HOW WELL CAN ONE RESOLVE THE STATE SPACE OF A CHAOTIC MAP?

A Thesis
Presented to
The Academic Faculty

by

Domenico Lippolis

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Physics

Georgia Institute of Technology
May 2010

HOW WELL CAN ONE RESOLVE THE STATE SPACE OF A CHAOTIC MAP?

Approved by:

Professor Predrag Cvitanović, Adviser
School of Physics
Georgia Institute of Technology

Professor Jean Bellissard
School of Physics
Georgia Institute of Technology

Professor Kurt Wiesenfeld
School of Physics
Georgia Institute of Technology

Professor Hao Min Zhou
School of Mathematics
Georgia Institute of Technology

Professor Howard Weiss
School of Mathematics
Georgia Institute of Technology

Date Approved: March 15 2010

ACKNOWLEDGEMENTS

I would like to thank Alexander Grigo and William H. Mather for useful discussions. I would also like to thank Carl P. Dettmann for checking one of my calculations, and Vattay Gabor for having examined the problem and given some opinions on the possible approaches to follow. Finally I would like to thank Rytis Paškauskas and Angelo Bongiorno for some help with the graphics.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

All physical systems are affected by some noise that limits the resolution that can be attained in partitioning their state space. For chaotic, locally hyperbolic flows, this resolution depends on the interplay of the local stretching/contraction and the smearing due to noise. My goal is to determine the 'finest attainable' partition for a given hyperbolic dynamical system and a given weak additive white noise. That is achieved by computing the local eigenfunctions of the Fokker-Planck evolution operator in linearized neighborhoods of the periodic orbits of the corresponding deterministic system, and using overlaps of their widths as the criterion for an optimal partition. The Fokker-Planck evolution is then represented by a finite transition graph, whose spectral determinant yields time averages of dynamical observables. The method applies in principle to both continuous- and discrete-time dynamical systems. Numerical tests of such optimal partitions on unimodal maps support my hypothesis.

CHAPTER I

INTRODUCTION

Long-term sensitivity to small perturbations is the best-known signature of chaos: a tiny change in the initial conditions of a dynamical system can result in dramatic changes in the long run [6, 31, 39, 34, 35].

Because of that, in order to fully characterize the phase space ('state space') of a set of equations, one should determine all its (countably infinite) solutions one by one, which is impractical, to say the least. One way to efficiently chart the space is to locate a few regions of the phase space, which differ by their short-term dynamics, and label each one with a symbol [26, 9]. These regions are in turn refined into smaller regions [13], in order to characterize longer trajectories and make predictions on their whereabouts. Not surprisingly, there is no end to this process, and the resolution of the symbolic space turns out to be infinite, just like that of the phase space. Still, the symbolic dynamics helps find or characterize precious invariants of the dynamics, such as periodic orbits, invariant manifolds or tori [15]. These in turn can be used to estimate long-term averages of observables [43], such as correlations, escape rates, diffusion coefficients, and Lyapunov exponents, a measure of the sensitivity of the system to initial conditions.

In reality, any physical system suffers background noise, any numerical prediction suffers computational roundoff errors, any set of equations models nature up to a given accuracy, since degrees of freedom are always neglected. As a result, no single chaotic trajectory can be predicted in the long term [32], and there must be a limit to the resolution of the state space, as trajectories can now cross due to noise.

In the present work, I propose an algorithm to determine the finest possible or *optimal* partition of the chaotic state space of a one-dimensional discrete-time dynamical system (map), with uncorrelated, Gaussian-distributed, background noise.

I find it convenient for my analysis to study the evolution of densities of orbits, rather

than of single, noisy trajectories. The Fokker-Planck equation is derived in Sect. 1.2 to that aim. Next, the Fokker-Planck evolution operator, a discrete-time analog of the homonymous equation, is introduced in Sect. 1.3. It will be used everywhere in what follows. Chapter 2 is focused on the evolution of densities surrounding periodic oribts of the deterministic system. The eigenspectra of the Fokker-Planck operator and of its adjoint are computed locally, in the neighborhood of periodic orbits of the map. In particular, the eigenfunctions found are used to partition the state space in a noisy environment, in Chapter 3. The optimal partition method is formulated and applied to a one-dimensional repeller. The next problem is how to test the optimal partition hypothesis. The idea is to use the partition obtained to calculate long-time averages (escape rate, diffusivity, etc.) from the spectrum of the Fokker-Planck evolution operator, by reducing the latter to a finite matrix (Sect. 4.1), whose entries are the transition rates between the intervals of the partition. Noise makes the borders of such intervals fuzzy, and errors can dangerously propagate from a much too rough approximation of the Fokker-Planck matrix, as one attempts to compute its spectrum. I bypassed that problem by evaluating the leading eigenvalue of the evolution operator by means of periodic orbit expansions, illustrated in the rest of Chapter 4 for a general setting, and narrowed down to a finite-dimensional Fokker-Planck operator in Sect. 5.1. Once all the tools have been developed, the optimal partition hypothesis is finally tested in Sect. 5.2, by computing the escape rate and the Lyapunov exponent of the repeller previously introduced in Sect. 3.2.

The local approximation of the Fokker-Planck operator has been formulated in a linearized neighborhood, yet it cannot be expected to work in any strongly nonlinear regime. For that reason, the whole construction, leading to the optimal partition hypothesis and its validity tests, is adapted in Chapter 6 to models with stronger nonlinearities, and exemplified on a unimodal map.

Summary and a short-term, down-to-the-point outlook in Chapter 7. multiply defined.

1.1 History and motivation

Crutchfield and Packard [8] are the first, in 1983, to raise the problem of an 'optimal partition', when dealing with a chaotic system in the presence of noise. More precisely, they

define the most efficient symbolic encoding of the dynamics, as the sequence of symbols that maximizes the metric entropy of the system. Once the maximum is found, they refine the partition until the entropy converges to some value. Their method is purely statistical, and works under the assumption that a generating partition already exists for the corresponding deterministic system. They also introduce the *attainable information*, as a limiting value for the probability to produce a certain sequence of symbols from the ensemble of all possible initial conditions. Once such limit is reached, no further refinements make sense.

Tang and co-workers [46] do something similar in the realm of chaotic data analysis, with the dynamics being unknown to start with. Their method is based on maximizing Shannon entropy and at the same time minimizing an error function with respect to the partition chosen. The same idea is used by Lehrman *et al.* [33] to encode chaotic signals in higher dimensions, where they also detect correlations between different signals by computing their *conditional entropy*. Daw, Finney and Tracy wrote a review of symbolic analysis of experimental data up to 2001 [17].

More recently, Kennel and Buhl [29, 5] proposed a method to estimate partitions from time-series data, which minimizes an energy-like functional with respect to the encoding chosen, and maximizes the correlation between distances in the state space and in the symbolic space. Once again, there is no regard for the interplay between noise and deterministic dynamics, as the latter is taken as unknown.

A different, indirect approach to the problem of the optimal resolution is that of the refinement of a transition matrix: given a chaotic, discrete-time dynamical system, the state space is partitioned, and the probabilities of points mapping between regions are estimated, so as to obtain a matrix, whose eigenvalues and eigenfunctions are then used to evaluate averages of observables defined on the chaotic set. This idea was first advanced by Ulam in 1960 [48], for a dynamical system with no noise, when he proposed a simple uniform-mesh grid as partition. Later on, Nicolis [36], and Rechester and White [41, 40] discussed different ways of constructing partitions for chaotic maps in one and two dimensions, which would make Ulam's method more efficient.

Dellnitz and Junge [19], Guder and Kreuzer [27], Froyland [22], and Keane *et al.* [28]

all come up with different algorithms of non-uniform refinement for such grid methods, summarized in a monograph by Froyland [23], who also treats their extension to random dynamical systems. In all cases, the ultimate threshold for every refinement seems to be determined by the convergence of the spectrum of the transition matrix.

Finally, Bollt *et al.* [4] show that the transition matrix of a stochastic dynamical system is finite under certain restrictions, and compute its entries using a new set of basis functions. They seem to get around the problem of determining the optimal resolution of their partition by choosing the size of the matrix *a priori*.

The objective of all the literature cited so far is either to construct a partition that encodes enough information from time-series data, with no attention to the dynamics and no emphasis on an intrinsic limit to its refinement, or to just optimize the diagonalization of a matrix. Devising a novel method to find this limit is, in my opinion, a problem of fundamental relevance. The first reason for that is efficiency: setting a limit to the resolution of a chaotic state space improves the computation of dynamical averages. In particular, I will show that exceeding the optimal resolution generates wrong results when attempting to estimate the escape rate of a repeller from the chaotic region. The second reason is that brute-force methods of diagonalization of an evolution operator can only be implemented in low-dimensional spaces, otherwise one can typically not afford to keep refining a grid until some cost function converges. More precisely, we need to know exactly *where* the resolution can be improved and up to how many iterations of the map. In order to achieve that knowledge, one must study the interplay of the noise with the deterministic dynamics everywhere *locally*. I will do that under the assumption that, if the noise is weak, the unstable periodic orbits of the deterministic map still constitute the skeleton of the dynamics for relatively short times, and I will use them as a starting point to look for local invariants of the noisy dynamics, *the eigenfunctions of the (adjoint) evolution operator*. The latter are then at the basis of an algorithm for the refinement of the state space, where the limit is set by a critical overlapping of the supports of the eigenfunctions. The advantage of a method based on periodic orbits is that, once again, it can be straightforwardly extended to higher dimensions, as discussed in the last chapter of this thesis. But the quest for

invariants is also a necessity: as I explain in Chapter 2, a noisy trajectory spreads in the state space by an amount, which depends on where it started, unless it is periodic. A bit like a wavepacket. So that, given a density of points in a chaotic state space, I would not know what iteration of what noisy orbit it would be, unless it has some periodicity with respect to the evolution operator.

In the next two sections I reproduce the derivation of the Fokker-Planck equation and of its path-integral solution, which can be found, for example, in the book by Risken [42].

1.2 The Fokker-Planck equation

Consider a Langevin-type set of equations [49]

$$\frac{dx}{dt} = f(x) + \xi(t)\,, \tag{1}$$

where the additive noise $\xi(t)$ is a Gaussian random variable such that

$$\langle \xi_j(t) \rangle = 0\,, \qquad \left\langle \xi_i(t)\xi_j(t') \right\rangle = 2D\,\delta(t-t')\,\delta_{ij}\,. \tag{2}$$

We seek an equation that express the evolution of a density of trajectories $\rho(x,t)$. Since (1) is a first-order ordinary differential equation, and the noise is uncorrelated ('white'), we have that (Appendix A)

$$\rho(x, t+\Delta t) = \int \rho(x',t) p(x, \Delta t | x') dx' \tag{3}$$

the evolved density $\rho(x, t+\Delta t)$ only depends on what it was a time Δt before and nothing else before time t, a property that defines *Markov* processes. We now seek a more explicit expression for (3). The following derivation is based on the awareness that Δt's are usually made to be sent to zero at some point, and that the noise is always meant to be weak with respect to the effects of the advection (I called it $f(x)$ in (1)). All that suggests the variation $\Delta x = x - x'$ must be relatively small in an interval Δt, and thus I change the variable in the integral to Δx in the perspective of an imminent Taylor expansion:

$$\rho(x, t+\Delta t) = \int \rho(x-\Delta x, t) p(\Delta x, x-\Delta x, \Delta t) d\Delta x \tag{4}$$

where $p(\Delta x, x-\Delta x, \Delta t)$ is the conditional probability of a change Δx, the initial variable been fixed at $x-\Delta x$. I now expand the product $\rho(x-\Delta x,t)p(\Delta x, x-\Delta x, \Delta t)$ in the previous integral, around $\Delta x = 0$, as promised:

$$\begin{aligned}\rho(x,t+\Delta t) &= \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\int \Delta x^n \partial_x^n\left(\rho(x,t)p(\Delta x,x,\Delta t)\right)d\Delta x = \\ &\sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\partial_x^n\left(\langle\Delta x^n\rangle\,\rho(x,t)\right) \end{aligned} \tag{5}$$

this is known as Kramers-Moyal expansion [42]. In order to go any further, we need to evaluate the moments $\langle\Delta x^n\rangle$. First moment, first:

$$\begin{aligned}\langle\Delta x\rangle &= \left\langle\frac{\Delta x}{\Delta t}\Delta t\right\rangle = \\ &\int f(x)\Delta t p(\Delta x,x,\Delta t)d\Delta x + \langle\xi\rangle = f(x)\Delta t\end{aligned} \tag{6}$$

where I used the Langevin equation (1) and the fact that the ensemble average of ξ is zero, from (2). The second moment reads

$$\begin{aligned}\left\langle\Delta x^2\right\rangle &= f(x)\Delta t^2 + \langle\xi\rangle\,\Delta t + \int_t^{t+\Delta t}dt'\int_t^{t+\Delta t}dt''\left\langle\xi(t')\xi(t'')\right\rangle = \\ &f(x)\Delta t^2 + \int_t^{t+\Delta t}dt' 2D = 2D\Delta t + O(\Delta t^2)\end{aligned} \tag{7}$$

where I used $\left\langle\xi(t'')\xi(t')\right\rangle = 2D\delta(t'-t'')$ from (2). One can realize, by just looking at the first identity in equation (7), that all higher moments will be at least proportional to Δt^2. Thus, I will stop the Kramers-Moyal expansion (5) at the second term and write

$$\rho(x,t+\Delta t) = \rho(x,t) - \partial_x\left(f(x)\rho(x,t)\right)\Delta t + D\partial_{xx}\rho(x,t)\Delta t \tag{8}$$

now move $\rho(x,t)$ to the left-hand side of (8), divide both sides by Δt and then send the latter to zero, so as to get the desired *Fokker-Planck* equation

$$\partial_t\rho(x,t) = -\partial_x\left(f(x)\rho(x,t)\right) + D\partial_{xx}\rho(x,t) \tag{9}$$

1.3 The Fokker-Planck operator

This thesis only deals with discrete-time dynamical systems, such as the map

$$x_{n+1} = f(x_n) + \xi_n \tag{10}$$

with the same assumptions on the noise that were previously made with the Langevin equation, only in discrete time. That means I also need a discrete-time version of the Fokker-Planck equation I just derived. Equation (8) is still in discrete time, and if I plug its left-hand side into the expression (3) for the evolution of a density in an interval Δt (our starting point in the derivation of (9)), I get that the conditional probability $p(x, \Delta t|x')$ must be written as

$$\begin{aligned} p(x, \Delta t|x') &= [1 - \Delta t \partial_x f(x) + D\Delta t \partial_{xx}]\, \delta(x - x') = \\ & \qquad \exp\left[-\Delta t \partial_x f(x) + D\Delta t \partial_{xx}\right] \delta(x - x') \end{aligned} \tag{11}$$

The previous is true since I have been throwing away all terms of $O(\Delta t^2)$ and higher. Now write the δ-function as a Fourier integral and move the exponential into it:

$$\begin{aligned} p(x, \Delta t|x') &= \exp\left[-\Delta t \partial_x f(x) + D\Delta t \partial_{xx}\right] \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iu(x'-x)} du = \\ & \qquad \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left[-iuf(x)\Delta t - u^2 D\Delta t + iu(x' - x)\right] du = \\ & \qquad \frac{1}{\sqrt{4\pi D\Delta t}} \exp\left(-\frac{[x' - x - f(x)\Delta t]^2}{4D\Delta t}\right) \end{aligned} \tag{12}$$

Now fit the previous expression to the notation of our map (10), that is $x - x' \to y \equiv x_{n+1}$, and $\Delta t = 1$, and the result is the Fokker-Planck evolution operator, as it will be used from now on:

$$\begin{aligned} \rho_{n+1}(y) &= (\mathcal{L} \circ \rho_n)(y) \\ &= \int [dx] \exp\left\{-\frac{1}{4D}[y - f(x)]^2\right\} \rho_n(x)\,, \end{aligned} \tag{13}$$

where $[dx] = (4\pi D)^{-1/2}$. In the noiseless limit, the Fokker-Planck operator reduces to its deterministic counterpart, the Perron-Frobenius operator [12]:

$$\lim_{D \to 0} \mathcal{L}\, \rho(y) = \mathcal{L}_{det} \rho(y) = \int [dx]\, \delta(y - f(x)) \rho(x)\,. \tag{14}$$

CHAPTER II

THE FOKKER-PLANCK OPERATOR, LOCALLY

In this chapter, I obtain the eigenspectra of the Fokker-Planck operator and of its adjoint, in the vicinity of any periodic point of a one dimensional map with white noise. It is the first new result of this thesis, and it will be used later on to estimate the optimal partition of the state space of the map.

Let me first define a periodic orbit or *cycle* of length (or period) n_p of a map $x' = f(x)$ as a set of points $x_1, x_2, ..., x_n$ such that $x_i = f^{n_p}(x_i)$. A periodic point of length one is called a *fixed point*. A periodic orbit is said to be *stable* if the matrix

$$J^p(x_i) = \left[\frac{df^{n_p}(x)}{dx}\right]_{x=x_i} \tag{15}$$

called *Jacobian*, is is such that the magnitude of all its eigenvlaues is less than one ($|\Lambda_p| < 1$). Otherwise the orbit is said to be *unstable*.

2.1 One-dimensional fixed point

Consider the 1-dimensional linear map

$$x_{n+1} = \Lambda x_n + \xi_n \,, \qquad |\Lambda| \neq 1 \,, \tag{16}$$

with additive white noise with variance $2D$:

$$\langle \xi_n \rangle \;=\; 0 \,, \qquad \langle \xi_n \xi_m \rangle \;=\; 2D\, \delta_{nm} \,. \tag{17}$$

This is the discrete-time version of the Ornstein-Uhlenbeck process [47]. The density $\rho(x)$ of trajectories evolves by the action of the Fokker-Planck operator (13):

$$\mathcal{L}\rho(x) = \int [dy]\, e^{-\frac{(x-\Lambda y)^2}{4D}} \rho(y) \,. \tag{18}$$

$|\Lambda| < 1$ **case:** In each iteration the map contracts the noisy trajectory points by factor Λ toward the $x = 0$ fixed point, while the noise smears them out with variance $2D$. The

normalized eigenfunctions $\tilde{\rho}_0$, $\tilde{\rho}_1$,$\cdots$ of (18) are [18, 25][1]

$$\begin{aligned} dx\,\tilde{\rho}_k(x) &= [dx]\,H_k(\mu x)\,e^{-x^2/2\sigma_0^2} \\ \mu^{-2} &= 2\sigma_0^2\,, \qquad \sigma_0^2 = \frac{2D}{1-\Lambda^2}\,, \end{aligned} \tag{19}$$

where $H_k(x)$ the kth Hermite polynomial, and $[dx] = dx/(4\pi D)^{1/2}$. Hermite polynomials pop up here [1, 45] as the linear fixed point of dynamical systems is the imaginary time version of the harmonic oscillator of QM. Note that the eigenvalues Λ^k are independent of the noise strength, so they are the same as for the $D \to 0$ deterministic case [2]. The unit-eigenvalue eigenfunction $\rho_0\,dx = [dx]\,\exp(-x^2/2\sigma_0^2)$ is the natural measure [12] for the Fokker-Planck operator, its variance $\sigma_0^2 = 2D/(1-\Lambda^2)$ a balance of the fixed-point contraction Λ and diffusive spread D.

$|\Lambda| = 1$ **case** is the marginal, pure diffusion case, and the behavior is not exponential, but power-law. If the map is nonlinear, one needs to go to the first nonlinear order to reestablish the control [25]. We will deal with this regime in Chapter 6.

$|\Lambda| > 1$ **case:**

$$\begin{aligned} dx\,\rho_k(x) &= [dx]\,H_k(\alpha x) \\ \alpha^{-2} &= 2\sigma_0^2\,, \qquad \sigma_0^2 = \frac{2D}{\Lambda^2-1}\,, \end{aligned} \tag{20}$$

with eigenvalues $1/|\Lambda|\Lambda^k$.

The eigenfunctions (19) and (20) are respectively the left and the right eigenfunctions of the Fokker-Planck operator with $|\Lambda| > 1$ (or the right and the left eigenfunctions of the same operator with $|\Lambda| < 1$). They are orthonormal:

$$\int [dx]\,\tilde{\rho}_k(x)\rho_j(x) = \delta_{kj}\,. \tag{21}$$

In the deterministic, noiseless limit, (18) reduces to the Perron-Frobenius operator:

$$\lim_{D\to 0} \mathcal{L}\,\rho(x) = \mathcal{L}_{det}\rho(x) = \int [dy]\,\delta(x-\Lambda y)\rho(y)\,. \tag{22}$$

[1] For a full derivation in continuous time, see Appendix B.

In the $|\Lambda| > 1$ expanding case the noiseless $D \to 0$ limit eigenfunctions (85) tend to the deterministic eigenfunctions [25, 2]

$$\rho_k(x) \to \frac{x^k}{k!} \tag{23}$$

while the contracting eigenfunctions (19) tend to distributions [25, 2]

$$\rho_k(x) \to (-1)^k \delta^{(k)}(x)\,. \tag{24}$$

2.2 Local eigenfunctions of a one-dimensional map

I now adapt the discrete Ornstein-Uhlenbeck fixed-point calculation of Sect. 2.1 to determination of the complete Fokker-Planck operator eigenspectrum in the neighborhood of every periodic point of any sufficiently smooth one-dimensional map

$$x_{n+1} = f(x_n) + \xi_n \tag{25}$$

by Taylor-expanding around an orbit point x_a.

I first introduce local coordinate systems z_a centered on the orbit points x_a, together with a notation for the map (25), its derivative, and, by the chain rule, the derivative of the kth iterate f^k evaluated at the point x_a,

$$\begin{aligned} x &= x_a + z_a\,, \quad f_a(z_a) = f(x_a + z_a) \\ f_a^{'} &= f'(x_a)\,, \quad f_a^{k\prime} = f_{a+k-1}^{'} \cdots f_{a+1}^{'} f_a^{'}\,, \quad k \geq 2\,. \end{aligned} \tag{26}$$

Here a is the label of point x_a, and the label $a{+}1$ is a shorthand for the next point b on the orbit of x_a, $x_b = x_{a+1} = f(x_a)$. For example, a period-3 periodic point might have label $a = 001$, and by $x_{010} = f(x_{001})$ the next point label is $b = 010$.

If the noise is weak, we can approximate (to leading order in D) the Fokker-Planck operator(13), which we now write in the local coordinates,

$$\mathcal{L}_a \circ \rho_n(x_{a+1} + z_{a+1}) = \int dz_a \mathcal{L}_a(z_{a+1}, z_a) \rho_n(x_a + z_a), \tag{27}$$

by linearization centered on x_a, the ath point along the orbit,

$$\mathcal{L}_a(z_{a+1}, z_a) \;=\; (4\pi D)^{-1/2}\, e^{-\frac{(z_{a+1} - f_a^{'} z_a)^2}{4D}}\,. \tag{28}$$

A Gaussian density

$$\rho_0(x_a + z_a) = c_a e^{-z_a^2/2\sigma_a^2} \,. \tag{29}$$

is transformed by (28) in a linearized neighborhood of the orbit, into another Gaussian

$$\rho_0(x_{a+1} + z_{a+1}) = c_{a+1} e^{-z_{a+1}^2/2(2D+(f_a'\sigma_a)^2)} \,, \tag{30}$$

centered around the next point of the periodic orbit, $x_{a+1} = f(x_a)$. The variances of the original and the transformed Gaussians are related by the following recursion relation:

$$\sigma_{a+1}^2 = 2D + (f_a^{'}\sigma_a)^2 \,. \tag{31}$$

After n steps,

$$\sigma_{a+n}^2 = \sigma_a^2 (f_a^{n'})^2 + 2D(1 + (f'_{a+n-1})^2 + (f^{2'}_{a+n-2})^2 + ... + (f^{n-1'}_{a+1})^2) \tag{32}$$

If the points $x_a, ..., x_{a+n-1}$, form a stable periodic orbit of length n, then (29) is a local eigenfunction of $\mathcal{L}^n{}_a$, when

$$\sigma_a^2 = \frac{2D}{1-\Lambda^2} \left(1 + \sum_{i=1}^{i=n-1} (f_{a+i}^{n-i'})^2 \right) \,, \qquad \Lambda = f_a^{n'} \,. \tag{33}$$

Expression (33) can also be derived in continuous time (see Appendix C), where $\Lambda \to J$ and $\sum_{i=1}^{i=n-1} (f_{a+i}^{n-i'})^2 \to \int dt J^2$, the jacobian J as defined in (15) for maps.

The rest of the spectrum of a periodic point belonging to a stable cycle of period n is

$$\tilde{\rho}_k(z_a) = \frac{1}{\sqrt{4\pi D}} H_k(\mu z_a) e^{-z_a^2/2\sigma_a^2} \,, \qquad \mu^{-2} = 2\sigma_a^2 \,. \tag{34}$$

with eigenvalues Λ^k. On the other hand, the local eigenfunctions of a periodic point of an unstable cycle of period n are

$$\tilde{\rho}_k(z_a) = \frac{1}{\sqrt{4\pi D}} H_k(\alpha z_a) \,, \qquad \alpha^{-2} = -2\sigma_a^2 \,. \tag{35}$$

Eigenvalues are in this case $\frac{1}{|\Lambda|\Lambda^k}$.

2.3 Adjoint of the Fokker-Planck operator

In one dimension, the adjoint $\mathcal{L}^\dagger$ of the Fokker-Planck operator satisfies, for any two densities $\rho(x)$ and $h(x)$:

$$\begin{aligned} \langle h|\mathcal{L}|\rho\rangle &= \int_{-\infty}^{\infty} h(x)dx \int_{-\infty}^{\infty} e^{-\frac{(x-f(y))^2}{4D}} \rho(y)[dy] \\ \left\langle \rho|\mathcal{L}^\dagger|h \right\rangle &= \int_{-\infty}^{\infty} e^{-\frac{(y-f(x))^2}{4D}} \rho(x)[dx] \int_{-\infty}^{\infty} h(y)dy \end{aligned} \tag{36}$$

so that

$$\mathcal{L}^\dagger \rho(x) = \int_{-\infty}^{\infty} e^{-\frac{(f(x)-y)^2}{4D}} \rho(y)[dy] \tag{37}$$

$\mathcal{L}$ carries a density $\rho(x)$, supported on some interval I, forward in time to a function supported on a subset of $f(I)$. The adjoint operator $\mathcal{L}^\dagger$ transports the density $\rho(x)$, supported on I, backward in time to a function supported on $f^{-1}(I)$ (see Appendix E for details).

2.3.1 Eigenfunctions of $\mathcal{L}^\dagger$ for a fixed point

Suppose $f(x) = \Lambda x$, with $|\Lambda| > 1$, then the previous reads

$$\mathcal{L}^\dagger \rho(x) = \int_{-\infty}^{\infty} e^{-\frac{(\Lambda x - y)^2}{4D}} \rho(y) dy \tag{38}$$

Eigenfunctions are

$$\begin{aligned} \rho_n(x) &= \frac{1}{\sqrt{4\pi D}} H_n(\mu x) e^{-x^2/2\sigma_0^2} \\ \mu^{-2} &= 2\sigma_0^2 \,, \qquad \sigma_0^2 = \frac{2D}{\Lambda^2 - 1} \end{aligned} \tag{39}$$

and eigenvalues $\frac{1}{|\Lambda|\Lambda^n}$.

On the other hand, if $|\Lambda| < 1$, the spectrum is given by

$$\rho_n(x) = \frac{1}{\sqrt{4\pi D}} H_n(\alpha x) \,, \qquad \alpha^{-2} = -2\sigma_0^2 \tag{40}$$

with eigenvalues Λ^n.

2.3.2 Evolution of a Gaussian and local approximation

Now take a Gaussian density

$$\rho_0(x_a + z_a) = c_a e^{-z_a^2/2\sigma_a^2} \tag{41}$$

The adjoint operator (37)

$$\begin{aligned} \mathcal{L}^\dagger \rho(x) &= \int_{-\infty}^{\infty} [dy] \, c_a e^{-\frac{(f(x)-y)^2}{4D}} e^{-(x-x_a)^2/2\sigma_a^2} \\ &= c_{a-1} e^{-\frac{(f(x)-x_a)^2}{2(\sigma_a^2+2D)}} \,. \end{aligned} \tag{42}$$

Given a nonlinear $f(x)$, I now approximate the density in the neighborhood of $f^{-1}(x_a)$ to linear order,

$$\begin{aligned}\mathcal{L}^\dagger \rho(x) &= c_{a-1} e^{-\frac{(f(x)-x_a)^2}{2(\sigma_a^2+2D)}} \simeq c_{a-1} e^{-\frac{(f'(f^{-1}(x_a))(x-f^{-1}(x_a)))^2}{2(\sigma_a^2+2D)}} \\ &= c_{a-1} e^{-(f'_{a-1} z_{a-1})^2/2(\sigma_a^2+2D)} \end{aligned} \tag{43}$$

using the notation introduced in Sect. 2.2. I can again obtain a recursion relation for the evolution of σ_a^2:

$$(f'_{a-1}\sigma_{a-1})^2 = \sigma_a^2 + 2D \tag{44}$$

which can be extended to the nth preimage of the point x_a :

$$(f^{n'}_{a-n}\sigma_{a-n})^2 = \sigma_a^2 + 2D(1 + (f'_{a-1})^2 + \cdots + (f^{n-1'}_{a-n+1})^2) \tag{45}$$

The initial density (41) is the leading eigenfunction of $\mathcal{L}^\dagger$ in the neighborhood of the unstable cycle, if:

$$\sigma_a^2 = \frac{2D}{\Lambda_p^2 - 1}\left(1 + \sum_{i=1}^{n-1} (f^{i'}_{a-i})^2\right) , \tag{46}$$

where $\Lambda_p = f_a^{n'_p}$. The rest of the spectrum is identical to that of a fixed point, with σ_a^2 as above.

CHAPTER III

OPTIMAL RESOLUTION OF A ONE-DIMENSIONAL MAP

Well-established techniques for partitioning and encoding the state space are reviewed in Sect. 3.1. Then, in Sect. 3.2, I illustrate my algorithm to determine the optimal partition on the example of a one-dimensional repeller in the presence of noise. This is the central result of my work.

3.1 Partitions, symbolic dynamics, transition graphs

I have been talking so far about regular dynamics, where the equations of motion (either deterministic, or Langevin, or Fokker-Planck) can be integrated, numerically or sometimes even analytically. My work focuses instead on chaotic systems, whose equations exhibit strong sensitivity to initial conditions, meaning one cannot integrate them for a long time due to roundoff errors. I will now introduce a well-established technique of characterizing orbits in a chaotic system, which consists of dividing the state space in regions (*partitioning*), assigning a symbol to each region, and *encoding* any trajectory with a sequence of symbols representing the regions of the state space it visits [1].

Everything is best explained with an example: consider the chaotic repeller

$$x_{n+1} = \Lambda_0 x_n(1 - x_n)(1 - bx_n)\,, \quad \Lambda_0 = 6, b = 0.6 \tag{47}$$

defined on the unit interval. In principle one can partition the unit interval arbitrarily, but if we want it to be useful, we need to do it cleverly, understanding in which points the dynamics changes qualitatively, and define those as borders between different regions. One can already guess, just by looking at Figure 1, that the only point at which anything in the dynamics can change is the maximum of the map, also called critical point x_c. More precisely, the map(47) *stretches* any interval on the left of the critical point to a larger

[1] Partitions are discussed comprehensively in `ChaosBook.org` [11], or in the monographs by Lefranc and Gilmore [26], and Kitchens [30]

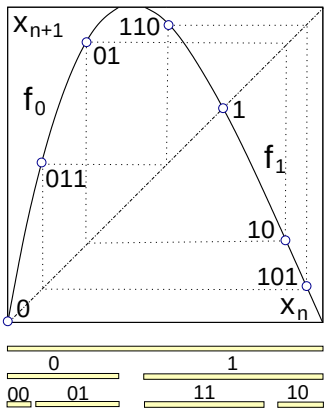


Figure 1: Partitioning the deterministic map (47) in intervals of one- and two-step memory (below). There is a periodic point in each interval (from `ChaosBook.org`).

interval, while it *contracts* any interval on the right of x_c to a smaller interval, or it *folds* it into two intervals of smaller length. This is true except for a small region (*gap*) around x_c, whose points are mapped outside the unit interval, to never go back inside, and therefore they can be left out of our analysis. Let any point on the left of x_c be encoded with a '0' and any point on its right be encoded with a '1'. Then a symbolic (in this case binary) sequence can be assigned to every orbit of (47). That is the simplest partition one can make, but we may want to know, for example, which points in the unit interval will map on which side of the critical point after two, three, m iterations. In which case the partition gets refined as in Figure 1 (1-step memory): '00' means the point is and will stay on the left of x_c after the next iteration, '01' that it will jump on the right side, and so on, to obtain the intervals $\mathcal{M}_{00}, \mathcal{M}_{01}, \mathcal{M}_{11}, \mathcal{M}_{10}$. Once again there is a set of points that escape the unit interval after two iterations, which our partition does not include. In other words, we are only concerned with the set of points that remain in the unit interval for as many memory steps as our partition accounts for, or *non-wandering set*. But how does one locate the intervals of a partition?

There are several different methods to do that. In the case of a one-dimensional repeller, the best way to obtain a refinement is to take the points that border the intervals of the partition and iterate them backwards, or find their pre-images. Take region '0' and try to find its refinements '00' and '01', for example: the point $x = 0$ is a fixed one, so that and

everything close enough to it still maps on the left of the critical point, more precisely the interval '00' begins at $x = 0$ and ends at that point x_{00} such that $f(f(x_{00})) = 1$ (looking at Figure 1 helps understand why). Therefore $x_{00} = f^{-1}(f^{-1}(1))$. The inverse map is double-valued, but it is easy to realize that one needs to take the value on the left branch twice: $x_{00} = f_0^{-1}(f_0^{-1}(1))$. One can do the exact same thing to work out the borders of the rest of the intervals of this and any n-th refinement of this partition, and find that all but two of them ($x = 0$ and $x = 1$) are n-th preimages of the two points that map into $x = 1$ (anything in between those points is in the *gap* and maps outside the unit interval and therefore we don't care about it).

Another way of partitioning is to use periodic orbits [2]. According to Devaney's definition [20], a chaotic map has a set of unstable periodic orbits that is dense in the non-wandering set. In particular, there is a periodic point in every region of a partition, which obviously bears the symbolic sequence of the interval it belongs to. In our four-interval partition, for instance, $\mathcal{M}_{01}$ and $\mathcal{M}_{10}$ both have one periodic point belonging to a periodic orbit of length two, while there is one fixed point in both regions $\mathcal{M}_{00}$ and $\mathcal{M}_{11}$, as illustrated by Figure 1. That gives us one possible way of making a partition, that is by looking for periodic orbits of the map of the same length as the memory refinement we are seeking, and draw and label a region around every periodic point found.

Once the partition has been determined, one can draw a *transition graph*, which says which regions are allowed to map where. Take our four-interval partition as an example: all points in region $\mathcal{M}_{11}$ are on the left of the critical point and will still map on the left (that is what the second '1' says), therefore they can either stay in $\mathcal{M}_{11}$ or move to $\mathcal{M}_{10}$, and so on. So in general the recipe is to discard the first digit and add, in turn, all the possible outcomes (in this case either '0' or '1') to the sequence left. Once this is done for all the regions in the partition, just draw a *node* for each interval and *links* to represent the possible transitions between nodes, as in Figure 2.

In principle, nothing prevents us from refining a partition more and more: the binary

[2] This approach was first introduced by Cvitanović and co-workers [15, 13]. More recently, the idea of using periodic orbits to partition the state space was also proposed by Davidchack *et al.* [16], as well as by Plumecoq and Lefranc [37, 38].

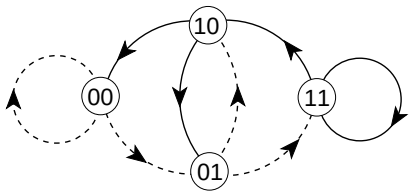


Figure 2: The transitions allowed in the two-step-memory partition of Figure 1, at a glance (from `ChaosBook.org`).

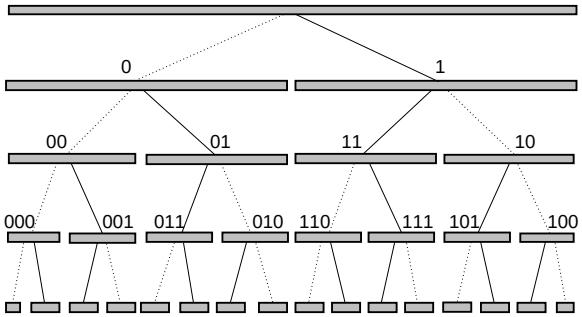


Figure 3: Evolution of successive refinements of the partition in Figure 1, each interval labeled with the symbolic sequence of all its points. (from `ChaosBook.org`).

tree in Figure 3 shows how every interval splits into two smaller regions at every new step of the refinement. In practice, a trajectory loses memory of where it has been, at some point, due to background noise, which sets a limit to the finest attainable resolution of the state space. This is the topic of the next section and the central result of my work.

3.2 The optimal resolution hypothesis

The idea behind *'the optimal resolution'* hypothesis is to partition the state space by means of periodic points, as seen in the last section. Technically, periodic points no longer exist, in the presence of noise, but, if the amplitude of the noise is relatively weak, periodic orbits still constitute the skeleton of the dynamics [43, 15], and can be used as a starting point to develop our algorithm. The effects of the noise are accounted for by switching to a Fokker-Planck picture, and by considering the local eigenfunctions of the evolution operator in the neighborhood of the periodic orbits of the noiseless system. As seen in the previous section, refining a partition means to trace points back to their pre-images. In the Fokker-Planck picture, that translates into using the adjoint operator $\mathcal{L}^\dagger$ in the vicinity of the periodic points of the deterministic system, and finding its local ground-state eigenfunctions, which

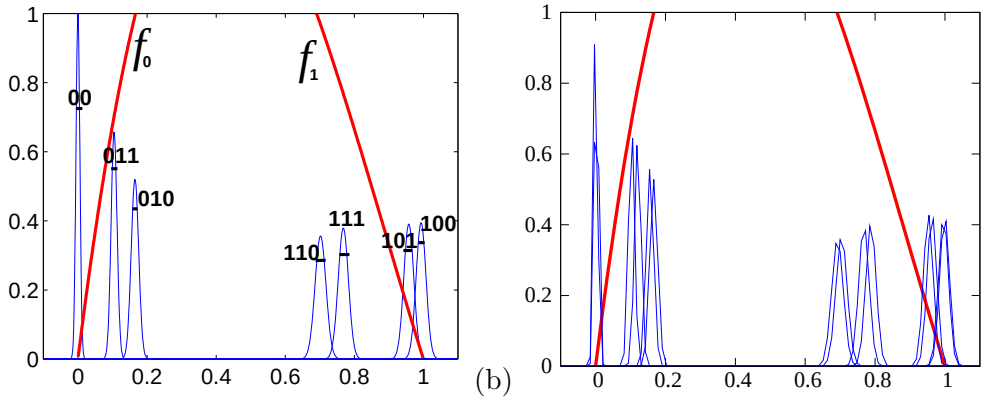


Figure 4: (a) f_0, f_1: branches of the deterministic map (48) for $\Lambda_0 = 8$ and $b = 0.6$. The local eigenfunctions $\tilde{\rho}_{a,0}$ with variances given by (46) provide a state space partitioning by neighborhoods of periodic points of period 3. These are computed for noise variance ($D =$ diffusion constant) $2D = 0.002$. (b) The next generation of eigenfunctions shows how the neighborhoods of the optimal partition cannot be resolved further. Only $\mathcal{M}_{011}$ can be split into $\mathcal{M}_{0110}$ and $\mathcal{M}_{0111}$ (second and third peak from the left), but that would not change the transition graph of Figure 6.

are then used to cover the non-wandering set of the state space. I derived the local spectrum of $\mathcal{L}^\dagger$ in Sect. 2.3.2, and found that the ground state is a Gaussian of width σ_a, such that (cf. equation (46))

$$\sigma_a^2 = \frac{2D}{\Lambda_p^2 - 1}\left(1 + \sum_{i=1}^{n-1}(f_{a-i}^{i'})^2\right) .$$

Every periodic point is assigned a neighborhood $[x_a - \sigma_a, x_a + \sigma_a]$, and the non-wandering set of the map is covered with neighborhoods of orbit points of higher and higher period n_p. I stop the refinement when adjacent neighborhoods, say of x_a and x_b, overlap in such a way that $|x_a - x_b| < \sigma_a + \sigma_b$.

As an example to illustrate the method, consider the chaotic repeller

$$x_{n+1} = \Lambda_0 x_n(1 - x_n)(1 - bx_n) + \xi_n\,, \quad \Lambda_0 = 8, b = 0.6 \tag{48}$$

with noise strength $2D = 0.002$, on the unit interval. The map is plotted in Figure 4 (a); this figure also shows the local eigenfunctions $\tilde{\rho}_{a,0}$ with variances given by (46). Each Gaussian is labeled by the $\{f_0, f_1\}$ branches visitation sequence of the corresponding deterministic periodic point. Figure 4 (b) illustrates the overlapping: $\mathcal{M}_{000}$ and $\mathcal{M}_{001}$ overlap, just like $\mathcal{M}_{0101}$ and $\mathcal{M}_{0100}$ and all the neighborhoods of the period $n_p = 4$ cycle points, except for $\mathcal{M}_{0110}$ and $\mathcal{M}_{0111}$. In this case the state space (the unit interval) can be resolved into 7

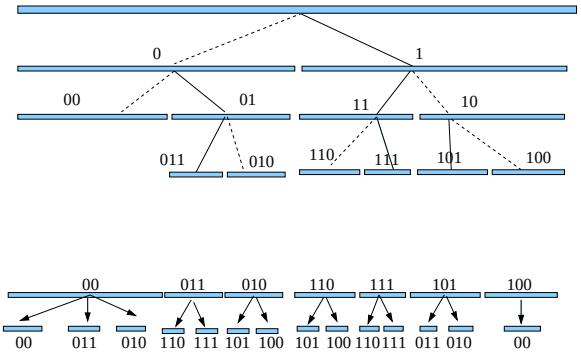


Figure 5: Binary tree of the refinement leading to the partition (49). Once the optimal resolution is found, the symbolic dynamics is recoded by relabeling the finite partition intervals.

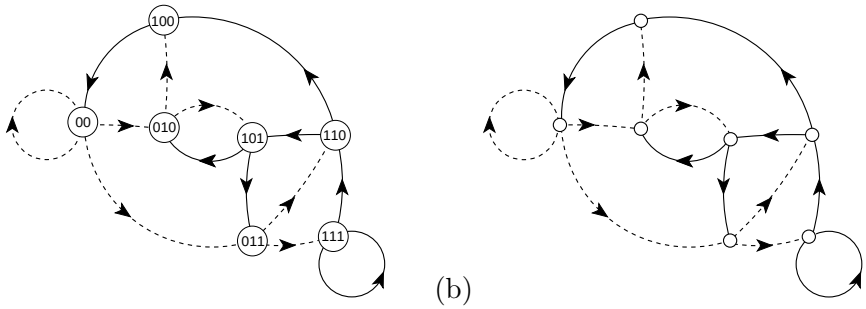


Figure 6: (a) Transition graph (graph whose links correspond to the nonzero elements of a transition matrix T_{ba}) describes which regions b can be reached from the region a in one time step. The 7 nodes correspond to the 7 regions of the optimal partition (49). Dotted links correspond to symbol 0, and the full ones to 1, indicating that the next region is reached by the f_0, respectively f_1 branch of the map plotted in Figure 4. (b) The region labels in the nodes can be omitted, with links keeping track of the symbolic dynamics. (from `ChaosBook.org`)

neighborhoods

$$\{\mathcal{M}_{00}, \mathcal{M}_{011}, \mathcal{M}_{010}, \mathcal{M}_{110}, \mathcal{M}_{111}, \mathcal{M}_{101}, \mathcal{M}_{100}\}\,. \tag{49}$$

It turns out that resolving $\mathcal{M}_{011}$ further into $\mathcal{M}_{0110}$ and $\mathcal{M}_{0111}$ is not essential, as it produces the same transition graph. Once the finest possible partition is determined, the finite binary tree in Figure 5 is drawn: Evolution in time maps the optimal partition interval $\mathcal{M}_{011} \to \{\mathcal{M}_{110}, \mathcal{M}_{111}\}$, $\mathcal{M}_{00} \to \{\mathcal{M}_{00}, \mathcal{M}_{011}, \mathcal{M}_{010}\}$, etc.. This is summarized in the transition graph (Figure 6), which we will use to estimate the escape rate and the Lyapunov exponent of the repeller.

How accurate is this algorithm? Everything that follows is aimed at testing the optimal partition hypothesis.

CHAPTER IV

PERIODIC ORBIT THEORY

What can a partition be used for, in practice? It can help, for example, find statistical properties of the system through long-time dynamical averages, e.g. correlations, Lyapunov exponents, or the escape rate from the chaotic region of the state space. In thus chapter I review the well-known methods of discretization for an evolution operator, and the periodic orbit theory.

4.1 The Fokker-Planck operator, discretized

In order to evaluate any dynamical average, we need to know about the transition rates between regions of the state space. Those will be, in a non-trivial way, the weights of the average. We then discretize the evolution operator on the state space, that is its own support.

This is best explained by dealing first with the deterministic evolution operator $\mathcal{L}_{det}$, and writing it as a matrix.

The simplest possible way of introducing a state space discretization is to partition the state space $\mathcal{M}$ with a non-overlapping collection of sets $\mathcal{M}_i$, $i = 1, \ldots, N$, and to consider piecewise constant densities, constant on each $\mathcal{M}_i$:

$$\rho(x) \;=\; \sum_{i=1}^{N} \rho_i \frac{\chi_i(x)}{|\mathcal{M}_i|}\,, \quad \chi_i(x) = \begin{cases} 1 & \text{if } x \in \mathcal{M}_i\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{50}$$

where $\chi_i(x)$ is the characteristic function of the set $\mathcal{M}_i$. The density ρ_i at a given instant is related to the densities at the previous step in time by the action of the Perron-Frobenius operator:

$$\begin{aligned} \rho'_j \;&=\; \int_{\mathcal{M}} dy\, \chi_j(y) \rho'(y) \;=\; \int_{\mathcal{M}} dx\, dy\, \chi_j(y)\, \delta(y - f(x))\, \rho(x) \\ &=\; \sum_{i=1}^{N} \rho_i \frac{|\mathcal{M}_i \cap f^{-1}(\mathcal{M}_j)|}{|\mathcal{M}_i|}\,. \end{aligned} \tag{51}$$

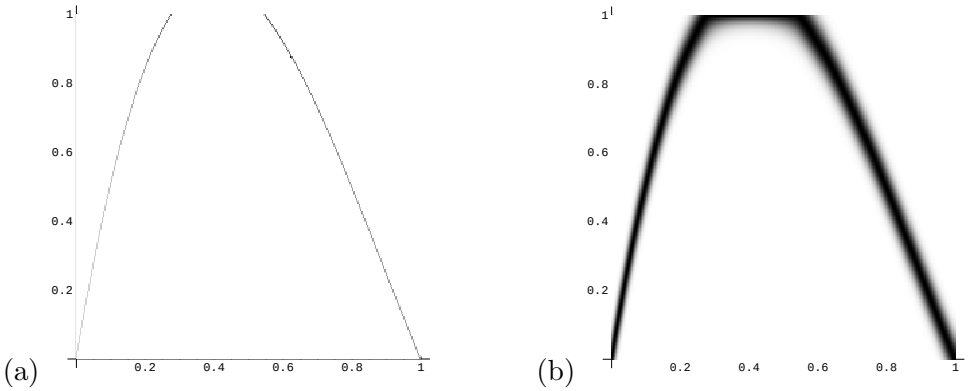


Figure 7: (a) Deterministic partition of the discretized Perron-Frobenius operatorfor the repeller (47) by a uniform mesh (256 intervals). Color/gray scale indicates the matrix element size. (b) The noisy version of (a), noise has variance $D = 10^{-3}$.

In this way

$$L_{ij} \;=\; \frac{|\mathcal{M}_i \cap f^{-1}(\mathcal{M}_j)|}{|\mathcal{M}_i|}\,, \quad \rho' = \rho\mathbf{L} \tag{52}$$

is the transition rate from $\mathcal{M}_j$ to $\mathcal{M}_i$, and the whole matrix is an approximation to the Perron-Frobenius operator(Figure 7). In the case of a repeller, the leading eigenvalue of the matrix yields the escape rate from the chaotic region of the state space and its leading left eigenvector is a piecewise constant approximation to the first eigenfunction of $\mathcal{L}_{det}$ (Figure 8). Then, the average of an observable $a(x)$ over the state space is

$$\langle a \rangle = \int dx\, e^{\gamma} \rho(x) a(x) \tag{53}$$

where $\rho(x)$ is the leading left eigenfunction of (52), γ is the escape rate, and $e^{\gamma}\rho$ is the normalized repeller measure, $\int dx\, e^{\gamma}\rho(x) = 1$.

In the presence of noise, the corresponding piecewise constant approximation to the Fokker-Planck operator (13) is

$$[\mathcal{L}]_{ij} \;=\; \frac{1}{|\mathcal{M}_i|}\frac{1}{\sqrt{4\pi D}} \int_{\mathcal{M}_i} dx \int_{f^{-1}(\mathcal{M}_j)} dy\, e^{-\frac{1}{4D}(y-f(x))^2}, \tag{54}$$

This method, proposed by Ulam in 1960 [48], has been shown to very accurately reproduce the spectrum for expanding maps, once finer and finer Markov partitions are used [21, 24]. The choice of the partition is indeed crucial to the convergence of the spectrum, meaning the better the partition, the fewer refinements are needed. A uniform mesh can always be

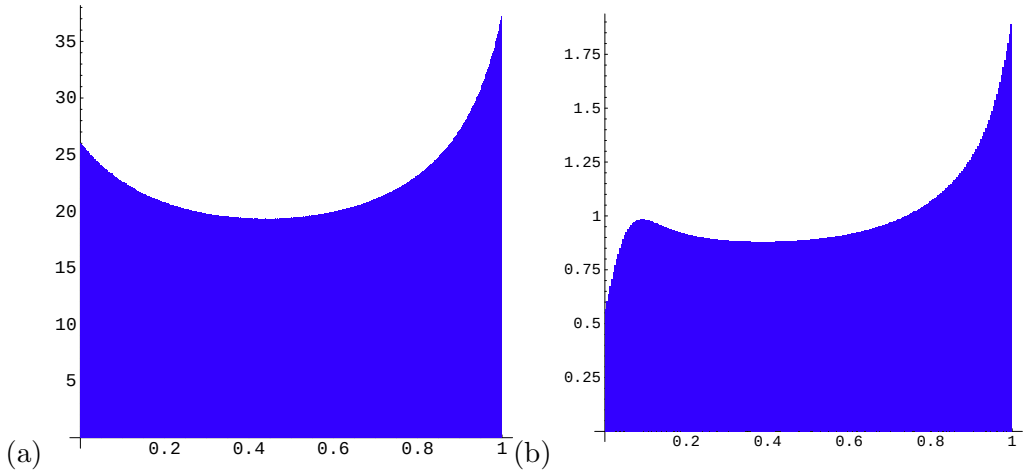


Figure 8: (a) The left leading eigenfunction ρ_0 of the uniformly discretized Perron-Frobenius operator Figure 7 for the repeller (47), $N = 256$. (b) Leading left eigenfunction of the uniformly discretized Fokker-Planck operator, in the presence of noise of variance $D = 10^{-3}$.

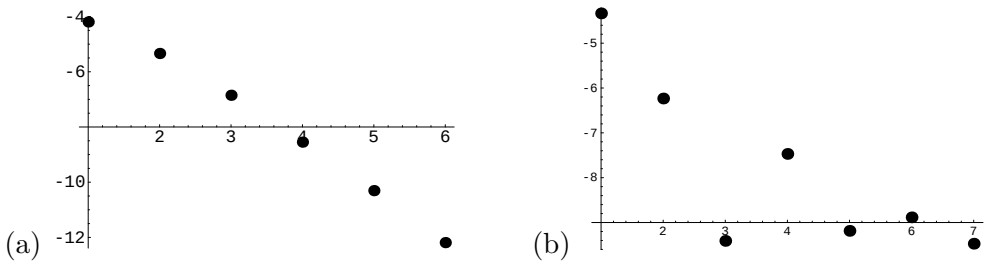


Figure 9: The convergence of the leading eigenvalue γ of the discretized Perron-Frobenius operator (Figure 7), using (a) a partition of the unit interval based on the preimages of the critical point, as explained in Sect. 3.1, and (b) a uniform-mesh discretization. Plotted: $\ln|\gamma^{(N)} - \gamma^{\infty}|$ as a function of $\ln N$, N being the number of partition intervals.

used, in case one has no clue. But the problem in that case is that the grid knows not what parts of the state space are more or less important, and the convergence is generally slower (Figure 9).

In the presence of noise, it is even more complicated to make an accurate non-overlapping partition of the state space. In our example of the repeller, we do not know exactly where the borders of the intervals that make the optimal partition are, meaning those seven neighborhoods (49) are not enough for the discretization (54) to provide an accurate estimate of the spectrum of $\mathcal{L}$. In that case it is still OK to use a uniform mesh and refine it until needed, and I will use that technique as a crosscheck in Chapter 5, but I am trying here to

develop a method that overlook the exact location of the borders of the intervals.

In order to do that, I shall adapt the well-known technique of *periodic orbit expansions*, outlined in the remainder of this chapter, to finite-dimensional operators like $\mathcal{L}$[1].

4.2 Evaluation of averages in a chaotic system

Given an observable $a = a(x)$, its expectation value is defined as as

$$\langle a \rangle = \lim_{n\to\infty} \frac{1}{\mathcal{M}} \int_{\mathcal{M}} dx \frac{1}{n} \int_0^n d\tau a(f^\tau(x)) \tag{55}$$

where $\mathcal{M}$ is the region of the state space we are taking into consideration. In reality (55) is not very practical when dealing with chaos. Let us define the quantity

$$A^n(x_0) = \sum_{k=0}^{n-1} a(f^k(x_0)) \tag{56}$$

as an *integrated observable* on a trajectory, and consider the spatial average

$$\left\langle e^{\beta A^n} \right\rangle = \frac{1}{\mathcal{M}} \int_{\mathcal{M}} dx e^{\beta A^n(x)} \tag{57}$$

where β is a dummy variable, which we take as a scalar. I now assume that every trajectory visits the whole phase space asymptotically (*ergodicity*), and that the system is also *mixing*, i.e.:

$$\lim_{n\to\infty} \langle a(x) b(f^n(x)) \rangle = \langle a(x) \rangle \langle b(x) \rangle \tag{58}$$

In that case, time averages can be replaced with space averages. We can then expect [10] the time average $\bar{a}$ to asymptotically tend to a constant, and the integrated observable A^n to tend to $n\bar{a}$. This way, the phase-space average (57) grows exponentially with the number of iterations, and its value is given asymptotically by the leading eigenvalue

$$\left\langle e^{\beta A^n} \right\rangle \propto e^{n s_0(\beta)}$$

and therefore

$$s_0(\beta) = \lim_{n\to\infty} \frac{1}{n} \ln \left\langle e^{\beta A^n} \right\rangle \tag{59}$$

[1] A full account of periodic orbit theory is given in `ChaosBook.org` [11].

which makes it possible to evaluate the moments of $a(x)$ as derivatives of $s_0(\beta)$, for example the expectation value (55) is:

$$\left.\frac{\partial s_0}{\partial \beta}\right|_{\beta=0} = \lim_{n\to\infty} \frac{1}{n} \langle A^n \rangle = \langle a \rangle \tag{60}$$

If the dynamics is confined to the region $\mathcal{M}$, definition (57) is correct, otherwise one must impose that the average be only on points of $\mathcal{M}$, so that (57) becomes

$$\left\langle e^{\beta A^n} \right\rangle = \frac{1}{\mathcal{M}} \int_{\mathcal{M}} dx \int_{\mathcal{M}} dy \delta(y - f^n(x)) e^{\beta A^n(x)} \tag{61}$$

in which we find the evolution operator

$$\mathcal{L}^n(y,x) = \delta(y - f^n(x)) e^{\beta A^n(x)} \tag{62}$$

I can now express the expectation value (61) as

$$\left\langle e^{\beta A^n} \right\rangle = \langle \mathcal{L}^n \rangle$$

and notice that $\mathcal{L}^n$ can be written as a linear combination of eigenvalues and eigenfunctions

$$\mathcal{L}^n a(x) = \sum_i c_i s_i^n \phi_i(x) =$$

$$c_0 s_0^n \phi_0(x) \left[1 + \sum \left(\frac{c_i s_i \phi_i(x)}{c_0 s_0 \phi_0(x)} \right)^n \right] \simeq c_0 s_0^n \phi_0(x) \tag{63}$$

for large n, where s_0 is the leading eigenvalue of $\mathcal{L}$, so that

$$\langle \mathcal{L} \rangle = e^{s_0} \tag{64}$$

and all one needs is to find s_0.

4.3 Traces, determinants, and dynamical $1/\zeta$ *function*

So far an idea was presented to estimate the expectation value of an observable by using an evolution operator, however it is not yet clear where periodic orbits fit in this picture. One way of finding the leading eigenvalue of $\mathcal{L}_{det}$ is to evaluate its trace

$$\mathrm{tr}\mathcal{L}^n = \int_{\mathcal{M}} dx \mathcal{L}^n(x,x) = \int_{\mathcal{M}} dx \delta(x - f^n(x)) e^{\beta A^n(x)} \tag{65}$$

since it tends to the leading eigenvalue, as $n \to \infty$. According to (65), the trace of $\mathcal{L}^n$ picks a contribution every time $x - f^n(x) = 0$, that is when x is a periodic point of $f(x)$, meaning we are summing over all the periodic orbits of period n of our system. Now that we know that, (65) is also

$$\mathrm{tr}\mathcal{L}^n = \int_{\mathcal{M}} dx \mathcal{L}^n(x,x) = \sum_{x_i \in \mathrm{Fix} f^n} \frac{e^{\beta A_i}}{|\det(\mathbf{1} - \mathbf{J}^n(x_i))|} \tag{66}$$

where Fix $f^n = \{x : f^n(x) = x\}$ is the set of periodic points of period n and the Jacobian in the denominator follows from a change of variable in (65) to evaluate the δ-function integral. One can immediately write the necessary condition $\mathbf{J}^n(x_i) \neq \mathbf{1}$ in order for (66) to exist. With that assumption in mind, let us now take a Laplace transform of $\mathrm{tr}\mathcal{L}_n$:

$$\sum_{n=1}^{\infty} z^n tr\mathcal{L}^n = \mathrm{tr}\frac{z\mathcal{L}}{1 - z\mathcal{L}} = \sum_p n_p \sum_{r=1}^{\infty} \frac{z^{n_p r} e^{r\beta A_p}}{|\det(\mathbf{1} - \mathbf{J}_p^r)|} \tag{67}$$

where p indicates a *prime* periodic orbit, *ie.* not a repetition of shorter cycles. The previous expression is known as *trace formula*. Let $\lambda_{p,e}$ and $\lambda_{p,c}$ be respectively the expanding and contracting eigenvalues of a periodic orbit p, and $\mathbf{J}_p$ the Jacobian of the same orbit. Then

$$|\det(\mathbf{1} - \mathbf{J}_p)|^{-1} = \frac{1}{|\Lambda_p|} \prod_e \frac{1}{1 - 1/\lambda_{p,e}} \prod_c \frac{1}{1 - \lambda_{p,c}} \tag{68}$$

where $\Lambda_p = \prod_e \lambda_{p,e}$, and

$$|\det(\mathbf{1} - \mathbf{J}_p^n)| \to \frac{1}{|\Lambda_p|^n} \tag{69}$$

when $n >> 1$. Now plug (69) into (67) to get an approximated expression of the trace formula.

The geometric series in (67) can be rewritten as a function of the eigenvalues of the evolution operator[2], in the form

$$\sum_{i=0}^{\infty} \frac{z e^{s_i}}{1 - z e^{s_i}} \tag{70}$$

whence it is evident that the trace diverges for $z = e^{-s_0}$. Thus, looking for the leading eigenvalue of $\mathcal{L}_{det}$ is the same as looking for the radius of convergence of (67), which in general is no easy task to fulfill. Rather, one can use the identity

$$\ln\det(1 - z\mathcal{L}) = \mathrm{tr}\ln(1 - z\mathcal{L}) = -\sum_{n=1}^{\infty} \frac{z^n}{n} \mathrm{tr}\mathcal{L}^n \tag{71}$$

[2]By definition, $\mathrm{tr}\mathcal{L}^n = \sum_{i=0}^{\infty} e^{n s_i}$

to obtain the *determinant* of the evolution operator, whose leading root (the closest solution of (71) to $z = 0$) is exactly the radius of convergence of the trace formula. We shall see that the power series (71) is to be truncated to a polynomial, at which point $z = e^{-s_0}$ is easy to find. If one gets the expression for $\det(1 - z\mathcal{L})$ from (71) and applies the limit (69), the outcome is the so-called *dynamical* $1/\zeta$ *function*:

$$1/\zeta(z, \beta) = \exp\left(-\sum_p \sum_{r=1}^{\infty} \frac{1}{r} \frac{z^{n_p r} e^{r\beta A_p}}{|\Lambda_p|^r}\right) \tag{72}$$

As already said,

$$1/\zeta(s_0, \beta) = 0$$

Equation (72) can be written in a more compact form by defining

$$t_p(z, \beta) = \frac{z^{n_p}}{|\Lambda_p|} e^{\beta A_p} \tag{73}$$

and considering that $\sum_r t_p^r/r = -\ln(1 - t_p)$, so as to obtain the *Euler product*

$$1/\zeta = \prod_p (1 - t_p) \tag{74}$$

4.4 Cycle expansions

The next step is to compute (74) and work out a formula to to estimate the average $\langle a \rangle$, introduced in Sect. 4.2. First, equation (74) can be rewritten as

$$1/\zeta = 1 - {\sum_{p_1,\ldots,p_k}}' (-1)^{k+1} t_{p_1} t_{p_2} \ldots t_{p_k}$$

where $\sum'$ means the sum is over all the distinct, non repeated combinations of the prime cycles. Now call $t_\pi = (-1)^{k+1} t_{p_1} t_{p_2} \ldots t_{p_k}$ any product of the weights t_p's of the prime orbits, we can write the previous expression as

$$1/\zeta = 1 - {\sum_{\pi}}' t_\pi \tag{75}$$

We are dealing here with an infinite series, and we need to truncate it properly. The idea is to arrange the terms of the sum in order of length, as the shorter cycles are normally the least unstable and therefore contribute the largest terms. Knowing about the symbolic

dynamics of the system (Sect. 3.1) also helps: in the case of a binary symbolic dynamics, for instance, the product

$$1/\zeta = (1-t_0)(1-t_1)(1-t_{01})(1-t_{001})(1-t_{011})$$

$$(1-t_{0001})(1-t_{0011})(1-t_{0111})...$$

is rewritten as

$$1/\zeta = 1 - t_0 - t_1 - t_{01} - t_{001} - t_{011} - t_{0001} - t_{0011} - t_{0111} - ...$$

$$+t_0t_1 + t_0t_{01} + t_{01}t_1 + t_0t_{001} + t_0t_{011} + t_{001}t_1 + t_{011}t_1$$

$$-t_0t_1t_{01} - ...$$

which we call *cycle expansion*. The next step is to regroup the terms of the sum in *fundamental* contributions and *curvature* corrections in the following way

$$1/\zeta = 1 - t_0 - t_1 - [(t_{01} - t_1t_0)] - [(t_{001} - t_{01}t_0) + (t_{011} - t_{01}t_1)]$$

$$-[(t_{0001} - t_0t_{001}) + (t_{0111} - t_{011}t_1)$$

$$+(t_{0011} - t_{001}t_1 - t_0t_{011} + t_0t_1t_{01})] - ...$$

$$= 1 - \sum_f t_f - \sum_n \hat{c} \tag{76}$$

The previous expansion is dominated by the first two terms (t_0 and t_1) and progressively corrected by the others, in which the weight of a cycle (say t_{001}) is typically *shadowed* by the weights of shorter cycles multiplied together (t_0 and t_{01}) to give the same symbolic sequence. In a hyperbolic system, where the stability eigenvalues grow exponentially with respect to the cycle period ($\Lambda_p \propto C^{n_p}$), the curvature corrections in the cycle expansions become exponentially smaller with the period of the truncation, which make the sum (76) converge very rapidly.

4.5 Cycle expansions and averages

Equation (72) says that the function $1/\zeta = 1/\zeta(z,\beta)$. We said in Sect. 4.3 that the leading root z_0 of $1/\zeta$ is nothing but e^{-s_0}, so that, at the end of the day, $z_0 = z_0(\beta)$. Then, in order to estimate the average $\langle a \rangle$ using (60)

$$\left.\frac{\partial s_0}{\partial \beta}\right|_{\beta=0} = \lim_{n\to\infty} \frac{1}{n} \langle A^n \rangle = \langle a \rangle$$

we may consider the equation $1/\zeta(s_0,\beta) = 0$ as an implicit function $F(\beta, s_0(\beta))$, and

$$\begin{aligned} 0 &= \frac{d}{d\beta} F(\beta, s_0(\beta)) \\ &= \frac{\partial F}{\partial \beta} + \frac{ds_0}{d\beta} \left.\frac{\partial F}{\partial s_0}\right|_{s_0=s_0(\beta)} \Rightarrow \frac{ds}{d\beta} = -\frac{\partial F}{\partial \beta} \Big/ \frac{\partial F}{\partial s_0} \end{aligned} \tag{77}$$

From (60), (75), and 73 we obtain

$$\langle a \rangle = \left[-\frac{\partial}{\partial \beta}(1/\zeta) \right]_{\beta=0} \Bigg/ \left[-z\frac{\partial}{\partial z}(1/\zeta) \right]_{\beta=0} \tag{78}$$

where

$$\langle A \rangle = {\sum_{\pi}}' (-1)^{k+1} \frac{A_{p_1} e^{-n_{p_1} s_0} + A_{p_2} e^{-n_{p_2} s_0} + ... + A_{p_k} e^{-n_{p_k} s_0}}{|\Lambda_{p_1} ... \Lambda_{p_k}|} \tag{79}$$

and

$$\langle n \rangle = {\sum_{\pi}}' (-1)^{k+1} \frac{n_{p_1} e^{-n_{p_1} s_0} + n_{p_2} e^{-n_{p_2} s_0} + ... + n_{p_k} e^{-n_{p_k} s_0}}{|\Lambda_{p_1} ... \Lambda_{p_k}|} \tag{80}$$

are respectively numerator and denominator of (78). Now we have everything we need to compute the average of an observable using cycle expansion.

CHAPTER V

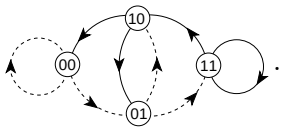
MATRIX APPROXIMATION TO THE FOKKER-PLANCK OPERATOR

I will now apply the cycle expansions described in the last chapter to a finite matrix, using the method of *local matrix approximation*, originally introduced by Cvitanović and co-workers [14] for an infinite-dimensional evolution operator. The adaptation of this method to a finite matrix and the use of a Hermite polynomials basis are both novel. In Sect. 5.2, I use this technique to validate the optimal partition hypothesis, formulated in Chapter 3.

5.1 Periodic orbit theory of a matrix

In the presence of noise, I claimed in Sect. 3.2 that the state space cannot be resolved beyond a certain optimal partition, meaning the evolution operator is now supported on a set of finite measure, and it is just a matrix. As an example, let us start from the Fokker-Planck operator for the simple partition of four intervals in Figure 1, Sect. 3.1:

$$\mathbf{L} = \begin{pmatrix} L_{00,00} & 0 & L_{00,10} & 0 \\ L_{01,00} & 0 & L_{01,10} & 0 \\ 0 & L_{10,01} & 0 & L_{10,11} \\ 0 & L_{11,01} & 0 & L_{11,11} \end{pmatrix} = \quad . \tag{81}$$



Now, if we knew the exact size of the intervals of the partition, a discretization like (54) would determine the entries L_{ij} with sufficient accuracy, but it turns out that, in the presence of noise, the borders of the intervals are fuzzy and we cannot just diagonalize the

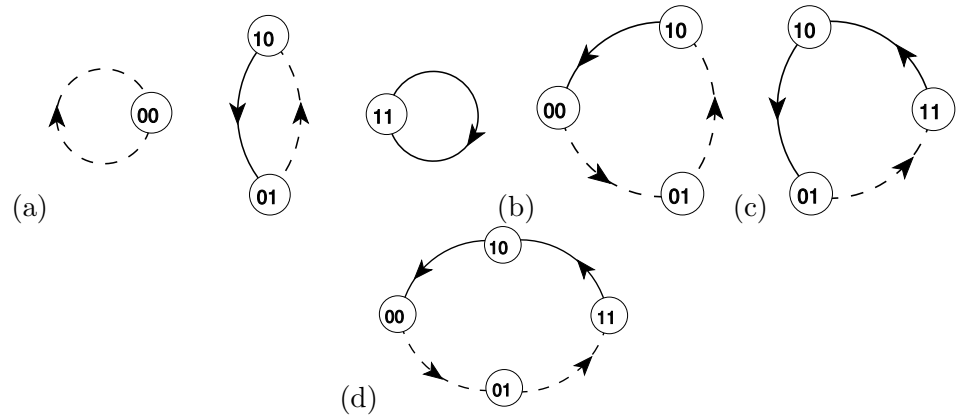


Figure 10: (a)-(d): all the non-self-intersecting loops of the four-node transition graph in Figure 81 (from `ChaosBook.org`).

matrix $\mathbf{L}$. The determinant (71) of the evolution operator is

$$\begin{aligned}\det(1-z\mathbf{L}) &= 1 - L_{00,00}z - L_{11,11}z + (-L_{01,10}L_{10,01} + L_{00,00}L_{11,11})z^2 + \\ &\quad (-L_{00,10}L_{10,01}L_{01,00} + L_{00,00}L_{01,10}L_{10,01} \\ &\quad -L_{01,10}L_{10,11}L_{11,01} + L_{01,10}L_{10,01}L_{11,11})z^3 + \\ &\quad (-L_{00,10}L_{10,11}L_{11,01}L_{01,00} + L_{00,00}L_{01,10}L_{10,11}L_{11,01} \\ &\quad +L_{00,10}L_{10,01}L_{01,00}L_{11,11} - L_{00,00}L_{01,10}L_{10,01}L_{11,11})z^4\,. \end{aligned} \tag{82}$$

Notice that every product of matrix elements in (82) describes the transition rate of a closed path of the graph, e.g. $L_{00,10}L_{10,01}L_{01,00}$ is the combined probability that an orbit starts in region $\mathcal{M}_{00}$, visits regions $\mathcal{M}_{10}$ and $\mathcal{M}_{01}$ and goes back to $\mathcal{M}_{00}$. Since there is a periodic orbit for each region of the partition (Sect. 3.1), we approximate the transition rate of every non-self-intersecting loop of the graph (Figure 10) with the contribution to trace formula (67) of the cycle that follows the same path, for instance

$$L_{00,10}L_{10,01}L_{01,00} \to t_{001} = \frac{1}{|1-\Lambda_{001}|} \tag{83}$$

and the expression for the determinant (82) becomes

$$\begin{aligned}\det(1-z\mathbf{L}) &= 1 - (t_0+t_1)z - [(t_{01} - t_1t_0)]z^2 - [(t_{001} - t_{01}t_0) + (t_{011} - t_{01}t_1)]z^3 \\ &\quad -[(t_{0011} - t_{001}t_1 - t_0t_{011} + t_0t_{01}t_1)]z^4 \end{aligned} \tag{84}$$

In the presence of weak noise, (83) is a zeroth-order approximation to the trace of the local Fokker-Planck operator, and higher order corrections can be included.

Let me now derive the approximation (83). As explained in Sect. 2.2, near a periodic point $x_a \in p$, the n_pth iterate ${\mathcal{L}^{n_p}}_a$ of the linearization (28) is the discrete-time version of the Ornstein-Uhlenbeck process [47], with left $\tilde{\rho}_0$, $\tilde{\rho}_1$, $\cdots$, respectively right ρ_0, ρ_1, $\cdots$ mutually orthogonal eigenfunctions [42] given by

$$\begin{aligned} \tilde{\rho}_{a,k}(z) &= \frac{\beta^{k+1}}{\sqrt{\pi}2^k k!} H_k(\beta z) e^{-(\beta z)^2} \\ \rho_{a,k}(z) &= \frac{1}{\beta^k} H_k(\beta z)\,, \end{aligned} \tag{85}$$

where $H_k(x)$ is the kth Hermite polynomial, $1/\beta = \sqrt{2}\sigma_a$, and the kth eigenvalue is $1/|\Lambda|\Lambda^k$. Given the finest possible partition, the Fokker-Planck operator now acts as a matrix with non-zero $a \to b$ entries. The idea is now to expand every matrix element in the Hermite basis,

$$\begin{aligned} [\mathbf{L}_{ba}]_{kj} &= \langle \tilde{\rho}_{b,k} | \mathcal{L} | \rho_{a,j} \rangle \\ &= \int \frac{dz_b dz_a\, \beta}{2^{j+1} j! \pi \sqrt{D}} e^{-(\beta z_b)^2 - \frac{(z_b - f_a(z_a))^2}{4D}} \\ &\qquad \times H_k(\beta z_b) H_j(\beta z_a)\,, \end{aligned} \tag{86}$$

where $1/\beta = \sqrt{2}\sigma_a$, and z_a is the deviation from the periodic point x_a. It is the number of resolved periodic points that determines the dimensionality of the Fokker-Planck matrix. Its eigenvalues are determined from the zeros of $\det(1 - z\mathbf{L})$ (equation (84)), expanded as a polynomial in z, with coefficients given by traces of powers of $\mathbf{L}$, as in the trace formula (67). As the trace of the nth iterate of the Fokker-Planck operator $\mathcal{L}^n$ is concentrated on periodic points $f^n(x_a) = x_a$, I evaluate the contribution of periodic orbit p to $\mathrm{tr}\,\mathbf{L}^{n_p}$ by centering $\mathbf{L}$ on the periodic orbit,

$$t_p = \mathrm{tr}_{\,p}\, \mathcal{L}^{n_p} = \mathrm{tr}\, \mathbf{L}_{ad} \cdots \mathbf{L}_{cb} \mathbf{L}_{ba}\,, \tag{87}$$

where $x_a, x_b, \cdots x_d \in p$ are successive periodic points. Now just Taylor-expand the exponential in (86) around the periodic point x_a,

$$\begin{aligned} e^{-\frac{(z_b - f_a(z_a))^2}{4D}} &= e^{-\frac{(z_b - f_a' z_a)^2}{4D}} \times \\ &\Big(1 - 2\sqrt{D}(f_a'' f_a' z_a^3 + f_a'' z_a^2 z_b) + O(D)\Big)\,. \end{aligned} \tag{88}$$

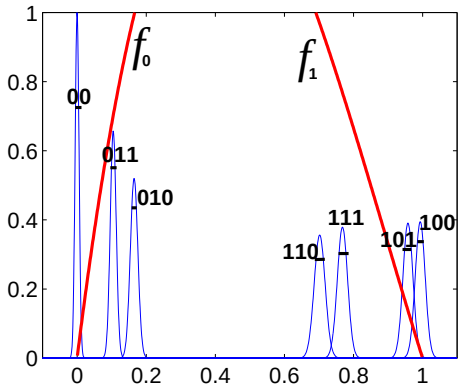


Figure 11: f_0, f_1: branches of the deterministic map (48) for $\Lambda_0 = 8$ and $b = 0.6$. The local eigenfunctions $\tilde{\rho}_a$ with variances given by (46) provide a state space partitioning by neighborhoods of periodic points of period 3. These are computed for noise variance ($D =$ diffusion constant) $2D = 0.002$. The neighborhoods $\mathcal{M}_{000}$ and $\mathcal{M}_{001}$ already overlap, so $\mathcal{M}_{00}$ cannot be resolved further. For periodic points of period 4, only $\mathcal{M}_{011}$ can be resolved further, into $\mathcal{M}_{0110}$ and $\mathcal{M}_{0111}$.

To leading order in the noise variance $2D$, t_p takes the deterministic value $t_p = 1/|\Lambda_p - 1|$, approximation (83). Higher order corrections will be needed in what follows for a sufficiently accurate comparison of different methods.

5.2 Testing the optimal partition hypothesis

It is now time to test the validity of the optimal partition method presented in Sect. 3.2. I do so by estimating the escape rate $\gamma = -\ln z_0$, where z_0^{-1} is the leading eigenvalue of Fokker-Planck operator $\mathcal{L}$, for the repeller considered in Sect. 3.2 and plotted again in Figure 11. The spectral determinant can be read off the transition graph of Figure 6, (dissected in all its non-self-intersecting loops in Figure 12):

$$\begin{aligned}
\det(1 - z\mathbf{L}) = 1 &- (t_0 + t_1)z - (t_{01} - t_0 t_1)\, z^2 \\
&-(t_{001} + t_{011} - t_{01}t_0 - t_{01}t_1)\, z^3 \\
&-(t_{0011} + t_{0111} - t_{001}t_1 - t_{011}t_0 - t_{011}t_1 + t_{01}t_0t_1)\, z^4 \\
&-(t_{00111} - t_{0111}t_0 - t_{0011}t_1 + t_{011}t_0t_1)\, z^5 \\
&-(t_{001011} + t_{001101} - t_{0011}t_{01} - t_{001}t_{011})\, z^6 \\
&-(t_{0010111} + t_{0011101} - t_{001011}t_1 - t_{001101}t_1 \\
&-t_{00111}t_{01} + t_{0011}t_{01}t_1 + t_{001}t_{011}t_1)\, z^7. \qquad (89)
\end{aligned}$$

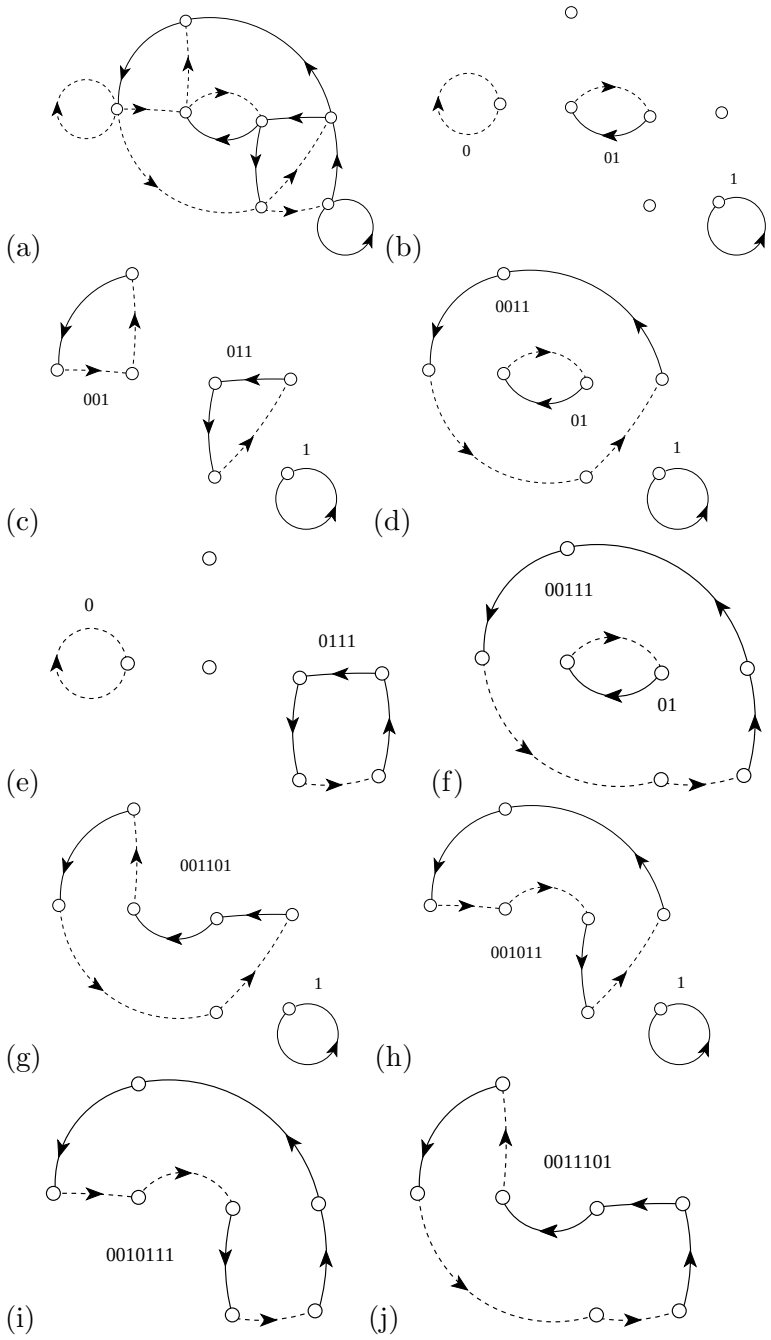


Figure 12: (a) The transition graph of the partition in Figure 11. (b)-(j) The fundamental cycles for the transition graph (a), i.e., the set of its non-self-intersecting loops. Each loop represents a local trace t_p, as in (87). (from `ChaosBook.org`)

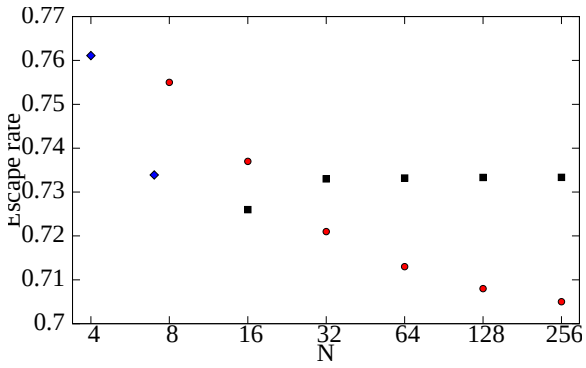


Figure 13: The escape rate γ of the repeller Figure 11 plotted as function of number of partition intervals N, estimated using: (⧫) under-resolved 4-interval and the 7-interval 'optimal partition', (●) all periodic orbits of periods up to $n = 8$ in the deterministic, binary symbolic dynamics, with $N_i = 2^n$ periodic-point intervals (the deterministic, noiseless escape rate is $\gamma_{<>} = 0.7011$), and (■) a uniform discretization (54) in $N = 16, \cdots, 256$ intervals. For $N = 512$ discretization yields $\gamma_{\rm num} = 0.73335(4)$.

The polynomial coefficients are given by products of non-intersecting loops of the transition graph [11], with the escape rate given by the leading root z_0^{-1} of the polynomial. Twelve periodic orbits $\overline{0}$, $\overline{1}$, $\overline{01}$, $\overline{001}$, $\overline{011}$, $\overline{0011}$, $\overline{0111}$, $\overline{00111}$, $\overline{001101}$, $\overline{001011}$, $\overline{0010111}$, $\overline{0011101}$ up to period 7 (out of the 41 contributing to the noiseless, deterministic cycle expansion up to cycle period 7) suffice to fully determine the spectral determinant of the Fokker-Planck operator. In the evaluation of traces (87) I include stochastic corrections up to order $O(D)$ (an order beyond the term kept in (88)). The escape rate of the repeller of Figure 11 so computed is reported in Figure 13, together with: (a) several deterministic, over-resolved partitions, and (b) a brute force numerical discretization of the Fokker-Planck operator. (a) If there is an optimal resolution, then any over-resolved periodic orbit expansions should give the wrong answer for the observable we want to estimate. I test such a statement by evaluating the escape rate using a cycle expansion of the kind I have described in Sect. 4.4, precisely equation (76), that is in terms of *all* deterministic periodic orbits of the map up to a given period, with t_p evaluated in terms of Fokker-Planck local traces (87), including stochastic corrections up to order $O(D)$. Figure 13 shows how the escape rate varies as I include all periodic orbits up to periods 2 through 8. Successive estimates of the escape rate appear to converge to a value different from the 'optimal partition' estimate. (b) I discretize the Fokker-Planck operator $\mathcal{L}$ by the piecewise-constant approximation on a uniform mesh

on the unit interval, introduced in Sect. 4.1, equation (54):

$$[\mathcal{L}]_{ij} \;=\; \frac{1}{|\mathcal{M}_i|}\frac{1}{\sqrt{4\pi D}}\int_{\mathcal{M}_i} dx \int_{f^{-1}(\mathcal{M}_j)} dy\, e^{-\frac{1}{4D}(y-f(x))^2},$$

where $\mathcal{M}_i$ is the ith interval in equipartition of the unit interval into N pieces. Empirically, $N = 128$ intervals suffice to compute the leading eigenvalue of the discretized $[128 \times 128]$ matrix $[\mathcal{L}]_{ij}$ to four significant digits. The latter turns out to be in excellent agreement with the escape rate calculated using the optimal partition, so that it now makes sense to test the method for an extensive range of values of the noise strength $2D$. The optimal partition method yields a different number of neighborhoods every time, the results are summarized in Table 1, and illustrated by Figure 14, with the estimates of the 'optimal partition' method within 2% of those given by the uniform discretization of Fokker-Planck. One can also see from the same table that the escape rates calculated with and without higher order corrections to the matrix elements (86) are consistent within less than 2%, meaning that the stochastic corrections (88) do not make a significant difference, as opposed to the choice of the partition, and need not be taken into account in this model.

Table 1: Escape rates of the repeller (48) from the unit interval, calculated from the determinant of the graph of the optimal partition: $(\gamma_{<>}^{O(D)})$ with stochastic corrections, $(\gamma_{<>})$ without stochastic corrections, and $(\gamma_{\rm num})$ by a uniform discretization of $\mathcal{L}$, for different values of D. n_r is the number of regions of the state space resolved by the optimal partition every time.

D	n_r	$\gamma_{<>}$	$\gamma_{<>}^{O(D)}$	$\gamma_{\rm num}$
0.01	4	0.763	0.748	0.773
0.008	5	0.763	0.751	0.769
0.005	5	0.763	0.755	0.759
0.003	6	0.736	0.732	0.747
0.001	7	0.735	0.734	0.733
0.0008	7	0.735	0.735	0.732
0.0005	9	0.736	0.735	0.729
0.0003	11	0.725	0.724	0.726
0.0001	14	0.722	0.722	0.718

Another interesting observable in the chaotic system in exam is the *Lyapunov exponent*, which measures how fast neighboring orbits separate, and therefore, how sensitive a dynamical system is to initial conditions: let x_0 and $x_0 + \delta_0$ be nearby points, then $\delta_n = f^n(x_0 + \delta_0) - f^n(x_0)$ is the separation of their nth iterates. If the orbits separate

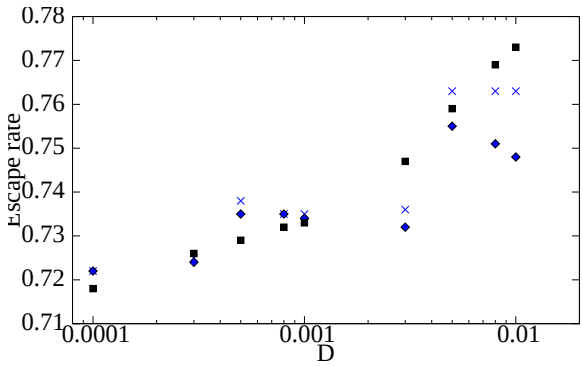


Figure 14: Escape rates of the repeller (48) vs. the noise strength D, using: the optimal partition method with (♦) and without (×) stochastic corrections; (■) a uniform discretization (54) in $N = 128$ intervals.

exponentially fast, then $|\delta_n| \propto |\delta_0| e^{n\lambda}$, with λ positive and asymptotically equal to

$$\lambda = \lim_{n\to\infty} \frac{1}{n} \ln \left| \frac{f^n(x_0+\delta_0) - f^n(x_0)}{\delta_0} \right| = \lim_{n\to\infty} \frac{1}{n} \ln \left| f^{n'}(x_0) \right| = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln \left| f'(x_i) \right| \quad (90)$$

for $\delta_0 \to 0$. In order to test the validity of the optimal partition, I compute the Lypaunov exponent as the ratio

$$\lambda = \langle \ln \Lambda \rangle \,/\, \langle n \rangle \,, \quad (91)$$

$\ln \Lambda = \sum \ln |f'(x_0)|$ being the integrated observable ('$A^n(x_0)$') defined in Sect. 4.2, and the two averages are given by the formulae (79) and (80) derived in Sect. 4.5:

$$\begin{aligned} \langle A \rangle \;\; = \;\; & A_0 t_0 + A_1 t_1 + (A_{01} t_{01} - (A_0 + A_1) t_0 t_1) + (A_{001} t_{001} - (A_{01} + A_0) t_{01} t_0) + \\ & + (A_{011} t_{011} - (A_{01} + A_1) t_{01} t_1) + \cdots \end{aligned} \quad (92)$$

with $t_p = e^\gamma / |\Lambda_p - 1|$. In this case the sums are finite, and over the loops of the transition graph generated by the optimal partition. On the other hand, I also use the uniform discretization (54) to crosscheck my estimate: this way the Lyapunov exponent is evaluated as the average

$$\lambda = \int dx \, e^\gamma \rho(x) \ln |f'(x)| \quad (93)$$

where $\rho(x)$ is the leading eigenfunction of (54), γ is the escape rate, and $e^\gamma \rho$ is the normalized repeller measure, $\int dx \, e^\gamma \rho(x) = 1$. Figure 15 shows close agreement ($< 1\%$) between the Lyapunov exponent estimated using the average (92), and the same quantity evaluated with (93), by the discretization method (54).

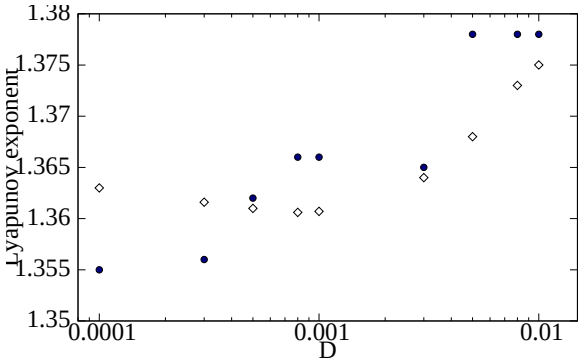


Figure 15: The Lyapunov exponent of the repeller (48) vs. the noise strength D, using: the optimal partition method ($\bullet$) without stochastic corrections, and ($\diamond$) a uniform discretization (54) over $N = 128$ intervals.

CHAPTER VI

WHEN THE GAUSSIAN APPROXIMATION FAILS

The last main result of this thesis is the formulation and validation of the optimal partition hypothesis for a non-hyperbolic map, and its test on a one-dimensional map. Take for example, the same map $f(x)$ as in (48), but with the parameter $\Lambda_0 = 1/f(x_c)$, where x_c is the maximum of the cubic parabola (Figure 16). As we can see from the figure, $f(x)$ maps the unit interval into itself, meaning there is no escape, besides it has a 'flat top' ($|f'(x)| << 1$) near its maximum, where the approximation (28) of the Fokker-Planck evolution operator

$$\mathcal{L}_a(z_{a+1}, z_a) = (4\pi D)^{-1/2}\, e^{-\frac{(z_{a+1} - f_a' z_a)^2}{4D}}$$

does not hold. Thus, I should first modify my choice of densities and neighborhoods, as the whole construction leading to the optimal partition algorithm was entirely based on the Gaussian approximation of the evolution operator.

Let us start recalling how the adjoint Fokker-Planck operator $\mathcal{L}^\dagger$ acts on a Gaussian density $\rho(x)$:

$$\begin{aligned} \mathcal{L}^\dagger \rho_a(x) &= \int_{-\infty}^{\infty} c_a e^{-\frac{(f(x)-y)^2}{4D}} e^{-(x-x_a)^2/2\sigma_a^2}[dy] \\ &= c_{a-1} e^{-\frac{(f(x)-x_a)^2}{\sigma_a^2+2D}} . \end{aligned} \tag{94}$$

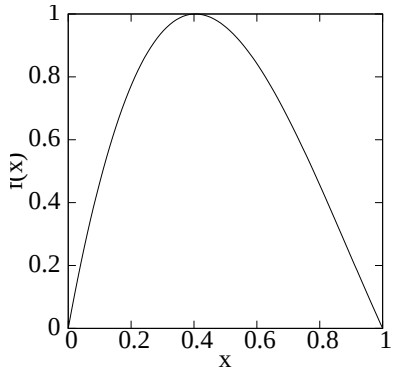


Figure 16: The 'skew' Ulam map $f(x) = \Lambda_0 x(1-x)(1-bx)$, with $b = 0.6$, and $\Lambda_0 = 1/f(x_c)$, x_c critical point of the map.

Suppose the point $f^{-1}(x_a)$ (cf. Sect. 2.3), around which I want to approximate the new density, is very close to the critical point (flat top) of the map, so that I can write

$$\begin{aligned}\rho_{a-1}(x) &= c_{a-1}e^{-\frac{(f(x)-x_a)^2}{2(\sigma_a^2+2D)}} \simeq c_{a-1}e^{-\frac{f''(f^{-1}(x_a))^2(x-f^{-1}(x_a))^4}{8(\sigma_a^2+2D)}} \\ &= c_{a-1}e^{-f''^2_{a-1}z^4_{a-1}/8(\sigma_a^2+2D)} \end{aligned} \tag{95}$$

the variance of the transformed density is

$$\sigma^2_{a-1} = \frac{\int z^2_{a-1}e^{-f''^2_{a-1}z^4_{a-1}/8(\sigma_a^2+2D)}dz_{a-1}}{\int e^{-f''^2_{a-1}z^4_{a-1}/8(\sigma_a^2+2D)}dz_{a-1}} = \frac{\Gamma(3/4)}{\Gamma(1/4)}\left(\frac{8(\sigma_a^2+2D)}{f''^2_{a-1}}\right)^{1/2} \tag{96}$$

that is no longer does the variance transform linearly, but as a square root, in the vicinity of the critical point of the map. Let now $\mathcal{L}^\dagger$ transform this new density

$$\mathcal{L}^\dagger e^{-\alpha^2 z^4_{a-1}} = \int e^{\frac{(y-f(z_{a-1}))^2}{4D}-\alpha^2 y^4}[dy] \tag{97}$$

where $\alpha^2 = f''^2_{a-1}/8(\sigma_a^2+2D)$. Now make the change of variable $\xi = y\sqrt{\alpha/4D}$, and write the density $\rho_{a-1}(y)$ as a power series, so that the previous integral reads

$$\begin{aligned}\mathcal{L}^\dagger e^{-\alpha^2 z^4_{a-1}} &= \sqrt{\frac{4D}{\alpha}}\int[d\xi]e^{-\left(\frac{\xi}{\sqrt{\alpha}}-\frac{f(z_{a-1})}{\sqrt{4D}}\right)^2}\sum_{n=0}^{\infty}(-1)^n\frac{\left[(4D)^2\xi^4\right]^n}{n!} \\ &= \sum_{n=0}^{\infty}\frac{(-1)^n(4n)!}{n!}(\sqrt{\alpha}f(z_{a-1}))^{4n}\sum_{k=0}^{2n}\frac{1}{(4n-2k)!k!}\left(\frac{4D}{4f^2(z_{a-1})}\right)^k \end{aligned} \tag{98}$$

I then group all the terms up to order $O(D)$ and neglect $O(D^2)$ and higher,

$$\begin{aligned}&\sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\left(\sqrt{\alpha}f(z_{a-1})\right)^{4n} + 4D\sum_{n=0}^{\infty}\frac{(-1)^n(4n)!}{4\left[n!(4n-2)!\right]}\alpha^2 f^{4n-2}(z_{a-1}) = \\ &\qquad e^{-\alpha^2 f^4(z_{a-1})} - 4D\left[3\alpha^2 f^2(z_{a-1})\right]\Phi\left(\frac{7}{4},\frac{3}{4},-\left[\sqrt{\alpha}f(z_{a-1})\right]^4\right)\end{aligned} \tag{99}$$

where Φ is a confluent hypergeometric function of the first kind, sometimes called [1] $1F_1$:

$$\Phi\left(\frac{7}{4},\frac{3}{4},-\left[\sqrt{\alpha}f(z_{a-1})\right]^4\right) = e^{-\alpha^2 f^4(z_{a-1})}\left(1-\frac{4\alpha^2 f^4(z_{a-1})}{3}\right) \tag{100}$$

I now want to consider the density (99) in the vicinity of the preimage of z_{a-1}, which is expected to be far enough from the flat top of the map so as the linear approximation $f(z_{a-1}) \simeq f'_{a-2}z_{a-2}$ to be valid, and evaluate the variance

$$\sigma_{a-2} = \frac{\int dz_{a-2}z^2_{a-2}\rho_{a-2}(z_{a-2})}{\int dz_{a-2}\rho_{a-2}(z_{a-2})} \tag{101}$$

it is useful to know, in computing the denominator of (101), that

$$\int f'^2_{a-2} z^2_{a-2} \Phi\left(\frac{7}{4}, \frac{3}{4}, -[\sqrt{\alpha} f'_{a-2} z_{a-2}]^4\right) dz_{a-2} = 0 \tag{102}$$

so that

$$\begin{aligned}
\sigma_{a-2} &= \frac{\int z_{a-2} e^{-\alpha^2 f'^4_{a-2} z^4_{a-2}} dz_{a-2} - 4D[3\alpha^2 f'^2_{a-2} \int z^4_{a-2} \Phi\left(\frac{7}{4}, \frac{3}{4}, -[\sqrt{\alpha} f'_{a-2} z_{a-2}]^4\right) dz_{a-2}}{\int e^{-\alpha^2 f'^4_{a-2} z^4_{a-2}}} \\
&= \frac{\Gamma(3/4)}{\Gamma(1/4)} (\alpha^2 f'^4_{a-2})^{-1/2} - 4D \frac{(\alpha^2 f'^4_{a-2})^1/4}{2\Gamma(5/4)} 3\alpha^2 f'^2_{a-2} \left[-\frac{4\Gamma(9/4)}{15(\alpha^2 f'^4_{a-2})^{5/4}} \right] \\
&= \frac{1}{f'^2_{a-2}} \left(\frac{\Gamma(3/4)}{\Gamma(1/4)} \frac{1}{\alpha} + 2D \right) = \frac{\sigma^2_{a-1} + 2D}{f'^2_{a-2}}
\end{aligned} \tag{103}$$

in the last identity I used the definition of α and (96). I have just shown that the variance of the density $\rho_{a-1}(z_{a-1})$ transforms again like the variance of a Gaussian, up to order $O(D)$ in the noise strength. By the same procedure, one can again assume the next preimage of the map x_{a-3} is such that the linear approximation is valid, and transform the density $\rho_{a-2}(z_{a-2})$ (99) up to $O(D)$ and obtain the same result for the variance, that is

$$\sigma^2_{a-3} = \frac{\sigma^2_{a-1} + 2D(1 + f'^2_{a-2})}{f'^2_{a-2} f'^2_{a-3}} \tag{104}$$

which is nothing but the expression (45) for the evolution of the variances of Gaussian densities, obtained in Sect. 2.3. In other words, the evolution of the variances goes back to be linear, to $O(D)$, although the densities transformed from the 'quartic Gaussian' (95) are hypergeometric functions.

The question is now how to modify the definition of neighborhoods given in Sect. 3.2, in order to fit the new approximation. Looking for eigenfunctions of $\mathcal{L}^\dagger$ seems to be a rather difficult task to fulfill, given the functional forms (95) and (99) involved. Since all I really care of are the variances, I define instead the following map

$$\sigma^2_{a-1} = \begin{cases} C \left(\frac{\sigma^2_a + 2D}{f'^2_{a-1}} \right)^{1/2} & |f'^2_{a-1} < 1| \\ \frac{\sigma^2_a + 2D}{f'^2_{a-1}} & \text{otherwise,} \end{cases} \tag{105}$$

$C = 2\sqrt{2}\Gamma(3/4)/\Gamma(1/4)$, for the evolution of the densities, and take its periodic points as our new neighborhoods. In practice, one can compute these numerically, but I will not need

orbits longer than $n_p = 4$ in my tests of the partition. Therefore I assume only one periodic point of $f(x)$ to be close to the flat top, and obtain analytic expressions for the periodic points of (105):

$$\tilde{\sigma}_a^2 \simeq C \left(\frac{2D\left(1 + f_{a-1}'^2 + ... + (f_{a-n+1}^{n-1'})^2\right)^2}{\tilde{\Lambda}_p^2} \right)^{1/2} \tag{106}$$

with $\tilde{\Lambda}_p = f_{a-n+1}^{n-1'} f_a''$, is valid when the cycle ends at a point x_a close to the flat top. Otherwise, take the periodic point x_{a-k}, that is the kth pre-image of the point x_a. The corresponding periodic variance has the form

$$\tilde{\sigma}_{a-k}^2 \simeq \frac{1}{(f_{a-1}^{k'})^2} \left(2D(1 + f_{a-1}'^2 + ... + (f_{a-1}^{k-1'})^2) + \tilde{\sigma}_a^2 \right) \tag{107}$$

both expressions (106) and (107) are approximated, as we further assumed $2D\tilde{\Lambda}_p^2 >> 1$, which is reasonable when $D \in [10^{-4}, 10^{-2}]$, our range of investigation for the noise strength. As before, a neighborhood of width $[x_a - \tilde{\sigma}_a, x_a + \tilde{\sigma_a}]$ is assigned to each periodic point x_a, and an optimal partition follows. However, due to the geometry of the map, such partitions as

$$\{\mathcal{M}_{000}, [\mathcal{M}_{001}, \mathcal{M}_{011}], \mathcal{M}_{010}, \mathcal{M}_{110}, \mathcal{M}_{111}, \mathcal{M}_{10}\}. \tag{108}$$

can occur. In this example the regions $\mathcal{M}_{001}$ and $\mathcal{M}_{011}$ overlap, and the partition results in a transition graph with three loops (cycles) of length one, while we know that our map only admits two fixed points. In this case I decided instead to follow the deterministic symbolic dynamics and ignore that particular overlap.

Let me now validate the method by estimating once again the escape rate of the noisy map in Figure 16. We notice that the matrix elements

$$\begin{aligned} [\mathbf{L}_{ba}]_{kj} &= \langle \tilde{\rho}_{b,k} | \mathcal{L} | \rho_{a,j} \rangle \\ &= \int \frac{dz_b dz_a\, \beta}{2^{j+1} j! \pi \sqrt{D}} e^{-(\beta z_b)^2 - \frac{(z_b - f_a' z_a)^2}{4D}} \\ &\qquad \times H_k(\beta z_b) H_j(\beta z_a), \end{aligned} \tag{109}$$

should be redefined in the neighborhood of the critical point of the map, where the Gaussian approximation to $\mathcal{L}$ fails. Following the approximation made in (95),

$$[\mathbf{L}_{ba}]_{kj} = \int \frac{dz_b dz_a\, \beta}{2^{j+1} j! \pi \sqrt{D}} e^{-(\beta z_b)^2 - \frac{(z_b - f_a'' \sqrt{4D} z_a^2/2)^2}{4D}} H_k(\beta z_b) H_j(\beta z_a). \tag{110}$$

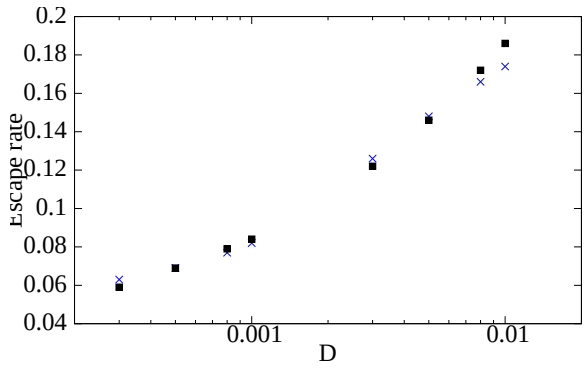


Figure 17: Escape rates of the 'skew' Ulam map vs. noise strength D, using: the optimal partition method ($\times$), and ($\blacksquare$) a uniform discretization (54) in $N = 128$ intervals.

However, as D decreases, it also reduces the quadratic term in the expansion of the exponential, so that the linear term $f'_a z_a$ must now be included in the matrix element:

$$[\mathbf{L}_{ba}]_{kj} = \int \frac{dz_b dz_a\, \beta}{2^{j+1} j! \pi \sqrt{D}} e^{-(\beta z_b)^2 - \frac{(z_b - f'_a z_a - f''_a \sqrt{4D} z_a^2/2)^2}{4D}} H_k(\beta z_b) H_j(\beta z_a)\,, \tag{111}$$

I find in the 'skew Ulam' model that the periodic orbits used in the expansion have x_a's near the flat top, such that $f'_a \sim 10^{-1}$ and $f''_a \sim 10$, and therefore the matrix element (110) would better be replaced with (111) when $D \sim 10^{-4}$. In order to know whether a cycle point is close enough to the flat top for the Gaussian approximation to fail, recall that the matrix element (109) is the zeroth-order term of a series in D, whose convergence can be probed by evaluating the higher order corrections (88): when the $O(\sqrt{D})$ and $O(D)$ corrections are of an order of magnitude comparable or bigger than the one of (109), I conclude that the Gaussian approximation fails and I use (110) or (111) instead. Everywhere else the usual matrix elements (109) are used, *without* the higher-order corrections, as they are significantly larger than in the case of the repeller, and they are not accounted for by the optimal partition method, which is entirely based on a zeroth-order Gaussian approximation of the evolution operator (cf. Sect. 2.2). Like before, I tweak the noise strength D within the range $[10^{-4}, 10^{-2}]$ and compare the escape rate evaluated with the optimal partition method and with the uniform discretization (54) The results are illustrated inFigure 17, the uniform discretization method and the method of the optimal partition are consistent within a $5-6\%$ margin. The results of Figure 17 are also reported in table 2, together with the number of intervals given by the optimal partition every time.

Table 2: Escape rates of the 'skew' Ulam map from the unit interval, calculated from the determinant of the graph of the optimal partition: $\gamma_{<>}$ is obtained using (109) and (111) for the matrix elements, with n_r indicating the number of regions of the corresponding optimal partition; γ_{num} is the escape rate obtained by uniform discretization of $\mathcal{L}$ ($N = 128$ intervals), for different values of D.

D	n_r	$\gamma_{<>}$	γ_{num}
0.01	7	0.174	0.186
0.008	7	0.166	0.172
0.005	7	0.148	0.146
0.003	7	0.126	0.122
0.001	13	0.082	0.084
0.0008	13	0.077	0.079
0.0005	14	0.069	0.069
0.0003	15	0.063	0.059

CHAPTER VII

SUMMARY AND OUTLOOK

I formulated the hypothesis of a finite resolution for the state space in the presence of white noise, and proposed an algorithm to determine the finest possible partition of a one-dimensional map. A Fokker-Planck kind of approach is at the basis of my analysis: a discrete-time evolution operator for densities of trajectories was derived, and then linearized in the neighborhood of periodic orbits of the deterministic system, as they still constitute the skeleton of the dynamics, if the noise is weak. The purpose of that is to obtain invariants of the Fokker-Planck operator and of its adjoint, to be used for partitioning the state space, in the same way as one uses periodic orbits of a deterministic map in the absence of noise. As it turns out, periodic points become Gaussian- or hypergeometric-shaped densities, which cover the state space of a chaotic map, until they overlap significantly within a 1σ confidence level: that sets the finest attainable resolution. The length of the longest orbits in the partition indicates the maximum number of iterations before the noisy system loses memory of where it has been. Not surprisingly, such memory depends on the interplay of the dynamics with the noise, and therefore is not uniform in the state space. The rest of the work presented is an attempt to test the optimal partition hypothesis, using periodic orbit expansions. Any evolution operator is forced into a finite-dimensional matrix, due to the finite resolution of its support. A local approximation of the Fokker-Planck operator allows us to compute such observables as the escape rate of the map or the Lyapunov exponent in a perturbative fashion, with the noise strength as order parameter. The results show satisfactory agreement with the ones obtained with a brute-force diagonalization of the evolution operator.

The future presents a number of challenges, both technical and conceptual. The first

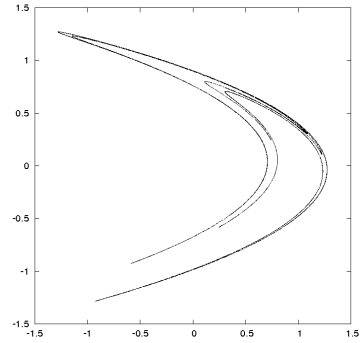
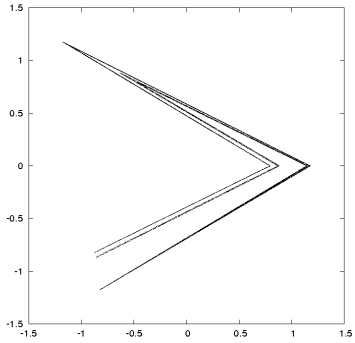
(a)  (b) 

Figure 18: (a)The Hénon attractor $(x', y') = (1 - ax^2 + y, bx)$, with $a = 1.4, b = 0.3$;(b) the Lozi attractor $(x', y') = (1 - a|x| + y, bx)$, with $a = 1.85, b = 0.3$.

thing that comes to mind is how the optimal partition hypothesis would extend to a higher-dimensional state space. Two main problems: (a) find the local spectrum of the Fokker-Planck evolution operator and (b) define an interval, now having to account for both expanding and *contracting* directions of the deterministic dynamics.

(a) Some preliminary work on the two-dimensional Hénon and Lozi maps (Figure 18) with isotropic noise, shows that both the local Fokker-Planck operator

$$\mathcal{L}\rho(\mathbf{y}) = \int \exp\left(-\frac{(\mathbf{y}-\mathbf{A}\mathbf{x})^T(\mathbf{y}-\mathbf{A}\mathbf{x})}{4D} - \mathbf{x}^T\mathbf{Q}\mathbf{x}\right) d^2x, \qquad A_{ij} = \frac{\partial f_i(x)}{\partial x_j} \tag{112}$$

and its adjoint admit Gaussian ground-state eigenfunctions, whose quadratic forms $\mathbf{x}^T\mathbf{Q}\mathbf{x}$ at the exponential are all degenerate, meaning

$$\mathbf{U}^T\mathbf{Q}\mathbf{U} = \mathbf{\Lambda}_\mu = \begin{pmatrix} 0 & 0 \\ 0 & \mu^2 \end{pmatrix} \tag{113}$$

and the eigenfunction is a Gaussian 'tube' of variance $\sigma^2 = 1/2\mu^2$, portrayed in Figure 19. The local eigenfunction of $\mathcal{L}$ ($\mathcal{L}^\dagger$) around a periodic point of a piecewise-linear map extends along the direction of the unstable (stable) eigenvector of the Jacobian of the cycle, namely the unstable (stable) manifold. In nonlinear maps these directions are somewhat skewed from the stable and unstable manifolds, but the picture remains similar.

(b) So now here comes an idea for how to define an interval of the partition: take both ground-state local eigenfunctions of $\mathcal{L}$ and $\mathcal{L}^\dagger$ and cross them as shown in Figure 19, then take the intersection of their supports (within a 1σ confidence) as the interval $\mathcal{M}$. The next

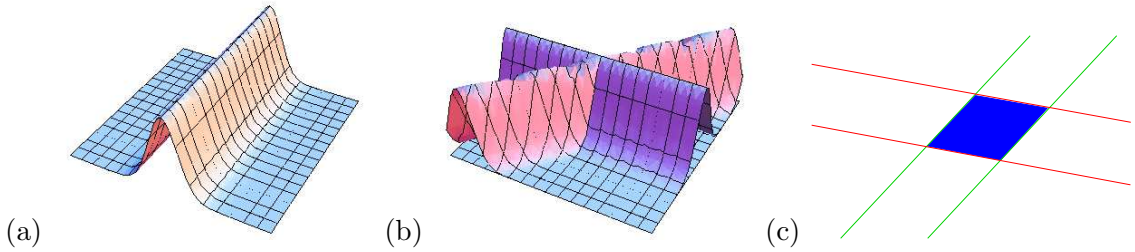


Figure 19: (a) A ground-state eigenfunction of the local two-dimensional Fokker-Planck operator (112); (b) the ground-state eigenfunctions of $\mathcal{L}$ and of its adjoint $\mathcal{L}^\dagger$, both operator linearized around the same cycle point; (c) my definition of partition interval in two dimensions: take the local densities in (b), cut off their supports at 1σ and take their intersections.

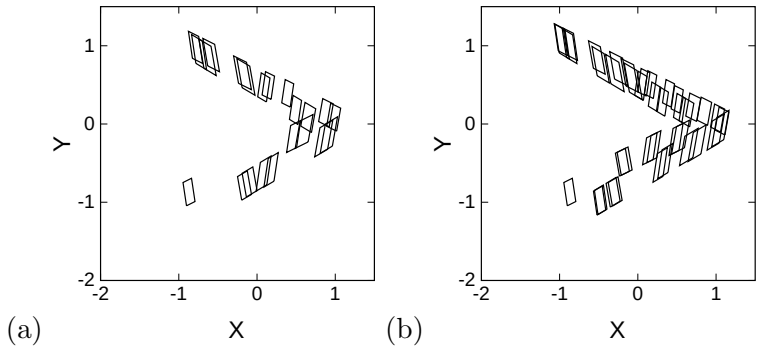


Figure 20: Intervals defined as in Figure 19 cover the Lozi attractor in Figure 18, using: (a) all periodic points of length five; (b) all periodic points of length six.

step is akin to what seen in one dimension, that is just cover the non-wandering set of the system with all the intervals found until they significantly overlap (Figure 20). A couple of technical issues still need to be addressed, before this idea can be tested in the same way as in the one-dimensional maps. First, general overlapping of a generation of periodic points (i.e. of a certain length) is not enough to determine the optimal partition: look at the Lozi attractor in Figure 18 and then at the overlapping neighborhoods in Figure 20: some parts of the attractor are still uncovered, even when the optimal resolution seems achieved, meaning one needs to find more periodic orbits. Unlike in the repeller, longer cycles are not always shadowed by shorter ones, due to forbidden sequences in the binary tree (cf. Sect. 3.1) that regulates the topology of the orbits. This phenomenon, known as *pruning* [13], is a science of its own. Arguably, one needs to understand how the evolution of forbidden sequences (*pruning fronts*) in symbolic space maps into the state space, before attempting

to fit the optimal partition method to two-dimensional attractors. The second issue concerns symmetries of the periodic orbits, such as, for instance, the exchange $(x, y) \to (y, x)$ for maps of the Hénon type, and whether those carry over to the local eigenfunctions of $\mathcal{L}$ and $\mathcal{L}^\dagger$. Taking symmetries into account would imply working in a reduced state space (*fundamental domain*), where fewer cycles suffice to understand the dynamics.

It will be necessary, sooner or later, to understand how noise convolves with continuous-time dynamical systems in d dimensions, from the perspective of a Poincaré-type of analysis. In other words, given the path integral:

$$\mathcal{L} \circ \rho(x_0, t_0) = \lim_{n\to\infty} \int [dx] e^{-[x_{i+1}-x_i-f(x_i)\delta t_i]^T \mathbf{\Delta}^{-1} [x_{i+1}-x_i-f(x_i)\delta t_i]} \rho(x_0, t_0) \tag{114}$$

where $\mathbf{\Delta}$ is the diffusion matrix, equal to just $4D\delta t_i$ is one dimension, and

$$[dx] = \prod_{i=0}^{n-1} \frac{dx_i^d}{(4\pi \det(\mathbf{\Delta})\, \delta t_i)^{d/2}}.$$

We want to establish whether, and to what extent, it is legal to split the previous into a product of two path integrals, one along the direction of the orbit, and the others along the directions normal to the orbit, in its *local* reference frame:

$$\begin{aligned}\mathcal{L} \circ \rho(x_0, t_0) \approx \lim_{n\to\infty} \int [dx_\parallel] e^{-[x_{i+1}-x_i-f(x_i)\delta t_i]_\parallel^T \mathbf{\Delta}_\parallel^{-1} [x_{i+1}-x_i-f(x_i)\delta t_i]_\parallel} \rho(x_{0\parallel}, t_0) \times \\ \lim_{n\to\infty} \int [dx_\perp] e^{-[x_{i+1}-x_i-f(x_i)\delta t_i]_\perp^T \mathbf{\Delta}_\perp^{-1} [x_{i+1}-x_i-f(x_i)\delta t_i]_\perp} \rho(x_{0\perp}, t_0)\,.\end{aligned} \tag{115}$$

One can reasonably hope that the previous is true for small, isotropic noise, when the diffusion matrix $\mathbf{\Delta}$ is diagonal, while the whole issue gets more complicated when the noise is anisotropic. Either way, the effect of the noise in the direction of traveling ($\parallel$) must be understood, whether it only affects the return time on a fixed Poincaré section, or it is entangled with the diffusion in the normal directions so as to alter the distribution of the noise on the surface of section.

I believe any further development of the optimal partition hypothesis must go through these steps.

APPENDIX A

MARKOV PROCESSES

In general, the joint probability of a stochastic process $p(x_1, t_1; x_2, t_2; ...; x_n, t_n)$ is expressed as a function of the initial probability $p(x_1, t_1)$ and the conditional probabilities for the successive steps:

$$p(x_1, t_1; x_2, t_2; ...; x_n, t_n) = p(x_1, t_1)p(x_2, t_2|x_1, t_1)...p(x_n, t_n|x_1, t_1; x_2, t_2; ...; x_{n-1}, t_{n-1}) \tag{116}$$

The process is said to be Markovian if, for any $t_1 < t_2 < ... < t_n$ and for any n,

$$p(x_n, t_n|x_1, t_1; x_2, t_2; ...; x_{n-1}, t_{n-1}) = p(x_n, t_n|x_{n-1}; t_{n-1}) \tag{117}$$

In words, once it has arrived at x_{n-1} at time t_{n-1}, a Markov process evolves further irrespective of its history before x_{n-1}. This way, equation (116) is rewritten as

$$p(x_1, t_1; x_2, t_2; ...; x_n, t_n) = p(x_1, t_1)p(x_2, t_2|x_1, t_1)...p(x_n, t_n|x_{n-1}, t_{n-1}) \tag{118}$$

Suppose we know x_1 and x_3, but nothing in between. Then what would be in general

$$p(x_3, t_3|x_1, t_1) = \int p(x_2, t_2|x_1, t_1)p(x_3, t_3|x_1, t_1; x_2, t_2)dx_2 \tag{119}$$

becomes, for a Markov process

$$p(x_3, t_3|x_1, t_1) = \int p(x_2, t_2|x_1, t_1)p(x_3, t_3|x_2, t_2)dx_2 \tag{120}$$

which is known as Chapman-Kolmogorov equation.

Now consider an initial distribution $\rho(x_0, t_0)$, which evolves into

$$\rho(x, t) = \int p(x, t|x_0, t_0)\rho(x_0, t_0)dx_0 \tag{121}$$

and, at a later time, into

$$\rho(x, t + \Delta t) = \int p(x, t + \Delta t|x_0, t_0)\rho(x_0, t_0)dx_0 \tag{122}$$

In the following, I will relate $\rho(x, t + \Delta t)$ to $\rho(x, t)$ using the fact that the process is markovian: the Chapman-Kolmogorov equation (120) reads, in this case,

$$p(x, t + \Delta t|x_0, t_0) = \int p(x', t|x_0, t_0)p(x, t + \Delta t|x', t)dx' \tag{123}$$

Now plug (123) into (122) and get the double integral

$$\begin{aligned} \rho(x, t + \Delta t) &= \int\int p(x', t|x_0, t_0)p(x, t + \Delta t|x', t)\rho(x_0, t_0)dx_0dx' = \\ &\int \rho(x', t)p(x, t + \Delta t|x', t)dx' \end{aligned} \tag{124}$$

If the process is stationary, that is $p(x, t) = p(x, t + T)$ for any fixed T, then

$$\rho(x, t + \Delta t) = \int \rho(x', t)p(x, t + \Delta t|x', t)dx' \tag{125}$$

as stated in Sect. 1.2

APPENDIX B

ORNSTEIN-UHLENBECK PROCESS

Here I reproduce results obtained by Dekker and Van Kampen [18] and later by Gaspard *et al.* [25].

Add weak uncorrelated noise $\hat{\xi}$ to the linear ODE

$$\frac{dx}{dt} = \lambda x$$

with a single equilibrium solution $x = 0$. The corresponding Langevin equation[1]

$$\frac{dx}{dt} = \lambda x + \hat{\xi} \tag{126}$$

leads to the Fokker-Planck equation [18, 25]

$$\partial_t \rho(x,t) + \partial_x[\lambda x \rho(x,t)] = D\,\partial_x^2 \rho(x,t)\,, \tag{127}$$

known as the Ornstein-Uhlenbeck process [47, 51, 42]. The analytical solution is obtained by rewriting ρ as

$$\rho = e^{-\frac{U}{2D}}\psi\,, \tag{128}$$

where $U = -\frac{\lambda}{2}x^2$ can be interpreted as the potential of the Langevin force in (126). The multi-dimensional Ornstein-Uhlenbeck process with potential

$$U = -\frac{1}{2}\sum_{i,j} A_{i,j} x_i x_j \tag{129}$$

is known in financial literature as the Vašiček model [50]. The equation for ψ has the Schrödinger form, with the quantum harmonic oscillator Hamiltonian:

$$\begin{aligned} -\partial_t \psi &= H\psi \\ H &= -D\partial_x^2 - \frac{1}{2}\partial_x^2 U + \frac{1}{4D}(\partial_x U)^2 = -D\partial_x^2 + \frac{\lambda}{2} + \frac{\lambda^2}{4D}x^2\,. \end{aligned} \tag{130}$$

[1] In papers by Ornstein and Uhlenbeck [47] and Chandrasekhar [7], the Langevin equation has a velocity as the 'x' variable.

A solution of the continuous time Fokker-Planck equation (127) can be expanded in the harmonic oscillator eigenfunction basis as

$$\rho(x,t) = \sum_{k=0}^{\infty} C_k \psi_k(x) e^{-s_k t}\,, \tag{131}$$

where ψ_k and s_k are respectively the eigenfunctions and eigenvalues of the time-independent Schrödinger equation

$$H\psi_k = -s_k \psi_k\,. \tag{132}$$

The solutions are [25]

$$\tilde{\psi}_k(x) = H_k(\mu x) e^{-(\mu x)^2}\,, \qquad \mu^2 = -\lambda/2D\,, \qquad s_k = -k\lambda \tag{133}$$

in the attracting case ($\lambda < 0$), and

$$\psi_k(x) = H_k(\mu x)\,, \qquad s_k = (k+1)\lambda\,, \tag{134}$$

in the repulsive case ($\lambda > 0$).

APPENDIX C

GAUSSIAN EVOLUTION IN CONTINUOUS TIME

Given a flow

$$\frac{dx}{dt} = f(x)\,, \tag{135}$$

the matrix of variations

$$A_{ij} = \frac{\partial f_i(x)}{\partial x_j} \tag{136}$$

describes the instantaneous rate of shearing of the infinitesimal neighborhood of $x(t)$ by the flow (135). Obviously one needs to integrate (136), together with the flow, in order to determine the amount of deformation $\mathbf{J}^t(x)$ of an infinitesimal neighborhood after a finite time t in the co-moving frame of $x(t)$:

$$\frac{d}{dt}\mathbf{J}^t(x) = \mathbf{A}(x)\mathbf{J}^t(x), \qquad\qquad \mathbf{J}^0(x) = \mathbf{1} \tag{137}$$

This concept will be used in the following, in order to shed some light on the meaning of formula (46),

$$\sigma_a^2 = \frac{2D}{1-\Lambda^2}\,\left(1 + \sum_{i=1}^{i=n-1} (f_{a+i}^{n-i'})^2\right)\,, \qquad \Lambda = f_a^{n'}$$

for the condition on the width of a Gaussian, in order for the latter to be an eigenfunction of the Fokker-Planck operator $\mathcal{L}$ in the neighborhood of a periodic point x_a. The idea is to see whether the formula becomes more familiar in the continuous-time limit. From Sect. 1.3, I can write the evolution of a Gaussian density by the Fokker-Planck operator of a flow, near a periodic point x_a, in a time step Δt:

$$\begin{aligned} \rho(z_{a+1}) \;\; &= \;\; \int \exp\left(-\frac{(f_a' z_a \Delta t + z_a - z_{a+1})^2}{4D\Delta t}\right) e^{-\frac{z_a}{2\sigma_a^2}} [dz_a] = \\ & \qquad c_{a+1} \exp\left(-\frac{z_{a+1}^2}{2(2D\Delta t + (f_a'\Delta t + 1)^2\sigma_a^2)}\right) \end{aligned} \tag{138}$$

where I have used the notation introduced in Sect. 2.2 for the local coordinates near the periodic point x_a and its image $x_{a+1} = f(x_a)$. Thus, widths map as

$$\sigma_{a+1}^2 = 2D\Delta t + ((f_a^{'}\Delta t + 1)\sigma_a)^2\,. \tag{139}$$

from which it is easy to obtain a condition on the width for the Gaussian to be an eigenfunction of the n-th iterate of $\mathcal{L}$:

$$\sigma_a^2 = \frac{2D\Delta t}{1-\Lambda^2}\left(1+\sum_{i=1}^{i=n-1}\prod_{1}^{i}(1+f'_{a+i}\Delta t)^2\right), \qquad \Lambda = \prod_{0}^{n-1}(1+f'_{a+i}\Delta t)\,. \tag{140}$$

Let me write out the term Λ after one, two, three iterations:

$$\begin{aligned}
\Lambda_1 &= 1+f'_a\Delta t \qquad (141)\\
\Lambda_2 &= 1+(f'_a+f'_{a+1})\Delta t + f'_a f'_{a+1}\Delta t^2\\
\Lambda_3 &= 1+(f'_a+f'_{a+1}+f'_{a+2})\Delta t + (f'_a f'_{a+1}+f'_a f'_{a+2}+f'_{a+1}f'_{a+2})\Delta t^2 + f'_a f'_{a+1}f'_{a+2}\Delta t^3
\end{aligned}$$

Recalling definitions (136), (137) and some simple algebra is all it takes to discover that (142) is nothing but the integration (137), discretized in Δt-time intervals to obtain the Jacobian of a one-dimensional flow, therefore

$$\Lambda \to J = e^{\int_0^t f'(x(\tau))d\tau} \tag{142}$$

Next, consider the expression 140 and rewrite the summation

$$\left(1+\sum_{i=1}^{i=n-1}\prod_{1}^{i}(1+f'_{a+i}\Delta t)^2\right)\Delta t = \left(1-\Lambda^2+\sum_{i=1}^{i=n}\prod_{1}^{i}(1+f'_{a+i}\Delta t)^2\right)\Delta t \to \int J(x(t))^2 dt \tag{143}$$

where I just took the continuous limit $\Delta t \to 0$ and $n \to \infty$: everything goes to zero except for the summation multiplied by the time step. Now I have an expression for the local eigenfunction (138) in continuous-time:

$$\rho(z_a(t)) = c(t)\exp\left(-\frac{1-J(t)^2}{4D\int J(t)^2 dt}\right) \tag{144}$$

As a rapid crosscheck, let me see what that exponent becomes for the Ornstein-Uhlenbeck process, where the matrix of variation is a constant λ (cf. Sect. B):

$$\frac{1-J(t)^2}{4D\int J(t)^2 dt} = \frac{1-e^{\lambda t}}{4D\int_0^t e^{2\lambda\tau}d\tau} = \frac{\lambda}{2D} \tag{145}$$

which is exactly the exponent (133) of the ground-state eigenfunction of the attracting case ($\lambda < 0$).

APPENDIX D

FOKKER-PLANCK OPERATOR AND TIME REVERSIBILITY

The Fokker-Planck evolution operator on the right-hand side of equation (9)

$$\mathcal{L}_D\rho(x,t) = -\partial_x[f(x)\rho(x,t)] + D\partial_{xx}\rho(x,t) \tag{146}$$

is not symmetric under the operation of time-reversal. That is apparent from the solution (131) of the one-dimensional Ornstein-Uhlenbeck problem: all the eigenfunctions (except for the invariant measure) decay for $t \to \infty$, but they would diverge to infinity if $t \to -t$.

Formally, a time-evolution operator $U(t, t_0)$ satisfies the equation (cf. [44])

$$\frac{\partial}{\partial t}[U(t,t_0)\rho] = \mathcal{L}_D[U(t,t_0)\rho] \tag{147}$$

$\mathcal{L}_D$ does not depend on time explicitly, therefore

$$U(t,t_0) = e^{-\mathcal{L}_D(t-t_0)} \tag{148}$$

In order to test whether $U(t, t_0)$ is symmetric under time reversal, let us write its expression for short times:

$$U(t_0+dt,t_0) = 1 - \mathcal{L}_D dt \tag{149}$$

Then $U(t_0 + dt, t_0)$ (and hence $U(t, t_0)$) is time-reversible iff

$$U^\dagger(t_0+dt,t_0)U(t_0+dt,t_0) = 1 \tag{150}$$

Plugging (148) into the previous condition, we obtain that $U(t, t_0)$ is T-symmetric iff

$$\mathcal{L}_D = -\mathcal{L}_D^\dagger \tag{151}$$

that is the Fokker-Planck operator should be *antihermitian*, in order for the evolution operator of the Fokker-Planck equation to be symmetric under time reversal. The adjoint of $\mathcal{L}_D$ is

$$\mathcal{L}_D^\dagger = D\partial_{xx} + \lambda x\partial_x \tag{152}$$

so that $\mathcal{L}_D$ is not antihermitian and the evolution operator (148) cannot be T-symmetric.

APPENDIX E

PERRON-FROBENIUS VS. KOOPMAN OPERATOR

The Perron-Frobenius operator

$$\mathcal{L}\rho(x) = \rho(f^{-1}(x)) \left| \frac{df^{-1}(x)}{dx} \right| = \rho(f^{-1}(x)) \left| \frac{1}{f'(f^{-1}(x))} \right| \tag{153}$$

transforms a density supported on a set A to a density supported on $f(A)$. In fact, if

$$\rho(f^{-1}(x)) = 0 \qquad \text{if} \qquad f^{-1}(x) \notin A,$$

then $x \in f(A)$.

On the other hand, the Koopman operator

$$\mathcal{L}^{\dagger}\rho(x) = \rho(f(x)) \tag{154}$$

transforms a density supported on A to a density supported on $f^{-1}(A)$. In fact, if

$$\rho(f(x)) = 0 \qquad \text{if} \qquad f(x) \notin A,$$

then $x \in f^{-1}(A)$. In this sense, the meaning of the Koopman operator is to map a density ρ(x) backwards in time.

This is different from applying the Perron-Frobenius operator associated to the inverse map $f^{-1}(x)$:

$$\mathcal{L}_{f^{-1}}\rho(x) = \rho(f(x)) \left| \frac{df(x)}{dx} \right| \tag{155}$$

if the map $f(x)$ is supported on A, then the transformed density is supported on $f^{-1}(A) \bigcap A$.

REFERENCES

[1] Abramowitz, M. and Stegun, I. A., eds., *Handbook of Mathematical Functions*. New York: Dover, 1964.

[2] Artuso, R., Rugh, H. H., and Cvitanović, P. Chapter *"Why does it work?"*, in ref. [11].

[3] Barrow-Green, J., *Poincaré and the Tree Body Problem*. Providence R.I.: Amer. Math. Soc., 1997.

[4] Bollt, E., Góra, P., Ostruszka, A., and Życzkowski, K., "Basis Markov partitions and transition matrices for stochastic systems," *SIAM J. Applied Dynam. Systems*, vol. 7, pp. 341–360, 2008. `arXiv:nlin/0605017`.

[5] Buhl, M. and Kennel, M. B., "Statistically relaxing to generating partitions for observed time-series data," *Phys. Rev. E*, vol. 71, p. 046213, 2005.

[6] Cartwright, M. L. and Littlewood, J. E., "On nonlinear differential equations of the second order," *J. London Math. Soc.*, vol. 20, pp. 180–189, 1945.

[7] Chandrasekhar, S., "Stochastic problems in physics and astronomy," *Rev. Mod. Phys.*, vol. 15, p. 1, 1943.

[8] Crutchfield, J. P. and Packard, N. H., "Symbolic dynamics of noisy chaos," *Physica D*, vol. 7, pp. 201–223, 1983.

[9] Cvitanović, P. Chapter *"Charting the state space"*, in ref. [11].

[10] Cvitanović, P. Chapter *"Averaging"*, in ref. [11].

[11] Cvitanović, P., Artuso, R., Mainieri, R., Tanner, G., and Vattay, G., *Chaos: Classical and Quantum*. Copenhagen: Niels Bohr Institute, 2010. `ChaosBook.org`.

[12] Cvitanović, P., Artuso, R., Rondoni, L., and Spiegel, E. A. Chapter *"Transporting densities"*, in ref. [11].

[13] Cvitanović, P., Gunaratne, G. H., and Procaccia, I., "Topological and metric properties of Hénon-type strange attractors," *Phys. Rev. A*, vol. 38, p. 1503, 1988.

[14] Cvitanović, P., Søndergaard, N., Palla, G., Vattay, G., and Dettmann, C. P., "Spectrum of stochastic evolution operators: local matrix representation approach," *Phys. Rev. E*, vol. 60, p. 3936, 1999. `arXiv:chao-dyn/9904027`.

[15] Cvitanović, P., "Invariant measurement of strange sets in terms of cycles," *Phys. Rev. Lett.*, vol. 61, p. 2729, 1988.

[16] Davidchack, R., Lai, Y., Bollt, E., and Dhamala, M., "Estimating generating partitions of chaotic systems by unstable periodic orbits," *Phys. Rev. E*, vol. 61, p. 1353, 2000.

[17] Daw, C. S., Finney, C. E. A., and Tracy, E. R., "A review of symbolic analysis of experimental data," *Rev. Sci. Instrum.*, vol. 74, pp. 915–930, 2003.

[18] Dekker, H. and Kampen, N. G. V., "Eigenvalues of a diffusion process with a critical point," *Phys. Lett. A*, vol. 73, p. 374, 1979.

[19] Dellnitz, M. and Junge, O., "An adaptive subdivision technique for the approximation of attractors and invariant measure," *Computing and Visualization in Science*, vol. 1, pp. 63–68, 1998.

[20] Devaney, R. L., *An Introduction to Chaotic Dynamical systems*. Red-wood City: Addison-Wesley, 1989.

[21] Froyland, G., "Computer-assisted bounds for the rate of decay of correlation," *Commun. Math. Phys.*, vol. 189, pp. 237–257, 1997.

[22] Froyland, G., "Approximating physical invariant measures of mixing dynamical systems in higher dimensions," *Nonlinear Analysis, Theory, Methods and Applications*, vol. 32, pp. 831–860, 1998.

[23] Froyland, G., "Extracting dynamical behaviour via Markov models," in *Nonlinear dynamics and statistics: Proc. Newton Inst., Cambridge 1998* (Mees, A., ed.), (Boston), pp. 281–321, Birkhäuser, 2001.

[24] Froyland, G., "On Ulam approximation of the isolated spectrum and eigenfunctions of hyperbolic maps," *Discrete and continuous dynamical systems*, vol. 17, pp. 671–689, 2007.

[25] Gaspard, P., Nicolis, G., Provata, A., and Tasaki, S., "Spectral signature of the pitchfork bifurcation: Liouville equation approach," *Phys. Rev. E*, vol. 51, p. 74, 1995.

[26] Gilmore, R. and Lefranc, M., *The topology of chaos*. New York: Wiley, 2002.

[27] Guder, R. and Kreuzer, E., "Control of adaptive refinement technique technique of generalized cell mapping by system dynamics," *Nonlinear Dynamics*, vol. 20, pp. 21–32, 1999.

[28] Keane, M., Murray, R., and Young, L.-S., "Computing invariant measures for expanding circle maps," *Nonlinearity*, vol. 11, pp. 27–46, 1998.

[29] Kennel, M. B. and Buhl, M., "Estimating good discrete partitions from observed data: Symbolic false nearest neighbors," *Phys. Rev. Lett.*, vol. 91, p. 084102, 2003. `arXiv:nlin/0304054`.

[30] Kitchens, B., ed., *Symbolic ynamics. One-sided, wo-sided, and ountable Markov shifts*. Berlin: Springer, 1998.

[31] Kovalevskaia, S., "Sur le problème de la rotation d'un corps solide d'un point fixe," *Acta Math.*, vol. 12, p. 177, 1889.

[32] Lasota, A. and MacKey, M., *Chaos, Fractals, and Noise; Stochastic Aspects of Dynamics*. Berlin: Springer, 1994.

[33] Lehrman, M., Rechester, A. B., and White, R. B., "Symbolic analysis of chaotic signals and turbulent fluctuations," *Phys. Rev. Lett.*, vol. 78, p. 54, 1997.

[34] Lorenz, E. N., "Deterministic nonperiodic flow," *J. Atmos. Sci.*, vol. 20, p. 130, 1963.

[35] May, R., "Simple mathematical models with very complicated dynamics," *Nature*, vol. 261, pp. 459–467, 1976.

[36] Nicolis, G., "Chaotic dynamics and markovian coarse-graining in nonlinear dynamical systems," 1990. in *Noise and Chaos in Nonlinear Dynamical Systems*, edited by F. Moss, L. A. Lugiato, and W. Schleich (Cambridge University Press).

[37] Plumecoq, J. and Lefranc, M., "From template analysis to generating partitions I: Periodic orbits, knots and symbolic encodings," *Physica D*, vol. 144, p. 231, 2000.

[38] Plumecoq, J. and Lefranc, M., "From template analysis to generating partitions II: Characterization of the symbolic encodings," *Physica D*, vol. 144, p. 259, 2000.

[39] Poincaré, H., *Les méthodes nouvelles de la méchanique céleste*. Paris: Guthier-Villars, 1899. For a very readable exposition of Poincaré's work and the development of the dynamical systems theory up to 1920's see ref. [3].

[40] Rechester, A. B. and White, R. B., "Symbolic kinetic analysis of two-dimensional maps," *Phys. Lett. A*, vol. 158, pp. 51–56, 1991.

[41] Rechester, A. B. and White, R. B., "Symbolic kinetic-equation for a chaotic attractor," *Phys. Lett. A*, vol. 156, pp. 419–424, 1991.

[42] Risken, H., *The Fokker-Planck Equation*. New York: Springer, 1996.

[43] Ruelle, D., *Statistical Mechanics, Thermodynamic Formalism*. Reading, MA: Addison-Wesley, 1978.

[44] Sakurai, J. J., *Modern Quantum Mechanics*. Reading, Mass.: Addison-Wesley-Longman, 1994.

[45] Takada, H., Kitaoka, Y., and Shimizua, Y., "Mathematical index and model in stabirometry," *Forma*, vol. 16, pp. 17–46, 2001.

[46] Tang, X. Z., Tracy, E. R., Boozer, A. D., deBrauw, A., and Brown, R., "Symbol sequence statistics in noisy chaotic signal reconstruction," *Phys. Rev. E*, vol. 51, pp. 3871–3889, 1995.

[47] Uhlenbeck, G. E. and Ornstein, L. S., "On the theory of the Brownian motion," *Phys. Rev.*, vol. 36, pp. 823–841, 1930.

[48] Ulam, S. M., *A Collection of Mathematical Problems*. New York: Interscience Publishers, 1960.

[49] van Kampen, N. G., *Stochastic Processes in Physics and Chemistry*. Amsterdam: North-Holland, 1992.

[50] Vasicek, O., "An equilibrium characterization of the term structure," *Journal of Financial Economics*, vol. 5, pp. 177–188, 1977.

[51] Wax, E., *Selected Papers on Noise and Stochastic Processes*. New York: Dover, 1954.