ldots$ and generate the future itinerary S^+ by checking whether $f^n(\gamma)$ is greater or less than 1/2. Then verify that (9.13) recovers γ.*

Chapter 10

Solution 10.4: Transition matrix and cycle counting.

e) The topological entropy is

$$h = \ln\left(1+\frac{\sqrt{5}}{2}\right) \ .$$

Solution 10.45: Alphabet {0,1}, prune $_1000_$, $_00100_$, $_01100_$.

step 1. *$_1000_$ prunes all cycles with a $_000_$ subsequence with the exception of the fixed point $\overline{0}$; hence we factor out $(1-t_0)$ explicitly, and prune $_000_$ from the rest. Physically this means that x_0 is an isolated fixed point - no cycle stays in its vicinity for more than 2 iterations. In the notation of exercise 10.19, the alphabet is $\{1, 2, 3; \overline{0}\}$, and the remaining pruning rules have to be rewritten in terms of symbols 2=10, 3=100:*

step 2. *alphabet $\{1, 2, 3; \overline{0}\}$, prune $_33_$, $_213_$,$_313_$. Physically, the 3-cycle $\overline{3} = \overline{100}$ is pruned and no long cycles stay close enough to it for a single $_100_$ repeat. As in exercise 10.7, prohibition of $_33_$ is implemented by dropping the symbol "3" and extending the alphabet by the allowed blocks 13, 23:*

step 3. *alphabet $\{1, 2, \underline{13}, \underline{23}; \overline{0}\}$, prune $_2\underline{13}_$, $_\underline{23}\,\underline{13}_$, $_\underline{13}\,\underline{13}_$, where $\underline{13} = 13$, $\underline{23} = 23$ are now used as single letters. Pruning of the repetitions $_\underline{13}\,\underline{13}_$ (the 4-cycle $\overline{\underline{13}} = \overline{1100}$ is pruned) yields the*

Result: *alphabet $\{1, 2, \underline{23}, \underline{113}; \overline{0}\}$, unrestricted 4-ary dynamics. The other remaining possible blocks $_\underline{213}_$, $_\underline{2313}_$ are forbidden by the rules of step 3. The topological zeta function is given by*

$$1/\zeta = (1-t_0)(1-t_1-t_2-t_{23}-t_{113}) \tag{M.17}$$

for unrestricted 4-letter alphabet $\{1, 2, \underline{23}, \underline{113}\}$.

Solution 10.12: Whence Möbius function? *Written out $f(n)$ line-by-line for a few values of n, (10.38) yields*

$$f(1) \quad = \quad g(1)$$

$$\begin{aligned} f(2) &= g(2)+g(1) \\ f(3) &= g(3)+g(1) \\ f(4) &= g(4)+g(2)+g(1) \\ &\cdots \\ f(6) &= g(6)+g(3)+g(2)+g(1) \\ &\cdots \end{aligned} \tag{M.18}$$

Now invert recursively this infinite tower of equations to obtain

$$\begin{aligned} g(1) &= f(1) \\ g(2) &= f(2)-f(1) \\ g(3) &= f(3)-f(1) \\ g(4) &= f(4)-[f(2)-f(1)]-f(1) = f(4)-f(2) \\ &\cdots \\ g(6) &= f(6)-[f(3)-f(1)]-[f(2)-f(1)]-f(1) \\ &\cdots \end{aligned}$$

We see that $f(n)$ contributes with factor -1 if n prime, and not at all if n contains a prime factor to a higher power. This is precisely the raison d'etre for the Möbius function, with whose help the inverse of (10.38) can be written as the Möbius inversion formula [?] (10.39).

Chapter 14

Solution 14.3: Stability of billiard cycles. *The 2-cycle $\overline{0}$ stability (5.5) is the solution to both problems (provided you evaluate correctly the hyperbola curvature on the diagonal).*

Solution 14.4: Numerical cycle routines. *A number of sample Fortran programs for finding periodic orbits is available on the homepage for this manuscript,* `www.nbi.dk/ChaosBook/`.

Solution 14.9: Billiard cycles by path length minimization *in a given order is to start with a guess path where each bounce is given some arbitrary position on the correct disk and then iteratively improve on the guess. To accomplish this an improvement cycle is constructed whereby each bouncing point in the orbit is taken in turn and placed in a new position so that it minimizes the path. Since the positions of all the other bounces are kept constant this involves choosing the new bounce position which minimizes the path from the previous bounce to the next bounce. This problem is schematically represented in fig. M.3*

Finding the point B involves a one dimensional minimization. We define the vectors $\vec{A} = \vec{OA}$, $\vec{B} = \vec{OB}$ and $\vec{C} = \vec{OC}$. We wish to minimize the length L_{ABC} by varying $\vec{B}$ subject to the constraint that $|\vec{B}| = a$. Clearly

$$L_{ABC} = \left|\vec{A}-\vec{B}\right| + \left|\vec{C}-\vec{B}\right|$$

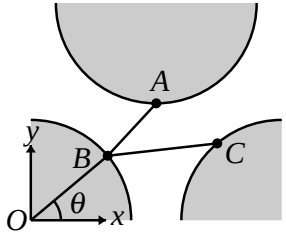


Figure M.3: Minimizing the path from the previous bounce to the next bounce.

$$= \sqrt{\vec{A}^2+\vec{B}^2-2\vec{A}\cdot\vec{B}} + \sqrt{\vec{C}^2+\vec{B}^2-2\vec{C}\cdot\vec{B}}$$

writing

$$\vec{B}(\theta) = a(\cos\theta, \sin\theta)$$

then the minima is given by

$$\frac{dL_{ABC}}{d\theta} = -\left(\frac{\vec{A}}{\sqrt{\vec{A}^2+\vec{B}^2-2\vec{A}\cdot\vec{B}}} + \frac{\vec{C}}{\sqrt{\vec{C}^2+\vec{B}^2-2\vec{C}\cdot\vec{B}}}\right)\cdot\vec{B}'(\theta) = 0.$$

The minima can then be found using a bisection algorithm or using Newton-Raphson. A simpler way is to observe that $\vec{B}'(\theta)$ is orthogonal to $\vec{B}(\theta)$ so that the vector

$$\vec{D} = \frac{\vec{A}}{\sqrt{\vec{A}^2+\vec{B}^2-2\vec{A}\cdot\vec{B}}} + \frac{\vec{C}}{\sqrt{\vec{C}^2+\vec{B}^2-2\vec{C}\cdot\vec{B}}}$$

will be proportional to $\vec{B}$. This then provides an iterative sequence for finding $\vec{B}$

- *Starting from your current guess for $\vec{B}$ calculate $\vec{D}$*
- *Put $\vec{B} = a\vec{D}/|\vec{D}|$*
- *Repeat the first step until you converge.*

At each iteration of the improvement cycle the total length of the orbit is measured. The minimization is complete when the path length stops improving. Although this algorithm is not as fast as the Newton-Raphson method, it nevertheless converges very rapidly.

(Adam Prügel-Bennet)

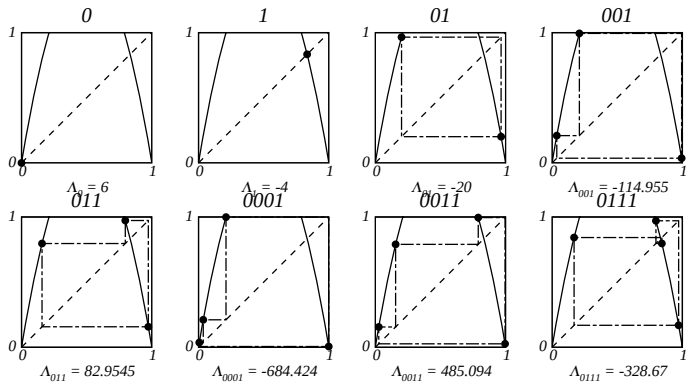


Figure M.4: Periodic orbits and stabilities for the logistics equation $x_{n+1} = 6x_n(1-x_n)$.

Chapter 15

Solution 15.2: Prime cycles for a 1-d repeller, analytic fromulas. *For the logistic map the prime cycles, ordered in terms of their symbolic dynamics, are listed in table 9.2*

$$\mathcal{P} = \{0, 1, 01, 001, 011, 0001, 0011, 0111, \ldots\}$$

The position of the prime cycles can be found by iterating the inverse mapping. If we wish to find the position of a prime orbit $p = b_1 b_2 \cdots b_{n_p}$, where $b_i \in \{0,1\}$, then starting from some initial point, $x = 1/2$ say, we apply one of the inverse mappings

$$f_\pm^{-1}(x) = \frac{1}{2} \pm \frac{1}{2}\sqrt{1 - x/4A}$$

where we choose f_-^{-1} if $b_1 = 0$ or f_+^{-1} if $b_1 = 1$. We then apply the inverse mapping again depending on the next element in the prime orbit. Repeating this procedure many times we converge onto the prime cycle. The stability Λ_p of a prime cycle p is given by the product of slopes of f around the cycle. The first eight prime cycles are shown in fig. M.4.

The stabilities of the first five prime orbits can be calculated for arbitrary A. We find that $\Lambda_0 = A$, $\Lambda_1 = 2 - A$, $\Lambda_{01} = 4 + 2A - A^2$, and

$$\Lambda_{\substack{001\\011}} = 8 + 2A - A^2 \pm A(2-A)\sqrt{A^2 - 2A - 7}. \qquad \text{(M.19)}$$

There is probably a closed form expression for the 4-cycles as well.

For crosschecking purposes: if $A = 9/2$, $\Lambda_0 = 9/2$ $\Lambda_1 = -5/2$ $\Lambda_{01} = -7.25$ $\Lambda_{011} = 19.942461\ldots$.

(Adam Prügel-Bennet)

Solution 15.2: Dynamical zeta function for a 1-d repeller *The escape rate can be estimated from the leading zero in the dynamical zeta function $1/\zeta(z)$, defined by*

$$1/\zeta(z) = \prod_p (1 - z^{n_p}/|\Lambda_p|) \ .$$

To compute the position of this pole we expand $1/\zeta(z)$ as a power series (15.5) in z

$$1/\zeta(z) = 1 - \sum_{i=1} \hat{c}_i z^i$$

where

$$\begin{aligned} \hat{c}_1 &= |\Lambda_0|^{-1} + |\Lambda_1|^{-1}\,, \qquad \hat{c}_2 = |\Lambda_{01}|^{-1} - |\Lambda_1\Lambda_0|^{-1} \\ \hat{c}_3 &= |\Lambda_{001}|^{-1} - |\Lambda_0\Lambda_{01}|^{-1} + |\Lambda_{011}|^{-1} - |\Lambda_{01}\Lambda_1|^{-1} \end{aligned}$$

etc.. Using the cycles up to length 6 we get

$$\begin{aligned} 1/\zeta(z) &= 1 - 0.416667z - 0.00833333z^2 \\ &\quad + 0.000079446z^3 - 9.89291\times 10^{-7} z^4 + \ldots \end{aligned}$$

The leading zero of this Taylor series is an estimate of $\exp(\gamma)$. Using $n = 1$, 2, 3 and 4 we obtain the increasingly accurate estimates for γ: 0.875469, 0.830597, 0.831519 and 0.831492 In a hope to improve the convergence we can use the Padé approximates $P_M^N(z) = \sum_{i=1}^N p_i z^i / (1 + \sum_{j=1}^M q_j z^j)$. Using the Padé approximates $P_1^{n-1}(z)$ for $n = 2$, 3 and 4 we obtain the estimates 0.828585, 0.831499 and 0.831493.

The above results correspond to $A = 6$; in the $A = 9/2$ case the leading zero is $1/z = 1.43549\ldots$ and $\gamma = 0.36150\ldots$. *(Adam Prügel-Bennet)*

Solution 15.2: Spectral determinant for a 1-d repeller *We are told the correct expression for the escape rate is also given by the logarithm of the leading zero of the spectral determinant (12.21), expanded as the Taylor series (15.8). The coefficients c_i should fall off super-exponentially so that truncating the Taylor series is expected to give a far more accurate estimate of the escape rate than using the dynamical zeta function. How do we compute the c_i coefficients in (15.8)? One straightforward method is to first compute the Taylor expansion of $\log(F(z))$*

$$\begin{aligned} \log(F(z)) &= \sum_p \sum_{k=0} \log\left(1 - \frac{t_p}{\Lambda_p^k}\right) = -\sum_p \sum_{k=0} \sum_{r=1} \frac{t_p^r}{\Lambda_p^{kr}} \\ &= -\sum_p \sum_{r=1} \frac{t_p^r}{1 - \Lambda_p{}^{-r}} = -\sum_p \sum_{r=1} B_p(r) z^{n_p r} \end{aligned}$$

where $B_p(r) = -1/r|\Lambda_p^r|(1+\Lambda_p^{-r})$ *. Writing* $\log(F(z))$ *as a power series*

$$\log(F(z)) = -\sum_{i=1} b_i z^i$$

we obtain

$$\begin{aligned} b_1 &= B_0(1) + B_1(1) \\ b_2 &= B_{01}(1) + B_0(2) + B_1(2) \\ b_3 &= B_{001}(1) + B_{011}(1) + B_0(3) + B_1(3) \\ b_3 &= B_{0001}(1) + B_{0011}(1) + B_{0111}(1) + B_{01}(2) + B_0(4) + B_1(4) \end{aligned} \tag{M.20}$$

etc.. To obtain the coefficients for the spectral determinant we solve

$$F(z) = 1 - \sum_{i=1} Q_i z^i = \exp\left(\sum_{i=1} b_i z^i\right)$$

for the Q_i*'s. This gives*

$$\begin{aligned} Q_1 &= b_1\,, \quad Q_2 = b_2 + b_1^2/2\,, \quad Q_3 = b_3 + b_1 b_2 + b_1^3/6 \\ Q_4 &= b_4 + b_1 b_3 + b_2^2/2 + b_2 b_1^2/2 + b_1^4/24 \end{aligned}$$

Using these formulas we find

$$F(z) = 1 - 0.4z - 0.0152381z^2 - 0.0000759784z^3 + 4.5311 \times 10^{-9} z^4 + \cdots$$

The logarithm of the leading zero of $F(z)$ *again gives the escape rate. Using the* $n = 1$*, 2, 3, and 4 truncations we find the approximation to* γ *of 0.916291, 0.832345, 0.83149289 and 0.8314929875. As predicted, the convergence is much faster for the spectral determinant than for the dynamical zeta function.*

In fig. M.5 we show a plot of the logarithm of the coefficients for the spectral determinant and for the dynamical zeta function.

(Adam Prügel-Bennet)

The above results correspond to $A = 6$*; in the* $A = 9/2$ *case all cycles up to length 10 yield* $\gamma = 0.36150966984250926\ldots$*.* *(Vadim Moroz)*

Solution 15.2: Functional dependence of escape rate for a 1-d repeller *We can compute an approximate functional dependence of the escape rate on the parameter* a

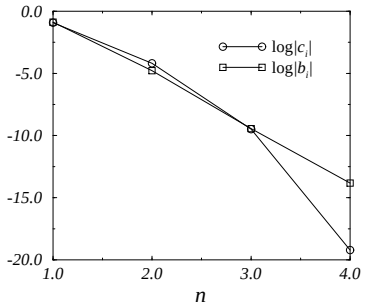


Figure M.5: Plot of the Taylor coefficients for the spectral determinant, c_i, and for the dynamical zeta function, b_i.

using the stabilities of the first five prime orbits computed above, see (M.19). The spectral determinant (for $a > 4$*) is*

$$\begin{aligned} F &= 1 - \frac{2z}{a-1} - \frac{8z^2}{(a-3)(a-1)^2(a+1)} \\ &\quad + \left(\frac{2(32 - 18a + 17a^2 - 16a^3 + 14a^4 - 6a^5 + a^6)}{(a-3)(a-1)^3(1+a)(a^2-5a+7)(a^2+a+1)} \right. \\ &\qquad \left. - \frac{2a(a-2)\sqrt{(a^2-2a-7)}}{(a^2-5a+7)(a^2-2a-7)(a^2+a+1)} \right) z^3 \end{aligned} \tag{M.21}$$

The leading zero is plotted in fig. M.6; it always remains real while the other two roots which are large and negative for $a > 5.13\ldots$ *become imaginary below this critical value. The accuracy of this truncation is clearly worst for* $a \to 4$*, the value at which the hyperbolicity is lost and the escape rate goes to zero.*

(Adam Prügel-Bennet)

Solution 15.3: Escape rate for the Ulam map. *The answer is given in ref. [2].*

Chapter 16

Solution 7.4: *The escape rate is the leading zero of the zeta function*

$$0 = 1/\zeta(\gamma) = 1 - e^\gamma/2a - e^\gamma/2a = 1 - e^\gamma/a.$$

So, $\gamma = \log(a)$ *if* $a > a_c = 1$ *and* $\gamma = 0$ *otherwise. For* $a \approx a_c$ *the escape rate behaves like*

$$\gamma(a) \approx (a - a_c).$$

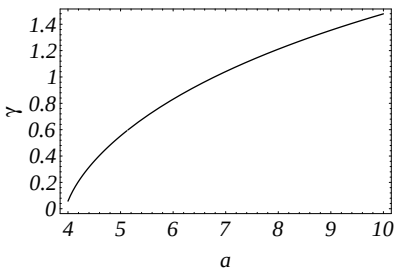


Figure M.6: Plot of the escape rate versus a for the logistic map $x_{n+1} = ax_n(1-x_n)$ calculated from the first five periodic orbits.

Solution 16.1: *The escape is controlled by the size of the primary hole of the repeller. All subholes in the repeller will be proportional with the main hole. The size of the main hole is $l = \sqrt{1-1/a}$. Near $a_c = 1$ the escape rate is*

$$\gamma(a) \sim (a - a_c)^{1/2}.$$

We can generalize this and the previous result and conclude that

$$\gamma(a) \sim (a - a_c)^{1/z},$$

where z is the order of the maximum of the single humped map.

Solution 16.2: *By direct evaluation we can calculate the zeta functions and the Fredholm determinant of this map. The zeta functions are*

$$1/\zeta_k(z) = \det(1 - z\mathbf{T}_k),$$

where

$$\mathbf{T}_k = \begin{pmatrix} T_{00}^{k+1} & T_{01}^{k+1} \\ T_{10}^{k+1} & T_{11}^{k+1} \end{pmatrix},$$

and $T_{00} = 1/a_1$, $T_{01} = (b - b/a_1)/(1-b)$, $T_{11} = (1 - b - b/a_2)/(1-b)$, $T_{10} = 1/a_2$ are inverses of the slopes of the map. The Fredholm determinant is the product of zeta functions

$$F(z) = \prod_{k=0}^{\infty} 1/\zeta_k(z).$$

The leading zeroes of the Fredholm determinant can come from the zeroes of the leading zeta functions.

The zeroes of $1/\zeta_0(z)$ are

$$\begin{aligned} 1/z_1 &= \frac{T_{00}+T_{11}+\sqrt{(T_{00}-T_{11})^2+4T_{01}T_{10}}}{2}, \\ 1/z_2 &= \frac{T_{00}+T_{11}-\sqrt{(T_{00}-T_{11})^2+4T_{01}T_{10}}}{2}. \end{aligned}$$

The zeroes of $1/\zeta_1(z)$ are

$$\begin{aligned} 1/z_3 &= \frac{T_{00}^2+T_{11}^2+\sqrt{(T_{00}^2-T_{11}^2)^2+4T_{01}^2T_{10}^2}}{2}, \\ 1/z_4 &= \frac{T_{00}^2+T_{11}^2-\sqrt{(T_{00}^2-T_{11}^2)^2+4T_{01}^2T_{10}^2}}{2}. \end{aligned}$$

By substituting the slopes we can show that $z_1 = 1$ is the leading eigenvalue. The next to leading eigenvalue, which is the correlation decay in discrete time, can be $1/z_3$ or $1/z_2$.

Chapter 17

Solution 17.1: *In the higher dimensional case there is no change in the derivation except Λ_p should be replaced with the product of expanding eigenvalues $\prod_j |\Lambda_{p,j}|$. The logarithm of this product is $\sum_j \log|\Lambda_{p,j}|$. The average of $\log|\Lambda_{,j}|$ is the j-th Lyapunov exponent.*

Solution 17.4: *The zeta function for the two scale map is*

$$1/\zeta(z,\beta) = 1 - z\left(\frac{1}{a^\beta} + \frac{1}{b^\beta}\right).$$

The pressure function is

$$P(\beta) = \log z_0(\beta) = -\log\left(\frac{1}{a^\beta} + \frac{1}{b^\beta}\right).$$

The escape rate is

$$\gamma = P(1) = -\log\left(\frac{1}{a} + \frac{1}{b}\right),$$

The topological entropy is

$$K_0 = h_{top} = -P(0) = \log 2.$$

The Lyapunov exponent is

$$\overline{\lambda} = P'(1) = \frac{\log a/a + \log b/b}{1/a + 1/b}.$$

The Kolmogorov entropy is

$$K_1 = \overline{\lambda} - \gamma = P'(1) - P(1) = \frac{\log a/a + \log b/b}{1/a + 1/b} + \log\left(\frac{1}{a} + \frac{1}{b}\right).$$

The Rényi entropies are

$$K_\beta = (P(\beta) - \beta\gamma)/(\beta - 1) = (\log\left(\frac{1}{a^\beta} + \frac{1}{b^\beta}\right) + \beta\log\left(\frac{1}{a} + \frac{1}{b}\right))/(1-\beta).$$

The box counting dimension is the solution of the implicit equation $P(D_0) = 0$, which is

$$1 = \frac{1}{a_0^D} + \frac{1}{b_0^D}.$$

The information dimension is

$$D_1 = 1 - \gamma/\overline{\lambda}.$$

The rest of the dimensions can be determined from equation $P(q-(q-1)D_q) = \gamma q$. Taking exp *of both sides we get*

$$\frac{1}{a^{q-(q-1)D_q}} + \frac{1}{b^{q-(q-1)D_q}} = \left(\frac{1}{a} + \frac{1}{b}\right)^q.$$

For a given q we can find D_q from this implicit equation.

Solution 17.5: *The zeta function is*

$$1/\zeta(z,\beta) = \det\,(1 - \mathbf{T}_{\beta-1}),$$

where we replaced k with $\beta - 1$ in Solution M. The pressure can be calculated from the leading zero which is (see Solution M)

$$P(\beta) = \log z_0(\beta) = -\log\left(\frac{T_{00}^\beta + T_{11}^\beta + \sqrt{(T_{00}^\beta - T_{11}^\beta)^2 + 4T_{01}^\beta T_{10}^\beta}}{2}\right).$$

Solution 17.6: *We can easily read off that $b = 1/2$, $a_1 = arcsin(1/2)/2\pi$ and $a_2 = a_1$ and do the steps as before.*

Chapter 20

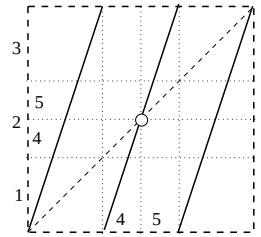
Solution 20.1: Diffusion for odd integer Λ. *Consider first the case $\Lambda = 3$, illustrated in fig. M.7. If $\beta = 0$, the dynamics in the elementary cell is simple enough; a partition can be constructed from three intervals, which we label $\{\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3\}$, with the alphabet ordered as the intervals are laid out along the unit interval. The Markov graph is fig. M.7(c), and the dynamical zeta function is*

$$1/\zeta|_{\beta=0} = 1 - (t_1 + t_2 + t_3) = 1 - 3z/\Lambda\,,$$

with eigenvalue $z = 1$ as required by the flow conservation.

However, description of global diffusion requires more care. As explained in the definition of the map (20.9), we have to split the partition $\mathcal{M}_2 = \mathcal{M}_4 \cup (\frac{1}{2}) \cup \mathcal{M}_5$, and exclude the fixed point $f(\frac{1}{2}) = \frac{1}{2}$, as the map $\hat{f}(\hat{x})$ is not defined at $\hat{f}(\frac{1}{2})$. (Are we to jump to the right or to the left at that point?) As we have $f(\mathcal{M}_4) = \mathcal{M}_1 \cup \mathcal{M}_4$, and similarly for $f(\mathcal{M}_5)$, the Markov graph fig. M.7(d) is infinite, and so is the dynamical zeta function:

$$1/\zeta = 1 - t_1 - t_{14} - t_{144} - t_{1444} \cdots - t_3 - t_{35} - t_{355} - t_{3555} \cdots .$$



(a) (b) (c) (d)

Figure M.7: (a) (b) A partition of the unit interval into three or five intervals, labeled by the order along the unit interval $\mathcal{A} = \{\mathcal{M}_1, \mathcal{M}_2 = \mathcal{M}_4 \cup (\frac{1}{2}) \cup \mathcal{M}_5, \mathcal{M}_3\}$. The partition is Markov, as the critical point is also a fixed point. (c) the Markov graph for this Markov partition.

The infinite alphabet $\mathcal{A} = \{1, 14, 144, 1444 \cdots 3, 35, 355, 3555 \cdots\}$ is a consequence of the exclusion of the fixed point(s) x_4, x_5. As is customary in such situations (see sect. ?? exercise 15.10, and chapter 18, inter alia*), we deal with this by dividing out the undesired fixed point from the dynamical zeta function. We can factorize and resum the weights using the piecewise linearity of (20.9)*

$$1/\zeta = 1 - \frac{t_1}{1-t_4} - \frac{t_3}{1-t_5}\,.$$

The diffusion constant is now most conveniently evaluated by evaluating the partial derivatives of $1/\zeta$ as in (15.16)

$$\begin{aligned} \langle T\rangle_\zeta &= -z\frac{\partial}{\partial z}\frac{1}{\zeta} = 2\left(\frac{t_1}{1-t_4} + \frac{t_1 t_4}{(1-t_4)^2}\right)\Big|_{z=1,\beta=0} = \frac{3}{4} \\ \langle \hat{x}^2\rangle_\zeta\Big|_{z=1,\beta=0} &= 2\left(\frac{\hat{n}_1(\hat{n}_1+\hat{n}_4)\Lambda^2}{(1-1/\Lambda)^2} + 2\frac{\hat{n}_4^2/\Lambda^3}{(1-1/\Lambda)^3}\right) = \frac{1}{2} \end{aligned} \tag{M.22}$$

yielding $D = 1/3$, in agreement with in (20.21) for $\Lambda = 3$.

Chapter 23

Solution 23.1: Lorentzian representation of the Dirac delta function. *To see that (23.15) is a delta function, express explicitely the imaginary part:*

$$-\lim_{\epsilon\to+0}\frac{1}{\pi}\mathrm{Im}\,\frac{E - E_n - i\epsilon}{(E-E_n)^2+\epsilon^2} = \lim_{\epsilon\to+0}\frac{1}{\pi}\frac{\epsilon}{(E-E_n)^2+\epsilon^2}. \tag{M.23}$$

This is a Lorenzian of width ϵ with a peak at $E = E_n$. It has a correct normalization for the delta function as

$$\frac{1}{\pi}\int_{-\infty}^{\infty} dE \frac{\epsilon}{(E-E_n)^2+\epsilon^2} = 1, \tag{M.24}$$

independently of the value of ϵ. Argue that in the $\epsilon \to \infty$ limit the support of the Lorentzian is concentrated at $E = E_n$, and providing that the function integrated over has a finite first derivative at $E = E_n$ and falls of sufficiently rapidly as $E \to \pm\infty$, this is a representation of the delta function.

Solution 23.6: Free particle action.

a) *a d-dimensional free particle:*
Lagrangian is the difference of kinetic and potential energy. For free motion the potential energy is $U(q) = 0$ and the velocity and the kinetic energy $m\dot{q}^2/2$ are constant. Integrating $\int_{t_0}^{t} d\tau\, L(q(\tau), \dot{q}(\tau), \tau)$ we obtain

$$R(q, q', t) = m(q-q')^2/2t$$

Solution 23.14: Stationary phase approximation. *The main contribution to such integrals comes from neighborhoods of values of x of stationary phase, the points for which the gradient of the phase vanishes*

$$\frac{\partial}{\partial x}\Phi(x) = 0.$$

Intuitively, these are the important contributions as for $\hbar \to 0$ the phase $\Phi(x)/\hbar$ grows large and the function $e^{i\Phi(x)/\hbar}$ oscillates rapidly as a function of x, with the negative and positive parts cancelling each other. More precisely, if the stationary points are well separated local extrema of $\Phi(x)$, we can deform the integration contour and approximate $\Phi(x)/\hbar$ up to the second order in x by

$$I \approx \sum_n A(x_n) e^{i\Phi(x_n)/\hbar} \int d^d x\, e^{\frac{i}{2\hbar}(x-x_n)^T \mathbf{D^2}\Phi(x_n)(x-x_n)}.$$

The second derivative matrix is a real symmetric matrix, so we can transform it to a diagonal matrix by a similarity transformation

$$Diag(\lambda_1, ..., \lambda_d) = \mathbf{O}\mathbf{D}^2\Phi\mathbf{O}^+,$$

where $\mathbf{O}$ is a matrix of an orthogonal transfomation. In the rotated coordinate system $u = \mathbf{O}(x - x_n)$ and the integral takes form

$$I \approx \sum_n A(x_n) e^{i\Phi(x_n)/\hbar} \int d^d u\, e^{\sum_{k=1}^{d} i\lambda_k u_k^2/2\hbar},$$

where we used the fact that the Jacobi determinant of an orthogonal transformation is $\det \mathbf{O} = 1$. Carrying out the Gauss integrals

$$\int du e^{i\lambda u^2/2\hbar} = \frac{(2\pi i\hbar)^{1/2}}{\sqrt{\lambda}} \tag{M.25}$$

and using $\det \mathbf{D}^2\Phi(x_n) = \prod_{k=1}^{d} \lambda_k$ we obtain the stationary phase estimate of (23.73).

A nice exposition of the subject is given in ref. [17].

Solution 23.18: A usefull determinant identity. *Divide out E in the last column of 23.75 and get the following matrix*

$$E\begin{bmatrix} x_{1,1} & \dots & x_{1,n} & y_1 E^{-1} \\ \vdots & \ddots & \vdots & \vdots \\ x_{n,1} & \dots & x_{n,n} & y_n E^{-1} \\ z_1 & \dots & z_n & 1 \end{bmatrix}$$

Now we subtract the last column multiplied with z_n from the second last column (these matrix operations does not change the determinant) to get

$$E\begin{bmatrix} x_{1,1} & \dots & x_{1,n-1} & x_{1,n} - z_n y_1 E^{-1} & y_1 E^{-1} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ x_{n,1} & \dots & x_{n,n-1} & x_{n,n} - z_n y_n E^{-1} & y_n E^{-1} \\ z_1 & \dots & z_{n-1} & 0 & 1 \end{bmatrix}$$

This continues eliminating all the z_i's in the bottom row getting the following matrix

$$E\begin{bmatrix} x_{1,1} - z_1 y_1 E^{-1} & \dots & x_{1,n} - z_n y_1 E^{-1} & y_1 E^{-1} \\ \vdots & \ddots & \vdots & \vdots \\ x_{n,1} - z_1 y_n E^{-1} & \dots & x_{n,n} - z_n y_n E^{-1} & y_n E^{-1} \\ 0 & \dots & 0 & 1 \end{bmatrix}$$

and we get (23.76) by expansion from the bottom row.

Chapter 24

Solution 24.2: Monodromy matrix from second variations of the action. *If we take two points in the configuration space q and q' connected with a trajectory with energy E and vary them in such a way that the variation of their initial and final points are transverse*

to the velocity of the orbit in that point, we can write the variations of the initial and final momenta as

$$\delta p_{\perp i} = \frac{\partial^2 S(q,q',E)}{\partial q_{\perp i}\partial q_{\perp k}}\delta q_{\perp k} + \frac{\partial^2 S(q,q',E)}{\partial q_{\perp i}\partial q'_{\perp k}}\delta q'_{\perp k} \tag{M.26}$$

and

$$\delta p'_{\perp i} = -\frac{\partial^2 S(q,q',E)}{\partial q'_{\perp i}\partial q_{\perp k}}\delta q_{\perp k} - \frac{\partial^2 S(q,q',E)}{\partial q'_{\perp i}\partial q'_{\perp k}}\delta q'_{\perp k}\,. \tag{M.27}$$

Next we express the variations of the final momenta and coordinates in terms of the initial ones. In the obvious shorthand we can write (M.27) as

$$\delta q_\perp = -S^{-1}_{q'q}S_{q'q'}\delta q'_\perp - S^{-1}_{q'q}\delta p'_\perp,$$

From (M.26) it then follows that

$$\delta p_\perp = (S_{qq'} - S_{qq}S^{-1}_{q'q}S_{q'q'})\delta q'_\perp - S_{qq}S^{-1}_{q'q}\delta p'_\perp. \tag{M.28}$$

These relations remain valid in the $q' \to q$ limit, with q on the periodic orbit, and can also be expressed in terms of the monodromy matrix of the periodic orbit. The monodromy matrix for a surface of section transverse to the orbit within the constant energy $E = H(q,p)$ shell is

$$\begin{aligned}\delta q_\perp &= \mathbf{J}_{qq}\delta q'_\perp + \mathbf{J}_{qp}\delta p'_\perp,\\ \delta p_\perp &= \mathbf{J}_{pq}\delta q'_\perp + \mathbf{J}_{pp}\delta p'_\perp.\end{aligned} \tag{M.29}$$

In terms of the second derivatives of the action the monodromy matrix is

$$\begin{aligned}\mathbf{J}_{qq} &= -S^{-1}_{q'q}S_{q'q'}\,, \qquad \mathbf{J}_{qp} = -S^{-1}_{q'q}\,,\\ \mathbf{J}_{pq} &= (S_{qq'} - S_{qq}S^{-1}_{q'q}S_{q'q'})\,, \qquad \mathbf{J}_{pp} = -S_{qq}S^{-1}_{q'q}\,,\end{aligned}$$

and vice versa

$$\begin{aligned}S_{qq} &= \mathbf{J}_{pp}\mathbf{J}^{-1}_{qp}\,, \qquad S_{qq'} = \mathbf{J}_{pq} - \mathbf{J}_{pp}\mathbf{J}^{-1}_{qp}\mathbf{J}_{qq},\\ S_{q'q} &= -\mathbf{J}^{-1}_{qp}\,, \qquad S_{q'q'} = -\mathbf{J}^{-1}_{qp}\mathbf{J}_{qq}.\end{aligned}$$

Now do exercise 24.3.

Solution 24.3: Jacobi gymnastics. *We express the Jacobi matrix elements in* $\det(\mathbf{1}-\mathbf{J})$ *with the derivative matrices of S*

$$\det(1-J) = \det\begin{pmatrix} I + S^{-1}_{q'q}S_{q'q'} & S^{-1}_{q'q} \\ -S_{qq'} + S_{qq}S^{-1}_{q'q}S_{q'q'} & I + S_{qq}S^{-1}_{q'q}\end{pmatrix}.$$

We can multiply the second column with $S_{q'q'}$ from the and substract from the first column, leaving the determinant unchanged

$$\det(1-J) = \det\begin{pmatrix} I & S^{-1}_{q'q} \\ -S_{qq'} - S_{q'q'} & I + S_{qq}S^{-1}_{q'q}\end{pmatrix}.$$

Then, we multiply the second column with $S_{q'q}$ from the right and compensate this by dividing the determinant with $\det S_{q'q}$

$$\det(1-J) = \det\begin{pmatrix} I & I \\ -S_{qq'} - S_{q'q'} & S_{q'q} + S_{qq}\end{pmatrix}/\det S_{q'q}.$$

Finally we subtract the first column from the second one

$$\det(1-J_j)) = \det\begin{pmatrix} I & 0 \\ S_{qq'} + S_{q'q'} & S_{qq'} + S_{q'q'} + S_{q'q} + S_{qq}\end{pmatrix}/\det S_{q'q}.$$

The last determinant can now be evaluated and yields the desired result (24.15)

$$\det(1-J_j) = \det(S_{qq'} + S_{q'q'} + S_{q'q} + S_{qq})/\det S_{q'q}.$$

Chapter ??

Solution ??: The one-disk scattering wave function.

$$\psi(\vec{r}) = \frac{1}{2}\sum_{m=-\infty}^{\infty}\left(H^{(2)}_m(kr) - \frac{H^{(2)}_m(ka)}{H^{(1)}_m(ka)}H^{(1)}_m(kr)\right)e^{im(\Phi_r - \Phi_k)}\,. \tag{M.30}$$

(For $r < a$, $\psi(\vec{r}) = 0$ of course.)

(Andreas Wirzba)

Solution ??: Ghosts do not exist. *In ref. [1] the ghost cancellation rule (??) is proved for the convolution of two $\mathbf{A}$-matrices and generalized to hold also inside an arbitrary (periodic) itinerary (with and without creeping sections).*

Consider the itinerary $(1,\underline{2},3,4,\underline{5},6)$ with ghost sections at disk 2 and 5 resulting from the sixth order trace. Its geometrical contribution cancels in the trace-log expansion against the geometrical reduction of the itineraries $(1,\underline{2},3,4,6)$, $(1,3,4,\underline{5},6)$ from the 5th-order trace with ghost sections at disk 2 or 5, respectively, and against the geometrical reduction of the itinerary $(1,3,4,6)$ of the 4th-order trace with no ghost contribution:

$$\begin{aligned}
&-\frac{1}{6}\left(6\,\underline{\mathbf{A}}^{1,\underline{2}}\underline{\mathbf{A}}^{\underline{2},3}\mathbf{A}^{3,4}\underline{\mathbf{A}}^{4,\underline{5}}\underline{\mathbf{A}}^{\underline{5},6}\mathbf{A}^{6,1}\right)\\
&\qquad-\frac{1}{5}\left(5\,\underline{\mathbf{A}}^{1,\underline{2}}\underline{\mathbf{A}}^{\underline{2},3}\mathbf{A}^{3,4}\mathbf{A}^{4,6}\mathbf{A}^{6,1}+5\,\mathbf{A}^{1,3}\mathbf{A}^{3,4}\underline{\mathbf{A}}^{4,\underline{5}}\underline{\mathbf{A}}^{\underline{5},6}\mathbf{A}^{6,1}\right)\\
&\qquad-\frac{1}{4}\left(4\,\mathbf{A}^{1,3}\mathbf{A}^{3,4}\mathbf{A}^{4,6}\mathbf{A}^{6,1}\right)\\
&=\quad(-1+2-1)\,\mathbf{A}^{1,3}\mathbf{A}^{3,4}\mathbf{A}^{4,6}\mathbf{A}^{6,1}=0\,.
\end{aligned}$$

The prefactors $-1/4$, $-1/5$, $-1/6$ result from the trace-log expansion, the factors 4, 5, 6 inside the brackets are due to the cyclic permutations, and the rule (??) was used. If there are two or more ghost segments adjacent to each other, the ghost rule (??) has to be generalized to

$$\begin{aligned}
&\cdots\underline{\mathbf{A}}^{i,\underline{i+1}}\underline{\mathbf{A}}^{\underline{i+1},\underline{i+2}}\cdots\underline{\mathbf{A}}^{\underline{i+k},\underline{i+k+1}}\cdots\underline{\mathbf{A}}^{\underline{i+n-1},i+n}\cdots\\
&=\quad\cdots\left(-\underline{\mathbf{A}}^{i,\underline{i+2}}\right)\cdots\underline{\mathbf{A}}^{\underline{i+k},\underline{i+k+1}}\cdots\underline{\mathbf{A}}^{\underline{i+n-1},i+n}\cdots\\
&=\quad\cdots\underline{\mathbf{A}}^{i,\underline{i+3}}\cdots\underline{\mathbf{A}}^{\underline{i+k},\underline{i+k+1}}\cdots\underline{\mathbf{A}}^{\underline{i+n-1},i+n}\cdots\\
&=\quad\cdots(-1)^{n-1}\mathbf{A}^{i,i+n}\cdots\,. \qquad\text{(M.31)}
\end{aligned}$$

Finally, let us discuss one case with a repeat, e.g. the itinerary $(1,\underline{2},3,4,1,\underline{2},3,4)$ with repeated ghost sections at disk 2 in the semiclassical limit. The cancellations proceed in the trace-log expansion as follows:

$$\begin{aligned}
&-\frac{1}{8}\left(4\,\underline{\mathbf{A}}^{1,\underline{2}}\underline{\mathbf{A}}^{\underline{2},3}\mathbf{A}^{3,4}\mathbf{A}^{4,1}\underline{\mathbf{A}}^{1,\underline{2}}\underline{\mathbf{A}}^{\underline{2},3}\mathbf{A}^{3,4}\mathbf{A}^{4,1}\right)\\
&\qquad-\frac{1}{7}\left(7\,\underline{\mathbf{A}}^{1,\underline{2}}\underline{\mathbf{A}}^{\underline{2},3}\mathbf{A}^{3,4}\mathbf{A}^{4,1}\mathbf{A}^{1,3}\mathbf{A}^{3,4}\mathbf{A}^{4,1}\right)\\
&\qquad-\frac{1}{6}\left(3\,\mathbf{A}^{1,3}\mathbf{A}^{3,4}\mathbf{A}^{4,1}\mathbf{A}^{1,3}\mathbf{A}^{3,4}\mathbf{A}^{4,1}\right)\\
&=\quad\left(-\frac{1}{2}+1-\frac{1}{2}\right)\,\left[\mathbf{A}^{1,3}\mathbf{A}^{3,4}\mathbf{A}^{4,1}\right]^2=0
\end{aligned}$$

Note that the cyclic permutation factors of the 8th and 6th order trace are halved because of the repeat. The occurrence of the ghost segment in the second part of the 7th order itinerary is taken care of by the weight factor 7.

(Andreas Wirzba)

Chapter H

Solution 16.3:

(d) In the $A=9/2$ case all cycles up to length 9 yield $\lambda=1.08569\ldots$. (Vadim Moroz)

Solution H.1: *Using the multiplicative property of the Jacobi matrix we can write*

$$\Lambda^{t'+t}(x_0,\mathbf{u}_0)=||\mathbf{J}^{t'+t}(x_0)\mathbf{u}_0||=||\mathbf{J}^{t'}(x(t))\mathbf{J}^{t}(x_0)\mathbf{u}_0||.$$

We can introduce the time evolved unit vector

$$\mathbf{u}(t)=\mathbf{J}^{t}(x_0)\mathbf{u}_0/||\mathbf{J}^{t}(x_0)\mathbf{u}_0||.$$

Then

$$||\mathbf{J}^{t'}(x(t))\mathbf{J}^{t}(x_0)\mathbf{u}_0||=||\mathbf{J}^{t'}(x(t))\mathbf{u}(t)||\,||\mathbf{J}^{t}(x_0)\mathbf{u}_0||,$$

which is the desired result.

We have to adjoin the tangent space, since the stretching factor depends on $\mathbf{u}$ and not just on x. The stretching factor is multiplicative along the entire trajectory $(x(t),\mathbf{u}(t))$. However, it is not multiplicative along the phase space trajectory $x(t)$ with a fixed $\mathbf{u}$.

Solution H.2: *If $b=a^2$ and $T_b=2T_a$ we can introduce the variable $y=e^{sT_a}$. The dynamo rate equation then reads*

$$0=1-x+x^2.$$

The solutions of this are $x_\pm=(1\pm i\sqrt{3})/2$. The dynamo rate is then a complex cojugate pair $\nu=\log x_\pm/T_a$.

The escape rate equation is

$$0=1-x/a-x^2/a^2.$$

The solutions are $x_\pm=a(-1\pm\sqrt{5})/2$. The escape rate is $\gamma=\log(x_+)/T_a$.

In the reverse case the escape rate remains unchanged, while the dynamo rate becomes $\nu=\log((\sqrt{5}+1)/2)/T_a$. In this case the advected field grows with an exponential rate. In the previous case it shows oscillations in addition to the exponential growth due to the imaginary part of the rate.

Appendix N

Projects

You are urged to try to work through the essential steps in a project that combines the techniques learned in the course with some application of interest to you for other reasons. It is OK to share computer programs and such, but otherwise each project should be distinct, not a group project. The essential steps are:

- **Dynamics**
  1. construct a symbolic dynamics
  2. count prime cycles
  3. prune inadmissible itineraries, construct Markov graphs if appropriate
  4. implement a numerical simulator for your problem
  5. compute a set of the shortest periodic orbits
  6. compute cycle stabilities
- **Averaging, numerical**
  1. estimate by numerical simulation some observable quantity, like the escape rate,
  2. or check the flow conservation, compute something like the Lyapunov exponent
- **Averaging, periodic orbits**
  1. implement the appropriate cycle expansions
  2. check flow conservation as function of cycle length truncation, if the system is closed
  3. implement desymmetrization, factorization of zeta functions, if dynamics possesses a discrete symmetry

4. compute a quantity like the escape rate as a leading zero of a spectral determinant or a dynamical zeta function.
5. or evaluate a sequence of truncated cycle expansions for averages, such as the Lyapunov exponent or/and diffusion coefficients
6. compute a physically intersting quantity, such as the conductance
7. compute some number of the classical and/or quantum eigenvalues, if appropriate

N.1 Deterministic diffusion, zig-zag map

To illustrate the main idea of chapter 20, tracking of a globally diffusing orbit by the associated confined orbit restricted to the fundamental cell, we consider a class of simple 1-d dynamical systems, chains of piecewise linear maps, where all transport coefficients can be evaluated analytically. The translational symmetry (20.10) relates the unbounded dynamics on the real line to the dynamics restricted to a "fundamental cell" - in the present example the unit interval curled up into a circle. An example of such map is the sawtooth map

$$\hat{f}(x) = \begin{cases} \Lambda x & x \in [0, 1/4 + 1/4\Lambda] \\ -\Lambda x + (\Lambda+1)/2 & x \in [1/4 + 1/4\Lambda, 3/4 - 1/4\Lambda] \\ \Lambda x + (1-\Lambda) & x \in [3/4 - 1/4\Lambda, 1] \end{cases} . \tag{N.1}$$

The corresponding circle map $f(x)$ is obtained by modulo the integer part. The elementary cell map $f(x)$ is sketched in fig. N.1. The map has the symmetry property

$$\hat{f}(\hat{x}) \;=\; -\hat{f}(-\hat{x})\,, \tag{N.2}$$

so that the dynamics has no drift, and all odd derivatives of the generating function (20.3) with respect to β evaluated at $\beta = 0$ vanish.

The cycle weights are given by

$$t_p = z^{n_p} \frac{e^{\beta \hat{n}_p}}{|\Lambda_p|}\,. \tag{N.3}$$

The diffusion constant formula for 1-d maps is

$$D = \frac{1}{2} \frac{\langle \hat{n}^2 \rangle_\zeta}{\langle n \rangle_\zeta} \tag{N.4}$$

where the "mean cycle time" is given by

$$\langle n \rangle_\zeta = z \frac{\partial}{\partial z} \frac{1}{\zeta(0,z)}\bigg|_{z=1} = -{\sum}' (-1)^k \frac{n_{p_1} + \cdots + n_{p_k}}{|\Lambda_{p_1} \cdots \Lambda_{p_k}|}\,, \tag{N.5}$$

the mean cycle displacement squared by

$$\langle \hat{n}^2 \rangle_\zeta = \frac{\partial^2}{\partial \beta^2} \frac{1}{\zeta(\beta,1)}\bigg|_{\beta=0} = -{\sum}' (-1)^k \frac{(\hat{n}_{p_1} + \cdots + \hat{n}_{p_k})^2}{|\Lambda_{p_1} \cdots \Lambda_{p_k}|}\,, \tag{N.6}$$

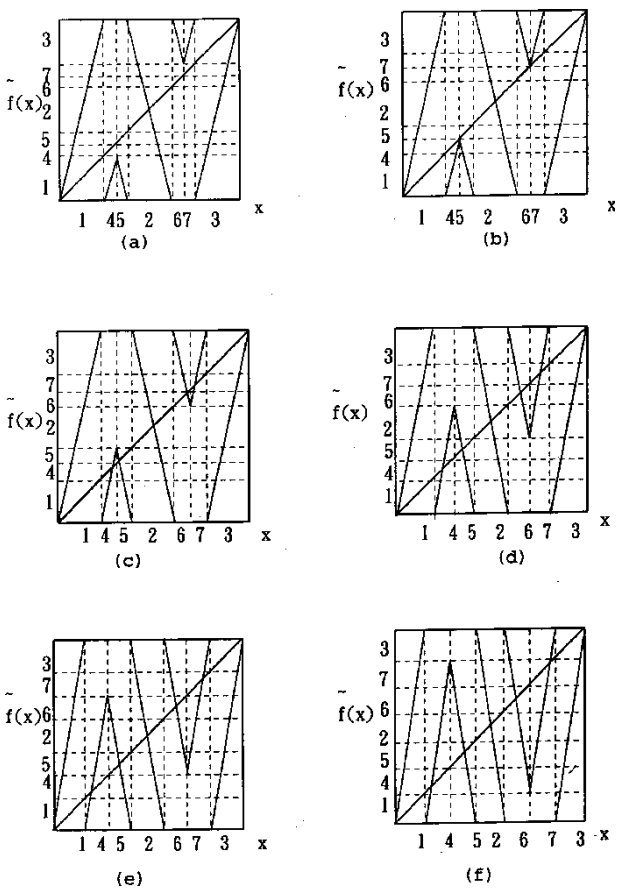


Figure N.1: (a)-(f) The sawtooth map (N.1) for the 6 values of parameter a for which the folding point of the map aligns with the endpoint of one of the 7 intervals and yields a finite Markov partition (from ref. [1]). The corresponding Markov graphs are given in fig. N.2.

and the sum is over all distinct non-repeating combinations of prime cycles. Most of results expected in this projects require no more than pencil and paper computations.

Implementing the symmetry factorization (20.35) is convenient, but not essential for this project, so if you find sect. 19.1.2 too long a read, skip the symmetrization.

N.1.1 The full shift

Take the map (N.1) and extend it to the real line. As in example of fig. 20.3, denote by a the critical value of the map (the maximum height in the unit cell)

$$a = \hat{f}(\frac{1}{4} + \frac{1}{4\Lambda}) = \frac{\Lambda+1}{4}\,. \tag{N.7}$$

Describe the symbolic dynamics that you obtain when a is an integer, and derive the formula for the diffusion constant:

$$D = \frac{(\Lambda^2-1)(\Lambda-3)}{96\Lambda} \qquad \text{for } \Lambda = 4a-1,\ a\in\mathbb{Z}\,. \tag{N.8}$$

If you are going strong, derive also the fromula for the half-integer $a = (2k+1)/2$, $\Lambda = 4a+1$ case and email it to DasBuch@nbi.dk. You will need to partition $\mathcal{M}_2$ into the left and right half, $\mathcal{M}_2 = \mathcal{M}_8 \cup \mathcal{M}_9$, as in the derivation of (20.21).

20.1 on p. 399

N.1.2 Subshifts of finite type

We now work out an example when the partition is Markov, although the slope is not an integer number. The key step is that of having a partition where intervals are mapped *onto* unions of intervals. Consider for example the case in which $\Lambda = 4a-1$, where $1 \le a \le 2$. A first partition is constructed from seven intervals, which we label $\{\mathcal{M}_1, \mathcal{M}_4, \mathcal{M}_5, \mathcal{M}_2, \mathcal{M}_6, \mathcal{M}_7, \mathcal{M}_3\}$, with the alphabet ordered as the intervals are laid out along the unit interval. In general the critical value a will not correspond to an interval border, but now we choose a such that the critical point is mapped onto the right border of $\mathcal{M}_1$, as in fig. N.1(a). The critical value of $f()$ is $f(\frac{\Lambda+1}{4\Lambda}) = a-1 = (\Lambda-3)/4$. Equating this with the right border of $\mathcal{M}_1$, $x = 1/\Lambda$, we obtain a quadratic equation with the expanding solution $\Lambda = 4$. We have that $f(\mathcal{M}_4) = f(\mathcal{M}_5) = \mathcal{M}_1$, so the transition matrix

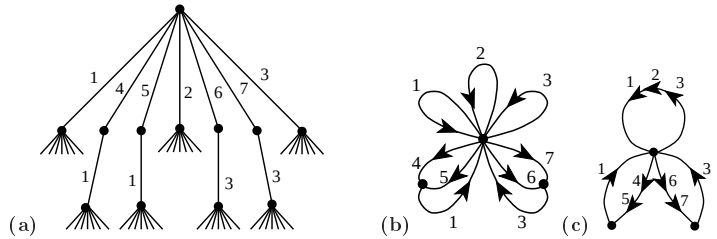


Figure N.2: (a) The sawtooth map (N.1) partition tree for fig. N.1(a); while intervals $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$ map onto the whole unit interval, $f(\mathcal{M}_1) = f(\mathcal{M}_2) = f(\mathcal{M}_3) = \mathcal{M}$, intervals $\mathcal{M}_4, \mathcal{M}_5$ map onto $\mathcal{M}_1$ only, $f(\mathcal{M}_4) = f(\mathcal{M}_5) = \mathcal{M}_1$, and similarly for intervals $\mathcal{M}_6, \mathcal{M}_7$. An initial point starting out in the interval $\mathcal{M}_1$, $\mathcal{M}_2$ or $\mathcal{M}_3$ can land anywhere on the unit interval, so the subtrees originating from the corresponding nodes on the partition three are similar to the whole tree and can be identified (as, for example, in fig. 9.12), yielding (b) the Markov graph for the Markov partition of fig. N.1(a). (c) the Markov graph in the compact notation of (20.26).

(9.2) is given by

$$\phi' = T\phi \;=\; \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_4 \\ \phi_5 \\ \phi_2 \\ \phi_6 \\ \phi_7 \\ \phi_3 \end{pmatrix} \tag{N.9}$$

and the dynamics is unrestricted in the alphabet

$$\{1,\, \underline{41},\, \underline{51},\, 2,\, \underline{63},\, \underline{73},\, 3,\}\,.$$

One could diagonalize (N.9) on the computer, but, as we saw in sect. 9.6, the Markov graph fig. N.2(b) corresponding to fig. N.1(a) offers more insight into the dynamics. The dynamical zeta function

$$\begin{aligned} 1/\zeta &= 1 - (t_1 + t_2 + t_3) - 2(t_{14} + t_{37}) \\ 1/\zeta &= 1 - 3\frac{z}{\Lambda} - 4\cosh\beta\frac{z^2}{\Lambda^2}\,. \end{aligned} \tag{N.10}$$

follows from the loop expansion (10.13) of sect. 10.3.

The material flow conservation sect. 16.2 and the symmetry factorization (20.35) yield

$$0 = \frac{1}{\zeta(0,1)} = \left(1 + \frac{1}{\Lambda}\right)\left(1 - \frac{4}{\Lambda}\right)$$

which indeed is satisfied by the given value of Λ. Conversely, we can use the desired Markov partition topology to write down the corresponding dynamical zeta function, and use the $1/\zeta(0,1) = 0$ condition to fix Λ. For more complicated transition matrices the factorization (20.35) is very helpful in reducing the order of the polynomial condition that fixes Λ.

The diffusion constant follows from (20.36) and (N.4)

$$\langle n\rangle_\zeta = -\left(1 + \frac{1}{\Lambda}\right)\left(-\frac{4}{\Lambda}\right)\,, \quad \langle \hat{n}^2\rangle_\zeta = \frac{4}{\Lambda^2}$$

$$D = \frac{1}{2}\,\frac{1}{\Lambda + 1} = \frac{1}{10}$$

Think up other non-integer values of the parameter for which the symbolic dynamics is given in terms of Markov partitions: in particular consider the cases illustrated in fig. N.1 and determine for what value of the parameter a each of them is realized. Work out the Markov graph, symmetrization factorization and the diffusion constant, and check the material flow conservation for each case. Derive the diffusion constants listed in table N.1. It is not clear why the final answers tend to be so simple. Numerically, the case of fig. N.1(c) appears to yield the maximal diffusion constant. Does it? Is there an argument that it should be so?

The seven cases considered here (see table N.1, fig. N.1 and (N.8)) are the 7 simplest complete Markov partitions, the criterion being that the critical points map onto partition boundary points. This is, for example, what happens for unimodal tent map; if the critical point is preperiodic to an unstable cycle, the grammar is complete. The simplest example is the case in which the tent map critical point is preperiodic to a unimodal map 3-cycle, in which case the grammar is of golden mean type, with _00_ substring prohibited (see fig. 9.12). In case at hand, the "critical" point is the junction of branches 4 and 5 (symmetry automatically takes care of the other critical point, at the junction of branches 6 and 7), and for the cases considered the critical point maps into the endpoint of each of the seven branches.

One can fill out parameter a axis arbitrarily densely with such points - each of the 7 primary intervals can be subdivided into 7 intervals obtained by 2-nd iterate of the map, and for the critical point mapping into any of those in 2 steps the grammar (and the corresponding cycle expansion) is finite, and so on.

fig. N.1	Λ	D
	3	0
(a)	4	$\frac{1}{10}$
(b)	$\sqrt{5}+2$	$\frac{1}{2\sqrt{5}}$
(c)	$\frac{1}{2}(\sqrt{17}+5)$	$\frac{2}{\sqrt{17}}$
(c')	5	$\frac{2}{5}$
(d)	$\frac{1}{2}(\sqrt{33}+5)$	$\frac{1}{8}+\frac{3}{88}\sqrt{33}$
(e)	$2\sqrt{2}+3$	$\frac{1}{2\sqrt{2}}$
(f)	$\frac{1}{2}(\sqrt{33}+7)$	$\frac{1}{4}+\frac{1}{4\sqrt{33}}$
	7	$\frac{2}{7}$

Table N.1: The diffusion constant as function of the slope Λ for the $a = 1, 2$ values of (N.8) and the 6 Markov partitions of fig. N.1

N.1.3 Diffusion coefficient, numerically

(optional:)
Attempt a numerical evaluation of

$$D = \frac{1}{2}\lim_{n\to\infty}\frac{1}{n}\langle \hat{x}_n^2\rangle \ . \tag{N.11}$$

Study the convergence by comparing your numerical results to the exact answers derived above. Is it better to use few initial $\hat{x}$ and average for long times, or to use many initial $\hat{x}$ for shorter times? Or should one fit the distribution of $\hat{x}^2$ with a gaussian and get the D this way? Try to plot dependence of D on Λ; perhaps blow up a small region to show that the dependance of D on the parameter Λ is fractal. Compare with fig. 20.5 and figures in refs. [1, 2, 7, 8].

N.1.4 D is a nonuniform function of the parameters

(optional:)
The dependence of D on the map parameter Λ is rather unexpected - even though for larger Λ more points are mapped outside the unit cell in one iteration, the diffusion constant does not necessarily grow. An interpretation of this lack of monotonicity would be interesting.

You can also try applying periodic orbit theory to the sawtooth map (N.1) for a random "generic" value of the parameter Λ, for example $\Lambda = 6$. The idea is to bracket this value of Λ by the nearby ones, for which higher and higher iterates of the critical value $a = (\Lambda+1)/4$ fall onto the partition boundaries, compute the exact diffusion constant for each such approximate Markov partition, and study

their convergence toward the value of D for $\Lambda = 6$. Judging how difficult such problem is already for a tent map (see sect. 10.6 and appendix E.1), this is too ambitious for a week-long exam.

References

[N.1] H.-C. Tseng, H.-J. Chen, P.-C. Li, W.-Y. Lai, C.-H. Chou and H.-W. Chen, "Some exact results for the diffusion coefficients of maps with pruned cycles", *Phys. Lett.* **A** 195, 74 (1994).

[N.2] C.-C. Chen, "Diffusion Coefficient of Piecewise Linear Maps", *Phys. Rev.* **E51**, 2815 (1995).

[N.3] H.-C. Tseng and H.-J. Chen, "Analytic results for the diffusion coefficient of a piecewise linear map", *Int. J. Mod. Phys.***B 10**, 1913 (1996).

N.2 Deterministic diffusion, sawtooth map

To illustrate the main idea of chapter 20, tracking of a globally diffusing orbit by the associated confined orbit restricted to the fundamental cell, we consider in more detail the class of simple 1-d dynamical systems, chains of piecewise linear maps (20.9). The translational symmetry (20.10) relates the unbounded dynamics on the real line to the dynamics restricted to a "fundamental cell" - in the present example the unit interval curled up into a circle. The corresponding circle map $f(x)$ is obtained by modulo the integer part. The elementary cell map $f(x)$ is sketched in fig. 20.3. The map has the symmetry property

$$\hat{f}(\hat{x}) \;=\; -\hat{f}(-\hat{x})\,, \tag{N.12}$$

so that the dynamics has no drift, and all odd derivatives of the generating function (20.3) with respect to β evaluated at $\beta = 0$ vanish.

The cycle weights are given by

$$t_p = z^{n_p}\frac{e^{\beta\hat{n}_p}}{|\Lambda_p|}\,. \tag{N.13}$$

The diffusion constant formula for 1-d maps is

$$D = \frac{1}{2}\frac{\langle \hat{n}^2\rangle_\zeta}{\langle n\rangle_\zeta} \tag{N.14}$$

where the "mean cycle time" is given by

$$\langle n\rangle_\zeta = z\frac{\partial}{\partial z}\frac{1}{\zeta(0,z)}\bigg|_{z=1} = -{\sum}'(-1)^k\frac{n_{p_1}+\cdots+n_{p_k}}{|\Lambda_{p_1}\cdots\Lambda_{p_k}|}\,, \tag{N.15}$$

the mean cycle displacement squared by

$$\langle \hat{n}^2\rangle_\zeta = \frac{\partial^2}{\partial\beta^2}\frac{1}{\zeta(\beta,1)}\bigg|_{\beta=0} = -{\sum}'(-1)^k\frac{(\hat{n}_{p_1}+\cdots+\hat{n}_{p_k})^2}{|\Lambda_{p_1}\cdots\Lambda_{p_k}|}\,, \tag{N.16}$$

and the sum is over all distinct non-repeating combinations of prime cycles. Most of results expected in this projects require no more than pencil and paper computations.

fig. 20.4	Λ	D
	4	$\frac{1}{4}$
(a)	$2+\sqrt{6}$	$1-\frac{3}{4}\sqrt{6}$
(b)	$2\sqrt{2}+2$	$\frac{15+2\sqrt{2}}{16+4\sqrt{2}}$
(c)	5	1
(d)	$3+\sqrt{5}$	$\frac{5}{2}\frac{\Lambda-1}{3\Lambda-4}$
(e)	$3+\sqrt{7}$	$\frac{5\Lambda-4}{3\Lambda-2}$
	6	$\frac{5}{6}$

Table N.2: The diffusion constant as function of the slope Λ for the $\Lambda = 4, 6$ values of (20.20) and the 5 Markov partitions like the one indicated in fig. 20.4.

N.2.1 The full shift

Reproduce the formulas of sect. 20.2.1 for the diffusion constant D for Λ both even and odd integer.

N.2.2 Subshifts of finite type

We now work out examples when the partition is Markov, although the slope is not an integer number. The key step is that of having a partition where intervals are mapped *onto* unions of intervals.

Start by reproducing the formula (20.28) of sect. 20.2.3 for the diffusion constant D for the Markov partition, the case where the critical point is mapped onto the right border of I_{1_+}.

Think up other non-integer values of the parameter Λ for which the symbolic dynamics is given in terms of Markov partitions: in particular consider the remaing four cases for which the critical point is mapped onto a border of a partion in one iteration. Work out the Markov graph symmetrization factorization and the diffusion constant, and check the material flow conservation for each case. Fill in the diffusion constants missing in table N.2. It is not clear why the final answers tend to be so simple. What value of Λ appears to yield the maximal diffusion constant?

The 7 cases considered here (see table N.2 and fig. 20.4) are the 7 simplest complete Markov partitions in the $4 \le \Lambda \le 6$ interval, the criterion being that the critical points map onto partition boundary points. In case at hand, the "critical" point is the highest point of the left branch of the map (symmetry automatically takes care of the other critical point, the lowest point of the left branch), and for the cases considered the critical point maps into the endpoint of each of the seven branches.

One can fill out parameter a axis arbitrarily densely with such points - each of the 6 primary intervals can be subdivided into 6 intervals obtained by 2-nd iterate of the map, and for the critical point mapping into any of those in 2 steps the grammar (and the corresponding cycle expansion) is finite, and so on. Some details of how this is accomplished are given in appendix **??** for a related problem, the pruned Bernulli shift.

N.2.3 Diffusion coefficient, numerically

(optional:)
Attempt a numerical evaluation of

$$D = \frac{1}{2} \lim_{n\to\infty} \frac{1}{n} \langle \hat{x}_n^2 \rangle \ . \tag{N.17}$$

Study the convergence by comparing your numerical results to the exact answers derived above. Is it better to use few initial $\hat{x}$ and average for long times, or to use many initial $\hat{x}$ for shorter times? Or should one fit the distribution of $\hat{x}^2$ with a gaussian and get the D this way? Try to plot dependence of D on Λ; perhaps blow up a small region to show that the dependance of D on the parameter Λ is fractal. Compare with fig. 20.5 and figures in refs. [1, 2, 7, 8].

N.2.4 D is a nonuniform function of the parameters

(optional:)
The dependence of D on the map parameter Λ is rather unexpected - even though for larger Λ more points are mapped outside the unit cell in one iteration, the diffusion constant does not necessarily grow. Fig. 20.5 taken from ref. [7] illustrates the fractal dependence of diffusion constant on the map parameter. An interpretation of this lack of monotonicity would be interesting.

You can also try applying periodic orbit theory to the sawtooth map (20.9) for a random "generic" value of the parameter Λ, for example $\Lambda = 4.5$. The idea is to bracket this value of Λ by the nearby ones, for which higher and higher iterates of the critical value $a = \Lambda/2$ fall onto the partition boundaries, compute the exact diffusion constant for each such approximate Markov partition, and study their convergence toward the value of D for $\Lambda = 4.5$. Judging how difficult such problem is already for a tent map (see sect. 10.6 and appendix E.1), this is too ambitious for a week-long exam.

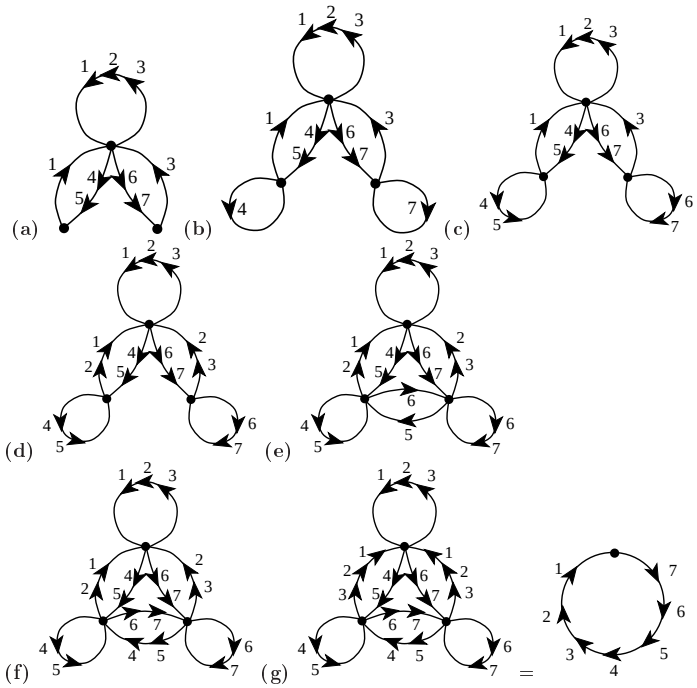


Figure N.3: (a)-(g) Markov graphs corresponding to the cases of fig. N.1 for which the sawtooth map (N.1) folding point aligns with an endpoint of one of the 7 intervals and yields a finite Markov partition. In the case (g) the large Markov graph is equivalent to the simple unrestricted shift in 7 symbols. Graphs (f), (e), ..., are obtained by pruning links from the unrestricted expanded graph (g).

Chapter N

Solution, project sect. N.1: Markov partitions for the zig-zag map

Fig. N.1(b) $\Lambda = \sqrt{5}+2$ with the corresponding Markov graph given in fig. N.3(b). As the 4 and 5 links (and other sets of links) are parallel, we have simplified the graph using the compact notation of (20.26). The loop expansion (10.13) yields

$$\begin{aligned} 1/\zeta \ &= \ 1 - (t_1 + t_2 + t_3) - (t_4 + t_7) \\ &\quad -2(t_{14} + t_{37}) + (t_1 + t_2 + t_3)(t_4 + t_7) + t_4 t_7 \end{aligned}$$

$$
\begin{aligned}
& +2(t_{14}t_7 + t_4t_{37}) - (t_1+t_2+t_3)t_4t_7 \\
= \;& 1 - 3\frac{z}{\Lambda} - 2\cosh\beta\frac{z}{\Lambda} + \frac{z^2}{\Lambda^2} + 2\cosh\beta\frac{z^2}{\Lambda^2} + \frac{z^3}{\Lambda^3}\,. \qquad \text{(N.18)}
\end{aligned}
$$

The Markov graphs are not unique, and extra terms cancel in the evaluation. Flow conservation (16.11) and symmetry factorization (**??**) yield

$$
0 = 1/\zeta(0,1) = \left(1-\frac{1}{\Lambda}\right)\left(1-\frac{4}{\Lambda}-\frac{1}{\Lambda^2}\right)
$$

The second factor indeed vanishes for the given value of Λ. The diffusion constant follows from

$$
\langle n\rangle_\zeta = -\left(1-\frac{1}{\Lambda}\right)\left(-\frac{4}{\Lambda}-\frac{2}{\Lambda^2}\right)\,, \quad \langle \hat{n}^2\rangle_\zeta = \frac{2}{\Lambda}-\frac{4}{\Lambda^2} = \frac{2}{\Lambda}\left(1-\frac{2}{\Lambda}\right)
$$

$$
D = \frac{1}{2\sqrt{5}}
$$

Fig. N.1(c) $\Lambda = (\sqrt{17}+5)/2$ with the corresponding Markov graph given in fig. N.3(c).

$$
\begin{aligned}
1/\zeta &= 1-(t_1+t_2+t_3)-2(t_4+t_7) \\
&\quad -2(t_{14}+t_{37}) + 2(t_1+t_2+t_3)(t_4+t_7) + 4t_4t_7 \\
&\quad +4(t_{14}t_7+t_4t_{37}) - 4(t_1+t_2+t_3)t_4t_7 \\
&= 1 - 3\frac{z}{\Lambda} - 4\cosh\beta\frac{z}{\Lambda} + 4\frac{z^2}{\Lambda^2} + 8\cosh\beta\frac{z^2}{\Lambda^2} - 4\frac{z^3}{\Lambda^3}\,, \qquad \text{(N.19)}
\end{aligned}
$$

Flow conservation (16.11) and symmetry factorization (**??**) yield

$$
0 = 1/\zeta(0,1) = \left(1-\frac{2}{\Lambda}\right)\left(1-\frac{5}{\Lambda}+\frac{2}{\Lambda^2}\right)
$$

The second factor indeed vanishes for the given value of Λ. The diffusion constant follows from

$$
\langle n\rangle_\zeta = -\left(1-\frac{2}{\Lambda}\right)\left(-\frac{5}{\Lambda}-\frac{4}{\Lambda^2}\right)\,, \quad \langle \hat{n}^2\rangle_\zeta = \frac{4}{\Lambda}-\frac{8}{\Lambda^2} = \frac{2}{\Lambda}\left(1-\frac{2}{\Lambda}\right)
$$

$$
D = \frac{2}{\sqrt{17}}
$$

It is not clear why the final answers tend to be so simple. Numerically this is the maximal diffusion constant. Is it? Why?

Fig. N.1(d) $\Lambda = (\sqrt{33}+5)/2$ with the corresponding Markov graph given in fig. N.3(d).

$$
\begin{aligned}
1/\zeta &= 1-(t_1+t_2+t_3)-2(t_4+t_7) \\
&\quad -4(t_{14}+t_{37}) + 2(t_1+t_2+t_3)(t_4+t_7) + 4t_4t_7 \\
&\quad +8(t_{14}t_7+t_4t_{37}) - 4(t_1+t_2+t_3)t_4t_7 \\
&= 1 - 3\frac{z}{\Lambda} - 4\cosh\beta\frac{z}{\Lambda} + 4\frac{z^2}{\Lambda^2} + 4\cosh\beta\frac{z^2}{\Lambda^2} + 4\frac{z^3}{\Lambda^3}\,. \qquad \text{(N.20)}
\end{aligned}
$$

Flow conservation (16.11) and symmetry factorization (**??**) yield

$$
0 = 1/\zeta(0,1) = \left(1-\frac{2}{\Lambda}\right)\left(1-\frac{5}{\Lambda}-\frac{2}{\Lambda^2}\right)
$$

The second factor indeed vanishes for the given value of Λ. The diffusion constant follows from

$$
\langle n\rangle_\zeta = -\left(1-\frac{2}{\Lambda}\right)\left(-\frac{5}{\Lambda}-\frac{4}{\Lambda^2}\right)\,, \quad \langle \hat{n}^2\rangle_\zeta = \frac{4}{\Lambda}-\frac{4}{\Lambda^2} = \frac{4}{\Lambda}\left(1-\frac{1}{\Lambda}\right)
$$

$$
D = \frac{1}{8} + \frac{5}{88}\sqrt{33}
$$

Fig. N.1(e) $\Lambda = 2\sqrt{2}+3$ with the corresponding Markov graph given in fig. N.3(e).

$$
\begin{aligned}
1/\zeta &= 1-3t_1-2(t_4+t_7) \\
&\quad -4(t_{14}+t_{37}) - t_{47} + 6t_1(t_4+t_7) + 4t_4t_7 \\
&\quad -4(t_{147}+t_{347}) + 8(t_{14}t_7+t_4t_{37}) + 3t_1t_{47} - 12t_1t_4t_7 \\
&= 1 - 3\frac{z}{\Lambda} - 4\cosh\beta\frac{z}{\Lambda} + 3\frac{z^2}{\Lambda^2} + 4\cosh\beta\frac{z^2}{\Lambda^2} - \frac{z^3}{\Lambda^3}\,. \qquad \text{(N.21)}
\end{aligned}
$$

Markov graphs are not unique, and the extra terms cancel in the evaluation. Flow conservation (16.11) and symmetry factorization (**??**) yield

$$
0 = 1/\zeta(0,1) = \left(1-\frac{1}{\Lambda}\right)\left(1-\frac{6}{\Lambda}+\frac{1}{\Lambda^2}\right)
$$

The second factor indeed vanishes for the given value of Λ. The diffusion constant follows from

$$\langle n\rangle_\zeta = -\left(1-\frac{1}{\Lambda}\right)\left(-\frac{6}{\Lambda}+\frac{2}{\Lambda^2}\right)\,, \quad \left\langle \hat{n}^2\right\rangle_\zeta = \frac{4}{\Lambda}-\frac{4}{\Lambda^2} = \frac{4}{\Lambda}\left(1-\frac{1}{\Lambda}\right)$$

$$D = \frac{1}{2\sqrt{2}}$$

Fig. N.1(f) $\Lambda = (\sqrt{33}+7)/2$ with the corresponding Markov graph given in fig. N.3(f). Go through the whole construction, check that:

$$\begin{aligned} 1/\zeta &= 1-3t_1-2(t_4+t_7) \\ &\quad -4(t_{14}+t_{37}+t_{47})+6t_1(t_4+t_7)+4t_4t_7 \\ &\quad -8(t_{417}+t_{437})+8(t_{14}t_7+t_4t_{37})+12t_1t_{47}-12t_1t_4t_7 \\ &= 1-3\frac{z}{\Lambda}-4\cosh\beta\frac{z}{\Lambda}+4\cosh\beta\frac{z^2}{\Lambda^2}\,. \end{aligned} \tag{N.22}$$

Flow conservation (16.11) and symmetry factorization (**??**) yield

$$0 = 1/\zeta(0,1) = \left(1-\frac{7}{\Lambda}+\frac{4}{\Lambda^2}\right)$$

The diffusion constant follows from

$$\langle n\rangle_\zeta = \left(-\frac{7}{\Lambda}+\frac{8}{\Lambda^2}\right) = \left(\frac{7}{\Lambda}-2\right)\,, \quad \left\langle \hat{n}^2\right\rangle_\zeta = \frac{4}{\Lambda}-\frac{4}{\Lambda^2} = -\frac{3}{\Lambda}+1$$

$$D = \frac{1}{4}+\frac{1}{4\sqrt{33}}$$

Fig. N.1(g) $\Lambda = 7$. The Markov graph corresponding to fig. N.1(g) is fig. N.3(g).

$$\begin{aligned} 1/\zeta &= 1-3t_1-2(t_4+t_7) \\ &= 1-3\frac{z}{\Lambda}-4\cosh\beta\frac{z}{\Lambda} \end{aligned} \tag{N.23}$$

The diffusion constant follows from (N.8)

$$\langle n\rangle_\zeta = \frac{7}{\Lambda}\,, \quad \left\langle \hat{n}^2\right\rangle_\zeta = \frac{4}{\Lambda}$$

$$D = \frac{2}{7}$$