STATE SPACE PARTITIONS OF STOCHASTIC CHAOTIC MAPS

A Thesis
Presented to
The Academic Faculty

by

Jeffrey M. Heninger

In Partial Fulfillment
of the Requirements for the Degree
Bachelor of Science in the
School of Physics

Georgia Institute of Technology
May 2014

STATE SPACE PARTITIONS OF STOCHASTIC CHAOTIC MAPS

Approved by:

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

Professor Brian Kennedy
School of Physics
Georgia Institute of Technology

Date Approved: 1 April 2014

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

First, I would acknowledge my advisor, Predrag Cvitanović, for getting me involved in this project and providing instruction and support throughout the project. Domenico Lippolis began the project and helped me to understand how to move forward with the research. My discussions with Jean Bellissard and Brian Kennedy also yielding useful insights. I would also like to thank Pavel M. Svetlichnyy for breaking an impasse on how to deal with a Gaussian basis.

SUMMARY

The finest resolution that can be achieved in any real chaotic system is limited by the presence of noise. This noise can be used to define neighborhoods of the deterministic periodic orbits using the local eigenfunctions of the Fokker-Planck operator and its adjoint. We extend the work of Domenico Lippolis [4, 7] to include hyperbolic periodic orbits. The dynamics along the stable and unstable directions are separated. The neighborhoods on the stable and unstable manifolds can be defined in the same way as the neighborhoods for entirely stable or entirely unstable orbits. The neighborhoods are then returned to the original coordinates. The Fokker-Planck evolution can be described as a finite Markov transition graph between these neighborhoods. Its spectral determinant is used to calculate the time averages of observables. We apply this technique to calculate long time observables of the Lozi map.

CHAPTER I

INTRODUCTION

1.1 Informal Overview of Our Process

Suppose in the course of your research you happen upon a deterministic system whose behavior is highly erratic. Although you can write down the equations of motion that the system should follow, its long-term behavior proves elusive. It continually returns to states similar to where it starts, although never exactly the same and it has sensitive dependence on initial conditions. In short, your system is deterministically chaotic.

What do you do? Since you have the equations of motion, a first guess is that you should put them on a computer and tell the computer to solve them. This works as long as you are only interested in short times. For longer times, the sensitive dependence on initial conditions defeats you. Any error in your initial conditions, noise in the physical system, or numerical imprecision is amplified until the uncertainty is as large as the system itself.

You need to use a different technique to approach the problem. The key to unraveling the dynamics lies in its recurrence: the system approximately repeats itself. This recurrence suggests that there are periodic solutions, but they are unstable, so you never see them in practice. If you look for these periodic solutions, you find that there are infinitely many of them, with arbitrarily long periods. The number of periodic solutions with a given period is roughly the exponential of that period. The behavior of a periodic solution is known for all time – it simply repeats itself – so the periodic orbits are the key to understanding the long-time behavior of the system.

Although these periodic orbits can never be seen directly, they form the skeleton of the rest of the dynamical system. A typical trajectory, although not periodic itself, can be described in terms of these periodic orbits. The trajectory starts close to one periodic orbit. Since there is continuity in the equations of motion, it will follow the periodic orbit for a ways. However, the neighborhood of the periodic orbit is unstable, so your trajectory will eventually leave the neighborhood the periodic orbit. It then finds itself in the neighborhood of another periodic orbit. The process repeats. Instead of calculating individual trajectories, we instead think of the behavior of a trajectory as the list of periodic orbits it follows and how long it follows each of them. [3]

There is an ambiguity in this approach: there is no notion of what it means for a trajectory to be close to another trajectory. In the deterministic system, the state space can be resolved to an arbitrary precision. All of the periodic orbits must be considered in order to get a complete description of the dynamics. Since there are infinitely many periodic orbits, and they are not completely trivial to calculate, this process would take infinite time.

To resolve this crisis, we return to the noise inherent to any physical system. When calculating individual trajectories, the noise was troublesome; it prevented us from

following any trajectory for a long time. In this new framework, the noise becomes essential because it imposes finite resolution to the state space. Trajectories separated by a short distance are indistinguishable because the noise can transfer a trajectory between them. The noise can be used to define what it means to be close to a periodic orbit.

We define the neighborhood of a periodic orbit using the local competition between the noise and the expansion or contraction of the dynamics. The deterministic dynamics will either make nearby trajectories contract onto or expand away from the periodic orbit. The noise smears out the trajectories. The neighborhood of a periodic orbit is a distribution for which the strength of the noise exactly balances the strength of the deterministic dynamics. This notion of a stationary distribution depends on whether the deterministic dynamics is contracting or expanding. [4, 7]

In this thesis, we consider the case when the deterministic dynamics is contracting in some directions and expanding in other directions. This is necessary for almost all physical applications. The dynamics must be separated into the expanding and contracting directions, the separate notions of stationary distributions are applied on each manifold, and then the resulting distributions are recombined to form the neighborhood in the full state space.

Shorter periodic orbits are more important. The noise has had less time to disrupt the deterministic dynamics. This suggests a systematic way to develop these neighborhoods. Start with the shortest periodic orbits and find their neighborhoods, then find the neighborhoods of increasingly longer periods. The finest possible resolution of the system is when the neighborhoods are just starting to overlap. These neighborhoods cover the attractor, the set of all of long-term behaviors of the system. We have created a finite partition of the attractor.

Once we have these neighborhoods, we are able to think of the dynamics in different terms. Instead of thinking of individual trajectories, we can think of the dynamics in terms of the probability of transferring between one neighborhood and another. The result is a finite Markov graph.

All of the techniques developed for Markov graphs can now be applied to your dynamical system. In particular, there is a theorem which states that ergodic systems – systems for which all of the long-time behavior looks the same – have a unique global stationary distribution, which corresponds to the leading eigenvector of the transition probability matrix. The global stationary distribution describes the probability of finding your system at a certain point on the attractor after the system has been allowed to evolve for a long time.

This global stationary distribution allows you to calculate long-time observables for your system. The long-time average of any variable which you are trying to measure is the sum of the value of the variable on each neighborhood weighted by the probability of being in the neighborhood given by the global stationary distribution.

In this thesis, we apply this technique to the Lozi map, a simple chaotic system in 2-dimensions. However, the technique is general and can be applied to any chaotic system. If your system is in discrete time, the formalism developed here is directly applicable. If your system is in continuous time, you have to first transfer the problem to a Poincaré section before applying this technique. We begin by describing how to

do this continuous to discrete time transformation before describing how to apply this technique in discrete time.

1.2 Literature Review

Our research focuses on the interplay between chaotic deterministic dynamics and weak stochastic variables. In a chaotic system, the long-term behavior of a typical trajectory neither approaches a fixed point nor a periodic orbit. Individual trajectories are difficult to follow because nearby trajectories separate at an exponential rate. Any numerical error in the calculation or initial uncertainty in the data gets magnified exponentially, making these orbits useless for long-time calculations. The individual trajectories approach and explore a set in the state space with dimension greater than one. This set is called the strange attractor of the system and is dense with unstable periodic orbits. Although these periodic orbits cannot be seen directly in a physical realization of the system, they provide the basis for the theoretical description of the dynamics. A typical trajectory will come close to a periodic orbit and follow it for a while. Since the periodic orbit is unstable, the trajectory will eventually move away from it. The trajectory then approaches another periodic orbit. The strange attractor is dense with periodic orbits, so the behavior of the trajectory can be well described by the series of periodic orbits that it follows. Long time expectation values of any observables are calculated in terms of these periodic orbits. Periodic orbit theory is discussed in depth in Chaosbook.org [3]. In order for this description to be made rigorous, there must be a precise definition of what it means for a trajectory to be close to a periodic orbit. To make this definition, we add a stochastic process to the chaotic system. This random noise smears the dynamics and introduces a minimum resolution below which the chaotic dynamics cannot be observed.

Stochastic processes describe systems that are inherently random. The canonical stochastic problem involves the motion of a particle which is being subject to random 'kicks' in different directions. These kicks are collectively called the noise of the system. The motion of the particle is random, so the best description is statistical. This behavior can then be combined with other dynamics to model more complicated random processes.

The interplay between noise and deterministic dynamics has mostly been studied in electrical engineering as control theory. In control theory, a circuit or other electrical system is described by a linear differential equation. The system is controlled by some number of inputs and measured by some number of outputs. Electrical engineers would like to be able to both completely measure and completely control the electrical dynamics of the circuit. This has led them to define Grammians, which describe how the system will respond to given inputs and measurements.

These Grammians were initially formulated for fully stable or fully unstable linear electrical systems in continuous time. Similar techniques were developed for hyperbolic linear systems in continuous time by Zhou *et al.* [12]. Varga extended these techniques for hyperbolic systems in discrete time as well [11]. This research focuses exclusively on linear systems; no work has been done on control theory for chaotic systems.

For our research, the inputs are random, corresponding to the weak noise. We assume that sufficient outputs are given to fully describe the system at each time; we are not concerned with the dynamics of the outputs. The Grammians describe the distributions which are 'stationary' under the combined action of the noise and the dynamics near individual periodic orbits. The definition of 'stationary' is different depending on the stability of the periodic orbit: for entirely contracting system, a stationary distribution is unchanged under the combined action of the dynamics and the noise; while for an entirely expanding system, a distribution is stationary under the adjoint of this process. We use these 'stationary' distributions to define the neighborhoods of the periodic orbits. These neighborhoods can then be used to partition the state space of the system.

The idea that noise can be used to form a partition of the state space was first put forward by Crutchfield and Packard [2]. They argue that the most efficient partition of the state space is the one that maximizes the metric and topological entropy. These entropies are measurements of how chaotic the system is. By choosing the partition that maximizes these entropies, the partition captures as much of the underlying chaos of the system as possible. This technique is an inherently global technique: the partition is chosen to maximize a global quantity. However, chaotic dynamics typically have high spatial variability. The balance of the deterministic dynamics and the background noise thus varies between different locations. The optimum partition we suggest is computed locally near each periodic orbit. It accurately captures the spatial variations of the chaotic dynamics.

Once the state space has been partitioned, the dynamics can be described by a transition matrix. This transition matrix describes the probability that a point in a given neighborhood makes the transition to another neighborhood. Methods for the evaluation of these transition matrices in stochastic systems were developed by Bollt *et al.* and Froyland [1, 5]. These problems do not attempt to develop an optimum partition. Instead, they assume a form for the partition and study the techniques for calculating the transition matrices.

The problem of developing an optimum partition for a stochastic chaotic system was formulated by Domenico Lippolis for his Ph.D. Thesis [4, 6, 7]. He focused on studying the neighborhoods of fixed points which are entirely contracting or entirely repelling. Consider first the situation where the dynamics are entirely contracting. An initial distribution is allowed to evolve under the action of both the noise and the dynamics. Since the noise is weak, the distribution remains localized near the periodic orbit, so the dynamics can be linearized. The distribution spreads out with the noise and contracts with the dynamics. These two processes counteract each other. After a long time, any distribution will converge to a Gaussian, with covariance given by the Lyapunov equation [**?**].

$$Q = MQM^{\top} + \Delta \tag{1}$$

The covariance can then be used to define the neighborhood of each periodic point. For entirely expanding dynamics, an initial distribution will not converge since both the noise and the dynamics cause it to expand. However, there is a similar definition

of the neighborhood using the adjoint of the Lyapunov equation.

$$MQM^\top = Q + \Delta \tag{2}$$

To extend this definition to hyperbolic periodic points, we separate the expanding and contracting directions and solve the Lyapunov equation or its adjoint on each subspace. Several theorems about the solutions of matrix equations are necessary to ensure that this process for finding the covariance matrices always results in well defined solutions. These theorems were proven by Ostrowski and Schneider [9] in 1962.

CHAPTER II

FROM CONTINUOUS TO DISCRETE TIME

2.1 Linear Flow and White Noise

In our applications, we assume that the noise is weak, so the deterministic dynamics dominate the behavior of the system. We focus our analysis on behavior in the vicinity of an equilibrium point or a periodic orbit, although it is equally valid for any orbit of the system. First, we will consider the action of the deterministic dynamics in a neighborhood of the periodic orbit. Then, we consider the action of the noise if the dynamics were absent. The noise and deterministic dynamics are then combined to describe the entire behavior.

Consider an initial condition a small displacement δx away from a periodic orbit. The deterministic dynamics in a locally co-moving frame is linear.

$$\dot{\delta x} = A\ \delta x \tag{3}$$

The deterministic solution is simple.

$$\delta x(t) = e^{(t-t_0)A}\ \delta x(t_0) \tag{4}$$

We now add noise to this system. The differential equation becomes

$$\dot{\delta x} = A\ \delta x + \xi \tag{5}$$

The ξ is a random vector which in general depends on the position $\xi = \xi(x)$. This must be specified in the statement of the problem. This system was first studied by Ornstein and Uhlenbeck in 1930 [10]. ξ is described by several properties:

$\langle \xi \rangle = 0$ – the noise does not have any inherent drift.

$\langle \xi_i\ {\xi_j}^\top \rangle = \Delta_{ij}$ – the random numbers are correlated according to a diffusion matrix.

$\langle \xi(t)\ \xi(t') \rangle \propto \delta(t - t')$ – the noise at each time is independent of the noise at any other time.

The action of the noise without the dynamics is a diffusion process. Small displacements evolve according to

$$\langle \delta x_i\ \delta x_j \rangle = \Delta_{ij}(x)\ \delta t. \tag{6}$$

We now consider how a displacement moves in an infinitesimal time step under the combined action of the noise and the deterministic dynamics.

$$\delta x_{n+1} = J^{\delta t}(x_n)\ \delta x_n + \xi(x_n) \tag{7}$$

The Jacobian for this infinitesimal time step is

$$J^{\delta t} = e^{\delta t A} = \mathbf{1} + \delta t\ A + O(\delta t^2). \tag{8}$$

There are many different possible initial displacements δx. Instead of thinking of each possible displacement, we think of a cloud of initial conditions close to the periodic orbit. The cloud is described by a covariance matrix:

$$Q_{ij} = \langle \delta x_i\ \delta x_j{}^\top \rangle \tag{9}$$

The covariance matrix is calculated by averaging over the different possible initial displacements. The covariance matrix evolves according to (7).

$$\begin{aligned} Q_{ij}(t_{n+1}) &= \Big\langle \big(J^{\delta t}(x(t_n))\delta x_i(t_n) + \xi_i(x(t_n))\big)\big(J^{\delta t}(x(t_n))\delta x_i(t_n) + \xi_i(x(t_n))\big)^\top \Big\rangle \\ &= \Big\langle \big((\mathbf{1} + \delta t\ A)\delta x_i(t_n) + \xi(x(t_n))\big)\big((\mathbf{1} + \delta t\ A)\delta x_i(t_n) + \xi(x(t_n))\big)^\top \Big\rangle \\ &= \Big\langle \delta x_i(t_n)\ \delta x_j(t_n)^\top + \delta t\big(A_{il}Q_{lj}(t_n) + Q_{im}(t_n)A_{mj} + \Delta_{ij}\big) + O(\delta t^2) \Big\rangle \\ &= Q_{ij}(t_n) + \delta t\ \big(A_{il}Q_{lj}(t_n) + Q_{im}(t_n)A_{mj} + \Delta_{ij}\big) + O(\delta t^2) \end{aligned} \tag{10}$$

We take the difference between the covariances at time t_{n+1} and time t_n, divide by δt, and take the limit as $\delta t \to 0$ to get Newton's definition of the derivative. The result is a differential equation for the covariance matrix.

$$\dot{Q} = AQ + QA^\top + \Delta \tag{11}$$

This is the Lyapunov equation [**?**]. Applications of this equation appear frequently in control theory, although with a different interpretation. In control theory, Δ is a matrix of inputs which can be used to influence the dynamics, not a measurement of the intrinsic noise.

When dealing with Lyapunov's equation, we are given $A = A(x(t))$ and $\Delta = \Delta(x(t))$ and we want to solve for $Q(t)$ for any $Q(t_0)$. An exact solution is available.

$$Q(t) = J(t, t_0)Q(t_0)J^\top(t, t_0) + \int_{t_0}^{t} d\tau J(t, \tau)\Delta(x(\tau))J^\top(t, t_0) \tag{12}$$

This can be checked directly by taking derivatives of this expression using $\dot{J} = AJ$.

2.2 Reduction to the Poincaré Section

The solution to Lyapunov's equation (11) applies for any finite time. In particular, we can take the time to be the time to return to a Poincaré section.

A Poincaré section is a subspace of dimension one less than the dimension of the state space chosen so the flow crosses it repeatedly. The dynamics can be considered on the Poincaré section instead of in the entire state space. A initial point on the Poincaré section will move forward in time through the entire state space until it

returns to the section. The dynamics can then be considered in discrete time. After one time step, the initial condition will flow one time around through the state space and return to another point on the section. This process is then repeated with the new point as the initial condition.

Shifting to a Poincaré section allows us to use a technique developed in discrete time to describe systems in continuous time. This is important because most physical systems are most naturally described using continuous time, while the simplest descriptions of chaotic dynamics are typically in discrete time.

When we switch to a description of the dynamics based on the Poincaré section, we must also convert Lyapunov's equation to discrete time. We want to be able to evolve a cloud of initial conditions in the neighborhood of a periodic point on the Poincaré section one step forward in time so it returns to a new neighborhood on the Poincaré section. This can be done by integrating the solution of the Lyapunov equation (12) until it returns to the Poincaré section.

The result is the discrete time Lyapunov equation:

$$Q(t_{n+1}) = J(x(t_n))\ Q(t_n)\ J^{\top}(x(t_n)) + \Delta(x(t_n)), \tag{13}$$

$$\text{where}\ \ \Delta(x(t_n)) = \frac{1}{t_{n+1} - t_n} \int_{t_n}^{t_{n+1}} d\tau\ J(x(\tau))\ \Delta\ J^{\top}(x(\tau)) \tag{14}$$

Even if Δ is uniform in the original state space, it will not be uniform on the Poincaré section because these integrals are carried out over different trajectories. The theory must deal with nonuniform Δ. If Δ has explicit x dependence, the only change is in evaluating the integral for $\Delta(x(t_n))$.

CHAPTER III

LOCAL FORMULATION

Consider a smooth stochastic map in d dimensions.

$$x_{n+1} = f(x_n) + \xi_n, \tag{15}$$

where ξ_n represents white noise. In general, the noise is anisotropic and characterized by a symmetric diffusion tensor $\Delta \equiv \xi\, \xi^\top$. The eigenvalues of Δ are strictly positive, so under one step of noisy dynamics an initial Dirac-δ localized density distribution is smeared out into a Gaussian ellipsoid whose widths and orientations are controlled by the eigenvalues and eigenvectors of Δ.

$$\delta_D(y) = \frac{1}{N} e^{-\frac{1}{2}(y-f(x))^\top \frac{1}{\Delta}(y-f(x))} \tag{16}$$

The normalization constant is $N = (2\pi)^{d/2}\sqrt{\det \Delta}$.

3.1 Attractive Fixed Points

Suppose the map has a fixed point at x^*. In a neighborhood, $x = x^* + z$, we can approximate the map by the linear map $M \equiv \partial f(x^*)$ with the fixed point at $z = 0$.

$$z_{n+1} = M z_n \tag{17}$$

This approximation is valid as long as the distribution is localized near the fixed point. In order for this assumption to always be true, the noise must be weak. Otherwise, in one time step, the distribution will diffuse enough to make the approximation invalid.

We start with an initial Gaussian distribution, centered at $z = 0$.

$$\rho(z, n) = \frac{1}{C_n} e^{-\frac{1}{2} z^\top \frac{1}{Q_n} z} \tag{18}$$

C_n is a normalization constant: $C_n = (2\pi)^{d/2}\sqrt{\det Q_n}$.

If the eigenvalues of Q_n are distinct, the distribution is a cigar-shaped ellipsoid, with the eigenvectors of Q_n giving the orientation of various axes. We would like to translate this one step forward in time. The action of the noisy Fokker-Planck operator is the convolution of the initial distribution with the action of the noise (16). Convolution of a Gaussian with a Gaussian is again a Gaussian, so the noisy Fokker-Planck operator (16) maps the ellipsoid $\rho(z, n)$ into an ellipsoid $\rho(y, n+1)$ one time step later.

$$\begin{aligned} \rho(y, n+1) &= \frac{1}{N C_n} \int dz^d e^{-\frac{1}{2}[(y-Mz)^\top \frac{1}{\Delta}(y-Mz) + z^\top \frac{1}{Q_n} z]} \\ &= \frac{1}{C_{n+1}} e^{-\frac{1}{2} y^\top \frac{1}{Q_{n+1}} y}\,. \end{aligned} \tag{19}$$

To determine Q_{n+1}, complete the square, integrate over z, and get the initial condition squished by the dynamics and spread out by the noise:

$$\text{Let } \frac{1}{J} \equiv M^\top \frac{1}{\Delta} M + \frac{1}{Q_n},\ k \equiv 2M^\top \frac{1}{\Delta} y,\ l \equiv \frac{1}{2} Jk,$$

$$K \equiv -\frac{1}{4} k^\top J k = -y^\top \frac{1}{\Delta} M \left(M \frac{1}{\Delta} M + \frac{1}{Q_n} \right)^{-1} M^\top \frac{1}{\Delta} y,\ L \equiv y^\top \frac{1}{\Delta} y\,.$$

$$\begin{aligned}
(y - Mz)^\top \frac{1}{\Delta}(y - Mz) + z^\top \frac{1}{Q_n} z &= y^\top \frac{1}{\Delta} y - y^\top \frac{1}{\Delta} Mz - z^\top M^\top \frac{1}{\Delta} y + z^\top M^\top \frac{1}{\Delta} Mz + z^\top \frac{1}{Q_n} z \\
&= z^\top \left(M^\top \frac{1}{\Delta} M + \frac{1}{Q_n} \right) z - 2y^\top \frac{1}{\Delta} Mz + y^\top \frac{1}{\Delta} y \\
&= z^\top \frac{1}{J} z - k^\top z + L \\
&= (z - l)^\top \frac{1}{J} (z - l) + K + L \qquad (20)
\end{aligned}$$

The first term integrates out and becomes part of the normalization coefficient.

$$e^{-\frac{1}{2} y^\top \frac{1}{Q_{n+1}} y} \propto e^{-\frac{1}{2} y^\top \left[-\frac{1}{\Delta} M \left(M^\top \frac{1}{\Delta} M + \frac{1}{Q_n} \right)^{-1} M^\top \frac{1}{\Delta} + \frac{1}{\Delta} \right] y} \qquad (21)$$

$$\begin{aligned}
Q_{n+1} &= \left[-\frac{1}{\Delta} M \left(M^\top \frac{1}{\Delta} M + \frac{1}{Q_n} \right)^{-1} M^\top \frac{1}{\Delta} + \frac{1}{\Delta} \right]^{-1} \\
&= \Delta \left[\mathbf{1} - \left(\mathbf{1} + \frac{1}{M^\top} \frac{1}{Q_n} \frac{1}{M} \Delta \right)^{-1} \right]^{-1} \\
&= \Delta \left(\mathbf{1} + \frac{1}{M^\top} \frac{1}{Q_n} \frac{1}{M} \Delta \right) \left(\mathbf{1} + \frac{1}{M^\top} \frac{1}{Q_n} \frac{1}{M} \Delta - \mathbf{1} \right)^{-1} \\
&= \Delta \left(\mathbf{1} + \frac{1}{M^\top} \frac{1}{Q_n} \frac{1}{M} \Delta \right) \frac{1}{\Delta} M Q_n M^\top \\
&= M Q_n M^\top + \Delta \qquad (22)
\end{aligned}$$

This says that the two variances (the noise matrix Δ and the deterministically transported $Q_n \to M Q_n M^\top$) add up as Gaussian variances, i.e., as sums of squares. This result is identical to the result calculated for the evolution of covariances on a Poincaré section (13).

If M has all eigenvalues strictly contracting, $|\Lambda_j| < 1$, any initial compact measure (not only an initial ρ, of Gaussian form) converges to the invariant natural measure ρ_0 whose variance satisfies the stationary condition $Q_n = Q_{n+1} = \cdots = Q_\infty = Q$.

$$Q = M Q M^\top + \Delta. \qquad (23)$$

We look at Q_∞:

$$Q_\infty = \Delta + M \Delta M^\top + M^2 \Delta (M^\top)^2 + \cdots + M^\infty Q_n (M^\top)^\infty \qquad (24)$$

Since M is strictly contracting, the higher terms shrink exponentially, so the series converges and the initial covariance becomes irrelevant. We conclude

$$\begin{aligned} Q &= \Delta + M\Delta M^\top + M^2\Delta(M^\top)^2 + \cdots && (25)\\ &= \sum_{m,n=0}^{\infty} M^n \Delta (M^\top)^m \delta_{mn}. && (26) \end{aligned}$$

To get rid of the δ_{mn} and put the expression in resolvent form, use the Fourier representation of Kronecker δ.

$$\delta_{mn} = \int_0^{2\pi} \frac{d\theta}{2\pi} e^{i\theta(m-n)} \tag{27}$$

$$\begin{aligned} Q &= \int_0^{2\pi} \frac{d\theta}{2\pi} \sum_{m,n=0}^{\infty} (e^{-i\theta}M)^n \Delta (e^{i\theta}M^\top)^m && (28)\\ &= \int_0^{2\pi} \frac{d\theta}{2\pi} \frac{1}{\mathbf{1} - e^{-i\theta}M} \Delta \frac{1}{\mathbf{1} - e^{i\theta}M^\top} && (29) \end{aligned}$$

This gives an integral expression for Q in a d-dimensional map, provided M is strictly contracting.

3.2 Repelling Fixed Points

In order to find the stationary condition when M is strictly expanding, we need to use the adjoint Fokker-Planck operator. After one time step, the adjoint Fokker-Planck operator maps the ellipsoid $\rho(y, n+1)$ into a new ellipsoid $\rho(z, n)$.

$$\begin{aligned} \rho(z,n) &= \frac{1}{NC_{n+1}} \int dy^d e^{-\frac{1}{2}\left[(y-Mz)^\top \frac{1}{\Delta}(y-Mz) + y^\top \frac{1}{Q_{n+1}} y\right]} && (30)\\ &= \frac{1}{C_n} e^{-\frac{1}{2} z^\top \frac{1}{Q_n} z} && (31) \end{aligned}$$

Complete the square, integrate over y as in (20), and get the initial condition squished by the reverse flow and spread out by the noise:

$$MQ_{n-1}M^\top = Q_n + \Delta\,. \tag{32}$$

The stationary condition is $Q_n = Q_{n+1} = Q$.

$$MQM^\top = Q + \Delta \tag{33}$$

This says that the stationary distribution in expanding dynamics is the distribution for which the effect of the deterministic dynamics is the same as the effect of the noise, if each acts on the distribution independently.

Since this is a stationary condition,

$$Q_{n+1} = Q_n = Q_{n-1} = \cdots = Q_{-\infty} = Q\,.$$

We look at $Q_{-\infty}$:

$$MQ_{-\infty}M^\top = \Delta + \frac{1}{M}\Delta\frac{1}{M^\top} + \frac{1}{M^2}\Delta\frac{1}{(M^\top)^2} + \cdots + \frac{1}{M^\infty}Q_n\frac{1}{(M^\top)^\infty}. \tag{34}$$

Since M is strictly expanding, the higher terms shrink exponentially, so the series converges and the initial covariance becomes irrelevant. We conclude

$$\begin{aligned} MQM^\top &= \Delta + \frac{1}{M}\Delta\frac{1}{M^\top} + \frac{1}{M^2}\Delta\frac{1}{(M^\top)^2} + \cdots & (35)\\ &= \sum_{m,n=0}^{\infty}\frac{1}{M^n}\Delta\frac{1}{(M^\top)^m}\delta_{mn} & (36)\end{aligned}$$

To get rid of δ_{mn} and put the expression in resolvent form, use the Fourier representation of Kronecker δ:

$$\delta_{mn} = \int_0^{2\pi}\frac{d\theta}{2\pi}e^{i\theta(n-m)} \tag{37}$$

$$\begin{aligned} MQM^\top &= \int_0^{2\pi}\frac{d\theta}{2\pi}\sum_{m,n=0}^{\infty}\left(e^{i\theta}/M\right)^n\Delta\left(e^{-i\theta}/M^\top\right)^m\\ &= \int_0^{2\pi}\frac{d\theta}{2\pi}\frac{1}{\mathbf{1}-e^{i\theta}/M}\Delta\frac{1}{\mathbf{1}-e^{-i\theta}/M^\top}\\ Q &= \int_0^{2\pi}\frac{d\theta}{2\pi}\frac{-e^{-i\theta}}{\mathbf{1}-e^{-i\theta}M}\Delta\frac{-e^{i\theta}}{\mathbf{1}-e^{i\theta}M^\top}\\ &= \int_0^{2\pi}\frac{d\theta}{2\pi}\frac{1}{\mathbf{1}-e^{-i\theta}M}\Delta\frac{1}{\mathbf{1}-e^{i\theta}M^\top} \end{aligned} \tag{38}$$

This gives us an integral expression for Q in a d-dimensional map, provided M is strictly expanding. While the forward (23) and the adjoint (33) stationarity conditions differ, this is the same expression we found for M strictly contracting.

3.3 Hyperbolic Fixed Points

We now consider the case when M has both expanding and contracting (but no marginal) eigenvalues. Neither the Fokker-Planck operator nor its adjoint can be used to derive an expression for the stationary covariance Q. Instead, we define Q as the integral expression which we found in both the expanding and contracting case.

$$Q \equiv \int_0^{2\pi}\frac{d\theta}{2\pi}\frac{1}{\mathbf{1}-e^{-i\theta}M}\Delta\frac{1}{\mathbf{1}-e^{i\theta}M^\top} \tag{39}$$

We would like to show that this is equivalent to separating out the expanding and contracting parts of M with a similarity transform, finding the stationary covariances for the expanding and contracting parts separately, and then recombining them and doing the inverse transformation to return to the original coordinates.

Let S be the transformation which brings M to Jordan form so the expanding and contracting subspaces are clearly separated.

$$\Lambda \equiv S^{-1}MS = \begin{bmatrix} \Lambda_e & 0 \\ 0 & \Lambda_c \end{bmatrix}. \tag{40}$$

Here Λ_e and Λ_c are the expanding and contracting portions of the monodromy matrix, respectively.

To determine how the diffusion tensor transforms, we have to look at how the stochastic variables ξ transform. The original equation of the map transforms under this similarity transformation as

$$\begin{aligned} z_{n+1} &= Mz_n + \xi_n & (41)\\ S^{-1}z_{n+1} &= S^{-1}MSS^{-1}z_n + S^{-1}\xi_n & (42) \end{aligned}$$

Since $\Delta = \xi\xi^\top$, we conclude that Δ transforms according to

$$\hat{\Delta} \equiv S^{-1}\Delta(S^{-1})^\top = \begin{bmatrix} \Delta_{ee} & \Delta_{ec} \\ \Delta_{ce} & \Delta_{cc} \end{bmatrix}. \tag{43}$$

Let $Q_e > 0, Q_c > 0$ be solutions of

$$Q_e = \int_0^{2\pi} \frac{d\theta}{2\pi} \frac{1}{\mathbf{1} - e^{-i\theta}\Lambda_e} \Delta_{ee} \frac{1}{\mathbf{1} - e^{i\theta}{\Lambda_e}^\top}, \tag{44}$$

$$Q_c = \int_0^{2\pi} \frac{d\theta}{2\pi} \frac{1}{\mathbf{1} - e^{-i\theta}\Lambda_c} \Delta_{cc} \frac{1}{\mathbf{1} - e^{i\theta}{\Lambda_c}^\top}. \tag{45}$$

This is equivalent to saying that Q_e satisfies $\Lambda_e Q_e {\Lambda_e}^\top = Q_e + \Delta_{ee}$, the stationary condition for expanding M, and Q_c satisfies $Q_c = \Lambda_c Q_c {\Lambda_c}^\top + \Delta_{cc}$, the stationary condition for contracting M.

We would like to prove for the hyperbolic case that

$$Q = S \begin{bmatrix} Q_e & 0 \\ 0 & Q_c \end{bmatrix} S^\top. \tag{46}$$

The stationary Q for hyperbolic M is equivalent to a contour integral around the unit circle in the complex plane.

$$Q = \oint_\Gamma \frac{ds}{2\pi} (\mathbf{1} - s^{-1}M)^{-1} \Delta (\mathbf{1} - sM)^{-1} \tag{47}$$

We do our similarity transform to make $M \to \Lambda$ be block diagonal. We can then express this contour integral in terms of the four blocks.

$$\begin{aligned} Q &= \frac{1}{2\pi} \oint_\Gamma ds (SS^{-1} - Ss^{-1}\Lambda S^{-1})^{-1} \Delta ((S^{-1})^\top S^\top - (S^{-1})^\top s \Lambda^\top S^\top)^{-1} \\ &= \frac{1}{2\pi} \oint_\Gamma ds S (\mathbf{1} - s^{-1}\Lambda)^{-1} \hat{\Delta} (\mathbf{1} - s\Lambda^\top)^{-1} S^\top & (48) \end{aligned}$$

$$
\begin{aligned}
& (\mathbf{1} - s^{-1}\Lambda)^{-1}\hat{\Delta}(\mathbf{1} - s\Lambda^\top)^{-1} \\
= & \begin{bmatrix} \mathbf{1} - s^{-1}\Lambda_e & 0 \\ 0 & \mathbf{1} - s^{-1}\Lambda_c \end{bmatrix}^{-1} \begin{bmatrix} \Delta_{ee} & \Delta_{ec} \\ \Delta_{ce} & \Delta_{cc} \end{bmatrix} \begin{bmatrix} \mathbf{1} - s\Lambda_e^\top & 0 \\ 0 & \mathbf{1} - s\Lambda_c^\top \end{bmatrix}^{-1} \\
= & \begin{bmatrix} (\mathbf{1} - s^{-1}\Lambda_e)^{-1} & 0 \\ 0 & (\mathbf{1} - s^{-1}\Lambda_c)^{-1} \end{bmatrix} \begin{bmatrix} \Delta_{ee} & \Delta_{ec} \\ \Delta_{ce} & \Delta_{cc} \end{bmatrix} \begin{bmatrix} (\mathbf{1} - s\Lambda_e^\top)^{-1} & 0 \\ 0 & (\mathbf{1} - s\Lambda_c^\top)^{-1} \end{bmatrix} \\
= & \begin{bmatrix} (\mathbf{1} - s^{-1}\Lambda_e)^{-1}\Delta_{ee}(\mathbf{1} - s\Lambda_e^\top)^{-1} & (\mathbf{1} - s^{-1}\Lambda_e)^{-1}\Delta_{ec}(\mathbf{1} - s\Lambda_c^\top)^{-1} \\ (\mathbf{1} - s^{-1}\Lambda_c)^{-1}\Delta_{ce}(\mathbf{1} - s\Lambda_e^\top)^{-1} & (\mathbf{1} - s^{-1}\Lambda_c)^{-1}\Delta_{cc}(\mathbf{1} - s\Lambda_c^\top)^{-1} \end{bmatrix} \qquad (49)
\end{aligned}
$$

$$
Q = S \oint_\Gamma \frac{ds}{2\pi} \begin{bmatrix} (\mathbf{1} - s^{-1}\Lambda_e)^{-1}\Delta_{ee}(\mathbf{1} - s\Lambda_e^\top)^{-1} & (\mathbf{1} - s^{-1}\Lambda_e)^{-1}\Delta_{ec}(\mathbf{1} - s\Lambda_c^\top)^{-1} \\ (\mathbf{1} - s^{-1}\Lambda_c)^{-1}\Delta_{ce}(\mathbf{1} - s\Lambda_e^\top)^{-1} & (\mathbf{1} - s^{-1}\Lambda_c)^{-1}\Delta_{cc}(\mathbf{1} - s\Lambda_c^\top)^{-1} \end{bmatrix} S^\top \qquad (50)
$$

The upper and lower blocks are Q_e and Q_c, respectively. We can show that the off-diagonal blocks are 0 using Cauchy's Residue Theorem.
All the poles of $(\mathbf{1} - s^{-1}\Lambda_e)^{-1}$ lie outside the unit circle.
All the poles of $(\mathbf{1} - s^{-1}\Lambda_c)^{-1}$ lie inside the unit circle.
All the poles of $(\mathbf{1} - s\Lambda_e^\top)^{-1}$ lie inside the unit circle.
All the poles of $(\mathbf{1} - s\Lambda_c^\top)^{-1}$ lie outside the unit circle.
$(\mathbf{1} - s^{-1}\Lambda_e)^{-1}\Delta_{ec}(\mathbf{1} - s\Lambda_c^\top)^{-1}$ has poles only outside of the unit circle, so you can integrate around Γ counterclockwise to get 0.
$(\mathbf{1} - s^{-1}\Lambda_c)^{-1}\Delta_{ce}(\mathbf{1} - s\Lambda_e^\top)^{-1}$ has poles only inside of the unit circle, so you can integrate around Γ clockwise to get 0.

$$
\therefore Q = S \begin{bmatrix} Q_e & 0 \\ 0 & Q_c \end{bmatrix} S^\top \qquad (51)
$$

Notice that we can deform Γ in any way which doesn't cross the eigenvalues of M and still get the same result.

This technique was developed for linear hyperbolic systems in continuous time by Zhou, *et al.* [12] and in discrete time by Varga [11].

If we are interested in the stationary expanding and contracting covariances individually, we can separate Q:

$$
Q_e \equiv S \begin{bmatrix} Q_e & 0 \\ 0 & 0 \end{bmatrix} S^\top \qquad (52)
$$

$$
Q_c \equiv S \begin{bmatrix} 0 & 0 \\ 0 & Q_c \end{bmatrix} S^\top \qquad (53)
$$

This gives us two stationary distributions, one in the expanding manifold and the other in the contracting manifold.

Thus far, we have taken the problem of finding stationary distributions at hyperbolic points and converted it into a sequence of simpler linear problems:

1. Linear the map at the fixed point.
2. Find a similarity transform S which takes M to Jordan form. $M \to S^{-1}MS = \Lambda$.

3. Apply the transform to diffusion tensor $\Delta \to S^{-1}\Delta(S^{-1})^\top = \hat{\Delta}$.

4. For the expanding parts of Λ, solve the stationarity condition $\Lambda_e Q_e {\Lambda_e}^\top = \Delta_{ee} + Q_e$ for Q_e.

5. For the contracting parts of Λ, solve the stationarity condition $Q_c = \Lambda_c Q_c {\Lambda_c}^\top + \Delta_{cc}$ for Q_c.

6. Return to the original coordinates to find the stationary distributions along the stable and unstable manifolds.

For the special case in which Λ is diagonal, we have explicit expressions for Q_e and Q_c:

$$(Q_e)_{ij} = \frac{1}{\Lambda_i\Lambda_j - 1}(\Delta_{ee})_{ij} \tag{54}$$

$$(Q_c)_{ij} = \frac{1}{1 - \Lambda_i\Lambda_j}(\Delta_{cc})_{ij} \tag{55}$$

Note that they are both positive definite, as covariance matrices should be.

3.4 Periodic Points

Now consider a point x_a on a periodic orbit p of period n_p. This point is a fixed point of the ${n_p}^{th}$ iterate of the map, so we can treat it using the same method developed for fixed points, provided we use the proper effective monodromy matrix and diffusion tensor.

In order to deal with periodic orbits, we need to be able to calculate how a Gaussian distribution evolves along a non-stationary orbit $\{\cdots, x_{a-1}, x_a, x_{a+1}, \cdots\}$ under the action of the Fokker-Planck operator and its adjoint. Since M and Δ vary along the orbit in general, we denote M and Δ at point x_a as M_a and Δ_a, respectively.

We have already found the result for one time step. The action of the Fokker-Planck operator is equivalent to acting on the distribution's covariance matrix:

$$Q_{a+1} = M_a Q_a M_a^\top + \Delta_a \tag{56}$$

The action of the inverse adjoint Fokker-Planck operator is equivalent to acting on the distributions covariance matrix:

$$Q_{a+1} = M_a Q_a M_a^\top - \Delta_a \tag{57}$$

In both cases, the new distribution has covariance Q_{a+1} and is centered at point x_{a+1}.

Introduce some new notation for the monodromy matrix after n steps along a non-stationary orbit.

$$M_a^n \equiv M_{a+n-1} \cdots M_{a+2} M_{a+1} M_a = \prod_{i=1}^{n} M_{a+n-i} \tag{58}$$

We can iterate the action of the Fokker-Planck operator and its inverse adjoint to find how the covariance evolves after multiple time steps.

After two time steps, the Fokker-Planck operator is equivalent to

$$\begin{aligned} Q_{a+2} &= M_{a+1} Q_{a+1} {M_{a+1}}^\top + \Delta_{a+1} && (59)\\ &= M_{a+1} \left(M_a Q_a {M_a}^\top + \Delta_a \right) {M_{a+1}}^\top + \Delta_{a+1} && (60)\\ &= M_a^2 Q_a {M_a^2}^\top + M_{a+1} \Delta_a {M_{a+1}}^\top + \Delta_{a+1}. && (61) \end{aligned}$$

Similarly, the inverse adjoint Fokker-Planck operator is equivalent to

$$Q_{a+1} = M_a^2 Q_a {M_a^2}^\top - M_{a+1} \Delta_a {M_{a+1}}^\top - \Delta_{a+1}. \quad (62)$$

After three time steps, the Fokker-Planck operator is equivalent to

$$\begin{aligned} Q_{a+3} &= M_{a+2} Q_{a+2} {M_{a+2}}^\top + \Delta_{a+2} && (63)\\ &= M_a^3 Q_a {M_a^3}^\top + M_{a+1}^2 \Delta_a {M_{a+1}^2}^\top + M_{a+2} \Delta_{a+1} {M_{a+2}}^\top + \Delta_{a+2}. && (64) \end{aligned}$$

Similarly, the inverse adjoint Fokker-Planck operator is equivalent to

$$Q_{a+3} = M_a^3 Q_a {M_a^3}^\top - M_{a+1}^2 \Delta_a {M_{a+1}^2}^\top - M_{a+2} \Delta_{a+1} {M_{a+2}}^\top - \Delta_{a+2}. \quad (65)$$

After n time steps, the Fokker-Planck operator is equivalent to

$$\begin{aligned} Q_{a+n} &= M_{a+n-1} Q_{a+n-1} {M_{a+n-1}}^\top + \Delta_{a+n-1} && (66)\\ &= M_a^n Q_a {M_a^n}^\top + M_{a+1}^{n-1} \Delta_a {M_{a+1}^{n-1}}^\top + \cdots + M_{a+n-1} \Delta_{a+n-2} {M_{a+n-1}}^\top \\ &\quad + \Delta_{a+n-1} && (67) \end{aligned}$$

Similarly, the inverse adjoint Fokker-Planck operator is equivalent to

$$\begin{aligned} Q_{a+n} &= M_a^n Q_a {M_a^n}^\top - M_{a+1}^{n-1} \Delta_a {M_{a+1}^{n-1}}^\top - \cdots - M_{a+n-1} \Delta_{a+n-2} {M_{a+n-1}}^\top \\ &\quad - \Delta_{a+n-1} && (68) \end{aligned}$$

Since we are primarily interested in periodic orbits, $x_{a+n_p} = x_a$, we can introduce the monodromy matrix for the periodic orbit,

$$M_{p,a} \equiv M_a^{n_p} \quad (69)$$

and the total accumulated noise for the periodic orbit.

$$\begin{aligned} \Delta_{p,a} &\equiv M_{a+1}^{n_p-1} \Delta_a {M_{a+1}^{n_p-1}}^\top + \cdots + M_{a+n_p-1} \Delta_{a+n_p-2} {M_{a+n_p-1}}^\top \\ &\quad + \Delta_{a+n_p-1} && (70)\\ &= \sum_{i=1}^{n_p} M_{a+i}^{n_p-i} \, \Delta_{a+i-1} \, {M_{a+i}^{n_p-i}}^\top && (71) \end{aligned}$$

The stationary distribution occurs when $Q_{a+n_p} = Q_a = Q$.

For the Fokker-Planck operator, the stationary condition is

$$Q = M_{p,a}\ Q\ {M_{p,a}}^{\top}\ +\ \Delta_{p,a}. \tag{72}$$

For the inverse adjoint Fokker-Planck operator, the stationary condition is

$$Q = M_{p,a}\ Q\ {M_{p,a}}^{\top}\ -\ \Delta_{p,a}. \tag{73}$$

This looks identical to the stationary conditions for fixed points, provided we replace M_a by $M_{p,a}$ and Δ_a by $\Delta_{p,a}$. We can therefore use the same techniques to find Q that we developed for fixed points.

CHAPTER IV

GOING GLOBAL

4.1 Partitioning the Attractor

The neighborhoods that we have just defined for each periodic point can be used to partition the attractor. Each neighborhood defines a finite resolution in the vicinity of its periodic point. Any details smaller than this are quickly smeared out by the noise. We can thus keep adding neighborhoods to the partition until they completely cover the attractor. Once the neighborhoods start to overlap, adding more neighborhoods will not give you any more information. The information that you get from the extra neighborhoods is completely smeared out by the noise.

Shorter periodic orbits are less affected by the noise than longer periodic orbits. For a longer periodic orbit, the noise has more time to push a trajectory off of the deterministic path. This implies that if there are two neighborhoods that overlap, the one with the shorter periodic orbit is more important to understanding the dynamics. The longer one is more likely to be disrupted by the noise before it returns to its initial condition.

The definition of the partition can be somewhat arbitrary, as long as it completely covers the attractor. Other partitions that also cover the attractor differ in the amount of overlap that exists between neighborhoods. Our technique for calculating the Markov graph takes into account these overlaps. The technique is robust to changes in the partition. The different amounts of overlap between neighborhoods may result in different representations of the global stationary distribution in terms of the partition, but they all correspond to the same global stationary distribution when we return to the coordinates of the state space.

4.2 Markov Graph on the Attractor

We have a set of N Gaussians which cover the noisy attractor, each of which takes the form

$$\rho_i(x) = \frac{1}{\sqrt{(2\pi)^d \det Q_i}} \; e^{-\frac{1}{2}(x-\mu_i)^\top \frac{1}{Q_i}(x-\mu_i)} \quad . \tag{74}$$

Suppose there is another normalized function $f(x)$ defined on the noisy attractor. We would like to find $\{f_j\}$, the coefficients of f in our Gaussian basis. We write

$$f(x) = \sum_{i=1}^{N} f_i \; \rho_i(x) \; . \tag{75}$$

In order to maintain normalization, the coefficients must satisfy

$$1 = \int f(x)dx = \int \sum_{i=1}^{N} f_i \; \rho_i(x)dx = \sum_{i=1}^{N} f_i \int \rho_i(x)dx = \sum_{i=1}^{N} f_i = 1. \tag{76}$$

We are interested in finding the f_i.

The f_i should be chosen so that the distance between the actual $f(x)$ and its representation in our basis is minimized. To measure the distance, we use the L^2 norm defined over the entire space.

We minimize

$$\frac{1}{2}\int \left(f(x) - \sum_{i=1}^{N} f_i \ \rho_i(x) \right)^2 dx \tag{77}$$

subject to the constraint that the coefficients form a probability distribution: $\sum_i f_i = 1$ and $f_i \geq 0$.

In particular, we can choose $f(x)$ to be one of the basis distributions that has been evolved one step forward in time. In order to evolve a distribution forward in time, we apply the Fokker-Planck operator (19). This is equivalent to moving the center of the Gaussian one step forward in time with the deterministic Lozi map and replacing the covariance matrix with $Q_{n+1} = MQ_nM^\top + \Delta$. The resulting distribution will be referred to as

$$\rho_{k,1}(x) = \mathcal{L}^1_{FP} \ \rho_k(x) \ , \tag{78}$$

where $\mathcal{L}^1_{FP}$ is the one time step Fokker-Planck operator.

The Fokker-Planck operator is used for the entire distribution. We do not have to separate the expanding and contracting directions and apply different operators to each of them. Although this was necessary to determine a stationary distribution, any physical realization of our system only moves forward in time. Because of this, we are only interested in how the distributions move forward in time. Moving exclusively forward in time is necessary to ensure that distributions approach the attractor and that motion on the attractor is ergodic.

We write the coefficients of the basis elements moved one step forward in time in terms of their own basis as $T_{i;k,1}$. These coefficients determine a transition matrix. This matrix contains the probability of transitioning between each pair of basis elements in one time step.

The time evolution of any distribution expressed in terms of the basis $\{\rho_i(x)\}$ can be written in terms of this transition basis.

$$\mathcal{L}^1_{FP} f(x) = \mathcal{L}^1_{FP} \sum_{k=1}^{N} f_k \ \rho_k(x) = \sum_{k=1}^{N} f_k \ \mathcal{L}^1_{FP} \rho_k(x) = \sum_{i=1}^{N} \left(\sum_{k=1}^{N} T_{i;k,1} \ f_k \right) \rho_i(x) \tag{79}$$

The one time step evolution of the coefficients of some distribution in our Gaussian basis can be found by applying the transition matrix on the original coefficients. Longer time dynamics can be found by repeated action of this transition matrix.

4.3 Stationary Distribution and Long-Time Observables

We now have a stochastic transition matrix. Its elements are the probability of transitioning between the neighborhoods of the periodic points in our partition in one time step. The original chaotic system has been reduced to a Markov graph.

We would like to apply the Perron-Frobenius Theorem to this system. In order to apply this theorem, we need the further condition that this system is ergodic. Here, the noise helps us again. Because the tails of the Gaussian neighborhoods extend to infinity, the noise gives us a nonzero transition probability between any two neighborhoods. Some of these probabilities may be extremely small, so they would be indistinguishable from numerical error. The chaotic system that this technique is applied to should at least be ergodic at length scales above the length scale of the noise. Any non-ergodic features smaller than this will be smeared out by the noise. If the system is not ergodic on a larger scale, our formalism can be applied to each attractor individually.

The Perron-Frobenius Theorem declares that the leading eigenvalue of a nonnegative ergodic matrix is isolated and the corresponding eigenvector consists of positive entries. The second eigenvalue is also real and isolated. The corresponding eigenvector can contain positive and negative real entries.

For a globally attracting chaotic system, the leading eigenvalue is one because the system is probability conserving. The leading eigenvector corresponds to the global stationary distribution. This distribution is unchanged under the action of the transition matrix. Any other initial distribution will converge to this. The second eigenvalue measures the exponential rate at which this convergence occurs. Its eigenvector represents how an initial distribution approaches the global stationary distribution. The last variations away from the global stationary distribution will take the form of this second distribution.

Any initial distribution in the full state space can be projected onto our basis of Gaussians and then evolved forward in time. The distribution will converge to the stationary distribution on our basis. If we take the stationary distribution on our basis and return it to the original coordinates, we find the global stationary distribution on the entire state space. Any initial distribution will converge to it under the repeated action of the Fokker-Planck operator.

This distribution can be used to calculate expectation values of long-time observables. The long-time average of the observable is equal to the average of the observable distributed on the state space according to the global stationary distribution. These averages can be calculate and then compared to the measured averages from experimental data.

For a chaotic repeller, the leading eigenvalue is less than one and measures the escape rate. Long-time observables are less relevant in these systems because the dynamics do not stay localized for long times. Instead, the measurable quantity that we are interested in is the escape rate.

4.4 Finding the Global Stationary Distribution Directly

The Markov process discussed above yields a complete characterization of dynamics of distributions on the partition. However, if we are interested in finding the long-time expectation values of long-time observables, we do not need to know the entire transition matrix. Instead, we can use a similar variational calculation to directly calculate the global stationary distribution.

The global stationary distribution $h(x)$ satisfies

$$\mathcal{L}^1_{FP}\ h(x) - h(x) = 0. \tag{80}$$

We can write $h(x)$ in terms of our Gaussian basis (75). The resulting distribution will not satisfy (80) exactly. Instead, we choose normalized coefficients $\{h_i\}$ so that the difference between the one time step evolution of the distribution and the original distribution is minimized. The difference between the functions is calculated using the L^2 norm and is called the error of our approximation.

$$E^2 = \frac{1}{2}\int \left(\mathcal{L}^1_{FP}\left(\sum_{i=1}^N h_i\rho_i(x)\right) - \left(\sum_{i=1}^N h_i\rho_i(x)\right)\right)^2 dx \tag{81}$$

$$= \frac{1}{2}\int \left(\sum_{i=1}^N h_i\left(\mathcal{L}^1_{FP}\rho_i(x) - \rho_i(x)\right)\right)^2 dx \tag{82}$$

We also want to satisfy the constraint that the h_i are a probability distribution.

4.4.1 Constrained Gradient Descent

In order to find this minimum, we use a constrained gradient descent. We begin with the initial guess that the distribution is uniform over the partition: $h_i = 1/N$. We follow the direction of fastest decrease of the error function subject to the constraint $\sum_i h_i = 1$. The minimum of E^2 might occur with all of the h_i positive. If this is the case, the procedure is finished. Otherwise, one of the h_i is zero before the minimum is reached. If this occurs, we enforce this additional constraint and then continue the gradient descent. Iterating this process finds the probability distribution with minimum error.

For this process, the gradient is linear in the coefficients.

$$\frac{\partial(E^2)}{\partial h_j} = \sum_{i=1}^N h_i \int \left(\mathcal{L}^1_{FP}\rho_i(x) - \rho_i(x)\right)\left(\mathcal{L}^1_{FP}\rho_j(x) - \rho_j(x)\right) dx \tag{83}$$

We consider the coefficients h_j to be dynamical variables that move in the direction of fastest decrease of E^2. There is also the constraint that $\sum_j \dot{h}_j = 0$. This can be accomplished by subtracting out the component of the gradient of E^2 which is normal to the surface of constraint.

$$\text{Let}\ \ \mathbf{L}_{ij} = \mathbf{L}_{ji} = \int \left(\mathcal{L}^1_{FP}\rho_i(x) - \rho_i(x)\right)\left(\mathcal{L}^1_{FP}\rho_j(x) - \rho_j(x)\right) dx \tag{84}$$

$$\dot{h}_j = -\sum_{i=1}^N h_i\mathbf{L}_{ij} + \frac{1}{N}\sum_{i,j=1}^N h_i\mathbf{L}_{ij} = -\sum_{i=1}^N h_i\left(\mathbf{L}_{ij} - \frac{1}{N}\sum_{j=1}^N \mathbf{L}_{ij}\right) \tag{85}$$

The dynamical process is linear, so there is a unique minimum. If this minimum has all coefficients positive, then it is the optimum probability distribution. Otherwise, we follow the solution until the first of the h_j is zero. The index of this coefficient

is k_1. At this point, we also enforce the constraint $h_{k_1} = 0$. The dynamical equations then become:

$$\dot{h}_{j\neq k_1} = -\sum_{i=1,\ i\neq k_1}^{N} h_i \left(\mathbf{L}_{ij} - \frac{1}{N-1} \sum_{j=1,\ j\neq k_1}^{N} \mathbf{L}_{ij} \right) \tag{86}$$

$$\dot{h}_{k_1} = 0 \tag{87}$$

This new linear system evolves until it reaches a minimum or until another of the $h_j = 0$. We repeat the same procedure. After m of the coefficients are zero, the dynamical system satisfies m additional non-negativity constraints. The indices of the constrained variables are given by $\{k_1, \cdots, k_m\}$.

$$\begin{aligned} \dot{h}_{j\notin\{k\}} &= -\sum_{i=1,\ i\notin\{k\}}^{N} h_i \left(\mathbf{L}_{ij} - \frac{1}{N-m} \sum_{j=1,\ j\notin\{k\}}^{N} \mathbf{L}_{ij} \right) \\ \dot{h}_{k_1} &= 0 \\ &\vdots \\ \dot{h}_{k_m} &= 0 \end{aligned} \tag{88}$$

After iterating this procedure several times, we obtain the probability distribution which minimizes the error. This procedure is applicable to optimizing any other quadratic function of the coefficients of a probability distribution. In particular, it could be used to determine the transition probabilities between neighborhoods which appear in the stochastic transition matrix.

4.4.2 Gaussian Overlap Integrals

In order to calculate the gradient, we have to perform the overlap integrals between two Gaussians. The integrals take the form:

$$\frac{1}{N_a N_b} \int e^{-\frac{1}{2}(x-\mu_a)^\top \frac{1}{Q_a}(x-\mu_a)} e^{-\frac{1}{2}(x-\mu_b)^\top \frac{1}{Q_b}(x-\mu_b)} dx. \tag{89}$$

This integral is solved by completing the square. A similar procedure is described in (20).

$$\begin{aligned} \text{Let} \quad \frac{1}{J} &\equiv \frac{1}{Q_a} + \frac{1}{Q_b} \\ k &\equiv 2\left(\frac{1}{Q_a}\mu_a + \frac{1}{Q_b}\mu_b \right) \\ K &\equiv -\frac{1}{4} k^\top J k \\ L &\equiv \mu_a^\top \frac{1}{Q_a}\mu_a + \mu_b^\top \frac{1}{Q_b}\mu_b \end{aligned}$$

The value for the overlap integral between two Gaussians is in terms of these quantities.

$$\sqrt{\frac{\det J}{(2\pi)^d \det Q_a \det Q_b}}\, e^{-\frac{1}{2}(K+L)} \tag{90}$$

4.5 Other Eigenfunctions of the Evolution Operator

The one time step Fokker-Planck operator, $\mathcal{L}^1_{FP}$, also has other eigenfunctions. These could have been found from the stochastic transition matrix. We can also use the technique developed for the global stationary distributions to find these eigenfunctions directly.

The eigenfunctions of $\mathcal{L}^1_{FP}$ satisfy

$$\mathcal{L}^1_{FP} h(x) - e^s h(x) = 0 \tag{91}$$

for some constants s. We again express $h(x)$ is terms of our basis and minimize the error. The error is now defined according to

$$E^2 = \frac{1}{2} \int \left(\sum_{i=1}^{N} h_i \left(\mathcal{L}^1_{FP} \rho_i(x) - e^s \rho_i(x) \right) \right)^2 dx. \tag{92}$$

We can again use a constrained gradient descent to find the minimum error. The eigenvalue s is also an unknown. It must therefore also be included as one of the dynamical variables in the gradient descent. The resulting differential equation for s is nonlinear. It has an attracting equilibrium for each eigenvalue of $\mathcal{L}^1_{FP}$. These are local minima of the error function. Which eigenvalue you arrive at as a result of the constrained normalization procedure depends on your initial guess of the parameters.

Since the other eigenfunctions do not have the interpretation that they are global probability distributions, we do not have to enforce the constraints that the h_i are always positive. We do maintain the constraint that these eigenvectors are normalized.

CHAPTER V

LOZI MAP

We now apply our procedure to the Lozi map,

$$\begin{aligned} x_{n+1} &= 1 - a|x_n| + by_n \\ y_{n+1} &= x_n \end{aligned} \tag{93}$$

In our calculations the noise is isotropic, $\Delta = 0.1\ \mathbf{1}$ or $0.01\ \mathbf{1}$, and the Lozi parameters are set to the values $a = 1.85$, $b = 0.3$, unless noted otherwise. For these values of the parameters, the noiseless Lozi map is chaotic [8].

We also include white noise characterized by a diffusion tensor Δ. For this example, we will use isotropic noise.

5.1 Deterministic Attractor

First, we numerically find the attractor of the Lozi map in the absence of noise. This can be accomplished using the following steps:

1. Pick an initial point that is not a fixed or periodic point. We used $(0.1, 0.1)$.
2. Calculate the action of the map for 1,000 iterations. This removes the effects of transient behavior.
3. Calculate the action of the map for the next 100,000 iterations. Store these data points. They are close to the strange attractor.
4. Plot these last 100,000 points. They trace out the noiseless attractor.

5.2 Periodic Points

Next, we need to be able to find all the periodic points of the Lozi attractor of a given period n. This is easier for the Lozi map than for most other dynamical systems since the Lozi map is piecewise linear. The most efficient method for finding the periodic points of the Lozi map is pruning.

The Lozi map can be written as either of the two linear maps, $x_{n+1} = f_s(x_n)$,

$$\begin{aligned} x_{n+1} &= 1 \mp ax_n + by_n \\ y_{n+1} &= x_n\,, \end{aligned}$$

where $s = 1$ if $x > 0$, and $s = 0$ if $x < 0$. Each iteration of the Lozi map is the action of either f_0 or f_1 (depending on whether $x > 0$ or $x < 0$). The Lozi map is thus a set of two linear maps. Which map to use depends on whether there is a 0 or 1 in the itinerary of the periodic point.

Instead of searching for all period-n periodic points, we first compute the set of 2^n possible periodic points, and then prune the points that are not periodic solutions of the Lozi map.

There are 2^n possible ways the Lozi map could act after n time steps, corresponding to all the possible sequences of n 0's and 1's. Each sequence corresponds to a different linear map, so each has exactly one fixed point. These points give us the possible locations of periodic points.

To prune the points down to only the admissible periodic points, we need to check whether each step has the proper 0 or 1 for its current location. To do this, we need to calculate the orbit for each possible periodic point. Since we are already calculating the periodic orbit, there is an easier way to prune. Simply calculate the orbit and check to see whether the orbit returns to itself after n steps. Any orbit that does not return to itself is pruned from the list, and we are left with only the admissible periodic points.

The periodic points of short periods are plotted in figure 1. They are all located on the attractor and, as the period increases, gradually cover it. If we were to include periodic orbits of all lengths, they would be dense on the attractor.

5.3 Local Stationary Distributions

We can now find the stationary covariances at the fixed points.

The one-time step Jacobian matrix for the Lozi map is given by

$$M = \begin{bmatrix} -a \text{ sign}(x) & b \\ 1 & 0 \end{bmatrix}.$$

We then diagonalize M as in (40). Δ is transformed according to (43). We can then solve for Q directly in each one-dimensional subspace.

We separate out the expanding and contracting directions and return to original coordinates (53). We now have two covariance matrices, Q_e and Q_c. Each has one singular direction with 0 variance and one singular direction with positive variance. We have two oriented Gaussian distributions at each fixed point. These distributions are the stationary distributions in the stable (unstable) manifold under the Fokker-Planck (adjoint Fokker-Planck) operator.

The Lozi map has two fixed points, both of which are hyperbolic for our parameter values. The stationary distributions at each are shown in figure 2.

The stationary covariances at the periodic points can be calculated in a similar way to the stationary covariance at the fixed points. We calculate the periodic orbit monodromy matrix $M_{p,a}$ and the total accumulated noise for the periodic orbit $\Delta_{p,a}$, as given in (69) and (70), respectively. We use the same code to calculate Q for the periodic orbits. We plot each neighborhood as a parallelogram, with their widths given by the standard deviations of the covariance matrices and their axis given by the corresponding singular vectors. The neighborhoods for all periods up to period 6, are shown in figure 3.

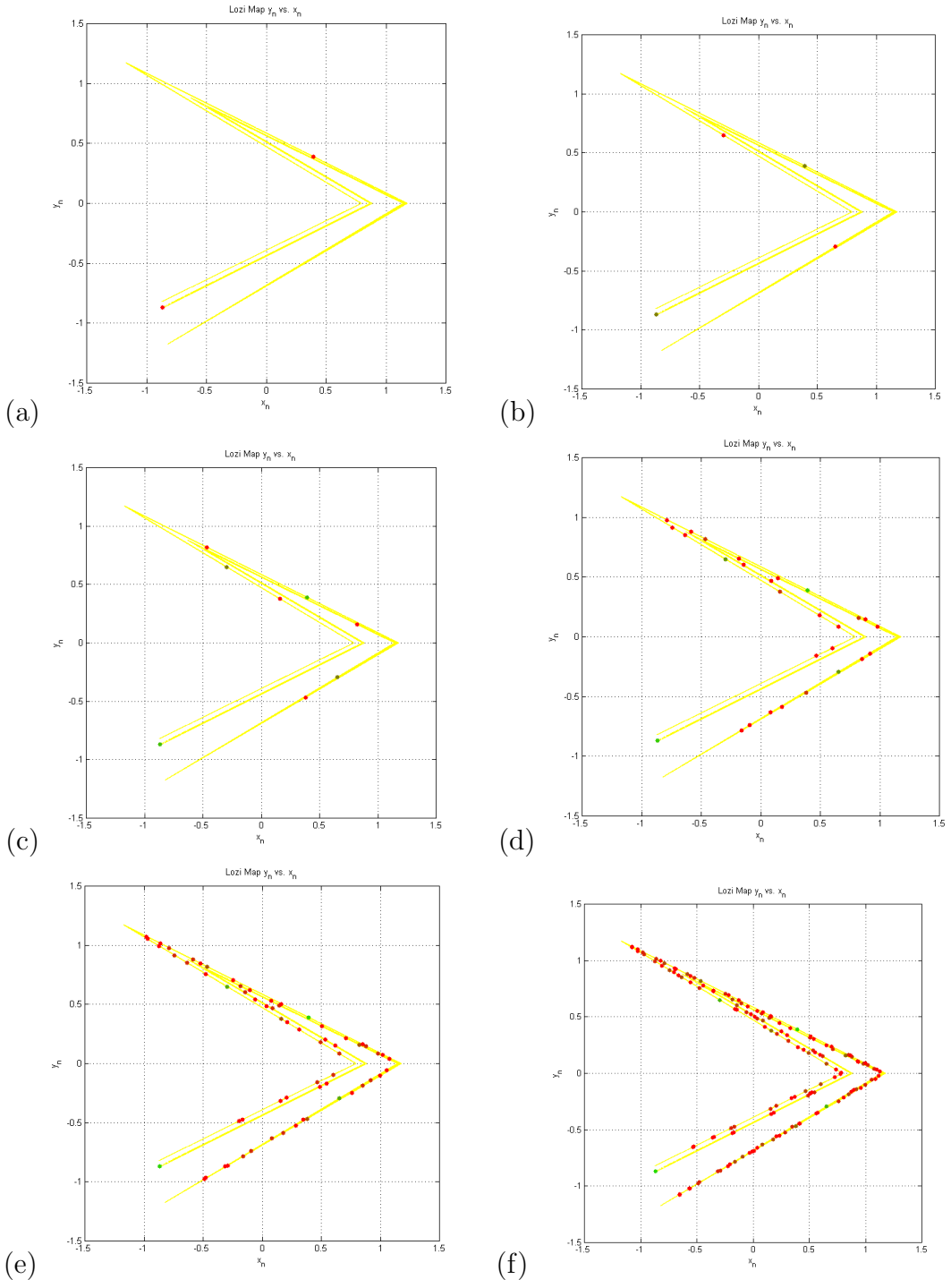

Figure 1: Lozi periodic points of all periods up to period $n = 7$. The color of a periodic point represents its period. The longest ones are red; the shortest ones are green. The noiseless Lozi attractor is plotted in yellow in the background. (a) The fixed points, $n = 1$. (b) $n = 2$. There are no periodic orbits of period 3. (c) $n = 4$. (d) $n = 5$. (e) $n = 6$. (f) $n = 7$.

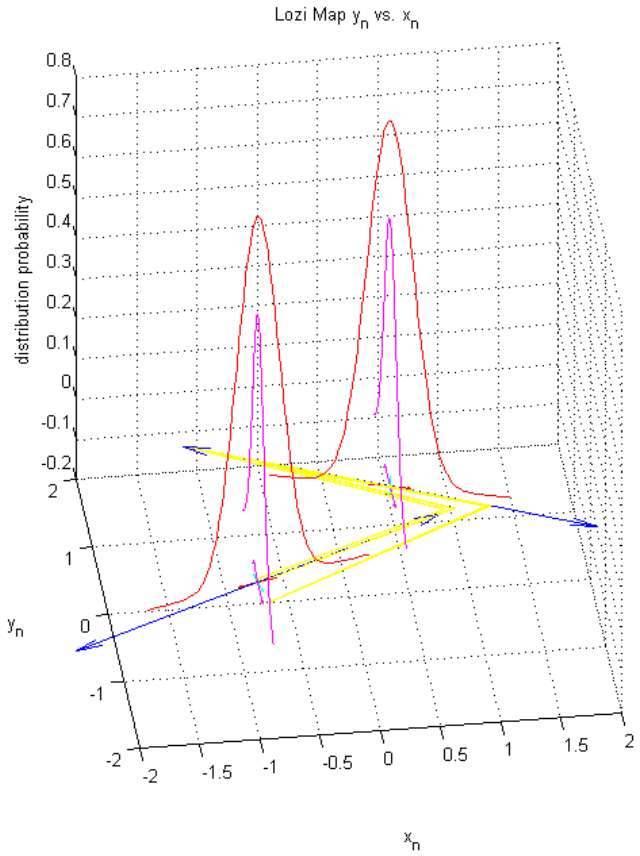

Figure 2: Stationary distributions at the fixed points of the Lozi map, with $\Delta = 0.1$ **1**. The Lozi attractor is shown in yellow in the $x_n - y_n$ plane. The fixed points are marked in red. Extending from the fixed point and in the same plane are the eigenvectors of M and the singular vectors of Q. The expanding eigenvector is shown in blue and the contracting eigenvector is shown in cyan. The lengths are given by the magnitudes of the corresponding eigenvectors. The expanding singular vector is magenta and the contracting singular vector is red. The lengths are given by the corresponding variances. The distributions are shown above the singular vectors in the corresponding colors.

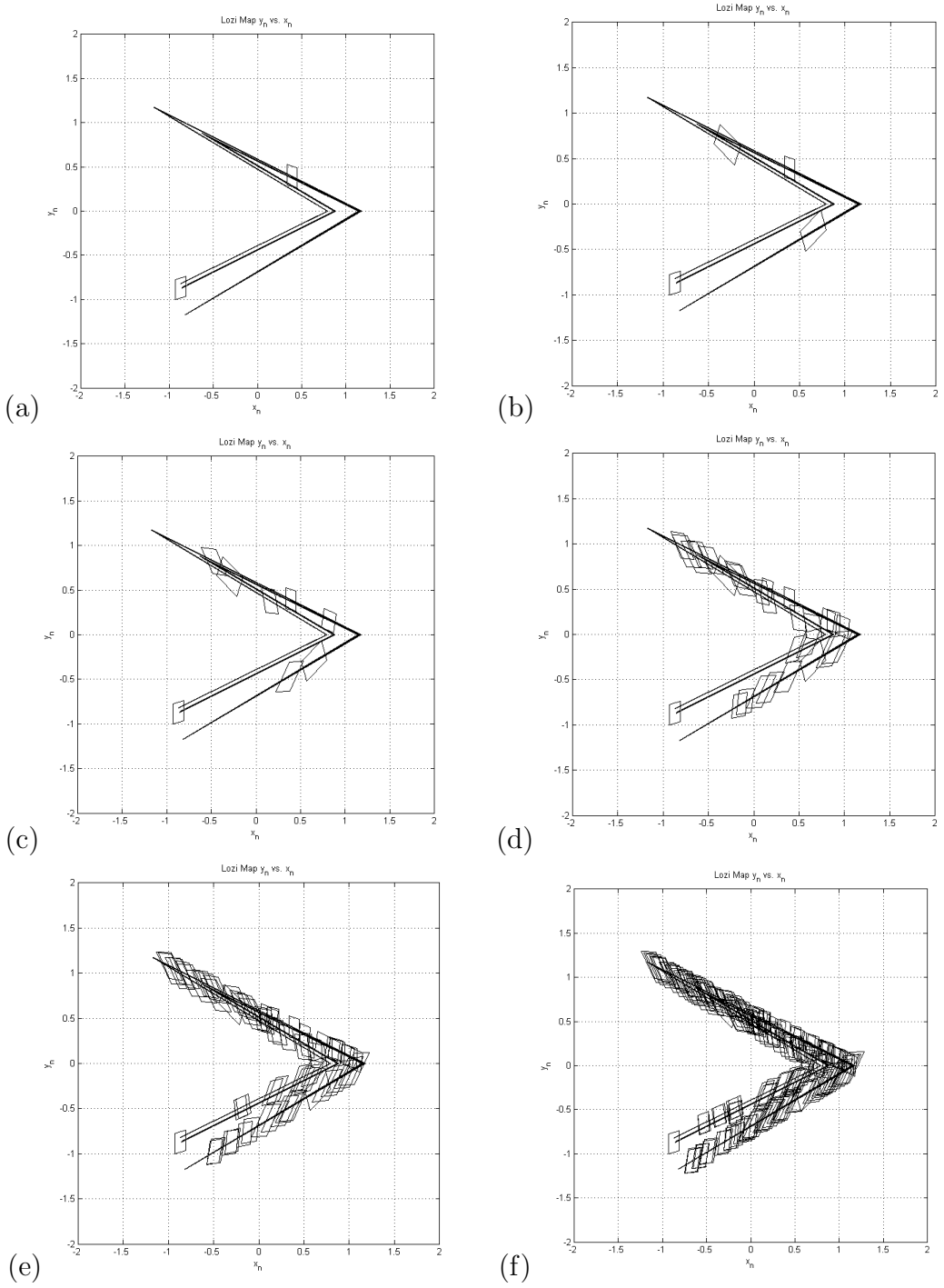

Figure 3: Neighborhoods of all periodic points in the Lozi map for periods up to n. For this figure, we used $\Delta = 0.01$ **1**. (a) $n = 1$. (b) $n = 2$. (c) $n = 4$. (d) $n = 5$. (e) $n = 6$ (f) $n = 7$.

5.4 Partitioning the Attractor

We choose a particular technique for forming the optimum partition of this attractor. We begin with the fixed points of the maps. These are the periodic points with the shortest possible orbit, so they are least effected by the noise. They are automatically included in the partition.

We then consider all points of period one longer than the previous period. If any of their neighborhoods overlap with the neighborhoods of points in the partition of lesser period, they are automatically discarded.

In general, there will also be some points of the same period with overlapping neighborhoods. For each set of these equal period overlapping points, remove the points one at a time until there is no more overlap. There will typically be more than one way to do this. Compare the different ways of removing points from this set and choose the one that covers the most possible volume of the state space. This ensures that there will be no overlap while still covering as much area of the attractor with these points as possible.

This process is repeated from each subsequent period until the entire attractor is covered by neighborhoods of periodic points. Including any more points will only cause overlap between the neighborhoods.

There is still the question of what it means that two neighborhoods 'overlap' and thus should be discarded. Each neighborhood is characterized by a Gaussian, with tails that extend to infinity. We are drawing each neighborhood as a parallelogram with widths given the standard deviations of the stationary covariances and the axes pointing along the expanding and contracting manifolds. Excessive overlap occurs when at least 50% of one of the parallelograms is contained within another parallelogram.

The resulting partitions found using this technique for two different noise strengths are shown in figure 4.

5.5 Global Stationary Distribution

Rather than calculating the entire transition matrix, we calculate the global stationary distribution directly using the constrained gradient descent procedure described in section 4.4. The resulting distributions for two different strengths of the noise are shown in figure 5.

These distributions are not entirely satisfying. Both of them had to enforce non-negativity constraints on some the coefficients. The corresponding Gaussians are thus not used in the global stationary distribution. We do not understand why some of the elements of the partition are automatically pruned from the global stationary distribution.

We also calculated the distributions using a brute force method, shown in figure 6. Because of this self-pruning, the global stationary distributions are more lumpy than they should be. For the weaker value of noise, our calculation misses one of the lower branches of the attractor that can be seen in the brute force calculation.

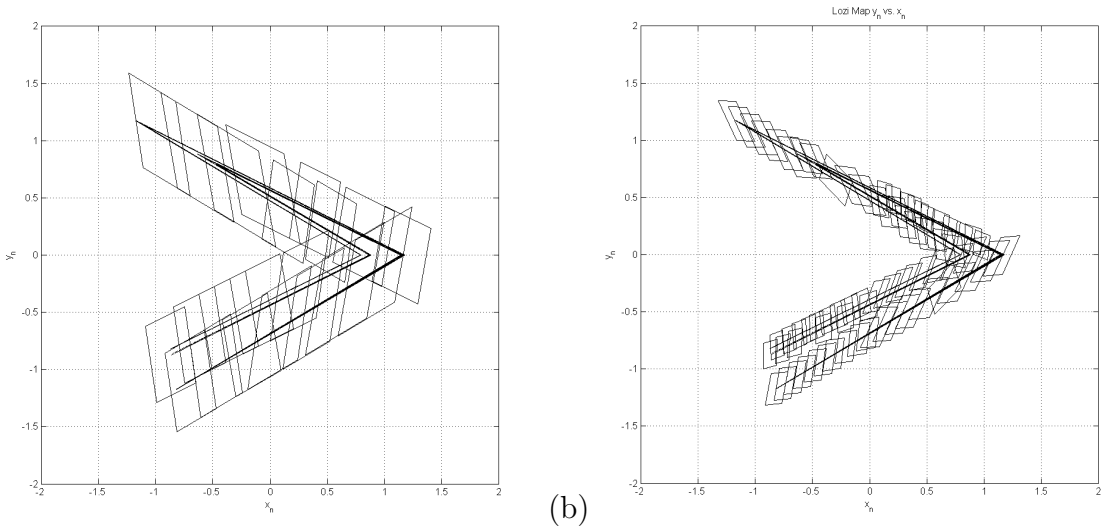

Figure 4: Optimum partition of the Lozi map with the noise strength given by (a) $\Delta = 0.1\ \mathbf{1}$ and (b) $\Delta = 0.01\ \mathbf{1}$.

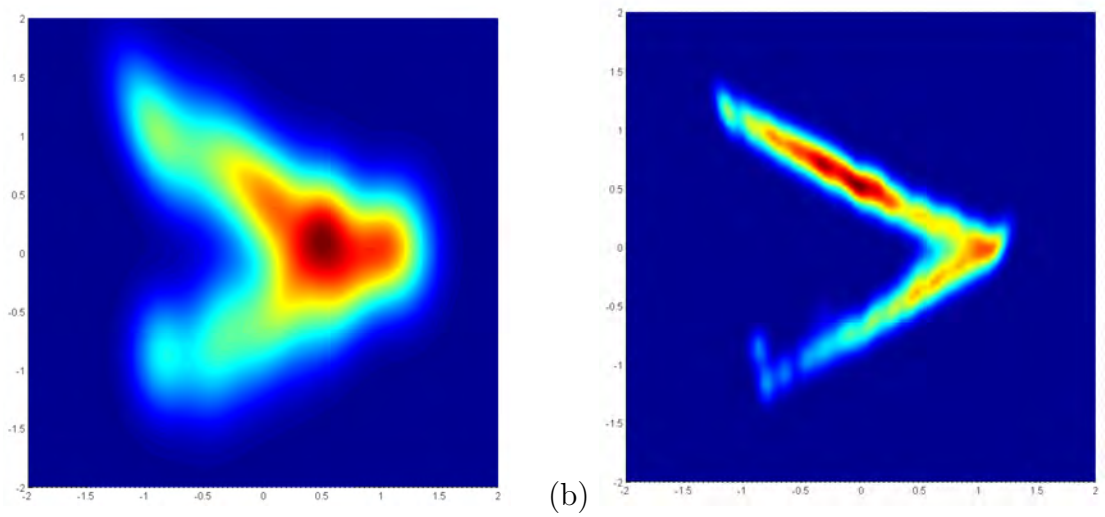

Figure 5: The global stationary distribution of the Lozi map calculated on our partition. The strength of the noise is different in the two images. In (a), $\Delta = 0.1\ \mathbf{1}$. In (b), $\Delta = 0.01\ \mathbf{1}$. The red regions are the regions where the probability of finding the particle is largest. The dark blue regions are the regions where the probability of finding the particle is almost zero.

The problems arise in our techniques because the Lozi map is not globally attracting. The stable manifold of the fixed point in the lower left corner of the attractor divides two basins of attraction. On one side of the manifold, the dynamics approach the deterministic attractor that we have been investigating. On the other side of the manifold, any initial condition approaches $(-\infty, -\infty)$.

The deterministic attractor is unaffected by the other basis of attraction. However, the noise can bump a trajectory across the stable manifold of the fixed point. There is thus a nonzero probability of escaping from the noisy attractor. Although the deterministic distribution is attracting, the noisy distribution is a repeller.

The leading eigenfunction of the one time step Fokker-Planck operator is thus less than one because the global stationary distribution does not conserve probability. In future work, we will use the technique of section 4.5 to determine both the escape rate from the repeller and the corresponding distribution. This should match more closely with the distribution observed in figure 6.

5.6 Brute Force Method

In order to assess the accuracy of the above technique, we compare it to a direct numerical calculation of the global stationary distribution of the Lozi attractor. The global stationary distribution, or natural measure, of a typical chaotic map is difficult to calculate numerically because it is everywhere singular. The noise smears out the singularities of the attractor and makes it amenable to numerical calculations.

We compare our results against a purely numerical calculation of the global stationary distribution of the noisy Lozi map. The algorithm proceeds a follows:

1. Choose the arbitrary initial condition (0.1,0.1).

2. Iterate the Lozi map for 100 iterations to remove transient behavior. This brings the trajectory close to the attractor, so we can begin the simulation of the noisy Lozi map.

3. Iterate the Lozi map for 10^8 time steps. At each time step, add a random number distributed according to a Gaussian with diffusion matrix Δ. Record the location of each point along the trajectory.

4. If the noise happens to kick the trajectory off of the attractor, choose another random initial condition, remove the transients again, and then continue iterating the noisy Lozi map.

5. Introduce an external finite resolution to the state space. Place square bins of side length 0.01 uniformly across the region of the state space containing the attractor. For the Lozi map, this is the box $[-1.5, 1.5] \times [-1.5, 1.5]$.

6. Count the number of points in each bin.

7. Normalize the distribution so the total sum over the entire attractor is one.

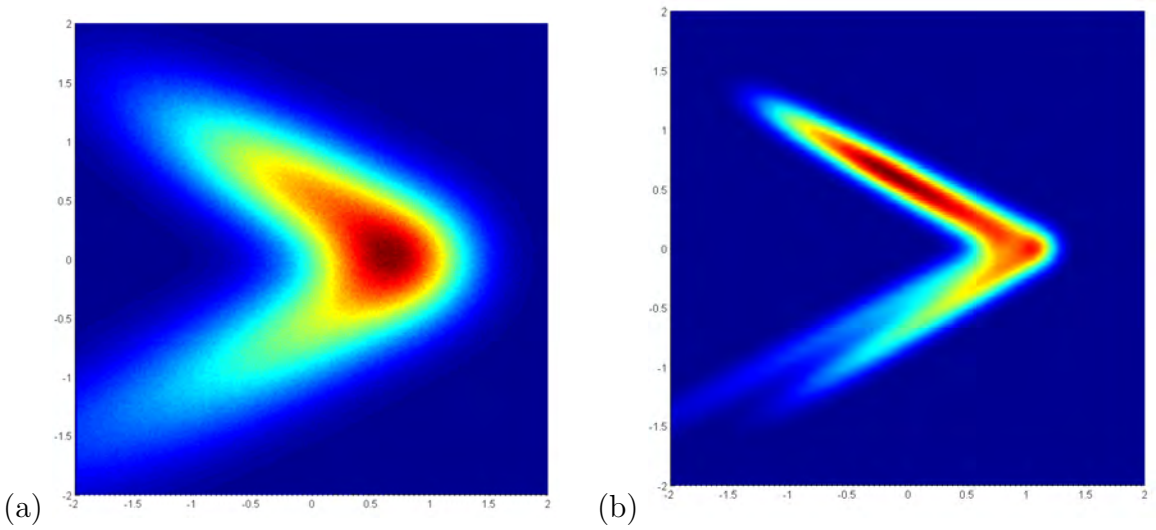

Figure 6: The global stationary distribution of the Lozi map calculated using brute force. The strength of the noise is different in the two images. In (a), $\Delta = 0.1\ \mathbf{1}$. In (b), $\Delta = 0.01\ \mathbf{1}$. The red regions are the regions where the probability of finding the particle is largest. The dark blue regions are the regions where the probability of finding the particle is almost zero.

8. Plot the resulting distribution.

Since the system is ergodic, after a long time, any trajectory will be distributed according to the global stationary distribution. This provides a direct technique for calculating the global stationary distribution of the Lozi map. The resulting distributions are shown in figure 6.

This technique is not easily generalized to systems which much larger dimensions. The number of bins that you need to use to cover the attractor increases as the exponential of the dimension of the state space. For two dimensions, this is not a problem. However, for models of fluid mechanics and other continuous systems which are typically described using state spaces with thousands of dimensions this brute force method is unfeasible.

Our technique of using a local description of the noise to partition the state space of a chaotic map depends primarily on the number of periodic orbits used to cover the attractor. Although it is more difficult to find periodic orbits in other systems, once the periodic orbits are found, the procedure described here is the same.

Considering the noise also gives a condition for when to stop looking for more periodic orbits. In a completely deterministic system, all of the periodic orbits are needed in order to calculate long-time averages of the observables of the state. The noise defines neighborhoods of each periodic point. Once these neighborhoods begin overlapping, finding more neighborhoods in that region will not improve the description of the dynamics. We only need to look for periodic orbits until their neighborhoods completely cover the attractor for the system. The presence of noise gives a condition that allows us to get a complete understanding of a chaotic system after finite computation time.

REFERENCES

[1] Bollt, E., Góra, P., Ostruszka, A., and Życzkowski, K., "Basis Markov partitions and transition matrices for stochastic systems," *SIAM J. Appl. Dyn. Syst.*, vol. 7, pp. 341–360, 2008.

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

[3] Cvitanović, P., Artuso, R., Mainieri, R., Tanner, G., and Vattay, G., *Chaos: Classical and Quantum*. Copenhagen: Niels Bohr Institute, 2016.

[4] Cvitanović, P. and Lippolis, D., "Knowing when to stop: How noise frees us from determinism," in *Let's Face Chaos through Nonlinear Dynamics* (Robnik, M. and Romanovski, V. G., eds.), (Melville, NY), pp. 82–126, American Institute of Physics, 2012.

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

[6] Lippolis, D., *How well can one resolve the state space of a chaotic map?* PhD thesis, School of Physics, Georgia Inst. of Technology, Atlanta, 2010.

[7] Lippolis, D. and Cvitanović, P., "How well can one resolve the state space of a chaotic map?," *Phys. Rev. Lett.*, vol. 104, p. 014101, 2010.

[8] Lyapunov, A. M., "The general problem of the stability of motion," *Int. J. Control*, vol. 55, pp. 531–534, 1992.

[9] Misiurewicz, M., "Strange attractors for the Lozi mappings," *Ann. New York Acad. Sci.*, vol. 357, pp. 348–358, 1980.

[10] Ostrowski, A. and Schneider, H., "Some theorems on the inertia of general matrices," *J. Math. Anal. Appl.*, vol. 4, pp. 72–84, 1962.

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

[12] Varga, A., "On balancing and order reduction of unstable periodic systems," in *Periodic Control Systems 2001* (Bittanti, S. and Colaneri, P., eds.), (New York), pp. 177–182, IFAC Workshop, Elsevier, 2001.

[13] Zhou, K., Salomon, G., and Wu, E., "Balanced realization and model reduction for unstable systems," *Intl. J. Robust Nonlin. Control*, vol. 9, pp. 183–198, 1999.