State space partitioning: Il Buono, il brutto, il cattivo

↓PRIVATE
↑PRIVATE

D. Lippolis and P. Cvitanović

Center for Nonlinear Science, School of Physics,

Georgia Institute of Technology, Atlanta, GA 30332-0430

(Dated: December 2, 2007)

Abstract

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. Our goal is to determine the finest possible partition of the state space for a given hyperbolic dynamical system and a given weak additive white noise of specified strength. We test these ideas on two models: the "skew Ulam" map and the Lozi attractor. In the Ulam case, we achieve an optimal partition by tracking noisy preimages of the critical point. In the Lozi case we compute the local eigenfunctions of the Fokker-Planck operator in the neighborhood of each periodic point, and use their widths to attain an optimal partition of the state space. In either case, the finest attainable partition for a given noise covers the state space by a finite tiling, and the Fokker-Planck evolution is represented by a finite Markov graph.

[add verbiage from nsf05th]

PACS numbers: 05.45.-a, 45.10.db, 45.50.pk, 47.11.4j

FIX these pacs - list their descriptions as well as a commented list. Identify keywords, list them.

I. INTRODUCTION

[66] [67] [68]

The set of unstable periodic orbits forms the skeleton of the chaotic state space explored by the long-time trajectories of an ergodic flow. This skeleton can be used to partition the state space into a set of regions, each region a neighborhood of a periodic point. For deterministic flows such partitions yield - via periodic orbit theory - all long-time dynamical averages, such as Lyapunov exponents, escape rates, and correlations [1, 2].

Longer and longer cycles yield finer and finer partitions as the neighborhood of each unstable cycle p shrinks exponentially as $1/|\Lambda_p|$, the product of cycle's expanding stability eigenvalues, while the number of partitions grows exponentially with the cycle length. Since there is an infinite number of cycles, with each neighborhood shrinking asymptotically to a point, a deterministic chaotic system can - in principle - be resolved arbitrarily finely. However, any physical system experiences a background noise, any numerical prediction is made with finite accuracy, with the finite computational precision playing the role of noise, and any set of equations models nature up to a given finite accuracy, since degrees of freedom are always neglected. In reality, a dynamical system cannot have infinitely fine state space resolution, as noise will always wash out the fine details. [69]

If the noise is weak and uncorrelated in time, the short-time dynamics is not altered significantly: short periodic orbits of the associated deterministic system can still be used to coarsely partition the state space and estimate dynamical averages. The weak-noise corrections to the spectrum of evolution operators were treated in a triptych of articles [3–5], by three different methods. They can be computed perturbatively to a remarkably high order [6] in the noise strength D. [70] However, it turns out that the eigenvalues of such operators offer no guidance to the "finite resolution" problem, which does not appear to have been addressed in literature (literature mostly deals with estimating partitions from

[66] PC: change style to Nonlinearity
[67] PC: fix bibstyle to Nonlinearity
[68] PC: introduce colors
[69] Domenico: Can the limit of resolution solve the inverse problem: if I know the resolution of the state space from an experiment, can I get the strength of the noise by tuning D in my simulation? PC: I cannot imagine how one would set up such experiment...?
[70] PC: does Gaspard state a noisy trace formula? DL: Gaspard does state a noisy trace formula, I'm still reading his 2002 paper on the subject. PC: should his results be incorporated here?

observed data [7–10]).

Intuitively, the noise smears out the neighborhood of each periodic point, whose size is now determined by the interplay between the diffusive spreading, linear with time and parameterized [11] by the diffusion constant D, and the deterministic exponential contraction/expansion, characterized by the stability eigenvalues $|\Lambda_{p,i}|$. In the long run the diffusion always wins, and every sequence of successive refinements of deterministic partition of the state space stops at the finest attainable partition, the partition beyond which the diffusive smearing of partition borders exceeds the size of any further deterministic subpartition. Ulam [12, 13] conjectured that successive refinement of the coarse-graining would provide a convergent sequence of finite-state Markov approximations to the Perron-Frobenius operator.

For nonlinear chaotic flows, the effective diffusion rate differs from trajectory to trajectory, meaning that different neighborhoods merge at different times, so there is no one "Ehrenfest time" beyond which noise takes over; rather, as we shall show here, the optimal partition has to be computed for each dynamical system and each given time-independent stochastic noise density distribution. [71]

This is work, but one is handsomely rewarded. First and foremost, as the optimal partition is always finite, the Fokker-Planck operator can always be represented by a finite matrix. The dynamics of this best possible of all partitions is encoded by a Markov graph of finite memory, which can be used to write a finite-dimensional spectral determinant. Second, while the state space of a generic deterministic flow is an infinitely interwoven hierarchy of chaotic, attracting, elliptic and parabolic regions, the noisy dynamics erases any structures finer than the optimal partition, thus curing both the affliction of long-period attractors/elliptic islands with very small immediate basins of attraction, and the power-law correlation decays caused by marginally stable regions of state space.

⇓PRIVATE

We are motivated in part by the parallel effort in the semiclassical periodic orbit theory, where the Heisenberg and Ehrenfest times play the key roles. In the deep classical limit (wavelength much less than system size) the Ehrenfest time (the time for a wavepacket to spread to a classical size, and for interference effects to show up in Wigner density

[71] PC: remark somewhere that noise breaks smooth conjugation invariance of dynamics. Looking for the effect in eigenvalue spectrum is not smart, because the map needs nonlinearity higher than quadratic for the effect to show up, but on the level of partitions we already see it in the local Gaussian approximation.

distributions) plays a crucial role (discussed in Haake [14]). Ehrenfest time characterizes the departure of quantum dynamics observables from their classical dynamics values. The size of wavepacket grows in unstable directions as $\Delta q \exp(\lambda t)$, where λ is the mean Lyapunov exponent. The Ehrenfest time

$$T_{Ehr} = \frac{1}{\lambda} \ln \frac{\Delta}{\hbar} \,. \tag{1}$$

gives an estimate for the limiting time above which the periodic orbits approximation fails (due to constructive interference between the periodic orbits).

The Ehrenfest time is to be contrasted with the 'Heisenberg time', the time needed to resolve the discreteness of the quasi-energy spectrum. The Heisenberg time is defined by the inverse of the mean level spacing:

$$T_H = \hbar/\Delta E \approx T_0/\hbar^{D-1} \,, \tag{2}$$

where $T_0 \approx$ period of the shortest periodic orbit. It is an optimistic estimate for the range of validity of the periodic orbits approximation, with no immediate analogue in the problem at hand.

The Ehrenfest time is proportional to the logarithm of Planck's constant while the Heisenberg time is proportional to the inverse of Planck's constant. Thus the Ehrenfest time is considerably shorter than the Heisenberg time.

The Gaussian density "packets" studied in what follows parallels the study of quantum "coherent states." Gaussian packets that saturate the Heisenberg uncertainty bound, see the discussion and references in ref. [15–19]. Cametti and Presilla [20] discuss what happens near classical equilibrium points, with marginal stability power-law behavior.

⇑PRIVATE

While the general idea is intuitive, nonlinear dynamics interacts with noise in a highly nontrivial way, and methods to implement the optimal partition for a given noise, especially in the higher dimensions, still need to be developed. In this paper we shall explore four approaches to partitioning, by

- coarse-grained Fokker-Planck operator
- preimages of a critical point
- kneading invariant of the critical point
- periodic orbit neighborhoods

in presence of weak, Gaussian noise. [72] In sect. II we introduce the discrete-time Fokker-Planck operator $\mathcal{L}_\sigma$, its adjoint $\mathcal{L}_\sigma^\dagger$, and in sect. III we derive the eigenfunctions of $\mathcal{L}_\sigma$ and $\mathcal{L}_\sigma^\dagger$ in linear neighborhoods of fixed points and cycle points. In sect. V, we describe the evolution of a Gaussian density of trajectories under the action of the adjoint operator $\mathcal{L}_\sigma^\dagger$. In sect. VII we determine the optimal partition of a noisy skew Ulam map state space, by (a) backward iteration of a Gaussian centered at the critical point of the map, (b) forward iteration of the critical point (kneading invariant), and (c) computing noise widths of unstable periodic orbits. The approaches (a) and (c) yield different optimal partitions. We show how to construct the finite memory Markov graph for either. The deterministic skew Ulam map has a complete binary symbolic dynamics and full measure, but is strongly non-hyperbolic in the vicinity of its critical point. We investigate how this non-hyperbolicity affects the convergence of cycle expansions.

⇓PRIVATE

⇑PRIVATE

In higher dimension there are no critical points, and the state space partitioning by their preimages is an open problem. However, the periodic orbits are unambiguously computable, and in sects. VIII and IX we use noisy periodic orbits to partition the state space of the Lozi map.

II. FOKKER-PLANCK OPERATOR

Consider a d-dimensional Langevin flow [73]

$$\frac{dx}{dt} = v(x) + \hat{\xi}(t)\,, \tag{3}$$

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

$$\left\langle \hat{\xi}_j(t) \right\rangle = 0\,, \qquad \left\langle \hat{\xi}_i(t)\hat{\xi}_j(t') \right\rangle = 2D\,\delta(t-t')\,\delta_{ij}\,. \tag{4}$$

The Fokker-Planck equation [21, 22] describes the time evolution of a density of Langevin trajectories (3):

$$\partial_t \rho(x,t) + \nabla\cdot(v(x)\rho(x,t)) = D\nabla^2\rho(x,t)\,. \tag{5}$$

In what follows we shall write down many Gaussian integrals, so it pays to introduce compact notation for the normalized d-dimensional integration measure on $n+1$ points

[72] PC: update this overview

[73] PC: why hat on $\hat{\xi}(t)$?

$x_0 = x(0), x_1 = x(\delta t), \cdots, x_k = x(k\delta t), \cdots, x_n = x(t)$, of a trajectory $x(t)$, separated by time increments δt. Write the intermediate integrations in a compact form by defining

$$[dx] = \prod_{i=0}^{n-1} \frac{dx_i^d}{(4\pi D\delta t)^{d/2}} \,. \tag{6}$$

The short time step $\delta t = t/n$, $x_j = x(j\delta t)$, limit of the stochastic path integral [23]

$$\rho(x,t) = \mathcal{L}_\sigma^t \, \rho(x,0) = \lim_{n\to\infty} \int [dx] e^{-\frac{1}{4D}(\dot{x}_i - v(x_i))^2 \delta t_i} \rho(x_0, 0) \,, \tag{7}$$

together with an appropriate [74] finite time-step definition of $\dot{x}_i$, defines the Fokker-Planck operator $\mathcal{L}_\sigma^t$ for a noisy flow.

The finite time step formulation of the Fokker-Planck operator motivates what follows: we shall consider here noisy *discrete time* dynamical systems, with $\delta t = 1$,

$$x_{n+1} = f(x_n) + \hat{\xi}_n \,, \tag{8}$$

The noise $\hat{\xi}_n$ at discrete time n will be assumed isotropic normally distributed random variable, with *standard normal* (*Gaussian*) probability distribution function [PDF],

$$P(\hat{\xi}_n) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\hat{\xi}_n^2}{2\sigma^2}\right) \,, \tag{9}$$

of mean 0, variance $\sigma^2 = 2D$, and standard deviation $\sigma = \sqrt{2D}$, uncorrelated in time:

$$\left\langle \hat{\xi}_n \right\rangle = 0 \,, \qquad \left\langle \hat{\xi}_n \cdot \hat{\xi}_m \right\rangle = 2dD \, \delta_{nm} \,. \tag{10}$$

The action of discrete time step Fokker-Planck operator is noise-additive, centered on the nonlinear deterministic step $f(x_n)$ and smeared out diffusively by $\hat{\xi}_n$:

$$\rho_{n+1}(y) = (\mathcal{L}_\sigma \circ \rho_n)(y) = \int [dx] \, \exp\left\{-\frac{1}{4D}[y - f(x)]^2\right\} \rho_n(x) \,. \tag{11}$$

Here $[dx]$ is the normalized measure (6) with the unit time step $\delta t = 1$. The nth iterate of the stochastic evolution operator $\mathcal{L}_\sigma$ is a $(n–1)d$-dimensional path integral over n noisy trajectory points, with the trajectory represented with by the n-array of d-dimensional vectors x_a with time labels $a, b, \ldots$ ranging from 0 to $n-1$,

$$\mathcal{L}_\sigma^n \;=\; \int [dx] \, \exp\left\{-\frac{1}{4D} \sum_a \left[x_{a+1} - f(x_a)\right]^2\right\} \,. \tag{12}$$

[74] PC: cite Ito, Stratanovich here

The exponent $x_{a+1} - f(x_a)$ is in general nonlinear, and the path integral has to be evaluated numerically. In the weak noise limit it can be evaluated perturbatively [3–5] as an asymptotic series in D^k.

In the rest of the paper we assume that the noise is weak, and study the leading, Gaussian approximation to $\mathcal{L}^n_\sigma$, obtained by keeping track only of dynamics linearized around the flow, Taylor expanding the iterated mapping at *discrete* times $n = 1, 2, \ldots$ to linear order. The linearized neighborhood is transported by the Jacobian matrix

$$\mathbf{M}^n(x_0)_{ij} = \left.\frac{\partial f^n_i(x)}{\partial x_j}\right|_{x=x_0} . \tag{13}$$

The multiplication by 1-time step Jacobian matrix $\mathbf{M}$ along the n points x_0, x_1, x_2, ..., x_{n-1} on the trajectory of x_0, for the stability of the nth iterate of a d-dimensional map

$$\begin{aligned} \mathbf{M}^n(x_0) &= \mathbf{M}(x_{n-1}) \cdots \mathbf{M}(x_1)\mathbf{M}(x_0)\,, \\ \mathbf{M}(x)_{kl} &= \frac{\partial}{\partial x_l} f_k(x)\,, \quad x_m = f^m(x_0) \end{aligned} \tag{14}$$

follows from the chain rule for matrix derivatives.

The eigenvalues and eigendirections of the Jacobian matrix describe the deformation of an initial infinitesimal sphere of neighboring trajectories into an ellipsoid a finite time t later. Nearby trajectories separate exponentially along unstable directions, approach each other along stable directions, and change slowly (algebraically) their distance along marginal directions. The Jacobian matrix $\mathbf{M}^t$ is in general neither symmetric, nor diagonalizable by a rotation, nor do its (left or right) eigenvectors define an orthonormal coordinate frame.

III. DISCRETE ORNSTEIN-UHLENBECK PROCESS

Consider the 1-dimensional linear map

$$x_{n+1} = \Lambda x_n + \xi_n\,, \qquad |\Lambda| \neq 1\,, \tag{15}$$

with additive white noise with variance $2D$:

$$\langle \xi_n \rangle = 0\,, \qquad \langle \xi_n \xi_m \rangle = 2D\,\delta_{nm}\,. \tag{16}$$

This is the discrete time version of the Ornstein-Uhlenbeck process (the continuous time case is reviewed in appendix A).

The density $\rho(x)$ of trajectories evolves by the action of the Fokker-Planck operator (11):

$$\mathcal{L}_\sigma \rho(x) = \int [dy]\, e^{-\frac{(x-\Lambda y)^2}{4D}} \rho(y)\,. \tag{17}$$

$|\Lambda| < 1$ **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 (17) are [23]

$$\begin{aligned} dx\, \tilde{\rho}_k(x) &= [dx]\, H_k(\mu x)\, e^{-x^2/4\hat{D}}\,, \qquad \text{eigenvalue } \Lambda^k \\ \mu^{-2} = 4\hat{D}\,, \quad \hat{D} &= \frac{D}{1-\Lambda^2}\,, \end{aligned} \tag{18}$$

where $H_k(x)$ the kth Hermite polynomial, and $[dx]$ defined in (6), $[dx] = dx/(4\pi D)^{1/2}$. Hermite polynomials pop up here [24, 25] as the linear fixed point of dynamical systems is the imaginary time version of the harmonic oscillator of QM, see appendix A. 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 [1]. The unit-eigenvalue eigenfunction $\rho_0\, dx = [dx]\, \exp(-x^2/2\hat{D})$ is the natural measure [1] for the Fokker-Planck operator, its variance $2\hat{D} = 2D/(1-\Lambda^2)$ a balance of the fixed-point contraction Λ and diffusive spread D.

$|\Lambda| = 1$ **case** the marginal, pure diffusion case, and the behavior is not exponential, but power-law. If the map is nonlinear, need to go to the first nonlinear order to reestablish the control. [75]

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

$$dx\, \rho_k(x) = [dx]\, H_k(\alpha x)\,, \qquad \alpha^{-2} = -4\hat{D}\sqrt{\frac{\Lambda^2-1}{4D}}\,, \tag{19}$$

with eigenvalues $\Lambda^{-(k+1)}$.

The eigenfunctions (18) and (19) 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:

$$\int dx\, \rho_k(x)\rho_j(x) = \delta_{kj}\,. \tag{20}$$

In the deterministic, noiseless limit (17) reduces to the Perron-Frobenius operator:

$$\lim_{D\to 0} \mathcal{L}_\sigma\, \rho(x) = \mathcal{L}\rho(x) = \int dy\, \delta(x-\Lambda y)\rho(y)\,. \tag{21}$$

[75] PC: presumably explained in Gaspard's article

In the $|\Lambda| > 1$ expanding case the noiseless $D \to 0$ limit eigenfunctions (19) tend to the deterministic eigenfunctions [1] [76]

$$\rho_k(x) \to \frac{x^k}{k!} \tag{22}$$

while the contracting eigenfunctions (18) tend to distributions [22]

$$\rho_k(x) \to (-1)^k \delta^{(k)}(x)\,. \tag{23}$$

IV. LOCAL CYCLE-POINT EIGENFUNCTIONS

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

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

by Taylor-expanding around a periodic point x_0, so that the evolution operator (17) now acts on a density $\rho(x)$ as

$$\mathcal{L}_\sigma\, \rho(x) = \int [dy]\, \exp\left[-\frac{1}{4D}(x - f(x_0) - f'(x_0)(y - x_0))^2\right] \rho(y)\,. \tag{24}$$

If x_0 belongs to a stable periodic orbit, in linear neighborhood, week noise approximation the invariant measure is Gaussian, as in the single fixed-point case,

$$\rho_0(x) = C_0 e^{-(\mu_0(x-x_0))^2}\,. \tag{25}$$

Equation (24) then turns (25) into

$$C_1 e^{(c_0 x - \mu_0^2 x^2)/(4D\mu_0^2 + f'^2(x_0))}\,. \tag{26}$$

Now let

$$\mu_1^2 = \frac{\mu_0^2}{4D\mu_0^2 + f'^2(x_0)}\,. \tag{27}$$

This way the evolution operator produces another exponential

$$C_2 e^{(c_1 x - \mu_1^2 x^2)/(4D\mu_1^2 + f'^2(x_1))}\,, \tag{28}$$

[76] PC: refer to specific chapter in ref. [1]

where $x_1 = f(x_0)$. Now one can see the general pattern:

$$\mathcal{L}_\sigma C_n e^{(c_n x - \mu_n^2 x^2)} = C_{n+1} e^{(c_{n+1} x - \mu_{n+1}^2 x^2)/(4D\mu_{n+1}^2 + f'^2(x_n))}\,, \tag{29}$$

so that the terms of order x^2 in the exponential are related through the recursion relation:

$$\mu_{n+1}^2 = \frac{\mu_n^2}{4D\mu_n^2 + f'^2(x_n)} \tag{30}$$

and so

$$\mu_{n+1}^2 = \frac{\mu_{n-1}^2}{4D\mu_{n-1}^2(1+f_n'^2) + f_n'^2 f_{n-1}'^2} = \frac{\mu_{n-2}^2}{4D\mu_{n-2}^2(1+f_n'^2+f_n'^2 f_{n-1}'^2) + f_{n-2}'^2 f_{n-1}'^2 f_n'^2} = \ldots = \frac{\mu_0^2}{4D\mu_0^2(1+f_n'^2+f_{n-1}'^2 f_n'^2 + \ldots + f_n'^2 \cdot \ldots \cdot f_0'^2) + f_0'^2 \cdot \ldots \cdot f_n'^2}$$

where f_j' stands for $f'(x_j)$. For a cycle point on an attractive cycle $x_1, ..., x_n$ of length n, the local eigenfunction of $\mathcal{L}_\sigma^n$ is of the form (25), with

$$\mu^2 = \frac{1-\Lambda^2}{4D}\,\frac{1}{1+\sum_{i=1}^{n-1}\prod_{j=i}^{n-1} f_j'^2}\,, \qquad \Lambda = \prod_{m=0} 0^{n_p} f_m'\,. \tag{31}$$

As for the fixed-point Ornstein-Uhlenbeck process, the local eigenfunctions of a periodic point belonging to a stable cycle p of period n_p are

$$\tilde{\rho}_k(x) = \frac{1}{\sqrt{4\pi D}} H_k(\mu(x-x_0)) e^{-(\mu(x-x_0))^2} \tag{32}$$

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

$$\tilde{\rho}_k(x) = \frac{1}{\sqrt{4\pi D}} H_k(\alpha(x-x_0)) \tag{33}$$

where $\alpha^2 = -\mu^2$, as in the case of a single fixed point. Eigenvalues are in this case Λ^{-k-1}.

V. ADJOINT FOKKER-PLANCK OPERATOR

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

$$\begin{aligned} \langle h|\mathcal{L}_\sigma|\rho\rangle &= \int_{-\infty}^{\infty} h(x)dx \int_{-\infty}^{\infty} e^{-\frac{(x-f(y))^2}{4D}} \rho(y)dy = \\ \left\langle \rho|\mathcal{L}_\sigma^\dagger|h\right\rangle &= \int_{-\infty}^{\infty} e^{-\frac{(y-f(x))^2}{4D}} \rho(x)dx \int_{-\infty}^{\infty} h(y)dy \end{aligned} \tag{34}$$

so that

$$\mathcal{L}_\sigma^\dagger \rho(x) = \int_{-\infty}^{\infty} e^{-\frac{(f(x)-y)^2}{4D}} \rho(y) dy \tag{35}$$

$\mathcal{L}_\sigma$ 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 $\mathcal{L}_\sigma^\dagger$ transports the density $\rho(x)$, supported on I, backward in time to a function supported on $f^{-1}(I)$.

A. Eigenfunctions of $\mathcal{L}_\sigma^\dagger$ for a fixed point

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

$$\mathcal{L}_\sigma^\dagger \rho(x) = \int_{-\infty}^{\infty} e^{-\frac{(\Lambda x-y)^2}{4D}} \rho(y) dy \tag{36}$$

Eigenfunctions are

$$\rho_n(x) = \frac{1}{\sqrt{4\pi D}} H_n(\alpha x) e^{-(\alpha x)^2}, \quad \alpha = \sqrt{\frac{\Lambda^2 - 1}{4D}} \tag{37}$$

and eigenvalues Λ^{-n-1}.

B. Evolution of a Gaussian and local approximation

Now take [77]

$$\rho(x) = C_0 e^{-\mu^2 (x-x_0)^2} \tag{38}$$

From (35) [78]

$$\begin{aligned} \mathcal{L}_\sigma^\dagger \rho(x) &= \int_{-\infty}^{\infty} C_0 e^{-\frac{(f(x)-y)^2}{4D}} e^{-\mu^2 (y-x_0)^2} dy \\ &= C_1 e^{-\frac{\mu^2 (f(x)-x_0)^2}{1+4D\mu^2}} . \end{aligned} \tag{39}$$

Given a nonlinear $f(x)$, we now evolve the density in the neighborhood of $f^{-1}(x_0)$ to linear order,

$$\mathcal{L}_\sigma^\dagger \rho(x) = C_1 e^{-\frac{\mu^2 (f(x)-x_0)^2}{1+4D\mu^2}} \simeq C_1 e^{-\frac{\mu^2 (f'(f^{-1}(x_0))(x-f^{-1}(x_0)))^2}{1+4D\mu^2}} \tag{40}$$

[77] PC: why introduce a new Gaussian parameter $\mu^2 = 1/2\sigma^2$ instead of using the standard notation σ for Gaussian probability distribution variance (9)?

[78] PC: why mystify this? Variances add, so $\sigma_{n+1}^2 = \sigma_n^2 + 2D$ and (39) is exp $\left(-\frac{1}{2}\frac{[f(x)-x_0]^2}{\sigma_n^2+2D}\right)$. Please change to the standard variances notation in all of your formulas, do not make me use time on that if you want me to help you with real problems. I have written this up for you in (B12).

which then reduces (39) to a Gaussian. As in sect. IV, we can obtain a recursion relation for the evolution of μ:

$$\mu_{-1}^2 = \frac{(f'(f^{-1}(x_0))\mu_0)^2}{1+4D\mu_0^2} \tag{41}$$

which can be extended to the nth preimage of the point x_0 :

$$\mu_{-n}^2 = \frac{\Lambda^2\mu_0^2}{1+4D\mu_0^2(1+\sum_{i=1}^{n-1}\prod_{j=1}^{i} f'^2(f^{-j}(x_0)))} \tag{42}$$

where $\Lambda = \prod_{i=1}^{n} f'(f^{-i}(x_0))$. The leading eigenfunction of $\mathcal{L}_\sigma^\dagger$ is:

$$\rho(x) = C_0 e^{-\mu^2(x-x_0)^2}, \quad \mu^2 = \frac{\Lambda^2-1}{4D} \tag{43}$$

as well as the expression for an eigenfunction of $\mathcal{L}_\sigma^n$:

$$\rho(x) = C_n e^{-\mu_n^2(x-x_0)^2}, \tag{44}$$

$$\mu_n^2 = \frac{\Lambda^2-1}{4D}\frac{1}{1+\sum_{i=1}^{n-1}\prod_{j=1}^{i} f'^2(f^{-j}(x_0))} \tag{45}$$

VI. OPTIMAL PARTITION FOR A NOISY UNIMODAL MAP

We now explore three approaches to construction of an optimal partition of the state space of a unimodal map, in the presence of weak, Gaussian noise, by adopting the deterministic partitions by

- preimages of the critical point
- kneading invariant of the critical point
- periodic orbit neighborhoods.

[79] We shall refer to any unimodal map for which the critical point x_c is mapped onto the unstable fixed point $\overline{0}$ as an "Ulam" map. For Ulam maps of the critical point is preperiodic to fixed point $\overline{0}$, so the kneading invariant is simply $1\overline{0}$, and it plays no further role in partitioning of the state space. As a model map on which to test different approaches we take the "skew Ulam" map [26], [80]

$$f(x) = \Lambda_0 x(1-x)(1-bx)\ , \quad 1/\Lambda_0 = x_c(1-x_c)(1-bx_c)\,, \tag{46}$$

[79] PC: draw skew_ulam.eps or maybe use Vaggelis's, if he has drawn it

[80] PC: Ref. [27] credits A. J. Lichtenberg and M. A. Lieberman, *Regular and Chaotic Dynamics* (Springer, New York, 1992) with introducing this map, also with $b = 0.6$. Could it be ref. [26] did not introduce it first?

a nonlinear deformation of the Ulam map

$$f(x) = 4x(1-x)\,. \tag{47}$$

The $x_0 = 0$ fixed point $\overline{0}$ stability multiplier Λ_0 is fixed by condition $f(x_c) = 1$, and x_c as a function of b fixed by $f'(x_c) = 0$ condition. The skew Ulam map is a convenient starting point for testing ideas about partitioning for several reasons:

- The symbolic dynamics is a complete binary symbolic dynamics, and the nth preimage of the critical point generates a complete binary cover of the unit interval.[81]
- All cycles can be easily determined through a combination of inverse iterations and the Newton method [1].
- The quadratic contraction around the critical point x_c makes this mapping *nonhyperbolic*.

In our numerical work we fix $b = 0.6$ (arbitrarily, it is the value used in ref. [26]), so $x_c = 0.40456\ldots$, and $\Lambda_0 = 5.4819\ldots$ in (46).

A. Fokker-Planck spectrum

In order to be able to gauge accuracy of different noise-limited finite partitions we start by a numerical calculation of the exact Fokker-Planck spectrum.

Think first of the deterministic evolution operator, $\mathcal{L}$, as a matrix. [82]

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 $\mathcal{M}_i$:

$$\rho(x) \;=\; \sum_{i=1}^{N} \rho_i \frac{\chi_i(x)}{|\mathcal{M}_i|}\,, \quad \chi_i(x) = \begin{cases} 1 \;\text{ if } x \in \mathcal{M}_i\,, \\ 0 \;\text{ otherwise}\,. \end{cases} \tag{48}$$

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

[81] PC: add references
[82] PC: rewrite this for the Fokker-Planck operator case

operator: [83]

$$\rho'_j = \int_{\mathcal{M}} dy\, \chi_j(y)\rho'(y) = \int_{\mathcal{M}} dx\, dy\, \chi_j(y)\, \delta(y - f(x))\, \rho(x)$$
$$= \sum_{i=1}^{N} \rho_i \frac{|\mathcal{M}_i \cap f^{-1}(\mathcal{M}_j)|}{|\mathcal{M}_i|}\,. \tag{49}$$

In this way

$$L_{ij} = \frac{|\mathcal{M}_i \cap f^{-1}(\mathcal{M}_j)|}{|\mathcal{M}_i|}, \quad \rho' = \rho\mathbf{L} \tag{50}$$

is a matrix approximation to the Perron-Frobenius operator, and its leading left eigenvector is a piecewise constant approximation to the invariant measure.

The corresponding piecewise constant approximation to Fokker-Planck operator (11) is [84]

$$[\mathcal{L}_\sigma]_{ij} = \frac{1}{|\mathcal{M}_i|} \frac{1}{\sqrt{4\pi D}} \int_{\mathcal{M}_i} dx \int_{f^{-1}(\mathcal{M}_j)} dy \exp\left\{-\frac{1}{4D}[y - f(x)]^2\right\}. \tag{51}$$

It is an old idea of Ulam [12] that such an approximation for the Perron-Frobenius operator is a meaningful one. [85] The approximation of the deterministic Perron-Frobenius operator defined by (50) has been shown to reproduce the spectrum for expanding maps, once finer and finer Markov partitions are used [28, 29]. The subtle point of choosing a state space partitioning for a "generic case" is discussed in ref. [30, 31]. [86]

We discretize the Fokker-Planck operator by using the deterministic subdivision of the unit interval by the 2^n preimages of the critical point, and compute the matrix elements (50), see figure 1. [87] The Fokker-Planck eigenfunctions extend outside the unit interval, where there is no deterministic partition. In order to guarantee that the discretization of the eigenfunction tails is as fine as within the unit interval, we extend the unit interval on each side by the three standard deviations $I = 3\sigma$ of a Gaussian density,

$$dx\, \rho_0(x) = [dx]\, e^{-(x-x_0)^2/4D} \tag{52}$$

centered respectively at the first and the second image of the critical point. We subdivide I into 2^n intervals of equal size. In figure 1 each column of rectangles represents a row of the L

[83] PC: recheck, missing f' jacobian here?
[84] PC: please recheck - what formula do you use?
[85] PC: you might want to read Ulam [12] to see what the master really wrote - I have not checked it
[86] PC: I think you would profit from reading Froyland - link to his homepage, with all his papers is in sect. C. In particular, we are defining the discretized Fokker-Planck without reference to natural measure, while he defines something that depend on it. I think we are right (but I might be wrong).
[87] PC: write up the correct formula for the noise case, refer to it here rather than to the deterministic (50)

matrix, with the size of the Fokker-Planck operator matrix element $-\ln L_{ij}$ is indicated by color/graytone scale (red/black the highest magnitude, down to blue/white for small/zero matrix elements). [88]

Figure 2 illustrates the convergence of the escape rate $\gamma = -s_0$, and the gap to the next eigenvalue s_1 for the skew Ulam map (46).

[89] The problem with such state space discretization approaches is that they are blind, the grid knows not what parts of the state space are more or less important. This observation motivates the next step in developing the theory of long-time dynamics of chaotic systems: we shall exploit the intrinsic topology of the flow to give us both an invariant partition of the state space and a measure of the partition volumes.

Furthermore, a piecewise constant ρ belongs to an unphysical function space, and with such approximations one is plagued by numerical artifacts such as spurious eigenvalues. We shall employ a more refined approach to extracting spectra, by expanding the initial and final densities ρ, ρ' in some basis φ_0, φ_1, $\varphi_2, \cdots$ (orthogonal polynomials, let us say), and replacing $\mathcal{L}(y,x)$ by its φ_α basis representation $\mathbf{L}_{\alpha\beta} = \langle\varphi_\alpha|\mathcal{L}|\varphi_\beta\rangle$. The art is then the subtle art of finding a "good" basis for which finite truncations of $\mathbf{L}_{\alpha\beta}$ give accurate estimates of the eigenvalues of $\mathcal{L}$.

Regardless of how sophisticated the choice of basis might be, the basic problem cannot be avoided - as illustrated by the natural measure for the Hénon map eigenfunctions of $\mathcal{L}$ are complicated, singular functions concentrated on fractal sets, and in general cannot be represented by a nice basis set of smooth functions. We shall resort to matrix representations of $\mathcal{L}$ and the φ_α basis approach only insofar this helps us prove that the spectrum that we compute is indeed the correct one, and that finite periodic orbit truncations do converge.

[90] In figure 3 we plot the leading eigenfuction5B ρ_0, that is, the natural measure for

[88] PC: Re: figure 1:
= divide the 3 standard deviations into subintervals, as described in the text
= state $2D$ =? in the caption
= tell your program to use a large font size for graph labels, so we have figures publication ready, if we use them
= experiment again with replacing color by grayscale (so one can see this in B/W paper printouts)

[89] PC: tone the remainder of this down - noisy case is better than the deterministic one. One would want to use Hermite basis in some form, but not for purposes of this paper

[90] PC: they are looking good. If I remember well the deterministic calculations, 2nd eigenvalue has about 1/2 significant digits when computed from cycle expansions, but from matrix eigenvalue computations we

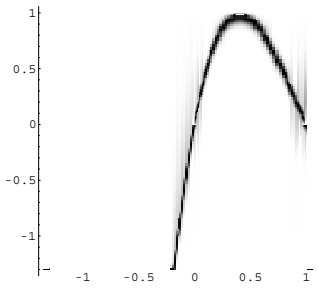


FIG. 1: Deterministic partition of the discretized Fokker-Planck operator for the skew Ulam map (46) by preimages of the critical point ($2^6 + 2^6 + 2^4$ intervals) with variance $2D = 0.002$. Color/gray scale indicates the matrix element size; resolving into intervals smaller than the local mean deviation $\sqrt{2D_i}$ leads to no further significant improvement in predictions of averages.

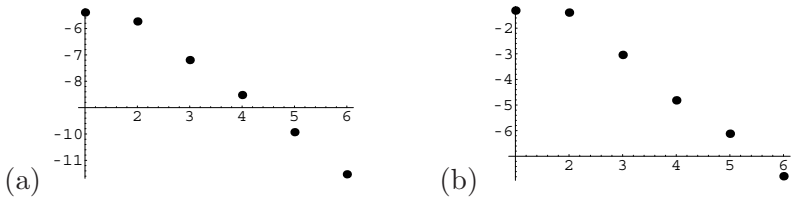


FIG. 2: The convergence of (a) the first $s_0 = -\gamma = 0.0690$ and (b) the second $s_1 = -1.204 - 0.857i$ eigenvalues of the discretized Fokker-Planck matrix figure 1, variance $2D = 0.002$. Plotted: $\ln|s_\alpha^{(N)} - s_\alpha^{\infty}|$ as a function of $\ln N$, the number of partition intervals N. In the deterministic Ulam map partition, with the 2^n intervals, $\ln N = n \ln 2$. s_α is estimated as a limit $s_\alpha^{(\infty)}$ from the computed $\ln|s_\alpha^{(N)}|$.

$2D = 0.002$ Fokker-Planck equation. is as smooth deformation of the Ulam map (47), for which the measure is known analytically the Ulam map [91]

$$\rho_0(x) = \frac{1}{\pi\sqrt{x(1-x)}}\,. \tag{53}$$

[92] [93]

used to get many eigenvalues to the machine precision, so this is a bit puzzling.

[91] PC: check textbooks for this. If you write it up as a problem and solution, I'll put it into ChaosBook, seems like a good exercise

[92] Domenico: The new figure (invariant measure of Ulam map) is obtained using the same exact code I used for skew Ulam map, so maybe there's no bug.

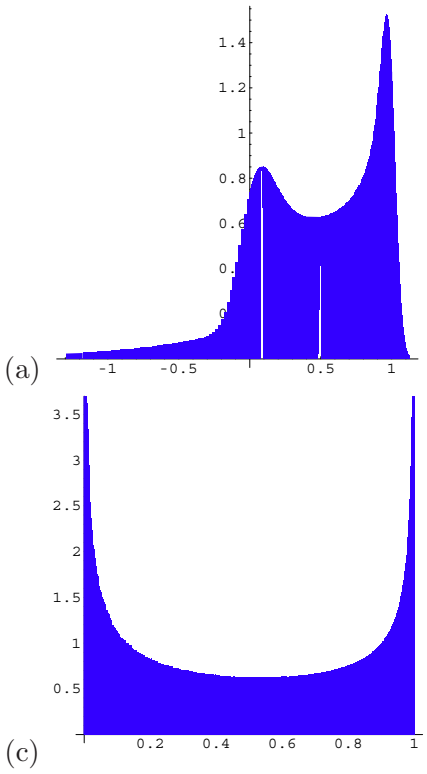


FIG. 3: (a) The leading eigenfunction (natural measure) ρ_0 of the Fokker-Planck operator figure 1 for the skew Ulam map, variance $2D = 0.002$, N=256+256 interval discretization. (c) Deterministic case, N=512.

PC: why did you not use (47), $b \to 0$ limit of (47), rather than using an equivalent but different form of the Ulam map? It is more consistent, and if this irregularity is due to skewness, you caould compare resulats for different b. Also, you should normalize (53) correctly, and plot it on the same graph figure 3(d), to check whether you agree with the analytic answer.

PC: work out the chaos course problem with skew Ulam tent map; the size of the partition is in that case given by $\Lambda_0^{n_0}\Lambda_1^{n_1}$, n_0 the number of zeros in the partition label, so that cannot be monotone. This is worked out in detail by Billings and Bolt [32]. So skew Ulam is not monotone, either, and that's the cause of these peaks; people talk about "fractality of eigenstates, but I did not realize it kicks in already here. A good potential problem set. These measures are indeed kinky, so probably not much point of computing them, unless you work out a tent example in detail, to illustrate why you will not plot them. You are right. But with noise, every measure should be smooth beyond the noise resolution. There is something weird about your measures in figure 3(a)-(c), because your calculation is not converging to anything. People have computed this before - have a book by Boyarski [33] and Driebe, D.J.: *Fully Chaotic Maps and Broken Time Symmetry* (Kluwer,c 1999) in the CNS library, ref. [34, 35] and many others that do nothing but $1d$ maps and their measures.

PC: Please check whether your definition agrees with my (51) - I am not sure I am right. The correct formula is presumably in Froyland [28, 29].

[93] PC: I guess next step is to try pruning to finite partitions, see how well you do. We need to decide how

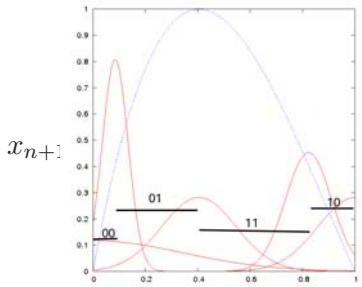


FIG. 4: Skew Ulam map and two-step-memory partition: in the deterministic case (left) the border-points are, from left to right, $f^2(x_c)$,$f_0^{-1}(x_c)$,x_c,$f_1^{-1}(x_c)$,$f(x_c)$; in the noisy case (right) the fuzzy borders of the intervals are Gaussians centered at the deterministic points.

B. Optimal partition by preimages of the critical point

For a unimodal map the borders between adjacent intervals are determined by the preimages of the critical point x_c, defined by condition $f'(x_c) = 0$. In the presence of noise, the critical point is replaced by a Gaussian density (16), centered at x_c:

$$dx\, \rho(x) = [dx]\, e^{-(x-x_c)^2/4D}\,. \tag{54}$$

Instead of iterating the critical point deterministically, by means of $f^n(x_c)$ and $f^{-n}(x_c)$, we iterate the density (54) either forward or backwards in time by using $\mathcal{L}_\sigma^n$ and $\mathcal{L}_\sigma^{n\dagger}$, respectively. At each iteration, we approximate the result with a Gaussian as explained in sect. V. Figure 4 shows a two-symbol-memory partition of the chaotic attractor of the skew Ulam map. [94]

We take noise of variance $2D = 0.02$: the deterministic intervals $\mathcal{M}_{00} = [f^2(x_c), f_0^{-1}(x_c)]$, $\mathcal{M}_{01} = [f_0^{-1}(x_c), x_c]$, $\mathcal{M}_{11} = [x_c, f_1^{-1}(x_c)]$, $\mathcal{M}_{10} = [f_1^{-1}(x_c), f(x_c)]$ now have fuzzy borders, with each border-point replaced by a distribution is a Gaussian peaked at the deterministic border point.

---

accurately it makes sense to compute things like γ, then 1) use this as a criterion for optimal partition = partition where you glue together all intervals which are too small too matter, and construct a coarser transition matrix. 2) compare with partition obtained from cyclepoints' standard deviations. Interestingly, cycle expansion cannot get your γ, as that should be converging to 0 (noise does not show up in the leading order, only as perturbative correction). The fact that we cut off at x=1 worries me - to get ρ_0 to converge well one needs to extend these partitions to several standard deviations outside the [0,1] interval.

[94] PC: skew_ulam.eps is damaged

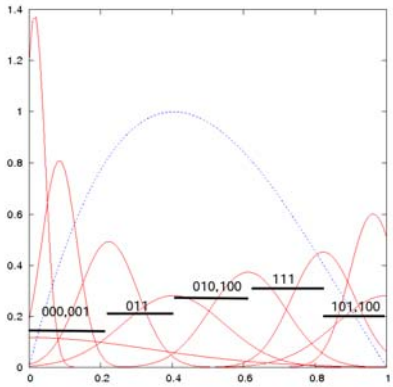


FIG. 5: Gaussian densities marking the borders of the intervals that make a three-step memory partition of the chaotic attractor, for the noisy skew Ulam map with variance $2D = 0.02$.

The optimal resolution of a region of the state space is reached once the borders of the intervals overlap significantly. Significant overlapping occurs when two probability densities of adjacent border point overlap at the 68.3% confidence level [36]. [95] There is no pruning in this model, however, the noise makes the system lose memory as soon as the Gaussians marking the fuzzy borders of the intervals overlap significantly. In order to reach the optimal resolution, different regions of the chaotic attractor may need a different level of partition refinement: figure 5 shows a three-symbol partition: we can see significant overlapping of the Gaussian densities marking the fuzzy borders of the outer intervals. However, borders in the central region of the attractor, closer to the critical point, do not overlap, so that a finer resolution is still needed there. As it can be inferred from (42), the regions of least instability overlap first.

The finest partition attained for $D = 0.01$ is given in figure 6. [96] The intervals 00 and 01 are distinct, but 000 and 001 are not, since their fuzzy borders overlap significantly. Consequently, the dynamics loses memory beyond that resolution. At the boundary marked by the critical point x_c, the intervals 010 and 110 overlap significantly and form a single partition region. The Markov graph corresponding to this finite memory partition is drawn

[95] PC: maybe discussion of overlap in mathworld.wolfram.com/Convolution.html is useful

[96] PC: I have not checked your numerical $D = 0.01$ calculation. What follows is a redrawing of your figures, and an attempt - preliminary - at the corresponding Markov graph. The error in figure 5 - where you indicate 010+100 as cojoined siblings - sure slowed me down. Fortunately they are labeled correctly in the text.

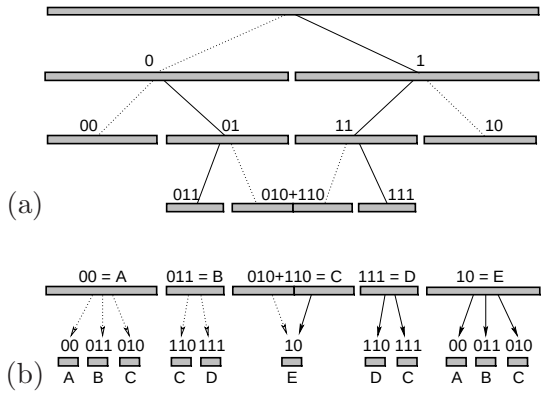


FIG. 6: Noisy skew Ulam map with variance $2D = 0.02$ (a) The optimal partition. (b) The 5 regions of the optimal partition relabeled, with the regions they map onto in the next iterate of the dynamics indicated. A dashed line represents the 0 symbol, and a full line the symbol 1.

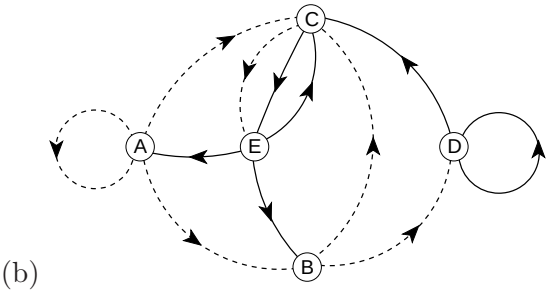


FIG. 7: Noisy skew Ulam map with variance $2D = 0.02$, the finite-memory Markov graph obtained from the optimal partition of figure 6 (b):

in (figure 7). For this Markov graph,

$$1/\zeta(0) = 1-(\frac{1}{\Lambda_A}+\frac{1}{\Lambda_B})z+(\frac{1}{\Lambda_A\Lambda_B}-\frac{1}{\Lambda_{CE}})z^2+(\frac{1}{\Lambda_A\Lambda_{CE}}+\frac{1}{\Lambda_D}\Lambda_{CE}-\frac{1}{\Lambda_{ACE}}-\frac{1}{\Lambda_{EBC}})z^3+(-\frac{1}{\Lambda_{ABCE}}-\frac{1}{\Lambda_{BDEC}}-\Lambda \tag{55}$$

whose first zero is $z_0 = 0.04018$, in perfect disagreement with the first eigenvalue of the discretized Fokker-Planck operator, $\gamma = 0.1467$. The main problem here is that the weak-noise approximation makes little sense for $2D = 0.02$, as figure 3 may suggest.

C. Optimal partition by cycle neighborhoods

We next partition the unit interval using the local eigenfunctions (37) of $\mathcal{L}^\dagger_\sigma$, determined in the neighborhood of each periodic point of the map. We define a neighborhood of a periodic

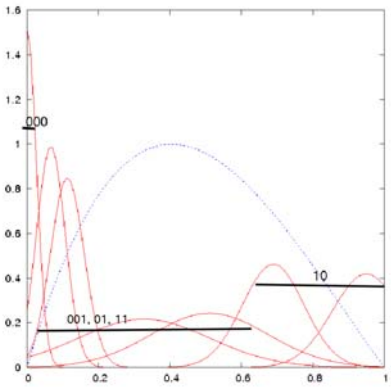


FIG. 8: Optimal partition for the same skew Ulam map with noise $2D = 0.02$, but defining the intervals by means of the local periodic orbit widths $\sqrt{2\hat{D}}$.

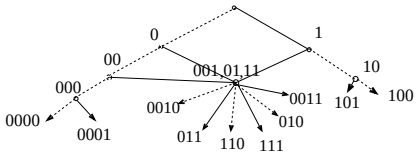


FIG. 9: Pruned binary tree for noisy Ulam skew map, this time obtained by partitioning with the method of the eigenfunctions.

point as the interval $[x_p - \sqrt{2}/\beta, x_p + \sqrt{2}/\beta]$ [97] of the support of its local eigenfunction, centered at the periodic point and spread over one standard deviation on each side of the point. We associate to each of these intervals the symbolic sequence of the corresponding periodic point, so that we can draw once again a binary tree, and stop resolving as soon as significant overlapping occurs between the Gaussians. Figure 8 and figure 9 show the optimal minimal resolution found with this second method, and the corresponding binary tree. Here the intervals labelled with 001,01,11 also overlap significantly, so that they are represented by one common node in the subsequent Markov graph (figure 10).

As we can see from the Markov graphs in figure 7 and figure 10, the two methods produce different results for the same model. This is expected, since the widths of the Gaussians in the backward evolution depend on the stability of all the points visited by the noisy sequence of iterations (cf. (42) with (45)). Obviously, the preimages of the critical point have different stabilities than the cycles involved in the partition of figure 8. If $000 =$ "A",

[97] PC: what "β"? You mean mean square deviation $\sqrt{2\hat{D}}$?

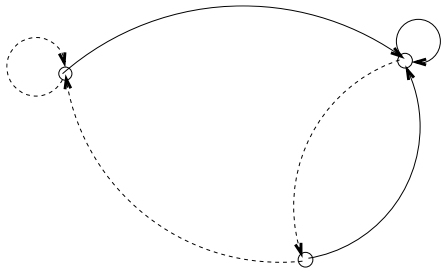


FIG. 10: Markov graph for the binary tree of figure 9.

$001 \cup 01 \cup 11 =$ "B", $10 =$ "C", we have

$$1/\zeta(0) = (-\frac{1}{\Lambda_A} - \frac{1}{\Lambda_B})z + (\frac{1}{\Lambda_A\Lambda_B} - \frac{1}{\Lambda_{BC}}) + (\frac{1}{\Lambda_{BC}\Lambda_A} - \frac{1}{\Lambda_{ABC}}) \tag{56}$$

whose leading zero is $z_0 = 0.03407$, also inconsistent with $\gamma = 0.1467$, obtained by discretization of the Fokker-Planck operator.

D. Optimal partition by preimages of the critical point, pruned map

We now apply the preimages of the critical point partition method to the noisy unimodal map (46) with $a = 5.4$, variance $2D = 0.002$. Now the map is not of the Ulam type, and infinitely many itineraries are inadmissible, as determined by the kneading theory. [98]

The shortest inadmissible symbol block is "000". Unlike in the previous example, the problem is now that there is no longer a one-to-one correspondence between each finite memory sequence and the periodic point obtained by repeating that sequence. For instance, the string 1100 is admissible, but the cycle $\overline{1100}$ is not, since its rightmost cycle point 1001 lies to the right of the kneading sequence $11001100 \succ 10011111$. Our strategy consists in covering the attractor with neighborhoods of all the periodic points that we can find, until the optimal resolution is reached. After that, the binary tree is reconstructed by knowing the pruning rules. In our example, we look for periodic points and their noisy neighborhoods starting from length $n = 1$: the fixed point "$\overline{0}$" is pruned, whereas "$\overline{1}$" exists. We go on to length $n = 2$, finding the cycle points 01, 10, 1. The sequence 00 is not represented, although it is allowed, since $\overline{0}$ is pruned, hence there is a void to fill. [99]

[98] PC: can you please define the map you re using? What is a?

[99] PC: $\overline{0}$ is not pruned - it is admissible, but it is transient and isolated - nearby trajectories cannot return

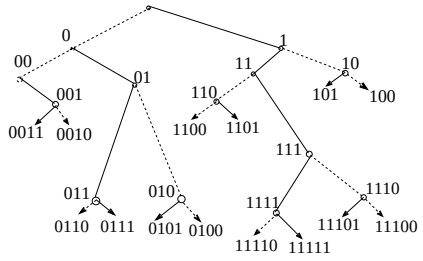


FIG. 11: Binary tree of finite memory obtained for the pruned unimodal map (**??**), with variance $2D = 0.002$.

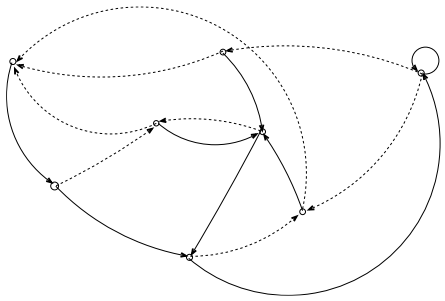


FIG. 12: Markov graph for the binary tree of figure 11.

We look for cycles of length $n = 3$: $\overline{001}$, $\overline{011}$, $\overline{010}$, $\overline{110}$, $\overline{1}$, $\overline{101}$, $\overline{100}$. Now the regions 0 and 00 are covered by the periodic point 001, whereas the sequence 000 is itself forbidden by the grammar. At this point, the neighborhoods that we have found cover the whole attractor. The corresponding Gaussian densities do not overlap significantly, though, except for $\overline{101}$ and $\overline{100}$, thus we can make the partition finer, by going on to $n = 4$: region 001 splits into 0011 and 0010, both admissible, but the cycles $\overline{0011}$ and $\overline{0010}$ are both pruned, so we just leave 001 at this step. Something similar happens with region 100, which should split into 1001 and 1000, but $\overline{1001}$ is pruned and 1000 is forbidden since it contains 000. The sequence 110 splits into 1100 and 1101, but again $\overline{1100}$ is pruned and therefore we leave 110 in our partition. The sequence 111 splits into 1111 and 1110, both of those contain a periodic point, and the partition can be refined. As a result, we now have the following neighborhoods: 001, 011, 010, 110, 1111, 1110, 10. The procedure can be iterated to periodic orbits of length $n = 5$, but it turns out that all the new neighborhoods found overlap with each other, so that the optimal partition is reached by level $n = 4$. Figure 11 shows the binary tree for this example: the neighborhoods listed above are the ends of the tree, the rest of the nodes above them are reconstructed following the pruning rules: the only pruned (inadmissible)

block is _000_. As in the previous examples, a Markov graph can be drawn from the binary tree (figure 12).

VII. STABILITY ORDERING

[100] Most dynamical systems of interest have no finite grammar, so at any order in z a cycle expansion may contain unmatched terms which do not fit neatly into the almost canceling curvature corrections. Similarly, for intermittent systems discussed in ChaosBook.org, curvature corrections are in general not small, so again the cycle expansions may converge slowly. For such systems schemes which collect the pseudocycle terms according to some criterion other than the topology of the flow may converge more quickly than expansions based on the topological length.

All chaotic systems exhibit some degree of shadowing, and a good truncation criterion should do its best to respect the shadowing at least approximately. If a long cycle is shadowed by two or more shorter cycles and the flow is smooth, the period and the action will be additive in sense that the period of the longer cycle is approximately the sum of the shorter cycle periods. Similarly, stability is multiplicative, so shadowing is approximately preserved by including all terms with pseudocycle stability

$$|\Lambda_{p_1} \cdots \Lambda_{p_k}| \leq \Lambda_{\max} \tag{57}$$

and ignoring all more unstable pseudocycles.

Two such schemes for ordering cycle expansions which approximately respect shadowing are truncations by the pseudocycle period (or action) and the stability ordering that we shall discuss here. In these schemes a dynamical zeta function or a spectral determinant is expanded keeping all terms for which the period, action or stability for a combination of cycles (pseudocycle) is less than a given cutoff.

Stability ordering was introduced by Dahlqvist and Russberg [37] in a study of chaotic dynamics for the $(x^2y^2)^{1/a}$ potential. The presentation here is extracted from ChaosBook.org,

[100] PC: this section is eventually going into your thesis - I put it here as an inspiration for writing up the optimal partition discussion, it is very similar - partition to the smallest resolvable intervals. In the deterministic stability ordering one does something similar, keeps only intervals whose size is larger than $1/\Lambda_{\max}$.

and follows the exposition of Dettmann and Morriss [38] for the Lorentz gas which is hyperbolic but the symbolic dynamics is highly pruned, and Dettmann and Cvitanović [39] for a family of intermittent maps. In the applications discussed in the above papers, the stability ordering yields a considerable improvement over the topological length ordering. In quantum chaos applications cycle expansion cancelations are affected by the phases of pseudocycles (their actions), hence *period ordering* rather than stability is frequently employed.

The two settings in which the stability ordering may be preferable to [101] the ordering by topological cycle length are the cases of bad grammar and of intermittency.

A. Stability ordering for bad grammars

For generic flows it is often not clear what partition of the state space generates the "optimal" symbolic dynamics. Stability ordering does not require understanding dynamics in such detail: if you can find the cycles, you can use stability ordered cycle expansions. Stability truncation is thus easier to implement for a generic dynamical system than the curvature expansions [102] which rely on finite subshift approximations to a given flow.

Cycles can be detected numerically by searching a long trajectory for near recurrences. The long trajectory method for detecting cycles [103] preferentially finds the least unstable cycles, regardless of their topological length. Another practical advantage of the method (in contrast to Newton method searches) is that it only finds cycles in a given connected ergodic component of state space, ignoring isolated cycles or other ergodic regions elsewhere in the state space.

Why should stability ordered cycle expansion of a dynamical zeta function converge better than the rude trace formula [104] (**??**)? [105] The argument has essentially already been laid out in sect. **??**: [106] in truncations that respect shadowing most of the pseudocycles appear in shadowing combinations and nearly cancel, while only the relatively small subset affected by the longer and longer pruning rules is not shadowed. So the error is typically of

[101] PC: add Berry-Keating somewhere
[102] PC: refer here to the thesis section
[103] PC: refer here to the thesis section
[104] PC: explain better
[105] PC: refer here to the thesis equation
[106] PC: refer here to the thesis equation

plot here

FIG. 13: The leading eigenvalue $\ln|s_0^{(N)} - s_0|$ convergence for the logistic map (??), plotted *vs.* lnN.

plot here

FIG. 14: The entropy $\ln|h^{(N)} - h|$ convergence for the logistic map (??), plotted *vs.* lnN for the binary and the stability (Kolmogorov) state space partitioned topological zeta functions.

the order of $1/\Lambda$, smaller by factor e^{hT} than the [107] trace formula (**??**) error, where h is the entropy and T typical cycle length for cycles of stability Λ. [108]

The results of a stability ordered expansion (57) should always be tested for robustness by varying the cutoff $\Lambda_{\max}$. If this introduces significant variations, smoothing is probably necessary.

The leading eigenvalue convergence $\ln|s_0^{(N)} - s_0|$ is plotted in figure 14 as a function of lnN, the number of intervals N is $N = 2^n$ for the Ulam map, $N = e^{nh_n}$ for a pruned logistic map, and $N = N(\Lambda_{\max})$ for the stability ordered cycle expansion.

B. Stability ordering for intermittent flows

[109] Longer but less unstable cycles can give larger contributions to a cycle expansion than short but highly unstable cycles. In such situation truncation by length may require an exponentially large number of very unstable cycles before a significant longer cycle is first included in the expansion. This situation is best illustrated by intermittent maps [110] that we shall study in detail in ChaosBook.org, the simplest of which is the Farey map [111]

$$f(x) = \begin{cases} f_0 = x/(1-x) & 0 \le x \le 1/2 \\ f_1 = (1-x)/x & 1/2 \le x \le 1 \end{cases} . \tag{58}$$

[107] PC: refer here to the thesis equation
[108] PC: Incorporate Grassberger's dimension of the Cantor set of pruning rules
[109] PC: use what you need from this section, remove the rest
[110] PC: add figure
[111] PC: replace by your logistic map

For this map the symbolic dynamics is of complete binary type, so lack of shadowing is not due to lack of a finite grammar, but rather to the intermittency caused by the existence of the marginal fixed point $x_0 = 0$, for which the stability equals $\Lambda_0 = 1$. This fixed point does not participate directly in the dynamics and is omitted from cycle expansions. Its presence is felt in the stabilities of neighboring cycles with n consecutive repeats of the symbol 0's whose stability falls of only as $\Lambda \sim n^2$, in contrast to the most unstable cycles with n consecutive 1's which are exponentially unstable, $|\Lambda_{01^n}| \sim [(\sqrt{5}+1)/2]^{2n}$.

The symbolic dynamics is of complete binary type. A quick count in the style of ChaosBook.orgleads to a total of 74,248,450 prime cycles of length 30 or less, not including the marginal point $x_0 = 0$. Evaluating a cycle expansion to this order would be no mean computational feat. However, the least unstable cycle omitted has stability of roughly $\Lambda_{10^{30}} \sim 30^2 = 900$, and so amounts to a 0.1% correction. The situation may be much worse than this estimate suggests, because the next, 10^{31} cycle contributes a similar amount, and could easily reinforce the error. Adding up all such omitted terms, we arrive at an estimated error of about 3%, for a cycle-length truncated cycle expansion based on more than 10^9 pseudocycle terms! On the other hand, truncating by stability at say $\Lambda_{\max} = 3000$, only 409 prime cycles suffice to attain the same accuracy of about 3% error, figure 15.

As the Farey map maps the unit interval onto itself, the leading eigenvalue of the Perron-Frobenius operator should equal $s_0 = 0$, so $1/\zeta(0) = 0$. Deviation from this exact result serves as an indication of the convergence of a given cycle expansion. The errors of different truncation schemes are indicated in figure 15. We see that topological length truncation schemes are hopelessly [112] bad in this case; stability length truncations are somewhat better, but still rather bad.

VIII. TWO-DIMENSIONAL MAPS

The local approximation to $\mathcal{L}_\sigma$ seen in sect. IV can be repeated with a map of arbitrary dimension d. For simplicity, we expand the map around the origin. Assuming isotropic Gaussian noise, $\mathcal{L}_\sigma$ acts on a density $\rho(v)$ as

$$\mathcal{L}_\sigma \rho(v) = \int \exp\left(-\frac{(v-\mathbf{M}x)^T(v-\mathbf{M}x)}{4D}\right)\rho(x)d^dx\,, \tag{59}$$

[112] PC: why is error in figure 15larger when smoothed?

plot here

FIG. 15: Comparison of cycle expansion truncation schemes for the logistic map (??); the deviation of the truncated cycles expansion for $|1/\zeta_N(0)|$ from the exact flow conservation value $1/\zeta(0) = 0$ is a measure of the accuracy of the truncation. The jagged line is logarithm of the stability ordering truncation error; the smooth line is smoothed according to sect. **??**; the diamonds indicate the error due the topological length truncation, with the maximal cycle length N shown. They are placed along the stability cutoff axis at points determined by the condition that the total number of cycles is the same for both truncation schemes.

where $\mathbf{M}$ is the jacobian matrix of the map calculated in the origin. If $\mathbf{M}$ is symmetric and can be brought to a diagonal form by an orthogonal matrix, then the problem can be easily decomposed and the eigenfunctions of $\mathcal{L}_\sigma$ are just the product of the one-dimensional eigenfunctions found in sect. IV, each one pointing along an eigenvector of $\mathbf{M}$. The above equation can be written as

$$\mathcal{L}_\sigma \rho(v) = \int \exp\left(-\frac{x^T J^T J x - 2x^T J^T v + v^T v}{4D}\right) \rho(x) d^d x \,. \tag{60}$$

Let

$$\rho(v) = \rho_1(U^T v)_1 ... \rho_d(U^T v)_d \tag{61}$$

where U is the matrix of the eigenvectors of J. With this guess, equation (60) reads

$$\mathcal{L}_\sigma \rho(v) = \int \exp\left(-\frac{z^T \Lambda^T \Lambda z - 2z^T \Lambda^T U^T v + v^T v}{4D}\right) \rho_1(z_1) ... \rho_d(z_d) d^d z \,, \tag{62}$$

where Λ is the diagonal matrix of eigenvalues of J. Now the previous expression can be split into d independent integrals. The functions $\rho_i(U^T v)_i$ are either of the form (19) or (18) depending on whether $|\Lambda_i|$ is greater or less than unity.

As a simple example, let us consider a two-dimensional map. In the vicinity of a point of an unstable cycle, with $|\Lambda_1| > 1, |\Lambda_2| < 1$, the local eigenfunctions of $\mathcal{L}_\sigma$ are

$$\rho_{k,l}(v) = \frac{1}{4\pi D} H_k[\alpha_1 (U^T v)_1] H_l[\alpha_2 (U^T v)_2] e^{-[\alpha_2 (U^T v)_2]^2} \tag{63}$$

and the local eigenfunctions of $\mathcal{L}_\sigma^\dagger$ are

$$\tilde{\rho}_{k,l}(v) = \frac{1}{4\pi D} H_k[\alpha_1 (U^T v)_1] e^{-[\alpha_1 (U^T v)_1]^2} H_l[\alpha_2 (U^T v)_2] \tag{64}$$

The first eigenfunction of the spectrum is a tube of Gaussian section that goes along the direction of the unstable eigenvector of J (the eigenfunction of the adjoint operator goes along the stable direction). *Thus isotropic noise does not affect the orientation of these eigenfunctions in the plane, but only the width of their Gaussian section.*

A. Neighborhoods

We define the neighborhood of a hyperbolic periodic point by analogy with the deterministic case [40], After having constructed pinning coordinates, he and take the intersection between the images of all the iterations of the map, that runs forward in the stable direction and backward in the unstable direction. For the product of two 1–d maps

$$f(x,y) = (\Lambda_1 x, \Lambda_2 y)\,, \qquad |\Lambda_1| < 1\,, \quad |\Lambda_2| > 1$$
$$f_x^{-n} \bigcap_{n\to\infty} f_y^n = (0,0)$$

One obtains the same result by taking the intersection between the sets of points for which the eigenfunctions of the corresponding evolution operator and its adjoint are non-zero. The zeroth order eigenfunctions of the evolution operator and its adjoint are

$$\mathcal{L}_f[h_0(x,y)] = |\Lambda_1^{-1}|\delta(x) \tag{65}$$

and

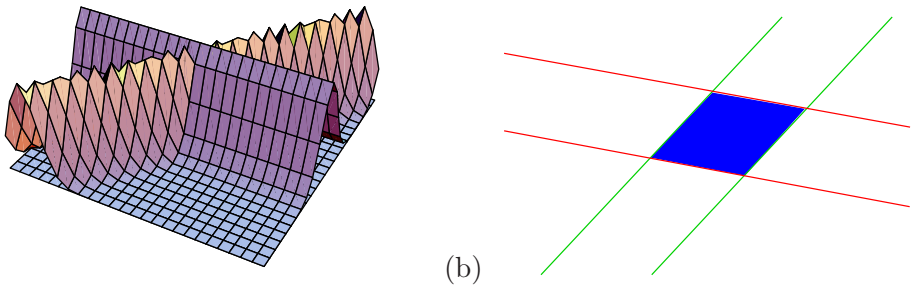
$$\mathcal{L}_f^\dagger[\tilde{h}_0(x,y)] = |\Lambda_2^{-1}|\delta(y)\,. \tag{66}$$

The subset of R^2 for which both h and $\tilde{h}$ are non-zero is the fixed point $(0,0)$. In the noisy case, the eigenfunctions involved are all smooth, so we define the noise-smeared neighborhood of the fixed point as the intersection of the width of two Gaussians, the the parallelepiped

$$\{(U^T v)_1 \in [-\sqrt{2}/\alpha_1, \sqrt{2}/\alpha_1]\} \bigcap \{s(U^T v)_2 \in [-\sqrt{2}/\alpha_2, \sqrt{2}/\alpha_2]\} \tag{67}$$

For a generic flow, the jacobian $\mathbf{M}$ cannot be diagonalized by an orthogonal matrix. I have not figured out how to find the whole eigenspectrum of $\mathcal{L}_\sigma$ and of its adjoint. Yet, $\rho_{0,0}$ and $\tilde{\rho}_{0,0}$ are all we need to define a neighborhood as in (67) and it turns out that they have the same form as (63) and (64):

$$\rho_{0,0}(v) = \frac{1}{4\pi D} H_0[\alpha_1 (U^T v)_1] H_0[\alpha_2 (U^T v)_2] e^{-[\alpha_2 (U^T v)_2]^2} \tag{68}$$



(a) (b)

FIG. 16: (a) The stable manifold (68) leading eigenfunction of $\mathcal{L}_\sigma$ and the unstable manifold (69) leading eigenfunction of the adjoint (backward in time) $\mathcal{L}_\sigma^\dagger$. (b) Visualization of the neighborhood of a periodic point as the intersection of the forward in time (68) and backward in time (69) Gaussians.

$$\tilde{\rho}_{0,0}(v) = \frac{1}{4\pi D} H_0[\alpha_1 (V^T v)_1] e^{-[\alpha_1 (f V^T v)_1]^2} H_0[\alpha_2 (V^T v)_2] \tag{69}$$

The H_0's are constants, whereas the matrices U and V are such that the "Gaussian tubes" (68) and (69) are extended respectively along the directions of the unstable and stable eigenvectors of $\mathbf{M}$. [113]

To summarize, my working definition of the noise-smeared neighborhood of a fixed point (and therefore, any cycle point as well) is: the intersection of (the leading eigenfunction of $\mathcal{L}_\sigma$ along the contracting direction) $\bigcap$ (the leading eigenfunction of the adjoint, backward in time $\mathcal{L}_\sigma^\dagger$ along the expanding direction).

IX. LOZI MAP

The method developed above is here applied to very simple uniformly hyperbolic examples of a chaotic repeller and a chaotic attractor in order to give explict examples of noise-limited finite partitions of the state space, and the corresponding evolution operators written in a finite matrix approximations.

As an illustration of the sense in which the set of unstable periodic orbits forms the skeleton of the chaotic state space explored by the long-time dynamics, compare the strange

[113] PC: save the mathematica program that generates them in .../lippolis/figsSrc/mathematica

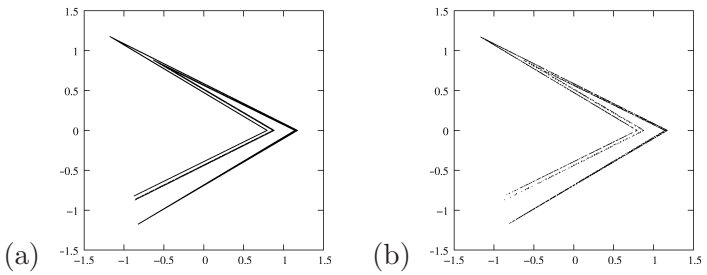


FIG. 17: Lozi map (a) strange attractor, (b) periodic points of length $n = 12$.

attractor of the Lozi map, figure 17(a)

$$\begin{aligned} x_{n+1} &= 1 - a|x_n| + by_n \\ y_{n+1} &= x_n \,. \end{aligned} \tag{70}$$

for parameter values $a = 1.85$, $b = 0.3$ with the set of all periodic points of period $n = 12$, figure 17(b). A few fixed points and shortest cycles already pin down the position and the shape of the strange attractor. Longer and longer unstable cycles partition the attractor arbitrarily finely. [114]

Consider now a noisy Lozi map [115]

$$\begin{aligned} x_{n+1} &= 1 - a|x_n| + by_n + \xi_{x,n} \\ y_{n+1} &= x_n + \xi_{y,n} \,. \end{aligned} \tag{71}$$

For the area contraction parameter $b = -1$ the deterministic version of this map is area-preserving, that is, Hamiltonian. For the "stretching" parameter a value $a = 6$, the non–wandering set of this Hamiltonian map is a repeller. For the purpose of illustrating the concept, the noise strength is set at $D = 0.01$. Compared to the size of the system, this is not very weak, but it is realistic for practical applications. By doing so, it is easier to write the transition matrix for the partition, as well as the corresponding Markov graph. The symbolic dynamics of the deterministic system is complete, there is no pruning in this case.

Figure 18 shows two consecutive partitions of the state space of the repeller. The first one is made of neighborhoods calculated around the eight cycles of length three, in the

[114] PC: always plot the return map (x_n, x_{n+1}), for Hénon and Lozi, not the (x, y) coordinates. Why? the return map is the object that connects discrete maps to continuous time flows. In addition, y coordinate shrinks with decreasing b, while (x_n, x_{n+1}) plot remain of the same size, and nicely connects to the 1-d parabola limit as $b \to 0$.

[115] PC: as in (3), (8), noise sequence carries a time label

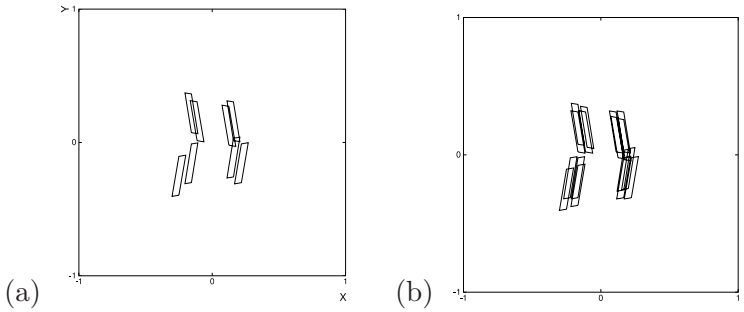


FIG. 18: Two consecutive partitions of the noisy Lozi repeller, with neighborhoods of (a) 3-cycles, and (b) 4-cycles.

second around the 16 cycles of length four. The size of each neighborhood is clearly exaggerated with respect to the size of the system, but, again, this is just for the purpose of illustration. We can see from the pictures that some of the regions do overlap, so that it is only possible to distinguish five regions in the first partition and three in the second. That leads to the conclusion that no more refined partition is possible with this level of noise. The cycles of the deterministic system (and their respective neighborhoods) are labeled $000, 001, 010, 011, 100, 101, 110, 111$. Label each cluster of overlapping neighborhoods by a letter: regions $(000, 001, 010) \equiv A$ overlap, as well as regions $(100, 110) \equiv B$, so that the transition matrix is

$$\begin{pmatrix} T_{A,A} & 0 & T_{A,B} & T_{A,101} & 0 \\ T_{011,A} & 0 & 0 & T_{011,101} & 0 \\ T_{B,A} & T_{B,011} & T_{B,B} & 0 & T_{B,111} \\ T_{101,A} & 0 & T_{101,B} & 0 & 0 \\ 0 & T_{111,011} & 0 & 0 & T_{111,111} \end{pmatrix}, \tag{72}$$

[116] from which a finite-memory Markov graph can be drawn. [117] Markov graph for the Lozi map chaotic repeller transition matrix (72).

The situation is very similar with a chaotic attractor. In the example of the Lozi map we set the parameters $a = 1.85$, $b = 0.3$, noise strength $D = 0.005$. This time there is heavy pruning.

[116] PC: Noch ein mal: If you do not find the Wolfram program (last page of current manuscript) useful, use xfig.

[117] PC: current figure **??** is a placeholder, replace by the correct Markov graph.

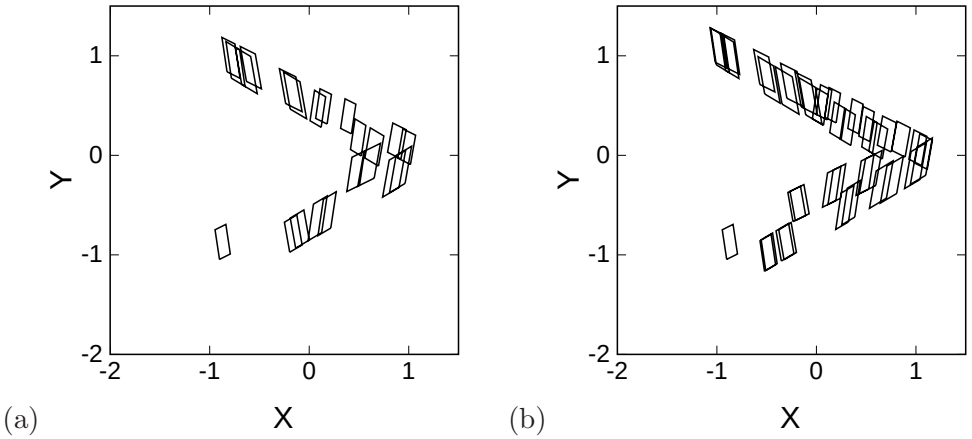


FIG. 19: Lozi $a = 1.85$, $b = 0.3$, noise strength $D = 0.005$ chaotic attractor, smeared out by noise: (a) partition with neighborhoods of period-5 cycles. (b) partition with period-6 cycles: no improvement in resolution.

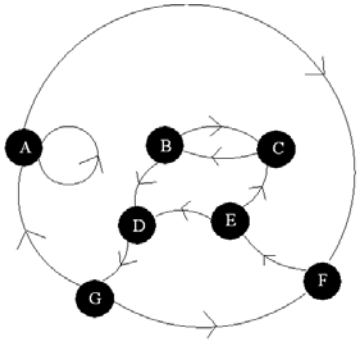


FIG. 20: The Markov graph generated by the partition in Figure 19 (a)

Figure 19 shows that the partition given by the neighborhoods of the 6-cycles not improve the resolution of the state space, compared to the partition given by the neighborhoods of the 5-cycles.

X. SUMMARY AND CONCLUSIONS

[118]

Acknowledgments.

We are grateful to William Mather for a careful reading of the manuscript and several

[118] PC: write a summary

good suggestions, and to Hans Fogedby, Ralf Metzler and Vattay Gábor for scholarly advice. P.C. thanks G. Robinson, Jr. for support.

APPENDIX A: ORNSTEIN-UHLENBECK PROCESS

Add weak uncorrelated noise $\hat{\xi}$ to the linear ODE [119]

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

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

$$\frac{dx}{dt} = \lambda x + \hat{\xi} \tag{A1}$$

leads to the Fokker-Planck equation [21, 22]

$$\partial_t \rho(x,t) + \partial_x[\lambda x \rho(x,t)] = D\,\partial_x^2 \rho(x,t)\,, \tag{A2}$$

known as the Ornstein-Uhlenbeck process [23, 41, 42]. The analytical solution is obtained by rewriting ρ as

$$\rho = e^{-\frac{U}{2D}}\psi\,, \tag{A3}$$

where $U = -\frac{\lambda}{2}x^2$ can be interpreted as the potential of the Langevin force in (A1). The multi-dimensional Ornstein-Uhlenbeck process with potential

$$U = -\frac{1}{2}\sum_{i,j} A_{i,j} x_i x_j \tag{A4}$$

is known in financial literature as the Vašiček model [43]. The equation for ψ has the Schrödinger form, with the quantum harmonic oscillator Hamiltonian:

$$\begin{aligned} -\partial_t \psi &= H\psi \\ H &= -D\partial_x^2 - \frac{1}{2}\partial_x^2 U + \frac{1}{4D}(\partial_x U)^2 = -D\partial_x^2 + \frac{\lambda}{2} + \frac{\lambda^2}{4D}x^2\,. \end{aligned} \tag{A5}$$

[120] A solution of the continuous time Fokker-Planck equation (A2) can be expanded in the harmonic oscillator eigenfunction basis as

$$\rho(x,t) = \sum_{k=0}^{\infty} C_k \psi_k(x) e^{-s_k t}\,, \tag{A6}$$

[119] PC: Lebowiz: for true OU process this is a velocity equation - there is also an equation for position. See Uhlenbeck and Chandrashekar reprint in Wax [41], CNS library.

[120] PC: make into ChaosBook problem

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

$$H\psi_k = -s_k\psi_k\,. \tag{A7}$$

The solutions are [22] [121]

$$\tilde{\psi}_k(x) = H_k(\mu x)e^{-(\mu x)^2}\,, \qquad \mu^2 = -\lambda/2D\,, \qquad s_k = -k\lambda \tag{A8}$$

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

$$\psi_k(x) = H_k(\mu x)\,, \qquad s_k = (k+1)\lambda\,, \tag{A9}$$

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

⇓PRIVATE

⇑PRIVATE

APPENDIX B: ROLL YOUR OWN CIGAR

1. Convolution of Gaussians

We want to evaluate the integral

$$\mathcal{L}_\sigma \cdot \rho_n(y) = \int dx^d\, e^{-\frac{1}{2}[(y-\mathbf{J}x)^T\frac{1}{\Delta}(y-\mathbf{J}x)+x^T\frac{1}{Q_n}x]} \tag{B1}$$

Introducing the variable $z = \mathbf{M}x$, the previous expression becomes

$$\mathcal{L}_\sigma \cdot \rho_n(y) = \int dz^d\, |\det \mathbf{M}|^{-1}\, e^{-\frac{1}{2}[(y-z)^T\frac{1}{\Delta}(y-z)+(\mathbf{M}^{-1}z)^T\frac{1}{Q_n}(\mathbf{M}^{-1}z)]} \tag{B2}$$

Now that can be regarded as as the convolution of two Gaussians

$$\begin{aligned} f(z) &= e^{-\frac{1}{2}z^T\frac{1}{\Delta}z} \\ g(z) &= e^{-\frac{1}{2}(\mathbf{M}^{-1}z)^T\frac{1}{Q_n}(\mathbf{M}^{-1}z)} \end{aligned} \tag{B3}$$

According to the convolution theorem, [122]

$$f * g = \frac{1}{2\pi}\int dp F(p)G(p)e^{ip\cdot x} \tag{B4}$$

[121] PC: Lebowiz: if x is velocity then $\exp(-x^2/2k_BT)$, one thinks of D as temperature.
[122] PC: reference to "convolution theorem" here

Now

$$\begin{aligned} F(p) &= \frac{1}{(2\pi)^d}\int d^d x f(x) e^{-ip\cdot x} = \sqrt{|\det \Delta|}\, e^{\frac{1}{2}p^T \Delta p} \\ G(p) &= \frac{1}{(2\pi)^d}\int d^d x g(x) e^{-ip\cdot x} = \sqrt{|\det\left[\mathbf{M}Q_n\mathbf{M}^T\right]|}\, e^{\frac{1}{2}p^T \mathbf{M}Q_n\mathbf{M}^T p} \end{aligned} \tag{B5}$$

so that

$$\begin{aligned} f * g &= \frac{1}{(2\pi)^d}\sqrt{|\det\left[\mathbf{M}Q_n\mathbf{M}^T\right]\det\Delta|}\int d^d p e^{\frac{1}{2}p^T(\Delta+\mathbf{M}Q_n\mathbf{M}^T)p+ip\cdot x} \\ &= \left|\frac{\det\left(\mathbf{M}Q_n\mathbf{M}^T\right)\det\Delta}{\det\left(\mathbf{M}Q_n\mathbf{M}^T+\Delta\right)}\right|^{1/2} e^{-\frac{1}{2}x^T(\Delta+\mathbf{M}Q_n\mathbf{M}^T)^{-1}x} \end{aligned} \tag{B6}$$

2. Fixed point Gaussian evolution

In general, the noise is anisotropic, with the isotropic diffusion weight $2D\delta_{ij}$ in (4) replaced by a symmetric diffusion matrix Δ_{ij}, and the kernel of the evolution operator in (11)

$$e^{-\frac{1}{2}(y-fx)^T\cdot\frac{1}{\Delta}\cdot(y-fx)}\,. \tag{B7}$$

We assume that the eigenvalues of Δ are strictly positive, so under one step of noisy dynamics an initial Dirac-δ localized density distribution is smeared out into a Gaussian ellipsoid whose widths and orientation are controlled by the eigenvalues and eigenvectors of Δ. [123] For isotropic noise $\Delta = 2D\,\mathbf{1}$, where D is the diffusion coefficient. The normalization constant c is fixed by $\int dx^d \delta_D(x) = 1$.

Consider a smooth map

$$x_{n+1} = f(x_n) + \xi_n \tag{B8}$$

with a fixed point at $x^* = 0$. In a neighborhood $x = x^* + y$ we approximate the map by the linear map (Jacobian matrix) $\mathbf{M} = \partial f(x^*)$ with the fixed point at $x = 0$:

$$y_{n+1} = \mathbf{M}y_n \tag{B9}$$

acting on a Gaussian density distribution, also centered on $x = 0$:

$$\rho_n = \frac{1}{C_n}e^{-\frac{1}{2}x^T\cdot\frac{1}{Q_n}\cdot x}\,. \tag{B10}$$

[123] PC: include Kadanoff book argument why eigs are non-negative

C_n is a normalization constant, fixed by $\int dx^d \rho_n = 1$. If the eigenvalues of Q_n are distinct, the distribution is a "cigar-shaped" ellipsoid, with eigenvectors of Q_n giving the orientation of various axes.

Convolution of a Gaussian with a Gaussian is again a Gaussian, so the noisy evolution operator (B7) maps the ellipsoid ρ_n into a ellipsoid ρ_{n+1} one time step later

$$\begin{aligned} \rho_{n+1}(y) &= \mathcal{L}_\sigma \rho_n(y) = \frac{1}{cC_n} \int dx^d \, e^{-\frac{1}{2}[(y-\mathbf{M}x)^T \frac{1}{\Delta}(y-\mathbf{M}x) + x^T \cdot \frac{1}{Q_n} \cdot x]} \\ &= \frac{1}{C_{n+1}} \, e^{-\frac{1}{2} y^T \cdot \frac{1}{Q_{n+1}} \cdot y} \, . \end{aligned} \tag{B11}$$

Complete the square, integrate over x, get the initial distribution squished by the flow and spread out by the noise:

$$Q_{n+1} = \mathbf{M} Q_n \mathbf{M}^T + \Delta \, . \tag{B12}$$

This says the two variances (the noise matrix Δ and the deterministically transported $Q_n \to \mathbf{M} Q_n \mathbf{M}^T$) add up as Gaussian variances, that is, sums of squares.

If $\mathbf{M}$ has all eigenvalues strictly contracting, $|\Lambda_j| < 1$, any initial compact measure (not only a initial distribution ρ_1 of Gaussian form) converges to the invariant natural measure ρ_0 whose variance satisfies the fixed point condition [124]

$$\begin{aligned} Q^* &= \mathbf{M} Q^* \mathbf{M}^T + \Delta \\ &= \Delta + \mathbf{M} \Delta \mathbf{M}^T + \mathbf{M}^2 \Delta (\mathbf{M}^T)^2 + \mathbf{M}^3 \Delta (\mathbf{M}^T)^3 + \cdots \\ &= \sum_{m,n=0}^{\infty} \mathbf{M}^n \Delta (\mathbf{M}^T)^m \delta_{m,n} \end{aligned} \tag{B13}$$

This is a convergent series as long as all eigenvalues of the Jacobian matrix are strictly contracting. To get rid of the $\delta_{m,n}$, use some integral representation of Kronecker δ (which one is natural?). For example, with a Fourier representation, get

$$\begin{aligned} \delta_{mn} &= \int \frac{d\theta}{2\pi} e^{i\theta(m-n)} \, , \\ Q^* &= \int \frac{d\theta}{2\pi} \sum_{m,n} (e^{-i\theta}\mathbf{M})^n \Delta (e^{i\theta}\mathbf{M}^T)^m \\ &= \int \frac{d\theta}{2\pi} \frac{1}{\mathbf{1} - e^{-i\theta}\mathbf{M}} \Delta \frac{1}{\mathbf{1} - e^{i\theta}\mathbf{M}^T} \end{aligned} \tag{B14}$$

gives an expression for the fixed point Q^* in d dimensions. But how do you evaluate it? Remember, we do not really need Q^*, we probably need its eigenvalues only. [125]

[124] PC: introduce expansion in terms of derivatives of Dirac δ's
[125] PC: not sure this is true

For $d = 1$ the width of the natural measure is

$$Q^* = \frac{2D}{1 - |\Lambda|^2}\,, \tag{B15}$$

a balance between exponential contraction by Λ and diffusive smearing out by $2D$. For strongly contracting Λ, the width is due to the noise only. As $|\Lambda| \to 1$ the width diverges, as the trajectories are no longer confined, but diffuse by Brownian motion.

3. Equilibrium Gaussian evolution

If you do not like maps, recast the problem back into the infinitesimal time step form:

$$\begin{aligned} \mathbf{M} &= e^{A\delta t}\,, \qquad \mathbf{M}^T = e^{A^T \delta t} \\ Q_{n+1} &= Q_n + B_n \delta t\,, \qquad \Delta \to \delta t \Delta\,. \end{aligned} \tag{B16}$$

Expanding to linear order yield equilibrium condition

$$0 = AQ + QA^T + \Delta\,. \tag{B17}$$

Not sure how to solve this in d-dimensions. In $1d$, $A \to \lambda$, $\Delta \to 2D$, and we reproduce the effective diffusion width (A8):

$$\hat{D} = -D/2\lambda\,. \tag{B18}$$

---

[1] P. Cvitanović, R. Artuso, R. Mainieri, G. Tanner, and G. Vattay, *Chaos: Classical and Quantum* (Niels Bohr Institute, Copenhagen, 2006), ChaosBook.org.

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

[3] P. Cvitanović, C. Dettmann, R. Mainieri, and G. Vattay, J. Stat. Phys. **93**, 981 (1998).

[4] P. Cvitanović, C. Dettmann, R. Mainieri, and G. Vattay, Nonlinearity **12**, 939 (1999).

[5] P. Cvitanović, C. Dettmann, G. Palla, N. Sondegard, and G. Vattay, Phys. Rev. E **60**, 3936 (1999).

[6] C. P. Dettmann, FIND **FIND**, FIND (FIND).

[7] M. B. Kennel and M. Buhl, Physical Review Letters **91**, 084102 (2003), arXiv:nlin/0304054.

[8] M. B. Kennel and M. Buhl, in *Experimental Chaos*, edited by V. in, L. Kocarev, T. L. Carroll, B. J. Gluckman, S. Boccaletti, and J. Kurths (2003), vol. 676 of *American Institute of Physics Conference Series*, pp. 380–380.

[9] C. S. Daw, C. E. A. Finney, and E. R. Tracy, Review of Scientific Instruments **74**, 915 (2003).

[10] M. Buhl and M. B. Kennel, Phys. Rev. E **71**, 046213 (2005).

[11] A. Einstein, Ann. Physik **17**, 549 (1905).

[12] S. M. Ulam, *A Collection of Mathematical Problems* (Interscience Publishers, New York, 1960).

[13] A. Lasota and M. MacKey, *Chaos, Fractals, and Noise; Stochastic Aspects of Dynamics* (Springer-Verlag, Berlin, 1994).

[14] F. Haake, *Quantum Signatures of Chaos* (Springer-Verlag, Berlin, 2001), 2nd ed.

[15] G. A. Hagedorn and A. Joye, *Exponentially accurate semiclassical dynamics: Propagation, localization, ehrenfest times, scattering and more general states* (1999), arXiv:math-ph/9911036.

[16] R. M. Angelo, L. Sanz, and K. Furuya, *Ordered quantization and the ehrenfest time scale* (2003), arXiv:quant-ph/0302020.

[17] L. Kaplan, Physical Review E **70**, 026223 (2004), arXiv:nlin/0406054.

[18] L. Kaplan, Physical Review E **72**, 036214 (2005), arXiv:nlin/0507046.

[19] P. W. Brouwer, *Semiclassical theory of the Ehrenfest-time dependence of quantum transport* (2007), arXiv:0706.1487.

[20] F. Cametti and C. Presilla, Phys. Rev. Lett. **89**, 040403 (2002), arXiv:quant-ph/0201147.

[21] H. Dekker and N. V. Kampen, Physics Letters **73A**, 374 (1979).

[22] P. Gaspard, G. Nicolis, A. Provata, and S. Tasaki, Phys. Rev. E **51**, 74 (1995).

[23] H. Risken, *The Fokker-Planck Equation* (Springer-Verlag, 1996).

[24] Abramowitz, M. and Stegun, I. A., eds., *Handbook of Mathematical Functions* (Dover Publications, New York, 1964).

[25] H. Takada, Y. Kitaoka, and Y. Shimizua, Forma **16**, 1746 (2001).

[26] R. Artuso, E. Aurell, and P. Cvitanović, Nonlinearity **3**, 361 (1990).

[27] L. Zhu, Y.-C. Lai, F. C. Hoppensteadt, and E. M. Bollt, Chaos **13**, 410 (2003).

[28] G. Froyland, Commun. Math. Phys. **189**, 237 (1997).

[29] G. Froyland, Discrete and Continuous Dynamical Systems **17**, 671 (2007).

[30] G. Froyland, Nonlinearity **12**, 79 (1999).

[31] G. Froyland, *Extracting dynamical behaviour via markov models* (2000), URL http://math-www.uni-paderborn.de/~froyland/froyland.ps.gz.

[32] L. Billings and E. Bolt, Chaos Solitons and Fractals **12**, 365 (2001), see also www.mathstat.concordia.ca/pg/bilbollt.html.

[33] A. Boyarsky and P. Góra, *Laws of Chaos: Invariant measures and dynamical systems in one dimension* (Birkhüser, Boston, 1997).

[34] A. Ostruszka and K. Zyczkowski, Phys. Lett. A **289**, 306 (2001), arXiv.org:nlin/0106003.

[35] P. Gaspard, G. Nicolis, A. Provata, and S. Tasaki, Phys. Rev. E **51**, 74 (1995).

[36] P. Bevington and D. Robinson, *Data Reduction and Error Analysis for the Physical Sciences* (McGraw-Hill, New York, 1992).

[37] P. Dahlqvist and G. Russberg, J. Phys. A **24**, 4763 (1991).

[38] C. P. Dettmann and G. P. Morriss, Phys. Rev. Lett. **78**, 4201 (1997).

[39] C. P. Dettmann and P. Cvitanović, Phys. Rev. E **56**, 6687 (1997).

[40] H. Rugh, Nonlinearity **5**, 1237 (1992).

[41] E. Wax, *Selected Papers on Noise and Stochastic Processes* (Dover, New York, 1954).

[42] G. E. Uhlenbeck and L. S. Ornstein, Physical Review **36**, 823 (1930).

[43] O. Vasicek, Journal of Financial Economics **5**, 177 (1977), URL http://dx.doi.org/10.1016/0304-405X(77)90016-2.

[44] I. Abu-Falahah and J. L. Torrea, Glasgow Math. J. **48**, 203 (2006).

[45] S. N. Evans (2007), www.stat.berkeley.edu/users/evans/206notes_Spring06.pdf.

[46] M. Jacobsen, Bernoulli J. of Math. Stat. and Prob. **2**, 271 (1997).

[47] O. Kallenberg, *Foundations of Modern Probability* (Springer, 1997).

[48] K. M. Lie and X. B. Liu, J. Appl. Mechanics **71**, 677 (2004).

[49] P. Sjögren, J. Fourier Anal. Appl. **3**, 813 (1997).

[50] P. F. Gora, *Stationary distributions of a noisy logistic process* (2005), arXiv.org:cond-mat/0502389.

[51] E. Bollt, P. Gora, A. Ostruszka, and K. Zyczkowski, *Basis markov partitions and transition matrices for stochastic systems* (2006), arXiv.org:nlin/0605017.

[52] P. Gaspard, J. Stat. Phys. **106**, 57 (2002).

[53] S. Subbiah and D. J. Driebe, Chaos **8**, 741 (1998).

[54] D. J. Driebe and G. E. Ordnez, Phys. Lett. A **211**, 204 (1996).

[55] G. E. Ordonez and D. J. Driebe, J. Stat. Phys. **84**, 269 (1996).

[56] A. J. Scott and G. J. Milburn, J. Phys. A **34**, 7541 (2001), URL http://arxiv.org/abs/nlin/0102039.

[57] I. García-Mata and M. Saraceno, Phys. Rev. E **69**, 056211 (2004), arXiv:nlin/0312062.

[58] J. Weber, F. Haake, and P. Šeba, Phys. Rev. Lett. **85**, 3620 (2000), arXiv:nlin/0001013.

[59] J. Weber, F. Haake, P. A. Braun, C. Manderfeld, and P. Seba, J. Phys. A **34**, 7195 (2001), arXiv:nlin/0105047.

[60] C. Manderfeld, J. Weber, and F. Haake, J. Phys. A **34**, 9893 (2001), arXiv:nlin/0107020.

[61] M. Dellnitz, G. Froyland, and S. Sertl, Nonlinearity **13**, 1171 (2000).

[62] F. Akesson and J. Lehoczky, *Discrete eigenfunction expansion of multi-dimensional brownian motion and the ornstein-uhlenbeck process* (1998), URL citeseer.ist.psu.edu/akesson98discrete.html.

[63] A. Caticha and C. Cafaro, *From information geometry to Newtonian dynamics* (2007), arXiv:0710.1071.

[64] S. V. Malinin and V. Y. Chernyak, *Classical nonlinear response of a chaotic system: Langevin dynamics and spectral decomposition* (2007), arXiv:nlin/0703014.

[65] C.-L. Ho and R. Sasaki, *Deformed multi-variable Fokker-Planck equations* (2007), cond-mat/0703291.

APPENDIX C: FLOTSAM

Flotsam

HISTORY

Gábor Vattay, 21 Oct 2006:

Your notes contain an important new piece of technical information, the basis functions in terms of Hermite polynomials. We overlooked this before and could lead to much more tractable formulas than the ones we had 10 years ago.

Since this result is so nice, I immediately checked who invented it for the OU Fokker-Planck. According to the attached paper by Takada et al it was invented by Abramowitz and Stegun in 1964.

Concerning the neighborhoods, I think it is too early to deal with their detailed definitions. I would concentrate now on what can be gained from this new technical development, such as an exact expression of the spectrum and then in the process of systematic approximation a useful definition of neighborhood can come out naturally.

Please have a look at refs. [44–49], perhaps of interest.

Sep 20 2007: Weinan E's work looks interesting - you might want to check out his stochastic publications on www.math.princeton.edu/∼weinan/publist.html

VISUALIZATION

Predrag, Aug 2006: S. Wolfram told me that you should download from mathematica.com package

discrete math graphplot

with which you can draw Markov diagrams automatically. Let me know if you run into problems.

The Geometry Center summer undergraduate project

www.geom.uiuc.edu/ megraw

offers some cute visualizations of stable/unstable manifolds.

RANDOM SOURCES

Predrag Nov 10, 2007:

Stationary solutions to a Fokker-Planck equation corresponding to a noisy logistic equation with correlated Gaussian white noises are constructed by Gora [50].

Bollt *et al.* [51]: " We analyze dynamical systems subjected to an additive noise and their deterministic limit. In this work, we will introduce a notion by which a stochastic system has something like a Markov partition for deterministic systems. For a chosen class of the noise profiles the Frobenius-Perron operator associated to the noisy system is exactly represented by a stochastic transition matrix of a finite size K. This feature allows us to introduce for these stochastic systems a basis–Markov partition, defined herein, irrespectively of whether the deterministic system possesses a Markov partition or not. We show that in the deterministic limit, corresponding to K –¿ infinity, the sequence of invariant measures of the noisy systems tends, in the weak sense, to the invariant measure of the deterministic system. Thus by introducing a small additive noise one may approximate transition matrices and invariant measures of deterministic dynamical systems."

Predrag Nov 6, 2007:

Include Gaspard's noisy trace formula derivation [52] into your thesis.

Predrag Oct 27, 2007:

Potentially of interest:

Subbiah and Driebe@ [53]: author = "S. Subbiah and D. J. Driebe", "Spectral decomposition of the tent map with varying height"

Predrag should check whether ref. [54, 55] "Using symmetries of the Frobenius-Perron operator to determine spectral decompositions" is of use for ChaosBook discrete.tex chapter.

Spina and Saraceno, 0710.4321 say:

" We consider the dicontinuous mapping introduced in [56] from the point of view of the spectral properties of the Liouville dynamics and study both the classical and the quantum Liouvillian at limited phase space resolution.

In pure hyperbolic systems the asymptotic decay of classical correlation functions is exponential and the decay rates can be rigorously obtained from the Ruelle-Pollicott (RP) resonances [**?**]. To study the long time behavior of general chaotic and mixed systems, which are beyond the validity of the RP theorem, non rigorous methods have been developed

to compute the resonances of the Frobenius-Perron(FP)propagator. All these approaches are based on a coarse grained Liouville dynamics of the density function in the limit of zero coarse graining [**?**].

In [57] the blurring of phase space structures is implemented by adding a diffusive noise in the Liouville equation, which results in a coarse graining of the FP propagator and the limit of vanishing noise is finally considered .

An alternative method developed in [58–60] uses a truncation of the infinite unitary FP operator to a finite dimension N in a basis of functions ordered by increasing resolution. The eigenvalues of this non unitary NXN operator are calculated and the the limit $N \to \infty$ is taken.

We compute the spectrum of the classical coarse grained FP and quantum (Husimi) propagator for the discontinuous map on the sphere presented in [56], using the truncation method of [58–60].

We follow closely the procedure that Haake and collaborators developed to study a system with mixed dynamics: the kicked top. We show that for low resolution classical and quantum mechanical propagators coincide, although the convergence is slow. We study the behavior of the eigenvalues and eigenfunctions of the truncated propagator with increasing resolution and compare our results to the ones obtained for a map with a mixed phase space [58, 59].
"

CREATE lippolis.bib entry: Ru Ruelle D. 1986 *Phys. Rev. Lett.* **56** 405

CREATE lippolis.bib entry: Fi Fishman S. and Rahav S. 2002 *Lectures Notes in Physics: Dynamics of Dissipation*, (eds.Garbaczewski P. and Olkiewicz R., Springer-Verlag, Berlin)

Gary Froyland, web.maths.unsw.edu.au/ froyland,
has many papers on Markov models, Ulam approximations to eigenfunctions of hyperbolic maps, etc. [29, 30, 61] I put ref. [31] into `ChaosBook.org/library`, seems quite useful.

Most Froyland transition matrix computations use the GAIO software package
www.math.uni.paderborn.de/ agdellnitz/gaio .

Akesson and Lehoczky [62] paper might have something to say about the eigenfunctions for multi-dimensional Ornstein-Uhlenbeck Process.

A wild paper [63]: derives Newtonian dynamics from information geometry. Might be worth a look.

D-optimal design measure for parameter estimation

optimal designs for model discrimination and parameter estimation

Atkinson & Fedorov (1975)

Ref. [64]: "We consider the classical response of a strongly chaotic Hamiltonian system. The spectrum of such a system consists of discrete complex Ruelle-Pollicott (RP) resonances which manifest themselves in the behavior of the correlation and response functions. We interpret the RP resonances as the eigenstates and eigenvalues of the Fokker-Planck operator obtained by adding an infinitesimal noise term to the first-order Liouville operator. We demonstrate how the deterministic expression for the linear response is reproduced in the limit of vanishing noise. For the second-order response we establish an equivalence of the spectral decomposition with infinitesimal noise and the long-time asymptotic expansion for the deterministic case."

Ref. [65]: too specific to the integrable Calogero model.

A blog entry on stochastic processes:

iefnncd.livejournal.com/2007/01/21/

with jewels such as: "predrag cvitanovic, C. P. dettmann, ronnie mainier and Gábor Vattay, 'Trace formulas for stochastic phylogenesis operators: glassy conjugationmethod,' chao-dyn/9811003 saint james the apostle davidson, stochastic point of accumulation hypothesis."

Yuri Bakhtin, Assistant Professor, School of Mathematics seems to have expertise in stochastic systems

TEXT CLIPS

follows from the chain rule for matrix derivatives.

$$\frac{\partial}{\partial x_i} f_j(f(x)) = \sum_{k=1}^{d} \frac{\partial}{\partial y_k} f_j(y)\Big|_{y=f(x)} \frac{\partial}{\partial x_i} f_k(x)\,.$$

COMMENTS AND SPECULATION (NOTES)

Here are some issues raised by Bartsch and Davidchack.

My criterion for overlapping is indeed arbitrary. I pick one half-width (68thought things may be made more rigorous by defining the probability of two adjacent border points to be swapped: given two Gaussians $\rho_a(x)$ and $\rho_b(x)$, centered respectively at x_a and x_b, $x_a < x_b$, the probability that a random point x, distributed according to $\rho_a(x)$, be greater or equal

than x_b is,

$$P(x > x_b) = \int_{x_b}^{\infty} \rho_a(x)dx = \int_{x_b}^{\infty} dx C_a e^{-\beta_a^2(x-x_a)^2}$$

On the other hand, the probability that a random point y, distributed as $\rho_b(x)$, be less or equal than x_a is

$$P(y < x_a) = \int_{-\infty}^{x_a} \rho_b(y)dy = \int_{-\infty}^{x_a} dy C_b e^{-\beta_b^2(x-x_b)^2}$$

Now I say that the sum of these two probabilities is the probability of two border points to be swapped, and that we could choose a certain threshold value as significant overlapping.

Now take a point x in the partition interval $[x_a, x_b]$. Its probability to be instead in either of the two adjacent intervals to $[x_a, x_b]$ is the sum

$$\int_{x}^{\infty} dy C_a e^{-\beta_a^2(y-x_a)^2} + \int_{-\infty}^{x} dy C_b e^{-\beta_b^2(x-x_b)^2}$$

So the probability of all the points of an interval $[x_a, x_b]$ of the deterministic partition to actually lie in the adjacent intervals is

$$\frac{\int_{x_a}^{x_b} dx \int_{x}^{\infty} dy C_a e^{-\beta_a^2(y-x_a)^2}}{\int_{x_a}^{x_b} dx} + \frac{\int_{x_a}^{x_b} dx \int_{-\infty}^{x} dy C_b e^{-\beta_b^2(x-x_b)^2}}{\int_{x_a}^{x_b} dx}$$

The formula might be wrong, but the idea is that, even when the optimal resolution is not yet reached, the Markov graph that arises from our analysis should be stochastic, that is every entry in the matrix of transitions should be a probability.

Davidchack suggests numerical tests to verify whatever we predict (take a bunch of local densities and iterate them, see where they end up, make a statistics). Bartsch says a "real calculation" would be needed, such as Lyapunov exponent or similar, before submitting any paper (we're rather far from that, anyway). The problem is that this has already been done in the famous trilogy of papers, the result does not depend on noise for, say, Ulam's map, so it would be a matter of stopping at some point and claiming a limit of sig figs, or deriving a different trace formula (using Hermite as basis?), which might improve the convergence but I believe the terms of the expansion would still be independent of the noise.

DL: I also talked to Eckhardt, he says another problem is that different regions overlap at different times, so when would we truncate the expansion? I think maybe stability ordering would make more sense, since the least unstable regions are the ones that overlap first, but it's just a thought.

PC Sep 20 2007: Domenico, I do not understand your conversation with Eckhardt. Our point is precisely that different regions overlap at different times, and this paper shows how

this is done. Am I missing something? Stability ordering is a *deterministic* partition up to given neighborhood size, here noise determines the optimal partition.