

Knowing when to stop: How noise frees us from determinism

Predrag Cvitanović* and Domenico Lippolis†

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

†Department of Physics, Pusan National University, Busan 609-735, South Korea

Abstract. Deterministic chaotic dynamics presumes that the state space can be partitioned arbitrarily finely. In a physical system, the inevitable presence of some noise sets a finite limit to the finest possible resolution that can be attained. Much previous research deals with what this attainable resolution might be, all of it based on a global averages over stochastic flow. We show how to compute the *locally* optimal partition, for a given dynamical system and given noise, in terms of local eigenfunctions of the Fokker-Planck operator and its adjoint. 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, on sets of unstable periodic orbits of the deterministic system. We test our hypothesis on unimodal maps.

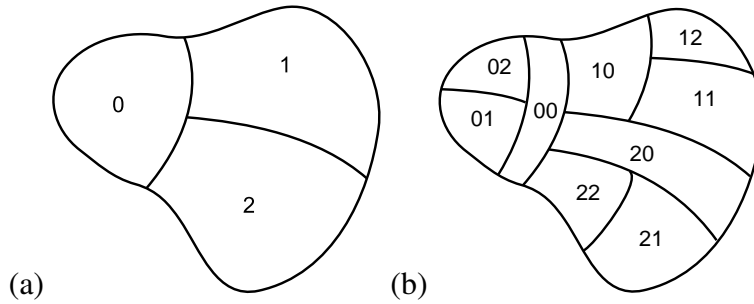
Keywords: noise, stochastic dynamics, Fokker-Planck operator, chaos, cycle expansions, periodic orbits, perturbative expansions, trace formulas, spectral determinants, zeta functions.

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

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*. In ref. [7] we have shown that the hypothesis enables us to define the optimal partition for a 1-dimensional map; here we explain how it is implemented for a high-dimensional state space flows with a few expanding directions, such as the transitional *Re* number Navier-Stokes flows. 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 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



Source: ChaosBook.org

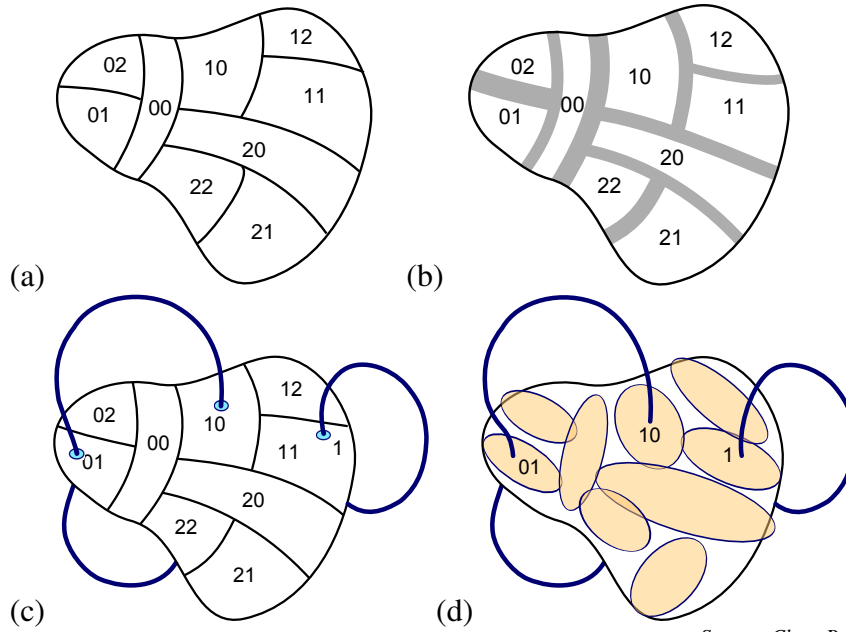
FIGURE 1. (a) A coarse partition of state space \mathcal{M} into regions \mathcal{M}_0 , \mathcal{M}_1 , and \mathcal{M}_2 , labeled by ternary alphabet $\mathcal{A} = \{1, 2, 3\}$. (b) A 1-step memory refinement of the partition of figure 1, with each region \mathcal{M}_i subdivided into \mathcal{M}_{i0}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

implement such partition of the state space, each region a neighborhood of a periodic point [8, 9]. 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, \dots) 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 (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. 5, 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 [10, 11, 12] by a diffusion constant, 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



Source: ChaosBook.org

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 adjacent neighborhoods. (c) The fixed point $\bar{1} = \{x_1\}$ and the two-cycle $\bar{01} = \{x_{01}, x_{10}\}$ 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.

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 remark 5.1) 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. [7].

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 is much simpler than it sounds: the Lyapunov equation

$$Q = MQM^T + \Delta$$

(and its generalizations to periodic points of hyperbolic flows), determines Q , the size of local neighborhood, as balance of the noise variance Δ and the linearized dynamics M . The effort of going local 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 of lower-dimensional dynamics, and a persuasive application 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’ should provide an optimal, computable 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. These formulas are known as Lyapunov equations, reviewed in appendix C. In order to understand the effect on noise on the hyperbolic, mixed expanding / contracting dynamics, we study the eigenfunctions of the discrete time Fokker-Planck operator in linear neighborhood of a fixed point 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 Fokker-Planck operator. The continuous time formulation of the same problem, known as the Ornstein-Uhlenbeck process, is reviewed in appendix B. 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 remark 5.1. We formulate our optimal partition hypothesis in sect. 5: 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. 6, we assess the accuracy of our method by computing the escape rate and the Lyapunov exponent, discuss weak noise corrections, and compare the results with a discretization of the Fokker-Planck operator on a uniform mesh. In sect. 7 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 results are summarized and the open problems discussed in sect. 8.

2. NOISY TRAJECTORIES AND THEIR DENSITIES

The literature on stochastic dynamical systems is vast, starting with the Laplace 1810 memoir [13]. The material reviewed in this section, sect. 4 and appendix B is standard [1, 3, 14], 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 derived here is the well known evolution law (14) for the covariance matrix Q_a of a linearly evolved Gaussian density,

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

To keep things simple we shall use only the discrete time dynamics in what follows, but we do discuss the continuous time formulation in appendix B, as our results apply both to the continuous and discrete time flows.

Consider a noisy *discrete time* dynamical system [15, 16, 17]

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

where x is a d -dimensional state vector, and $x_{n,j}$ is its j th component at time n . In the Fokker-Planck description individual noisy trajectories are replaced by the evolution of the density of noisy trajectories, with the $x_{n+1} - f(x_n)$ probability distribution of zero mean and covariance matrix (diffusion tensor) Δ ,

$$\langle \xi_{n,j} \rangle = 0, \quad \langle \xi_{n,i} \xi_{m,j}^T \rangle = \Delta_{ij} \delta_{nm}, \quad (2)$$

where $\langle \dots \rangle$ stands for ensemble average over many realizations of the noise.

The general case of a diffusion tensor $\Delta(x)$ which is a state space position dependent but time independent can be treated along the same lines. In this case the stochastic flow (1) is written as $x_{n+1} = x_n + \sigma(x) \xi_n$, where $\langle \xi_n \xi_m^T \rangle = \mathbf{1} \delta_{nm}$ is white noise, $\Delta = \sigma \sigma^T$, $\sigma(x)$ is called the ‘diffusion matrix’, and the noise is referred to as ‘multiplicative’ (see Kuehn [18]).

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

$$\begin{aligned} \rho_{k+1}(y) &= [\mathcal{L} \rho_k](y) = \int dx \mathcal{L}(y, x) \rho_k(x) \\ \mathcal{L}(y, x) &= \frac{1}{N} e^{-\frac{1}{2}(y-f(x))^T \frac{1}{\Delta}(y-f(x))}, \end{aligned} \quad (3)$$

is centered on the deterministic step $f(x)$ and smeared out diffusively by noise. Were diffusion uniform and isotropic, $\Delta(x) = 2D\mathbf{1}$, the Fokker-Planck operator would be proportional to $\exp(-\{y - f(x)\}^2/2\Delta)$, i.e., the penalty for straying from the deterministic path is just a quadratic error function. The k th iterate of \mathcal{L} is a d -dimensional path integral over the $k - 1$ intermediate noisy trajectory points,

$$\mathcal{L}^k(x_k, x_0) = \int [dx] e^{-\frac{1}{2} \sum_n (x_{n+1} - f(x_n))^T \frac{1}{\Delta} (x_{n+1} - f(x_n))}, \quad (4)$$

where the Gaussian normalization factor in (3) is absorbed into intermediate integrations by defining

$$[dx] = \prod_{n=1}^{k-1} \frac{dx_n^d}{N}, \quad N = \sqrt{2\pi^d \det \Delta}. \quad (5)$$

We shall also need to determine the effect of noise accumulated along the trajectory points *preceding* x . As the noise is additive forward in time, one cannot simply invert the Fokker-Planck operator; instead, the past is described by the *adjoint Fokker-Planck operator*,

$$\tilde{\rho}_{k-1}(x) = [\mathcal{L}^\dagger \tilde{\rho}_k](x) = \int [dy] e^{-\frac{1}{2}(y-f(x))^T \frac{1}{\Delta}(y-f(x))} \tilde{\rho}_k(y), \quad (6)$$

which transports a density concentrated around the point $f(x)$ to a density concentrated around the previous point x and adds noise to it. In the deterministic, vanishing noise limit this is the Koopman operator (76).

The Fokker-Planck operator (3) is non-hermitian and non-unitary. For example, if the deterministic flow is contracting, the natural measure (the leading right eigenvector of the Fokker-Planck operator) will be concentrated and peaked, but then the corresponding left eigenvector has to be broad and flat, as backward in time the deterministic flow is expanding. We shall denote by ρ_α the right eigenvectors of \mathcal{L} , and by $\tilde{\rho}_\alpha$ its left eigenvectors, i.e., the right eigenvectors of the adjoint operator \mathcal{L}^\dagger .

3. ALL NONLINEAR NOISE IS LOCAL

Our first goal is to convince the reader that the diffusive dynamics of nonlinear flows is *fundamentally different from Brownian motion*, with the flow inducing a local, history dependent noise. In order to accomplish this, we go beyond the standard stochastic literature and generalize the notion of invariant deterministic recurrent solutions, such as fixed points and periodic orbits, to noisy flows. While a Langevin trajectory (1) cannot be periodic, in the Fokker-Planck formulation (4) a recurrent motion can be defined as one where a peaked distribution returns to the initial neighborhood after time n . Recurrence so defined not only coincides with the classical notion of a recurrent orbit in the vanishing noise limit, but it also enables us to derive exact formulas for how this local, history dependent noise is to be computed.

As the function $x_{n+1} - f(x_n)$ is a nonlinear function, in general the path integral (4) 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 (75). For weak noise the Fokker-Planck operator can be evaluated perturbatively [19, 20, 21, 22] as an asymptotic series in powers of the diffusion constant, centered on the deterministic trajectory. Here we retain only the linear term in this series, which has a particularly simple dynamics given by a covariance matrix evolution formula (see (14) and (22) below) that we now derive.

We shift local coordinates to the deterministic trajectory $\{\dots, x_{-1}, x_0, x_1, x_2, \dots\}$ centered coordinate frames $x = x_a + z_a$, Taylor expand $f(x) = f_a(z_a) = x_{a+1} + M_a z_a + \dots$, and approximate the noisy map (1) by its linearization,

$$z_{a+1} = M_a z_a + \xi_a, \quad M_{ij} = \partial f_i / \partial x_j \quad (7)$$

with the deterministic trajectory points at $z_a = z_{a+1} = 0$, and M_a the one step Jacobian matrix (see appendix A). The corresponding linearized Fokker-Planck operator (3) is

given in the local coordinates $\rho_a(z_a) = \rho(x_a + z_a, a)$ by

$$\rho_{a+1}(z_{a+1}) = \int dz_a \mathcal{L}_a(z_{a+1}, z_a) \rho_a(z_a) \quad (8)$$

by the linearization (7) centered on the deterministic trajectory

$$\mathcal{L}_a(z_{a+1}, z_a) = \frac{1}{N} e^{-\frac{1}{2}(z_{a+1} - M_a z_a)^T \frac{1}{\Delta} (z_{a+1} - M_a z_a)}. \quad (9)$$

The subscript ‘ a ’ in \mathcal{L}_a distinguishes the local, linearized Fokker-Planck operator from the full operator (4).

The kernel of the linearized Fokker-Planck operator (9) 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

$$\rho_a(z) = \frac{1}{C_a} e^{-\frac{1}{2} z^T \frac{1}{Q_a} z}, \quad C_a = (2\pi)^{d/2} (\det Q_a)^{1/2}, \quad (10)$$

and the action of the linearized adjoint Fokker-Planck operator on density

$$\tilde{\rho}_a(z) = \frac{1}{C_a} e^{-\frac{1}{2} z^T \frac{1}{\tilde{Q}_a} z}, \quad C_a = (2\pi)^{d/2} (\det \tilde{Q}_a)^{1/2}, \quad (11)$$

also centered on the deterministic trajectory, but with its own strictly positive $[d \times d]$ covariance matrices Q, \tilde{Q} . Label ‘ a ’ plays a double role, and $\{a+1, a\}$ stands both for the {next, initial} space partition and for the times the trajectory lands in these partitions (see appendix A.1). The linearized Fokker-Planck operator (9) maps the Gaussian $\rho_a(z_a)$ into the Gaussian

$$\rho_{a+1}(z_{a+1}) = \frac{1}{C_a} \int [dz_a] e^{-\frac{1}{2} [(z_{a+1} - M_a z_a)^T \frac{1}{\Delta} (z_{a+1} - M_a z_a) + z_a^T \frac{1}{Q_a} z_a]} \quad (12)$$

one time step later. Likewise, linearizing the adjoint Fokker-Planck operator (6) around the x_a trajectory point yields:

$$\tilde{\rho}_a(z_a) = \frac{1}{C_{a+1}} \int [dz_{a+1}] e^{-\frac{1}{2} [(z_{a+1} - M_a z_a)^T \frac{1}{\Delta} (z_{a+1} - M_a z_a) + z_{a+1}^T \frac{1}{\tilde{Q}_{a+1}} z_{a+1}]} \quad (13)$$

Completing the squares, integrating and substituting (10), respectively (11) we obtain the formula for Q covariance matrix evolution forward in time,

$$Q_{a+1} = M_a Q_a M_a^T + \Delta_a, \quad (14)$$

and in the adjoint case, the evolution of the \tilde{Q} is given by

$$M_a \tilde{Q}_a M_a^T = \tilde{Q}_{a+1} + \Delta_a. \quad (15)$$

The two covariance matrices differ, as the adjoint evolution \tilde{Q}_a is computed by going backwards along the trajectory. These *covariance evolution* rules are the basis of all that follows, except for the ‘flat-top’ of sect. 7.

Think of the initial covariance matrix (10) as an error matrix describing the precision of the initial state, a cigar-shaped probability distribution $\rho_a(z_a)$. In one time step this density is deterministically advected and deformed into density with covariance MQM^T , and then the noise Δ is added: the two kinds of independent uncertainties add up as sums of squares, hence the covariance evolution law (14), resulting in the Gaussian ellipsoid whose widths and orientation are given by the singular values and singular vectors (68) of the covariance matrix. After n time steps, 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.

The pleasant surprise is 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. We have attached label ‘ a ’ to $\Delta_a = \Delta(x_a)$ in (14) to account for the noise distributions that are inhomogeneous, state space dependent, but time independent multiplicative noise. As we shall show, in nonlinear dynamics the noise is *never* isotropic and/or homogeneous. For example, if the iterative system we are studying is obtained by Poincaré sections of a continuous time flow, the accumulated noise integrated over one Poincaré section return depends on the return trajectory segment, even when the infinitesimal time step noise (87) is homogenous.

Remark 3.1 Covariance evolution. In quantum mechanics the linearized evolution operator corresponding to the linearized Fokker-Planck operator (9) is known as the Van Vleck propagator, the basic block in the semi-classical periodic orbit quantization [23, 24]. Q covariance matrix composition rule (14) or its continuous time version (106) is called ‘covariance evolution’ in ref. [25], for example, but it goes all the way back to Lyapunov’s 1892 thesis [26], see appendix C. In the Kalman filter literature [27, 28] it is called ‘prediction’.

3.1. The attractive, the repulsive and the noisy

For Brownian dynamics $x_{n+1} = x_n + \xi_n$, with $M = \mathbf{1}$, we obtain $Q_n = Q_0 + n\Delta$, i.e., the variance of a Gaussian packet of $\rho_n(z)$ noisy trajectories grows linearly in time, as expected for the Brownian diffusion. What happens for nontrivial, $M \neq \mathbf{1}$, dynamics? The formulas (14) and (15) are exact for finite numbers of time steps, but whether they have a long time limit depends on the stability of the deterministic trajectory.

Here we shall derive the $n \rightarrow \infty$ limit for deterministic flows that are either contracting or expanding in all eigen-directions, with asymptotic stationary distributions concentrated either on fixed points or periodic points. We shall consider the general hyperbolic flows, where some of the eigen-directions are unstable, and other stable in another publication [29]. In this context the 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 eigen-directions, backward in time along the unstable, expanding eigen-directions. The short cycles will be the most important ones, and only finite time, single cycle period calculations will be required.

If M is contracting, with the multipliers $1 > |\Lambda_1| \geq |\Lambda_2| \geq \dots \geq |\Lambda_d|$, in n time steps the memory of the covariance Q_{a-n} of the starting density is forgotten at exponential rate $\sim |\Lambda_1|^{-2n}$, with iteration of (14) leading to a limit distribution:

$$Q_a = \Delta_a + M_{a-1}\Delta_{a-1}M_{a-1}^T + M_{a-2}^2\Delta_{a-2}(M_{a-2}^2)^T + \dots \quad (16)$$

For fixed and periodic points we can give an explicit formula for the $n \rightarrow \infty$ covariance.

Consider a noisy map (1) with a deterministic fixed point at x_q . In a neighborhood $x = x_q + z$ we approximate the map f by its linearization (7) with the fixed point at $z = 0$, acting on a Gaussian density distribution (10), 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 11 (b). If the fixed point is *attractive*, with all multipliers of M strictly contracting, any compact initial measure (not only initial distributions of Gaussian form) converges under applications of (14) to the unique invariant natural measure $\rho_0(z)$ whose covariance matrix satisfies the condition

$$Q = MQM^T + \Delta. \quad (17)$$

For a *repelling* fixed point the condition (15) on the adjoint eigenvector $\tilde{\rho}_0$ yields

$$M\tilde{Q}M^T = \tilde{Q} + \Delta, \quad (18)$$

with a very different interpretation: as the Jacobian matrix M has only *expanding* Floquet multipliers, the deterministic dynamics expands the fixed-point neighborhood exponentially, with no good notion of a local neighborhood in the large forward time limit. Instead, its past defines the neighborhood, with \tilde{Q} the covariance of the optimal distribution of points that can reach the fixed point in one time step, given the diffusion tensor Δ .

These conditions are central to control theory, where the attracting fixed point condition (17) is called the *Lyapunov equation* (see appendix C), Q and \tilde{Q} are known respectively as controllability and observability Gramians, and there is much wisdom and open source code available to solve these (see remark C.1), as well as the more general hyperbolic equations. In order to develop some intuition about the types of solutions we shall encounter, we assume first, for illustrative purposes, that $[d \times d]$ Jacobian matrix M has distinct real contracting Floquet multipliers $\{\Lambda_1, \Lambda_2, \dots, \Lambda_d\}$ and right eigenvectors $M\mathbf{e}^{(j)} = \Lambda_j\mathbf{e}^{(j)}$. Construct from the d column eigenvectors a $[d \times d]$ similarity transformation

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

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 (17) now takes form $\hat{Q} - \Lambda\hat{Q}\Lambda = \hat{\Delta}$. The matrix elements are $\hat{Q}_{ij}(1 - \Lambda_i\Lambda_j) = \hat{\Delta}_{ij}$, so

$$\hat{Q}_{ij} = \frac{\hat{\Delta}_{ij}}{1 - \Lambda_i\Lambda_j}, \quad (19)$$

and the attracting fixed point covariance matrix in the original coordinates is given by

$$Q = S\hat{Q}S^T. \quad (20)$$

For the adjoint case, the same algebra yields

$$\hat{Q}_{ij} = \frac{\hat{\Delta}_{ij}}{\Lambda_i\Lambda_j - 1}, \quad (21)$$

for the matrix elements of \hat{Q} , with the covariance matrix in the fixed coordinates again given by $\hat{Q} = S\hat{Q}S^T$.

As (20) is not a similarity transformation, evaluation of the covariance matrix Q requires a numerical diagonalization, which yields the singular values and singular vectors (principal axes) of the equilibrium Gaussian ‘cigar’ (see appendix A). The singular vectors of this symmetric matrix have their own orientations, distinct from the left/right eigenvectors of the non-normal Jacobian matrix M .

Remark 3.2 Hyperbolic flows. The methods to treat the cases where some of the eigen-directions are unstable, and other stable are implicit in the Oseledec [30] definition of Lyapunov exponents, the rigorous proof of existence of classical spectral (Fredholm) determinants by Rugh [31], and the controllability and observability Gramians of control theory [32]: the flow at a hyperbolic fixed point or cycle point can be locally factorized into stable and unstable directions, and for unstable directions one needs to study noise evolution in the past, by means of the adjoint operator (6).

3.2. In nonlinear world noise is never isotropic

Now that we have established the exact formulas (14), (15) for the extent of the noise-smearred out neighborhood of a fixed point, we turn to the problem of computing them for periodic orbits. An attractive feature of the deterministic periodic orbit theory is that certain properties of periodic orbits, such as their periods and Floquet multipliers, are *intrinsic*, independent of where they are measured along the the periodic 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 [33]). Each cycle point has a different memory and differently distorted neighborhood, so we need to compute the Fokker-Planck eigenfunction ρ_a at each cycle point x_a .

The basic idea is simple: A periodic point of an n -cycle is a fixed point of the n th iterate of the map (1). Hence the formula (16) for accumulated noise, together the fixed point condition (17) also yields the natural measure covariance matrix at a periodic point x_a on a periodic orbit p ,

$$Q_a = M_{p,a}Q_aM_{p,a}^T + \Delta_{p,a}, \quad (22)$$

where

$$\begin{aligned}\Delta_{p,a} = & \Delta_a + M_{a-1}\Delta_{a-1}M_{a-1}^T + M_{a-2}^2\Delta_{a-2}(M_{a-2}^2)^T \\ & + \cdots + M_{a-n_p+1}^{n_p-1}\Delta_{a-n_p+1}(M_{a-n_p+1}^{n_p-1})^T\end{aligned}\quad (23)$$

is the noise accumulated per a single transversal of the periodic orbit, $M_{p,a} = M_p(x_a)$ is the cycle Jacobian matrix (70) evaluated on the periodic point x_a , and we have used the periodic orbit condition $x_{a+n_p} = x_a$. Similarly, for the adjoint evolution the fixed point condition (18) generalizes to

$$M_{p,a}\tilde{Q}_aM_{p,a}^T = \tilde{Q}_a + \tilde{\Delta}_{p,a}, \quad (24)$$

where

$$\begin{aligned}\tilde{\Delta}_{p,a} = & \Delta_a + M_{a+1}\Delta_{a+1}M_{a+1}^T + M_{a+2}^2\Delta_{a+2}(M_{a+2}^2)^T \\ & + \cdots + M_{a+n_p-1}^{n_p-1}\Delta_{a+n_p-1}(M_{a+n_p-1}^{n_p-1})^T\end{aligned}\quad (25)$$

is the noise accumulated per a single transversal of the periodic orbit backward in time.

As there is no single coordinate frame in which different $M_{a-k}^k\Delta_{a-k}(M_{a-k}^k)^T$ 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. ONE-DIMENSIONAL INTUITION

The very general, exact formulas that we have obtained so far (and so easily), valid in any dimension, might be elegant, but it is a bit hard to get one's head around a formula such as the expression for the accumulated cycle noise (23). These results are easier to grasp by studying the effect of noise on 1-dimensional systems, such as the noisy linear map (7),

$$z_{n+1} = f(z_n) + \xi_n, \quad f(z_n) = \Lambda z_n, \quad (26)$$

with the deterministic fixed point at $f(z) = z = 0$, and additive white noise (2) with variance Δ . The density $\rho(x)$ of trajectories evolves by the action of the Fokker-Planck operator (3):

$$[\mathcal{L}\rho](x) = \int [dy] e^{-\frac{1}{2}\frac{(x-\Lambda y)^2}{\Delta}} \rho(y). \quad (27)$$

If a 1-dimensional noisy linear map (26) is *contracting*, any initial compact measure converges under applications of (27) to the unique invariant natural measure $\rho_0(z)$ concentrated at the deterministic fixed point $z = 0$ whose variance (10) is given by (17):

$$Q = \frac{\Delta}{1-\Lambda^2}, \quad \rho_0(z) = \frac{1}{\sqrt{2\pi Q}} e^{-z^2/2Q}. \quad (28)$$

The variance (22) of a periodic point x_a on an attractive n -cycle p is

$$Q_a = \frac{\Delta_{p,a}}{1-\Lambda^2}, \quad \Lambda = f_a^{n'}, \quad (29)$$

where the accumulated noise per a cycle traversal (23) is given by

$$\Delta_{p,a} = \Delta(1 + (f'_{a-1})^2 + (f'_{a-2})^2 + \cdots + (f'^{n-1}_{a-n_p+1})^2). \quad (30)$$

Variance (28) expresses a balance between contraction by Λ and diffusive smearing by Δ at each time step. For strongly contracting Λ , the width is due to the noise only. As $|\Lambda| \rightarrow 1$ the width diverges: the trajectories are only weakly confined and diffuse by Brownian motion into a broad Gaussian.

Consider next the adjoint operator acting on a *repelling* noisy fixed point, $|\Lambda| > 1$. The stationary measure condition (18) yields

$$\tilde{Q} = \frac{\Delta}{\Lambda^2 - 1}, \quad \tilde{\rho}_0(z) = \frac{1}{\sqrt{2\pi\tilde{Q}}} e^{-z^2/2\tilde{Q}}. \quad (31)$$

While the dominant feature of the attracting fixed point variance (28) was the diffusion strength Δ , weakly modified by the contracting multiplier, for the unstable fixed point the behavior is dominated by the expanding multiplier Λ ; the more unstable the fixed point, the smaller is the neighborhood one step in the past that can reach it.

The variance (24) of a periodic point x_a on an unstable n -cycle p is

$$\tilde{Q}_a = \frac{\Delta}{1 - \Lambda_p^{-2}} \left(\frac{1}{(f'_a)^2} + \frac{1}{(f'_{a+1})^2} \cdots + \frac{1}{\Lambda_p^2} \right). \quad (32)$$

For an unstable cycle typically all derivatives along the cycle are expanding, $|f'_{a+k}| > 1$, so the dominant term in (32) is the most recent one, $\tilde{Q}_a \approx \Delta/(f'_a)^2$. By contrast, forward in time (29) the leading estimate of variance of an attractive periodic point is $Q_a \approx \Delta$. These leading estimates are not sensitive to the length of the periodic orbit, so all trajectories passing through a neighborhood of periodic point x_a will have comparable variances.

4.1. Ornstein-Uhlenbeck spectrum

The variance (17) is stationary under the action of \mathcal{L} , and the corresponding Gaussian is thus an eigenfunction. Indeed, as we shall now show, for the linear flow (27) 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 (see appendix B.1) in one dimension. The linearized Fokker-Planck operator is a Gaussian, so it is natural to consider the set of Hermite polynomials, $H_0(x) = 1$, $H_1(x) = 2x$, $H_2(x) = 4x^2 - 2$, \cdots , as candidates for its eigenfunctions. $H_n(x)$ is an n th-degree polynomial, orthogonal with respect to the Gaussian kernel

$$\frac{1}{2^n n! \sqrt{\pi}} \int dx H_m(x) e^{-x^2} H_n(x) = \delta_{mn}. \quad (33)$$

There are three cases to consider:

$|\Lambda| > 1$ **expanding case:** The form of the left $\tilde{\rho}_0$ eigenfunction (31) suggests that we rescale $x \rightarrow x/\sqrt{2\tilde{Q}}$ and absorb the Gaussian kernel in (33) into left eigenfunctions $\tilde{\rho}_0, \tilde{\rho}_1, \dots$,

$$\tilde{\rho}_k(z) = \frac{1}{\sqrt{2\pi} 2^{3k/2} k! \tilde{Q}^{(k+1)/2}} H_k((2\tilde{Q})^{-1/2} z) e^{-z^2/2\tilde{Q}}, \quad (34)$$

The right eigenfunctions are then

$$\rho_k(z) = (2\tilde{Q})^{k/2} H_k((2\tilde{Q})^{-1/2} z), \quad (35)$$

By construction the left, right eigenfunctions are orthonormal to each other:

$$\int dx \tilde{\rho}_k(x) \rho_j(x) = \delta_{kj}. \quad (36)$$

One can verify [3] that for the fixed point $z = 0$, these are the right, left eigenfunctions of the adjoint Fokker-Planck operator (6), where the k th eigenvalue is $1/|\Lambda|\Lambda^k$. Note that the Floquet multipliers Λ^k are independent of the noise strength, so they are the same as for the $\Delta \rightarrow 0$ deterministic Perron-Frobenius operator (75).

$|\Lambda| = 1$ **marginal case:** This is the 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 (71) to reestablish the control [12]. This we do in sect. 7.

$|\Lambda| < 1$ **contracting 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 Δ . Now what was the left eigenfunction for the expanding case (34) is the peaked right eigenfunction of the Fokker-Planck operator, $\{\rho_0, \rho_1, \rho_2, \dots\}$, with eigenvalues $\{1, \Lambda, \Lambda^2, \dots\}$ [11, 12]

$$\rho_k(x) = N_k^{-1} H_k((2Q)^{-1/2} x) e^{-x^2/2Q}, \quad Q = \Delta/(1 - \Lambda^2), \quad (37)$$

where $H_k(x)$ is the k th Hermite polynomial, and N_k^{-1} follows from the prefactor in (34).

These discrete time results can be straightforwardly generalized to continuous time flows of sect. B, as well as to higher dimensions. So far we have used only the leading eigenfunctions (the natural measure), but in sect. 6 we shall see that knowing the whole spectrum in terms of Hermite polynomial is a powerful tool for the computation of weak-noise corrections.

Remark 4.1 Ornstein-Uhlenbeck process. The simplest example of a continuous time stochastic flow (86) is the Langevin flow (99) in one dimension. In this case, nothing is lost by considering discrete-time dynamics which is strictly equivalent to the continuous time Ornstein-Uhlenbeck process (100) discussed in appendix B.1.

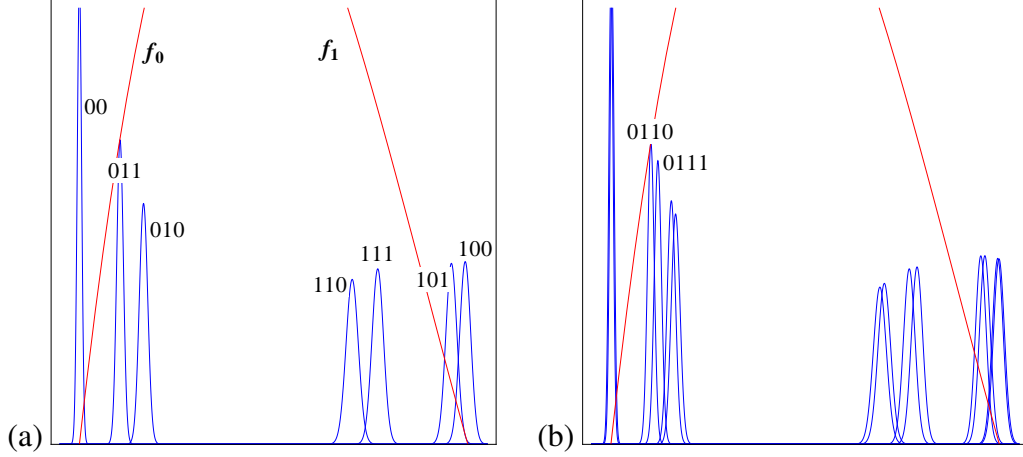


FIGURE 3. (a) f_0, f_1 : branches of the deterministic map (38) for $\Lambda_0 = 8$ and $b = 0.6$. The local eigenfunctions $\tilde{\rho}_{a,0}$ with variances given by (32) provide a state space partitioning by neighborhoods of periodic points of period 3. These are computed for noise variance $\Delta = 0.002$. 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.

5. HAVE NO FEAR OF GLOBALIZATION

We are now finally in position to address our challenge: *Determine the finest possible partition for a given noise.*

We shall explain 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 \tilde{Q}_a of the leading Gaussian eigenfunction of the local adjoint Fokker-Planck operator \mathcal{L}^\dagger . Every periodic point is assigned a one-standard deviation neighborhood $[x_a - \sqrt{\tilde{Q}_a}, x_a + \sqrt{\tilde{Q}_a}]$. 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 $|x_a - x_b| < \sqrt{\tilde{Q}_a} + \sqrt{\tilde{Q}_b}$. As an illustration of the method, consider the chaotic repeller on the unit interval

$$x_{n+1} = \Lambda_0 x_n (1 - x_n) (1 - b x_n) + \xi_n, \quad \Lambda_0 = 8, \quad b = 0.6, \quad (38)$$

with noise strength $\Delta = 0.002$.

The map is plotted in figure 3 (a), together with the local eigenfunctions $\tilde{\rho}_a$ with variances given by (32). 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 of partition intervals: $\{\mathcal{M}_{000}, \mathcal{M}_{001}\}$, $\{\mathcal{M}_{0101}, \mathcal{M}_{0100}\}$ overlap and so do all other 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 neigh-

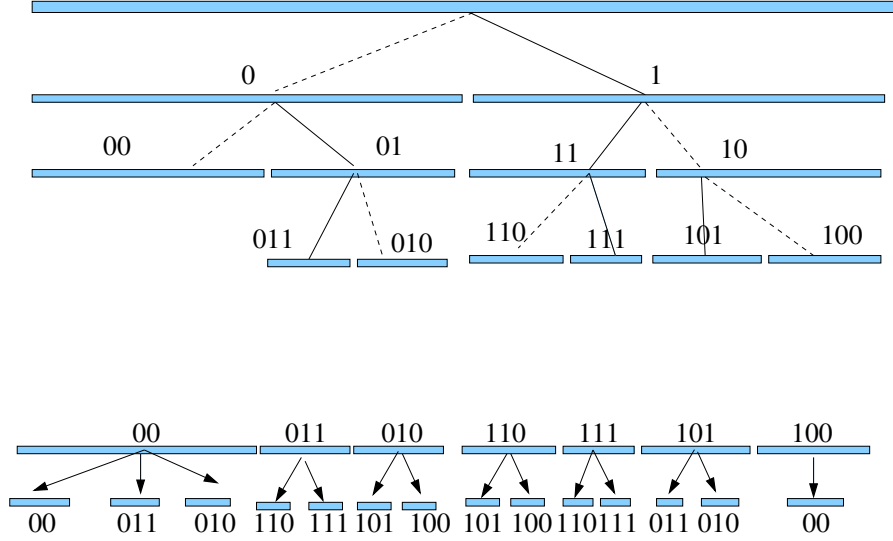


FIGURE 4. (upper panel) The unit interval partitioned deterministically by a binary tree. Due to the noise, the partitioning stops where the eigenfunctions of figure 3 overlap significantly. (lower panel) Once the optimal partition is found, the symbolic dynamics is recoded by relabeling the finite partition intervals, and refashioned into the transition graphs of figure 5.

borhoods

$$\{\mathcal{M}_{00}, \mathcal{M}_{011}, \mathcal{M}_{010}, \mathcal{M}_{110}, \mathcal{M}_{111}, \mathcal{M}_{101}, \mathcal{M}_{100}\}. \quad (39)$$

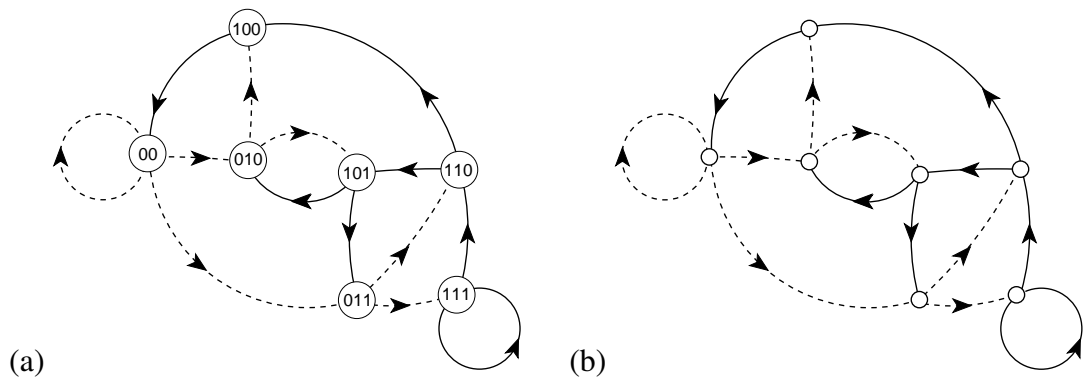
It turns out that resolving \mathcal{M}_{011} further into \mathcal{M}_{0110} and \mathcal{M}_{0111} would not affect our estimates, as it would produce the same transition graph.

Once 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 $\mathcal{M}_{011} \rightarrow \{\mathcal{M}_{110}, \mathcal{M}_{111}\}$, $\mathcal{M}_{00} \rightarrow \{\mathcal{M}_{00}, \mathcal{M}_{011}, \mathcal{M}_{010}\}$, etc.. This is summarized in the transition graph in figure 5, which we will use to estimate the escape rate and the Lyapunov exponent of the repeller.

Remark 5.1 A brief history of state space partitions. 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 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 Shannon 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 do here.

The engineering literature on optimal experimental design [34, 35, 36, 37] 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 [38] 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



Source: ChaosBook.org

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 (39). 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.

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 [39]. Tang and co-workers [40] 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.* [41] 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 [39].

Kennel and Buhl [42, 43, 44] 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. [42] 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 [45] 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 [46] 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 [47, 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. Rechester and White [48, 49] 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.* [50] 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.

In ref. [44] Kennel and Buhl approximate the distribution of the points in each symbolic region by a Gaussian with mean μ and variance τ ,

$$f(x|\mu, \tau) = \frac{1}{(2\pi\tau)^{n/2}} \exp \left[\sum_i^n -\frac{1}{2\tau}(x_i - \mu)^2 \right], \quad (40)$$

and estimate “code length” by the *ad hoc* Rissanen prior on τ , defined with no reference to dynamics, and thus morally unrelated to our periodic orbits based optimal partition.

Dellnitz and Junge [51], Guder and Kreuzer [52], Froyland [53], and Keane *et al.* [54] propose a variety of non-uniform refinement algorithms for such grids, reviewed in a monograph by Froyland [55], who also treats their extension to random dynamical systems. In all cases, the ultimate threshold for every refinement is determined by the convergence of the spectrum of the transition matrix.

Theoretical investigations mostly focus on deterministic limits of stochastic models. The Sinai-Ruelle-Bowen [56, 57, 58] or natural measure (also called equilibrium measure, SRB measure, physical measure, invariant density, natural density, or natural invariant) is singled out amongst all invariant measures by its robustness to weak-noise perturbations, so there is considerable literature that studies it as the deterministic limit of a stochastic process.

6. 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 (32) 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. 4.1, for the linearized flow the entire eigenspectrum is available analytically. For a periodic point $x_a \in p$, the n_p th iterate $\mathcal{L}_a^{n_p}$ of the linearization (9) is the discrete time version of the Ornstein-Uhlenbeck process, with left $\tilde{\rho}_0, \tilde{\rho}_1, \dots$, respectively right ρ_0, ρ_1, \dots mutually orthogonal eigenfunctions (34).

The number of resolved periodic points determines the dimensionality of the Fokker-Planck matrix. Partition (39) being the finest possible partition, the Fokker-Planck oper-

ator now acts as $[7 \times 7]$ matrix with non-zero $a \rightarrow b$ entries expanded in the Hermite basis,

$$\begin{aligned} [\mathbf{L}_{ba}]_{kj} &= \langle \tilde{\rho}_{b,k} | \mathcal{L} | \rho_{a,j} \rangle \\ &= \int \frac{dz_b dz_a \beta}{2^{j+1} j! \pi \sqrt{\Delta/2}} e^{-(\beta z_b)^2 - \frac{(z_b - f'_a(z_a))^2}{2\Delta}} \\ &\quad \times H_k(\beta z_b) H_j(\beta z_a), \end{aligned} \quad (41)$$

where $1/\beta = \sqrt{2Q_a}$, and z_a is the deviation from the periodic point x_a .

Periodic orbit theory (summarized in appendix A) 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. 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 n th 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 $\text{tr} \mathbf{L}^{np}$ by centering \mathbf{L} on the periodic orbit,

$$t_p = \text{tr}_p \mathcal{L}^{np} = \text{tr} \mathbf{L}_{ad} \cdots \mathbf{L}_{cb} \mathbf{L}_{ba}, \quad (42)$$

where $x_a, x_b, \dots, x_d \in p$ are successive periodic points. To leading order in the noise variance Δ , t_p takes the deterministic value $t_p = 1/|\Lambda_p - 1|$. The nonlinear diffusive effects in (41) can be accounted for [19] by the weak-noise Taylor series expansion around the periodic point x_a ,

$$e^{-\frac{(z_b - f_a(z_a))^2}{2\Delta}} = e^{-\frac{(z_b - f'_a z_a)^2}{2\Delta}} \left(1 - \sqrt{2\Delta} (f''_a f'_a z_a^3 + f''_a z_a^2 z_b) + O(\Delta) \right). \quad (43)$$

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 and its loop expansion in figure 6,

$$\begin{aligned} \det(1 - z\mathbf{L}) &= 1 - (t_0 + t_1)z - (t_{01} - t_0 t_1)z^2 \\ &\quad - (t_{001} + t_{011} - t_{01} t_0 - t_{01} t_1)z^3 \\ &\quad - (t_{0011} + t_{0111} - t_{001} t_1 - t_{011} t_0 - t_{011} t_1 + t_{01} t_0 t_1)z^4 \\ &\quad - (t_{00111} - t_{0111} t_0 - t_{0011} t_1 + t_{011} t_0 t_1)z^5 \\ &\quad - (t_{001011} + t_{001101} - t_{0011} t_{01} - t_{001} t_{011})z^6 \\ &\quad - (t_{0010111} + t_{0011101} - t_{001011} t_1 - t_{001101} t_1 \\ &\quad \quad - t_{00111} t_{01} + t_{0011} t_{01} t_1 + t_{001} t_{011} t_1)z^7. \end{aligned} \quad (44)$$

The polynomial coefficients are given by products of non-intersecting loops of the transition graph [24], with the escape rate given by the leading root z_0^{-1} of the polynomial.

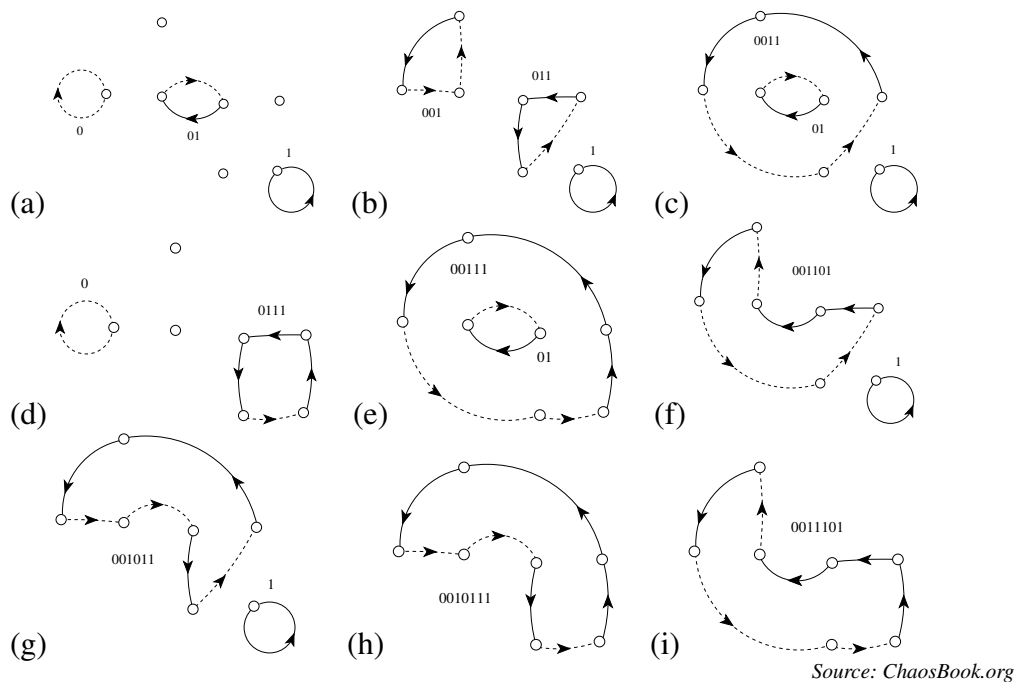


FIGURE 6. (a)-(i) The fundamental cycles for the transition graph figure 5 (b), i.e., the set of its non-self-intersecting loops. Each loop represents a local trace t_p ; together they form the determinant (44).

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 (42) we include stochastic corrections up to order $O(\Delta)$ (an order beyond the term kept in (43)). The escape rate of the repeller of figure 3 so computed is reported in figure 7.

Since our optimal partition algorithm is based on a sharp overlap criterion, small changes in noise strength Δ can lead to transition graphs of different 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 direct 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 7. (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 (42), including stochastic corrections up to order $O(\Delta)$. Figure 7 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}

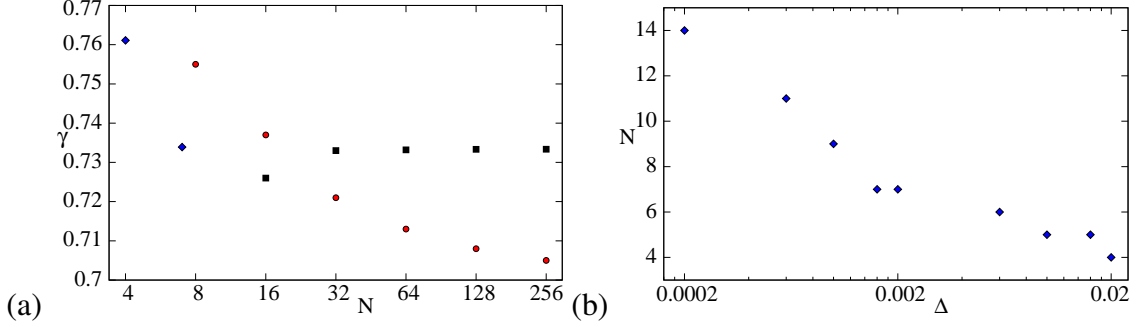


FIGURE 7. (a) The escape rate γ of the repeller in figure 3 plotted as function of number of partition intervals N , estimated using: (◆) under-resolved 4-interval and the 7-interval optimal partition, (●) all periodic orbits of periods up to $n = 8$ in the deterministic, binary symbolic dynamics, with $N_i = 2^n$ periodic-point intervals (the deterministic, noiseless escape rate is $\gamma_{<>} = 0.7011$), and (■) a uniform discretization (45) in $N = 16, \dots, 256$ intervals. For $N = 512$ discretization yields $\gamma_{\text{num}} = 0.73335(4)$. (b) Number of neighborhoods required by the optimal partition method vs. the noise strength Δ .

by a piecewise-constant approximation on a uniform mesh on the unit interval [47],

$$[\mathcal{L}]_{ij} = \frac{1}{|\mathcal{M}_i|} \frac{1}{\sqrt{2\pi\Delta}} \int_{\mathcal{M}_i} dx \int_{f^{-1}(\mathcal{M}_j)} dy e^{-\frac{1}{2\Delta}(y-f(x))^2}, \quad (45)$$

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

We estimate the escape rate of the repeller (38) for a range of values of the noise strength Δ . The optimal partition method requires a different numbers of neighborhoods every time for different noise strengths. The results are illustrated by figure 7 (b) and 8, 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 (41) are consistent within less than 2%, meaning that the stochastic corrections (43) do not make a significant difference, compared to the effect of the optimal choice of the partition, and need not be taken into account in this example.

The optimal partition estimate of the Lyapunov exponent is given by $\lambda = \langle \ln |\Lambda| \rangle / \langle n \rangle$, where the cycle expansion average of an integrated observable A [24]

$$\begin{aligned} \langle A \rangle &= A_0 t_0 + A_1 t_1 + [A_{01} t_{01} - (A_0 + A_1) t_0 t_1] \\ &\quad + [A_{001} t_{001} - (A_{01} + A_0) t_{01} t_0] + \dots \\ &\quad + [A_{011} t_{011} - (A_{01} + A_1) t_{01} t_1] + \dots \end{aligned} \quad (46)$$

is the finite sum over cycles contributing to (44), 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 (45) to cross check our estimate: this way the Lyapunov exponent is evaluated as the average

$$\lambda = \int dx e^{\gamma} \rho(x) \ln |f'(x)|, \quad (47)$$

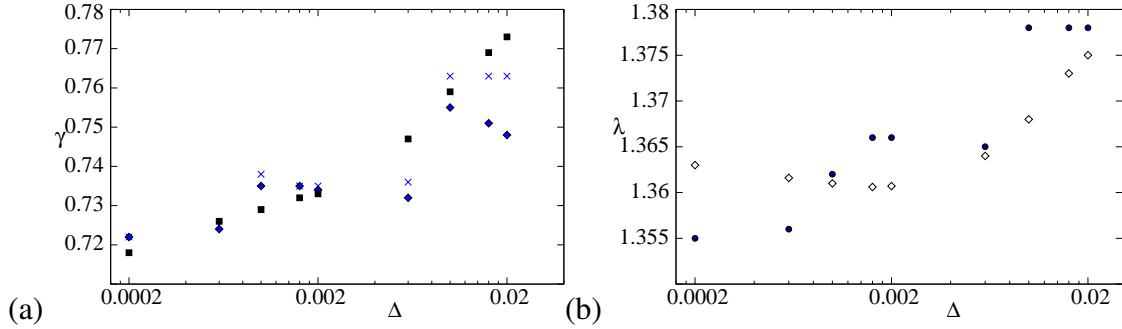


FIGURE 8. (a) Escape rates of the repeller (38) vs. the noise strength Δ , using: the optimal partition method with (\blacklozenge) and without (\times) stochastic corrections; (\blacksquare) a uniform discretization (45) in $N = 128$ intervals. (b) The Lyapunov exponent of the repeller (38) vs. the noise strength Δ , using: the optimal partition method (\bullet) without stochastic corrections, and (\diamond) a uniform discretization (45) over $N = 128$ intervals.

where $\rho(x)$ is the leading eigenfunction of (45), 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 (46), where $t_p = 1/|\Lambda_p - 1|$ (no higher-order stochastic corrections), and the same quantity evaluated with (47), by the discretization method (45).

Remark 6.1 Weak noise corrections. The weak-noise corrections to the spectrum of evolution operators were first treated by Gaspard [4] for continuous-time systems, and in a triptych of articles [19, 20, 21] for discrete-time maps: they can be computed perturbatively to a remarkably high order [59] in the noise strength Δ . However, as we have shown here, the *eigenvalues* of such operators offer no guidance to the ‘optimal partition’ problem; one needs to compute the *eigenfunctions*.

7. 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 deterministically either disconnected or transient (“noise induced escape,” “noise induced chaos”) and second, it smoothens 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 relatively simple example, consider the skew Ulam map [60], i.e., the cubic map (38) with the parameter $\Lambda_0 = 1/f(x_c)$. 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 $(x_b - x)^{-1/2}$ near the critical value $f(x_c) = x_b$ (see, for example, ref. [61] for a discussion). As explained in ref. [60], a close passage to the critical point effectively replaces the accumulated Floquet multiplier by its square root. For example, for the skew Ulam map (38) n -cycles whose itineraries are of form $0^{n-1}1$ spend long time in the neighborhood of $x_0 = 0$, and then 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 $\Lambda_{0\dots 01} \propto \Lambda_0^{n/2}$, 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 naïvely expect [62, 63, 16].

For the corresponding noisy map (1) the critical point is extended into the ‘flat top’ region where $|f'(x)| \ll 1$, and the linearized, Gaussian approximation (9) 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 acts on a Gaussian density centered at x_a , as in (10):

$$\begin{aligned} [\mathcal{L}^\dagger \tilde{\rho}_a](x) &= \frac{1}{C_a} \int_{-\infty}^{\infty} e^{-\frac{(f(x)-y)^2}{2\Delta}} e^{-(y-x_a)^2/2Q_a} [dy] \\ &= \frac{1}{C_{a-1}} e^{-\frac{(f(x)-x_a)^2}{2(Q_a+\Delta)}}. \end{aligned} \quad (48)$$

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

$$\rho_{a-1}(x) = \frac{1}{C_{a-1}} e^{-\frac{(f(x)-x_a)^2}{2(Q_a+\Delta)}} = \frac{1}{C_{a-1}} e^{-f_{a-1}'' z_{a-1}^4 / 8(Q_a+\Delta)}. \quad (49)$$

During a close passage to the critical point, the variance does not transform linearly, but as a square root:

$$Q_{a-1} = \frac{\int z^2 e^{-f_{a-1}'' z^4 / 8(Q_a+\Delta)} dz}{\int e^{-f_{a-1}'' z^4 / 8(Q_a+\Delta)} dz} = \frac{\Gamma(3/4)}{\Gamma(1/4)} \left(\frac{8(Q_a+\Delta)}{f_{a-1}''^2} \right)^{1/2}. \quad (50)$$

We show in appendix D 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 $O(\Delta)$ 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 $\rho_{a-2}(z_{a-2})$ (Eq. (115), appendix D) up to $O(\Delta)$ and obtain the same result for the variance, that is

$$Q_{a-3} = \frac{Q_{a-1} + \Delta(1 + f_{a-2}'^2)}{f_{a-2}'^2 f_{a-3}'^2} \quad (51)$$

which is again the evolution of the variances in the Gaussian approximation. In other words, the evolution of the variances goes back to be linear, to $O(\Delta)$, although the densities transformed from the ‘quartic Gaussian’ (49) are no longer Gaussians.

The question is now how to modify the definition of neighborhoods given in sect. 5, 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 (49) and (115) involved. Since we only care about the variances, we define instead the following map

$$Q_{a-1} = \begin{cases} C \left(\frac{Q_a + \Delta}{f_{a-1}'^2} \right)^{1/2} & |f_{a-1}'^2| < 1 \\ \frac{Q_a + \Delta}{f_{a-1}'^2} & \text{otherwise} \end{cases}, \quad (52)$$

$C = 2\sqrt{2}\Gamma(3/4)/\Gamma(1/4)$, 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 (52):

$$\tilde{Q}_a \simeq C \left(\frac{\Delta \left(1 + f_{a-1}'^2 + \dots + (f_{a-n+1}'^2)^2 \right)^{1/2}}{\tilde{\Lambda}_p} \right)^{1/2} \quad (53)$$

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

$$\tilde{Q}_{a-k} \simeq \frac{1}{(f_{a-1}^{k'})^2} \left(\Delta (1 + f_{a-1}'^2 + \dots + (f_{a-1}^{k-1'})^2) + \tilde{Q}_a \right) \quad (54)$$

both expressions (53) and (54) are approximate, as we further assumed $\Delta \tilde{\Lambda}_p^2 \gg 1$, which is reasonable when $\Delta \in [10^{-4}, 10^{-2}]$, our range of investigation for the noise strength. As before, a neighborhood of width $[x_a - \sqrt{\tilde{Q}_a}, x_a + \sqrt{\tilde{Q}_a}]$ is assigned to each periodic point x_a , and an optimal partition follows. However, due to the geometry of the map, such partitions as

$$\{\mathcal{M}_{000}, [\mathcal{M}_{001}, \mathcal{M}_{011}], \mathcal{M}_{010}, \mathcal{M}_{110}, \mathcal{M}_{111}, \mathcal{M}_{10}\} \quad (55)$$

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 the deterministic symbolic dynamics and ignore the overlap.

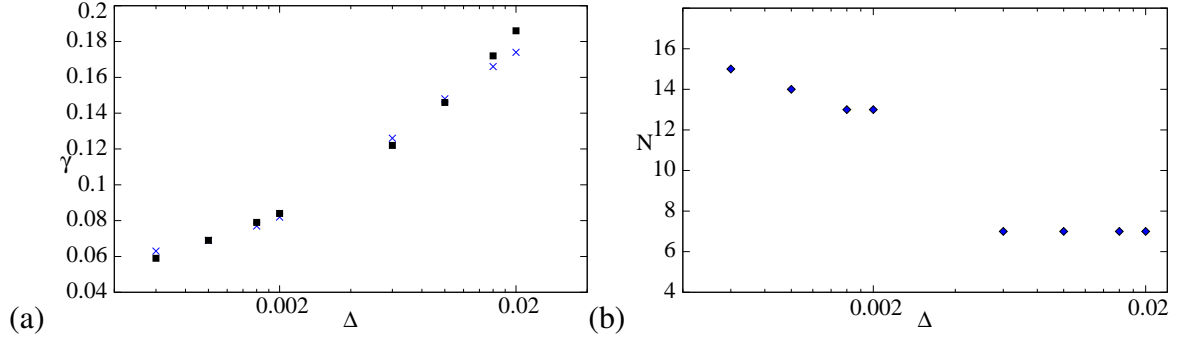


FIGURE 9. (a) Escape rate γ of the ‘skew Ulam’ map vs. noise strength Δ , using: (×) the optimal partition method; (■) a uniform discretization (45) in $N = 128$ intervals. (b) Number of neighborhoods required by the optimal partition method vs. the noise strength Δ .

Let us now test the method by estimating once again the escape rate of the noisy map (38). We note that the matrix elements

$$\begin{aligned}
 [\mathbf{L}_{ba}]_{kj} &= \langle \tilde{\rho}_{b,k} | \mathcal{L} | \rho_{a,j} \rangle \\
 &= \int \frac{dz_b dz_a \beta}{2^j j! \pi \sqrt{2\Delta}} e^{-(\beta z_b)^2 - \frac{(z_b - f'_a z_a)^2}{2\Delta}} \\
 &\quad \times H_k(\beta z_b) H_j(\beta z_a),
 \end{aligned} \tag{56}$$

$1/\beta = \sqrt{2Q_a}$, 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 (49):

$$[\mathbf{L}_{ba}]_{kj} = \int \frac{dz_b dz_a \beta}{2^j j! \pi \sqrt{2\Delta}} e^{-(\beta z_b)^2 - \frac{(z_b - f'_a z_a - f''_a \sqrt{2\Delta} z_a / 2)^2}{2\Delta}} H_k(\beta z_b) H_j(\beta z_a), \tag{57}$$

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

$$[\mathbf{L}_{ba}]_{kj} = \int \frac{dz_b dz_a \beta}{2^j j! \pi \sqrt{2\Delta}} e^{-(\beta z_b)^2 - \frac{(z_b - f'_a z_a - f''_a \sqrt{2\Delta} z_a / 2)^2}{2\Delta}} H_k(\beta z_b) H_j(\beta z_a), \tag{58}$$

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 (57) better be replaced with (58) when $\Delta \sim 10^{-4}$. 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 (56) is the zeroth-order term of a series in Δ , whose convergence can be probed by evaluating the higher order corrections (43): when the $O(\sqrt{\Delta})$ and $O(\Delta)$ corrections are of an order of magnitude comparable or bigger than the one of (56), we conclude that the Gaussian approximation fails and we use (57) or (58) instead. Everywhere else we use our usual matrix elements (56), *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. Like before, we tweak the noise strength Δ within the range $[10^{-4}, 10^{-2}]$ and compare the escape rate evaluated with the optimal partition method and with the uniform discretization (45). The results are illustrated in figure 9: the uniform discretization and the method of the optimal partition are consistent within a 5% margin.

8. SUMMARY AND CONCLUSIONS

Physicists tend to believe that with time Brownian motion leads to $x(t)^2 \approx \Delta t$ broadening of a noisy trajectory neighborhood. In nonlinear dynamics nothing of the sort happens; the noise broadening is balanced by non-linear stretching and contraction, and infinite length recurrent Langevin trajectories have finite noise widths, not widths that spread \propto time. Here periodic orbits play special role: computable in finite time, they persist for infinite time, and are thus natural objects to organize state space partitions around. On the other hand, computation of unstable periodic orbits in high-dimensional state spaces, such as Navier-Stokes, is at the border of what is currently feasible numerically [64, 65], 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 deterministic 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 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 method for partitioning the state space of a chaotic repeller in presence of weak Gaussian noise first introduced in ref. [7]. 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. [19, 20, 21], with the stochastic trace formulas and determinants [19, 4] expressed as finite sums, truncated at orbit periods corresponding to the local turnover times. A novel aspect of the work presented here is its representation in terms of the Hermite basis (sect. B.1), eigenfunctions of the linearized Fokker-Planck operator (9), and the finite dimensional matrix representation of the Fokker-Planck operator.

It should be noted that our linearization of Fokker-Planck operators does not imply that the nonlinear dynamics is being modeled by a linear one. Our description is fully nonlinear, with periodic orbits providing the nonlinear backbone of chaotic dynamics, dressed up stochastically by Fokker-Planck operators local to each cycle. This is a stochastic cousin of Gutzwiller's WKB approximation based semi-classical quanti-

zation [66] of classically chaotic systems, where in a parallel effort to utilize quantum-mechanical \hbar ‘graininess’ of the quantum phase space to terminate periodic orbit sums, Berry and Keating [67] have proposed inclusion of cycles of periods up to a single ‘Heisenberg time’. In light of the stochastic dynamics insights gained here, this proposal merits a reexamination - each neighborhood is likely to have its own Heisenberg time.

The work of Abarbanel *et al.* [68, 69, 28, 70] 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 (14) is a step in the Kalman filter procedure. One could combine the state space charts of turbulent flows of ref. [71] (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 a $\approx 10^5$ -dimensional ODE representation, and then track the experimental 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.’

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. In the $1d$ setting numerical tests indicate that the ‘optimal partition’ method can be as accurate as the much finer grained direct numerical Fokker-Planck operator calculations. In higher dimensions (and especially in the extreme high dimensions required by fluid dynamics stimulations) such direct Fokker-Planck PDE integrations are not feasible, while the method proposed here is currently the only implementable approach.

The success of the optimal partition hypothesis in a one-dimensional setting is encouraging, and use of noise as a smoothing device that eliminates singularities and pathologies from clusterings of orbits is promising. 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; the method has not yet been tested in a high-dimensional hyperbolic setting. A limiting factor to applications of the periodic orbit theory to high-dimensional problems ranging from fluid flows to chemical reactions might be 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.

In summary: Each periodic point owns a cigar, which for a high-dimensional dissipative flow is shaped along a handful of expanding and least-contracting directions. The remaining large (even infinite!) number of the strongly contracting directions is limited by the noise; the cigar always has the dimensionality of the full state space. Taken together, the set of overlapping cigars, or the optimal partition, weaves the carpet (of the full dimensionality of state space) which envelops the entire ‘inertial manifold’ explored by turbulent dynamics.

Acknowledgments.

We are grateful to D. Barkley, T. Bartsch, C.P. Dettmann, H. Fogedby, A. Grigo, A. Jackson, R. Mainieri, W.H. Mather, R. Metzler, E. Ott, S.A. Solla, N. Sørndergaard, and G. Vattay for many stimulating discussions. Special thanks go to M. Robnik who made it possible to present this work at the 8th Maribor Summer School ‘Let’s Face Chaos through Nonlinear Dynamics’, which brought us together with T. Prosen, from whom we have learned that fixed points of the covariance evolution equations are the Lyapunov equations. P.C. thanks Glen P. Robinson, Jr. for support. D.L. was supported by NSF grant DMS-0807574 and G.P. Robinson, Jr..

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. [24] for further details.

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 10. Consider dynamics induced by iterations of a d -dimensional map $f : \mathcal{M} \rightarrow \mathcal{M}$, 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 k th iterate of map f by composition

$$f^k(x) = f\left(f^{k-1}(x)\right), \quad f^0(x) = x. \quad (59)$$

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

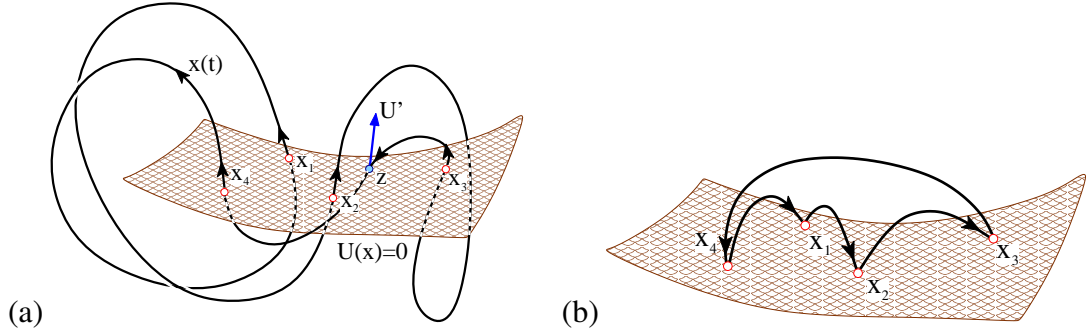
$$\{x_0, x_1, x_2, \dots, x_k\} = \left\{x, f(x), f^2(x), \dots, f^k(x)\right\}, \quad (60)$$

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

$$M_{ij}^k(x_0) = \left. \frac{\partial f_i^k(x)}{\partial x_j} \right|_{x=x_0}. \quad (61)$$

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

$$M^k(x_0) = M(x_{k-1}) \cdots M(x_1)M(x_0), \quad M_{ij} = \partial f_i / \partial x_j, \quad (62)$$



Source: ChaosBook.org

FIGURE 10. (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 is 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) .

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

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

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

$$\Lambda_\ell = e^{k\lambda^{(\ell)}} = e^{k(\mu^{(\ell)} + i\omega^{(\ell)})}. \quad (63)$$

Jacobian matrix $M^k(x_0)$ and its eigenvalues (Floquet multipliers) depend on the initial point x_0 and the elapsed time k . For notational brevity we tend to omit this dependence, but in general

$$\Lambda = \Lambda_\ell = \Lambda_\ell(x_0, k), \quad \lambda = \lambda^{(\ell)}(x_0, k), \quad \omega = \omega^{(\ell)}(x_0, k), \dots \text{ etc.},$$

depend on the traversed trajectory.

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

$$f^n(x_k) = f(f(\dots f(x_k) \dots)) = x_k, \quad k = 0, 1, 2, \dots, n-1. \quad (64)$$

For example, the orbit of x_1 in figure 10 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 $\{\Lambda_{p,1}, \Lambda_{p,2}, \dots, \Lambda_{p,d}\}$ 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 $|\Lambda_1| \geq |\Lambda_2| \geq \dots \geq |\Lambda_d|$, sort them into sets $\{e, m, c\}$

$$\begin{aligned} \text{expanding:} \quad \{\Lambda\}_e &= \{\Lambda_{p,j} : |\Lambda_{p,j}| > 1\} \\ \text{marginal:} \quad \{\Lambda\}_m &= \{\Lambda_{p,j} : |\Lambda_{p,j}| = 1\} \\ \text{contracting:} \quad \{\Lambda\}_c &= \{\Lambda_{p,j} : |\Lambda_{p,j}| < 1\}, \end{aligned} \quad (65)$$

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

$$\Lambda_p = \prod_e \Lambda_{p,e}. \quad (66)$$

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

$$\mu_p^{(i)} = \frac{1}{n_p} \ln |\Lambda_{p,i}|. \quad (67)$$

They can be loosely 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*.

The Jacobian matrix M is in general non-normal: it neither symmetric, nor diagonalizable by a rotation, nor do its (left or right) eigenvectors define an orthonormal coordinate frame (for brevity we omit in what follows the time superscript, $M^k \rightarrow M$). As any matrix with real elements, M can be expressed in the singular value decomposition form

$$M = UDV^T, \quad (68)$$

where D is diagonal and real, and U, V are orthogonal matrices. The diagonal elements $\sigma_1, \sigma_2, \dots, \sigma_d$ of D are called the *singular values* of M , namely the square root of the eigenvalues of $M^T M = VD^2V^T$ (or $MM^T = UD^2U^T$), which is a symmetric, positive semi-definite matrix (and thus admits only real, non-negative eigenvalues).

Singular values $\{\sigma_j\}$ are *not related* to the M eigenvalues $\{\Lambda_j\}$ in any simple way. From a geometric point of view, when all singular values are non-zero, M maps the unit sphere into an ellipsoid, figure 11 (b): the singular values are then the lengths of the semiaxes of this ellipsoid. Note however that the singular vectors of $M^T M$ that determine the orientation of the semiaxes are distinct from the M eigenvectors $\{\mathbf{e}^{(j)}\}$, and that $M^T M$ satisfies no semigroup property along the flow. For this reason the M eigenvectors $\{\mathbf{e}^{(j)}\}$ are sometimes called ‘covariant’ or ‘covariant Lyapunov vectors’, in

order to emphasize the distinction between them and the singular value decomposition semiaxes directions.

Eigenvectors / eigenvalues are suited to study of iterated forms of a matrix, such as M^k or exponentials $\exp(tA)$, and are thus a natural tool for study of dynamics. Singular vectors are not. They are suited to study of M itself, and the singular value decomposition is convenient for numerical work (any matrix, square or rectangular, can be brought to this form), as a way of estimating the effective rank of matrix M by neglecting the small singular values.

A.1. 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 (59), its derivative, and, by the chain rule, the derivative (62) of the k th iterate f^k evaluated at the point x_a ,

$$\begin{aligned} x &= x_a + z_a, & f_a(z_a) &= f(x_a + z_a) \\ z_{a+1} &= M_a z_a + \dots & & \\ M_a &= f'(x_a), & M_a^k &= M_{a+k-1} \dots M_{a+1} M_a, \quad k \geq 2. \end{aligned} \quad (69)$$

The monodromy (or Floquet) matrix,

$$M_{p,a} = M^{n_p}(x_a), \quad (70)$$

evaluated on the periodic point x_a is position dependent, but its eigenvalues, the Floquet multipliers $\{\Lambda_{p,1}, \Lambda_{p,2}, \dots, \Lambda_{p,d}\}$ 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

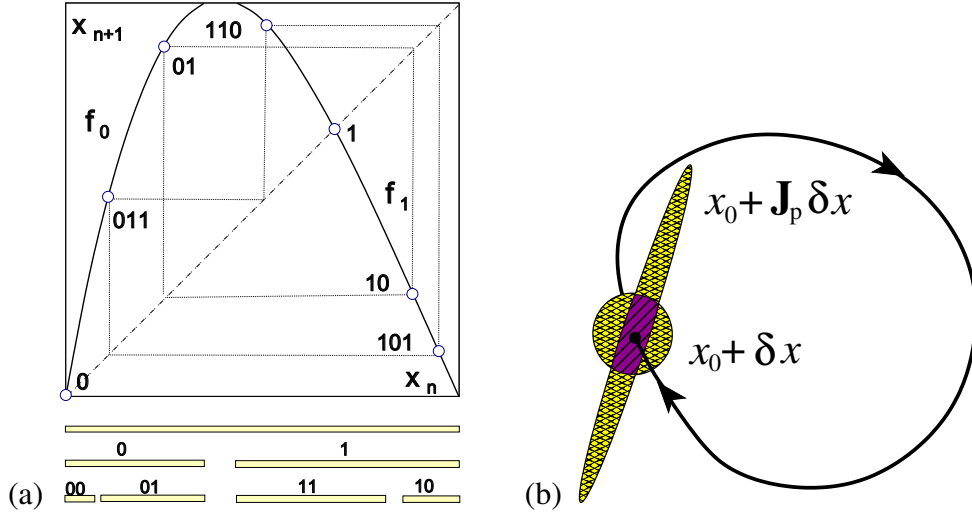
$$f(x) = x_{a+1} + f'_a z_a + \frac{1}{2} f''_a z_a^2 + \dots, \quad (71)$$

where

$$\begin{aligned} f'_a &= f'(x_a), & f''_a &= f''(x_a) \\ f_a^{k'} &= f'_{a+k-1} \dots f'_{a+1} f'_a, & k &\geq 2. \end{aligned} \quad (72)$$

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

We label trajectory points by either x_n , $n = 1, 2, \dots$, 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 11 (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 $01\bar{1} = (x_{011}, x_{110}, x_{101})$ is traversed in 3 iterations.



Source: ChaosBook.org

FIGURE 11. (a) Periodic points of the map $f(x) = 6x(1-x)(1-0.6x)$ 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 eigen-directions $\mathbf{e}^{(i)}$ of M_p , the monodromy matrix (70).

A.2. Periodic orbit theory

Since its initial formulations by Ruelle [58] and Gutzwiller [66], 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. [24, 72]), 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

$$\langle a \rangle = \lim_{n \rightarrow \infty} \frac{1}{n} \langle A^n \rangle, \quad A^n(x_0) = \sum_{k=0}^{n-1} a(x_k). \quad (73)$$

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 $t \rightarrow \infty$. Hence the theory is recast in the language of *linear* evolution operators (Liouville, Perron-Frobenius, Ruelle-Araki, ...)

$$\rho(y, t) = [\mathcal{L}_{det}^t \rho](y) = \int_{\mathcal{M}} dx \delta(y - f^t(x)) \rho(x, 0), \quad (74)$$

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 (74)

$$\mathcal{L}_{det}^t(x, y) = \delta(x - f^t(y)) \quad (75)$$

as the *Perron-Frobenius operator*, with the subscript *det* indicating that this is deterministic, to distinguish it from the noisy Fokker-Planck operator (3).

The *Koopman operator* action on a state space function $a(x)$ is to replace it by its downstream value time t later, $a(x) \rightarrow a(x(t))$ evaluated at the trajectory point $x(t)$:

$$\begin{aligned} [\mathcal{L}_{det}^{t\dagger} a](x) &= a(f^t(x)) = \int_{\mathcal{M}} dy \mathcal{L}_{det}^{t\dagger}(x,y) a(y) \\ \mathcal{L}_{det}^{t\dagger}(x,y) &= \delta(y - f^t(x)). \end{aligned} \quad (76)$$

Given an initial density of representative points $\rho(x)$, the average value of $a(x)$ evolves as

$$\begin{aligned} \langle a \rangle(t) &= \frac{1}{|\rho_{\mathcal{M}}|} \int_{\mathcal{M}} dx a(f^t(x)) \rho(x) = \frac{1}{|\rho_{\mathcal{M}}|} \int_{\mathcal{M}} dx [\mathcal{L}_{det}^{t\dagger} a](x) \rho(x) \\ &= \frac{1}{|\rho_{\mathcal{M}}|} \int_{\mathcal{M}} dx dy a(y) \delta(y - f^t(x)) \rho(x). \end{aligned}$$

The ‘propagator’ $\delta(y - f^t(x))$ can equally well be interpreted as belonging to the Perron-Frobenius operator (75), so the two operators are adjoint to each other,

$$\int_{\mathcal{M}} dx [\mathcal{L}_{det}^{t\dagger} a](x) \rho(x) = \int_{\mathcal{M}} dy a(y) [\mathcal{L}_{det}^t \rho](y). \quad (77)$$

This suggests an alternative point of view, which is to push dynamical effects into the density. In contrast to the Koopman operator which advances the trajectory by time t , the Perron-Frobenius operator depends on the trajectory point time t in the past. Koopman operators are so cool, that it is no wonder that Igor Mezić is so enamored with Koopmania [73, 74].

The Perron-Frobenius operators are non-normal, not self-adjoint operators, so their left and right eigenvectors differ. The right eigenvectors of a Perron-Frobenius operator are the left eigenvectors of the Koopman, and vice versa. While one might think of a Koopman operator as an ‘inverse’ of the Perron-Frobenius operator, the notion of ‘adjoint’ is the right one, especially in settings where flow is not time-reversible, as is the case for infinite dimensional flows contracting forward in time and for stochastic flows.

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

$$\mathcal{L}_{det} \circ \rho(x) = \int dy \delta(x - \Lambda y) \rho(y) = \frac{1}{|\Lambda|} \rho\left(\frac{x}{\Lambda}\right). \quad (78)$$

In the $|\Lambda| > 1$ expanding case the right, expanding deterministic eigenfunctions [12, 75] are monomials

$$\rho_k(x) \rightarrow x^k / k!, \quad k = 0, 1, 2, \dots, \quad (79)$$

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

$$\rho_k(x) \rightarrow (-1)^k \delta^{(k)}(x). \quad (80)$$

In discretizations $\mathcal{L}_{det}^t(y, x)$ 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 $\Lambda_1 = -\Lambda_0/(\Lambda_0 - 1) < -1$:

$$f(x) = \begin{cases} f_0(x) = \Lambda_0 x, & x \in \mathcal{M}_0 = [0, 1/\Lambda_0] \\ f_1(x) = \Lambda_1(1-x), & x \in \mathcal{M}_1 = (1/\Lambda_0, 1]. \end{cases} \quad (81)$$

As in figure 11 (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

$$\rho(x) = \begin{cases} \rho_0 & \text{if } x \in \mathcal{M}_0 \\ \rho_1 & \text{if } x \in \mathcal{M}_1 \end{cases}, \quad (82)$$

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

$$\begin{pmatrix} \rho_0 \\ \rho_1 \end{pmatrix} \rightarrow \mathbf{L}\rho = \begin{pmatrix} \frac{1}{|\Lambda_0|} & \frac{1}{|\Lambda_1|} \\ \frac{1}{|\Lambda_0|} & \frac{1}{|\Lambda_1|} \end{pmatrix} \begin{pmatrix} \rho_0 \\ \rho_1 \end{pmatrix}, \quad (83)$$

stretching both ρ_0 and ρ_1 over the whole unit interval Λ . Ulam [47, 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 (73) from the leading eigenvalue $z_0 = z(\beta)$ of an evolution operator by means of the *classical trace formula* [76, 24], which for map f , takes form

$$\langle a \rangle = \left. \frac{\partial z_0}{\partial \beta} \right|_{\beta=0}, \quad \sum_{\alpha=0}^{\infty} \frac{1}{z - z_\alpha} = \sum_p n_p \sum_{r=1}^{\infty} \frac{z^{rn_p} e^{r\beta \cdot A_p}}{|\det(\mathbf{1} - M_p^r)|}. \quad (84)$$

or, even better, by deploying the associated spectral determinant

$$\det(1 - z\mathcal{L}_{det}) = \exp \left(- \sum_p \sum_{r=1}^{\infty} \frac{1}{r} \frac{z^{rn_p} e^{r\beta \cdot A_p}}{|\det(\mathbf{1} - M_p^r)|} \right). \quad (85)$$

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 (62), evaluated in the periodic orbit p , the deterministic exponential contraction/expansion is characterized by its Floquet multipliers $\{\Lambda_{p,1}, \dots, \Lambda_{p,d}\}$, and p contribution to (84) 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 [9, 77], with cycles ordered hierarchically; short cycles give dominant contributions to (84), 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 [31]. 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* [9, 77, 72, 24].

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 (84).

B. FOKKER-PLANCK OPERATOR, CONTINUOUS TIME FORMULATION

The material reviewed in this appendix is standard [1, 3, 14], 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.

Consider a d -dimensional stochastic flow

$$\frac{dx}{dt} = v(x) + \hat{\xi}(t), \quad (86)$$

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 $\delta t \rightarrow 0$. This holds if the noise is Brownian, i.e., the probability that the trajectory reaches x_{n+1} is given by a normalized Gaussian

$$\mathcal{L}^{\delta t}(x_{n+1}, x_n) = \frac{1}{N} \exp \left[-\frac{1}{2\delta t} (\xi_n^T \frac{1}{\Delta} \xi_n) \right]. \quad (87)$$

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 $\{x_n \rightarrow x_{n+1}, x_n \rightarrow f^{\delta t}(x_n)\}$ (discrete time formulation),

$$\delta x_n = x_{n+1} - x_n \simeq \dot{x}_n \delta t, \quad f^{\delta t}(x_n) - x_n \simeq v(x_n) \delta t, \quad (88)$$

where

$$\{x_0, x_1, \dots, x_n, \dots, x_k\} = \{x(0), x(\delta t), \dots, x(n\delta t), \dots, x(t)\} \quad (89)$$

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)

$$\langle \xi_j(t_n) \rangle = 0, \quad \langle \xi_i(t_m) \xi_j^T(t_n) \rangle = \Delta_{ij} \delta_{nm}, \quad (90)$$

where $\langle \dots \rangle$ 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,

$$\mathcal{L}^t(x, x_0) = \frac{1}{\sqrt{2\pi\Delta t}} \exp\left[-\frac{(x - x_0)^2}{2\Delta t}\right], \quad (91)$$

of mean 0, variance Δt , and standard deviation $\sqrt{\Delta t}$, uncorrelated in time:

$$\langle x_{n+1} - x_n \rangle = 0, \quad \langle (x_{m+1} - x_m)(x_{n+1} - x_n) \rangle = \Delta \delta_{mn}. \quad (92)$$

$\mathcal{L}^t(x, x_0)$ describes the diffusion at any time, including the integer time increments $\{t_n\} = \{\delta t, 2\delta t, \dots, n\delta t, \dots\}$, and thus provides a bridge between the continuous and discrete time formulations of noisy evolution. We have set $\delta t = 1$ in (92) anticipating the discrete time formulation of sect. 2.

In physical problems the diffusion tensor Δ is almost always anisotropic: for example, the original Langevin flow [78] is a continuous time flow in {configuration, velocity} phase space, with white noise probability distribution $\exp(-\mathbf{x}^2/2k_B T)$ modeling random Brownian force kicks applied only to the velocity variables \mathbf{x} . In this case one thinks of diffusion coefficient $D = k_B T/2$ as temperature. For sake of simplicity we shall sometimes assume that diffusion in d dimensions is uniform and isotropic, $\Delta(x) = 2D\mathbf{1}$. The more general case of a tensor Δ which is a state space position dependent but time independent can be treated along the same lines, as we do in (14). In this case the stochastic flow (86) is written as [79] $dx = v(x) dt + \sigma(x) d\hat{\xi}(t)$, $\sigma(x)$ is called ‘diffusion matrix,’ and the noise is referred to as ‘multiplicative.’

The distribution (87) 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 (87)

$$\mathcal{L}^t(x_k, x_0) = \int [dx] \exp\left\{-\frac{1}{2\Delta\delta t} \sum_{n=1}^{k-1} [x_{n+1} - f^{\delta t}(x_n)]^2\right\}, \quad (93)$$

where $t = k \delta t$, and the Gaussian normalization factor in (87) is absorbed into intermediate integrations by defining as

$$\begin{aligned} [dx] &= N^{-1} \prod_{n=1}^{k-1} dx_n^d \\ N &= (2\pi\delta t)^{d/2} (\det \Delta)^{1/2} \quad \text{anisotropic diffusion tensor } \Delta \\ &= (2\pi\Delta\delta t)^{d/2} \quad \text{isotropic diffusion,} \end{aligned} \quad (94)$$

The stochastic flow (86) can now be understood as the continuous time, $\delta t \rightarrow 0$ limit, with the velocity noise $\hat{\xi}(t)$ a Gaussian random variable of zero mean and covariance matrix

$$\langle \hat{\xi}_j(t) \rangle = 0, \quad \langle \hat{\xi}_i(t) \hat{\xi}_j(t') \rangle = \Delta_{ij} \delta(t - t'). \quad (95)$$

It is worth noting that the continuous time flow noise $\hat{\xi}(t)$ in (86) and (95) is dimensionally a velocity $[x]/[t]$, while the discrete time noise ξ_n in (87), (90) is dimensionally a length $[x]$. The continuous time limit of (93), $\delta t = t/k \rightarrow 0$, defines formally the Fokker-Planck operator

$$\mathcal{L}^t(x, x_0) = \int [dx] \exp \left\{ -\frac{1}{2\Delta} \int_0^t [\dot{x}(\tau) - v(x(\tau))]^2 d\tau \right\} \quad (96)$$

as a stochastic path (or Wiener) integral [63, 80, 3] for a noisy flow, and the associated continuous time Fokker-Planck (or forward Kolmogorov) equation [1, 3, 81] describes the time evolution of a density of noisy trajectories (86),

$$\partial_t \rho(x, t) + \nabla \cdot (v(x)\rho(x, t)) = D \nabla^2 \rho(x, t). \quad (97)$$

The $\delta t \rightarrow 0$ limit and the proper definition of $\dot{x}(\tau)$ are delicate issues [82, 83, 14, 84] of no import for the applications of stochasticity studied here.

In probabilist literature [85] 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 (96). The exponent

$$-\frac{1}{2\Delta\delta t} [x_{n+1} - f^{\delta t}(x_n)]^2 \simeq -\frac{1}{2\Delta} [\dot{x}(\tau) - v(x(\tau))]^2 \delta t \quad (98)$$

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 [86, 87], with velocity $\dot{x}(\tau)$ along a trial path varied with aim of minimizing its distance to the target $v(x(\tau))$.

The finite time step formulation (93) of the Fokker-Planck operator motivates the exposition of sect. 2, which starts by setting $\delta t = 1$. In the linearized setting, the two formulations are fully equivalent.

Remark B.1 A brief history of noise. The cost function (98) appears to have been first introduced by Wiener as the exact solution for a purely diffusive Wiener-Lévy process in one

dimension, see (91). Onsager and Machlup [88, 89] 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 (96), and the associated symplectic Hamiltonian were first written down in 1970’s by Freidlin and Wentzell [89], whose formulation of the ‘large deviation principle’ was inspired by the Feynman quantum path integral [90]. Feynman, in turn, followed Dirac [91] who was the first to discover that in the short-time limit the quantum propagator (imaginary time, quantum sibling of the Wiener stochastic distribution (91)) is exact. Gaspard [4] thus refers to the ‘pseudo-energy of the Onsager-Machlup-Freidlin-Wentzell scheme.’ M. Roncadelli [22, 92] refers to the Fokker-Planck exponent in (96) as the ‘Wiener-Onsager-Machlup Lagrangian,’ constructs weak noise saddle-point expansion and writes transport equations for the higher order coefficients.

B.1. Ornstein-Uhlenbeck process

The variance (17) 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 (86) is the Langevin flow in one dimension,

$$\frac{dx}{dt} = \lambda x + \hat{\xi}(t), \quad (99)$$

with ‘drift’ $v(x)$ linear in x , and the single deterministic equilibrium solution $x = 0$. The associated Fokker-Planck equation (97) is known as the Ornstein-Uhlenbeck process [93, 11, 12, 94, 3]:

$$\partial_t \rho(x, t) + \partial_x (\lambda x \rho(x, t)) = D \partial_x^2 \rho(x, t). \quad (100)$$

(Here $\Delta = 2D$, $D =$ Einstein diffusion constant.) One can think of this equation as the linearization of the Fokker-Planck equation (97) around an equilibrium point. 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 [95, 96] discussed here in sect. 4.1.

The key ideas are easier to illustrate by the noisy, strictly equivalent discrete-time dynamics of sect. 4.1, rather than by pondering the meaning of the stochastic differential equation (100).

Remark B.2 Quantum mechanical analogue. The relation between Ornstein-Uhlenbeck process and the Schrödinger equation for the quantum harmonic oscillator is much older than

quantum mechanics: Laplace [13] wrote down in 1810 what is now known as the Fokker-Planck equation and computed the Ornstein-Uhlenbeck process eigenfunctions [97] in terms of Hermite polynomials (33). According to L Arnold [14] review of the original literature, the derivations are much more delicate: the noise is *colored* rather than Dirac delta function in time. He refers only to the linear case (99) as the ‘Langevin equation’.

C. LYAPUNOV EQUATION

In his 1892 doctoral dissertation [26] A. M. Lyapunov defined a dynamical flow to be “stable in the sense of Lyapunov” if the scalar Lyapunov function

$$V(x) \geq 0, \quad V(0) = 0$$

computed on the state space of dynamical flow $\dot{x} = v(x)$ satisfies the ‘inflow’ condition

$$\dot{V} = \frac{\partial V}{\partial x_j} v_j \leq 0. \quad (101)$$

While there is no general method for constructing Lyapunov functions, for a linear d -dimensional autonomous flow, $\dot{x} = Ax$, the Lyapunov function can be taken quadratic,

$$V(x) = x^T \frac{1}{Q} x, \quad Q = Q^T > 0, \quad (102)$$

and (101) takes form

$$\dot{V} = x^T \left(A^T \frac{1}{Q} + \frac{1}{Q} A \right) x.$$

Here A^T is the transpose of the stability matrix A , and $Q > 0$ is the shorthand for matrix being positive definite (or Hurwitz), i.e., having the entire eigenvalue spectrum, $\sigma(Q) \in \mathbb{C}_+$, in the right-hand half of the complex plane. Strict positivity $Q > 0$ guarantees that Q is invertible.

The Lyapunov differential equation for a time varying system is

$$\dot{Q} = AQ + QA^T + \Delta, \quad Q(t_0) = Q_0. \quad (103)$$

For steady state, time invariant solutions $\dot{Q} = 0$, (103) becomes the Lyapunov matrix equation

$$AQ + QA^T + \Delta = 0. \quad (104)$$

The Lyapunov theorem [26, 98] states that a $[d \times d]$ matrix Q has all its characteristic roots with real parts positive if, and only if, for any positive definite symmetric matrix $\Delta = \Delta^T > 0$, there exists a unique positive definite symmetric matrix Q that satisfies the *continuous Lyapunov equation* (104). The flow is then said to be asymptotically stable. In our application we are given a noise correlation matrix Δ , and the theorem states the obvious: the stationary state with a covariance matrix $Q > 0$ exists provided that the

stability matrix A is stable, $A < 0$. We have to require strict stability, as $A \leq 0$ would allow for a noncompact, Brownian diffusion along the marginal stability eigen-directions.

For a linear discrete-time system $x_{n+1} = Mx_n$ the quadratic Lyapunov function (102) must satisfy

$$V(x_{n+1}) - V(x_n) = x_n^T \left(M^T \frac{1}{Q} M - \frac{1}{Q} \right) x_n \leq 0,$$

leading to the *discrete Lyapunov equation*

$$Q = MQM^T + \Delta, \quad \text{for any } \Delta = \Delta^T > 0. \quad (105)$$

Lyapunov theorem now states that there is a unique solution Q , provided the Jacobian matrix M is a *convergent matrix*, i.e., a matrix whose eigenvalues (Floquet multipliers) are all less than unity in magnitude, $|\Lambda_i| < 1$.

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 appendix B:

$$M = e^{A\delta t}, \quad M^T = e^{A^T\delta t}, \quad \Delta \rightarrow \delta t \Delta, \quad (106)$$

where $A = \partial v / \partial x$ is the stability matrix. Expanding to linear order yields the differential version of the equilibrium condition (17)

$$0 = AQ + QA^T + \Delta \quad (107)$$

(with the proviso that now Δ is covariance matrix (95) for the velocity fluctuations). The condition (107) is well known [14, 99, 3] and widely used, for example, in the molecular and gene networks literature [100, 101, 102, 103]. In one dimension, $A \rightarrow \lambda$ (see (63)), Δ diffusion tensor $\rightarrow \Delta$, and the diffusive width is given by [14]

$$Q = -\Delta / 2\lambda, \quad (108)$$

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

If A has eigenvalues of both signs, the necessary and sufficient condition for the existence of a unique solution to time invariant case (104) is that no two eigenvalues add up to zero [104],

$$\lambda^{(i)} + \lambda^{(j)} \neq 0, \quad i, j = 1, 2, \dots, d. \quad (109)$$

For the discrete Lyapunov equation (105) the condition for the existence of a unique solution is that no two eigenvalues have product equal to one,

$$\Lambda_i \Lambda_j \neq 1, \quad i, j = 1, 2, \dots, d. \quad (110)$$

This condition is obviously satisfied in the Lyapunov case, with Jacobian matrix M asymptotically stable in the discrete-time domain (all eigenvalues of M are strictly inside of a unit circle).

If A is stable, the continuous Lyapunov equation (104) has a unique solution

$$Q = \int_0^{\infty} dt e^{At} \Delta e^{A^T t}, \quad (111)$$

and

$$Q = \sum_{k=0}^{\infty} M^k \Delta (M^T)^k$$

is the unique solution of the discrete Lyapunov equation (105).

In chaotic dynamics we are interested in saddles, i.e., hyperbolic points with both expanding and contracting eigen-directions. Given a $[d \times d]$ square matrix A with real elements, let the numbers of eigenvalues in the left half of the complex plane, the imaginary axis and the right half of the complex plane be denoted by the integer triple

$$\text{In}A = (\text{In}A_-, \text{In}A_0, \text{In}A_+)$$

which counts (stable, marginal, unstable) eigen-directions. Following J. J. Sylvester (1852), this triple is called the *inertia* of matrix A . In the case of a nondegenerate symmetric bilinear form (such as the symmetric matrix Q) the numbers of positive and negative eigenvalues are also known as the signature of the matrix. A is said to be (negative) stable if $A < 0$, i.e., $\text{In}A = (d, 0, 0)$, and positive stable if $A > 0$, i.e., $\text{In}A = (0, 0, d)$.

The Lyapunov theorem generalized to hyperbolic fixed points is known as the Main Inertia Theorem [105, 106, 107]: For a given stability matrix A and a noise correlation matrix $\Delta = \Delta^T > 0$, there exists a symmetric Q such that $AQ + QA^T + \Delta = 0$ and

$$(\text{In}Q_-, 0, \text{In}Q_+) = (\text{In}A_+, 0, \text{In}A_-),$$

if and only if A has no marginal eigenvalues [108], $\text{In}A_0 = 0$. The Lyapunov theorem is a special case: it states that $Q > 0$ if and only if $\text{In}A = (d, 0, 0)$.

We shall solve (104) and diagonalize Q numerically. The Main Inertia Theorem guarantees that the covariance matrix Q will have the same number of expanding / contracting semi-axes as the number of the expanding / contracting eigenvalues of the stability matrix A . The contracting ones define the semi-axes for covariance evolution forward in time, and the expanding ones the semi-axes for the adjoint evolution.

Remark C.1 Lyapunov equation. The continuous Lyapunov equation (104) is a special case of the Sylvester equation,

$$AQ + QB = C \quad (112)$$

where A, B, Q, C are $[d \times d]$ matrices, and the discrete Lyapunov equation (105) is a special case of Stein's equation. The Sylvester equation can be solved numerically with the Bartels-Stewart algorithm [109], in full generality, with no assumptions on degenerate eigenvalues or defective matrices (matrices for which there are fewer eigenvectors than dimensions). It is implemented in LAPACK, Matlab and GNU Octave. Kuehn [18] reviews the available numerical methods for solving the Lyapunov equation. Discrete Lyapunov equation is solved numerically with the Kitagawa algorithm [110], using, for example, `Mathematica DiscreteLyapunovSolve`.

The Mathematica solvers for the continuous- and discrete-time Lyapunov equation implement the Schur method by Bartels and Stewart [109], based on the decomposition of to the real Schur form; and the Hessenberg-Schur method by Golub, Nash, and Van Loan [111] to solve the Sylvester equation, based on the decomposition of the smaller of two matrices and to the real Schur form and the other matrix to the Hessenberg form [112]. For continuous Lyapunov equation see Matlab function `lyap`, which cites ref. [113, 114, 115, 116, 117, 111] as sources. Related is Matlab function `covar` for output and state covariance of a system driven by white noise. For a more practical criterion that matrix has roots with negative real parts, see Routh-Hurwitz stability criterion [118] on the coefficients of the characteristic polynomial. The criterion provides a necessary condition that a fixed point is stable, and determines the numbers of stable/unstable eigenvalues of a fixed point.

D. 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

$$\mathcal{L}^\dagger e^{-\alpha^2 z_{a-1}^4} = \int e^{-\frac{(y-f'_{a-2}z_{a-2})^2}{2\Delta} - \alpha^2 y^4} [dy] \quad (113)$$

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

$$\begin{aligned} \mathcal{L}^\dagger e^{-\alpha^2 z_{a-1}^4} &= \sqrt{\frac{2\Delta}{\alpha}} \int [d\xi] e^{-\left(\frac{\xi}{\sqrt{\alpha}} - \frac{f'_{a-2}z_{a-2}}{\sqrt{2\Delta}}\right)^2} \sum_{n=0}^{\infty} (-1)^n \frac{[(2\Delta)^2 \xi^4]^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n (4n)!}{n!} (\sqrt{\alpha} f'_{a-2} z_{a-2})^{4n} \sum_{k=0}^{2n} \frac{1}{(4n-2k)! k!} \left(\frac{2\Delta}{4(f'_{a-2} z_{a-2})^2}\right)^k \end{aligned} \quad (114)$$

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

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} (\eta)^{4n} + 2\Delta \sum_{n=0}^{\infty} \frac{(-1)^n (4n)!}{4[n!(4n-2)!]} \alpha^{1-2n} \eta^{4n-2} = \\ e^{-\eta^4} - 2\Delta [3\alpha\eta^2] \Phi\left(\frac{7}{4}, \frac{3}{4}, -\eta^4\right) \end{aligned} \quad (115)$$

where Φ (also sometimes called M or ${}_1F_1$ in the literature) is a confluent hypergeometric function of the first kind [119], which can be expressed as

$$\Phi\left(\frac{7}{4}, \frac{3}{4}, -\eta^4\right) = \frac{4}{3} e^{-\eta^4} \left(\frac{3}{4} - \eta^4\right) \quad (116)$$

We now want to evaluate the variance of the density (115)

$$Q_{a-2} = \frac{\int dz_{a-2} z_{a-2}^2 \rho_{a-2}(z_{a-2})}{\int dz_{a-2} \rho_{a-2}(z_{a-2})}. \quad (117)$$

It is useful to know, when computing the denominator of (117), that

$$\int \eta^2 \Phi\left(\frac{7}{4}, \frac{3}{4}, \eta^4\right) d\eta = 0 \quad (118)$$

so that

$$\begin{aligned} Q_{a-2} &= \frac{(\alpha^2 f_{a-2}'^4)^{-1/2} \int \eta^2 e^{-\eta^4} d\eta - 4D[3f_{a-2}'^{-2}] \int \eta^4 \Phi\left(\frac{7}{4}, \frac{3}{4}, -\eta^4\right) d\eta}{\int e^{-\eta^4} d\eta} \\ &= \frac{1}{f_{a-2}'^2} \left(\frac{\Gamma(3/4)}{\Gamma(1/4)} \frac{1}{\alpha} + \Delta \right) = \frac{Q_{a-1} + \Delta}{f_{a-2}'^2} \end{aligned} \quad (119)$$

in the last identity we used the definition of α and (50).

Next we derive (116), which expresses the a confluent hypergeometric function in terms of an exponential and a Laguerre polynomial. Start with the identity [119]:

$$\Phi(a, b, z) = e^z \Phi(b - a, b, -z) \quad (120)$$

in particular, the hypergeometric function in (115) becomes

$$\Phi\left(\frac{7}{4}, \frac{3}{4}, -\eta^4\right) = e^{-\eta^4} \Phi\left(-1, \frac{3}{4}, \eta^4\right) \quad (121)$$

A confluent hypergeometric function can be written in terms of a Laguerre polynomial [119]:

$$L_n^\alpha(x) = \binom{n+\alpha}{n} \Phi(-n, \alpha+1, x), \quad L_n^\alpha(x) = \sum_{m=0}^n (-1)^m \binom{n+\alpha}{n-m} \frac{x^m}{m!} \quad (122)$$

Thus, in our case

$$L_1^{-1/4}(\eta^4) = \binom{1-1/4}{1} \Phi\left(-1, \frac{3}{4}, \eta^4\right) \quad (123)$$

and

$$\Phi\left(-1, \frac{3}{4}, \eta^4\right) e^{-\eta^4} = \frac{4}{3} \left(\frac{3}{4} - \eta^4\right) e^{-\eta^4}. \quad (124)$$

REFERENCES

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

2. A. Lasota, and 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, *J. Stat. Phys.* **106**, 57 (2002).
5. H. C. Fogedby, *Phys. Rev. Lett.* **94**, 195702 (2005).
6. H. C. Fogedby, *Phys. Rev. E* **73**, 031104 (2006).
7. D. Lippolis, and P. Cvitanović, *Phys. Rev. Lett.* **104**, 014101 (2010), [arXiv:0902.4269](https://arxiv.org/abs/0902.4269).
8. D. Ruelle, *Statistical Mechanics, Thermodynamic Formalism*, Addison-Wesley, Reading, MA, 1978.
9. P. Cvitanović, *Phys. Rev. Lett.* **61**, 2729 (1988).
10. A. Einstein, *Ann. Physik* **17**, 549 (1905).
11. H. Dekker, and N. G. V. Kampen, *Phys. Lett. A* **73**, 374–376 (1979).
12. P. Gaspard, G. Nicolis, A. Provata, and S. Tasaki, *Phys. Rev. E* **51**, 74 (1995).
13. P. S. Laplace, *Mem. Acad. Sci. (I), XI, Section V*. pp. 375–387 (1810).
14. L. Arnold, *Stochastic Differential Equations: Theory and Applications*, Wiley, New York, 1974.
15. Y. Kifer, *Math. USSR-Izv.* **8**, 1083–1107 (1974).
16. M. J. Feigenbaum, and B. Hasslacher, *Phys. Rev. Lett.* **49**, 605–609 (1982).
17. A. Boyarsky, *Physica D* **11**, 130–146 (1984).
18. C. Kuehn, Deterministic continuation of stochastic metastable equilibria via Lyapunov equations and ellipsoids (2011), [arXiv:1106.3479](https://arxiv.org/abs/1106.3479).
19. P. Cvitanović, C. P. Dettmann, R. Mainieri, and G. Vattay, *J. Stat. Phys.* **93**, 981 (1998), [arXiv:chao-dyn/9807034](https://arxiv.org/abs/chao-dyn/9807034).
20. P. Cvitanović, C. P. Dettmann, R. Mainieri, and G. Vattay, *Nonlinearity* **12**, 939 (1999), [arXiv:chao-dyn/9811003](https://arxiv.org/abs/chao-dyn/9811003).
21. P. Cvitanović, N. Søndergaard, G. Palla, G. Vattay, and C. P. Dettmann, *Phys. Rev. E* **60**, 3936 (1999), [arXiv:chao-dyn/9904027](https://arxiv.org/abs/chao-dyn/9904027).
22. M. Roncadelli, *Phys. Rev. E* **52**, 4661 (1995).
23. M. Gutzwiller, *Chaos in Classical and Quantum Mechanics*, Springer, Berlin, 1990.
24. P. Cvitanović, R. Artuso, R. Mainieri, G. Tanner, and G. Vattay, *Chaos: Classical and Quantum*, Niels Bohr Institute, Copenhagen, 2012, ChaosBook.org.
25. M. K. Tippett, and S. E. Cohn, *Nonlinear Processes in Geophysics* **8**, 331–340 (2001).
26. A. M. Lyapunov, *Int. J. Control* **55**, 531–534 (1992).
27. R. E. Kalman, *Trans. ASME – J. Basic Engineering* **82**, 35–45 (1960).
28. H. D. I. Abarbanel, D. R. Creveling, R. Farsian, and M. Kostuk, *SIAM J. Appl. Math.* **8**, 1341–1381 (2009).
29. P. Cvitanović, and D. Lippolis, Optimal resolution of the state space of a chaotic flow in presence of noise (2012), in preparation.
30. V. I. Oseledec, *Trans. Moscow Math. Soc.* **19**, 197–221 (1968).
31. H. H. Rugh, *Nonlinearity* **5**, 1237 (1992).
32. K. Zhou, G. Salomon, and E. Wu, *Int. J. Robust and Nonlin. Contr.* **9**, 183–198 (1999).
33. O. Cepas, and J. Kurchan, *European Physical Journal B* **2**, 221 (1998), [arXiv:cond-mat/9706296](https://arxiv.org/abs/cond-mat/9706296).
34. V. Fedorov, *Theory of Optimal Experiments*, Academic Press, New York, 1972.
35. L. D. Brown, I. Olkin, J. Sacks, and H. P. Wynn, editors, *Jack Carl Kiefer Collected Papers*, Springer, Berlin, 1985, vol. III, chap. Design of Experiments, p. 718.
36. A. C. Atkinson, and V. V. Fedorov, *Biometrika* **62**, 57–70 (1975).
37. V. Fedorov, and V. Khabarov, *Biometrika* **73**, 183–190 (1986).
38. J. P. Crutchfield, and N. H. Packard, *Physica D* **7**, 201–223 (1983).
39. C. S. Daw, C. E. A. Finney, and E. R. Tracy, *Rev. Sci. Instrum.* **74**, 915–930 (2003).
40. X. Z. Tang, E. R. Tracy, A. D. Boozer, A. deBrauw, and R. Brown, *Phys. Rev. E* **51**, 3871–3889 (1995).
41. M. Lehrman, A. B. Rechester, and R. B. White, *Phys. Rev. Lett.* **78**, 54 (1997).
42. M. B. Kennel, and M. Buhl, *Phys. Rev. Lett.* **91**, 084102 (2003), [arXiv:nlin/0304054](https://arxiv.org/abs/nlin/0304054).
43. M. B. Kennel, and M. Buhl, “Estimating good discrete partitions from observed data: symbolic false nearest neighbors,” in *Experimental Chaos: Seventh Experimental Chaos Conference*, edited by L. Kočarev, T. L. Carroll, B. J. Gluckman, S. Boccaletti, and J. Kurths, Am. Inst. of Phys.,

- Melville, New York, 2003, pp. 380–380.
44. M. Buhl, and M. B. Kennel, *Phys. Rev. E* **71**, 046213 (2005).
 45. D. Holstein, and H. Kantz, *Phys. Rev. E* **79**, 056202 (2009), [arXiv:0808.1513](https://arxiv.org/abs/0808.1513).
 46. R. P. Boland, T. Galla, and A. J. McKane, Limit cycles, complex Floquet multipliers and intrinsic noise (2009), [arXiv:0903.5248](https://arxiv.org/abs/0903.5248).
 47. S. M. Ulam, *A Collection of Mathematical Problems*, Interscience Publishers, New York, 1960.
 48. A. B. Rechester, and R. B. White, *Phys. Lett. A* **156**, 419–424 (1991).
 49. A. B. Rechester, and R. B. White, *Phys. Lett. A* **158**, 51–56 (1991).
 50. E. Bollt, P. Góra, A. Ostruszka, and K. Życzkowski, *SIAM J. Applied Dynam. Systems* **7**, 341–360 (2008), [arXiv:nlin/0605017](https://arxiv.org/abs/nlin/0605017).
 51. M. Dellnitz, and O. Junge, *Computing and Visualization in Science* **1**, 63–68 (1998).
 52. R. Guder, and E. Kreuzer, *Nonlinear Dynamics* **20**, 21–32 (1999).
 53. G. Froyland, *Nonlinear Analysis, Theory, Methods and Applications* **32**, 831–860 (1998).
 54. M. Keane, R. Murray, and L.-S. Young, *Nonlinearity* **11**, 27–46 (1998).
 55. G. Froyland, “Extracting dynamical behaviour via Markov models,” in *Nonlinear Dynamics and Statistics: Proc. Newton Inst., Cambridge 1998*, edited by A. Mees, Birkhäuser, Boston, 2001, pp. 281–321.
 56. Y. G. Sinai, *Russian Math. Surveys* **166**, 21 (1972).
 57. R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Springer, Berlin, 1975.
 58. D. Ruelle, *Invent. Math.* **34**, 231 (1976).
 59. C. P. Dettmann, *Physica D* **187**, 214–222 (2004).
 60. R. Artuso, E. Aurell, and P. Cvitanović, *Nonlinearity* **3**, 361 (1990).
 61. D. Ruelle, *Nonlinearity* **22**, 855–870 (2009), [arXiv:arXiv.org:0901.0484](https://arxiv.org/abs/0901.0484).
 62. J. R. Crutchfield, M. Nauenberg, and J. Rudnick, *Phys. Rev. Lett.* **46**, 933 (1981).
 63. B. Shraiman, C. E. Wayne, and P. C. Martin, *Phys. Rev. Lett.* **46**, 935–939 (1981).
 64. G. Kawahara, and S. Kida, *J. Fluid Mech.* **449**, 291–300 (2001).
 65. P. Cvitanović, and J. F. Gibson, Geometry of turbulence in wall-bounded shear flows: Periodic orbits (2010), in S. I. Abarzhi and K. R. Sreenivasan, eds., *Turbulent Mixing and Beyond, Proc. 2nd Int. Conf. ICTP, Trieste*,
 66. M. C. Gutzwiller, *J. Math. Phys.* **12**, 343–358 (1971).
 67. M. V. Berry, and J. P. Keating, *J. Phys. A* **23**, 4839–4849 (1990).
 68. H. D. I. Abarbanel, R. Brown, and M. B. Kennel, *Int. J. Mod. Phys. B* **5**, 1347–1375 (1991).
 69. H. D. I. Abarbanel, R. Brown, and M. B. Kennel, *J. Nonlinear Sci.* **1**, 175–199 (1991).
 70. J. C. Quinn, and H. D. I. Abarbanel, *Quart. J. Royal Meteor. Soc.* **136**, 19 (2009), [arXiv:0912.1581](https://arxiv.org/abs/0912.1581).
 71. J. F. Gibson, J. Halcrow, and P. Cvitanović, *J. Fluid Mech.* **611**, 107–130 (2008), [arXiv:0705.3957](https://arxiv.org/abs/0705.3957).
 72. P. Gaspard, *Chaos, Scattering and Statistical Mechanics*, Cambridge Univ. Press, Cambridge, 1997.
 73. Z. Levnajić, and I. Mezić, Ergodic theory and visualization I: Visualization of ergodic partition and invariant sets (2008), [arXiv:0805.4221](https://arxiv.org/abs/0805.4221).
 74. C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, *J. Fluid Mech.* **641**, 115 (2009).
 75. R. Artuso, H. H. Rugh, and P. Cvitanović (2012), chapter “Why does it work?”, in ref. [24].
 76. P. Cvitanović, and B. Eckhardt, *J. Phys. A* **24**, L237 (1991).
 77. R. Artuso, E. Aurell, and P. Cvitanović, *Nonlinearity* **3**, 325–359 (1990).
 78. S. Chandrasekhar, *Rev. Mod. Phys.* **15**, 1 (1943).
 79. J. L. Doob, *Ann. Math.* **43**, 351–369 (1942).
 80. P. C. Martin, E. D. Siggia, and H. A. Rose, *Phys. Rev. A* **8**, 423 (1973).
 81. B. Øksendal, *Stochastic Differential Equations*, Springer, New York, 2003.
 82. K. Itô, *Proc. Imp. Acad. Tokyo* **20**, 519–524 (1944).
 83. R. L. Stratonovich, *Conditional Markov Processes and Their Application to the Theory of Control*, Elsevier, New York, 1968.
 84. R. F. Fox, *Phys. Rep.* **48**, 179–283 (1978).
 85. N. Berglund, and B. Gentz, *Noise-Induced Phenomena in Slow-Fast Dynamical Systems: A Sample-Paths Approach*, Springer, Berlin, 2005.

86. L. S. Pontryagin, V. Boltyanskii, R. Gamkrelidze, and E. Mishenko, *Mathematical Theory of Optimal Processes*, Interscience Publishers, New York, 1962.
87. R. Bellman, *Dynamic Programming*, Princeton Univ. Press, Princeton, 1957.
88. L. Onsager, and S. Machlup, *Phys. Rev.* **91**, 1505, 1512 (1953).
89. M. I. Freidlin, and A. D. Wentzel, *Random Perturbations of Dynamical Systems*, Springer, Berlin, 1998.
90. R. Feynman, and A. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill Book Co., New York, 1965.
91. P. A. M. Dirac, *Phys. Z. Sowjetunion* **3**, 64–72 (1933).
92. A. Defendi, and M. Roncadelli, *J. Phys. A* **28**, L515–L520 (1995).
93. G. E. Uhlenbeck, and L. S. Ornstein, *Phys. Rev.* **36**, 823–841 (1930).
94. E. Wax, *Selected Papers on Noise and Stochastic Processes*, Dover, New York, 1954.
95. M. Abramowitz, and I. A. Stegun, editors, *Handbook of Mathematical Functions*, Dover, New York, 1964.
96. H. Takada, Y. Kitaoka, and Y. Shimizua, *Forma* **16**, 17–46 (2001).
97. M. Jacobsen, *Bernoulli* **2**, 271–286 (1997).
98. R. Bellman, *Introduction to Matrix Analysis*, McGraw-Hill, New York, 1960.
99. J. Keizer, *Statistical Thermodynamics of Nonequilibrium Processes*, Springer, Berlin, 1987.
100. J. Paulsson, *Nature* **427**, 415–418 (2004).
101. J. Paulsson, *Phys. Life Rev.* **2**, 157–175 (2005).
102. G. Hornung, and N. Barkai, *PloS Computational Biology* **4**, e80055 (2008).
103. F. J. Bruggeman, N. Blüthgen, and H. V. Westerhoff, *PloS Computational Biology* **5**, e1000506 (2009).
104. P. Lancaster, and M. Tismenetsky, *The Theory of Matrices*, Academic Press, New York, 1985.
105. O. Taussky, *J. SIAM* **9**, 640–643 (1961).
106. A. Ostrowski, and H. Schneider, *J. Math. Anal. Appl.* **4**, 72–84 (1962).
107. D. Carlson, and H. Schneider, *Bull. Amer. Math. Soc.* **68**, 479–484 (1962).
108. A. C. Antoulas, and D. C. Sorensen, *Linear Algebra and its Applications* **326**, 137–150 (2001).
109. R. H. Bartels, and G. W. Stewart, *Comm. ACM* **15**, 820–826 (1972).
110. G. Kitagawa, *Int. J. Control* **25**, 745–753 (1977).
111. G. Golub, S. Nash, and C. Van Loan, *IEEE Trans. Auto. Contr.* **24**, 909–913 (1979).
112. K. Datta, and Y. Hong, *IEEE Trans. Auto. Contr.* **48**, 2223–2228 (2003).
113. A. E. Bryson, and Y. C. Ho, *Applied Optimal Control*, Hemisphere Publ., 1975.
114. A. Y. Barraud, *IEEE Trans. Auto. Contr.* **22**, 883–885 (1977).
115. S. J. Hammarling, *IMA J. Num. Anal.* **2**, 303–325 (1982).
116. N. J. Higham, *A.C.M. Trans. Math. Soft.* **14**, 381–396 (1988).
117. T. Penzl, *Advances in Comp. Math.* **8**, 33–48 (1998).
118. G. Meinsma, *Systems and Control Letters* **25**, 237–242 (1995).
119. I. S. Gradshteyn, and I. M. Ryzhik, *Table of Integrals. Series, and Products*, Academic Press, New York, 1965.