Optimal resolution of the state space of a chaotic flow in presence of noise

D. Lippolis[1], P. Cvitanović

Center for Nonlinear Science, School of Physics, Georgia Institute of Technology, Atlanta, GA 30332-0430

Abstract

Understanding chaotic dynamics often requires us to make arbitrarily fine partitions a fractal state space. In reality, the inevitable presence of some background noise sets a finite scale in every system, as well as a limit to the finest possible resolution that can be attained. We first analyze the interplay of the deterministic dynamics with the noise in the neighborhood of a periodic orbit of a map, by using a discretized version of Fokker-Planck formalism. Then we propose a method to determine the 'optimal resolution' of the state space, based on solving Fokker-Planck's equation locally in the non-wandering set (i.e. near the unstable periodic orbits) of the deterministic system, and test our hypothesis on unimodal maps.

1. Introduction

The effect of noise on the behavior of a nonlinear dynamical system is a fundamental problem in many areas of science [1, 2, 3], and the interplay of noise and chaotic dynamics is of particular interest [4, 5, 6]. Our purpose here is two-fold. First, we address operationally the fact that weak noise limits the attainable resolution of the state space of a chaotic system by formulating the *optimal partition hypothesis*. Second, we show that the optimal partition hypothesis replaces the Fokker-Planck PDEs by finite, low-dimensional matrix Fokker-Planck operators, with finite cycle expansions, optimal for a given level of precision, and whose eigenvalues give good estimates of long-time observables (escape rates, Lyapunov exponents, etc.).

A chaotic trajectory explores a strange attractor, and for chaotic flows evaluation of long-time averages requires effective partitioning of the state space into smaller regions. In a hyperbolic, everywhere unstable deterministic dynamical

[1]Corresponding author
Email address: domenico@gatech.edu (P. Cvitanović)

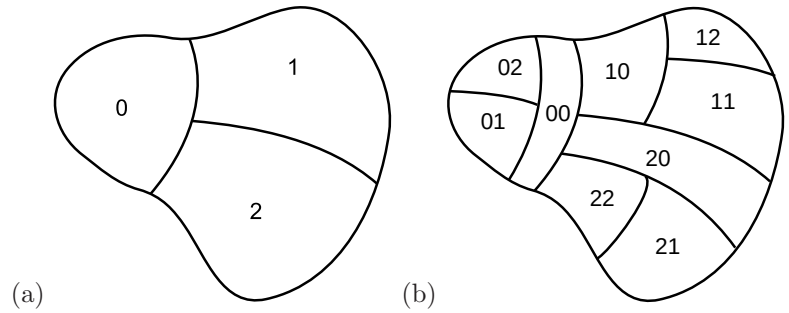

Figure 1: (a) A coarse partition of state space $\mathcal{M}$ into regions $\mathcal{M}_0$, $\mathcal{M}_1$, and $\mathcal{M}_2$, labeled by ternary alphabet $\mathcal{A} = \{1, 2, 3\}$. (b) A 1-step memory refinement of the partition of figure 1, with each region $\mathcal{M}_i$ subdivided into $\mathcal{M}_{i0}$, $\mathcal{M}_{i1}$, and $\mathcal{M}_{i2}$, labeled by nine 'words' $\{00, 01, 02, \cdots, 21, 22\}$.

system, consecutive Poincaré section returns subdivide the state space into exponentially growing number of regions, each region labeled by a distinct finite symbol sequence, as in figure 1. In the unstable directions these regions stretch, while in the stable directions they shrink exponentially. The set of unstable periodic orbits forms a 'skeleton' that can be used to implement such partition of the state space, each region a neighborhood of a periodic point [7, 8]. Longer and longer cycles yield finer and finer partitions as the neighborhood of each unstable cycle p shrinks exponentially with cycle period as $1/|\Lambda_p|$, where Λ_p is the product of cycle's expanding Floquet multipliers. As there is an exponentially growing infinity of longer and longer cycles, with each neighborhood shrinking asymptotically to a point, a deterministic chaotic system can - in principle - be resolved arbitrarily finely. But that is a fiction for any of the following reasons:

- any physical system experiences (background, observational, intrinsic, measurement, $\cdots$) noise
- any numerical computation is a noisy process due to the finite precision of each step of computation
- any set of dynamical equations models nature up to a given finite accuracy, since degrees of freedom are always neglected
- any prediction only needs to be computed to a desired finite accuracy

The problem we address here is sketched in figure 2; while a deterministic partition can, in principle, be made arbitrarily fine, in practice any noise will blur the boundaries and render the best possible partition finite. Thus our task is to determine the optimal attainable resolution of the state space of a given hyperbolic dynamical system, affected by a given weak noise. This we do by formulating the *optimal partition* hypothesis which we believe determines the best possible state space partition for a desired level of predictive precision. We know of no practical way of computing the 'blurred' partition boundaries of

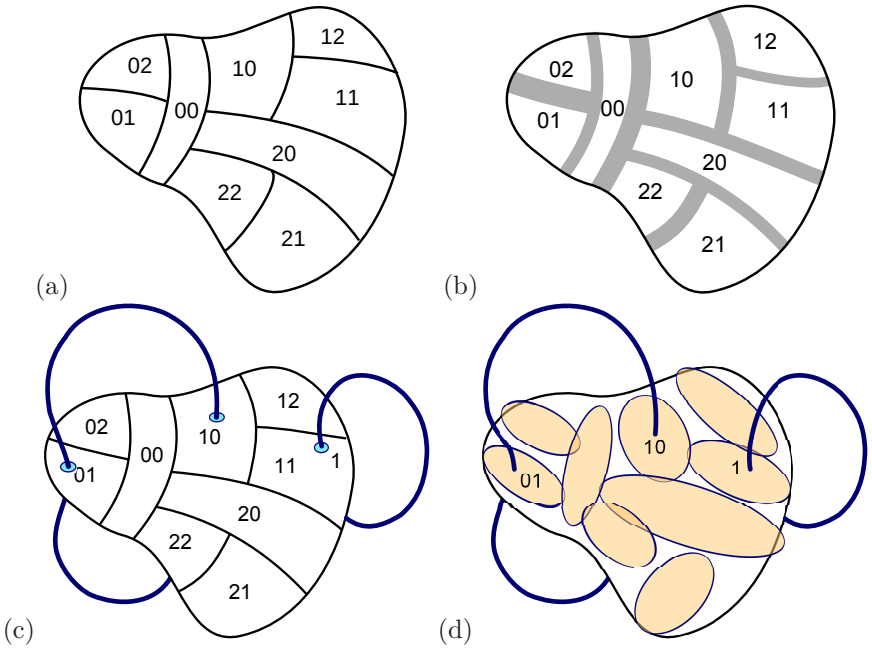

Figure 2: (a) A deterministic partition of state space $\mathcal{M}$. (b) Noise blurs the partition boundaries. At some level of deterministic partitioning some of the boundaries start to overlap, preventing further local refinement of adjecent neighborhoods. (c) The fixed point $\overline{1} = \{x_1\}$ and the two-cycle $\overline{01} = \{x_{01}, x_{10}\}$ are examples of the shortest period periodic points within the partition regions of (a). The *optimal partition* hypothesis: (d) Noise blurs periodic points into cigar-shaped trajectory-centered densities explored by the Langevin noise. The optimal partition hypothesized in this paper consists of the maximal set of resolvable periodic point neighborhoods.

figure 2 (b). Instead, we propose to determine the optimal partition in terms of blurring of periodic point neighborhoods, as in figure 2 (d). As we demonstrate in sect. 6.1, our implementation requires determination of only a small set of solutions of the deterministic equations of motion.

Intuitively, the noise smears out the neighborhood of a periodic point, whose size is now determined by the interplay between the diffusive spreading parameterized [11, 12, 13] by a diffusion constant D, and its exponentially shrinking deterministic neighborhood. If the noise is weak, the short-time dynamics is not altered significantly: short periodic orbits of the deterministic flow still coarsely partition the state space. As the periods of periodic orbits increase, the diffusion always wins, and successive refinements of a deterministic partition of the state space stop at the finest attainable partition, beyond which the diffusive smearing exceeds the size of any deterministic subpartition.

There is a considerable literature (reviewed here in sect. 5) on interplay of noise and chaotic deterministic dynamics, and the closely related problem of limits on the validity of the semi-classical periodic orbit quantization. All of this literature implicitly assumes uniform hyperbolicity and seeks to define a single, globally averaged diffusion induced average resolution (Heisenberg time, in the context of semi-classical quantization). However, the local diffusion rate differs from a trajectory to a trajectory, as different neighborhoods merge at different times, so there is no one single time beyond which noise takes over. Nonlinear dynamics interacts with noise in a nonlinear way, and methods for implementing the optimal partition for a given noise still need to be developed. This paper is an attempt in this direction. Here we follow and expand upon the Fokker-Planck approach to the 'optimal partition hypothesis' introduced in ref. [14].

What is novel here is that we show how to compute the *locally* optimal partition, for a given dynamical system and given noise, in terms of local eigenfunctions of the forward-backward actions of the Fokker-Planck operator and its adjoint. This effort brings a handsome reward: as the optimal partition is always finite, the dynamics on this 'best possible of all partitions' is encoded by a finite transition graph of finite memory, and the Fokker-Planck operator can be represented by a finite matrix. In addition, while the state space of a generic deterministic flow is an infinitely interwoven hierarchy of attracting, hyperbolic, 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/ellipticity, and the power-law correlation decays caused by marginally stable regions of state space.

The dynamical properties of high-dimensional flows are not just simple extensions from lower dimensional systems, and an extension of the Ruelle / Gutzwiller periodic orbit theory to high-dimensional dynamics would be an important advance. If such flow has only a few expanding directions, the above set of overlapping stochastic 'cigars' might provide an optimal, manageable cover of the long-time chaotic attractor embedded in a state space of arbitrarily high dimension.

The requisite Langevin / Fokker-Planck description of noisy flows is reviewed in sect. 2. This discussion leans heavily on the deterministic dynamics and periodic orbit theory notation, summarized in Appendix A. In sect. 3 we derive the formulas for the size of noise-induced neighborhoods of attractive fixed and periodic points, for maps and flows in arbitrary dimension. In order to understand the effect on noise on the hyperbolic, mixed expanding / contracting dynamics, we study the eigenfunctions of the Fokker-Planck operator in linear neighborhood of a fixed point (the Ornstein-Uhlenbeck process) of a noisy one-dimensional map in sect. 4, and show that the neighborhood along unstable directions is fixed by the evolution of a Gaussian density of trajectories under the action of the adjoint operator $\mathcal{L}^\dagger$, sect. 4.1. Having defined the local neighborhood of every periodic point, we turn to the global partition problem. Previous attempts at state space partitioning are reviewed in sect. 5. We formulate our optimal partition hypothesis in sect. 6: track the diffusive widths of unstable periodic orbits until they start to overlap. We test the approach by applying it to a 1-dimensional repeller, and in sect. 7, we assess the accuracy of our method by computing the escape rate and the Lyapunov exponent of the map, discuss weak noise corrections, and compare the results with a discretization of the Fokker-Planck operator on a uniform mesh. In sect. 8 we address the problem of estimating the optimal partition of a non-hyperbolic map, where the linear approximation to the Fokker-Planck operator fails; the requisite calculations are relegated to Appendix B. The results are summarized and the open problems discussed in sect. 9.

2. Noisy trajectories and their densities

The literature on stochastic dynamical systems is vast, starting with the Laplace 1810 memoir [15]. The material reviewed in this section and sect. 3 is standard [1, 3, 16], but needed in order to set the notation for what is new here, the role that Fokker-Planck operators play in defining stochastic neighborhoods of periodic orbits. The key result rederived here is the well known composition law (22) for the covariance matrix Q_a of a linearly evolved Gaussian density, $Q_{a+1} = M_a Q_a M_a^T + \Delta_a$.

2.1. Fokker-Planck operator, continuous time formulation

Consider a d-dimensional stochastic flow

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

where the deterministic velocity field $v(x)$ is called 'drift' in the stochastic literature, and $\hat{\xi}(t)$ is additive noise, uncorrelated in time. A way to make sense of $\hat{\xi}(t)$ is to first construct the corresponding probability distribution for additive noise ξ at a short but finite time δt. In time δt the deterministic trajectory advances by $v(x_n)\,\delta t$. As δt is arbitrary, it is desirable that the diffusing cloud of noisy trajectories is given by a distribution that keeps its form as $\delta t \to 0$. This

holds if the noise is Brownian, i.e., the probability that the trajectory reaches x_{n+1} is given by a normalized Gaussian

$$\mathcal{L}^{\delta t}(x_{n+1}, x_n) \,=\, \frac{1}{N} \exp\left[-\frac{1}{2\,\delta t}(\xi_n^T \frac{1}{\Delta}\xi_n)\right] . \tag{2}$$

Here $\xi_n = \delta x_n - v(x_n)\,\delta t$, the deviation of the noisy trajectory from the deterministic one, can be viewed either in terms of velocities $\{\dot{x}, v(x)\}$ (continuous time formulation), or finite time maps $\{x_n \to x_{n+1}, x_n \to f^{\delta t}(x_n)\}$ (discrete time formulation),

$$\delta x_n = x_{n+1} - x_n \simeq \dot{x}_n\,\delta t\,, \qquad f^{\delta t}(x_n) - x_n \simeq v(x_n)\,\delta t\,, \tag{3}$$

where

$$\{x_0, x_1, \cdots, x_n, \cdots, x_k\} = \{x(0), x(\delta t), \cdots, x(n\delta t), \cdots, x(t)\} \tag{4}$$

is a sequence of $k+1$ points $x_n = x(t_n)$ along the noisy trajectory, separated by time increments $\delta t = t/k$, and the superfix ^{T} indicates a transpose. The probability distribution $\xi(t_n)$ is characterized by zero mean and covariance matrix (diffusion tensor)

$$\langle \xi_j(t_n) \rangle = 0\,, \qquad \left\langle \xi_i(t_m)\,\xi_j^T(t_n) \right\rangle = \Delta_{ij}\,\delta_{nm}\,, \tag{5}$$

where $\langle \cdots \rangle$ stands for ensemble average over many realizations of the noise. For example, in one dimension the white noise $\xi_n = x_{n+1} - x_n$ for a pure diffusion process (no advection, $v(x_n) = 0$) is a normally distributed random variable, with standard normal (Gaussian) probability distribution function,

$$\mathcal{L}^t(x, x_0) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left[-\frac{(x-x_0)^2}{2\sigma^2 t}\right] , \tag{6}$$

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

$$\langle x_{n+1} - x_n \rangle = 0\,, \qquad \langle (x_{m+1} - x_m)(x_{n+1} - x_n) \rangle = 2\,D\,\delta_{mn}\,. \tag{7}$$

$\mathcal{L}^t(x, x_0)$ describes the diffusion at any time, including the integer time increments $\{t_n\} = \{\delta t, 2\delta t, \cdots, n\delta t, \cdots\}$, and thus provides a bridge between the continuous and discrete time formulations of noisy evolution. We have set $\delta t = 1$ in (7) anticipating the discrete time formulation of the next section.

In physical problems the diffusion tensor $\mathbf{\Delta}$ is almost always anisotropic: for example, the original Langevin flow [17] is a continuous time flow in {configuration, velocity} phase space, with white noise probability distribution $\exp(-\mathbf{x}^2/2k_BT)$ modeling random Brownian force kicks applied only to the velocity variables $\mathbf{x}$. In this case one thinks of diffusion coefficient $D = k_BT/2$ as temperature. For sake of simplicity we shall often assume that diffusion in d dimensions is uniform and isotropic, $\mathbf{\Delta}(x) = 2\,D\,\mathbf{1}$. The more general case of a tensor $\mathbf{\Delta}$ which is a

state space position dependent but time independent can be treated along the same lines, as we do in Eq. (22). In this case the stochastic flow (1) is written as [**?**] $dx = v(x)\,dt + \sigma(x)\,d\hat{\xi}(t)$, $\sigma(x)$ is called 'diffusion matrix,' and the noise is referred to as 'multiplicative.'

The distribution (2) describes how an initial density of particles concentrated in a Dirac delta function at x_n spreads in time δt. In the Fokker-Planck description individual noisy trajectories are replaced by the evolution of the density of noisy trajectories. The finite time Fokker-Planck evolution $\rho(x,t) = \mathcal{L}^t \circ \rho(x,0)$ of an initial density $\rho(x_0,0)$ is obtained by a sequence of consecutive short-time steps (2)

$$\mathcal{L}^t(x_k,x_0) = \int [dx]\, \exp\left\{ -\frac{1}{4D\delta t} \sum_{n=1}^{k-1} [x_{n+1} - f^{\delta t}(x_n)]^2 \right\}, \tag{8}$$

where $t = k\,\delta t$, and the Gaussian normalization factor in (2),

$$\begin{aligned} N &= (2\pi\delta t)^{d/2} (\det \Delta)^{1/2} \qquad \text{anisotropic diffusion tensor } \Delta \\ &= (4\pi D\delta t)^{d/2} \qquad\quad \text{isotropic diffusion}\,, \end{aligned} \tag{9}$$

is absorbed into intermediate integrations by defining $[dx] = \prod_{n=1}^{k-1} (dx_n^d/N)$.

The stochastic flow (1) can now be understood as the continuous time, $\delta t \to 0$ limit, with the velocity noise $\hat{\xi}(t)$ a Gaussian random variable of zero mean and covariance matrix

$$\left\langle \hat{\xi}_j(t) \right\rangle = 0\,, \qquad \left\langle \hat{\xi}_i(t)\, \hat{\xi}_j(t') \right\rangle = \Delta_{ij}\, \delta(t-t')\,. \tag{10}$$

It is worth noting that the continuous time flow noise $\hat{\xi}(t)$ in (1) and (10) is dimensionally a velocity $[x]/[t]$, while the discrete time noise ξ_n in (2), (5) is dimensionally a length $[x]$. The continuous time limit of (8), $\delta t = t/k \to 0$, defines formally the Fokker-Planck operator

$$\mathcal{L}^t(x,x_0) = \int [dx]\, \exp\left\{ -\frac{1}{4D} \int_0^t [\dot{x}(\tau) - v(x(\tau))]^2 d\tau \right\} \tag{11}$$

as a stochastic path (or Wiener) integral [18, 19, 3] for a noisy flow, and the associated continuous time Fokker-Planck equation [1, 3] describes the time evolution of a density of noisy trajectories (1),

$$\partial_t \rho(x,t) + \nabla \cdot (v(x)\rho(x,t)) \;=\; D\,\nabla^2 \rho(x,t)\,. \tag{12}$$

The $\delta t \to 0$ limit and the proper definition of $\dot{x}(\tau)$ are delicate issues [20, 21, 16**?**] of no import for the applications of stochasticity studied here.

In probabilist literature [**?**] the differential operator $-\nabla \cdot (v(x)\rho(x,t)) + D\,\nabla^2 \rho(x,t)$ is called 'Fokker-Planck operator;' here we reserve the term exclusively for the finite time, 'Green function' integral operator (11). The exponent

$$-\frac{1}{4D\delta t} \left[x_{n+1} - f^{\delta t}(x_n) \right]^2 \;\simeq\; -\frac{1}{4D} \left[\dot{x}(\tau) - v(x(\tau)) \right]^2 \delta t$$

can be interpreted as a cost function which penalizes deviation of the noisy trajectory δx from its deterministic prediction $v\,\delta t$, or, in the continuous time limit, the deviation of the noisy trajectory tangent $\dot{x}$ from the deterministic velocity v. Its minimization is one of the most important tools of the optimal control theory [**?** **?**], with velocity $\dot{x}(\tau)$ along a trial path varied with aim of minimizing its distance to the target $v(x(\tau))$. This cost function appears to have been first introduced by Wiener as the exact solution for a purely diffusive Wiener-Lévy process in one dimension, see (6). Onsager and Machlup [22, 23] use it in their variational principle to study thermodynamic fluctuations in a neighborhood of single, linearly attractive equilibrium point (i.e., without any dynamics). The dynamical 'action' Lagrangian in the exponent of (11), and the associated symplectic Hamiltonian were first written down in 1970's by Freidlin and Wentzell [23], whose formulation of the 'large deviation principle' was inspired by the Feynman quantum path integral [24]. Feynman, in turn, followed Dirac [25] who was the first to discover that in the short-time limit the quantum propagator (imaginary time, quantum sibling of the Wiener stochastic distribution (6)) is exact. Gaspard [4] thus refers to the 'pseudo-energy of the Onsager-Machlup-Freidlin-Wentzell scheme.' M. Roncadelli [26, 27] refers to the Fokker-Planck exponent in (11) as the 'Wiener-Onsager-Machlup Lagrangian,' constructs weak noise saddle-point expansion and writes transport equations for the higher order coefficients.

To keep things simple we shall use only the discrete time dynamics in what follows, but we have here introduced the continuous time formulation first, in order to emphasize that all our results apply both to the continuous and discrete time flows, with the periodic orbits covariance matrices computed using the continuous time Fokker-Planck operator (8).

2.2. Discrete time Fokker-Planck operator, adjoint operator

The finite time step formulation (8) of the Fokker-Planck operator motivates what follows: set $\delta t = 1$, and consider a noisy *discrete time* dynamical system [28, 29, 30]

$$x_{k+1} = f(x_k) + \xi_k \,. \tag{13}$$

The action of discrete one-time step *Fokker-Planck operator* on the density distribution at time k,

$$\begin{aligned} \rho(y, k+1) &= \mathcal{L} \circ \rho(y,k) \,=\, \int dx\, \mathcal{L}(y,x)\rho(x,k) \\ \mathcal{L}(y,x) &= \frac{1}{N} \exp\left\{ -\frac{1}{2}(y - f(x))^T \frac{1}{\Delta}(y - f(x)) \right\} , \end{aligned} \tag{14}$$

is centered on the deterministic step $f(x)$ and smeared out diffusively by noise. The kth iterate of $\mathcal{L}$ is a d-dimensional path integral over the $k-1$ intermediate noisy trajectory points,

$$\mathcal{L}^k(x_k, x_0) \,=\, \int [dx]\, \exp\left\{ -\frac{1}{2} \sum_n (x_{n+1} - f(x_n))^T \frac{1}{\Delta}(x_{n+1} - f(x_n)) \right\} . \tag{15}$$

In order to estimate the size of a noisy neighborhood of a trajectory point x_a along its *unstable* directions, we need to determine the effect of noise also on the points preceding x_a. This is described by the *adjoint Fokker-Planck operator*

$$\begin{aligned}\tilde{\rho}(y,k-1) &= \mathcal{L}^\dagger \circ \tilde{\rho}(y,k) \\ &= \int [dy]\, \exp\left\{-\frac{1}{2}(y-f(x))^T \frac{1}{\Delta}(y-f(x))\right\} \tilde{\rho}(y,k)\,, \quad (16)\end{aligned}$$

which carries a density concentrated around the previous point x_{n-1} to a density concentrated around x_n. The Fokker-Planck operator is non-selfadjoint. We denote by ρ_α its right eigenvectors, and by $\tilde{\rho}_\alpha$ its left eigenvectors.

3. Linearized flow

The function $x_{a+1} - f(x_a)$ is in general nonlinear, and the path integral (15) can only be evaluated numerically. In the vanishing noise limit the Gaussian kernel sharpens into the Dirac δ-function, and the Fokker-Planck operator reduces to the deterministic Perron-Frobenius operator (A.16). For weak noise the Fokker-Planck operator can be therefore evaluated perturbatively [31, 32, 33, 26] as an asymptotic series in powers of D^k, centered on the deterministic trajectory. The linear term in this series has a particulary simple dynamics given by a covariance matrix addition formula (see (22) and (33) below) that we now derive, the basis of almost all the follows, except for the 'flat-top' of sect. 8.

As in (A.10), we shift locally the coordinates to the deterministic trajectory centered coordinate frames $x = x_a + z_a$, Taylor expand $f(x) = f_a(z_a) = x_{a+1} + M_a z_a + \cdots$, and approximate the noisy map (13) by its linearized action,

$$z_{a+1} = M_a z_a + \xi_a\,, \quad (17)$$

with the deterministic trajectory points at $z_a = z_{a+1} = 0$, and M_a the Jacobian matrix (A.10). The corresponding linearized Fokker-Planck operator (14) is given in the local coordinates

$$\begin{aligned}\rho_{a+1}(z_{a+1}) &= \int dz_a\, \mathcal{L}_a(z_{a+1}, z_a)\, \rho_a(z_a) \\ \rho_a(z_a) &= \rho(x_a + z_a, a) \quad (18)\end{aligned}$$

by the linearization (17) centered on the deterministic trajectory points $\{\ldots, x_{-1}, x_0, x_1, x_2, \ldots,\}$,

$$\mathcal{L}_a(z_{a+1}, z_a) = \frac{1}{N}\, e^{-\frac{1}{2}(z_{a+1}-M_a z_a)^T \frac{1}{\Delta}(z_{a+1}-M_a z_a)}\,. \quad (19)$$

In quantum mechanics the corresponding linearized evolution operator is known as the Van Vleck propagator, the basic block in the semi-classical periodic orbit quantization [**?** 61]. The linearized Fokker-Planck operator (19) is a Gaussian. As a convolution of a Gaussian with a Gaussian is again a Gaussian, we investigate the action of the linearized Fokker-Planck operator on a normalized, cigar-shaped Gaussian density distribution

$$\rho_a(z) = \frac{1}{C_a} e^{-\frac{1}{2} z^T \frac{1}{Q_a} z}\,, \qquad C_a = (2\pi)^{d/2} (\det Q_a)^{1/2}\,, \quad (20)$$

also centered on the deterministic trajectory, but with its own strictly positive covariance matrix Q_a. As explained in Appendix A.2, label 'a' plays a double role, and $\{a, a+1\}$ stands both for the {initial,next} space partition and for the times the trajectory lands in these partitions. The linearized noisy evolution operator (18) maps the ellipsoid $\rho_a(z_a)$ into ellipsoid $\rho_{a+1}(z_{a+1})$ one time step later

$$\begin{aligned}\rho_{a+1}(z_{a+1}) &= \frac{1}{C_a}\int [dz_a]\, e^{-\frac{1}{2}[(z_{a+1}-M_a z_a)^T \frac{1}{\Delta}(z_{a+1}-M_a z_a)+z_a^T \frac{1}{Q_a} z_a]} \\ &= \frac{1}{C_{a+1}}\, e^{-\frac{1}{2} z_{a+1}^T \frac{1}{Q_{a+1}} z_{a+1}}\,. \end{aligned} \tag{21}$$

Completing the square and integrating over z_a we obtain the key formula for all that follows: the noise covariance matrix Δ and the deterministically transported local density covariance matrix $Q \to MQM^T$ add up as Gaussian variances, i.e., sums of squares, and one time step later a Gaussian density distribution $\rho_a(z_a)$ is smeared into a Gaussian ellipsoid whose widths and orientation are given by the eigenvalues and eigenvectors of the covariance matrix

$$Q_{a+1} = M_a Q_a M_a^T + \Delta_a\,. \tag{22}$$

This covariance is an interplay of the Brownian noise and the deterministic nonlinear contraction/amplification; in the Kalman filter literature [**?** 34] it is called 'prediction'. The diffusive dynamics of a nonlinear system is thus fundamentally different from Brownian motion, as the flow induces a history dependent effective noise.

We have attached label 'a' to $\Delta_a = \Delta(x_a)$ in (22) to account for the possibility that the noise is inhomogeneous, state space dependent, but time independent multiplicative noise. As explained in ref. [31], in the field-theoretic interpretation of the path integral (15), every term in the Taylor series (A.12) is a 'vertex,' with the kth derivative corresponding to a power of $D^{k/2}$, or noisy trajectories with k noise kicks. In the leading, 'single kick' order, the variance Q_a is built up from the deterministically propagated $M_a^n Q_{a-n} M_a^{nT}$ initial distribution, and the sum of noise kicks at intervening times, $M_a^k \Delta_{a-k} M_a^{kT}$, also propagated deterministically. If M is contracting, over time the memory of the covariance Q_{a-n} of the starting density is lost, with iteration of (22) leading to a limit distribution:

$$Q_a = \Delta_a + M_{a-1}\Delta_{a-1}M_{a-1}^T + M_{a-2}^2\Delta_{a-2}(M_{a-2}^2)^T + \cdots\,. \tag{23}$$

Note that the evaluation of this noise requires no Fokker-Planck PDE formalism. The width of a Gaussian packet centered on a trajectory is fully specified by a deterministic computation that is already a pre-computed byproduct of the periodic orbit computations; the deterministic orbit and its linear stability.

For $x_{k+1} = x_k + \xi_k$, Browniam dynamics $M = \mathbf{1}$, and we obtain $Q_n = n\,\Delta$, i.e., the variance of a Gaussian packet of $\rho_n(z)$ noisy trajectories grows linearly in time, as expected for Brownian diffusion (6). What happens for nontrivial,

$M \neq \mathbf{1}$, dynamics? This is easiest to illustrate by studying the effect of noise on a deterministically attracting equilibrium, following Langevin, Ornstein and Uhlenbeck.

3.1. Attractive fixed point: Deterministic contraction vs. stochastic diffusion

Consider a noisy map (13) with a deterministic (zero noise strength) fixed point at x_q. In a neighborhood $x = x_q + z$ we approximate the map f by its linearized action (17) with the fixed point at $z = 0$, acting on a Gaussian density distribution (20), also centered on $z = 0$. The distribution is cigar-shaped ellipsoid, with eigenvectors of Q_n giving the orientation of various axes at time n, see figure A.12 (b).

If all eigenvalues (Floquet multipliers) of M are strictly contracting, any initial compact measure (not only initial distributions of Gaussian form) converges under applications of (22) to the unique invariant natural measure $\rho_0(z)$ whose covariance matrix satisfies the fixed point condition

$$\begin{aligned} Q &= MQM^T + \Delta \\ &= \Delta + M\Delta M^T + M^2\Delta(M^T)^2 + M^3\Delta(M^T)^3 + \cdots . \end{aligned} \tag{24}$$

We solve for Q as follows. Assume that $[d\times d]$ Jacobian matrix M has distinct contracting Floquet multipliers in $\{\Lambda_1, \Lambda_2, \cdots, \Lambda_d\}$ and right eigenvectors $\mathbf{e}^{(j)}$

$$M\,\mathbf{e}^{(j)} = \Lambda_j\,\mathbf{e}^{(j)} . \tag{25}$$

Construct from the d column eigenvectors a $[d\times d]$ similarity matrix

$$S = \left(\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \cdots, \mathbf{e}^{(d)}\right)$$

that diagonalizes M, $S^{-1}MS = \Lambda$ and its transpose $S^T M^T (S^{-1})^T = \Lambda$. Define $\hat{Q} = S^{-1}Q(S^{-1})^T$ and $\hat{\Delta} = S^{-1}\Delta(S^{-1})^T$. The fixed point condition (24) now takes form

$$\hat{Q} - \Lambda\hat{Q}\Lambda = \hat{\Delta} . \tag{26}$$

The matrix elements are $\hat{Q}_{ij}(1 - \Lambda_i\Lambda_j) = \hat{\Delta}_{ij}$, so

$$\hat{Q}_{ij} = \frac{\hat{\Delta}_{ij}}{1 - \Lambda_i\Lambda_j} , \tag{27}$$

and the attracting fixed point covariance matrix is given by

$$Q = S\hat{Q}S^T . \tag{28}$$

As (28) is not a similarity transformation, evaluation of the covariance matrix Q requires a numerical diagonalization, which yields the principal axis of the equilibrium Gaussian 'cigar'. The eigenvectors of this symmetric matrix have their own orientations, distinct from the left/right eigenvectors of the non-normal Jacobian matrix M.

The simplest case is illuminating: for a contracting noisy 1-dimensional linear map $z_{n+1} = \Lambda z_n + \xi_n$, $|\Lambda| < 1$, the width of the natural measure (20) concentrated at the deterministic fixed point $z = 0$ is

$$Q = \frac{2D}{1-|\Lambda|^2}\,, \qquad \rho_0(z) = \frac{1}{\sqrt{2\pi\, Q}}\, \exp\left(-\frac{z^2}{2\,Q}\right)\,, \tag{29}$$

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

We note in passing that the effective diffusive width is easily recast from the discrete map formulation back into the infinitesimal time step form of sect. 2.1:

$$M = e^{A\delta t}\,, \qquad M^T \,=\, e^{A^T\delta t}\,, \qquad \Delta \,\to\, \delta t \Delta\,, \tag{30}$$

where $A = \partial v/\partial x$ is the stability matrix. Expanding to linear order yields the differential version of the equilibrium condition (24)

$$0 = AQ + QA^T + \Delta \tag{31}$$

(with the proviso that now Δ is covariance matrix (10) for the velocity fluctuations). The condition (31) is well known [16**?** , 3], and currently widely used in the molecular and gene networks literature [**? ? ? ?**]. In one dimension, $A \to \lambda$ (see (A.5)), $\Delta \to 2D$, and the diffusive width is given by [16]

$$Q = -D/2\lambda\,, \tag{32}$$

a balance between the diffusive spreading D and the deterministic contraction rate $\lambda < 0$. If $\lambda \to 0$, the measure spreads out diffusively.

3.2. Attractive periodic points: For nonlinear flow noise is never isotropic

An attractive feature of periodic orbit theory is that certain properties of periodic orbits, such as their periods and Floquet multipliers, are *intrinsic* to the periodic orbit, independent of where they are measure along the orbit, and invariant under all smooth conjugacies, i.e., all smooth nonlinear coordinate transformations. Noise, however, is specified in a given coordinate system and breaks such invariances (for an exception, a canonically invariant noise, see Kurchan [**?**]). Nevertheless, it still suffices to compute the effect of noise on a single periodic point, with the succeeding points given by the noise composition rule (22) in case of discrete time, or integrating 31 for the continuous time case.

The Gaussian natural measure characterized by cycle-point covariance (33) is intrinsic to the cycle. It can be computed equally well for discrete and periodic points, see (2). Once computed at a cycle point x_a it can be propagated to all other cycle points by iterating (22).

A periodic point of an n-cycle is a fixed point of the nth iterate of the map (13). Hence the formula (23) for accumulated noise, together the fixed point

condition (24) also yields the natural measure covariance matrix at a periodic point x_a on a periodic orbit p,

$$Q_{p,a} = \Delta_{p,a} + M_{p,a}\Delta_{p,a}(M_{p,a})^T + \cdots + M^k_{p,a}\Delta_{p,a}(M^k_{p,a})^T + \cdots , \quad (33)$$

where

$$\begin{aligned}\Delta_{p,a} &= \Delta_a + M_{a-1}\Delta_{a-1}M^T_{a-1} + M^2_{a-2}\Delta_{a-2}(M^2_{a-2})^T \\ &\quad + \cdots + M^{n_p-1}_{a-n_p+1}\Delta_{a-n_p+1}(M^{n_p-1}_{a-n_p+1})^T \end{aligned} \quad (34)$$

is the noise accumulated per a single transversal of the periodic orbit, $M_{p,a}$ is the monodromy matrix (A.11) evaluated on the periodic point x_a, and we have used the periodic orbit condition $x_{a+n_p} = x_a$. Note that the fixed (24) and periodic point (33) conditions can be imposed only if both the dynamical system and the noise are autonomous, i.e., the noise covariance matrix depends only on the position but not on time.

For example, for 1-dimensional dynamics, the accumulated noise per a cycle traversal is

$$2D_{p,a} = 2D(1 + (f'_{a-1})^2 + (f^{2'}_{a-2})^2 + \cdots + (f^{n-1'}_{a-n_p+1})^2) \quad (35)$$

As there is no single coordinate frame in which different $M^k_{a-k}(M^k_{a-k})^T$ can be simultaneously diagonalized, the accumulated noise is *never* isotropic. So the lesson is that regardless of whether the external noise Δ is isotropic or anisotropic, the nonlinear flow always renders the effective noise anisotropic and spatially inhomogeneous.

4. Eigenfunctions of the Ornstein-Uhlenbeck process

So far we have assumed that the deterministic flow is either contracting or expanding, with asymptotic stationary distributions concentrated either on fixed points or periodic points. How are we to deal with hyperbolic flows, where some of the eigen-directions are unstable, with $|\Lambda_j| > 1$? The answer is implicit in the Oseledec [**?**] definition of Lyapunov exponents, and in the rigorous proof of existence of classical spectral (Fredholm) determinants by Rugh [36]: the flow on each hyperbolic orbit can be locally factorized into stable and unstable directions, and for unstable directions one needs to study noise evolution backwards in time, by means of the adjoint operator (16). Here description in terms of periodic orbits is very useful; the neighborhood of a periodic point will be defined as the noise contracting neighborhood forward in time along contracting eigendirections, backward in time along the unstable, expanding eigendirections, and we only need to do a finite time calculation in either direction.

The variance (24) is stationary under the action of $\mathcal{L}$, and the corresponding Gaussian is thus an eigenfunction. Indeed, for the linearized flow the entire eigenspectrum is available analytically, and as Q_a can always be brought to a diagonal, factorized form in its orthogonal frame, it suffices to understand the

simplest case, the Ornstein-Uhlenbeck process in one dimension. This simple example will enable us to show that the noisy measure along unstable directions is described by the eigenfunctions of the adjoint Fokker-Planck operator.

The simplest example of a stochastic flow (1) is the Langevin flow in one dimension,

$$\frac{dx}{dt} = \lambda\, x + \hat{\xi}(t)\,, \tag{36}$$

with 'drift' $v(x)$ linear in x, and the single deterministic equilibrium solution $x = 0$ (L. Arnold [16] refers only to this linear case as the "Langevin equation). The associated Fokker-Planck equation (12) is known as the Ornstein-Uhlenbeck process [37, 12, 13, 38, 3]:

$$\partial_t \rho(x,t) + \partial_x(\lambda\, x\, \rho(x,t)) \;=\; D\, \partial_x^2 \rho(x,t)\,. \tag{37}$$

For negative constant λ the spreading of noisy trajectories by random kicks is balanced by the linear damping term (linear drift) $v(x) = \lambda\, x$ which contracts them toward zero. For this choice of $v(x)$, and this choice only, the Fokker-Planck equation can be rewritten as the Schrödinger equation for the quantum harmonic oscillator, with its well-known Hermite polynomial eigenfunctions [39, 40]. The observation is much older than quantum mechanics: Laplace [15] wrote down in 1810 what is now known as the Fokker-Planck equation and computed the Ornstein-Uhlenbeck process eigenfunctions [41] in terms of Hermite polynomials $H_0(x) = 1$, $H_1(x) = 2\,x$, $H_2(x) = 4\,x^2 - 2$, $\cdots$. $H_n(x)$ is an nth-degree polynomial, orthogonal with respect to the Gaussian kernel

$$\frac{1}{2^n n!\,\sqrt{2\pi}} \int dx\, H_m(x)\, e^{-x^2}\, H_n(x) \;=\; \delta_{mn}\,. \tag{38}$$

The key ideas are perhaps more easily illustrated by the noisy, strictly equivalent discrete-time dynamics than by pondering the meaning of the stochastic differential equation (37). Consider the noisy 1-dimensional linear map (13)

$$z_{n+1} = \Lambda z_n + \xi_n\,, \tag{39}$$

with with the deterministic fixed point at $f(x_a) = x_a = 0$, and additive white noise (7) with variance $2D$. This is the discrete time version of the Langevin (36) and Ornstein-Uhlenbeck process (37). The density $\rho(x)$ of trajectories evolves by the action of the Fokker-Planck operator (14):

$$\mathcal{L} \circ \rho(x) = \int [dy]\, e^{-\frac{(x-\Lambda y)^2}{4D}} \rho(y)\,. \tag{40}$$

There are three cases to consider:

Floquet multiplier $|\Lambda| < 1$ case. In each iteration the map contracts the cloud of noisy trajectories by Floquet multiplier Λ toward the $x = 0$ fixed point, while the noise smears them out with variance $2D$. The Gaussian evolution operator kernel implies that the eigenfunctions are of Gaussian form, with Hermite

polynomial factors, and normalized by (38). The right eigenfunctions $\{\tilde{\rho}_0,\, \tilde{\rho}_1,\, \tilde{\rho}_2,\cdots\}$, eigenvalues $\{1,\, \Lambda,\, \Lambda^2,\cdots\}$ of (40) are [12, 13]

$$\tilde{\rho}_k(x) \;\; = \;\; N^{-1} H_k(\mu x)\, e^{-x^2/2\sigma_0^2}\,, \quad \mu^{-2} = 2\sigma_0^2\,, \;\; \sigma_0^2 = 2D/(1-\Lambda^2)\,, \tag{41}$$

where $H_k(x)$ is the kth Hermite polynomial, and $N = (4\pi D)^{1/2}$ is the Gaussian normalization. Note that the Floquet multipliers Λ^k are independent of the noise strength, so they are the same as for the $D \to 0$ deterministic Perron-Frobenius operator (A.16). The unit-eigenvalue eigenfunction $\rho_0 = N^{-1}\exp(-x^2/2\sigma_0^2)$ is the natural measure [42] for the Fokker-Planck operator, its variance $\sigma_0^2 = 2D/(1-\Lambda^2)$ a balance of the fixed-point contraction Λ and diffusive spread D.

$|\Lambda| = 1$ *case.* This is the marginal, pure diffusion limit, and the behavior is not exponential, but power-law. If the map is nonlinear, one needs to go to the first non-vanishing nonlinear order in Taylor expansion (A.12) to reestablish the control [13].

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

$$\rho_k(x) \;\; = \;\; N^{-1} H_k(\alpha x)\,, \qquad \alpha^{-2} = -2\sigma_0^2\,, \tag{42}$$

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

The eigenfunctions (41) and (42) are respectively the left and the right eigenfunctions of the Fokker-Planck operator with $|\Lambda| > 1$. They are orthonormal:

$$\int dx\, \tilde{\rho}_k(x)\rho_j(x) = \delta_{kj}\,. \tag{43}$$

Analogously, the adjoint of the Fokker-Planck operator for the same map(39) acts on a density by

$$\mathcal{L}^\dagger \circ \tilde{\rho}(x) = \int e^{-\frac{(\Lambda x - y)^2}{4D}} \tilde{\rho}(y)[dy] \tag{44}$$

If $|\Lambda| > 1$, the eigenfunctions are

$$\begin{aligned} \tilde{\rho}_n(x) \;\; &= \;\; N^{-1} H_n(\mu x)\, e^{-x^2/2\sigma_0^2} \\ \mu^{-2} \;\; &= \;\; 2\sigma_0^2\,, \qquad \sigma_0^2 = \frac{2D}{\Lambda^2 - 1} \end{aligned} \tag{45}$$

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

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

$$\rho_n(x) = N^{-1} H_n(\alpha x)\,, \qquad \alpha^{-2} = -2\sigma_0^2 \tag{46}$$

with eigenvalues Λ^n.

These discrete time results can be straightforwardly generalized to continuous time flows, where Ornstein-Uhlenbeck eigenfunctions describe linearized neighborhoods of periodic orbits, as well as to higher dimensions.

4.1. Adjoint action on a Gaussian, 1-dimensional case

Now take a Gaussian density

$$\rho_0(x_a + z_a) = c_a e^{-z_a^2/2\sigma_a^2} \tag{47}$$

The adjoint operator (16)

$$\begin{aligned} \mathcal{L}^\dagger \circ \tilde{\rho}(x) &= \int_{-\infty}^{\infty} [dy]\, c_a e^{-\frac{(f(x)-y)^2}{4D}} e^{-(x-x_a)^2/2\sigma_a^2} \\ &= C_{a-1} e^{-\frac{(f(x)-x_a)^2}{2(\sigma_a^2+2D)}} . \end{aligned} \tag{48}$$

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

$$\begin{aligned} \mathcal{L}^\dagger \circ \tilde{\rho}(x) &= c_{a-1} e^{-\frac{(f(x)-x_a)^2}{2(\sigma_a^2+2D)}} \simeq c_{a-1} e^{-\frac{(f'(f^{-1}(x_a))(x-f^{-1}(x_a)))^2}{2(\sigma_a^2+2D)}} \\ &= c_{a-1} e^{-(f'_{a-1} z_{a-1})^2/2(\sigma_a^2+2D)} \end{aligned} \tag{49}$$

using the notation introduced in section sect. **??**. For an unstable (expanding) map, variances of these densities shrink at every application of $\mathcal{L}^\dagger$. We obtain a recursion relation for the evolution of σ_a^2:

$$(f'_{a-1}\sigma_{a-1})^2 = \sigma_a^2 + 2D \tag{50}$$

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

$$(f^{n'}_{a-n}\sigma_{a-n})^2 = \sigma_a^2 + 2D(1 + (f'_{a-1})^2 + \cdots + (f^{n-1'}_{a-n+1})^2) \tag{51}$$

The initial density (47) is the leading eigenfunction of $\mathcal{L}^\dagger$ in the neighborhood of the unstable cycle, if:

$$\sigma_a^2 = \frac{2D}{1-\Lambda_p^{-2}} \left(\frac{1}{(f'_a)^2} + \cdots + \frac{1}{\Lambda_p^2} \right) , \tag{52}$$

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

5. Approaches to state space partitioning

There is considerable prior literature that addresses various aspects of the 'optimal partition' problem. Before reviewing it, let us state what is novel about the optimal partition hypothesis to be formulated here: Our estimates of limiting resolution are *local*, differing from region to region, while all of the earlier limiting resolution estimates known to us are *global*, based on global averages such as Shennon entropy or quantum-mechanical $\hbar$ 'granularity' of phase space. We know of no published algorithm that sets a limit to the resolution

of a chaotic state space by studying the interplay of the noise with the local stretching/contracting directions of the deterministic dynamics, as we shall do here.

The engineering literature on optimal experimental design [43, 44, 45, 46] employs criteria such as 'D-optimality,' the maximization of the Shannon information content of parameter estimates. Purely statistical in nature, these methods have little bearing on the dynamical approach that we pursue here.

In 1983 Crutchfield and Packard [47] were the first to study the problem of an optimal partition for a chaotic system in the presence of noise, and formulate a state space resolution criterion in terms of a globally averaged "attainable information." The setting is the same that we assume here: the laws governing deterministic dynamics are given, and one studies the effects of noise (be it intrinsic, observational or numerical) on the dynamics. They define the most efficient symbolic encoding of the dynamics as the sequence of symbols that maximizes the metric entropy of the entire system, thus their resolution criterion is based on a global average. Once the maximum for a given number of symbols is found, they refine the partition until the entropy converges to some value. They formulate their resolution criterion in terms of *attainable information*, a limiting value for the probability to produce a certain sequence of symbols from the ensemble of all possible initial conditions. Once such limit is reached, no further refinements are possible.

Most of the dynamical systems literature deals with estimating partitions from observed data [48]. Tang and co-workers [49] assume a noisy chaotic data set, but with the laws of dynamics assumed unknown. Their method is based on maximizing Shannon entropy and at the same time minimizing an error function with respect to the partition chosen. The same idea is used by Lehrman *et al.* [50] to encode chaotic signals in higher dimensions, where they also detect correlations between different signals by computing their conditional entropy. For a review of symbolic analysis of experimental data up to 2001, see Daw, Finney and Tracy [48].

Kennel and Buhl [51, 52, 53] estimate partitions for (high-dimensional) flows from noisy time-series data by minimizing a cost function which maximizes the correlation between distances in the state space and in the symbolic space, and indicates when to stop adjusting their partitions and therefore what the optimal partition is. In ref. [51] their guiding principle for a good partition is that short sequences of consecutive symbols ought to localize the corresponding continuous state space point as well as possible. They embed symbol sequences into the unit square, and minimize the errors in localizing the corresponding state space points under candidate partitions. Holstein and Kantz [54] present an information-theoretic approach to determination of optimal Markov approximations from time series data based on balancing the modeling and the statistical errors in low-dimensional embedding spaces. Boland, Galla and McKane [55] study the effects of intrinsic noise on a class of chemical reaction systems which in the deterministic limit approach a limit cycle in an oscillatory manner.

A related approach to the problem of the optimal resolution is that of the refinement of a transition matrix: given a chaotic, discrete-time dynamical system,

the state space is partitioned, and the probabilities of points mapping between regions are estimated, so as to obtain a transition matrix, whose eigenvalues and eigenfunctions are then used to evaluate averages of observables defined on the chaotic set. The approach was first proposed in 1960 by Ulam [56, 2], for deterministic dynamical systems. He used a uniform-mesh grid as partition, and conjectured that successive refinements of such coarse-grainings would provide a convergent sequence of finite-state Markov approximations to the Perron-Frobenius operator. Nicolis [57], and Rechester and White [58, 59] have proposed dynamics-based refinement strategies for constructing partitions for chaotic maps in one and two dimensions that would improve convergence of Ulam's method.

Bollt *et al.* [60] subject a dynamical system to a small additive noise, define a finite Markov partition, and show that the Perron-Frobenius operator associated to the noisy system is represented by a finite-dimensional stochastic transition matrix. Their focus, however, is on approximating the natural measure of a deterministic dynamical system by the vanishing noise limit of a sequence of invariant measures of the noisy system.

6. Optimal partition hypothesis

6.1. Resolution of a one-dimensional chaotic repeller

We now explain in detail our *'the best possible of all partitions'* hypothesis by formulating it as an algorithm. For every unstable periodic point x_a of a chaotic one-dimensional map, we calculate the corresponding width σ_a Gaussian eigenfunction of the local adjoint Fokker-Planck operator ($\mathcal{L}^\dagger$). Every periodic point is assigned a one-standard deviation neighborhood $[x_a - \sigma_a, x_a + \sigma_a]$. We cover the state space with neighborhoods of orbit points of higher and higher period n_p, and stop refining the local resolution whenever the adjacent neighborhoods, say of x_a and x_b, overlap in such a way that $|x_a - x_b| < \sigma_a + \sigma_b$.

As an illustration of the method, consider the chaotic repeller on the unit interval

$$x_{n+1} = \Lambda_0\, x_n(1 - x_n)(1 - bx_n) + \xi_n\,, \quad \Lambda_0 = 8,\ b = 0.6\,, \tag{53}$$

with noise strength $2D = 0.002$.

The map is plotted in figure 3 (a); together with the local eigenfunctions $\tilde{\rho}_a$ with variances given by (52). Each Gaussian is labeled by the $\{f_0, f_1\}$ branches visitation sequence of the corresponding deterministic periodic point (a symbolic dynamics, however, is not a prerequisite for implementing the method). Figure 3 (b) illustrates the overlapping: $\{\mathcal{M}_{000}, \mathcal{M}_{001}\}$, $\{\mathcal{M}_{0101}, \mathcal{M}_{0100}\}$ overlap and all the neighborhoods of the period $n_p = 4$ cycle points, except for $\mathcal{M}_{0110}$ and $\mathcal{M}_{0111}$. We find that in this case the state space (the unit interval) can be resolved into 7 neighborhoods

$$\{\mathcal{M}_{00}, \mathcal{M}_{011}, \mathcal{M}_{010}, \mathcal{M}_{110}, \mathcal{M}_{111}, \mathcal{M}_{101}, \mathcal{M}_{100}\}\,. \tag{54}$$

It turns out that resolving $\mathcal{M}_{011}$ further into $\mathcal{M}_{0110}$ and $\mathcal{M}_{0111}$ would not affect our estimates, as it would produce the same transition graph. Once

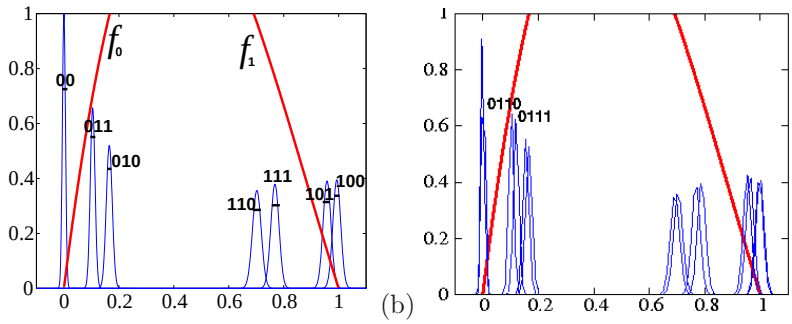

Figure 3: (a) f_0, f_1: branches of the deterministic map (53) for $\Lambda_0 = 8$ and $b = 0.6$. The local eigenfunctions $\tilde{\rho}_{a,0}$ with variances given by (52) provide a state space partitioning by neighborhoods of periodic points of period 3. These are computed for noise variance ($D =$ diffusion constant) $2D = 0.002$. The neighborhoods $\mathcal{M}_{000}$ and $\mathcal{M}_{001}$ already overlap, so $\mathcal{M}_{00}$ cannot be resolved further. (b) The next generation of eigenfunctions shows how the neighborhoods of the optimal partition cannot be resolved further. For periodic points of period 4, only $\mathcal{M}_{011}$ can be resolved further, into $\mathcal{M}_{0110}$ and $\mathcal{M}_{0111}$ (second and third peak from the left), but that would not change the transition graph of figure 5.

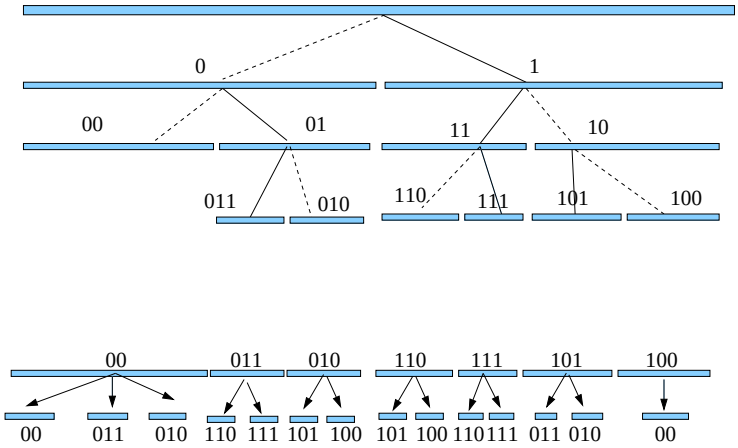

Figure 4: (a) The unit interval partitioned deterministically by a binary tree. Due to the noise, the partitioning stops where the eigenfunctions of figure 3 overlap significantly; (b) once the optimal partition is found, the symbolic dynamics is recoded by relabeling the finite partition intervals.

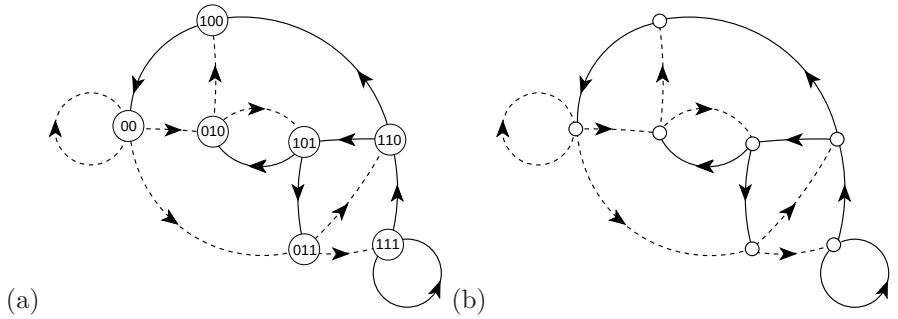

Figure 5: (a) Transition graph (graph whose links correspond to the nonzero elements of a transition matrix T_{ba}) describes which regions b can be reached from the region a in one time step. The 7 nodes correspond to the 7 regions of the optimal partition (54). Dotted links correspond to symbol 0, and the full ones to 1, indicating that the next region is reached by the f_0, respectively f_1 branch of the map plotted in figure 3. (b) The region labels in the nodes can be omitted, with links keeping track of the symbolic dynamics.

the finest possible partition is determined, a finite binary tree like the one in figure 4 is drawn: Evolution in time maps the optimal partition interval $\mathcal{M}_{011} \to \{\mathcal{M}_{110}, \mathcal{M}_{111}\}$, $\mathcal{M}_{00} \to \{\mathcal{M}_{00}, \mathcal{M}_{011}, \mathcal{M}_{010}\}$, etc.. This is summarized in the transition graph (figure 5), which we will use to estimate the escape rate and the Lyapunov exponent of the repeller.

7. Finite Fokker-Planck operator, and stochastic corrections

Next we show that the optimal partition enables us to replace Fokker-Planck PDEs by finite-dimensional matrices. The variance (52) is stationary under the action of $\mathcal{L}_a^{\dagger n_p}$, and the corresponding Gaussian is thus an eigenfunction. Indeed, as we showed in sect. 4, for the linearized flow the entire eigenspectrum is available analytically. For a periodic point $x_a \in p$, the n_pth iterate $\mathcal{L}_a^{n_p}$ of the linearization (19) is the discrete time version of the Ornstein-Uhlenbeck process [37], with left $\tilde{\rho}_0$, $\tilde{\rho}_1$, $\cdots$, respectively right ρ_0, ρ_1, $\cdots$ mutually orthogonal eigenfunctions [3] given by (**??**).

Partition (54) being the finest possible partition, the Fokker-Planck operator now acts as [7×7] matrix with non-zero $a \to b$ entries expanded in the Hermite basis,

$$\begin{aligned}
[\mathbf{L}_{ba}]_{kj} &= \langle \tilde{\rho}_{b,k} | \mathcal{L} | \rho_{a,j} \rangle \\
&= \int \frac{dz_b dz_a \, \beta}{2^{j+1} j! \pi \sqrt{D}} e^{-(\beta z_b)^2 - \frac{(z_b - f_a(z_a))^2}{4D}} \\
&\qquad \times H_k(\beta z_b) H_j(\beta z_a) \,, \qquad (55)
\end{aligned}$$

where $1/\beta = \sqrt{2}\sigma_a$, and z_a is the deviation from the periodic point x_a. It is the number of resolved periodic points that determines the dimensionality of the Fokker-Planck matrix.

Periodic orbit theory [61, 62] expresses the long-time dynamical averages, such as Lyapunov exponents, escape rates, and correlations, in terms of the leading eigenvalues of the Fokker-Planck operator $\mathcal{L}$. In our optimal partition approach, $\mathcal{L}$ is approximated by the finite-dimensional matrix $\mathbf{L}$, and its eigenvalues are determined from the zeros of $\det(1 - z\mathbf{L})$, expanded as a polynomial in z, with coefficients given by traces of powers of $\mathbf{L}$. As the trace of the nth iterate of the Fokker-Planck operator $\mathcal{L}^n$ is concentrated on periodic points $f^n(x_a) = x_a$, we evaluate the contribution of periodic orbit p to $\mathrm{tr}\,\mathbf{L}^{n_p}$ by centering $\mathbf{L}$ on the periodic orbit,

$$t_p = \mathrm{tr}_{\,p}\,\mathcal{L}^{n_p} = \mathrm{tr}\,\mathbf{L}_{ad}\cdots\mathbf{L}_{cb}\mathbf{L}_{ba}\,, \tag{56}$$

where $x_a, x_b, \cdots x_d \in p$ are successive periodic points. To leading order in the noise variance $2D$, t_p takes the deterministic value $t_p = 1/|\Lambda_p - 1|$. The nonlinear diffusive effects in (55) can be accounted for [31] by the weak-noise Taylor series expansion around the periodic point x_a,

$$e^{-\frac{(z_b - f_a(z_a))^2}{4D}} = e^{-\frac{(z_b - f_a' z_a)^2}{4D}} \left(1 - 2\sqrt{D}(f_a'' f_a' z_a^3 + f_a'' z_a^2 z_b) + O(D)\right). \tag{57}$$

Such higher order corrections will be needed in what follows for a sufficiently accurate comparison of different methods.

We illustrate the method by calculating the escape rate $\gamma = -\ln z_0$, where z_0^{-1} is the leading eigenvalue of Fokker-Planck operator $\mathcal{L}$, for the repeller plotted in figure 3. The spectral determinant can be read off the transition graph of figure 5,

$$\begin{aligned}
\det(1 - z\mathbf{L}) = 1 &- (t_0 + t_1)z - (t_{01} - t_0 t_1)\, z^2 \\
&-(t_{001} + t_{011} - t_{01}t_0 - t_{01}t_1)\, z^3 \\
&-(t_{0011} + t_{0111} - t_{001}t_1 - t_{011}t_0 - t_{011}t_1 + t_{01}t_0t_1)\, z^4 \\
&-(t_{00111} - t_{0111}t_0 - t_{0011}t_1 + t_{011}t_0t_1)\, z^5 \\
&-(t_{001011} + t_{001101} - t_{0011}t_{01} - t_{001}t_{011})\, z^6 \\
&-(t_{0010111} + t_{0011101} - t_{001011}t_1 - t_{001101}t_1 \\
&\quad -t_{00111}t_{01} + t_{0011}t_{01}t_1 + t_{001}t_{011}t_1)\, z^7.
\end{aligned} \tag{58}$$

The polynomial coefficients are given by products of non-intersecting loops of the transition graph [61], with the escape rate given by the leading root z_0^{-1} of the polynomial. Twelve periodic orbits $\overline{0}$, $\overline{1}$, $\overline{01}$, $\overline{001}$, $\overline{011}$, $\overline{0011}$, $\overline{0111}$, $\overline{00111}$, $\overline{001101}$, $\overline{001011}$, $\overline{0010111}$, $\overline{0011101}$ up to period 7 (out of the 41 contributing to the noiseless, deterministic cycle expansion up to cycle period 7) suffice to fully determine the spectral determinant of the Fokker-Planck operator. In the evaluation of traces (56) we include stochastic corrections up to order $O(D)$ (an order beyond the term kept in (57)). The escape rate of the repeller of figure 3 so computed is reported in figure 6.

Since our optimal partition algorithm is based on a sharp overlap criterion, small changes in noise strength D can lead to transition graphs of different

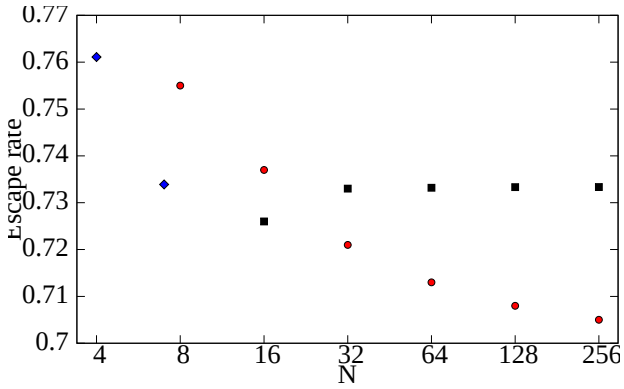

Figure 6: The escape rate γ of the repeller figure 3 plotted as function of number of partition intervals N, estimated using: (♦) under-resolved 4-interval and the 7-interval optimal partition, (●) all periodic orbits of periods up to $n = 8$ in the deterministic, binary symbolic dynamics, with $N_i = 2^n$ periodic-point intervals (the deterministic, noiseless escape rate is $\gamma_{<>} = 0.7011$), and (■) a uniform discretization (59) in $N = 16, \cdots, 256$ intervals. For $N = 512$ discretization yields $\gamma_{\rm num} = 0.73335(4)$.

topologies, and it is not clear how to assess the accuracy of our finite Fokker-Planck matrix approximations. We make three different attempts, and compute the escape rate for: (a) an under-resolved partition, (b) several deterministic, over-resolved partitions, and (c) a brute force numerical discretization of the Fokker-Planck operator.

(a) In the example at hand, the partition in terms of periodic points $\overline{00}$, $\overline{01}$, $\overline{11}$ and $\overline{10}$ is under-resolved; the corresponding escape rate is plotted in figure 6. (b) We calculate the escape rate by over-resolved periodic orbit expansions, in terms of *all* deterministic periodic orbits of the map up to a given period, with t_p evaluated in terms of Fokker-Planck local traces (56), including stochastic corrections up to order $O(D)$. Figure 6 shows how the escape rate varies as we include all periodic orbits up to periods 2 through 8. Successive estimates of the escape rate appear to converge to a value different from the optimal partition estimate. (c) Finally, we discretize the Fokker-Planck operator $\mathcal{L}$ by a piecewise-constant approximation on a uniform mesh on the unit interval [56],

$$[\mathcal{L}]_{ij} \;=\; \frac{1}{|\mathcal{M}_i|}\frac{1}{\sqrt{4\pi D}}\int_{\mathcal{M}_i} dx \int_{f^{-1}(\mathcal{M}_j)} dy\, e^{-\frac{1}{4D}(y-f(x))^2}, \tag{59}$$

where $\mathcal{M}_i$ is the ith interval in equipartition of the unit interval into N equal segments. Empirically, $N = 128$ intervals suffice to compute the leading eigenvalue of the discretized $[128\times128]$ matrix $[\mathcal{L}]_{ij}$ to four significant digits. This escape rate, figure 6, is consistent with the $N = 7$ optimal partition estimate to three significant digits.

We estimate the escape rate of the repeller (53) for a range of values of

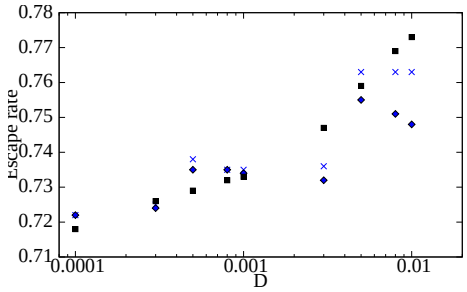

Figure 7: Escape rates of the repeller (53) vs. the noise strength D, using: the optimal partition method with (♦) and without (×) stochastic corrections; (■) a uniform discretization (59) in $N = 128$ intervals.

the noise strength $2D$. The optimal partition method requires a different numbers of neighborhoods every time for different noise strengths. The results are summarized in Table 1, and illustrated by figure 7, with the estimates of the optimal partition method within 2% of those given by the uniform discretization of Fokker-Planck. One can also see from the same table that the escape rates calculated with and without higher order corrections to the matrix elements (55) are consistent within less than 2%, meaning that the stochastic corrections (57) do not make a significant difference, as opposed to the choice of the partition, and need not be taken into account in this example.

Table 1: Escape rates of the repeller (53) from the unit interval, calculated from the determinant of the graph of the optimal partition: ($\gamma_{<>}^{O(D)}$) with stochastic corrections, ($\gamma_{<>}$) without stochastic corrections, and ($\gamma_{\rm num}$) by discretization of $\mathcal{L}$, for different values of D. n_r is the number of regions of the state space resolved by the optimal partition every time.

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

The optimal partition estimate of the Lyapunov exponent is given [61] by $\lambda = \langle \ln |\Lambda| \rangle / \langle n \rangle$, where the cycle expansion average of an observable A

$$\begin{aligned} \langle A \rangle &= A_0 t_0 + A_1 t_1 + (A_{01} t_{01} - (A_0 + A_1)\, t_0 t_1) \\ &\quad + (A_{001} t_{001} - (A_{01} + A_0) t_{01} t_0) + \cdots \\ &\quad + (A_{011} t_{011} - (A_{01} + A_1) t_{01} t_1) + \cdots \end{aligned} \tag{60}$$

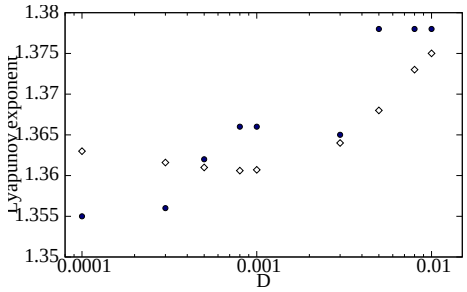

Figure 8: The Lyapunov exponent of the repeller (53) vs. the noise strength D, using: the optimal partition method ($\bullet$) without stochastic corrections, and ($\diamond$) a uniform discretization (59) over $N = 128$ intervals.

is the finite sum over cycles contributing to (58), and $\ln|\Lambda_p| = \sum \ln|f'(x_a)|$, the sum over the points of cycle p, is the cycle Lyapunov exponent. On the other hand, we also use the discretization (59) to crosscheck our estimate: this way the Lyapunov exponent is evaluated as the average

$$\lambda = \int dx\, e^{\gamma}\rho(x) \ln|f'(x)| \tag{61}$$

where $\rho(x)$ is the leading eigenfunction of (59), γ is the escape rate, and $e^{\gamma}\rho$ is the normalized repeller measure, $\int dx\, e^{\gamma}\rho(x) = 1$. Figure 8 shows close agreement ($< 1\%$) between the Lyapunov exponent estimated using the average (60), where $t_p = 1/|\Lambda_p - 1|$ (no higher-order stochastic corrections), and the same quantity evaluated with (61), by the discretization method (59).

8. When the Gaussian approximation fails

The state space of a generic deterministic flow is an infinitely interwoven hierarchy of attracting, hyperbolic and marginal regions, with highly singular invariant measures. Noise has two types of effects. First, it feeds trajectories into state space regions that are determinsitically either disconnected or transient ("noise induced escape," "noise induced chaos") and second, it smooths out the natural measure. Here, we are mostly concerned with the latter. Intuitively, the noisy dynamics erases any structures finer than the optimal partition, thus -in principle- curing both the affliction of long-period attractors/elliptic islands with very small immediate basins of attraction/ellipticity, and the slow, power-law correlation decays induced by marginally stable regions of state space. So how does noise regularize nonhyperbolic dynamics?

As a simple example, consider the skew Ulam map [63], i.e., the cubic map (53) with the parameter $\Lambda_0 = 1/f(x_c)$, see figure 9. The critical point x_c is the maximum of f on the unit interval, with vanishing derivative $f'(x_c) = 0$. As this map sends the unit interval into itself, there is no escape, but due to the quadratic maximum the (deterministic) natural measure exhibits a spike

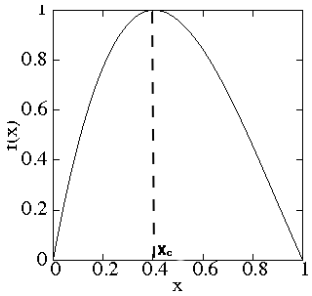

Figure 9: The skew Ulam map (53), with with the parameter $\Lambda_0 = 1/f(x_c)$ chosen so that the critical point maps into the left fixed point $(x_0, y_0) = (0, 0)$.

$(x_b - x)^{-1/2}$ near the critical value $f(x_c) = x_b$ (see, for example, ref. [64] for a discussion). As explained in ref. [63], a close passage to the critical point effectively replaces the accumulated Floquet multiplier by its square root. For example, for the skew Ulam map of figure 9 n-cycles whose itineraries are of form $\overline{0^{n-1}1}$ spend long time in the neighborhood of $x_0 = 0$, and than pass close to x_c. In the neighborhood of $x_0 = 0$ the Floquet multiplier gains a factor $\sim \Lambda_0 = f'(x_0)$ for each of the first $n-1$ iterations, and then experiences a strong, square root contraction during the close passage to the critical point x_c, resulting in the Floquet multiplier $\Lambda_{0\cdots 01} \propto \Lambda_0^{n/2}$, and a Lyapunov exponent that converges to $\lambda_0/2$, rather than λ_0 that would be expected in a hyperbolic flow for a close passage to a fixed point x_0. The same strong contraction is experienced by the noise accumulated along the trajectory prior to the passage by the critical point, rendering, for example, the period-doubling sequences more robust to noise than one would naively expect [65, 18, 29].

For the corresponding noisy map (13) the critical point is extended into 'flat top' region where $|f'(x)| \ll 1$, and the linearized, Gaussian approximation (19) to the Fokker-Planck operator does not hold. Thus, we should first modify our choice of densities and neighborhoods, as the whole construction leading to the optimal partition algorithm was based on the Gaussian approximation.

The adjoint Fokker-Planck operator $\mathcal{L}^\dagger$ acts on a Gaussian density centered at x_a, as in (20):

$$\begin{aligned} \mathcal{L}^\dagger \circ \rho_a(x) &= \int_{-\infty}^{\infty} \frac{1}{C_a} e^{-\frac{(f(x)-y)^2}{4D}} e^{-(y-x_a)^2/2\sigma_a^2} [dy] \\ &= \frac{1}{C_{a-1}} e^{-\frac{(f(x)-x_a)^2}{\sigma_a^2+2D}} . \end{aligned} \tag{62}$$

Suppose the point $x_{a-1} = f^{-1}(x_a)$ around which we want to approximate the new density, is very close to the critical point, so that we can write

$$\rho_{a-1}(x) = \frac{1}{C_{a-1}} e^{-\frac{(f(x)-x_a)^2}{2(\sigma_a^2+2D)}} = \frac{1}{C_{a-1}} e^{-{f''_{a-1}}^2 z_{a-1}^4/8(\sigma_a^2+2D)} , \tag{63}$$

with $f_a^{''}$ defined in (A.13). During a close passage to the critical point, the variance does not transform linearly, but as a square root:

$$\sigma_{a-1}^2 = \frac{\int z^2 e^{-{f_{a-1}^{''}}^2 z^4/8(\sigma_a^2+2D)} dz}{\int e^{-{f_{a-1}^{''}}^2 z^4/8(\sigma_a^2+2D)} dz} = \frac{\Gamma(3/4)}{\Gamma(1/4)} \left(\frac{8(\sigma_a^2+2D)}{{f_{a-1}^{''}}^2} \right)^{1/2} . \tag{64}$$

We show in Appendix B that in the next iteration the variance of the density $\rho_{a-1}(z_{a-1})$ transforms again like the variance of a Gaussian, up to order $O(D)$ in the noise strength. By the same procedure, one can again assume the next preimage of the map x_{a-3} is such that the linear approximation is valid, and transform the density $\rho_{a-2}(z_{a-2})$ (B.3) up to $O(D)$ and obtain the same result for the variance, that is

$$\sigma_{a-3}^2 = \frac{\sigma_{a-1}^2 + 2D(1+{f_{a-2}^{'}}^2)}{{f_{a-2}^{'}}^2 {f_{a-3}^{'}}^2} \tag{65}$$

which is nothing but the expression (51) for the evolution of the variances of Gaussian densities. In other words, the evolution of the variances goes back to be linear, to $O(D)$, although the densities transformed from the 'quartic Gaussian' (63) are no longer Gaussians.

The question is now how to modify the definition of neighborhoods given in sect. 6.1, in order to fit the new approximation. Looking for eigenfunctions of $\mathcal{L}^\dagger$ seems to be a rather difficult task to fulfill, given the functional forms (63) and (B.3) involved. Since we only care about the variances, we define instead the following map

$$\sigma_{a-1}^2 = \begin{cases} C \left(\frac{\sigma_a^2+2D}{{f_{a-1}^{''}}^2} \right)^{1/2} & |{f_{a-1}^{'}}^2 < 1| \\ \frac{\sigma_a^2+2D}{{f_{a-1}^{'}}^2} & \text{otherwise} \end{cases} , \tag{66}$$

$C = 2\sqrt{2}\Gamma(3/4)/\Gamma(1/4)$, for the evolution of the densities, and take its periodic points as our new neighborhoods. In practice, one can compute these numerically, but we will not need orbits longer than length $n_p = 4$ in our tests of the partition, therefore we can safely assume only one periodic point of $f(x)$ to be close to the flat top, and obtain analytic expressions for the periodic points of (66):

$$\tilde{\sigma}_a^2 \simeq C \left(\frac{2D\left(1 + {f_{a-1}^{'}}^2 + ... + (f_{a-n+1}^{n-1'})^2\right)}{\tilde{\Lambda}_p} \right)^{2^{1/2}} \tag{67}$$

with $\tilde{\Lambda}_p = f_{a-n+1}^{n-1'} {f_{a-1}^{''}}^2$, is valid when the cycle starts and ends at a point x_a close to the flat top. Otherwise, take the periodic point x_{a-k}, that is the $k-$th preimage of the point x_a. The corresponding periodic point variance has the form

$$\tilde{\sigma}_{a-k}^2 \simeq \frac{1}{(f_{a-1}^{k'})^2} \left(2D(1 + {f_{a-1}^{'}}^2 + ... + (f_{a-1}^{k-1'})^2) + \tilde{\sigma}_a^2 \right) \tag{68}$$

both expressions (67) and (68) are approximate, as we further assumed $2D\tilde{\Lambda}_p^2 \gg 1$, which is reasonable when $D \in [10^{-4}, 10^{-2}]$, our range of investigation for the noise strength. As before, a neighborhood of width $[x_a - \tilde{\sigma}_a, x_a + \tilde{\sigma}_a]$ is assigned to each periodic point x_a, and an optimal partition follows. However, due to the geometry of the map, such partitions as

$$\{\mathcal{M}_{000}, [\mathcal{M}_{001}, \mathcal{M}_{011}], \mathcal{M}_{010}, \mathcal{M}_{110}, \mathcal{M}_{111}, \mathcal{M}_{10}\} \tag{69}$$

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

Let us now test the method by estimating once again the escape rate of the noisy map in figure 9. We note that the matrix elements

$$\begin{aligned} [\mathbf{L}_{ba}]_{kj} &= \langle \tilde{\rho}_{b,k} | \mathcal{L} | \rho_{a,j} \rangle \\ &= \int \frac{dz_b dz_a \, \beta}{2^{j+1} j! \pi \sqrt{D}} e^{-(\beta z_b)^2 - \frac{(z_b - f'_a z_a)^2}{4D}} \\ &\qquad \times H_k(\beta z_b) H_j(\beta z_a) \,, \end{aligned} \tag{70}$$

should be redefined in the neighborhood of the critical point of the map, where the Gaussian approximation to $\mathcal{L}$ fails. We follow the approximation made in (63):

$$[\mathbf{L}_{ba}]_{kj} = \int \frac{dz_b dz_a \, \beta}{2^{j+1} j! \pi \sqrt{D}} e^{-(\beta z_b)^2 - \frac{(z_b - f''_a \sqrt{4D} z_a^2/2)^2}{4D}} H_k(\beta z_b) H_j(\beta z_a) \,, \tag{71}$$

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

$$[\mathbf{L}_{ba}]_{kj} = \int \frac{dz_b dz_a \, \beta}{2^{j+1} j! \pi \sqrt{D}} e^{-(\beta z_b)^2 - \frac{(z_b - f'_a z_a - f''_a \sqrt{4D} z_a^2/2)^2}{4D}} H_k(\beta z_b) H_j(\beta z_a) \,, \tag{72}$$

We find in our model that the periodic orbits we use in our expansion have x_a's within the flat top, such that $f'_a \sim 10^{-1}$ and $f''_a \sim 10$, and therefore (71) better be replaced with (72) when $D \sim 10^{-4}$. In order to know whether a cycle point is close enough to the flat top for the Gaussian approximation to fail, we recall that the matrix element (55) is the zeroth-order term of a series in D, whose convergence can be probed by evaluating the higher order corrections (57): when the $O(\sqrt{D})$ and $O(D)$ corrections are of an order of magnitude comparable or bigger than the one of (55), we conclude that the Gaussian approximation fails and we use (71) or (72) instead. Everywhere else we use our usual matrix elements (55), *without* the higher-order corrections, as they are significantly larger than in the case of the repeller, and they are not accounted for by the optimal partition method, which is entirely based on a zeroth-order Gaussian approximation of the evolution operator (see sect. **??**). Like before, we tweak

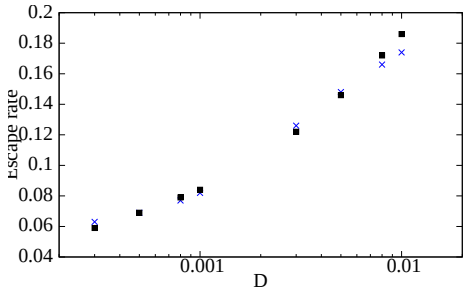

Figure 10: Escape rate γ of the 'skew Ulam' map vs. noise strength D, using: ($\times$) the optimal partition method; ($\blacksquare$) a uniform discretization (59) in $N = 128$ intervals.

the noise strength D within the range $[10^{-4}, 10^{-2}]$ and compare the escape rate evaluated with the optimal partition method and with the uniform discretization (59). The results are illustrated in figure 10: the uniform discretization method and the method of the optimal partition are consistent within a $5-6\%$ margin. The results of figure 10 are also reported in table 2, together with the number of intervals given by the optimal partition for each noise strength D.

Table 2: Escape rates of the 'Ulam skew' map from the unit interval, calculated from the determinant of the graph of the optimal partition: $\gamma_{<>}$ is obtained using (55) and (72) for the matrix elements, with n_r indicating the number of regions of the corresponding optimal partition; $\gamma_{\rm num}$ is the escape rate obtained by uniform discretization of $\mathcal{L}$ ($N = 128$ intervals), for different values of D.

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

9. Summary and conclusions

Computation of unstable periodic orbits in high-dimensional state spaces, such as Navier-Stokes, is at the border of what is currently feasible numerically [9, 10], and criteria to identify finite sets of the most important solutions are very much needed. Where are we to stop calculating orbits of a given hyperbolic flow? Intuitively, as we look at longer and longer periodic orbits, their neighborhoods shrink exponentially with time, while the variance of the noise-induced orbit smearing remains bounded; there has to be a *turnover time*, a time at which the noise-induced width overwhelms the exponentially shrinking

deterministic dynamics, so that no better resolution is possible. Given a specified (possibly state space dependent) noise, we need to find, periodic orbit by periodic orbit, whether a further sub-partitioning is possible.

We have described here the *optimal partition hypothesis*, a new method for partitioning the state space of a chaotic repeller in presence of weak Gaussian noise, and tested the method in a 1-dimensional setting against direct numerical Fokker-Planck operator calculation.

The key idea is that the width of the linearized adjoint Fokker-Planck operator $\mathcal{L}^\dagger$ eigenfunction computed on an unstable periodic point x_a provides the scale beyond which no further local refinement of state space is feasible. This computation enables us to systematically determine the optimal partition, the finest state space resolution attainable for a given chaotic dynamical system and a given noise. Once the optimal partition is determined, we use the associated transition graph to describe the stochastic dynamics by a *finite dimensional* Fokker-Planck matrix. An expansion of the Fokker-Planck operator about periodic points was already introduced in refs. [31, 32, 33], with the stochastic trace formulas and determinants [31, 4] expressed as finite sums, truncated at orbit periods corresponding to the local turnover times. The novel aspect of work presented here is its representation in terms of the Hermite basis (sect. 4), eigenfunctions of the linearized Fokker-Planck operator (19), and the finite dimensional matrix representation of the Fokker-Planck operator.

We have tested our optimal partition hypothesis by applying it to evaluation of the escape rates and the Lyapunov exponents of a $1d$ repeller in presence of additive noise. Numerical tests indicate that the 'optimal partition' method can be as accurate as the much finer grained discretizations of the Fokker-Planck operator.

The success of the optimal partition hypothesis in a 1-dimensional setting is encouraging. However, higher-dimensional hyperbolic maps and flows, for which an effective optimal partition algorithm would be very useful, present new, as yet unexplored challenges of disentangling the subtle interactions between expanding, marginal and contracting directions. A limiting factor to applications of the periodic orbit theory to high-dimensional problems ranging from fluid flows to chemical reactions is the lack of a good understanding of periodic orbits in more than three dimensions, of their stability properties, their organization and their impact on the dynamics. Use of noise as a smoothing device that eliminates singularities and pathologies from clusterings of orbits is promising.

The work of Abarbanel *et al.* [66, 67, 34, 68] suggest one type of important application beyond the low-dimensional Fokker-Planck calculations undertaken here. In data assimilation in weather prediction the convolution of noise variance and trajectory variance (22) is a step in the Kalman filter procedure. One could combine the state space charts of turbulent flows of ref. [69] (computed in the full 3-dimensional Navier-Stokes) with partial information obtained in experiments (typically a full 3-dimensional velocity field, fully resolved in time, but measured only on a 2-dimensional disk section across the pipe). The challenge is to match this measurement of the turbulent flow with a state space point in our 10^5-dimensional ODE representation, and then track the experi-

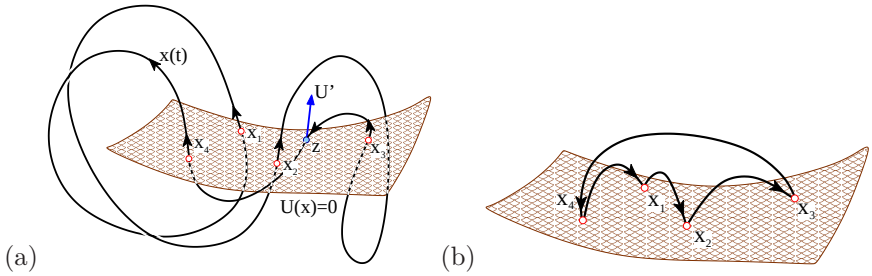

Figure A.11: (a) A Poincaré hypersurface $\mathcal{P}$, defined by a condition $U(x) = 0$, is intersected by the $x(t)$ orbit at times t_1, t_2, t_3, t_4, and closes a cycle (x_1, x_2, x_3, x_4), $x_k = x(t_k) \in \mathcal{P}$, of topological length 4 with respect to this section. The crossing z does not count, as it in the wrong direction. (b) The same orbit reduced to a Poincaré return map that maps points in the Poincaré section $\mathcal{P}$ as $x_{n+1} = f(x_n)$. In this example the orbit of x_1 is periodic and consists of the four periodic points (x_1, x_2, x_3, x_4).

mental observation to improve our theoretical prediction for the trajectory in the time ahead. That would be the absolutely best 'weather prediction' attainable for a turbulent pipe flow, limited by a combination of Lyapunov time and observational noise. In our parlance, the 'optimal partition of state space.'

Acknowledgments.

We are grateful to W.H. Mather and S.A. Solla for critical reading of early drafts of this manuscript, and to A. Grigo, R. Metzler and G. Vattay for many stimulating discussions. P.C. thanks Glen P. Robinson, Jr. for support. D.L. was supported by NSF grant DMS-0807574 and G.P. Robinson, Jr..

Appendix A. Periodic orbit theory, deterministic dynamics

We offer here a brief review of deterministic dynamics and periodic orbit theory. All of this is standard, but needed to set the notation used above The reader might want to consult ref. [61] for further details.

Appendix A.1. Deterministic dynamics - a notational prelude

Though the main applications we have in mind are to continuous flows, for purposes at hand it will suffice to consider discrete time dynamics obtained, for example, by reducing a continuous flow to mappings between successive Poincaré sections, as in figure A.11. Consider dynamics induced by iterations of a d-dimensional map $f : \mathcal{M} \to \mathcal{M}$, where $\mathcal{M} \subset \mathbb{R}^d$ is the state space (or 'phase space') of the system under consideration. The discrete 'time' is then an integer, the number of applications of a map. We denote the kth iterate of map f by composition

$$f^k(x) = f\left(f^{k-1}(x)\right) , \qquad f^0(x) = x\,. \tag{A.1}$$

The *trajectory* of $x = x_0$ is the finite set of points $x_j = f^j(x)$,

$$\{x_0, x_1, x_2, \ldots, x_k\} = \left\{x, f(x), f^2(x), \ldots, f^k(x)\right\} , \tag{A.2}$$

traversed in time k, and the *orbit* of x is the subset $\mathcal{M}_x$ of all points of $\mathcal{M}$ that can be reached by iterations of f. Here x_k is a point in the d-dimensional state space $\mathcal{M}$, and the subscript k indicates time. While a trajectory depends on the initial point x, an orbit is a set invariant under dynamics. The transformation of an infinitesimal neighborhood of an orbit point x under the iteration of a map follows from Taylor expanding the iterated mapping at finite time k. The linearized neighborhood is transported by the $[d\times d]$ Jacobian matrix

$$M^k_{ij}(x_0) = \left.\frac{\partial f^k_i(x)}{\partial x_j}\right|_{x=x_0} . \tag{A.3}$$

($J(x)$ for Jacobian, or derivative notation $M(x) \to Df(x)$ is frequently employed in the literature.) The formula for the linearization of kth iterate

$$M^k(x_0) = M(x_{k-1}) \cdots M(x_1)M(x_0)\,, \quad M_{ij} = \partial f_i/\partial x_j\,, \tag{A.4}$$

in terms of unit time steps M follows from the chain rule for functional composition,

$$\begin{aligned} \frac{\partial}{\partial x_i} f_j(f(x_0)) &= \sum_{k=1}^{d} \left.\frac{\partial}{\partial x_k} f_j(y)\right|_{y=f(x_0)} \frac{\partial}{\partial x_i} f_k(x_0) \\ M^2(x_0) &= M(x_1)M(x_0)\,. \end{aligned}$$

We denote by Λ_ℓ the ℓth *eigenvalue* or *multiplier* of the Jacobian matrix $M^k(x_0)$, and by $\lambda^{(\ell)}$ the ℓth *Floquet* or *characteristic* exponent, with real part $\mu^{(\ell)}$ and phase $\omega^{(\ell)}$:

$$\Lambda_\ell = e^{k\lambda^{(\ell)}} = e^{k(\mu^{(\ell)}+i\omega^{(\ell)})}\,. \tag{A.5}$$

In general the eigenvalues (Floquet multipliers) of $M^k(x_0)$ depend on the initial point x_0 and the elapsed time k.

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

$$f^n(x_k) = f(f(\ldots f(x_k)\ldots)) = x_k\,, \quad k = 0, 1, 2, \ldots, n-1\,. \tag{A.6}$$

For example, the orbit of x_1 in figure A.11 is the 4-cycle (x_1, x_2, x_3, x_4). The time-dependent n-periodic vector fields, such as the flow linearized around a periodic orbit, are described by Floquet theory. Hence we shall refer to the Jacobian matrix $M_p(x) = M^n(x)$ evaluated on a periodic orbit p as the *monodromy* or *Floquet* matrix, and to its eigenvalues $\{\Lambda_{p,1}, \Lambda_{p,2}, \ldots, \Lambda_{p,d}\}$ as Floquet multipliers. They are flow-invariant, independent of the choice of coordinates and the initial point in the cycle p, so we label them by their p label. We number the eigenvalues in order of decreasing magnitude $|\Lambda_1| \geq |\Lambda_2| \geq \ldots \geq |\Lambda_d|$, sort them into sets $\{e, m, c\}$

$$\begin{aligned} \text{expanding:} \qquad \{\Lambda\}_e &= \{\Lambda_{p,j} : |\Lambda_{p,j}| > 1\} \\ \text{marginal:} \qquad \{\Lambda\}_m &= \{\Lambda_{p,j} : |\Lambda_{p,j}| = 1\} \\ \text{contracting:} \qquad \{\Lambda\}_c &= \{\Lambda_{p,j} : |\Lambda_{p,j}| < 1\}\,, \end{aligned} \tag{A.7}$$

and denote by Λ_p (no jth eigenvalue index) the product of *expanding* Floquet multipliers

$$\Lambda_p = \prod_e \Lambda_{p,e} \,. \tag{A.8}$$

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

The stretching/contraction rates per unit time are given by the real parts of Floquet exponents

$$\mu_p^{(i)} = \frac{1}{n_p} \ln |\Lambda_{p,i}| \ . \tag{A.9}$$

They can be interpreted as Lyapunov exponents evaluated on the prime cycle p.

A periodic orbit p is *stable* if real parts of all of its Floquet exponents are strictly negative, $\mu_p^{(i)} < 0$. If all Floquet exponents are strictly positive, $\mu^{(i)} \geq \mu_{min} > 0$, the periodic orbit is *repelling*, and unstable to any perturbation. If some are strictly positive, and rest strictly negative, the periodic orbit is said to be *hyperbolic* or a *saddle*, and unstable to perturbations outside its stable manifold. Repelling and hyperbolic periodic orbits are unstable to generic perturbations, and thus said to be *unstable*. If all $\mu^{(i)} = 0$, the orbit is said to be *elliptic*, and if $\mu^{(i)} = 0$ for a subset of exponents, the orbit is said to be *partially hyperbolic*. If *all* Floquet exponents (other than the vanishing longitudinal exponent) of *all* periodic orbits of a flow are strictly bounded away from zero, the flow is said to be *hyperbolic*. Otherwise the flow is said to be *nonhyperbolic*.

Appendix A.2. Deterministic state space partitions

We streamline the notation by introducing local coordinate systems z_a centered on the trajectory points x_a, together with a trajectory-centered notation for the map (A.1), its derivative, and, by the chain rule, the derivative (A.4) of the kth iterate f^k evaluated at the point x_a,

$$\begin{aligned} x &= x_a + z_a \,, \quad f_a(z_a) = f(x_a + z_a) \\ z_{a+1} &= M_{a+1} z_{a+1} + \cdots \\ M_a &= f'(x_a) \,, \ \ M_a^k = M_{a+k-1} \cdots M_{a+1} M_a \,, \ \ k \geq 2 \,. \end{aligned} \tag{A.10}$$

The monodromy (or Floquet) matrix,

$$M_{p,a} = M^{n_p}(x_a) \,, \tag{A.11}$$

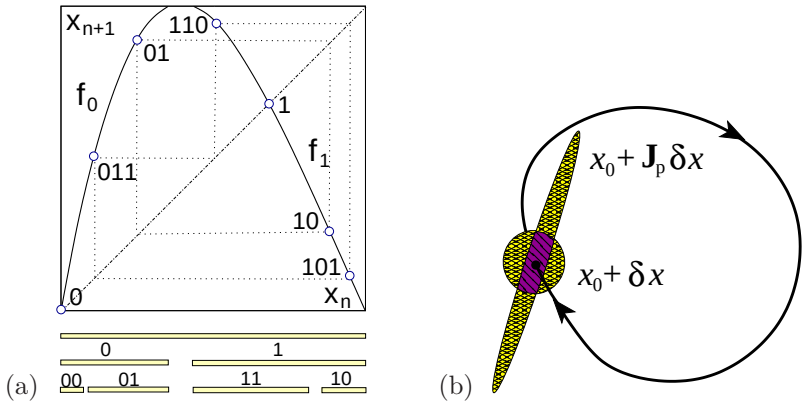

Figure A.12: (a) Periodic points of the map $f(x) = 6\,x\,(1-x)(1-0.6\,x)$ labeled according to the partition $\mathcal{M} = \{\mathcal{M}_0, \mathcal{M}_1\}$. (b) For a prime cycle p, Floquet matrix M_p returns an infinitesimal spherical neighborhood of $x_0 \in \mathcal{M}_p$ stretched into a cigar-shaped ellipsoid, with principal axes given by the eigendirections $\mathbf{e}^{(i)}$ of M_p, the monodromy matrix (A.11).

evaluated on the periodic point x_a is position dependent, but its eigenvalues, the Floquet multipliers $\{\Lambda_{p,1}, \Lambda_{p,2}, \ldots, \Lambda_{p,d}\}$ are invariant, intrinsic to the periodic orbit.

For example, if f is a 1-dimensional map, its Taylor expansion about $x_a = x - z_a$ is

$$f(x) = x_{a+1} + f_a' z_a + \frac{1}{2} f_a'' z_a^2 + \cdots , \tag{A.12}$$

where

$$\begin{aligned} f_a' &= f'(x_a)\,, \qquad f_a'' = f''(x_a) \\ f_a^{k\prime} &= f_{a+k-1}' \cdots f_{a+1}' f_a'\,, \quad k \geq 2\,. \end{aligned} \tag{A.13}$$

A cycle point for which $f_a' = 0$ is called a critical point.

We label trajectory points by either x_n, $n = 1, 2, \cdots$, in order to emphasize as time evolution of x_0, or by x_a, to emphasize that the trajectory point lies in the state space region labeled 'a.' Then the label $a+1$ is a shorthand for the next state space region b on the orbit of x_a, $x_b = x_{a+1} = f(x_a)$. For example, in figure A.12 (a) a periodic point is labeled $a = 011$ by the itinerary with which it visits the regions of the partitioned state space $\mathcal{M} = \{\mathcal{M}_0, \mathcal{M}_1\}$, and as $x_{110} = f(x_{011})$, the next point label is $b = 110$. The whole periodic orbit $\overline{011} = (x_{011}, x_{110}, x_{101})$ is traversed in 3 iterations.

Appendix A.3. Periodic orbit theory

Since its initial formulations by Ruelle [70] and Gutzwiller [71], the periodic orbit theory has developed into a powerful theoretical and computational tool for prediction of quantities measurable in chaotic dynamics. Schematically (a

detailed exposition can be found in refs. [61, 62]), one of the tasks of a theory of chaotic systems is to predict the long-time average of an experimentally measurable quantity $a(x)$ from the spatial and time averages

$$\langle a \rangle = \lim_{n\to\infty} \frac{1}{n} \langle A^n \rangle \ , \qquad A^n(x_0) = \sum_{k=0}^{n-1} a(x_k) \, . \tag{A.14}$$

What makes evaluation of such averages difficult is chaotic dynamics' sensitivity to initial conditions; exponentially unstable trajectories can be tracked accurately only for finite times. The densities of trajectories $\rho(x,t)$, however, can be well behaved for $t \to \infty$. Hence the theory is recast in the language of *linear* evolution operators (Liouville, Perron-Frobenius, Ruelle-Araki, $\cdots$)

$$\rho(y,t) = \mathcal{L}^t_{det} \circ \rho(y) = \int_{\mathcal{M}} dx \, \delta\big(y - f^t(x)\big) \, \rho(x,0) \, , \tag{A.15}$$

This evolution operator assembles the density $\rho(y,t)$ at time t by going back in time to the density $\rho(x,0)$ at time $t = 0$. Here we shall refer to the integral operator with singular kernel (A.15)

$$\mathcal{L}^t_{det}(x,y) = \delta\big(x - f^t(y)\big) \tag{A.16}$$

as the *Perron-Frobenius operator*.

If the map is linear, $f(x) = \Lambda x$, the Perron-Frobenius operator is

$$\mathcal{L}_{det} \circ \rho(x) = \int dy \, \delta(x - \Lambda y) \rho(y) = \frac{1}{|\Lambda|} \rho(\frac{x}{|\Lambda|}) \, . \tag{A.17}$$

In the $|\Lambda| > 1$ expanding case the right, expanding deterministic eigenfunctions [13, 72] are monomials

$$\rho_k(x) \to x^k/k! \, , \qquad k = 0, 1, 2, \cdots \, , \tag{A.18}$$

with eigenvalues $1/|\Lambda|\Lambda^n$, while the left, contracting eigenfunctions are distributions

$$\rho_k(x) \to (-1)^k \delta^{(k)}(x) \, . \tag{A.19}$$

In discretizations $\mathcal{L}^t_{det}(y,x)$ is represented by a matrix with y, x replaced by discrete indices, and integrals over x replaced by index summation in matrix multiplication. Indeed, for piece-wise linear mappings Perron-Frobenius operator can be a finite-dimensional matrix. For example, consider the expanding 1-dimensional 2–branch map $f(x)$ with slopes $\Lambda_0 > 1$ and $\Lambda_1 = -\Lambda_0/(\Lambda_0 - 1) < -1$:

$$f(x) = \begin{cases} f_0(x) = \Lambda_0 x \, , & x \in \mathcal{M}_0 = [0, 1/\Lambda_0) \\ f_1(x) = \Lambda_1(1-x) \, , & x \in \mathcal{M}_1 = (1/\Lambda_0, 1] \, . \end{cases} \tag{A.20}$$

As in figure A.12 (a), the state space (i.e., the unit interval) is partitioned into two regions $\mathcal{M} = \{\mathcal{M}_0, \mathcal{M}_1\}$. If density $\rho(x)$ is a piecewise constant on each partition

$$\rho(x) = \begin{cases} \rho_0 & \text{if } x \in \mathcal{M}_0 \\ \rho_1 & \text{if } x \in \mathcal{M}_1 \end{cases} \ , \tag{A.21}$$

the Perron-Frobenius operator acts as a $[2\times 2]$ Markov matrix $\mathbf{L}$ with matrix elements

$$\begin{pmatrix}\rho_0\\ \rho_1\end{pmatrix} \to \mathbf{L}\,\rho = \begin{pmatrix} \frac{1}{|\Lambda_0|} & \frac{1}{|\Lambda_1|} \\ \frac{1}{|\Lambda_0|} & \frac{1}{|\Lambda_1|} \end{pmatrix} \begin{pmatrix}\rho_0\\ \rho_1\end{pmatrix}, \tag{A.22}$$

stretching both ρ_0 and ρ_1 over the whole unit interval Λ. Ulam [56, 2] had conjectured that successive refinements of such piece-wise linear coarse-grainings would provide a convergent sequence of finite-state Markov approximations to the Perron-Frobenius operator.

The key idea of the periodic orbit theory is to abandon explicit construction of natural measure (the density functions typically observed in chaotic systems are highly singular) and instead compute chaotic spatial and time average (A.14) from the leading eigenvalue $z_0 = z(\beta)$ of an evolution operator by means of the *classical trace formula* [73, 61], which for map f, takes form

$$\langle a\rangle = \left.\frac{\partial z_0}{\partial\beta}\right|_{\beta=0}, \qquad \sum_{\alpha=0}^{\infty}\frac{1}{z-z_\alpha} = \sum_p n_p \sum_{r=1}^{\infty}\frac{z^{rn_p}e^{r\beta\cdot A_p}}{\left|\det\left(\mathbf{1}-M_p^r\right)\right|}. \tag{A.23}$$

or, even better, by deploying the associated spectral determinant

$$\det\left(1-z\mathcal{L}_{det}\right) = \exp\left(-\sum_p\sum_{r=1}^{\infty}\frac{1}{r}\frac{z^{rn_p}e^{r\beta\cdot A_p}}{\left|\det\left(\mathbf{1}-M_p^r\right)\right|}\right). \tag{A.24}$$

These formulas replace the chaotic, long-time uncontrollable flow by its *periodic orbit skeleton*, decomposing the dynamical state space into regions, with each region $\mathcal{M}_p$ centered on an unstable periodic orbit p of period n_p, and the size of the p neighborhood determined by the linearization of the flow around the periodic orbit. Here M_p is the monodromy matrix (A.4), evaluated in the periodic orbit p, the deterministic exponential contraction/expansion is characterized by its Floquet multipliers $\{\Lambda_{p,1},\cdots,\Lambda_{p,d}\}$, and p contribution to (A.23) is inversely proportional to its exponentiated return time (cycle period n_p), and to the product of expanding eigenvalues of M_p. With emphasis on *expanding:* in applications to dissipative systems such as fluid flows there will be only several of these, with the contracting directions - even when their number is large or infinite - playing only a secondary role.

Periodic solutions (or 'cycles') are important because they form the skeleton of the invariant set of the long time dynamics [8, 74], with cycles ordered hierarchically; short cycles give dominant contributions to (A.23), longer cycles corrections. Errors due to neglecting long cycles can be bounded, and for hyperbolic systems they fall off exponentially or even super-exponentially with the cutoff cycle length [36]. Short cycles can be accurately determined and global averages (such as transport coefficients and Lyapunov exponents) can be computed from short cycles by means of *cycle expansions* [8, 74, 62, 61].

A handful of very special, completely hyperbolic flows are today mathematically fully and rigorously under control. Unfortunately, very few physically interesting systems are of that type, and the full picture is more sophisticated than the cartoon (A.23).

Appendix B. Confluent hypergeometric functions and Laguerre polynomials

Let now $\mathcal{L}^\dagger$ transform this new density, around the next pre-image $x_{a-2} = f^{-1}(x_{a-1})$, as

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

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

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

We then group all the terms up to order $O(D)$ and neglect $O(D^2)$ and higher, and call $\eta = \sqrt{\alpha}f'_{a-2}z_{a-2}$

$$\begin{aligned}\sum_{n=0}^{\infty}\frac{(-1)^n}{n!}(\eta)^{4n} + 4D\sum_{n=0}^{\infty}\frac{(-1)^n(4n)!}{4\left[n!(4n-2)!\right]}\alpha^{1-2n}\eta^{4n-2} = \\ e^{-\eta^4} - 4D\left[3\alpha\eta^2\right]\Phi\left(\frac{7}{4},\frac{3}{4},-\eta^4\right)\end{aligned} \tag{B.3}$$

where Φ (also sometimes called M or ${}_1F_1$ in the literature) is a confluent hypergeometric function of the first kind [75], which can be expressed as (see Appendix B):

$$\Phi\left(\frac{7}{4},\frac{3}{4},-\eta^4\right) = \frac{4}{3}e^{-\eta^4}\left(\frac{3}{4}-\eta^4\right) \tag{B.4}$$

We now want to evaluate the variance of the density (B.3)

$$\sigma_{a-2}^2 = \frac{\int dz_{a-2}z_{a-2}^2\rho_{a-2}(z_{a-2})}{\int dz_{a-2}\rho_{a-2}(z_{a-2})}\,. \tag{B.5}$$

It is useful to know, when computing the denominator of (B.5), that

$$\int \eta^2\Phi\left(\frac{7}{4},\frac{3}{4},\eta^4\right)d\eta = 0 \tag{B.6}$$

so that

$$\begin{aligned}\sigma_{a-2}^2 &= \frac{(\alpha^2 f'^4_{a-2})^{-1/2}\int\eta^2 e^{-\eta^4}d\eta - 4D[3f'^{-2}_{a-2}]\int\eta^4\Phi\left(\frac{7}{4},\frac{3}{4},-\eta^4\right)d\eta}{\int e^{-\eta^4}d\eta} \\ &= \frac{1}{f'^2_{a-2}}\left(\frac{\Gamma(3/4)}{\Gamma(1/4)}\frac{1}{\alpha}+2D\right) = \frac{\sigma_{a-1}^2+2D}{f'^2_{a-2}}\end{aligned} \tag{B.7}$$

in the last identity we used the definition of α and (64).

Next we derive (B.4), which expresses the a confluent hypergeometric function in terms of an exponential and a Laguerre polynomial. Start with the identity [75]:

$$\Phi\left(a,b,z\right)=e^{z}\Phi\left(b-a,b,-z\right) \tag{B.8}$$

in particular, the hypergeometric function in (B.3) becomes

$$\Phi\left(\frac{7}{4},\frac{3}{4},-\eta^{4}\right)=e^{-\eta^{4}}\Phi\left(-1,\frac{3}{4},\eta^{4}\right) \tag{B.9}$$

A confluent hypergeometric function can be written in terms of a Laguerre polynomial [75]:

$$L_{n}^{\alpha}(x)=\binom{n+\alpha}{n}\Phi\left(-n,\alpha+1,x\right),\,,\qquad L_{n}^{\alpha}(x)=\sum_{m=0}^{n}(-1)^{m}\binom{n+\alpha}{n-m}\frac{x^{m}}{m!} \tag{B.10}$$

Thus, in our case

$$L_{1}^{-1/4}(\eta^{4})=\binom{1-1/4}{1}\Phi\left(-1,\frac{3}{4},\eta^{4}\right) \tag{B.11}$$

and

$$\Phi\left(-1,\frac{3}{4},\eta^{4}\right)e^{-\eta^{4}}=\frac{4}{3}\left(\frac{3}{4}-\eta^{4}\right)e^{-\eta^{4}}\,. \tag{B.12}$$

References

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

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

[3] H. Risken, The Fokker-Planck Equation, Springer, New York, 1996.

[4] P. Gaspard, Trace formula for noisy flows, J. Stat. Phys. 106 (2002) 57.

[5] H. C. Fogedby, Localized growth modes, dynamic textures, and upper critical dimension for the Kardar-Parisi-Zhang equation in the weak noise limit, Phys. Rev. Lett. 94 (2005) 195702.

[6] H. C. Fogedby, Kardar-Parisi-Zhang equation in the weak noise limit: Pattern formation and upper critical dimension, Phys. Rev. E 73 (2006) 031104.

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

[8] P. Cvitanović, Invariant measurement of strange sets in terms of cycles, Phys. Rev. Lett. 61 (1988) 2729.

[11] A. Einstein, Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen, Ann. Physik 17 (1905) 549.

[12] H. Dekker, N. G. V. Kampen, Eigenvalues of a diffusion process with a critical point, Phys. Lett. A 73 (1979) 374–376.

[13] P. Gaspard, G. Nicolis, A. Provata, S. Tasaki, Spectral signature of the pitchfork bifurcation: Liouville equation approach, Phys. Rev. E 51 (1995) 74.

[14] D. Lippolis, P. Cvitanović, How well can one resolve the state space of a chaotic map?, Phys. Rev. Lett. 104 (2010) 014101, arXiv:0902.4269.

[15] P. S. Laplace, Mémoire sur les intégrales définies et leur application aux probabilités, et spécialement à la recherche du milieu qu'il faut choisir entre les résultats des observations, Mem. Acad. Sci. (I), XI, Section V. (1810) 375–387.

[16] L. Arnold, Stochastic Differential Equations: Theory and Applications, Wiley, New York, 1974.

[17] S. Chandrasekhar, Stochastic problems in physics and astronomy, Rev. Mod. Phys. 15 (1943) 1.

[] J. L. Doob, The Brownian movement and stochastic equations, Ann. Math. 43 (1942) 351–369.

[18] B. Shraiman, C. E. Wayne, P. C. Martin, Scaling theory for noisy period-doubling transitions to chaos, Phys. Rev. Lett. 46 (1981) 935–939.

[19] P. C. Martin, E. D. Siggia, H. A. Rose, Statistical mechanics of classical systems, Phys. Rev. A 8 (1973) 423.

[20] K. Itô, Stochastic integral, Proc. Imp. Acad. Tokyo 20 (1944) 519–524.

[21] R. L. Stratonovich, Conditional Markov Processes and Their Application to the Theory of Control, Elsevier, New York, 1968.

[] R. F. Fox, Gaussian stochastic processes in physics, Phys. Rep. 48 (1978) 179–283.

[] N. Berglund, B. Gentz, Noise-Induced Phenomena in Slow-Fast Dynamical Systems: A Sample-Paths Approach, Springer, Berlin, 2005.

[] L. S. Pontryagin, V. Boltyanskii, R. Gamkrelidge, E. Mishenko, Mathematical Theory of Optimal Processes, Interscience Publishers, New York, 1962.

[] R. Bellman, Dynamic Programming, Princeton Univ. Press, Princeton, 1957.

[22] L. Onsager, S. Machlup, Fluctuations and irreversible processes, Phys. Rev. 91 (1953) 1505, 1512.

[23] M. I. Freidlin, A. D. Wentzel, Random Perturbations of Dynamical Systems, Springer, Berlin, 1998.

[24] R. Feynman, A. Hibbs, Quantum Mechanics and Path Integrals, McGraw-Hill Book Co., New York, 1965.

[25] P. A. M. Dirac, The Lagrangian in quantum mechanics, Phys. Z. Sowjetunion 3 (1933) 64–72.

[26] M. Roncadelli, Small-fluctuation expansion of the transition probability for a diffusion process, Phys. Rev. E 52 (1995) 4661.

[27] A. Defendi, M. Roncadelli, Semiclassical approximation as a small-noise expansion, J. Phys. A 28 (1995) L515–L520.

[28] Y. Kifer, On small random perturbations of some smooth dynamical systems, Math. USSR-Izv. 8 (1974) 1083–1107.

[29] M. J. Feigenbaum, B. Hasslacher, Irrational decimations and path-integrals for external noise, Phys. Rev. Lett. 49 (1982) 605–609.

[30] A. Boyarsky, On the significance of absolutely continuous invariant measures, Physica D 11 (1984) 130–146.

[31] P. Cvitanović, C. P. Dettmann, R. Mainieri, G. Vattay, Trace formulas for stochastic evolution operators: weak noise perturbation theory, J. Stat. Phys. 93 (1998) 981,
arXiv:chao-dyn/9807034.

[32] P. Cvitanović, C. P. Dettmann, R. Mainieri, G. Vattay, Trace formulae for stochastic evolution operators: smooth conjugation method, Nonlinearity 12 (1999) 939,
arXiv:chao-dyn/9811003.

[33] P. Cvitanović, N. Søndergaard, G. Palla, G. Vattay, C. P. Dettmann, Spectrum of stochastic evolution operators: local matrix representation approach, Phys. Rev. E 60 (1999) 3936,
arXiv:chao-dyn/9904027.

[] M. Gutzwiller, Chaos in Classical and Quantum Mechanics, Springer, Berlin, 1990.

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

[] R. E. Kalman, A new approach to linear filtering and prediction problems, Trans. ASME – J. Basic Engineering 82 (1960) 35–45.

[34] H. D. I. Abarbanel, D. R. Creveling, R. Farsian, M. Kostuk, Dynamical state and parameter estimation, SIAM J. Appl. Math. 8 (2009) 1341–1381.

[] J. Keizer, Statistical Thermodynamics of Nonequilibrium Processes, Springer, Berlin, 1987.

[] J. Paulsson, Summing up noise in gene expression, Nature 427 (2004) 415–418.

[] J. Paulsson, Models of stochastic gene expression, Phys. Life Rev. 2 (2005) 157–175.

[] G. Hornung, N. Barkai, Noise propagation and signaling sensitivity in biological networks, PloS Computational Biology 4 (2008) e80055.

[] F. J. Bruggeman, N. Blüthgen, H. V. Westerhoff, Noise management by molecular networks, PloS Computational Biology 5 (2009) e1000506.

[] O. Cepas, J. Kurchan, Canonically invariant formulation of Langevin and Fokker-Planck equations, European Physical Journal **B** 2 (1998) 221, `arXiv:cond-mat/9706296`.

[] V. I. Oseledec, A multiplicative ergodic theorem. liapunov characteristic numbers for dynamical systems, Trans. Moscow Math. Soc. 19 (1968) 197–221.

[36] H. H. Rugh, The correlation spectrum for hyperbolic analytic maps, Nonlinearity 5 (1992) 1237.

[37] G. E. Uhlenbeck, L. S. Ornstein, On the theory of the Brownian motion, Phys. Rev. 36 (1930) 823–841.

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

[39] M. Abramowitz, I. A. Stegun (Eds.), Handbook of Mathematical Functions, Dover, New York, 1964.

[40] H. Takada, Y. Kitaoka, Y. Shimizua, Mathematical index and model in stabirometry, Forma 16 (2001) 17–46.

[41] M. Jacobsen, Laplace and the origin of the Ornstein-Uhlenbeck process, Bernoulli 2 (1997) 271–286.

[42] P. Cvitanović, R. Artuso, L. Rondoni, E. A. Spiegel, chapter *"Transporting densities"*, in ref. [61].

[43] V. Fedorov, Theory of Optimal Experiments, New York, Academic Press, 1972.

[44] L. D. Brown, I. Olkin, J. Sacks, H. P. Wynn (Eds.), Jack Carl Kiefer Collected Papers, Vol. III, Springer, Berlin, 1985, Ch. Design of Experiments, p. 718.

[45] A. C. Atkinson, V. V. Fedorov, The design of experiments for discriminating between two rival models, Biometrika 62 (1975) 57–70.

[46] V. Fedorov, V. Khabarov, Duality of optimal designs for model discrimination and parameter estimation, Biometrika 73 (1986) 183–190.

[47] J. P. Crutchfield, N. H. Packard, Symbolic dynamics of noisy chaos, Physica D 7 (1983) 201–223.

[48] C. S. Daw, C. E. A. Finney, E. R. Tracy, A review of symbolic analysis of experimental data, Rev. Sci. Instrum. 74 (2003) 915–930.

[49] X. Z. Tang, E. R. Tracy, A. D. Boozer, A. deBrauw, R. Brown, Symbol sequence statistics in noisy chaotic signal reconstruction, Phys. Rev. E 51 (1995) 3871–3889.

[50] M. Lehrman, A. B. Rechester, R. B. White, Symbolic analysis of chaotic signals and turbulent fluctuations, Phys. Rev. Lett. 78 (1997) 54.

[51] M. B. Kennel, M. Buhl, Estimating good discrete partitions from observed data: Symbolic false nearest neighbors, Phys. Rev. Lett. 91 (2003) 084102, `arXiv:nlin/0304054`.

[52] M. B. Kennel, M. Buhl, Estimating good discrete partitions from observed data: symbolic false nearest neighbors, in: L. Kočarev, T. L. Carroll, B. J. Gluckman, S. Boccaletti, J. Kurths (Eds.), Experimental Chaos: Seventh Experimental Chaos Conference, Am. Inst. of Phys., Melville, New York, 2003, pp. 380–380.

[53] M. Buhl, M. B. Kennel, Statistically relaxing to generating partitions for observed time-series data, Phys. Rev. E 71 (2005) 046213.

[54] D. Holstein, H. Kantz, Optimal Markov approximations and generalized embeddings, Phys. Rev. E 79 (2009) 056202, `arXiv:0808.1513`.

[55] R. P. Boland, T. Galla, A. J. McKane, Limit cycles, complex Floquet multipliers and intrinsic noise, `arXiv:0903.5248` (2009).

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

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

[58] A. B. Rechester, R. B. White, Symbolic kinetic-equation for a chaotic attractor, Phys. Lett. A 156 (1991) 419–424.

[59] A. B. Rechester, R. B. White, Symbolic kinetic analysis of two-dimensional maps, Phys. Lett. A 158 (1991) 51–56.

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

[62] P. Gaspard, Chaos, Scattering and Statistical Mechanics, Cambridge Univ. Press, Cambridge, 1997.

[63] R. Artuso, E. Aurell, P. Cvitanović, Recycling of strange sets: II. Applications, Nonlinearity 3 (1990) 361.

[64] D. Ruelle, A review of linear response theory for general differentiable dynamical systems, Nonlinearity 22 (2009) 855–870,
arXiv:arXiv.org:0901.0484.

[65] J. R. Crutchfield, M. Nauenberg, J. Rudnick, Scaling for external noise at the onset of chaos, Phys. Rev. Lett. 46 (1981) 933.

[9] G. Kawahara, S. Kida, Periodic motion embedded in plane Couette turbulence: regeneration cycle and burst, J. Fluid Mech. 449 (2001) 291–300.

[10] P. Cvitanović, J. F. Gibson, Geometry of turbulence in wall-bounded shear flows: Periodic orbits, in S. I. Abarzhi and K. R. Sreenivasan, eds., *Turbulent Mixing and Beyond, Proc. 2nd Int. Conf. ICTP, Trieste*, *Phys. Scr.*, to appear (2010).

[66] H. D. I. Abarbanel, R. Brown, M. B. Kennel, Lyapunov exponents in chaotic systems: their importance and their evaluation using observed data, Int. J. Mod. Phys. B 5 (1991) 1347–1375.

[67] H. D. I. Abarbanel, R. Brown, M. B. Kennel, Variation of Lyapunov exponents on a strange attractor, J. Nonlinear Sci. 1 (1991) 175–199.

[68] J. C. Quinn, H. D. I. Abarbanel, State and parameter estimation using Monte Carlo evaluation of path integrals,
arXiv:0912.1581.

[69] J. F. Gibson, J. Halcrow, P. Cvitanović, Visualizing the geometry of state-space in plane Couette flow, J Fluid Mech. 611 (2008) 107–130,

arXiv:0705.3957.

[70] D. Ruelle, Zeta functions for expanding maps and Anosov flows, Invent. Math. 34 (1976) 231.

[71] M. C. Gutzwiller, Periodic orbits and classical quantization conditions, J. Math. Phys. 12 (1971) 343–358.

[72] R. Artuso, H. H. Rugh, P. Cvitanović, chapter *"Why does it work?"*, in ref. [61].

[73] P. Cvitanović, B. Eckhardt, Periodic orbit expansions for classical smooth flows, J. Phys. A 24 (1991) L237.

[74] R. Artuso, E. Aurell, P. Cvitanović, Recycling of strange sets: I. Cycle expansions, Nonlinearity 3 (1990) 325–359.

[75] I. S. Gradshteyn, I. M. Ryzhik, Table of Integrals. Series, and Products, Academic Press, New York, 1965.