LBNL-53522

ERNEST ORLANDO LAWRENCE BERKELEY NATIONAL LABORATORY

Random Perturbations of SRB Measures and Numerical Studies of Chaotic Dynamics

Kevin Kwei-Yu Lin

Computing Sciences Directorate

July 2003

Ph.D. Thesis

REFERENCE COPY
Does Not
Circulate

LBNL-53522
Copy 1

Library Annex Reference
Lawrence Berkeley National Laboratory

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBNL-53522

Random Perturbations of SRB Measures and Numerical Studies of Chaotic Dynamics

Kevin Kwei-Yu Lin
Ph.D. Thesis

Mathematics Department
University of California, Berkeley

and

Computing Sciences Directorate
Ernest Orlando Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720

July 2003

This work was supported by the Director, Office of Science, Computational and Technology Research, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Random Perturbations of SRB Measures and Numerical Studies of Chaotic Dynamics

Copyright © 2003

by

Kevin Kwei-Yu Lin

The U.S. Department of Energy has the right to use this document
for any purpose whatsoever including the right to reproduce
all or any part thereof.

Random Perturbations of SRB Measures and Numerical Studies of Chaotic Dynamics

by

Kevin Kwei-Yu Lin

B.S. (Massachusetts Institute of Technology) 1996
B.S. (Massachusetts Institute of Technology) 1997
M.Eng. (Massachusetts Institute of Technology) 1997

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Mathematics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
Professor Alexandre J. Chorin, Chair
Professor Ole H. Hald
Professor Eugene D. Commins

Fall 2003

Abstract

Random Perturbations of SRB Measures and Numerical Studies of Chaotic Dynamics

by

Kevin Kwei-Yu Lin
Doctor of Philosophy in Mathematics

University of California, Berkeley

Professor Alexandre J. Chorin, Chair

Chaotic behavior occurs naturally in a variety of physical situations governed by deterministic equations of motion. Deterministic chaos is characterized by the exponential separation of nearby initial conditions. Thus chaotic systems are intrinsically unstable, and this intrinsic instability makes the equations of motion computationally intractable over long times. In contrast, the frequencies with which a solution visits different states is generally stable under small perturbations of the solution and of the equations of motion, so the corresponding statistical averages can be extracted from long time numerical simulations.

This thesis proposes a number of simple algorithms which compute statistical averages of observables by adding noise to deterministic chaotic systems. The results of this study show that the addition of noise can be beneficial in numerical studies of the statistics of chaotic systems: in addition to covering up numerical artifacts which arise from round-off errors, a moderate amount of noise can help accelerate the convergence of computed time averages of observable quantities. A simple scaling argument is used to derive a rough error estimate and the unique ergodicity of the perturbed process is proved. The effect of noise on the statistical properties of the Kuramoto-Sivashinsky equation, in particular the Lyapunov exponents and the statistical dependence of Fourier modes, is also studied numerically.

To Mom, Dad, Holly, and Tanya.

Contents

List of Figures iv

List of Tables viii

1 Introduction 1
- 1.1 Examples of random perturbations . . . 2
- 1.2 Why chaotic systems? . . . 8
- 1.3 Motivation and related work . . . 8

2 Review of Hyperbolic Dynamics and Ergodic Theory 10
- 2.1 Basic theory of dynamical systems . . . 11
  - 2.1.1 Flows and maps . . . 11
  - 2.1.2 Invariant measures . . . 13
- 2.2 Hyperbolic dynamics . . . 15
  - 2.2.1 Stable manifold theorem . . . 17
  - 2.2.2 Other models of dynamical systems . . . 20
- 2.3 Lyapunov exponents & entropy . . . 21
  - 2.3.1 Entropy and information theory . . . 23
  - 2.3.2 Kolmogorov-Sinai entropy . . . 24
  - 2.3.3 Markov partitions and symbolic dynamics . . . 25
- 2.4 Sinai-Ruelle-Bowen measures . . . 27
- 2.5 Transfer operators . . . 32
  - 2.5.1 Finite state space Markov chains . . . 32
  - 2.5.2 Monte Carlo algorithms . . . 34
  - 2.5.3 Transfer operators of dynamical systems . . . 37
- 2.6 Zero-noise limits of uniformly hyperbolic systems . . . 39
- 2.7 The limits of hyperbolic theory . . . 41

3 Random Perturbations in Numerical Simulations 44
- 3.1 Asymptotically unstable subspaces & Lyapunov exponents . . . 44
- 3.2 Time averaging . . . 48
- 3.3 Random kicks in unstable directions . . . 49
- 3.4 General random kicks . . . 62
- 3.5 An intermittent interval map . . . 70

3.6 Scaling and white noise 75

4 Convergence and Related Problems **78**
4.1 Rate of mixing in Anosov systems 79
4.2 Error estimates 80
4.3 Unique ergodicity 83

5 The Kuramoto-Sivashinsky Equation **89**
5.1 Properties of the KSE 89
5.2 Parameters & initial conditions 92
5.3 Lyapunov exponents: comparison with existing results 93
5.4 Mode-mode mutual information 102

6 Future Work **110**

Bibliography **112**

List of Figures

1.1 Two ways of computing the solenoid attractor. 3
1.2 The Lorenz attractor. 4
1.3 The autocovariance functions $C_{1,1}(t)$, $C_{2,2}(t)$, and $C_{3,3}(t)$ for the Lorenz system (1.2). Two of the curves ($C_{1,1}$ and $C_{2,2}$) practically coincide and cannot be distinguished in this plot. The third curve $C_{3,3}$ decays much more slowly than the other two. 5
1.4 The expectation value $\int x_3 \, d\mu_\epsilon(x)$ as a function of the noise amplitude ϵ. . . 6
1.5 Autocovariance function $C_{3,3}(t)$ for the Lorenz flow, for ϵ ranging from 0 (slowest-decaying curve) to 4.3 (fastest-decaying curve). 6
1.6 The spectral power densities of the covariance functions shown in Figure 1.5. The spectral densities become more smooth as the noise amplitude ϵ increases. 7
1.7 The exponential decay times for each component of the randomly-perturbed Lorenz flow as functions of the noise amplitude ϵ. 7

2.1 Exponential separation of nearby trajectories in the Lorenz flow. 12
2.2 The section $K_{\text{Lorenz}} \cap P(z = 24)$ of the Lorenz flow. 14
2.3 The construction of holonomy maps. 19
2.4 A (generator for a) Markov partition of the cat map (2.3). 26
2.5 The "figure 8" attractor. 30
2.6 The Hénon attractor. 42

3.1 Estimates of Lyapunov exponents as a function of simulation time interval T. 50
3.2 Estimates of Lyapunov exponents as a function of simulation time interval T. 53
3.3 Comparison of estimated Lyapunov exponents as functions of simulation time interval. The solid black line represents the results computed with $\epsilon = 0$, while the dashed blue line represents the results with $\epsilon = 0.01$. 54
3.4 Estimated Lyapunov exponents for the Lorenz flow kicked by uniform noise of amplitude ϵ in the asymptotically unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$. 55
3.5 Decay times of the Lorenz flow periodically kicked by uniform noise of amplitude ϵ in the asymptotically unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$. 56

3.6 Integrated correlation times of the Lorenz flow periodically kicked by uniform noise of amplitude ϵ in the asymptotically unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$. 56
3.7 Spectral power densities $\log_{10}(\hat{C}(f))$ associated with $\hat{\lambda}_3$ for the Lorenz flow, plotted against the frequency f, for $\epsilon \in \{0.0, 1.0, 4.0\}$. 57
3.8 Variances of $\hat{\lambda}_k$ for the Lorenz flow periodically kicked by uniform noise of amplitude ϵ in the asymptotically unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$. 58
3.9 Approximate errors for the estimated Lyapunov exponents of the Lorenz flow periodically kicked by uniform noise of amplitude ϵ in the asymptotically unstable direction. The solid curve is associated with $\hat{\lambda}_1$, the dashed curve with $\hat{\lambda}_2$, and the dashed-dot curve with $\hat{\lambda}_3$. 58
3.10 Estimated Lyapunov exponents for the Lorenz flow kicked by perturbations which are uniformly distributed over cubes of side length ϵ. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$. 62
3.11 Decay times of the Lorenz flow periodically kicked by perturbations which are uniformly distributed over cubes of side length ϵ. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$. 63
3.12 Integrated correlation times of the Lorenz flow periodically kicked by perturbations which are uniformly distributed over cubes of side length ϵ. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$. 63
3.13 Variances of $\hat{\lambda}_k$ for the Lorenz flow periodically kicked by perturbations which are uniformly distributed over cubes of side length ϵ. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$. 64
3.14 Approximate errors for the estimated Lyapunov exponents of the Lorenz flow periodically kicked by perturbations which are uniformly distributed over cubes of side length ϵ. The solid curve is associated with $\hat{\lambda}_1$, the dashed curve with $\hat{\lambda}_2$, and the dashed-dot curve with $\hat{\lambda}_3$. 64
3.15 Estimated Lyapunov exponents for the Lorenz flow kicked by degenerate perturbations of amplitude ϵ in a linearly unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$. 66
3.16 Decay times of the Lorenz flow periodically kicked by degenerate perturbations of amplitude ϵ in a linearly unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$. 67
3.17 Integrated correlation times of the Lorenz flow periodically kicked by degenerate perturbations of amplitude ϵ in a linearly unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$. 67
3.18 Variances of $\hat{\lambda}_k$ for the Lorenz flow periodically kicked by degenerate perturbations of amplitude ϵ in a linearly unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$. 68
3.19 Approximate errors for the estimated Lyapunov exponents of the Lorenz flow periodically kicked by degenerate perturbations of amplitude ϵ in a linearly unstable direction. The solid curve is associated with $\hat{\lambda}_1$, the dashed curve with $\hat{\lambda}_2$, and the dashed-dot curve with $\hat{\lambda}_3$. 68

3.20 The intermittent map G_α for $\alpha = 0.999$. 71
3.21 Estimated Lyapunov exponents for the intermittent interval map with $\alpha = 0.49$ kicked by uniform noise of amplitude ϵ. 72
3.22 Decay times of the intermittent interval map with $\alpha = 0.49$ periodically kicked by uniform noise of amplitude ϵ. 72
3.23 Integrated correlation times of the intermittent interval map with $\alpha = 0.49$ periodically kicked by uniform noise of amplitude ϵ. 73
3.24 Variances of $\hat{\lambda}_k$ for the intermittent interval map with $\alpha = 0.49$ periodically kicked by degenerate perturbations of amplitude ϵ. 73
3.25 Approximate errors for the estimated Lyapunov exponents of the intermittent interval map with $\alpha = 0.49$ periodically kicked by uniform noise of amplitude ϵ. 74
3.26 Estimates of Lyapunov exponents as functions of ϵ, for various values of T. In the second column, the data points marked by squares come from solving the SDE (3.8). 76
3.27 Exponential decay times for the autocovariance functions of $\hat{\lambda}_k$. Again, the data points marked by squares are computed from the SDE (3.8). 77

5.1 The exponential separation of nearby initial conditions in the Kuramoto-Sivashinsky equation, for $\alpha = 1$ and $\beta = 0.085$. 92
5.2 Estimated Lyapunov exponents for the KS flow kicked by gaussian noise of variance ϵ in the asymptotically unstable direction. 94
5.3 Decay times of the KS flow periodically kicked by gaussian noise of variance ϵ in the asymptotically unstable direction. 94
5.4 Integrated correlation times of the KS flow periodically kicked by gaussian noise of variance ϵ in the asymptotically unstable direction. 95
5.5 Variances of the local expansion rates for the KS flow periodically kicked by gaussian noise of variance ϵ in the asymptotically unstable direction. 95
5.6 Approximate errors for the estimated Lyapunov exponents of the KS flow periodically kicked by gaussian noise of variance ϵ in the asymptotically unstable direction. 96
5.7 Autocovariance functions for the local expansion rates of the unperturbed KS flow. 96
5.8 Autocovariance functions for the local expansion rates of the KSE. The flow is periodically kicked in asymptotically unstable directions by gaussian random variables of variance $\epsilon = 0.005$. 100
5.9 Autocovariance functions for the local expansion rates of the KSE. The flow is periodically kicked in asymptotically unstable directions by gaussian random variables of variance $\epsilon = 0.01$. 100
5.10 Spectral power densities $\hat{C}(f)$ associated with the largest estimated Lyapunov exponent $\hat{\lambda}_1$ of the KS flow, plotted against the frequency f, for $\epsilon \in \{0.0, 0.005, 0.01\}$. 101
5.11 Some autocovariance functions of the histograms used in the mutual information calculation. The latter is quite noisy but does show faster initial decay. 107

5.12 Spectral power densities $\log_{10}(\hat{C}(f))$ for the correlation functions shown in Figure 5.11. 108

List of Tables

3.1 Correlation times of the estimated Lyapunov exponents of the Lorenz flow periodically kicked by uniform noise of amplitude ϵ in the asymptotically unstable direction. Estimated Lyapunov exponents are stated with their standard deviations. 61

3.2 Correlation times of the estimated Lyapunov exponents of the Lorenz flow periodically kicked by perturbations which are uniformly distributed over cubes of side length ϵ. Estimated Lyapunov exponents are stated with their standard deviations. 65

3.3 Correlation times of the estimated Lyapunov exponents of the Lorenz flow periodically kicked by degenerate perturbations of amplitude ϵ in a linearly unstable direction. Estimated Lyapunov exponents are stated with their standard deviations. 69

3.4 Correlation times of the estimated Lyapunov exponents of the intermittent interval map with $\alpha = 0.49$ periodically kicked by uniform noise of amplitude ϵ. Estimated Lyapunov exponents are stated with their standard deviations. 74

5.1 Correlation times of the estimated Lyapunov exponents of the KS flow periodically kicked by gaussian noise of variance ϵ in the asymptotically unstable direction. Estimated Lyapunov exponents are stated with their standard deviations. 97

5.2 Correlation times of the estimated Lyapunov exponents of the KS flow periodically kicked by gaussian noise of variance ϵ in the asymptotically unstable direction. Estimated Lyapunov exponents are stated with their standard deviations. 98

5.3 Correlation times of the estimated Lyapunov exponents of the KS flow periodically kicked by gaussian noise of variance ϵ in the asymptotically unstable direction. Estimated Lyapunov exponents are stated with their standard deviations. 99

5.4 The coarse-grained entropy for each mode of the KSE flow. 103

5.5 Error estimates for Table 5.4. 104

5.6 The coarse-grained mutual information the modes of the unperturbed KS flow. 104

5.7 Error estimates for Table 5.6. 105
5.8 The coarse-grained mutual information the modes of the KS flow at noise level $\epsilon = 0.01$. 105
5.9 Error estimates for Table 5.8. 106
5.10 Partitions used for this calculation. 106

Acknowledgments

I am grateful to the Fannie and John Hertz Foundation for its generous financial support over the last four years, and to Dr. Tom Weaver for his insightful comments during our meetings. I am also grateful to Professors Predrag Cvitanović, Charles Pugh, Michael Shub, Lai-Sang Young, Stéphane Nonnenmacher, and Viviane Baladi for helping me understand hyperbolic dynamics and for providing valuable references, and to Professors Fraydoun Rezakhanlou and Craig Evans for teaching me the rudiments of large deviations theory. I am fortunate to have met Professors Eugene Commins, Eyvind Wichmann, and G. I. Barenblatt at Berkeley, and to have had the opportunity to participate in their inspiring lectures.

I learned about the coupling method from Jonathan Mattingly, both from personal conversations and from his papers on applications of the technique to far more complex situations. What little I know about resonances, both quantum and classical, have come from interaction with Maciej Zworski. Many other friends, including Eugene Ingerman, Mayya Tokman, Panos Stinis, Yelena Shvets, Daniel Markiewicz, Laura Matusevich, and Michael Anshelevich, put up with my incoherent attempts to explain this material to them and to myself. Also, early drafts of this thesis benefited greatly from conversations with Professor Gerald Jay Sussman. To all of them I owe many thanks.

Jackie Scoggins at Lawrence Berkeley National Lab deserves a special Thank You: computing at LBL would not be the same without her diligence and dedication. I would also like to thank Valerie Heatlie for helping me navigate the bureaucracy of LBL over the last few years, and to thank Carol Morrison for her help over the past year.

I owe a great debt of gratitude to Ole Hald for not letting me get away with sloppy thinking and sloppier writing. While I still managed to make a few sections incomprehensible, the exposition would have been far less clear without Ole's help. And I am very grateful that I had the opportunity to study under the guidance of Alexandre Chorin: his patience, generosity, willingness to endure countless questions, and deep insight into (and love for) statistical physics and computation have taught me so much. I have made quite a few missteps over the last few years, and Alexandre has always been there to help me recover and learn from the process.

Finally, to my wife Tanya and to all of my family: I could not have arrived at this point without your support and understanding.

Chapter 1

Introduction

Chaotic behavior occurs naturally in a variety of physical situations governed by deterministic equations of motion, ranging from the dynamics of molecules [Gas98] to the motion of planets [SW92]. Chaos in deterministic systems is characterized by the exponential separation of solutions with nearby initial conditions. Thus chaotic systems are intrinsically unstable, and this intrinsic instability makes the equations of motion computationally intractable over moderately long time scales because of the truncation and round-off errors inherent in floating point calculations. In contrast, the frequencies with which a solution visits different states is generally stable under small perturbations of the solution and of the equations of motion, so these frequencies and the statistical averages of observables can be extracted from long time numerical simulations.

This thesis proposes a number of simple algorithms which compute statistical averages of observables by simulating random perturbations of deterministic chaotic systems. The results of this study show that the addition of noise can be beneficial in numerical studies of the statistics of chaotic systems: in addition to covering up numerical artifacts which arise from round-off errors, a moderate amount of noise can help accelerate the convergence of computed time averages of observable quantities.

This chapter provides an informal overview of the main ideas in this thesis. Precise definitions can be found in Chapter 2. Chapter 3 introduces the random perturbation algorithm for dissipative chaotic systems, providing a discussion (based on explicit numerical examples) of algorithmic issues. Chapter 4 discusses error estimates, and Chapter 5 applies these techniques to an infinite dimensional system, the Kuramoto-Sivashinsky equation.

1.1 Examples of random perturbations

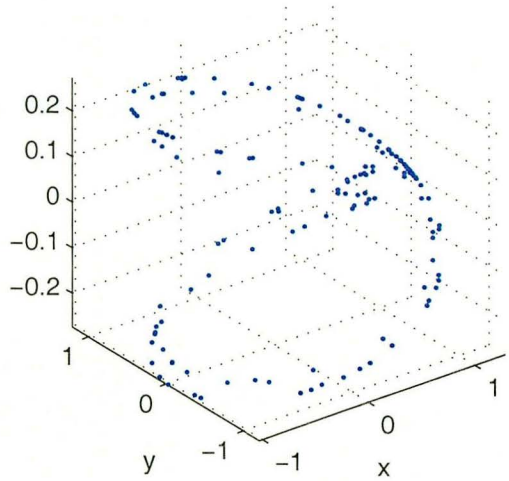
Let us begin by considering the map

$$F(\theta, r, s) = \begin{bmatrix} 2\theta \pmod{2\pi} \\ \epsilon_1 \cos(\theta) + \epsilon_2 r \\ \epsilon_1 \sin(\theta) + \epsilon_2 s \end{bmatrix} \tag{1.1}$$

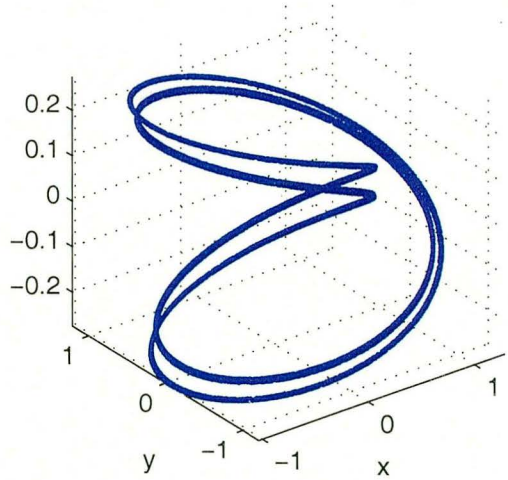
and its iterates on the domain $X = \{(\theta, r, s) : 0 \leq \theta < 2\pi, r^2 + s^2 \leq 1\}$. The map F is usually called the "solenoid map" for a reason which will become clear. It is easy to check that when $\theta \neq \theta'$, the n-fold iterates $F^n(\theta, r, s)$ and $F^n(\theta', r', s')$ will separate exponentially until the distance between them becomes ~ 1 regardless of the values of r, r', s, and s'. On the other hand, if $\theta = \theta'$ then $F^n(\theta, r, s)$ and $F^n(\theta', r', s')$ will converge exponentially, again regardless of r, r', s, and s'. Thus, while most pairs of initial conditions will separate exponentially fast, some special initial conditions will instead converge exponentially. It is natural to define the stable manifold of (θ, r, s) to be the set $\{(\theta', r', s') : \theta' = \theta\}$ and to call the directions transverse to the stable manifolds the unstable directions. (See Chapter 2 for precise definitions.)

It is not difficult to prove that this map has an "attractor": its domain X contains a compact subset K to which all trajectories converge as $n \to \infty$. Furthermore, it is easy to show that this attractor has dimension ≥ 1. However, a naive numerical implementation of the solenoid map yields Figure 1.1(a). A closer examination of the resulting numerical data shows that all numerically generated iterates of the solenoid map converge to a unique fixed point though the system has no stable fixed points. This problem is easy to understand: The first coordinate of the solenoid map simply iterates the one-dimensional map $\theta \mapsto 2\theta \pmod{2\pi}$. Clearly, in finite precision arithmetic, iterating this map will eventually produce a sequence of zeros.

There is an easy solution to this problem: At every step, add a small amount of noise to the angle θ. The resulting picture is shown in Figure 1.1(b). The addition of a small amount of noise helps cover up the fact that digital computers can only manipulate a finite amount of information. Moreover, this simple example points out another important property of chaotic systems: noise added in the direction of the stable manifold is simply damped out, while noise injected in the unstable directions is fed back into the system, making more efficient use of randomness. It is therefore not necessary to add noise in



(a)



(b)

Figure 1.1: Two ways of computing the solenoid attractor.

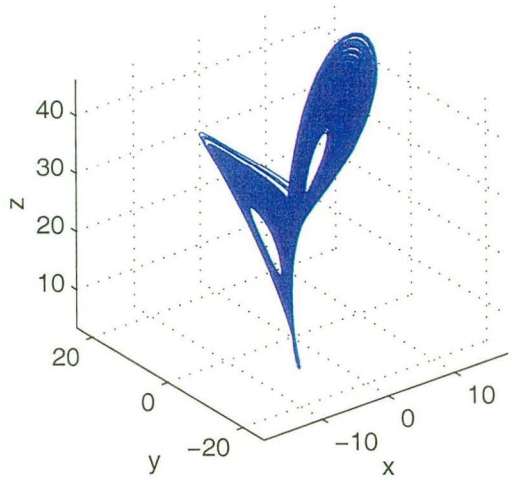


Figure 1.2: The Lorenz attractor.

all directions. This observation becomes important when we consider dynamical systems defined by partial differential equations, as the addition of noise may impose constraints on the efficiency of numerical integrators.

A second reason for adding noise to a dynamical system can be illustrated by the Lorenz equations:

$$\begin{aligned} \dot{x}_1 &= \sigma (x_2 - x_1) \\ \dot{x}_2 &= r x_1 - x_2 - x_1 x_3 \\ \dot{x}_3 &= -b x_3 + x_1 x_2 \end{aligned} \quad . \tag{1.2}$$

Let us assume, for the moment, that $\sigma = 10$, $r = 28$, and $b = \frac{8}{3}$. It has long been observed [Lor63] that with these parameters, almost every initial condition generates trajectories converging to the "strange attractor" in Figure 1.2). Recent work of Tucker [Tuc99] has shown decisively that the Lorenz attractor really exists and is not a numerical artifact.

As will be explained in Chapter 2, the existence of a robust compact attractor corresponds to the existence of a well-defined statistical steady state (invariant probability measures). To extract statistical information from the equations of motion, it is natural to perform long time numerical simulations and compute time averages of the results. The decay rate of autocovariance functions plays a fundamental role in determining the

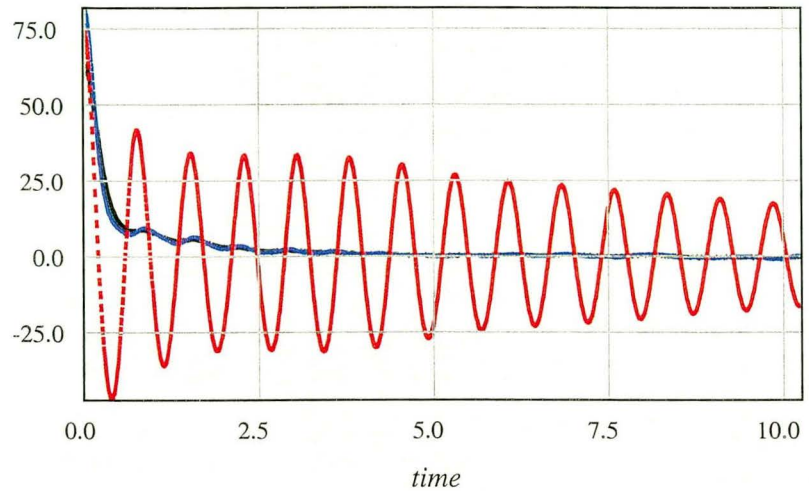


Figure 1.3: The autocovariance functions $C_{1,1}(t)$, $C_{2,2}(t)$, and $C_{3,3}(t)$ for the Lorenz system (1.2). Two of the curves ($C_{1,1}$ and $C_{2,2}$) practically coincide and cannot be distinguished in this plot. The third curve $C_{3,3}$ decays much more slowly than the other two.

accuracy of this procedure. Roughly speaking, a faster decay rate means that one can obtain more "effectively independent" samples per unit of computation time. (See §2.5.2.)

The autocovariance functions of the coordinate maps of the Lorenz system are shown in Figure 1.3. One can see that correlations decay much more slowly along one axis than along the others. This is because vertical motions along the attractor occur on a fundamentally slower time scale than motion in the other two directions. This slower rate of decay limits the efficiency of time average computations. In high- and infinite-dimensional systems, this problem is expected to be worse.

Suppose now we can estimate the unstable directions for the Lorenz flow, and periodically kick the system in such directions with a uniformly-distributed noise of amplitude ϵ. The corresponding average values of the x_3 coordinate is shown in Figure 1.4. Clearly, the expectation value of the third coordinate x_3 in the perturbed system converges smoothly to the expectation value in the unperturbed system as $\epsilon \to 0$. More importantly, the correlation decay times are affected by the perturbations as well, as shown in Figures

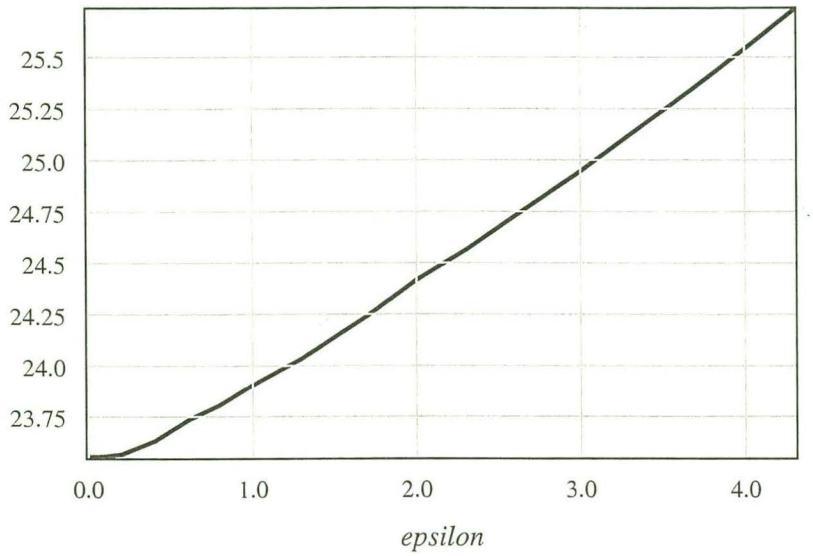


Figure 1.4: The expectation value $\int x_3 \, d\mu_\epsilon(x)$ as a function of the noise amplitude ϵ.

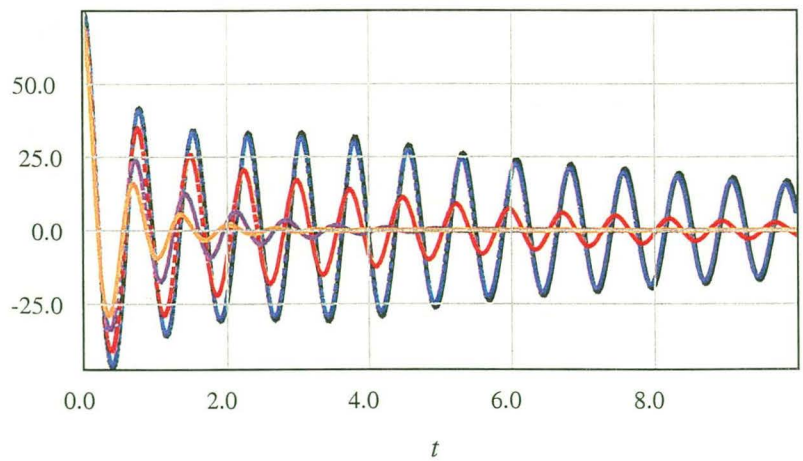


Figure 1.5: Autocovariance function $C_{3,3}(t)$ for the Lorenz flow, for ϵ ranging from 0 (slowest-decaying curve) to 4.3 (fastest-decaying curve).

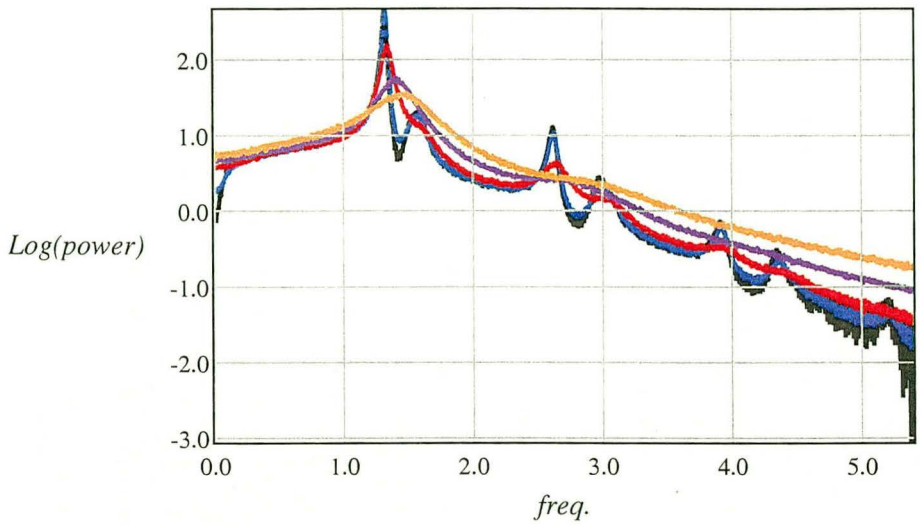


Figure 1.6: The spectral power densities of the covariance functions shown in Figure 1.5. The spectral densities become more smooth as the noise amplitude ϵ increases.

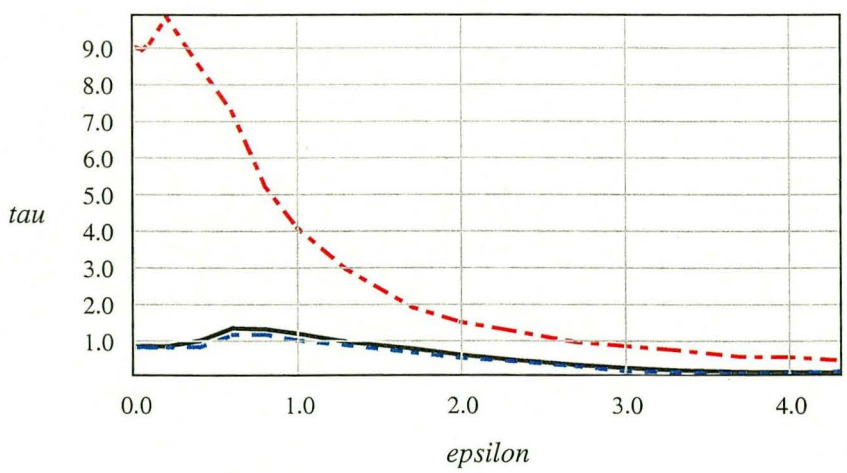


Figure 1.7: The exponential decay times for each component of the randomly-perturbed Lorenz flow as functions of the noise amplitude ϵ.

1.5 and 1.7. These results suggest that one can decrease the correlation times of a chaotic system with the introduction of noise, while incurring only small and controllable errors in the desired expectation values.

1.2 Why chaotic systems?

The idea of using random perturbations to reduce correlation times is very natural and is apparently applicable to a wide range of deterministic and stochastic dynamical systems and their invariant measures. Nevertheless, the effectiveness of random perturbations in numerical simulations depends strongly on the dynamical system and the invariant measure under study.

For example, the Langevin equation

$$\begin{aligned} dx(t) &= -ax(t) + \epsilon dw(t), \\ x(0) &\in N(0, \epsilon) \end{aligned} \tag{1.3}$$

has a unique invariant measure: it is the standard distribution with variance ϵ. This Itô stochastic differential equation defines a stationary Markov process whose autocorrelation function is

$$C(t) = \epsilon \cdot e^{-a|t|}. \tag{1.4}$$

This clearly has a decay rate independent of the noise amplitude ϵ.

Remark 1.1. If we think of the Langevin equation as an effective equation of motion for finite-dimensional system coupled to an infinite-dimensional heat bath, then the decay rate a and the noise size ϵ are related through the Einstein-Smoluchowski relations, and an increase in noise is accompanied by an increase in decay rate.

1.3 Motivation and related work

As stated above, one motivation for computing statistical properties of chaotic dynamical systems is their sensitive dependence on initial conditions: only statistical properties are effectively computable. Familiar geometric structures associated with chaotic dynamics, such as strange attractors and invariant manifolds, are naturally related to the statistical description of dynamical systems as well.

A second, longer-range motivation is that many physical problems involve an extremely large number of degrees of freedom, only a small subset of which is of real scientific interest. However, the "uninteresting" degrees of freedom may involve such small time and length scales as to make direct numerical simulations difficult. Thus, there is a great need for dimension-reduction techniques which can predict the dynamics of a system using only partial knowledge of its state. A simple approach is to compute the conditional expectation of the future behavior of a system using known initial data: this is the minimum mean-square error estimator and is optimal in the sense of least squares. This problem of optimal prediction involves the transient statistical behavior of physical systems and is thus closely related to nonequilibrium statistical mechanics [CHK02, Cho02]. It presents a far more challenging set of mathematical and computational problems than those outlined in the preceding paragraph, and a deeper understanding of statistical steady states is necessary for solving the problem of optimal prediction in dissipative systems. This is because such statistics supply the prior information needed for computing conditional expectations, and because the effectiveness of existing optimal prediction algorithms depends on the statistical structure of the dynamical system under study.

Chapter 2

Review of Hyperbolic Dynamics and Ergodic Theory

This chapter reviews some background material needed in the rest of the thesis. With the exception of §2.5.2, most of this material consists of standard results and definitions from dynamical systems theory, with a special focus on the statistical steady states of hyperbolic systems. For more balanced introductions to this subject, see Eckmann and Ruelle [ER85] and Young [You95]. See also the text by Katok and Hasselblatt [KH95] for a thorough development of the geometric theory of hyperbolic systems and Guckenheimer and Holmes [GH83] for applications of geometric methods to concrete examples. Section 2.5.2, which provides a quick introduction to Markov chain Monte Carlo methods, follows Sokal [Sok97] closely.

The following discussion assumes some familiarity with the basic concepts of differential geometry [Spi79] and measure theory [Fol84], in addition to some knowledge of elementary probability theory.

Some notes on notation:

1. Following standard practice in differential geometry, tangent vectors are identified with differential operators. Thus, a unit vector tangent to a submanifold of $\mathbb{R}^n$ in the direction of x_1 is written as $\frac{\partial}{\partial x_1}$.

2. Let X and Y be measure spaces, let μ be a measure on X, and let $F : X \to Y$ be a measurable map. Then the push-forward $F_*\mu$ of μ by F is the measure on Y given by $A \mapsto \mu(F^{-1}(A))$ for all measurable sets $A \subset Y$.

3. We will frequently need to consider the n-fold composition of a map F with itself. This is denoted by F^n.

2.1 Basic theory of dynamical systems

2.1.1 Flows and maps

Let X be a C^∞ manifold and f a smooth vector field on X. Consider the initial value problem

$$\begin{aligned} \dot{x}(t) &= f(x(t)), \\ x(0) &= x_0 \in X. \end{aligned} \tag{2.1}$$

Under suitable conditions, for example when X is compact and has no boundary, Equation (2.1) has a unique solution for all time $t \geq 0$. Equation (2.1) thus defines a map $\phi_t^f(x_0)$ via $\phi_t^f(x_0) = x(t)$. (The superscript f in ϕ_t^f will be dropped whenever no confusion is likely to arise.) The map $\phi_t^f(x_0)$ is a smooth function of (t, x_0) and satisfies $\phi_0^f(x_0) = x_0$. Replacing f with $-f$ shows that ϕ_t^f is defined for all $t \in \mathbb{R}$, that ϕ_t^f is invertible with smooth inverse, and that $(\phi_t^f)^{-1} = \phi_{-t}^f$. This family of diffeomorphisms forms a one-parameter group: $\phi_{t+s}^f = \phi_t^f \circ \phi_s^f$. The solution map ϕ_t^f is also called the time-t flow map, and the vector field f is said to generate the flow ϕ_t^f. A subset K of X is invariant if $\phi_t^f(K) = K$ for all $t \in \mathbb{R}$.

Dynamical systems theory focuses on geometric and topological structures induced by a flow ϕ_t on its phase space X rather than on the analytical properties of individual solutions. As an example, consider the Lorenz system (1.2) from the Introduction:

$$\begin{aligned} \dot{x}_1 &= \sigma(x_2 - x_1) \\ \dot{x}_2 &= rx_1 - x_2 - x_1x_3 \\ \dot{x}_3 &= -bx_3 + x_1x_2 \end{aligned}$$

For suitable parameter values like $\sigma = 10$, $r = 28$, and $b = \frac{8}{3}$, almost every initial condition produces a picture like Figure 1.2. Trajectories rapidly approach a butterfly-shaped compact set known as the Lorenz attractor, denoted here by K_{Lorenz}. The set K_{Lorenz} and a small neighborhood around it thus contain all the information relevant to the long-time behavior of the Lorenz system. On the other hand, Figure 2.1 demonstrates that nearby solutions of the Lorenz system tend to diverge exponentially fast in time (the vertical scale shows the natural logarithm of the distance). Thus, individual trajectories of the Lorenz system are unpredictable over moderately long time intervals, even though the existence

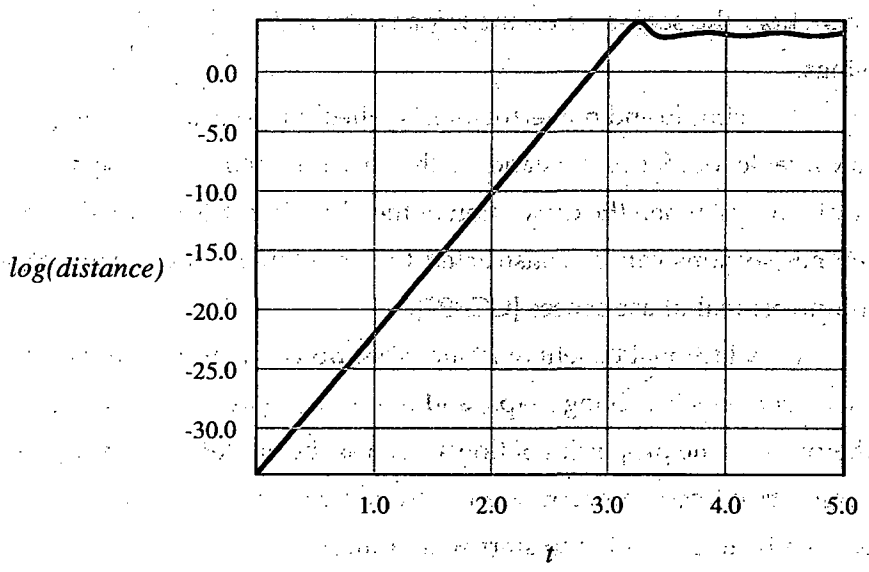


Figure 2.1: Exponential separation of nearby trajectories in the Lorenz flow.

of the attractor and an associated invariant probability measure (invariant measures are defined later in this chapter) constraints trajectories to visit regions of phase space with definite frequencies, facilitating the definition and computation of statistical properties of trajectories over long time intervals.

Let $P(z)$ denote the set

$$P(z) = \{(x_1, x_2, x_3) : x_3 = z\}.$$

If the constant z is chosen appropriately, $P(z)$ and K_{Lorenz} will intersect in a way which guarantees that almost every trajectory of the Lorenz flow intersects $P(z)$ infinitely many times (see Figure 2.2(a)). Poincaré realized that in such a situation, one can reduce the dimension of the dynamical system by recording only the intersections of individual trajectories with the plane $P(z)$. This defines a return map (or Poincaré map) $R_{f,z} : P(z) \to P(z)$ and a return time function $T_{f,z} : P(z) \to [0, +\infty]$, so that

$$T_{f,z}(x) = \inf\{t > 0 : \phi_t(x) \in P(z)\},$$
$$R_{f,z}(x) = \phi_{T_{f,z}(x)}(x).$$

The result is Figure 2.2(b). The two segments in the center represent downward crossings of the section; the segments on the left and the right side of the figure represent upward crossings.

This simple and powerful idea is called the Poincaré section [GH83, Ott00, SW01]. It is a valuable tool for understanding the structure of invariant sets. Poincaré sections are especially useful when the dimension of the phase space is relatively small (~ 3), although sometimes sections can be constructed for low-dimensional invariant sets embedded in high-dimensional phase spaces [CCP97].

An additional benefit of Poincaré sections is that they reduce problems involving flows to problems involving maps, and maps are sometimes easier to analyze than flows. Furthermore, some properties of flows, such as the existence of natural invariant measures and ergodicity (these concepts are developed below), can be deduced from the Poincaré map. Finally, it is possible to start with a map $F : X \to X$ and construct a manifold Y and a vector field f on Y such that X can be identified with a subset of Y and F arises as the Poincaré map of f. This construction is called a suspension. It is not needed here; interested readers can consult [KH95].

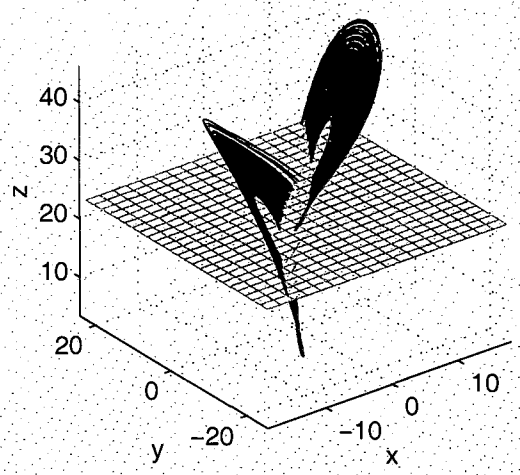
Another way to obtain maps from flows is by sampling a flow at periodic intervals. That is, simply fix a constant $T > 0$ and set $F = \phi_T$. Maps define discrete-time dynamical systems in the same way vector fields define continuous-time dynamical systems:

$$x_{n+1} = F(x_n).$$

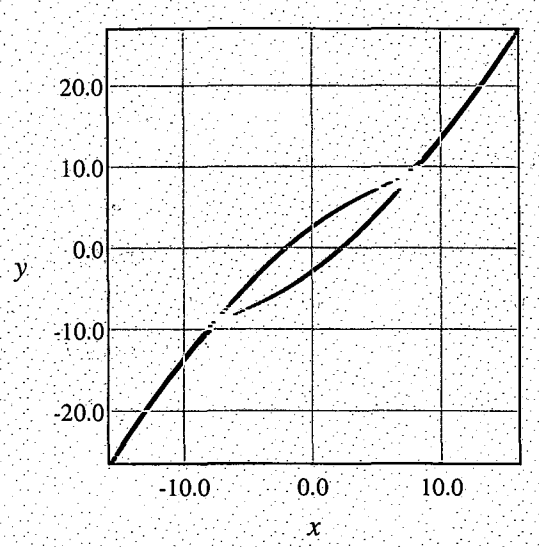
Let us denote the n-fold composition of F with itself by F^n, so that $x_n = F^n(x_0)$. Observe that random perturbations of the time-T map of a flow correspond to periodically "kicking" the flow by a random displacement. A large class of the random perturbations studied here fall into this category. For simplicity of notation, almost all of the remainder of this chapter concerns maps.

2.1.2 Invariant measures

Let μ be a Borel probability measure on the manifold X and suppose we are given a vector field f on X. The measure μ is invariant under ϕ_t if $\mu(\phi_t^{-1}(A)) = \mu(A)$ for all Borel sets $A \subset X$. Equivalently, one can say that the flow ϕ_t preserves the measure μ. If, instead of a flow, we have a map $F : X \to X$, then μ is said to be F-invariant if $\mu(F^{-1}(A)) = \mu(A)$



(a)



(b)

Figure 2.2: The section $K_{\text{Lorenz}} \cap P(z = 24)$ of the Lorenz flow.

for all Borel sets A. It is easy to check that the support of an invariant measure is an invariant set. The converse is not quite true: invariant sets do not always support invariant probability measures. Compact invariant sets, however, do support invariant measures. For example, suppose x_0 is a periodic point of F with period N, so that $F^N(x_0) = x_0$. Then the measure $\frac{1}{N}\sum_{k=0}^{N-1}\delta_{x_k}$ is an invariant measure of F. As a special case, every fixed point of F supports an invariant measure.

Recall that the space of probability measures with the total variation norm is isometrically isomorphic to the space of bounded linear functionals on the space $C_b(X)$ of bounded continuous functions with the sup norm. This result is the Riesz Representation Theorem, sometimes also called the Riesz-Markov Theorem [Fol84]. One consequence of the Riesz-Markov Theorem is that a measure μ is invariant under the flow ϕ_t if and only if

$$\int g \circ \phi_t \, d\mu = \int g \, d\mu \tag{2.2}$$

for all bounded continuous functions g and $t \in \mathbb{R}$. In later sections we will discuss invariant measures of Markov chains, which can be defined in a similar way: a measure μ is invariant for the Markov chain (x_n) if $\mathbb{E}\left[\int g(x_n)\, d\mu\right] = \int g \, d\mu$ for all bounded measurable functions g and all $n \geq 0$.

Poincaré sections also have a probabilistic interpretation: they provide visualizations of the support of invariant measures conditioned by a given constraint. For example, the picture in Figure 2.2(b) shows the support of an invariant measure of the Lorenz flow conditioned on $x_3 = z$. As explained in the Introduction, part of the motivation for this work is a desire to understand the applicability of optimal prediction methods to highly nonlinear PDEs. Thus conditional probability measures, and by extension Poincaré sections, will be of intrinsic interest in later discussions of the Kuramoto-Sivashinsky equation.

2.2 Hyperbolic dynamics

A common (though not universally accepted) definition of chaos is that almost all trajectories with nearby initial conditions diverge exponentially fast as $t \to \infty$. The Lorenz system is chaotic by this definition, as was shown by the results in Figure 2.1. However, despite the apparent simplicity of the Lorenz equations, it is very difficult to prove rigorous results concerning the properties of invariant measures and invariant sets

of the Lorenz flow. For this reason, a simpler class of deterministic chaotic systems called hyperbolic dynamical systems has come to occupy a central place in mathematical studies of deterministic chaos.

The simplest example of a hyperbolic dynamical system is the diffeomorphism $F_{\text{cat}} : \mathbb{T}^2 \to \mathbb{T}^2$ defined by

$$F_{\text{cat}}(x_1, x_2) = \begin{bmatrix} 2x_1 + x_2 \\ x_1 + x_2 \end{bmatrix} \pmod 1 \tag{2.3}$$

where the 2-torus $\mathbb{T}^2$ is the unit square $[0,1] \times [0,1]$ with periodic boundary conditions. The map F_{cat} is the Arnold cat map and the simplest example of a "hyperbolic toral automorphism." It has a very simple dynamical behavior. First, note that the Jacobian matrix of F_{cat} is simply

$$DF_{\text{cat}}(x_1, x_2) = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}. \tag{2.4}$$

It has eigenvalues

$$\lambda^{\pm} = \frac{3 \pm \sqrt{5}}{2} \tag{2.5}$$

and normalized eigenvectors

$$e^{\pm} = \frac{1}{\sqrt{4 - \lambda^{\pm}}} \begin{bmatrix} 1 \\ \lambda^{\pm} - 2 \end{bmatrix}. \tag{2.6}$$

For each $x \in \mathbb{T}^2$, define

$$W^u(x) = \left\{ x' \in \mathbb{T}^2 : (x' - x) \propto e^+ \pmod 1 \right\} \tag{2.7}$$

and

$$W^s(x) = \left\{ x' \in \mathbb{T}^2 : (x' - x) \propto e^- \pmod 1 \right\}. \tag{2.8}$$

As $0 < \lambda^- < 1 < \lambda^+$, if two nearby but distinct points x' and x'' belong to the same set $W^u(x)$, then $d(F^n_{\text{cat}}(x'), F^n_{\text{cat}}(x''))$ will increase exponentially until it is of order 1, while two nearby points on $W^s(x)$ will converge to each other exponentially fast. The set $W^u(x)$ is naturally called the unstable manifold (and $W^s(x)$ the stable manifold) of F_{cat} at x. Note that the sets W^u and W^s are invariant under F_{cat} in the sense that $F_{\text{cat}}(W^u(x)) \subset W^u(F_{\text{cat}}(x))$ and $F_{\text{cat}}(W^s(x)) \subset W^s(F_{\text{cat}}(x))$; this is why they are commonly referred to as invariant manifolds or invariant "foliations."

More generally, let X be a compact smooth n-manifold and $F : X \to X$ a smooth diffeomorphism. Suppose $K \subset X$ is invariant under F. The map F is said to be uniformly hyperbolic on K if there exist two continuous families of vector spaces E^u and E^s satisfying the following conditions:

1. The relations $E^u_x \subset T_xX$ and $E^s_x \subset T_xX$ hold for all $x \in K$.
2. The tangent space T_xX of X at x is equal to the sum of the vector spaces E^u_x and E^s_x for all $x \in K$.
3. For all $x \in K$, $DF(x)E^u_x = E^u_{F(x)}$ and $DF(x)E^s_x = E^s_{F(x)}$.
4. There exist constants $c > 0$ and $0 < \lambda < 1$ such that $||DF^{-n}(x) \restriction_{E^u_x}|| \leq c\lambda^n$ and $||DF^n(x) \restriction_{E^s_x}|| \leq c\lambda^n$ for all $x \in \Lambda$ and $n \geq 0$.

An F-invariant set K on which F is hyperbolic is a hyperbolic set for F. If there exists an open neighborhood U of K such that F maps U into itself and for all $x \in U$, $d(F^n(x), K) \to 0$ as $n \to \infty$, then K is an attractor. A hyperbolic attractor of F such that the set of periodic points of F are dense in K is an Axiom A attractor. In the special case when F is uniformly hyperbolic on the set $K = X$, the map F is said to be an Anosov diffeomorphism.

2.2.1 Stable manifold theorem

Suppose K is a hyperbolic set for F. The global stable set $W^u(x)$, for $x \in K$, is defined to be the set of points x' such that $\limsup_{n\to-\infty} d(F^n(x), F^n(x')) = 0$. Similarly, the global stable set $W^s(x)$ is the set of points x' such that $\limsup_{n\to+\infty} d(F^n(x), F^n(x')) = 0$.

Theorem 2.1 (Local Stable Manifolds). *Let X be a compact manifold and $F : X \to X$ a C^r diffeomorphism. Suppose $K \subset X$ is a closed invariant set and F is uniformly hyperbolic on K. Then there exists an $\epsilon > 0$ such that the sets*

$$\begin{aligned} W^s_\epsilon(x) &= \{x' \in X : \limsup_{n\to+\infty} d(F^n(x), F^n(x')) = 0, d(F^n(x), F^n(x')) \leq \epsilon, n \geq 0\} \\ W^u_\epsilon(x) &= \{x' \in X : \limsup_{n\to-\infty} d(F^n(x), F^n(y)) = 0, d(F^n(x), F^n(y)) \leq \epsilon, n \leq 0\} \end{aligned} \tag{2.9}$$

are embedded C^r-submanifolds of X. Furthermore, $W^u_\epsilon(x)$ is diffeomorphic to Euclidean space, $\dim W^u_\epsilon(x) = \dim E^u_x$, and $W^u_\epsilon(x)$ depends continuously on x. The analogous result holds for stable manifolds. If the metric $||\cdot||$ on X is adapted so that there is a constant $0 < \lambda < 1$ such that

$$||DF \restriction_{E^s}||, ||DF^{-1} \restriction_{E^u}|| < \lambda, \tag{2.10}$$

then for all $x \in K$, $y \in W^u(x)$, and $z \in W^s(x)$, the following holds:

$$\begin{aligned} \sup_{n\geq 0}\left(\lambda^{-n} d(F^n(x), F^n(y))\right) &\leq d(x,y) \\ \sup_{n\geq 0}\left(\lambda^{-n} d(F^{-n}(x), F^{-n}(z))\right) &\leq d(x,z) \end{aligned} \tag{2.11}$$

Note that we can always construct an adapted metric satisfying (2.10); see Shub [Shu87] or Young [You95] for details.

Theorem 2.2. *The global unstable and stable sets $W^u(x)$ and $W^s(x)$ satisfy*

$$\begin{aligned} W^u(x) &= \cup_{n\geq 0} F^{+n}\left(W^u_\epsilon\left(F^{-n}(x)\right)\right) \\ W^s(x) &= \cup_{n\geq 0} F^{-n}\left(W^s_\epsilon\left(F^{+n}(x)\right)\right) \end{aligned} \tag{2.12}$$

That is, the global unstable and stable sets defined earlier have the property that they are locally submanifolds of X with the same regularity as F. Thus, we may speak of stable and unstable manifolds instead of mere sets.

Theorem 2.3. *If K is a hyperbolic set for F, then the stable and unstable subspaces E^s and E^u are Hölder continuous vector bundles on K.*

Theorem 2.4. *Suppose K is a hyperbolic set for F. Then for every sufficiently small $\epsilon > 0$ there exists $\delta > 0$ such that $d(x,y) < \delta$ implies that $W^u_\epsilon(x)$ intersects $W^s_\epsilon(y)$ transversely at exactly one point (denoted by $[x,y]$). The map $(x,y) \mapsto [x,y]$ is continuous on its domain.*

These results imply that on every hyperbolic invariant set K, the stable and unstable manifolds form local coordinate systems (called the local product structure of K). These local coordinate systems are usually not smooth except in special cases — they are at best Hölder continuous in general.

The structure of hyperbolic attractors is intimately connected with the geometry of stable and unstable manifolds: a hyperbolic attractor K is the union of the unstable manifolds $W^u(x)$ for $x \in K$, and the union of the stable manifolds $W^s(x)$ forms a set of positive Riemann measure. For precise statements and proofs, see Shub [Shu87] and Young [You95]. In any case, the Hölder continuity of the local product structure is the best one can expect in general; for Anosov maps in two dimensions, one can prove that the local product structure is as smooth as F. The Hölder continuity result means that while stable and unstable manifolds are locally as smooth as the map F itself, they do not fit together in a very regular way.

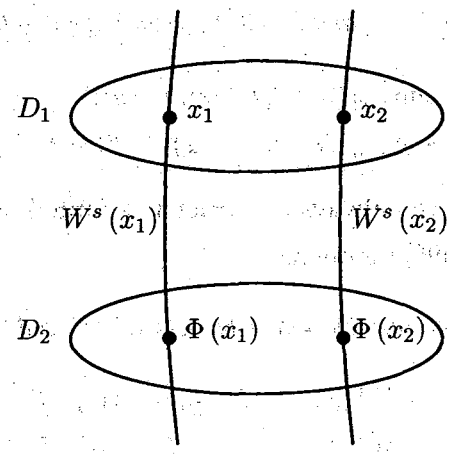


Figure 2.3: The construction of holonomy maps.

For later analysis, a different formulation of the preceding result is needed. Let D_1 and D_2 be two k-dimensional discs embedded in X near a point p in a hyperbolic set K. Suppose they are embedded in such a way that they both intersect the stable manifold $W^s(p)$ transversely, and that $k = \dim X - \dim W^s(p)$. Then the stable manifold defines a map $\Phi : D_1 \to D_2$ called the holonomy map, as illustrated in Figure 2.3. The Hölder continuity of the stable subspace E^s is equivalent to the statement that the holonomy map Φ is Hölder continuous. A stronger result is true: recall that if ν is a measure on Y, then $\Phi_*\mu$ is absolutely continuous with respect to ν if there exists a function $h \in L^1(\nu)$ such that $(\Phi_*\mu)(E) = \int_E h\, d\nu$.

Theorem 2.5. *Let D_1 and D_2 be embedded discs transverse to the stable manifolds, and let the holonomy map $\Phi : D_1 \to D_2$ be defined as above. Then it is absolutely continuous in the sense that if m_i denotes the intrinsic volume measure of the embedded disc D_i, then the measure $\Phi_* m_1$ is absolutely continuous with respect to the measure m_2. Furthermore, the Radon-Nikodym derivative of $\Phi_* m_1$ with respect to m_2 satisfies the inequality $|\frac{d\Phi_* m_1}{dm_2} - 1| \leq c < 1$ for some constant c depending only on the map F and the distance (in the C^1 metric) between the discs D_1 and D_2.*

See Barreira and Pesin [BP02] for more general versions of this result.

Example 2.6. Consider the solenoid map from the Introduction:

$$F(\theta, r, s) = \begin{bmatrix} 2\theta \pmod{2\pi} \\ \epsilon_1 \cos(\theta) + \epsilon_2 r \\ \epsilon_1 \sin(\theta) + \epsilon_2 s \end{bmatrix}, \tag{2.13}$$

where $0 \leq \theta < 2\pi$, $r^2 + s^2 \leq 1$, and ϵ_1, ϵ_2 are constants such that $0 < \epsilon_2 < \epsilon_1 < 1/2$. Differentiation yields

$$DF(\theta, r, s) = \begin{bmatrix} 2 & 0 & 0 \\ -\epsilon_1 \sin(\theta) & \epsilon_2 & 0 \\ \epsilon_1 \cos(\theta) & 0 & \epsilon_2 \end{bmatrix}, \tag{2.14}$$

so that the subspaces E_x^s of the tangent space $T_x X$ spanned by $\frac{\partial}{\partial r}$ and $\frac{\partial}{\partial s}$ are invariant in the sense that $DF(x)E_x^s = E_{F(x)}^s$ and DF is uniformly contracting when restricted to E_x^s. Similarly, DF is expanding on the subspaces E_x^u, which are transverse to E_x^s (see [KH95] or [Shu87] for details). Now, the discs $W^s(x) = \{x' : \theta(x) = \theta(x')\}$ (θ denotes the angle coordinate map) are clearly invariant in the sense that $F(W^s(x)) \subset W^s(F(x))$ and the contraction property above implies that $d(F^n(x), F^n(x')) \to 0$ as $n \to +\infty$ whenever $x \in X$ and $x' \in W^s(x)$. The discs $W^s(x)$ thus form the stable manifolds of F and the subspaces E_x^s the associated stable subspaces.

2.2.2 Other models of dynamical systems

Before continuing, note that not every hyperbolic dynamical system arises as a diffeomorphism F on a compact manifold X: there are many natural examples of dynamical systems which do not fall into this category. These include:

1. *Continuous-time dynamical systems:* As explained before, flows arise naturally in physical problems. The definition of hyperbolicity given above can be applied to flows. However, in studying ergodic properties a flow $\{\phi_t : t \in \mathbb{R}\}$ on a compact manifold X, it suffices to study the maps at discrete intervals $\{nT : n \in \mathbb{Z}\}$ for some fixed $T > 0$. In particular, if μ is an invariant probability measure of the continuous-time system ϕ_t, $t \in \mathbb{R}$, then it is also an invariant measure of the discrete-time system $\{\phi_T^k, k \in \mathbb{Z}\}$, and if μ is ergodic with respect to the flow, then it is also ergodic with respect to the discrete-time dynamics. Furthermore, if the flow has an SRB measure

(SRB measures are defined below), then so does its discrete-time version, and the SRB measures coincide.

2. *Non-invertible maps on manifolds with boundary:* Many mathematical and physical examples of dynamical systems are naturally described on domains with boundaries and defined by maps which are injective but not surjective. Most of the definitions given here will work with little modification in this setting. The solenoid map is a good example of a non-invertible map on a manifold with boundary.

3. *Flows defined by vector fields on* $\mathbb{R}^n$*:* The Lorenz flow from the introduction is an example of a flow with partial hyperbolic behavior. Such physically realistic examples are generally defined by ordinary differential equations on Euclidean n-space, which is unbounded and hence non-compact. The interesting motion, however, often takes place on and near a compact attractor K. One can often find a bounded open neighborhood U of K such that U is mapped into itself by the flow (i.e. the defining vector field point inward at every point of ∂U) and every point of U generates a trajectory which converges to K exponentially fast. The flow map ϕ_t then defines an injective (but not surjective!) mapping of the closure $\bar{U}$ of U into itself.

4. *Systems with singularities:* Discrete-time systems modeled by diffeomorphisms F or injective maps with nonsingular Jacobians $DF(x)$ are the simplest to analyze. This condition can fail even in the simplest examples, such as the quadratic map $x \mapsto 4x(1-x)$ on the unit interval $[0, 1]$.

While all these other classes of dynamical systems are important, a complete survey is beyond the scope of the current work. Instead, the focus here is on the simplest case, to explain the relevant theory and provide a language for describing the structure of hyperbolic systems and their invariant measures.

2.3 Lyapunov exponents & entropy

The preceding sections focus on the geometric properties of dynamical systems. This section focuses on the metrical aspects of dynamics, in particular on the construction of Lyapunov exponents, which measure the rates at which a dynamical system expands and contracts phase space volume.

More precisely, let $x \in X$ be any point and $v \in T_xX$ a nonzero tangent vector. Define

$$\lambda(x,v) = \lim_{n\to+\infty} \frac{1}{n} \log ||D(F^n)(x) \cdot v||. \tag{2.15}$$

For example, consider the cat map F in Equation (2.3): its Jacobian DF is constant, so the Lyapunov exponents are simply the eigenvalues $\lambda^{\pm}$.

More generally, Oseledec's theorem [You95] guarantees that for every invariant measure μ of F, the limit above exists for μ-almost every x and v. The quantity $\lambda(x,v)$ measures the asymptotic exponential rate of separation of two initial points near x separated infinitesimally in the direction given by v. Clearly, $\lambda(x,av) = \lambda(x,v)$ for every nonzero real number a — the Lyapunov exponent depends only on the subspace from which v is taken. Oseledec's theorem further states that for each x, there exists an increasing sequence of vector spaces $E^i(x)$ and functions $\lambda_i(x)$, $i = 1, ..., n$, such that

$$E^1(x) \subset ... \subset E^n(x) = T_xX$$

and

$$\inf\{\lambda(x,v) : v \in E^i(x)\} = \lambda_i(x).$$

One can thus speak of the Lyapunov spectrum $\{\lambda_i(x) : i = 1, ..., n\}$ of a map F with respect to the invariant measure μ.

The quantity $j_F^u(x) = \log ||DF(x) \restriction_{E^u(x)}||$ is the local expansion rate of F at x. It has the property that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} j_F^u(F^k(x)) = \sum_i \lambda_i^+(x) \tag{2.16}$$

for almost every x.

Lyapunov exponents are only defined for μ-almost every $x \in X$. Thus, the Lyapunov exponents of a dynamical system depend on both the map F and the invariant measure μ. For convenience, their dependence on F and μ is usually suppressed.

It is easy to see that the Lyapunov exponents of uniformly hyperbolic maps are all nonzero. However, in the setting of flows one must deal with invariant manifolds, called center manifolds, along which the flow is neither exponentially expanding nor exponentially contracting. That is, the flow has a zero Lyapunov exponent in directions tangent to center manifolds. The reason that center manifolds arise naturally in the setting of flows is that every bounded solution curve which does not coincide with the stable manifold of a fixed point must be a center manifold.

2.3.1 Entropy and information theory

Intuitively, the exponential separation of trajectories corresponds to an exponential stretching of phase space volume. This is accompanied by a loss of information about the initial condition of a given trajectory. A precise way to measure this loss of information or generation of randomness is provided by the notion of entropy from information theory [CT91, Sha48].

Let $\Omega = \{\omega_1, ..., \omega_N\}$ be a finite set and let μ be a probability distribution on Ω. The entropy of μ is defined by

$$H(\mu) = -\sum_{i=1}^{N} \mu(\omega_i) \log_2(\mu(\omega_i)). \tag{2.17}$$

Among other things, it is an estimate of the minimum number of random bits needed to generate a "typical" element of (Ω, μ). The function H (really a family of functions, one for each N) has some remarkable properties. These properties include:

1. The function H is symmetric in its arguments: permuting the entries of the probability vector μ does not change the value of H.

2. $H(1, 0, 0, ...) = 0$: the amount of information needed to simulate the absence of randomness is 0.

3. $H(0, \mu(\omega_1), ..., \mu(\omega_N)) = H(\mu(\omega_1), ..., \mu(\omega_N))$: null events do not add to the randomness of a system.

4. $H(\mu) \leq H(\frac{1}{N}, ..., \frac{1}{N})$, with equality if and only if $\mu(\omega_i) = \frac{1}{N}$ for all i: this means that the uniform distribution is the most random possible.

5. H is nonnegative and continuous (for each fixed N).

Furthermore, Shannon proved that H is the only function, up to a positive multiplicative constant, which has these properties. It is thus a natural measure of information. If $Z : \Omega \to \mathbb{R}$ is a random variable, $H(Z)$ can be defined by applying the definition above to the probability distribution of Z.

One basic property of entropy is that it is invariant under transformations. For finite sets, this simply means that if Z is a random variable and ψ is a function, then $H(\psi(Z)) \leq H(Z)$, with equality if and only if ψ is injective on the range of Z. This has

an amusing implication: let Z_N be a uniform random variable with range $\{1, ..., N\}$, and let $(p_1, p_2, p_3, ...) = (2, 3, 5, ...)$ denote the sequence of prime numbers. One can then define random variables $(E_1, E_2, E_3, ...)$ implicitly through the prime factorization of the random integer Z_N:

$$Z_N = p_1^{E_1} \cdot p_2^{E_2} \cdot ...$$

Since $Z_N \leq N$, we must have $E_k = 0$ whenever $k \geq \pi(N)$, the number of primes $\leq n$. Thus, by the convexity of the function $x \mapsto x \log_2(x)$,

$$\begin{aligned} \log_2(N) = H(Z_N) &= H(E_1, ..., E_{\pi(N)}) \\ &\leq H(E_1) + ... + H(E_{\pi(N)}). \end{aligned}$$

But $E_k \leq \log_2(N)$ with probability 1, so $H(E_k) \leq \log_2(\log_2(N))$. Therefore

$$\log_2(N) \leq \pi(N) \log_2(\log_2(N))$$

so that

$$\pi(N) \geq \frac{\log_2(N)}{\log_2(\log_2(N))}.$$

This estimate is certainly not sharp: it is well known that $\pi(N) \geq \log_2(N)$.

2.3.2 Kolmogorov-Sinai entropy

We now use concepts developed in the previous section to quantify the amount of randomness generated by a dynamical system. Let μ be an F-invariant measure. A μ-partition is a collection Ξ of measurable subsets of X such that $\mu\left(X - (\cup_{\xi \in \Xi} \xi)\right) = 0$ and $\mu(\xi \cap \xi') = 0$ for all $\xi, \xi' \in \Xi$ such that $\xi \neq \xi'$. Given two μ-partitions Ξ_1 and Ξ_2, their common refinement $\Xi_1 \vee \Xi_2$ is the partition $\{\xi_1 \cap \xi_2 : \xi_i \in \Xi_i\}$. Let $F^{-1}(\Xi)$ denote the partition $\{F^{-1}(\xi) : \xi \in \Xi\}$. The entropy of a μ-partition Ξ is

$$H(\mu, \Xi) = -\sum_{\xi \in \Xi} \mu(\xi) \log \mu(\xi). \tag{2.18}$$

Physically, we can think of partitions as providing a coarse-grained description of the system. Thus, each set $\xi \in \Xi$ specifies a coarse-grained state. The entropy $H(\mu, \Xi)$ then measures the average amount of information carried by each coarse-grained state $\xi \in \Xi$ when the system is in the statistical steady state μ.

Now consider $H\left(\mu, \bigvee_{k=0}^{n-1} F^{-k}(\Xi)\right)$. This measures the amount of information carried by each length-n sequence $(\xi_0, \xi_1, \ldots, \xi_{n-1})$ of coarse-grained states in Ξ when the system is in the statistical steady state μ. In chaotic systems, this quantity is expected to grow linearly as $n \to +\infty$. Thus, one can measure the rate of entropy production by

$$h_\mu(\Xi) = \lim_{n \to +\infty} \frac{1}{n} H\left(\mu, \bigvee_{k=0}^{n-1} F^{-k}(\Xi)\right). \tag{2.19}$$

The Kolmogorov-Sinai entropy (also called the metric entropy) of (F, μ) is then

$$h_\mu = \sup_{\Xi} h_\mu(\xi). \tag{2.20}$$

The metric entropy and the Lyapunov exponents of a dynamical system are not independent: Ruelle proved that

$$h_\mu \leq \sum_i \int_X \lambda_i^+ \, d\mu \tag{2.21}$$

holds for all invariant measures μ of F, where $g^+(x) = \max(g(x), 0)$.

Again, the basic property of entropy is that it measures the intrinsic amount of information in a random signal. For the Kolmogorov-Sinai entropy, this means that h_μ is invariant under bi-Lipschitz homeomorphisms. It is thus an example of a dynamical invariant.

2.3.3 Markov partitions and symbolic dynamics

It is possible to choose μ-partitions which dramatically simplify the analysis of uniformly hyperbolic systems. These nice partitions are called Markov partitions. Roughly speaking, a Markov partition has elements whose boundaries coincide with stable or unstable manifolds. One basic use of Markov partitions is to construct symbolic dynamics, reducing many questions in ergodic theory to corresponding questions in the statistical mechanics of one-dimensional lattices with rapidly decaying potentials. This approach has certain limitations (Markov partitions do not generally exist in nonuniformly hyperbolic systems) and is not necessary in any case (one can develop smooth ergodic theory without symbolic dynamics), but it does provide a useful conceptual link between hyperbolic dynamics and statistical mechanics. It is presented here in terms of the cat map example introduced earlier. Interested readers should see Bowen [Bow75] for the general construction.

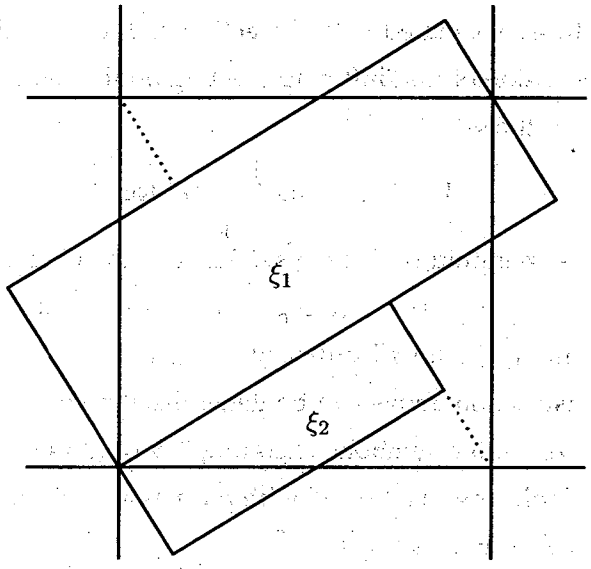


Figure 2.4: A (generator for a) Markov partition of the cat map (2.3).

The basic idea of Markov partitions is illustrated in Figure 2.4, where a Markov partition for the cat map (2.3) has been drawn. Let $\Xi = \{\xi_1, \xi_2\}$. For each point $p \in \mathbb{T}^2$, let $s = (..., s_{-1}, s_0, s_1, ...)$ be a bi-infinite symbolic sequence such that $s_k = \xi_i$ if ξ_i contains $F^k(p)$. Clearly, ambiguities arise whenever $p \in \partial\xi_i$ for some i, but such points form a set of Lebesgue measure zero. Thus the map $\Phi : p \mapsto s$ is well-defined almost everywhere. It is in fact invertible, and its inverse is given by $\Phi^{-1}(s) = \cap_{k=-\infty}^{\infty} F^{-k}(s_k)$. The map Φ identifies points in $\mathbb{T}^2$ with symbolic sequences in the space of bi-infinite sequences with symbols taken from the set Ξ; this space of bi-infinite symbolic sequences is denoted $\Xi^{\mathbb{Z}}$.

Let σ now denote the shift operator acting on symbolic sequences $\Xi^{\mathbb{Z}}$, so that

$$\sigma(s)(k) = s(k+1), \tag{2.22}$$

i.e. it shifts a symbolic sequence to the left. Then the following equation holds:

$$F \circ \Phi^{-1} = \Phi^{-1} \circ \sigma. \tag{2.23}$$

Heuristically, the map Φ allows one to identify F-invariant measures μ on $\mathbb{T}^2$ with shift-invariant measures on the sequence space $\Xi^{\mathbb{Z}}$. If we equip Ξ with the discrete topology and associate with the sequence space $\Xi^{\mathbb{Z}}$ the corresponding product topology, then $\Xi^{\mathbb{Z}}$ is a compact metrizable space and the map Φ^{-1} is a continuous mapping of $\Xi^{\mathbb{Z}}$ onto X.

Geometric structures associated with the action of F on X have natural counterparts associated with the action of the shift map σ acting on the sequence space $\Xi^{\mathbb{Z}}$. In particular, given a sequence s, the set

$$W_n^u \stackrel{\text{def}}{=} \{s' : s(k) = s'(k) \text{ for all } k < n\} \tag{2.24}$$

is a local unstable manifold of s: each sequence s' in W_n^u agrees with s to the *left* of the nth position, so that $d(\sigma^{-n}(s), \sigma^{-n}(s')) \to 0$ as $n \to +\infty$. Furthermore, each local unstable manifold of s contains W_n^u for all sufficiently negative n.

The construction above can be generalized to hyperbolic invariant sets. In the general case, however, the symbolic encoding Φ would identify the restriction $F\restriction_K$ of a map F to a hyperbolic invariant set K with the action of σ on a shift-invariant subspace of $\Xi^{\mathbb{Z}}$ (instead of the full sequence space $\Xi^{\mathbb{Z}}$).

2.4 Sinai-Ruelle-Bowen measures

This section is based mostly on Bowen [Bow75] and Young [You95].

Let X be a compact manifold and $F : X \to X$ a C^∞ diffeomorphism. A Borel measure μ on X is invariant if $F_*\mu = \mu$, where $(F_*\mu)(A)$ is defined to be $\mu(F^{-1}(A))$ for all Borel sets A. The pair (F, μ) is ergodic if every F-invariant Borel set A satisfies $\mu(A) \in \{0, 1\}$. The pair (F, μ) is mixing if $\lim_{n\to+\infty} \mu(F^{-n}(A) \cap B) = \mu(A) \cdot \mu(B)$. It is easy to check that if (F, μ) is mixing then it is ergodic as well and that the converse does not hold. There are equivalent formulations of these conditions: it is easy to check that (F, μ) is ergodic if and only if for every pair of Borel sets A and B in X such that $\mu(A)$ and $\mu(B)$ are both positive, there exists an integer $n \geq 0$ such that $\mu(F^{-n}(A) \cap B) > 0$, and that (F, μ) is mixing if and only if for every pair of Borel sets A and B with positive μ-measure, there exists an n_0 such that $\mu(F^{-n}(A) \cap B)$ is positive for all $n \geq n_0$.

By analogy, a map F is said to be topologically transitive if for all open sets A and B, there exists an integer $n \geq 0$ such that $F^{-n}(A) \cap B$ is nonempty. Similarly, a map F is topologically mixing if for all open sets A and B, there exists an integer $n_0 \geq 0$ such that $F^{-n}(A) \cap B$ is nonempty for all $n \geq n_0$.

The fundamental theorem of ergodic theory is the following:

Theorem 2.7 (Birkhoff). *Let X be a locally compact Hausdorff space, $F : X \to X$ a continuous map, and μ a Borel measure. If μ is invariant under F, then for every function $g \in L^1(\mu)$ and*

μ-almost every point $x_0 \in X$, the limit

$$\lim_{N\to+\infty} \frac{1}{N} \sum_{k=0}^{N-1} g(F^k(x_0)) \tag{2.25}$$

exists.

Because X is a compact metric space, every Borel measure on X is also a Radon measure. Thus, the Riesz-Markov theorem applies and there exists a bijective correspondence between Borel measures on X and bounded linear functions on $C(X)$ [Fol84]. This fact and Theorem 2.7 imply that for every diffeomorphism F, there exists a plethora of Borel measures on X which are invariant under F.

Some special dynamical systems have unique invariant measures. Most systems, however, have large families of invariant measures. For example, the cat map (2.3) has periodic orbits of all periods, and each periodic orbit supports an invariant measure. For systems with uniformly hyperbolic attractors, however, there is a particularly natural class of invariant measures called SRB (Sinai-Ruelle-Bowen) measures.

Let μ be an invariant measure, K the support of μ, and Ξ a μ-partition of K. For each $x \in K$, let $\Xi(x)$ denote the element of Ξ containing x (this is well-defined for μ-almost all $x \in K$). The μ-partition Ξ is said to be subordinate to the unstable foliation W^u if for every $x \in K$, $\Xi(x) \subset W^u(x)$. As X is compact Hausdorff, for every Borel measure μ and μ-partition Ξ there corresponds a canonical family of conditional measures $\{\mu_{\Xi(x)}\}$ such that $\mu_{\Xi(x)}(\Xi(x)) = 1$ and for every Borel set $B \subset X$, the function

$$x \mapsto \mu_{\Xi(x)}(B) \tag{2.26}$$

is measurable and

$$\int_X \mu_{\Xi(x)}(B)\mu(dx) = \mu(B). \tag{2.27}$$

The measure μ is absolutely continuous with respect to W^u if for every μ-partition Ξ of K and almost every $x \in K$ the conditional measure $\mu_{\Xi(x)}$ is absolutely continuous with respect to the normalized Lebesgue measure $(m_{\text{Leb}})_{\Xi(x)}$ of $\Xi(x)$, and if there exists a constant $c > 0$ such that

$$c^{-1} \leq \frac{d\mu_{\Xi(x)}}{d(m_{\text{Leb}})_{\Xi(x)}} \leq c$$

for almost all $x \in K$. Furthermore, we say that μ is absolutely continuous with respect to W^u with Hölder continuous densities if in addition to the condition above, there exist

constants $c_1 > 0$ and $0 < \alpha \leq 1$ such that the Radon-Nikodym derivatives $\frac{d\mu_{\Xi(x)}}{d(m_{\text{Leb}})_{\Xi(x)}}$ are α-Hölder continuous with Hölder norms bounded above by c_1.

Remark 2.8. Given a measure space (X, μ, Σ), a random variable Z on X, and a σ algebra $\Sigma' \subset \Sigma$, the conditional expectation of Z with respect to Σ' can be defined rather generally using the Radon-Nikodym theorem. One may be tempted to define conditional measures for elements of Σ' using conditional expectations. To construct genuine conditional measures (conditioned on sets of μ-measure zero), however, requires some care. This is because in order for each μ_x to be a measure, we must choose versions of conditional expectations which ensure that the set functions μ_x are countably additive. This results in an uncountable number of conditions. See Rogers and Williams [RW00a] for a thorough discussion of these issues.

An invariant measure μ supported on a compact invariant set K is said to be physically relevant if K has a measurable neighborhood U_K with positive Lebesgue measure such that for every $x_0 \in U$ and every bounded continuous function $g : X \to \mathbb{R}$,

$$\lim_{N \to +\infty} \frac{1}{N} \sum_{k=0}^{N-1} g(F^k(x_0)) = \int_X g \, d\mu_{\text{SRB}}. \tag{2.28}$$

An SRB measure is an ergodic invariant measure μ_{SRB} which is physically relevant and satisfies the following conditions:

1. The measure μ_{SRB} is absolutely continuous with respect to the unstable foliation W^u with Hölder continuous densities.

2. $(F \restriction_K, \mu_{\text{SRB}})$ has a positive Lyapunov exponent.

Much is known about the properties of SRB measures when they exist. To determine whether a given dynamical system actually possesses an SRB measure is much more difficult. See Young [You02] for a discussion of relevant issues.

It is not the case that every hyperbolic attractor possesses an SRB measure. For example, consider the "figure 8" attractor in Figure 2.5. It is easy to see that the physically relevant invariant measure is a Dirac delta mass located at the hyperbolic fixed point in the center, which we identify with the origin 0. The measure δ_0 has one positive and one negative Lyapunov exponent and is not absolutely continuous along its unstable manifold. Figure 2.5 also shows that not every invariant measure is physically relevant; a δ mass

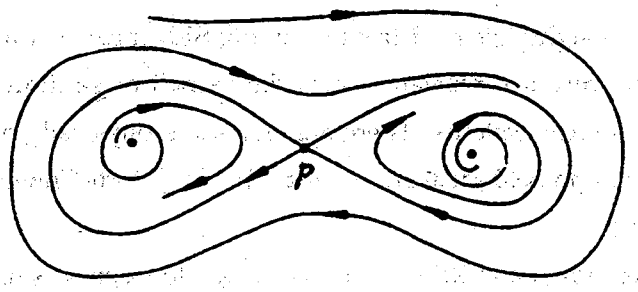


Figure 2.5: The "figure 8" attractor.

supported by either of the unstable fixed points in the "figure 8" flow is invariant but not physically relevant.

Here is the main result concerning SRB measures:

Theorem 2.9. *Let X be a compact Riemannian manifold and $F : X \to X$ a C^2 diffeomorphism. Suppose $K \subset X$ is a uniformly hyperbolic attractor and F is topologically transitive on K. Then there exists a unique SRB measure μ_{SRB} supported on K.*

See Bowen [Bow75] and Young [You95] for two different proofs of this theorem.

Remark 2.10. Theorem 2.7 asserts that the limit in Equation (2.28) exists for μ_{SRB}-almost every initial condition x_0. Theorem 2.9, on the other hand, guarantees the convergence of (2.28) for a set of positive Lebesgue measure. Thus, if one takes the view that sets of positive Lebesgue measure are physically observable, then SRB measures are precisely the physically relevant statistical steady states associated with hyperbolic attractors.

Remark 2.11. Ruelle's inequality (2.21) becomes an equality when the invariant measure μ is an SRB measure [You95]:

$$h_{\mu_{\text{SRB}}} = \sum_i \int_X \lambda_i^+ \, d\mu_{\text{SRB}}. \tag{2.29}$$

This is Pesin's formula. It gives a precise connection between the rate of phase space expansion and the rate of loss of information about initial conditions. It is instructive to check that Pesin's formula is satisfied by the cat map $(F_{\text{cat}}, m_{\text{Leb}})$.

We have already seen examples of systems with SRB measures: the first is the solenoid map, the second is the Lebesgue measure m_{Leb} for the cat map. To be sure, these are abstract mathematical examples. The existence or nonexistence of SRB measures outside the mathematically pristine setting of Axiom A attractors is difficult to establish.

Nonetheless, insights gleaned from studying SRB measures help illuminate many questions of direct physical relevance, including small-noise limits of random perturbations limits of dynamical systems. There are also some physically observable consequences of the existence of an SRB measure, such as the Gallavotti-Cohen Fluctuation Theorem (see Maes [Mae99]).

Even when an SRB measure is known to exist, it is usually not possible to compute its density along unstable manifolds. For that matter, it is generally impossible to determine unstable manifolds exactly. It is relatively easy, however, to compute the ratio of the probability density at distinct two points in the same unstable manifold [Dol01]:

Proposition 2.12. *Suppose μ is absolutely continuous along the unstable manifolds of a C^2 diffeomorphism F and has Hölder continuous densities. Let ρ denote the density of μ on a local unstable manifold $W^u_\epsilon(x)$ and let $J^u_F(x)$ be the jacobian determinant of F on the unstable subspace $E^u(x)$. Then*

$$\frac{\rho(x)}{\rho(x')} = \prod_{n=1}^{\infty} \frac{J^u_F(F^{-n}(x))^{-1}}{J^u_F(F^{-n}(x'))^{-1}} \tag{2.30}$$

for all $x' \in W^u_\epsilon(x)$.

Proof. Let m_k be the measure on $F^{-k}(W^u_\epsilon(x))$ induced by the metric of X and set $\rho_k\, dm_k = F^{-k}_*(\rho_0\, dm_0)$. Then

$$\begin{aligned}
\rho_1(x) =& \rho_0(F(x)) \cdot J^u_F(x) \\
\rho_2(x) =& \rho_0(F^2(x)) \cdot J^u_F(x) \cdot J^u_F(F(x)) \\
&\vdots \\
\rho_n(x) =& \rho_0(F^n(x)) \cdot J^u_F(x) \cdot \ldots \cdot J^u_F(F^{n-1}(x)).
\end{aligned}$$

Thus

$$\frac{\rho_n(F^{-n}(x))}{\rho_n(F^{-n}(x'))} = \frac{\rho_0(x)}{\rho_0(x')} \prod_{k=1}^{n} \frac{J^u_F(F^{-k}(x))}{J^u_F(F^{-k}(x'))}. \tag{2.31}$$

The unstable distribution $E^u(x)$ is Hölder continuous and the jacobian DF is C^1, so the local expansion coefficient $J^u_F(x)$ is also Hölder continuous. As x' and x belong to the same unstable manifold, $d(F^n(x), F^n(x'))$ must converge to 0 when we let $n \to \infty$. Therefore the right hand side of (2.31) converges uniformly, which implies the left hand side of (2.31) also converges and is a continuous function of x and x'. But the left hand side of (2.31) is independent of x' as long as x and x' belong to the same unstable manifold. Setting $x' = x$ finishes the proof. □

The quantity $\log J_F^u(x)$ is called the local expansion rate of F. We will also refer to $J_F^u(x)$ as the local expansion coefficient of F at x.

2.5 Transfer operators

Before continuing with a review of existing small-noise results on invariant measures of dynamical systems, it is useful to take a little detour through the theory of transfer operators. This links ergodic theory with statistical mechanics and yields insights into the properties of invariant measures. Almost everything here is taken from Baladi [Bal00], Bowen [Bow75], and Sokal [Sok97]. The emphasis is on the parallel between dynamical systems theory and the theory of discrete state space Markov chains.

2.5.1 Finite state space Markov chains

Let us begin by reviewing the theory of finite state space Markov chains, as transfer operators arise naturally in this setting and a basic understanding of Markov chains can provide valuable intuition about transfer operators in the dynamical systems setting. Furthermore, this theory forms the foundation for Markov chain Monte Carlo methods, which is part of the subject of this thesis.

Let S be a finite set, which will be called the state space in what follows. The set S may be, for example, the collection of all possible internal states of a digital computer or the collection of all valid occupation vectors of a lattice gas. Without loss of generality, identify S with the set of integers $\{1, 2, 3, ..., m\}$. A real $m \times m$ matrix P is a stochastic matrix if all its entries are nonnegative and each of its rows sums to 1. The latter condition is equivalent to the statement that the column vector $e = (1, 1, 1, ..., 1)^t$ is a right eigenvector of P with eigenvalue 1: $Pe = e$. A stochastic matrix P is irreducible if for every pair of indices x and x' such that $1 \leq x, x' \leq N$, there exists a nonnegative integer k such that the xx'th entry $(P^k)_{xx'}$ of P^k is positive; it is aperiodic if there exists a k such that $(P^k)_{xx'} > 0$ for all x and x'.

Each stochastic matrix P defines a family of Markov chains with values in S: $P_{xx'}$ simply specifies the transition probability from state i to state j. The basic result is the following:

Theorem 2.13 (Perron-Frobenius). *Every stochastic matrix P has a left eigenvector μ with*

nonnegative real entries and eigenvalue 1*, and the eigenvector occurs with multiplicity* 1 *if and only if* P *is irreducible. Furthermore,* 1 *is the only eigenvalue of* P *on the unit circle if and only if* P *is aperiodic.*

The proof can be found in Katok and Hasselblatt [KH95]. The Perron-Frobenius theorem guarantees the existence of invariant measures for finite state Markov chains, and gives conditions for their uniqueness. One uniqueness proof for irreducible aperiodic chains is of particular relevance here, so it is reproduced below:

Suppose P defines an irreducible aperiodic Markov chain on a finite state space S. Given two probability measures μ_1 and μ_2 on S, define a new chain $\tilde{x} = (x^1, x^2)$ on $S \times S$ as follows:

1. Set $\tilde{x}_0 = (x_0^1, x_0^2)$, where the initial conditions are drawn from the measure $\mu_1 \times \mu_2$.
2. For each $k \geq 0$, if $x_k^1 \neq x_k^2$, then let x_k^1 and x_k^2 evolve independently to obtain x_{k+1}^1 and x_{k+1}^2.
3. If $x_k^1 = x_k^2$, then let x_l^1 evolve according to P and set $x_l^2 = x_l^1$ for all $l > k$. The time k_{coup} is the coupling time of the chain; it is a random variable (and in fact a stopping time since it is the first hitting time of the diagonal in $S \times S$).

Now let $g : S \to \mathbb{R}$ be any observable, and suppose μ_1 and μ_2 are invariant measures of P. Then

$$\begin{aligned} \left|\int g\, d\mu_1 - \int g\, d\mu_2\right| &= \left|\mathbb{E}\left[g(x_k^1) - g(x_k^2)\right]\right| \\ &\leq \left|\mathbb{E}\left[g(x_k^1) - g(x_k^2) | k \leq k_{\text{coup}}\right]\right| P(k \leq k_{\text{coup}}) + \\ &\quad \left|\mathbb{E}\left[g(x_k^1) - g(x_k^2) | k > k_{\text{coup}}\right]\right| P(k > k_{\text{coup}}) \\ &\leq 2 \sup\{|g(s)| : s \in S\} \cdot P(k \leq k_{\text{coup}}). \end{aligned} \tag{2.32}$$

(Note that $g(x_k^1) = g(x_k^2)$ for $k > k_{\text{coup}}$!) Since the chain is irreducible and aperiodic, $P(k_{\text{coup}} < \infty) = 1$, which implies that $P(k \leq k_{\text{coup}}) \to 0$ as $k \to \infty$. Thus $\mu_1 = \mu_2$. Furthermore, an upper bound on $\lim_{k\to\infty} \frac{P(k \leq k_{\text{coup}})}{k}$ yields a lower bound on the rate of exponential decay of correlations. This proof illustrates the coupling method [Gri78], which has been applied to dynamical systems in recent years as well. See [Bal00] and [BL02] for examples.

2.5.2 Monte Carlo algorithms

The computer simulation of Markov chains has become ubiquitous in computational physics and computer science. Typically, one begins with a measure μ of interest and designs a stochastic matrix P (equivalently a Markov chain $\{x_k\}$) which preserves μ, and which is irreducible and aperiodic. One can then use the stochastic matrix P to generate sample paths $\{x_k\}$ of the corresponding Markov chain, with suitably generated initial conditions x_0, and compute empirical averages $\frac{1}{N}\sum_{k=0}^{N-1} g(x_k)$. As $N \to +\infty$, this is expected to converge to the weighted sum $\sum_{x\in S} g(x)\mu(x)$, where the probability vector μ plays the role of a probability measure. This way of estimating averages is known as the Markov chain Monte Carlo method.

Typically, the initial state x_0 in a Monte Carlo calculation is either chosen from an easily accessible distribution or a deterministic state. The important task is to ensure that P preserves μ. That is, $\mu = \mu \cdot P$, which means

$$\mu(x') = \sum_x \mu(x) \cdot P_{xx'}. \tag{2.33}$$

Clearly this condition is implied by

$$\mu(x') \cdot P_{x'x} = \mu(x) \cdot P_{xx'}. \tag{2.34}$$

This is the detailed balance condition, and is equivalent to the statement that P is a self-adjoint operator on the space $L^2(\mu)$. Given a probability vector μ, an easy way to construct a stochastic matrix P which preserves μ is to start with an arbitrary irreducible aperiodic matrix $Q = (Q_{xx'})$ and set

$$\begin{aligned} A_{xx'} &= \min\left(\frac{\mu(x')Q_{x'x}}{\mu(x)Q_{xx'}}, 1\right), \\ P_{xx'} &= Q_{xx'} \cdot A_{xx'}. \end{aligned} \tag{2.35}$$

In fact, it is easy to check that P satisfies the detailed balance condition with respect to μ. The matrix Q is the proposal matrix. To estimate averages with respect to μ using P constructed as above, simply follow this recipe:

Algorithm 2.14 (Metropolis).

Input:

proposal matrix Q

matrix $A = (A_{xx'})$ constructed as in (2.35)
initial state x_0
integer N *(we will generate $N+1$ steps of the Markov chain)*

1: **for** $k := 1$ to N **do**
2: generate a new state x' with probability $Q_{xx'}$
3: $x_k := x'$ with probability $A_{xx'}$ and $x_k := x$ with probability $1 - A_{xx'}$.
4: **end for**

Output: $\{x_0, ..., x_N\}$ *($N+1$ steps of the Markov chain)*

That is, to generate each step of the chain, one begins by generating a new sample using the proposal matrix Q and applying a simple test (2.35) to decide whether to accept the proposed new state or to keep the old one. This Markov chain Monte Carlo algorithm is called the Metropolis algorithm. Note that the Metropolis algorithm requires only the relative probability $\mu(x')/\mu(x)$, not the probability itself. This property is essential in applications: since it is difficult to estimate the partition function accurately in most problems of statistical mechanics, one can only compute relative probabilities $\mu(x')/\mu(x)$, not the absolute probabilities $\mu(x)$ and $\mu(x')$.

It is easy to give an upper bound on the errors in any Monte Carlo calculation. It is given by

$$\left\| \frac{1}{N} \sum_{k=0}^{N-1} g(x_k) \right\|_{L^2(\mu)} \leq \sqrt{\frac{2 \cdot C_{gg}(0) \cdot \tau_{\text{int}}(g)}{N}}, \tag{2.36}$$

where N is the number of samples used,

$$\tau_{\text{int}}(g) = \frac{1}{2} \sum_{n=-\infty}^{+\infty} C_{gg}(n)/C_{gg}(0) \tag{2.37}$$

is the integrated autocorrelation time of the observable g,

$$C_{gg}(n) = \mathbb{E}_\mu [g(x_n)g(x_0)] - \mathbb{E}_\mu [g(x_n)] \cdot \mathbb{E}_\mu [g(x_0)] \tag{2.38}$$

is the autocovariance function of g, and $\mathbb{E}_\mu [h(x_n)]$ is the expectation value of x_n when x_0 is chosen from the measure μ. That is, the error is roughly the standard deviation of g divided by the square root of the number of effectively independent samples. This is consistent with the intuition one derives from the Law of Large Numbers. The rate of decay of correlations thus determines the convergence rate of a Monte Carlo calculation.

This is clearly related to the spectrum of the stochastic matrix P. Note that the discrete-time Fourier transform

$$\hat{C}_{gg}(f) = \sum_{n=-\infty}^{+\infty} C_{gg}(n)\, e^{i2\pi nf} \tag{2.39}$$

must be real and positive for all frequencies f; it is the spectral power density of C_{gg} [Hel91].

A related issue is the rate of relaxation to equilibrium: suppose the stochastic matrix P is irreducible and aperiodic and we choose a probability measure m_0 which is not equal to the unique invariant measure μ. Let $(x_n : n \geq 0)$ be the Markov chain generated by P and m_0, and let m_n denote the distribution of x_n. Then the measures m_n converge to μ exponentially fast. More precisely, there exist constants $C > 0$ and $0 < \rho < 1$ such that $||m_n - \mu|| \leq c\rho^n$. (The choice of the norm $||\cdot||$ is irrelevant as the state space is finite here and probability measures are simply vectors.) The constant ρ is given by

$$\rho = \sup\{|\lambda| : \lambda \in \sigma(P) \setminus \{1\}\}. \tag{2.40}$$

A simple calculation shows that

$$-\log(\rho) \leq \lim_{n\to\infty} \left(-\frac{\log\left(C_{gg}(n)/C_{gg}(0)\right)}{n} \right) \tag{2.41}$$

for all g; the quantity on the right hand side is the exponential autocorrelation time $\tau_{\text{exp}}(g)$. Furthermore, it is easy to check that $\rho = \sup_g e^{-\tau_{\text{exp}}(g)}$. The exponential autocorrelation time thus determines the rate of relaxation to equilibrium.

The exponential and integrated autocorrelation times play important roles in the analysis of Monte Carlo algorithms. In numerical studies of chaotic systems, the convergence of time averages is also determined by these autocorrelation times. There is, however, a third factor which affects the convergence of long time averages: this is the rate at which an initial condition converges to the relevant attractor. This is denoted by τ_{att} in the present discussion. Note that both τ_{exp} and τ_{int} are properties of invariant measures, while τ_{att} is related to transient phenomena which differ dramatically from system to system. Some systems can exhibit metastability, that is, the system may stay near a bounded region of phase space for a long time before converging to an attracting set. This can arise, for example, for systems in parameter regimes near some types of bifurcations [GH83, Ott00]. Numerical methods which work with one or a few trajectory at a time can only obtain local information about the phase space structure of a dynamical system and

has no hope of detecting metastability or estimating τ_{att} in general. Global methods, such as the ones described in [DH97], can obtain such information, but are computationally intensive and impractical for systems with more than a few degrees of freedom.

2.5.3 Transfer operators of dynamical systems

Analogously, one can analyze deterministic hyperbolic dynamical systems using transfer operators, which generalize stochastic matrices to the setting of dynamical systems. Let $F : X \to X$ be a map with an invariant measure μ. The covariance function of two observables $g, h : X \to \mathbb{R}$ is

$$C_{gh}(n) = \int (g \circ F^n) \cdot h \, d\mu - \int g \, d\mu \int h \, d\mu. \tag{2.42}$$

Covariance functions are similarly defined for flows. When F is hyperbolic, μ is an SRB measure for F, and g and h are both Hölder continuous with a sufficiently large exponent α, it can be shown that $C_{gh}(n)$ decays exponentially as $n \to \pm\infty$. Furthermore, the Fourier transform

$$\hat{C}_{gh}(\omega) = \sum_{n=-\infty}^{\infty} C_{gh}(n) e^{-in\omega} \tag{2.43}$$

of C_{gh} can be extended to a meromorphic function in a strip about the real line [Bal00]. The poles and their multiplicities depend only on the map F and do not depend on the functions g and h. These poles are Ruelle-Pollicott resonances. They determine the decay rates and oscillation frequencies of covariance and correlation functions, and the set of resonances is called the correlation spectrum $\sigma_{\text{corr}}(F)$ of F. Numerical evidence clearly shows that the exponential decay of correlations can continue to hold in nonuniformly hyperbolic systems, as shown by Figure 1.3. In that figure, the autocovariance functions for the coordinate functions x_1, x_2, and x_3 of the Lorenz flow are seen to decay exponentially.

To understand why $C_{gh}(n)$ decays exponentially fast as $n \to +\infty$, let us introduce the operator

$$(T_F \cdot g)(x) = g(F(x)). \tag{2.44}$$

That is, $(T_F \cdot g)(x)$ can be interpreted as the value of the observable g when the system is in the state $F(x)$, so the operator T_F evolves observables g forward in time. Without loss of generality, let us assume that $g = h$ and $\int g \, d\mu = 0$. Then $C_{gg}(n) = \int g \cdot T_F^n g \, d\mu$, and one might hope to mimic the theory of finite state Markov chains by finding an appropriate Banach space B with the following properties:

1. The operator T_F is bounded on B.

2. The spectral radius of $T_F : B \to B$ is equal to 1.

3. $\sigma(T_F)$ consists of a simple eigenvalue at 1 and $\rho = \sup\{|\lambda| : \lambda \in \sigma(T_F) \setminus \{1\}\}$ is strictly less than 1.

The conditions above would imply the exponential decay of correlations for the dynamical system in question.

The spectrum of a bounded operator $T : B \to B$ can depend strongly on the choice of the space B, and the choice of an appropriate function space on which to study T_F is a nontrivial task. Some seemingly natural choices, for example, will not reveal information on correlation decay. Consider the operator T_F acting on the Hilbert space $L^2(\mu)$: it is easy to check that T_F is unitary and hence its spectrum is a subset of the unit circle. (T_F is a non-unitary isometry if F preserves the measure μ but is not invertible.) This means that one cannot deduce the decay of correlations from the spectrum of T_F as an operator on $L^2(\mu)$. Note that this choice does yield some information about the dynamics — for example, applying the spectral theorem to $T_F : L^2(\mu) \to L^2(\mu)$ tells us that for every $g \in L^2(\mu)$ in the orthogonal complement of $\{h \in L^2(\mu) : h \circ F = h\}$,

$$\begin{aligned}
\frac{1}{N}\sum_{k=0}^{N-1} T_F^k g &= \frac{1}{N}\sum_{k=0}^{N-1} \int_{\sigma(T_F)} \lambda^k P(d\lambda) \cdot g \\
&= \frac{1}{N}\sum_{k=0}^{N-1} \int_{\sigma(T_F)\setminus\{1\}} \lambda^k P(d\lambda) \cdot g \\
&= \frac{1}{N}\int_{\sigma(T_F)\setminus\{1\}} \frac{1-\lambda^{N+1}}{1-\lambda} P(d\lambda) \cdot g,
\end{aligned}$$

since T_F is unitary and its spectrum lies in the unit circle. As $N \to \infty$, the right hand side converges to 0. Thus the limit

$$\lim_{N\to\infty}\left(\frac{1}{N}\sum_{k=0}^{N-1} g \circ F^k\right) \tag{2.45}$$

always exists in the sense of L^2, and equals the orthogonal projection of g onto the subspace of F-invariant functions. This beautiful result, the Mean Ergodic Theorem, was first proved by John von Neumann [Fol84]. However, the decay of correlations is a more subtle property and cannot be deduced this way.

Recent work of Blank, Keller, and Liverani [BKL02] shows that for Anosov systems, i.e. maps F which are uniformly hyperbolic on all of the manifold X, one can indeed construct a Banach space so that T_F has properties (1) through (3) listed above. The space they construct consist of functions which are smooth along unstable manifolds and are allowed to be much rougher along stable manifolds. The spectrum and eigenfunctions of the operator T_F then give detailed information about the decay of correlations and the existence of SRB measures. Earlier work on the exponential decay of correlations in dynamical systems relied on "weighted" transfer operators and symbolic dynamics to reduce hyperbolic dynamical systems to to purely expanding systems. Roughly speaking, this corresponds to chopping off half of each bi-infinite symbolic sequence so that it becomes a one-sided sequence. The one-sided shift map is then non-invertible but purely expanding, which simplifies the analysis and allows one to prove the existence of a smooth invariant measure. This measure is then transferred back to the original hyperbolic system via the symbolic coding. In contrast, the approach of Blank, Keller, and Liverani is more direct and natural. The weighted transfer operator technique does play a crucial role in the so-called thermodynamic formalism [Bal00, Bow75] and in constructing large deviations functionals for dynamical systems [Tak84].

2.6 Zero-noise limits of uniformly hyperbolic systems

Let $F : X \to X$ be a map, and let $\{p_\epsilon(x,\cdot) : x \in X\}$ be a family of transition probability measures on X depending on a parameter $\epsilon > 0$. Let (x_n) satisfy $x_{n+1} = F(x_n)$ (with a possibly random initial condition x_0), and define the measure $F_* p_\epsilon(x,\cdot)$ by setting $F_* p_\epsilon(x, A) = p_\epsilon(x, F^{-1}(A))$ for all Borel sets A. Let (x_n^ϵ) be the chain with transition probabilities $F_* p_\epsilon(x,\cdot)$ and initial condition $x_0^\epsilon = x_0$. The chain (x_n^ϵ) is then then a random perturbation of (x_n). This interpretation is particularly meaningful if $p_\epsilon(x,\cdot)$ is in some sense small when ϵ is small. Now let $g : X \to \mathbb{R}$ be a continuous function. Then it is natural to ask whether

$$\lim_{\epsilon \to 0} \int g \, d\mu_\epsilon = \int g \, d\mu_0, \tag{2.46}$$

where μ_ϵ is an invariant measure for the chain (x_n^ϵ) and μ_0 is an invariant measure for F.

The following simple result provides a partial answer (see [Kif88]):

Proposition 2.15. *Suppose the transition measures $p_\epsilon(x,\cdot)$ converge to δ_x uniformly in probabil-*

ity, i.e. for all $a > 0$,

$$\lim_{\epsilon \to 0} \sup_{x \in X} p_\epsilon(x, X \setminus B(x, a)) = 0. \tag{2.47}$$

Let $\{\mu_\epsilon\}$ denote a family of invariant measures for F_ϵ for $\epsilon > 0$. Then any weak limit μ_0 of $\{\mu_\epsilon\}$ as $\epsilon \to 0$ is F-invariant.

Proof. Choose a subsequence $\epsilon_k \to 0$ such that $\mu_{\epsilon_k} \to \mu_0$ weakly as $k \to \infty$. As X is compact Hausdorff, every Borel probability measure on X is a Radon measure and it suffices to verify that $\int g \circ F \, d\mu_0 = \int g \, d\mu_0$ for every $g \in C(X)$:

$$\begin{aligned} \left|\int g \circ F \, d\mu_0 - \int g \, d\mu_0\right| &= \lim_k \left|\int g \circ F \, d\mu_{\epsilon_k} - \int g \, d\mu_{\epsilon_k}\right| \\ &= \lim_k \left|\int g \circ F \, d\mu_{\epsilon_k} - E \int g \circ F_{\epsilon_k} \, d\mu_{\epsilon_k}\right| \\ &\leq \lim_k E \int |g \circ F - g \circ F_{\epsilon_k}| \, d\mu_{\epsilon_k} \\ &= \lim_k \Big[E_{(d(F,F_{\epsilon_k})<a)} \left(\int |g \circ F - g \circ F_{\epsilon_k}| \, d\mu_{\epsilon_k}\right) \\ &\quad + E_{(d(F,F_{\epsilon_k})\geq a)} \left(\int |g \circ F - g \circ F_{\epsilon_k}| \, d\mu_{\epsilon_k}\right)\Big]. \end{aligned} \tag{2.48}$$

Choosing a small enough so that $|g(x) - g(y)| < b$ whenever $d(x, y) < a$ yields

$$\left|\int g \circ F \, d\mu_0 - \int g \, d\mu_0\right| \leq b + \lim_{k \to \infty} 2P\left(d(F, F_{\epsilon_k}) > a\right). \tag{2.49}$$

The second term vanishes by assumption and $b > 0$ was arbitrary, so μ_0 is invariant. □

Thus, the real problem is to figure out which invariant measures of F can arise as weak limits of $\{\mu_\epsilon\}$ as $\epsilon \to 0$. There is one special case which is well understood mathematically: Suppose the unperturbed map F has an Axiom A attractor on which it is topologically transitive and the random perturbation is given by transition probability measures which have continuous probability densities. Then the perturbed systems have unique invariant measures μ_ϵ and the limiting measure μ_0 must be the SRB measure of the original map F [Kif88]. More precisely:

Theorem 2.16. *Let $F : X \to X$ be a smooth diffeomorphism on a compact Riemannian manifold X and $\{p_\epsilon(x, \cdot) : x \in X\}$ a family of measures. Suppose the following conditions hold:*

1. *For each $g \in C(X, \mathbb{R})$, $\int g(y) \, p_\epsilon(x, dy)$ converges uniformly to $g(x)$ as $\epsilon \to 0$.*
2. *For each $\epsilon > 0$, $p_\epsilon(x, \cdot)$ is absolutely continuous with respect to the Lebesgue measure of X.*

Then the Markov chain defined by the transition probability density $F_* p_\epsilon(x, \cdot)$ has a unique invariant measure μ_ϵ, and the weak limit of μ_ϵ as $\epsilon \to 0$ is a convex combination of SRB measures of F.

Kifer proved that by showing that the weak limit μ_0 must be absolutely continuous with respect to μ_{SRB}. The ergodicity of μ_{SRB} then shows that $\mu_0 = \mu_{\text{SRB}}$.

The absolute continuity of transition probability measures means that the dynamics is perturbed in all spatial directions (though in a possibly inhomogeneous way). However, as explained in the Introduction, it is frequently desirable to use random perturbations which are limited to the unstable directions only. In the solenoid example, we can check directly that perturbations in the stable directions (spanned by $\frac{\partial}{\partial r}$ and $\frac{\partial}{\partial s}$) do not perform any useful function in the computation. A small amount of noise (uniform noise on the order of 10^{-15}) in the unstable direction $\frac{\partial}{\partial \theta}$ suffices to produce a much reliable picture of the attractor. Precise theorems in the case of singular random perturbations are not known to the author.

It is easy to show that if μ_ϵ is an invariant measure for the perturbed system, then any weak limit of μ_ϵ is invariant under F (see Chapter 4). The bulk of Kifer's proof is involved in showing that any weak limit is absolutely continuous with respect to the SRB measure μ_{SRB}. The ergodicity of μ_{SRB} then implies that the weak limit must coincide with μ_{SRB} itself. The key point is that this strategy can only show that the limiting measure is a probability measure; it does not give rates of convergence. There are, in fact, very few such estimates.

The use of noise in numerical studies of chaotic systems is not a new idea; additive noise is a natural way to ensure that simulated orbits visit different parts of the attractor more quickly with approximately the correct frequencies. The fundamental mathematical results date back to the work of Kifer [Kif74, Kif88] and Freidlin and Wentzell [FW98]. Building on Kifer's theorem, Hunt [Hun96], Kifer [Kif97], and Dellnitz et. al. [DJ99] have suggested the use of random perturbations as the basis of reliable numerical algorithms. However, these papers consider only perturbations which have absolutely continuous (with respect to Lebesgue) transition probability measures. The analysis of degenerate perturbations is much more challenging.

2.7 The limits of hyperbolic theory

It is important to remember that real physical systems rarely satisfy the axioms of uniform hyperbolicity. Most real physical systems have parameter regimes in which they experience *intermittency*, alternating between chaotic and periodic or quasiperiodic behav-

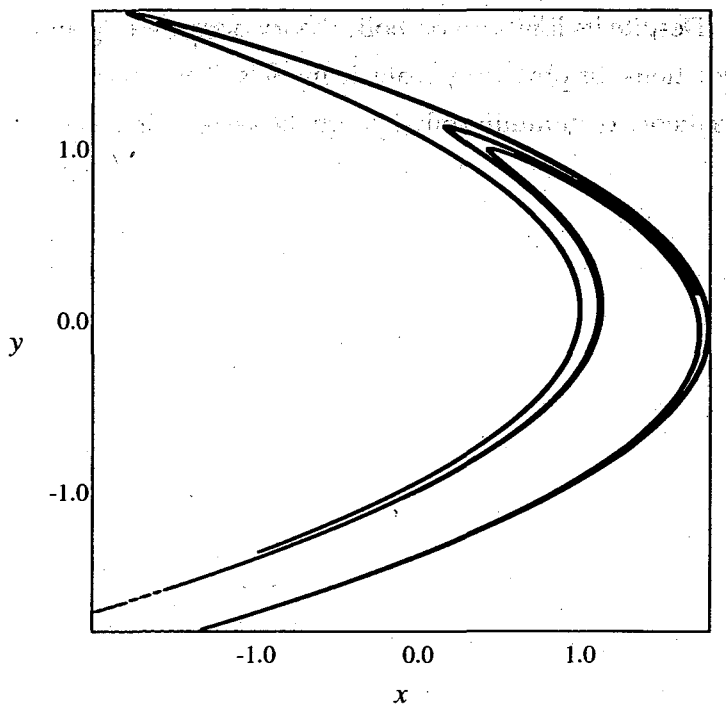


Figure 2.6: The Hénon attractor.

ior[1] [Ott00]. Even very simple mathematical models can exhibit nonuniform hyperbolic behavior and become very difficult to analyze. For example, consider the Hénon map

$$F_{\text{hénon}}(x_1, x_2) = \begin{bmatrix} a - x_1^2 + bx_2 \\ x_1 \end{bmatrix}, \tag{2.50}$$

with $a = 1.4$ and $b = 0.3$. It has two hyperbolic fixed points and an apparent attractor; it is shown in Figure 2.6. The strange attractor has well-known self-similar structures [GH83]. If the attractor were uniformly hyperbolic, the unstable manifolds of the fixed points would be dense inside the attractor. But the attractor folds back on itself ever more densely, so it is impossible to have stable manifolds which are everywhere *transverse* to the unstable manifolds with angles bounded away from 0, as is the case for hyperbolic attractors [KH95]. See Guckenheimer and Holmes [GH83] for a more extensive discussion of strange attractors.

[1] I am grateful to Dr. Tom Weaver of the Hertz Foundation for pointing this out.

Despite its limits, hyperbolic theory does yield qualitative insights and quantitative predictions for physically realistic models. There has also been much progress in the rigorous theory of nonuniformly hyperbolic systems in recent years [You95].

Chapter 3

Random Perturbations in Numerical Simulations

This chapter explores the use of noise in numerical algorithms for computing statistical averages of observables. Most of this discussion is based on numerical studies of the Lorenz system (1.2) [GH83, Lor63] and other low-dimensional systems. Analytical results remain far from optimal and have been summarized in Chapter 4.

3.1 Asymptotically unstable subspaces & Lyapunov exponents

The calculations described in this chapter focus on the computation of Lyapunov exponents because of their intrinsic interest as dynamical invariants. Furthermore, in order to compute Lyapunov exponents, we need to first compute the invariant subspaces $E^i(x)$ (see §2.3), and the computation of these invariant subspaces provides valuable insight into the local geometric structure of dynamical systems.

The easiest way to estimate $E^u(x)$ is to try to lift the flow to the orthonormal frame bundle. In concrete terms, this means that given a solution $x : [0, +\infty) \to X$ of the equations of motion

$$\begin{aligned} \dot{x}(t) &= f(x(t)) \\ x(0) &= x_0, \end{aligned}$$

we want to find equations of motion

$$\begin{aligned}
\dot{x}(t) &= f(x(t)),\\
\dot{v}_1(t) &= \xi_1(v_1(t), ..., v_n(t)),\\
\dot{v}_2(t) &= \xi_2(v_1(t), ..., v_n(t)),\\
&\vdots\\
\dot{v}_n(t) &= \xi_2(v_1(t), ..., v_n(t)),
\end{aligned}$$

so that the vectors $(v_1(t), ..., v_k(t))$, $1 \leq k \leq n$, form an orthonormal basis for the subspace $E^k(x(t))$ associated with the k largest Lyapunov exponents. In fact, it is possible to find the desired equations in a "triangular" form

$$\begin{aligned}
\dot{x}(t) &= f(x(t)),\\
\dot{v}_1(t) &= \xi_1(v_1(t)),\\
\dot{v}_2(t) &= \xi_2(v_1(t), v_2(t)),\\
&\vdots\\
\dot{v}_n(t) &= \xi_2(v_1(t), ..., v_n(t)),
\end{aligned}$$

so that if one is only interested in computing $E^k(x(t))$ for some $k < n$, then not all $n^2 + n$ equations are needed.

The vector field for the orthonormal basis $(v_1(t), ..., v_k(t))$ of $E^k(x(t))$ is easiest to explain in the form of a simple algorithm:

Algorithm 3.1 (Flow Field for Orthonormal Frames).

Input:

vector field f

point $x \in X$ *(system state; dimension of X is n)*

integer k

1: $\dot{x} := f(x)$

2: **for** $i := 1$ to k **do**

3: $\quad \dot{v}_i := Df(x) \cdot v_i$

4: **for** $j := 1$ to $i-1$ **do**

5: $\dot{v}_i := \dot{v}_i - (\dot{v}_i \cdot v_j + v_i \cdot \dot{v}_j)v_j$

6: **end for**

7: $\dot{v}_i := \dot{v}_i - (\dot{v}_i \cdot v)v$

8: **end for**

Output: vectors $\dot{x}, \dot{v}_1, ..., \dot{v}_k$ *(time derivatives)*

The algorithm is a straightforward consequence of the orthonormality conditions and the requirement that $v_1, ..., v_k$ span E^k. The inner loop enforces the orthogonality condition, while the last statement enforces the normality of the basis vectors. In hyperbolic flows, the vector $v_1(x)$ will converge fairly quickly to the maximally unstable direction $E^1(x)$ associated with the largest Lyapunov exponent. Because $v_2(x)$ is kept orthogonal to $v_1(x)$, it will not point in the direction associated with the second largest exponent. The vectors $v_1(x)$ and $v_2(x)$ together, however, span the subspace $E^2(x)$ associated with λ_1 and λ_2. A simple induction argument establishes the correctness of this algorithm.

We can compute the largest k Lyapunov exponents using Algorithm 3.1. We only need to compute the long-time averages of the partial traces of the Jacobian Df along the spans of $(v_1, ..., v_i)$ for $i = 1, ..., k$ [WSSV85]: since

$$\begin{aligned}\lambda_1 + ... + \lambda_k &= \lim_{T\to\infty} \frac{1}{T}\left(\int_0^T \operatorname{tr}\left(Df(x(t)) \restriction_{E^k(x(t))}\right) dt\right) \\ &= \lim_{T\to\infty} \frac{1}{T}\left(\sum_{i=1}^k \int_0^T v_i(t)^* \cdot Df(x(t)) \cdot v_i(t)\, dt\right),\end{aligned}$$

it follows that

$$\lambda_i = \lim_{T\to\infty} \frac{1}{T}\int_0^T v_i(t)^* \cdot Df(x(t)) \cdot v_i(t)\, dt. \tag{3.1}$$

(The row vector v^* is the transpose of the column vector v.)

In practice, we cannot simply insert Algorithm 3.1 into an ODE solver and expect it to work. This is because Algorithm 3.1, when implemented as part of an ODE solver, contains a vicious feedback mechanism: because of numerical errors, $(v_1, ..., v_k)$ will become slightly nonorthogonal after a finite number of steps. This forces the ODE integrator to take ever smaller steps to maintain orthogonality. Eventually, the computation grinds to a halt. Thus, we must periodically reorthonormalize $(v_1, ..., v_k)$. This can be quite expensive to do for higher dimensional systems. If one is only interested in computing the

Lyapunov exponents $\{\lambda_1, ..., \lambda_k\}$ and not the invariant subspace $E^k(x)$, then it is more efficient to simply remove the constraint that the basis $(v_1, ..., v_k)$ stays orthonormal. This is equivalent to computing the Lyapunov exponents for the time-T map ϕ_T for a fixed $T > 0$ instead of directly for the flow [BGS76, GPL90]. More precisely:

Algorithm 3.2.

Input:

flow map $\phi_T : X \to X$
point $x \in X$ *(initial condition)*
integer k
integer N_{run} *(running time)*
orthonormal vectors $\{v_1, ..., v_k\}$ *(initial frame)*

```
for i := 1 to k do
    λ̂_i := 0  (initialize)
end for
for n := 1 to N_run do
    for i := 1 to k do
        v_i := Dφ_T(x) · v_i
    end for
    for i := 1 to k do
        for j := 1 to i − 1 do
            v_i := v_i − (v_i · v_j) v_j  (reorthogonalize)
        end for
        λ̂_i := λ̂_i + log ||v_i|| / (N_run · T)
        v_i := v_i / ||v_i||  (renormalize)
    end for
    x := φ_T(x)  (continue with the next step)
end for
```

Output: $\{\hat{\lambda}_1, ..., \hat{\lambda}_k\}$ *(estimates of Lyapunov exponents)*

This is the standard algorithm one finds in [GPL90] and references therein.

Algorithm 3.2 can also be used to compute the Lyapunov exponents of an arbitrary map F by replacing ϕ_T with F. Algorithm 3.2 effectively performs a QR factorization

of the Jacobian DF using Gram-Schmidt orthogonalization. There is a more sophisticated variant of this method which uses Householder transformations in place of Gram-Schmidt orthogonalization; see [GPL90] and references therein.

3.2 Time averaging

Traditionally, the computation of the average value of an observable g with respect to an ergodic invariant measure proceeds by performing long-time simulations of the system in question and computing time averages. This simple procedure is formalized below in pseudo code:

Algorithm 3.3 (Time Averaging).

Input:
function $g : X \to \mathbb{R}$ *(observable of interest)*
map $F : X \to X$ *(dynamics)*
point $x_0 \in X$ *(initial condition)*
integer N_{init} *(initialization time)*
integer N_{run} *(running time)*

```
x := x_0
for n := 0 to N_init − 1 do
    x := F(x)
end for
ĝ := 0
for n := 1 to N_run do
    x := F(x)
    ĝ := ĝ + g(x)/N_run
end for
```

Output: real $\hat{g}$ *(an estimate of $\int g \, d\mu_{SRB}$)*

Note that the initialization loop is nearly identical to the averaging loop. For the sake of clarity, descriptions of algorithms in later sections of this chapter will abbreviate the initialization loop.

In order to apply Algorithm 3.3 to realistic problems reliably, we must be able to

estimate the numerical error in the result. As

$$\left|\hat{g} - \int g \, d\mu_{\text{SRB}}\right| \leq \sqrt{\frac{2\, C_{gg}(0)\, \tau_{\text{int}}(g)}{N_{\text{run}}}},$$

where $\tau_{\text{int}}(g)$ is the integrated autocorrelation time (see §2.5.2), the choice of the initialization time N_{init} and the running time N_{run} depend on knowledge of the decay rate of the autocorrelation function of the observable g. In exceptional cases, it is possible to obtain analytical bounds on this rate of decay. In most cases, however, analytical estimates are unavailable, and we must calculate the decay rate from empirical data.

A practical problem which arises frequently is that the autocorrelation function may decay very slowly, either with a small exponential time constant or even subexponentially. This corresponds to $\left|\sigma\left(T_F\right) \setminus \left\{1\right\}\right| \sim 1$. This is rather common in statistical mechanics and quantum field theory, as correlation lengths in lattice models diverge to infinity near phase transitions. For classical dynamical systems, a similar phenomenon exists. For example, polynomial decay is often expected in systems with nonuniform hyperbolicity or intermittency, such as the intermittent interval maps (see §3.5 and [LSV99, Ott00]). These difficulties will be addressed later in this chapter.

Figure 3.1 illustrates the result of applying Algorithm 3.3 to the Lorenz flow to compute Lyapunov exponents. As can be seen, the time averages $\left\{\hat{\lambda}_i(T)\right\}$ converge to well-defined limits as T increases. In the figure, the estimates

$$\hat{\lambda}_i(T) = \frac{1}{T}\int_0^T v_i^*(t) Df(x(t)) v_i(t)\, dt,$$

are plotted against the time interval T. The initial condition is taken to be a small ($\sim 10^{-15}$) perturbation of the origin. The estimates $\hat{\lambda}_i(T)$ will converge almost surely to the Lyapunov exponents as $T \to \infty$. Furthermore, $\lambda_2(T)$ will converge to 0 as $T \to \infty$, since flow lines are center manifolds (see §2.3).

3.3 Random kicks in unstable directions

At first sight, Algorithm 3.3 may appear to be the only practical way to compute the expectation value $\int g \, d\mu_{\text{SRB}}$ of an observable g with respect to μ_{SRB}. However, when the number of dimensions is sufficiently small, one can apply the subdivision algorithm of Dellnitz and coworkers [DH97, DHJR97, DJ99] to discretize the transfer operator. In many cases, it is also possible to apply periodic orbit theory and cycle expansions

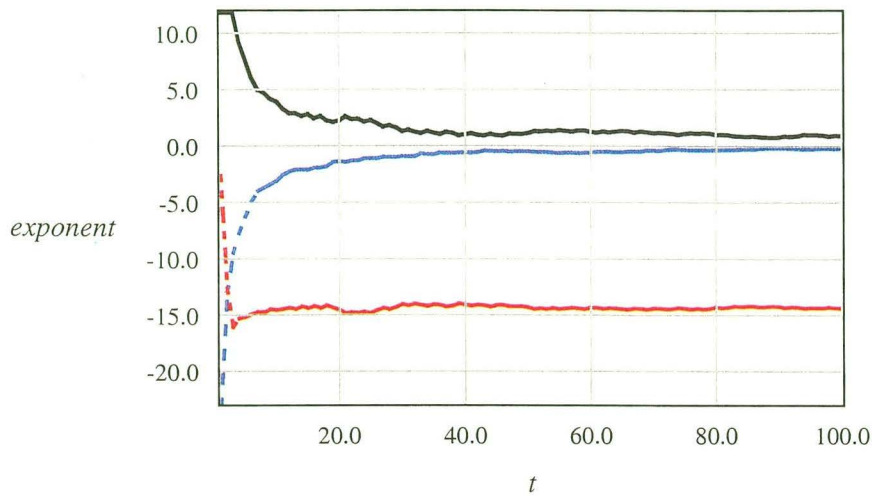


Figure 3.1: Estimates of Lyapunov exponents as a function of simulation time interval T.

[CCP97, Dor99, Gas98]. In moderately high dimensions, these alternative methods become less efficient and their implementation requires overcoming a host of theoretical and algorithmic challenges.

Notice that the Metropolis algorithm 2.14 extends naturally to continuous state space Markov chains with probability densities instead of probability vectors. Its implementation requires only that one can compute the relative probability of two different states. But Equation (2.30) provides a formula for the relative probability of two states lying on the same unstable manifold. Is there a way to use Equation (2.30) to construct a Metropolis algorithm for SRB measures? Specifically, can we perform Metropolis sampling by proposing moves x' along $W^u(x)$ and then comparing the probabilities of x and x' by evolving both points forward in time and computing the local expansion rates along each trajectory?

Let F denote a map with a hyperbolic attractor K and an associated SRB measure μ_{SRB}. Let x be a point in K. Fix $\epsilon > 0$ and let $m^u_{x)}$ be the normalized surface measure on $W^u_\epsilon(x)$ induced by the Riemann metric of the phase space X. In order to construct a

Metropolis algorithm using (2.30), we need to understand the behavior of proposal measures as they evolve along unstable manifolds. Let $q_x(\cdot)$, then, be a probability density on $W^u_\epsilon(x)$ (defined with respect to m^u_x) and let ρ_x denote the density of μ_{SRB} on $W^u_\epsilon(x)$. Then the acceptance ratio $A_q(x, x')$ associated with the family of proposal measures $\{q_x\}$ is

$$A_q(x, x') = \min\left(\frac{q_{x'}(x) \cdot \rho_x(x')}{q_x(x') \cdot \rho_x(x)}, 1\right). \tag{3.2}$$

Now, for each x we can define a new proposal measure $q^{(1)}_x$ by

$$q^{(1)}_x(x') = \frac{q_{F^{-1}(x)}(F^{-1}(x'))}{J^u_F(F^{-1}(x'))}. \tag{3.3}$$

The measure $q^{(1)}_x$ is obtained by taking the distribution q_x on the local unstable manifold $W^u_\epsilon(F^{-1}(x))$ and pushing it forward to $W^u_\epsilon(x)$ using the map F. Equations (3.3) and (2.30) imply that

$$A_{q^{(1)}}(x, x') = A_q(F^{-1}(x), F^{-1}(x')). \tag{3.4}$$

But this means that if we repeat the construction and define

$$q^{(n)}_x(x') = \frac{q_{F^{-n}(x)}(F^{-n}(x'))}{J^u_{F^n}(F^{-n}(x'))}.$$

By induction,

$$A_{q^{(n)}}(x, x') = A_q(F^{-n}(x), F^{-n}(x')).$$

Since x and x' belong to the same unstable manifold, $d\left(F^{-n}(x), F^{-n}(x')\right) \to 0$ as $n \to \infty$. It follows that pushing a proposal distribution forward under the map F improves the acceptance probability. Heuristically, then, we expect to be able to accept every proposal if we use the family of measures $\{q^{(N)}_x\}$ as our proposal measures for sufficiently large N.

The discussion above leads to the following algorithm (for simplicity, set $q_x(\cdot) \equiv 1$ and $N = 1$):

Algorithm 3.4 (Random Perturbations in Unstable Directions).

Input:

function $g : X \to \mathbb{R}$ *(observable of interest)*

map $F : X \to X$ *(dynamics)*

real number $\epsilon > 0$ *(noise amplitude)*

point $x_0 \in X$ *(initial condition)*

integer N_{init} *(initialization time)*

integer N_{run} *(running time)*

vector subspace $\hat{E}^u$ *(approximate unstable direction represented e.g. by basis vectors)*

$x := x_0$

$\hat{E}^u :=$ random orthonormal frame

do initialization loop for $N_{\text{init}} - 1$ steps *(this is nearly identical to the averaging loop below)*

$\hat{g} := g(x)$

for $n := 0$ to $N_{\text{run}} - 1$ **do**

  $\hat{E}^u := DF(x) \cdot \hat{E}^u$

  $x := x + \delta x$ *(δx is a small uniform random sample drawn from uniform probability measure supported on the ball $B^u(x, \epsilon) = B(x, \epsilon) \cap (x + \hat{E}^u)$.)*

  $x := F(x)$

  $\hat{g} := \hat{g} + g(x)/N_{\text{run}}$

end for

Output: $\hat{g}$ *(an estimate of $\int g \, d\mu_\epsilon \approx \int g \, d\mu_{SRB}$)*

Note that this is not, strictly speaking, a Monte Carlo algorithm. One may be tempted to use backward iterates of F to really implement a Monte Carlo method with the proposal measure $q^{(n)}$. That is unlikely to work because the unstable manifolds of F^{-1} are the stable manifolds of F, which (when F is dissipative) generally fill up a set of positive measure, so iterating F^{-1} is an extremely unstable process. Note also that it is generally impossible to propose moves that lie exactly on the unstable manifold; Algorithm 3.4 proposes moves along approximately unstable directions $\hat{E}^u$. The action of the map F then pushes the proposals closer to the unstable manifolds.

Algorithm 3.4 is extremely simple: it applies periodic random perturbations to the flow under study. The correctness of Algorithm 3.4, then, is equivalent to the stability of SRB measures under small random perturbations in the sense that the invariant measures μ_ϵ of the perturbed chain (x_n^ϵ) converge weakly to μ_{SRB} as $\epsilon \to 0$. Unfortunately, Theorem 2.16 does not apply here, as the random perturbations used here were directed in (approximately) unstable directions and cannot be absolutely continuous unless F is everywhere expanding (i.e. $\dim E^u = \dim X$). However, results appearing in the next chapter imply that Algorithm 3.4 is consistent in the following sense: if we first let $N_{\text{run}} \to \infty$ and then let $\epsilon \to 0$, then the output $\hat{g}$ of Algorithm 3.4 converges to $\int g \, d\mu_{\text{SRB}}$. Also, as noted

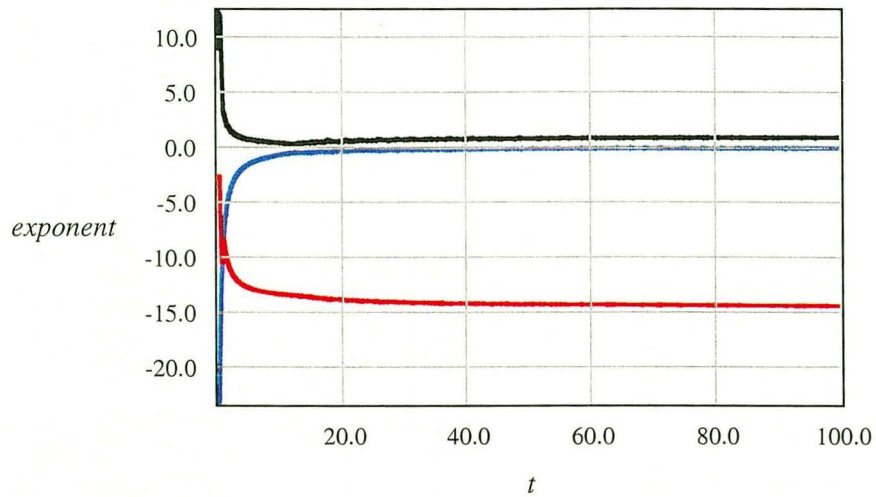


Figure 3.2: Estimates of Lyapunov exponents as a function of simulation time interval T.

in the Introduction, the random perturbations introduced in Algorithm 3.4 may help reduce the correlation time of the perturbed Markov chain.

Let us now combine Algorithms 3.2 and 3.4 and compute the Lyapunov exponents for the Lorenz flow with the standard parameters $\sigma = 10$, $b = \frac{8}{3}$, and $r = 28$. Figure 3.2 shows the partial sums $\hat{\lambda}_i$, defined in Algorithm 3.2. The flow is sampled at intervals of $T = 0.02$, and after each sample the state variable x is perturbed by uniformly distributed random variables of amplitude $\epsilon = 0.01$ in the estimated unstable direction $\hat{E}^u$. As in Figure 3.1, the initial condition x_0 is a small ($\sim 10^{-15}$) perturbation of the origin.

Because x_0 is near a hyperbolic fixed point, it takes some time for the computed trajectory to leave the vicinity of the origin and explore the rest of the attractor. This is the reason for the appearance of the "plateau" in the top curve of Figure 3.1 for $t < 5$. In contrast, in Figure 3.2, the addition of noise helped the partial sums converge much more quickly. One may object that the choice of x_0 is unnecessarily pessimistic, but it is important to remember that frequently, the phase space of a dynamical system is partitioned into multiple ergodic components. Choosing initial conditions near a known fixed point is

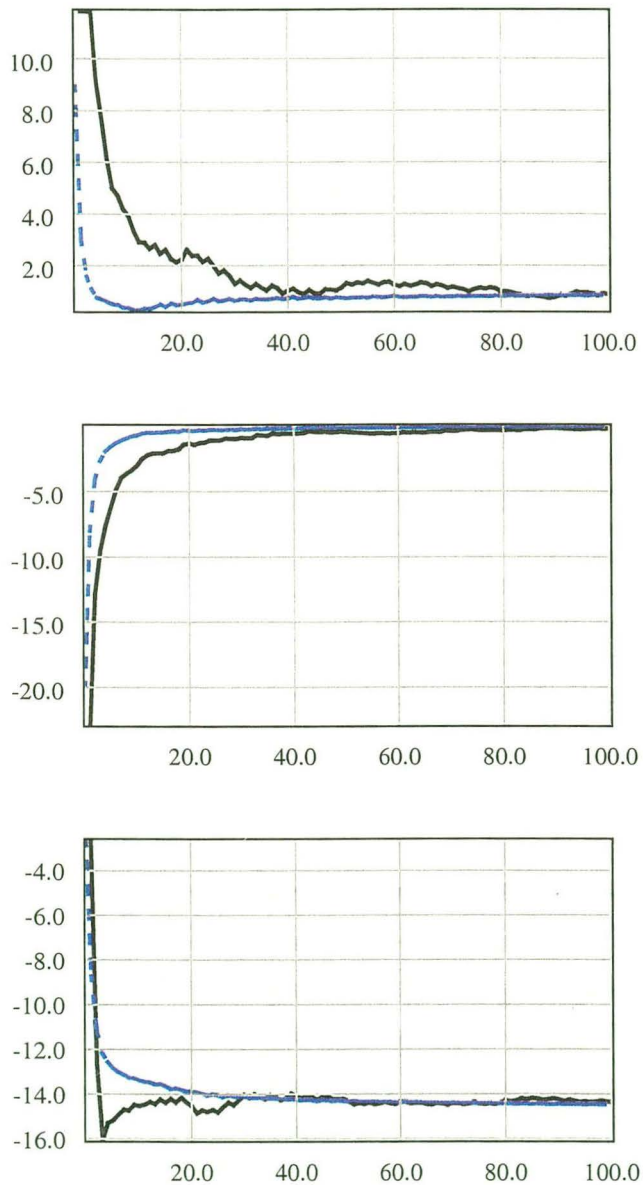


Figure 3.3: Comparison of estimated Lyapunov exponents as functions of simulation time interval. The solid black line represents the results computed with $\epsilon = 0$, while the dashed blue line represents the results with $\epsilon = 0.01$.

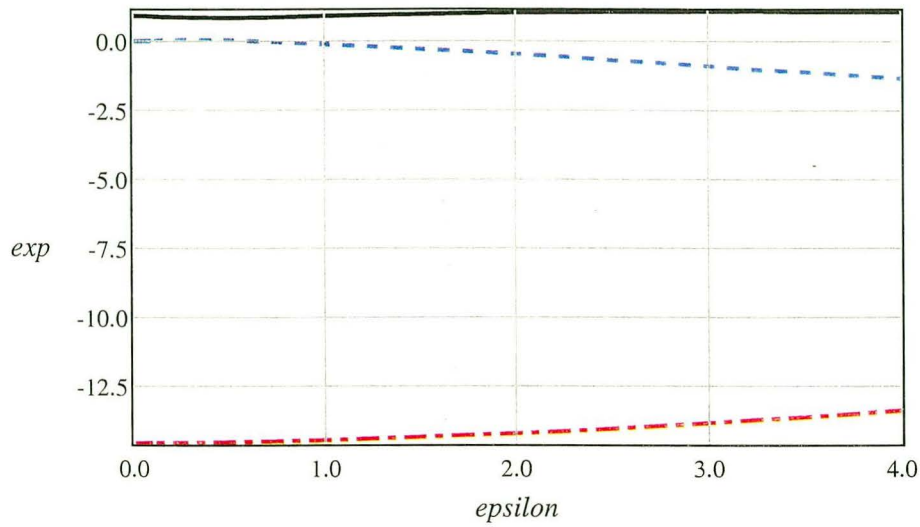


Figure 3.4: Estimated Lyapunov exponents for the Lorenz flow kicked by uniform noise of amplitude ϵ in the asymptotically unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$.

a simple way of ensuring (with high probability) that the statistical information we compute pertain to a single ergodic component rather than a convex combination of ergodic invariant measures.

Figure 3.4 shows the estimated Lyapunov exponents, computed by sampling the perturbed Lorenz flow at intervals of $T = 0.02$ for 10^6 steps, as a function of the noise amplitude ϵ. Figure 3.5 shows the corresponding empirical decay times $\hat{\tau}_{\text{exp}}$ of the autocorrelation functions for the local expansion rates. The empirical decay time $\hat{\tau}_{\text{exp}}$ is defined here to be the time required for the correlation function to decay to $\frac{1}{20}$ of its value at $t = 0$. Figure 3.6 shows the empirical estimate $\hat{\tau}_{\text{int}}$ of the integrated autocorrelation times as a function of ϵ; $\hat{\tau}_{\text{int}}$ is computed by numerically integrating the normalized autocovariance function from 0 to $\hat{\tau}_{\text{exp}}$ using the trapezoid rule. Finally, Figure 3.8 shows the variances of the local expansion rates.

The exponential decay time decreases with increasing ϵ, as expected, though $\hat{\tau}_{\text{int}}$ increases with increasing ϵ. These observations are confirmed by the spectral power density $\hat{C}(f) = \int_{-\infty}^{+\infty} C(t)\, e^{-i2\pi ft}\, dt$, shown in Figure 3.7: as ϵ increases, the spectral densities (1) become more smooth, which reflects a faster decay of correlations; (2) the peaks be-

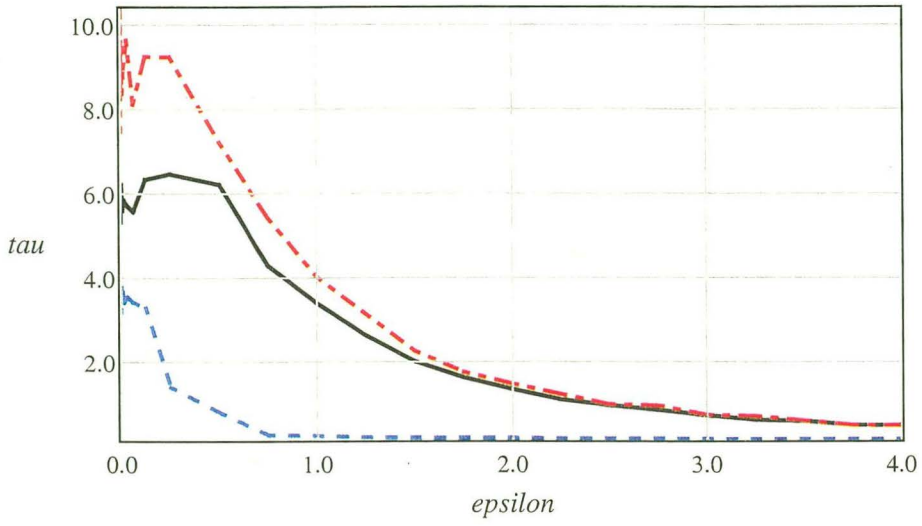


Figure 3.5: Decay times of the Lorenz flow periodically kicked by uniform noise of amplitude ϵ in the asymptotically unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$.

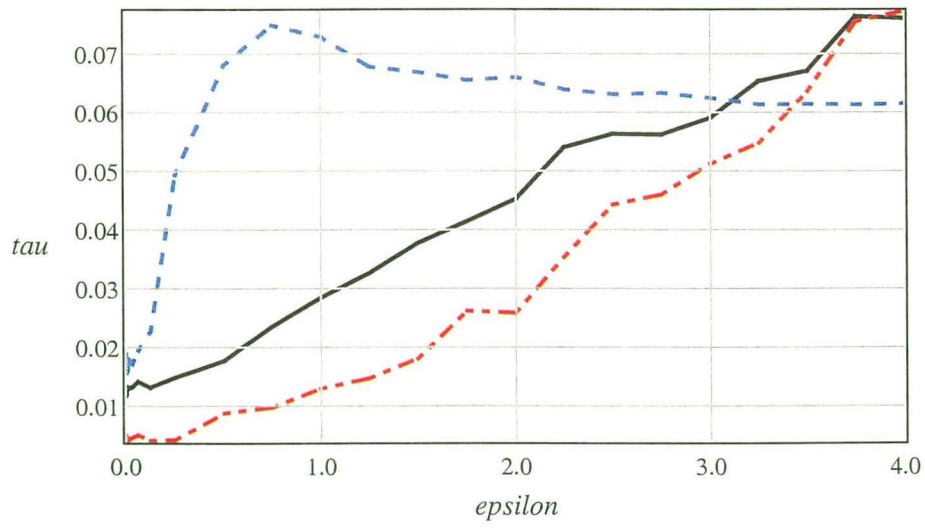


Figure 3.6: Integrated correlation times of the Lorenz flow periodically kicked by uniform noise of amplitude ϵ in the asymptotically unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$.

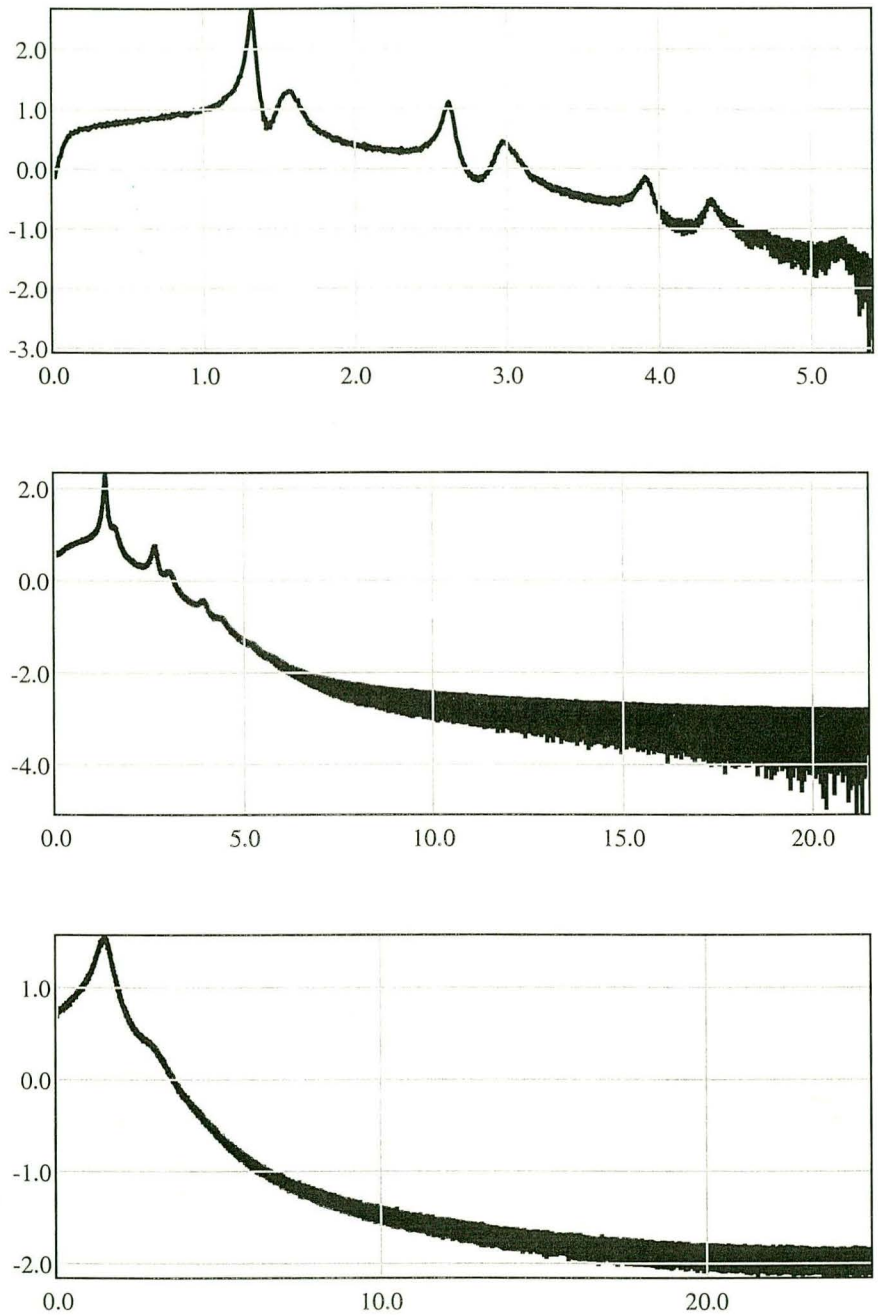


Figure 3.7: Spectral power densities $\log_{10}(\hat{C}(f))$ associated with $\hat{\lambda}_3$ for the Lorenz flow, plotted against the frequency f, for $\epsilon \in \{0.0, 1.0, 4.0\}$.

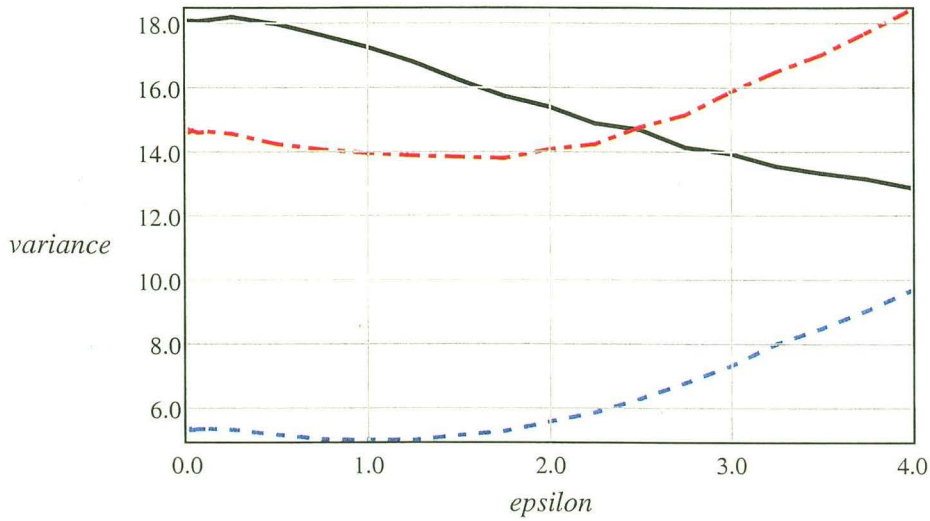


Figure 3.8: Variances of $\hat{\lambda}_k$ for the Lorenz flow periodically kicked by uniform noise of amplitude ϵ in the asymptotically unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$.

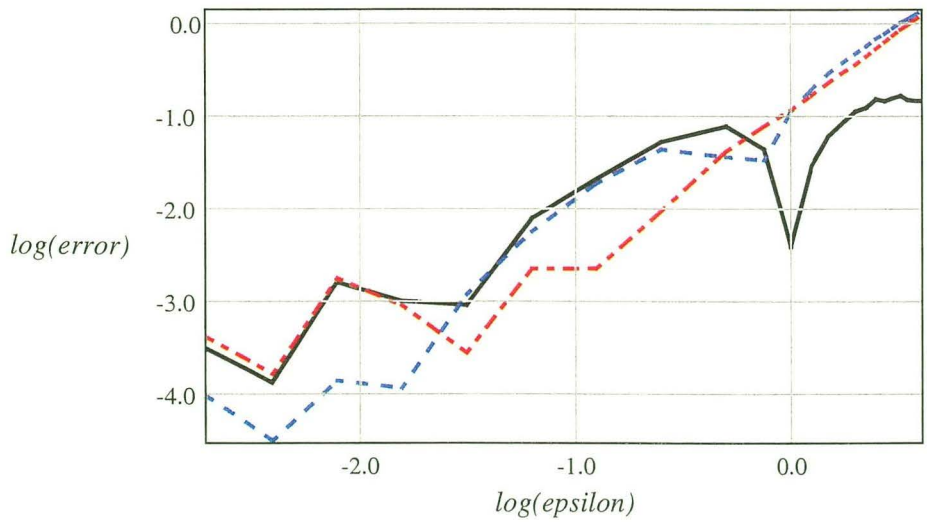


Figure 3.9: Approximate errors for the estimated Lyapunov exponents of the Lorenz flow periodically kicked by uniform noise of amplitude ϵ in the asymptotically unstable direction. The solid curve is associated with $\hat{\lambda}_1$, the dashed curve with $\hat{\lambda}_2$, and the dashed-dot curve with $\hat{\lambda}_3$.

come broader, which reflects the movement of resonances away from the real axis; and (3) $\hat{C}(0) = \tau_{\text{int}}$ increases slightly. Note that the numerical computation of τ_{int} can be somewhat subtle; see Sokal [Sok97] and references given there for a more detailed discussion. The qualitative features of Figure 3.6 do not depend sensitively on the numerical quadrature scheme or the integration limits. The roughness of $\hat{\tau}_{\text{int}}$ as a function of ϵ may reflect the instability of parts of the correlation spectrum with respect to random perturbations, as suggested in [BK98].

It is possible to extract some resonances from the correlation data, for example using Prony's method [MSW89, WMS88]. However, a straightforward application of Prony's method to the numerical data produced only an estimate of the dominant, slowest-decaying resonance, which can be extracted by the more naive means described above. It is an open problem to develop numerical algorithms which can reliably estimate fast-decaying ("broad") overlapping resonances, either by processing correlation data (or equivalent their Fourier transforms) or by other means.

In statistical mechanics, the integrated correlation time τ_{int} is usually expected to be comparable to the exponential decay time τ_{exp} except when the system is near a phase transition [Sok97]. In contrast, we see here that the τ_{exp} is much larger than τ_{int} when no noise is added. With the addition of noise, τ_{exp} quickly decreases while τ_{int} increases slightly. Thus with a suitable amount of noise, we can decrease τ_{exp} without increasing the already-small τ_{int} or changing $\{\hat{\lambda}_i\}$ too much.

Covariance functions can be expanded as

$$C_{gh}(n) = \sum_{\omega \in \sigma_{\text{corr}}(F)} a_{gh}(\omega) e^{in\omega} + R(n) \tag{3.5}$$

where $R(n) \sim \lambda_R^n$ is an exponentially decaying error term with decay rate $\lambda_R > \text{Im}\,[\omega]$ for all $\omega \in \sigma_{\text{corr}}(F)$. The exponential decay time τ_{exp} is therefore determined by the element of $\sigma_{\text{corr}}(F)$ with the smallest imaginary part. In contrast, the integrated correlation time is roughly equal to

$$\begin{aligned} \frac{1}{2} + \sum_{n=1}^{\infty} C_{gh}(n) &= \frac{1}{2} + \sum_{n=1}^{\infty} \left(\sum_{\omega \in \sigma_{\text{corr}}(F)} a_{gh}(\omega) e^{in\omega} \right) \\ &= \frac{1}{2} + \sum_{\omega \in \sigma_{\text{corr}}(F)} \frac{a_{gh}(\omega) e^{i\omega}}{1 - e^{i\omega}}. \end{aligned}$$

Thus, while τ_{exp} depends only on the part of $\sigma_{\text{corr}}(F)$ closest to the real line, τ_{int} depends on the rest of the correlation spectrum. Since τ_{exp} and τ_{int} differ so much when $\epsilon = 0$ and respond to random perturbations in qualitatively distinct ways, it is likely that random perturbations have a much larger effect on resonances which lie near the real axis than on those which lie farther away, and that the correlation spectrum of the Lorenz flow is concentrated away from the real axis but contains a few resonances with small imaginary parts. Whether this effect of random perturbations is true for other dynamical systems is unclear, but it appears to be consistent with the results of Blank and Keller [BK98]. The conclusion we can draw from these figures and from Table 3.1 is that for the Lorenz flow and other systems with relatively strong mixing properties, the addition of noise can help decrease the exponential decay time but not the integrated autocorrelation time. In fact, the standard deviations of the estimated Lyapunov exponents (see Table 3.1) actually increase with increasing ϵ.

In Monte Carlo studies of lattice models, the exponential decay time τ_{exp} is usually associated with the amount of time needed for equilibration. The integrated autocorrelation time, on the other hand, determines the number of steps we need to simulate the Markov chain before obtaining a new effectively independent sample. For dynamical systems, standard results about transfer operators (see §2.5.2, §2.5.3, and [Bal00]) show that this understanding of τ_{exp} and τ_{int} is still correct, though if the initial condition x_0 is far from the support of μ_{SRB} then we also need to take into account the attraction time τ_{att}.

In order to compute Lyapunov exponents using Algorithm 3.4, one must work with the orthonormal frame bundle lift of the flow (see Algorithm 3.1). This is because the calculation of Lyapunov exponents requires information about the jacobian Df. Thus, a straightforward implementation of Algorithm 3.4 for computing Lyapunov exponents may appear to require $D^2 f$. However, it is easy to see that the bundle degrees of freedom are enslaved to motion in the base: for subspaces E_1 and E_2 in general position, with $\dim E_1 = \dim E_2$, $d(DF^n E_1, DF^n E_2) \to 0$ as $n \to \infty$. Thus perturbations in the bundle degrees of freedom are unlikely to have a significant effect on the decay rate of correlations. We can thus modify Algorithm 3.4 when computing Lyapunov exponents to perturb only the state variable x, not the orthonormal frame $\hat{E}^u$. This requires only f and Df, not $D^2 f$.

ϵ	$\hat{\lambda}^{\epsilon}$	$\hat{\tau}_{\text{exp}}$	$\hat{\tau}_{\text{int}}$	$\lvert\hat{\lambda}^{\epsilon} - \hat{\lambda}^{0}\rvert$
0.0	$0.906 \pm 4.8 \times 10^{-3}$	5.334	0.013	0.000
0.5	$0.828 \pm 5.6 \times 10^{-3}$	6.208	0.017	0.077
1.0	$0.902 \pm 7.0 \times 10^{-3}$	3.398	0.028	0.003
1.5	$0.966 \pm 7.8 \times 10^{-3}$	2.009	0.037	0.060
2.0	$1.018 \pm 8.3 \times 10^{-3}$	1.341	0.045	0.112
2.5	$1.057 \pm 9.0 \times 10^{-3}$	0.968	0.056	0.151
3.0	$1.063 \pm 9.0 \times 10^{-3}$	0.734	0.059	0.157
3.5	$1.057 \pm 9.4 \times 10^{-3}$	0.594	0.067	0.151
4.0	$1.052 \pm 9.8 \times 10^{-3}$	0.487	0.076	0.146

(a) Estimates of largest exponent $\hat{\lambda}_1$

ϵ	$\hat{\lambda}^{\epsilon}$	$\hat{\tau}_{\text{exp}}$	$\hat{\tau}_{\text{int}}$	$\lvert\hat{\lambda}^{\epsilon} - \hat{\lambda}^{0}\rvert$
0.0	$0.00 \pm 3.1 \times 10^{-3}$	3.164	0.018	0.000
0.5	$0.035 \pm 5.9 \times 10^{-3}$	0.787	0.068	0.036
1.0	$-0.11 \pm 6.0 \times 10^{-3}$	0.247	0.072	0.112
1.5	$-0.28 \pm 5.8 \times 10^{-3}$	0.200	0.066	0.288
2.0	$-0.47 \pm 6.0 \times 10^{-3}$	0.193	0.065	0.470
2.5	$-0.68 \pm 6.2 \times 10^{-3}$	0.173	0.063	0.680
3.0	$-0.89 \pm 6.7 \times 10^{-3}$	0.166	0.062	0.896
3.5	$-1.10 \pm 7.2 \times 10^{-3}$	0.160	0.061	1.102
4.0	$-1.33 \pm 7.7 \times 10^{-3}$	0.160	0.061	1.339

(b) Estimates of second largest exponent $\hat{\lambda}_2$

ϵ	$\hat{\lambda}^{\epsilon}$	$\hat{\tau}_{\text{exp}}$	$\hat{\tau}_{\text{int}}$	$\lvert\hat{\lambda}^{\epsilon} - \hat{\lambda}^{0}\rvert$
0.0	$-14.5 \pm 2.6 \times 10^{-3}$	7.470	0.004	0.000
0.5	$-14.5 \pm 3.5 \times 10^{-3}$	7.210	0.008	0.041
1.0	$-14.4 \pm 4.2 \times 10^{-3}$	4.012	0.012	0.116
1.5	$-14.3 \pm 4.9 \times 10^{-3}$	2.256	0.018	0.227
2.0	$-14.2 \pm 6.0 \times 10^{-3}$	1.475	0.025	0.358
2.5	$-14.0 \pm 8.0 \times 10^{-3}$	0.988	0.044	0.529
3.0	$-13.8 \pm 9.0 \times 10^{-3}$	0.741	0.051	0.739
3.5	$-13.6 \pm 1.0 \times 10^{-2}$	0.600	0.063	0.951
4.0	$-13.3 \pm 1.1 \times 10^{-2}$	0.487	0.077	1.192

(c) Estimates of smallest exponent $\hat{\lambda}_3$

Table 3.1: Correlation times of the estimated Lyapunov exponents of the Lorenz flow periodically kicked by uniform noise of amplitude ϵ in the asymptotically unstable direction. Estimated Lyapunov exponents are stated with their standard deviations.

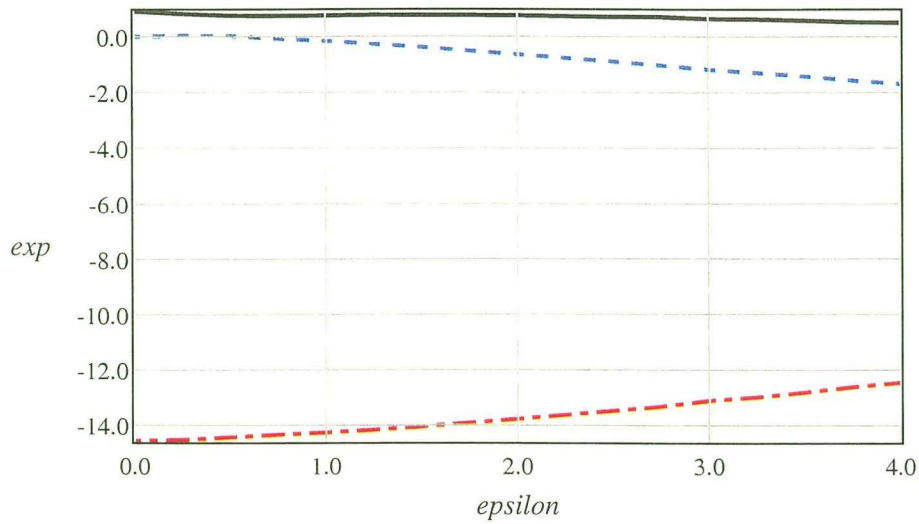


Figure 3.10: Estimated Lyapunov exponents for the Lorenz flow kicked by perturbations which are uniformly distributed over cubes of side length ϵ. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$.

3.4 General random kicks

Our attempt to construct hybrid Markov chain Monte Carlo algorithms for SRB measures has led to Algorithm 3.4. The results in the previous section show that this algorithm can be rather effective at reducing the initialization time (at least for the Lorenz system) while introducing only relatively small errors in computed averages. Algorithm 3.4 is, however, rather expensive, as it requires estimating the asymptotically unstable directions. Are there variations of Algorithm 3.4 which can be more efficiently implemented?

One natural idea is to simply use random perturbations of F with absolutely continuous transition measures. This class of perturbations has the advantage of a firmer mathematical foundation and has been advocated by a number of authors [Hun96, Kif97]. Theorem 2.16 implies the consistency of random perturbation algorithms based on absolutely continuous transition measures in the limit $\epsilon \to 0$, and it is fairly straightforward to establish the existence of a unique ergodic invariant measure when $\epsilon > 0$. Figures 3.10 - 3.14 and Table 3.2 contain the corresponding results. The data is qualitatively similar to those of the previous section.

However, as we are interested in applying these algorithms to nonlinear evo-

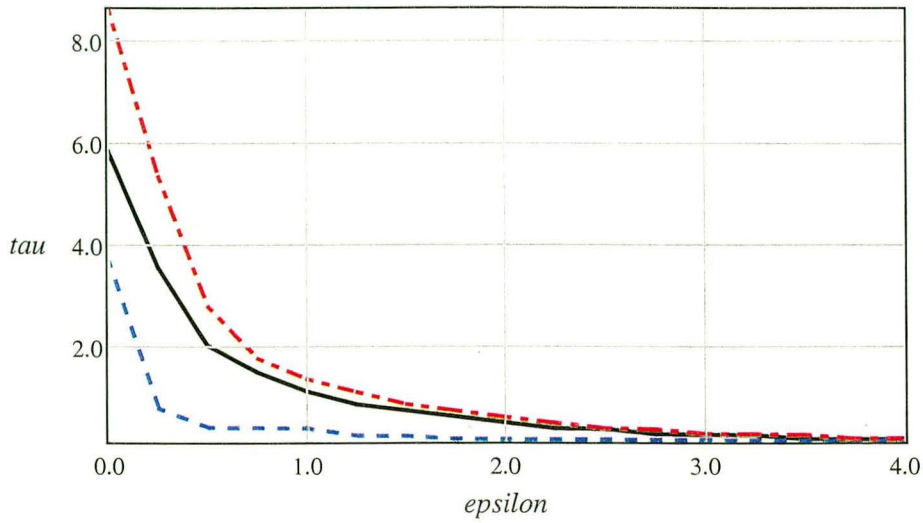


Figure 3.11: Decay times of the Lorenz flow periodically kicked by perturbations which are uniformly distributed over cubes of side length ϵ. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$.

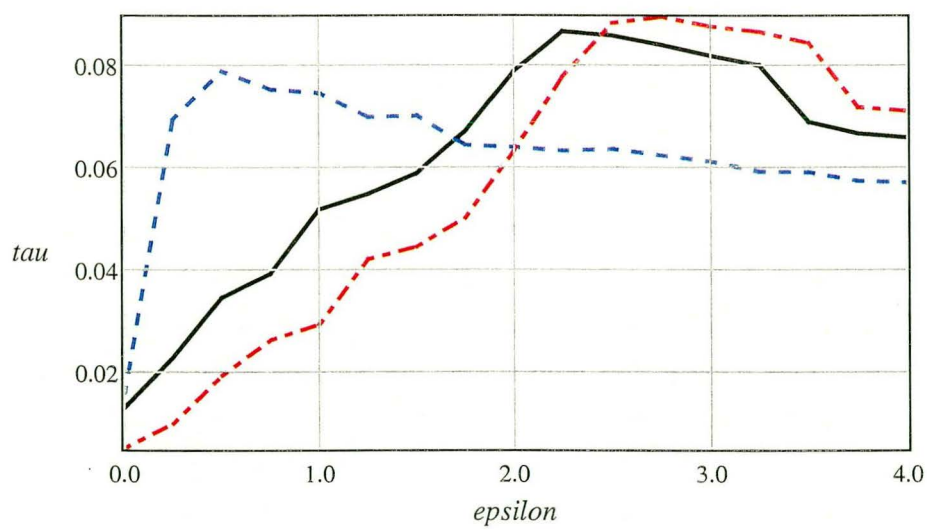


Figure 3.12: Integrated correlation times of the Lorenz flow periodically kicked by perturbations which are uniformly distributed over cubes of side length ϵ. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$.

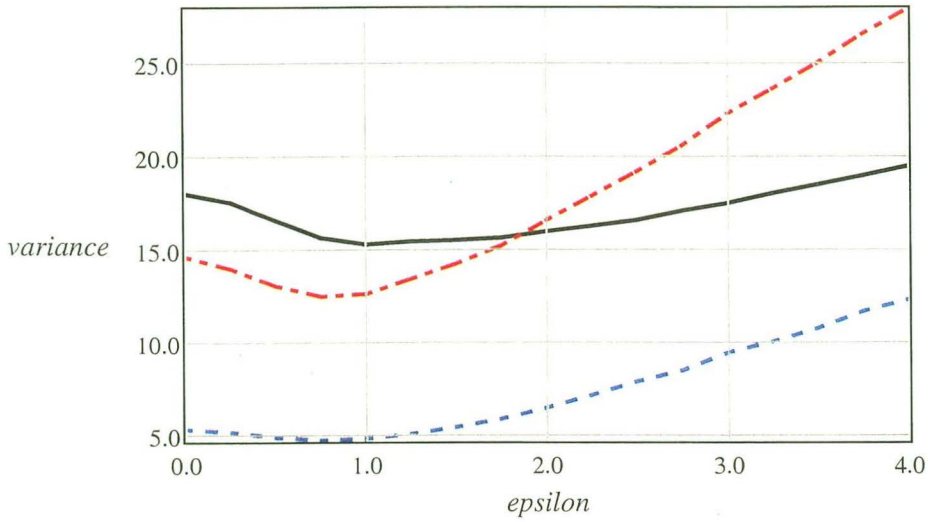


Figure 3.13: Variances of $\hat{\lambda}_k$ for the Lorenz flow periodically kicked by perturbations which are uniformly distributed over cubes of side length ϵ. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$.

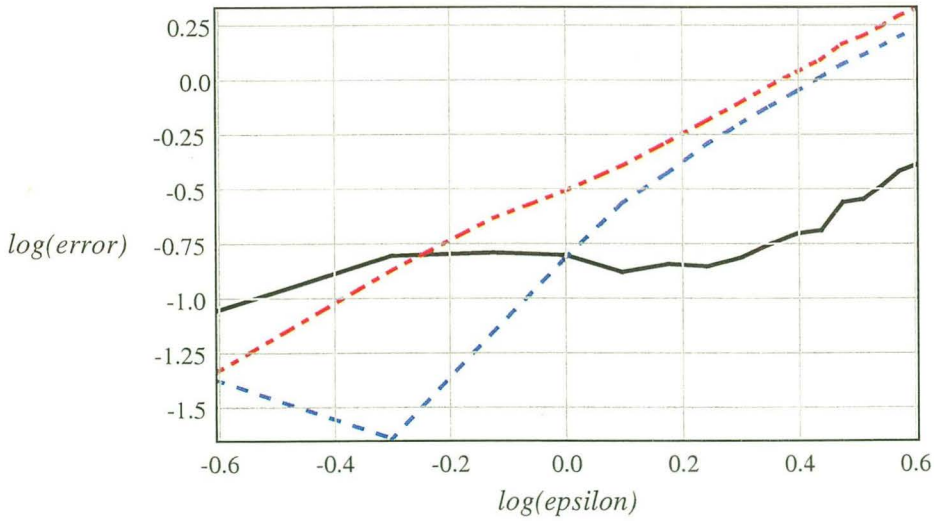


Figure 3.14: Approximate errors for the estimated Lyapunov exponents of the Lorenz flow periodically kicked by perturbations which are uniformly distributed over cubes of side length ϵ. The solid curve is associated with $\hat{\lambda}_1$, the dashed curve with $\hat{\lambda}_2$, and the dashed-dot curve with $\hat{\lambda}_3$.

ϵ	$\hat{\lambda}^\epsilon$	$\hat{\tau}_{\text{exp}}$	$\hat{\tau}_{\text{int}}$	$\lvert\hat{\lambda}^\epsilon - \hat{\lambda}^0\rvert$
0.0	$0.906 \pm 4.8 \times 10^{-3}$	5.841	0.012	0.000
0.5	$0.749 \pm 7.5 \times 10^{-3}$	2.029	0.034	0.156
1.0	$0.748 \pm 8.8 \times 10^{-3}$	1.121	0.051	0.157
1.5	$0.762 \pm 9.5 \times 10^{-3}$	0.747	0.058	0.143
2.0	$0.752 \pm 1.1 \times 10^{-2}$	0.507	0.078	0.154
2.5	$0.708 \pm 1.1 \times 10^{-2}$	0.380	0.085	0.197
3.0	$0.631 \pm 1.1 \times 10^{-2}$	0.273	0.081	0.275
3.5	$0.580 \pm 1.1 \times 10^{-2}$	0.186	0.068	0.326
4.0	$0.498 \pm 1.1 \times 10^{-2}$	0.180	0.065	0.408

(a) Estimates of largest exponent $\hat{\lambda}_1$

ϵ	$\hat{\lambda}^\epsilon$	$\hat{\tau}_{\text{exp}}$	$\hat{\tau}_{\text{int}}$	$\lvert\hat{\lambda}^\epsilon - \hat{\lambda}^0\rvert$
0.0	$0.00 \pm 2.9 \times 10^{-3}$	3.665	0.016	0.000
0.5	$0.022 \pm 6.1 \times 10^{-3}$	0.413	0.078	0.022
1.0	$-0.15 \pm 5.9 \times 10^{-3}$	0.393	0.074	0.153
1.5	$-0.37 \pm 6.1 \times 10^{-3}$	0.240	0.069	0.375
2.0	$-0.63 \pm 6.4 \times 10^{-3}$	0.180	0.063	0.635
2.5	$-0.88 \pm 7.0 \times 10^{-3}$	0.173	0.063	0.889
3.0	$-1.17 \pm 7.5 \times 10^{-3}$	0.160	0.060	1.177
3.5	$-1.42 \pm 7.9 \times 10^{-3}$	0.153	0.058	1.424
4.0	$-1.69 \pm 8.3 \times 10^{-3}$	0.146	0.056	1.697

(b) Estimates of second largest exponent $\hat{\lambda}_2$

ϵ	$\hat{\lambda}^\epsilon$	$\hat{\tau}_{\text{exp}}$	$\hat{\tau}_{\text{int}}$	$\lvert\hat{\lambda}^\epsilon - \hat{\lambda}^0\rvert$
0.0	$-14.5 \pm 2.7 \times 10^{-3}$	8.618	0.005	0.000
0.5	$-14.4 \pm 5.0 \times 10^{-3}$	2.777	0.019	0.133
1.0	$-14.2 \pm 6.0 \times 10^{-3}$	1.355	0.029	0.310
1.5	$-14.0 \pm 7.9 \times 10^{-3}$	0.861	0.044	0.518
2.0	$-13.7 \pm 1.0 \times 10^{-2}$	0.607	0.063	0.789
2.5	$-13.4 \pm 1.2 \times 10^{-2}$	0.387	0.088	1.086
3.0	$-13.1 \pm 1.3 \times 10^{-2}$	0.280	0.087	1.452
3.5	$-12.8 \pm 1.4 \times 10^{-2}$	0.260	0.084	1.750
4.0	$-12.4 \pm 1.4 \times 10^{-2}$	0.186	0.070	2.105

(c) Estimates of smallest exponent $\hat{\lambda}_3$

Table 3.2: Correlation times of the estimated Lyapunov exponents of the Lorenz flow periodically kicked by perturbations which are uniformly distributed over cubes of side length ϵ. Estimated Lyapunov exponents are stated with their standard deviations.

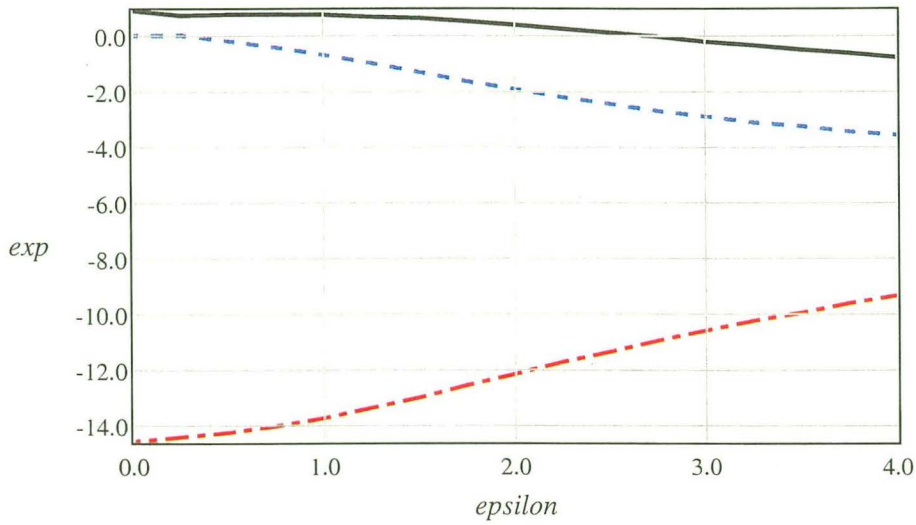


Figure 3.15: Estimated Lyapunov exponents for the Lorenz flow kicked by degenerate perturbations of amplitude ϵ in a linearly unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$.

lutionary PDEs, the use of absolutely continuous perturbations may increase the spatial roughness of the solution and reduce the efficiency of numerical integrators. In many physical examples, we are interested in differential equations of the form

$$\dot{x} = f(x) = Ax + \tilde{f}(x), \tag{3.6}$$

where $\tilde{f}$ is a smooth vector field which vanishes quadratically at $x = 0$. If the relevant dynamics is localized near the fixed point $x = 0$, then it would be dominated by the structure of the Jacobian matrix A. When A is hyperbolic, that is when it is diagonalizable and its spectrum does not intersect the imaginary axis, the stable and unstable manifolds of the fixed point 0 are tangent to the eigenspaces of A. These eigenspaces offer a convenient substitute for the asymptotically unstable subspaces in numerical studies of flows. This is especially relevant for numerical studies of nonlinear parabolic PDEs, such as reaction-diffusion equations and the Kuramoto-Sivashinsky equations, whose linear terms can often be exactly diagonalized by the usual Fourier basis.

Figures 3.15 - 3.19 and Table 3.3 contain the results for such a calculation for the Lorenz equations, where the flow is periodically perturbed by a degenerate perturbation, uniformly distributed with amplitude ϵ, in the linearly unstable direction of $Df_{\text{Lorenz}}(0)$.

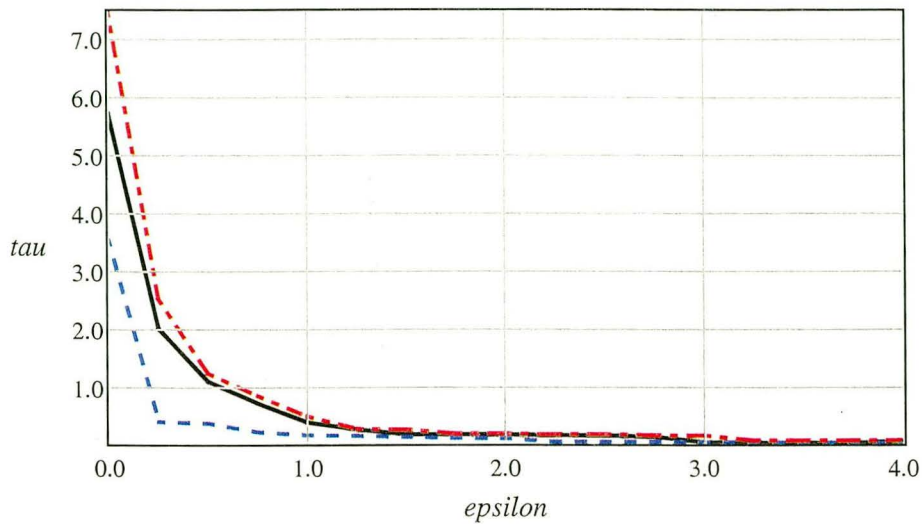


Figure 3.16: Decay times of the Lorenz flow periodically kicked by degenerate perturbations of amplitude ϵ in a linearly unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$.

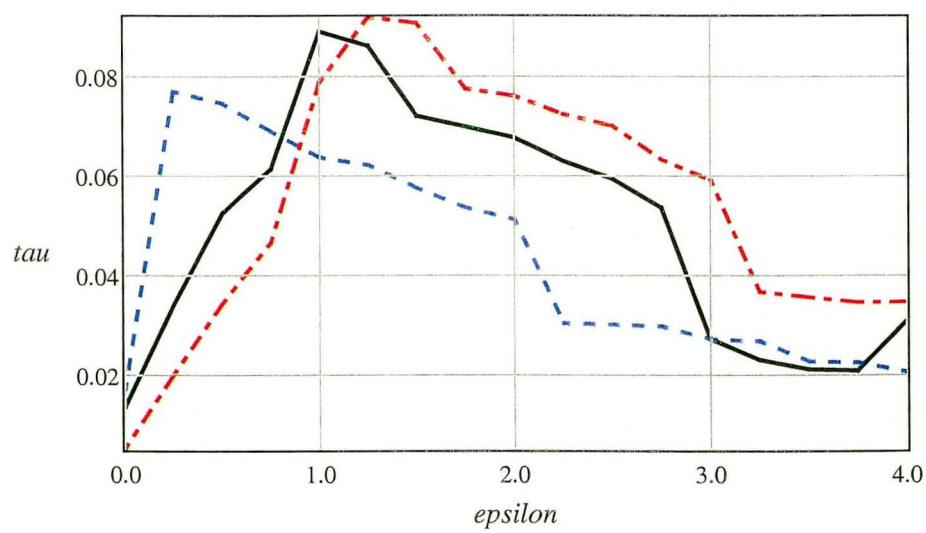


Figure 3.17: Integrated correlation times of the Lorenz flow periodically kicked by degenerate perturbations of amplitude ϵ in a linearly unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$.

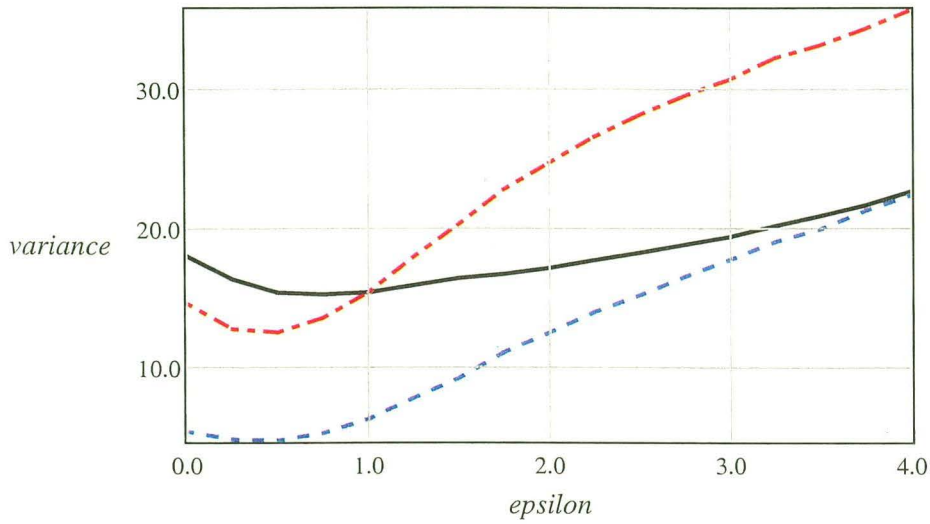


Figure 3.18: Variances of $\hat{\lambda}_k$ for the Lorenz flow periodically kicked by degenerate perturbations of amplitude ϵ in a linearly unstable direction. The solid curve is $\hat{\lambda}_1$, the dashed curve $\hat{\lambda}_2$, and the dashed-dot curve $\hat{\lambda}_3$.

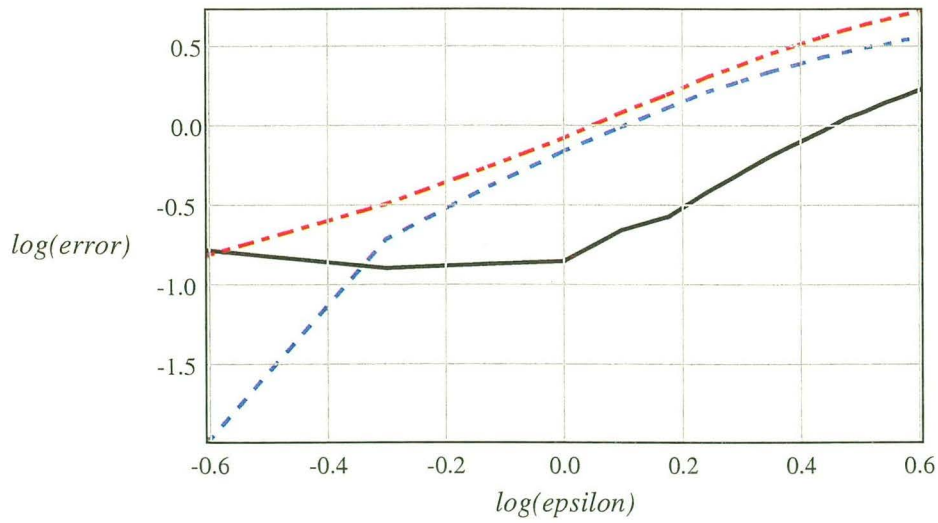


Figure 3.19: Approximate errors for the estimated Lyapunov exponents of the Lorenz flow periodically kicked by degenerate perturbations of amplitude ϵ in a linearly unstable direction. The solid curve is associated with $\hat{\lambda}_1$, the dashed curve with $\hat{\lambda}_2$, and the dashed-dot curve with $\hat{\lambda}_3$.

ϵ	$\hat{\lambda}^\epsilon$	$\hat{\tau}_{\text{exp}}$	$\hat{\tau}_{\text{int}}$	$\lvert\hat{\lambda}^\epsilon - \hat{\lambda}^0\rvert$
0.0	$0.904 \pm 4.9 \times 10^{-3}$	5.721	0.013	0.000
0.5	$0.778 \pm 8.9 \times 10^{-3}$	1.108	0.052	0.126
1.0	$0.764 \pm 1.1 \times 10^{-2}$	0.400	0.089	0.140
1.5	$0.637 \pm 1.0 \times 10^{-2}$	0.193	0.072	0.267
2.0	$0.398 \pm 1.0 \times 10^{-2}$	0.193	0.067	0.506
2.5	$0.120 \pm 1.0 \times 10^{-2}$	0.173	0.059	0.784
3.0	$-0.20 \pm 7.2 \times 10^{-3}$	0.060	0.027	1.105
3.5	$-0.48 \pm 6.6 \times 10^{-3}$	0.046	0.021	1.389
4.0	$-0.77 \pm 8.3 \times 10^{-3}$	0.080	0.030	1.680

(a) Estimates of largest exponent $\hat{\lambda}_1$

ϵ	$\hat{\lambda}^\epsilon$	$\hat{\tau}_{\text{exp}}$	$\hat{\tau}_{\text{int}}$	$\lvert\hat{\lambda}^\epsilon - \hat{\lambda}^0\rvert$
0.0	$0.00 \pm 3.0 \times 10^{-3}$	3.538	0.017	0.000
0.5	$-0.19 \pm 5.9 \times 10^{-3}$	0.387	0.074	0.192
1.0	$-0.68 \pm 6.3 \times 10^{-3}$	0.180	0.063	0.689
1.5	$-1.30 \pm 7.2 \times 10^{-3}$	0.146	0.057	1.300
2.0	$-1.91 \pm 8.0 \times 10^{-3}$	0.126	0.051	1.915
2.5	$-2.44 \pm 6.7 \times 10^{-3}$	0.066	0.030	2.441
3.0	$-2.89 \pm 6.9 \times 10^{-3}$	0.060	0.027	2.891
3.5	$-3.23 \pm 6.7 \times 10^{-3}$	0.053	0.022	3.239
4.0	$-3.56 \pm 6.7 \times 10^{-3}$	0.046	0.020	3.566

(b) Estimates of second largest exponent $\hat{\lambda}_2$

ϵ	$\hat{\lambda}^\epsilon$	$\hat{\tau}_{\text{exp}}$	$\hat{\tau}_{\text{int}}$	$\lvert\hat{\lambda}^\epsilon - \hat{\lambda}^0\rvert$
0.0	$-14.5 \pm 2.7 \times 10^{-3}$	7.483	0.005	0.000
0.5	$-14.2 \pm 6.5 \times 10^{-3}$	1.241	0.034	0.318
1.0	$-13.7 \pm 1.0 \times 10^{-2}$	0.500	0.078	0.829
1.5	$-13.0 \pm 1.3 \times 10^{-2}$	0.273	0.090	1.567
2.0	$-12.1 \pm 1.3 \times 10^{-2}$	0.200	0.076	2.421
2.5	$-11.3 \pm 1.4 \times 10^{-2}$	0.193	0.069	3.225
3.0	$-10.5 \pm 1.3 \times 10^{-2}$	0.166	0.059	3.997
3.5	$-9.94 \pm 1.0 \times 10^{-2}$	0.086	0.035	4.628
4.0	$-9.32 \pm 1.1 \times 10^{-2}$	0.093	0.034	5.247

(c) Estimates of smallest exponent $\hat{\lambda}_3$

Table 3.3: Correlation times of the estimated Lyapunov exponents of the Lorenz flow periodically kicked by degenerate perturbations of amplitude ϵ in a linearly unstable direction. Estimated Lyapunov exponents are stated with their standard deviations.

Note that the intuition behind the use of the linearly unstable directions is, strictly speaking, not applicable to the Lorenz equations, as the behavior of the Lorenz flow is not clearly dominated by its behavior near the unstable fixed point at the origin. Nevertheless perturbations in the linearly unstable direction at the origin appear to be as effective as perturbations in asymptotically unstable directions, possibly because all that is really needed is random perturbations which are transverse to the stable directions with high probability.

3.5 An intermittent interval map

The Lorenz flow is rather special in some ways: although it is nonuniformly hyperbolic and possesses a complicated strange attractor, it exhibits very robust mixing properties and strong dissipation. In this section we examine a simple family of maps which have properties that are expected to cause more serious difficulties.

Liverani, Saussol, and Vaienti [LSV99] have studied the family of maps

$$G_\alpha(x) = \begin{cases} x[1+(2x)^\alpha], & 0 \le x < 1/2, \\ 2x-1, & 1/2 \le x \le 1 \end{cases} \tag{3.7}$$

When $\alpha = 0$, G_α is the "angle-doubling" map $x \mapsto 2x \pmod 1$ and is uniformly expanding (i.e. uniformly hyperbolic with $\dim E^u = 1$ and $\dim E^s = 0$). When $\alpha > 0$, something more complicated happens: Figure 3.20 shows a plot of the map G_α for $\alpha = 0.999$ superimposed with a plot of the identity function $x \mapsto x$. The map is tangent to the diagonal at the origin and fails to be expanding there. Fix x_0 and a so that $0 < x_0 < a < \frac{1}{2}$ and let n be the smallest positive integer such that $G_\alpha^n(x_0) > a$. Linearizing yields

$$\frac{\ln a - \ln x_0}{\ln\left(1+(1+\alpha)(2a)^\alpha\right)} \lesssim n \lesssim \frac{\ln a - \ln x_0}{\ln\left(1+(1+\alpha)(2x_0)^\alpha\right)}.$$

A straightforward numerical calculation suggests that, in fact, $n \sim O\left(x_0^{-\alpha} \cdot (-\ln x_0)\right)$, so that the amount of time required for the system to escape from a small neighborhood of the origin is much larger when $\alpha > 0$ than when $\alpha = 0$.

Liverani, Saussol, and Vaienti proved that for $0 < \alpha < 1$, the map G_α has an absolutely continuous invariant probability measure. Autocorrelation functions will not decay exponentially fast, however; they can only decay at the rate $O(n^{-1/\alpha+1})$. When $\alpha < 1/2$, the covariance functions of G_α are absolutely summable and one can prove a central limit theorem. This map provides a simple example of intermittency coexisting

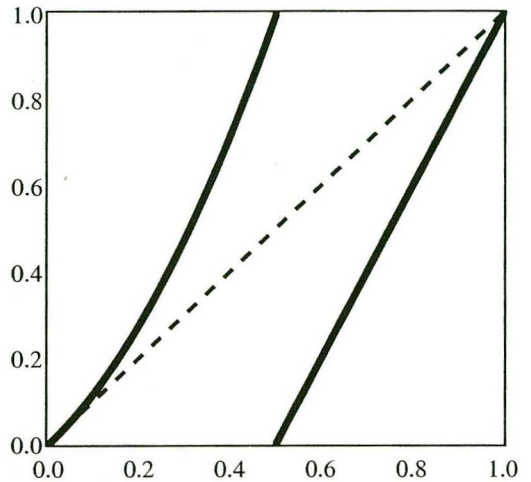


Figure 3.20: The intermittent map G_α for $\alpha = 0.999$.

with an absolutely continuous invariant measure. Note that the main idea of the proofs in [LSV99] is to use random perturbations of G_α to establish results on mixing rate, etc., in the zero noise limit.

In Figures 3.21 - 3.25, numerical estimates of the Lyapunov exponent of G_α is shown for $\alpha = 0.49$. As can be seen, the addition of noise significantly decreases the correlation time, as one might expect. Note that while $\tau_{\text{exp}} = \infty$ when $\epsilon = 0$, $\hat{\tau}_{\text{exp}}(\epsilon = 0)$ is finite because of its definition. It is also not surprising that in this case, τ_{exp} and τ_{int} have similar behavior as the noise amplitude ϵ is varied: the correlation functions of G_α, unlike those of the Lorenz flow, are simple decaying functions and do not appear to oscillate. This corresponds to a pure imaginary correlation spectrum $\sigma_{\text{corr}}(G_\alpha)$.

More dramatically, if we set $x_0 = 10^{-12}$ and $\alpha = 0.99$, x_n is only on the order of 10^{-9} after 10^6 iterations. These calculations suggest that the addition of a small amount of noise can help avoid some problems associated with intermittency in numerical simulations of low dimensional maps, at the same time perturbing the desired averages by only a small amount. A word of caution: this conclusion may not hold in more complex systems in higher dimensions, as realistic physical models are likely to have more complicated mechanisms for intermittent behavior.

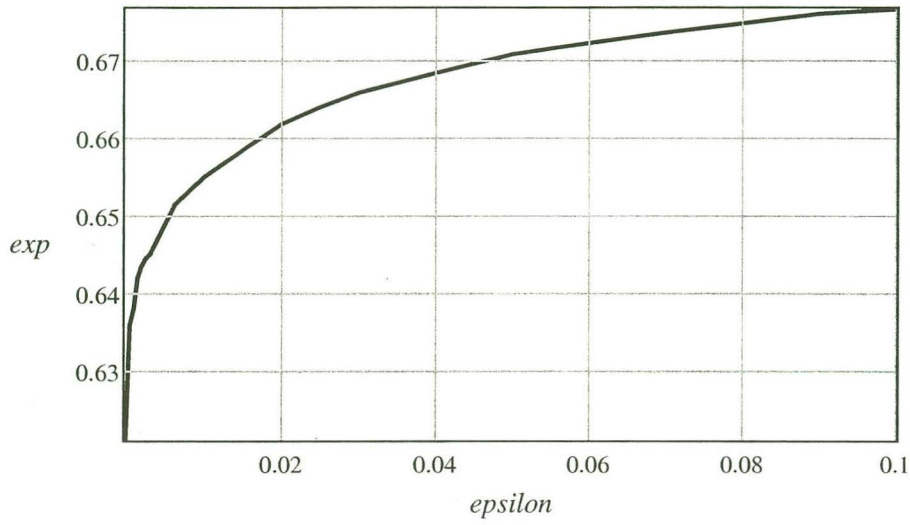


Figure 3.21: Estimated Lyapunov exponents for the intermittent interval map with $\alpha = 0.49$ kicked by uniform noise of amplitude ϵ.

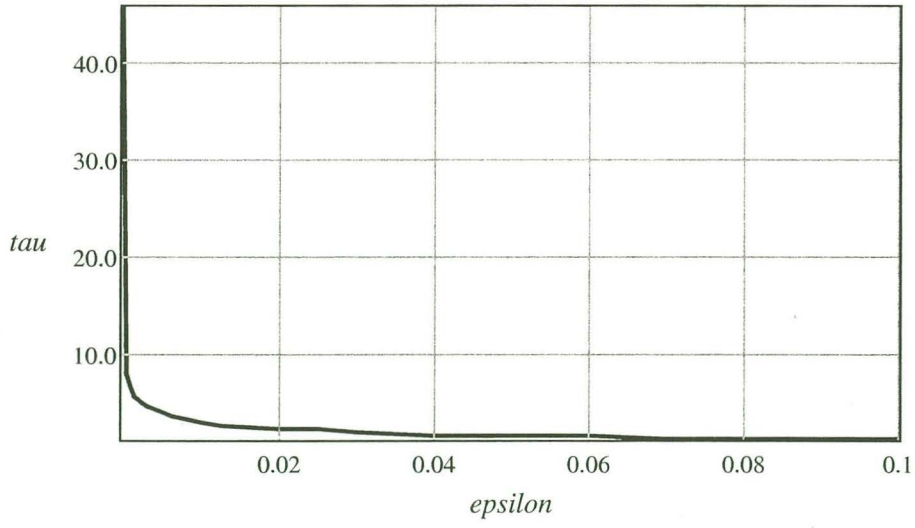


Figure 3.22: Decay times of the intermittent interval map with $\alpha = 0.49$ periodically kicked by uniform noise of amplitude ϵ.

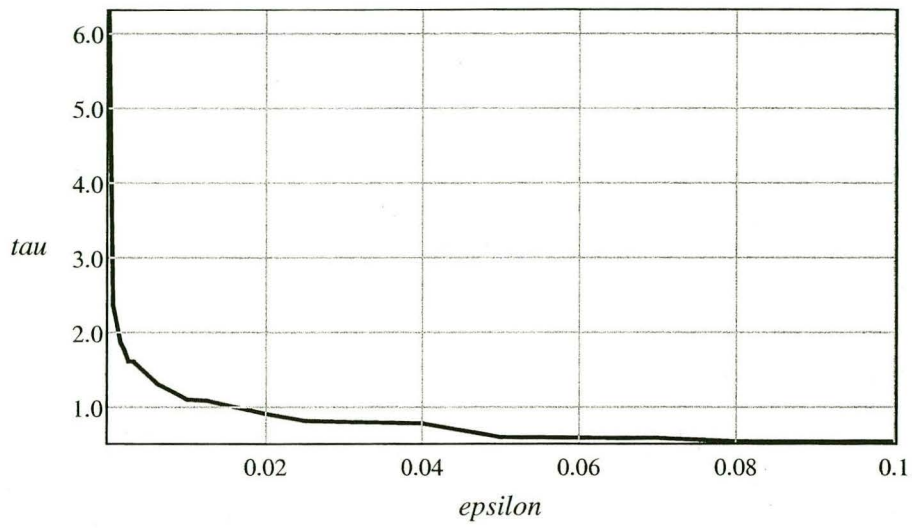


Figure 3.23: Integrated correlation times of the intermittent interval map with $\alpha = 0.49$ periodically kicked by uniform noise of amplitude ϵ.

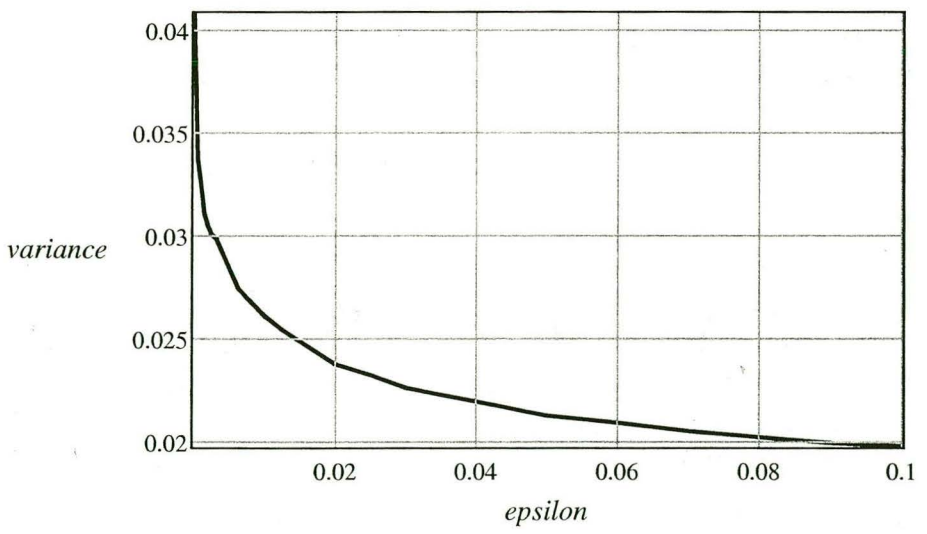


Figure 3.24: Variances of $\hat{\lambda}_k$ for the intermittent interval map with $\alpha = 0.49$ periodically kicked by degenerate perturbations of amplitude ϵ.

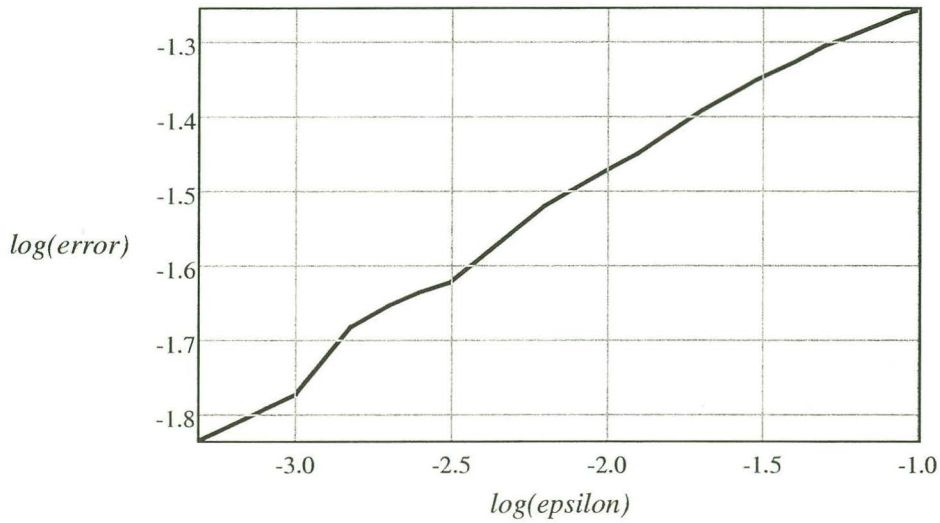


Figure 3.25: Approximate errors for the estimated Lyapunov exponents of the intermittent interval map with $\alpha = 0.49$ periodically kicked by uniform noise of amplitude ϵ.

ϵ	$\hat{\lambda}^{\epsilon}$	$\hat{\tau}_{\text{exp}}$	$\hat{\tau}_{\text{int}}$	$\lvert\hat{\lambda}^{\epsilon} - \hat{\lambda}^{0}\rvert$
0.0	$0.621 \pm 7.1 \times 10^{-4}$	45.73	6.298	0.000
0.01	$0.655 \pm 2.4 \times 10^{-4}$	3.004	1.105	0.033
0.02	$0.661 \pm 2.0 \times 10^{-4}$	2.336	0.910	0.040
0.03	$0.665 \pm 1.9 \times 10^{-4}$	2.002	0.806	0.044
0.04	$0.668 \pm 1.8 \times 10^{-4}$	1.669	0.782	0.047
0.05	$0.670 \pm 1.5 \times 10^{-4}$	1.669	0.598	0.049
0.06	$0.672 \pm 1.5 \times 10^{-4}$	1.669	0.592	0.050
0.07	$0.673 \pm 1.5 \times 10^{-4}$	1.335	0.586	0.052
0.08	$0.674 \pm 1.4 \times 10^{-4}$	1.335	0.544	0.053
0.09	$0.676 \pm 1.4 \times 10^{-4}$	1.335	0.541	0.054
0.1	$0.676 \pm 1.4 \times 10^{-4}$	1.335	0.538	0.055

Table 3.4: Correlation times of the estimated Lyapunov exponents of the intermittent interval map with $\alpha = 0.49$ periodically kicked by uniform noise of amplitude ϵ. Estimated Lyapunov exponents are stated with their standard deviations.

Figure 3.25 suggests that $\mathbb{E}\left|\hat{\lambda}(\epsilon) - \lambda\right| = O(\epsilon^{\theta})$ for some $\theta > 0$ as $\epsilon \to 0$. The analysis in Chapter 4 may offer an explanation for this behavior.

3.6 Scaling and white noise

When Algorithm 3.4 is applied to flows, there are two parameters in addition to the choice of the form of the random perturbation: the noise amplitude ϵ and the perturbation period T. Let us again compute the Lyapunov exponents $\hat{\lambda}_i$ for the Lorenz flow, this time periodically kicking the flow with period T using isotropic gaussian perturbations of variance ϵ. Figure 3.26 shows the estimated exponents $\hat{\lambda}_i$ as functions of the noise amplitude ϵ, for various values of the perturbation period T. As one can see, there is no apparent relation between the curves corresponding to different values of T. However, if we rescale the curves to plot $\hat{\lambda}_i$ against $\ln\sqrt{\epsilon/T}$, then we obtain the plots on the right column of Figure 3.26: the curves collapse onto a single curve. Figure 3.27 confirms a similar scaling for the exponential decay times.

Thus the estimated Lyapunov exponent $\hat{\lambda}_i$ is a function of $\ln\sqrt{\epsilon/T}$. This scaling behavior is not really surprising, as the Markov chain generated by adding periodic gaussian perturbations to the Lorenz flow, in the limit when $\epsilon \to 0$, $T \to 0$, and $\epsilon/T \to \epsilon_1 \in (0, +\infty)$ converges to the solution of the stochastic differential equation

$$dx = f_{\text{Lorenz}}(x)dt + \epsilon_1 dw \tag{3.8}$$

where w denotes standard Wiener process [KP92]. A similar scaling behavior is expected when the periodic gaussian perturbations are directed in asymptotically unstable directions, though only when $\ln\sqrt{\epsilon/T}$ is very small. An immediate implication of the existence of this limit is that when we apply period-T random perturbations of *variance* ϵ to hyperbolic flows, we must have $\epsilon/T \to 0$ to guarantee the weak convergence of invariant measures as $\epsilon \to 0$.

Computationally, it is much easier to apply Algorithm 3.4 and its variants to the Lorenz flow than to solve the stochastic differential equation (3.8). This is because SDE solvers are generally more complicated than their deterministic cousins [KP92]. There is, however, a theoretical advantage in relating periodic random perturbations to white noise: there exists a substantial machinery for manipulating and analyzing stochastic differential

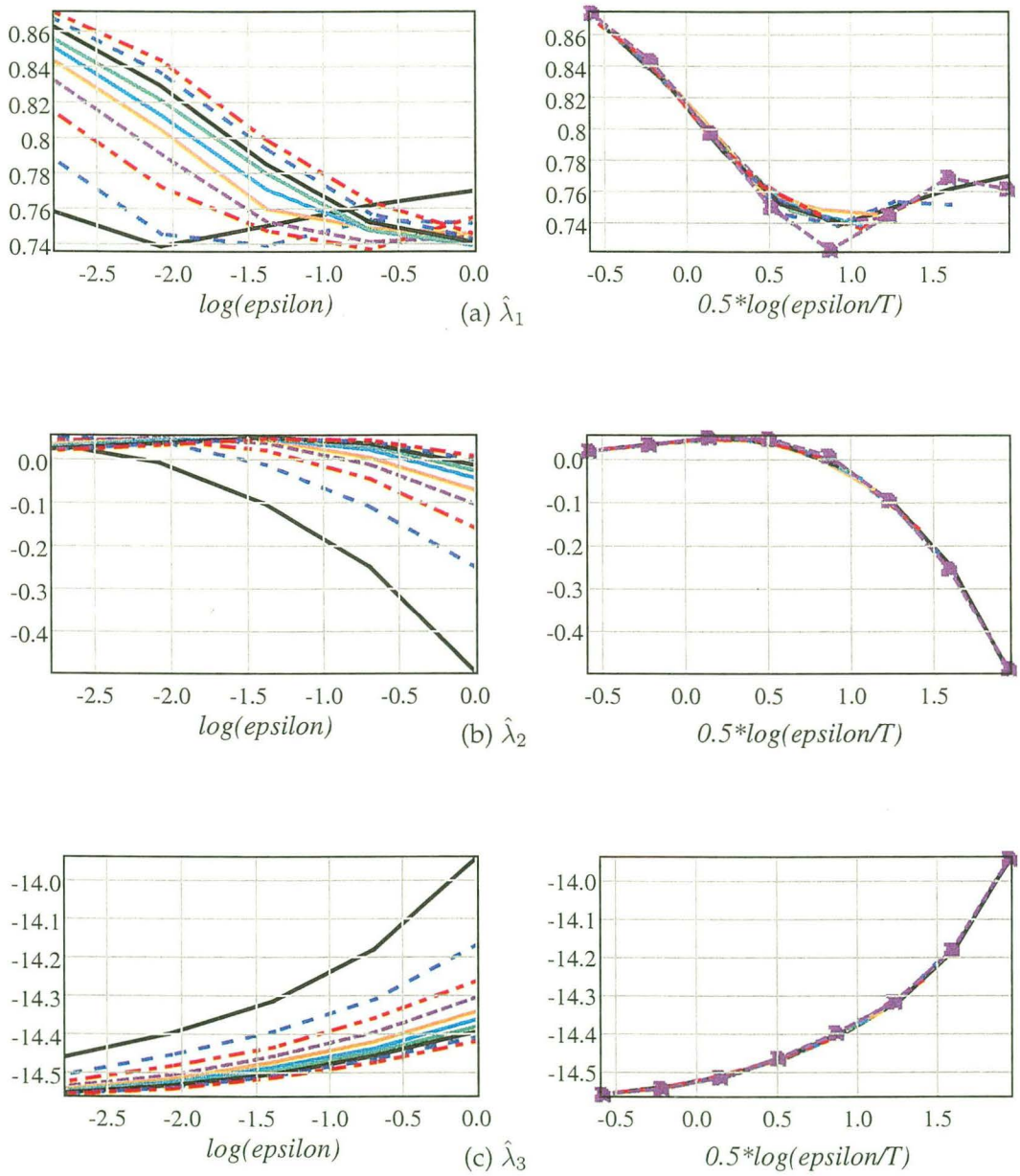


Figure 3.26: Estimates of Lyapunov exponents as functions of ϵ, for various values of T. In the second column, the data points marked by squares come from solving the SDE (3.8).

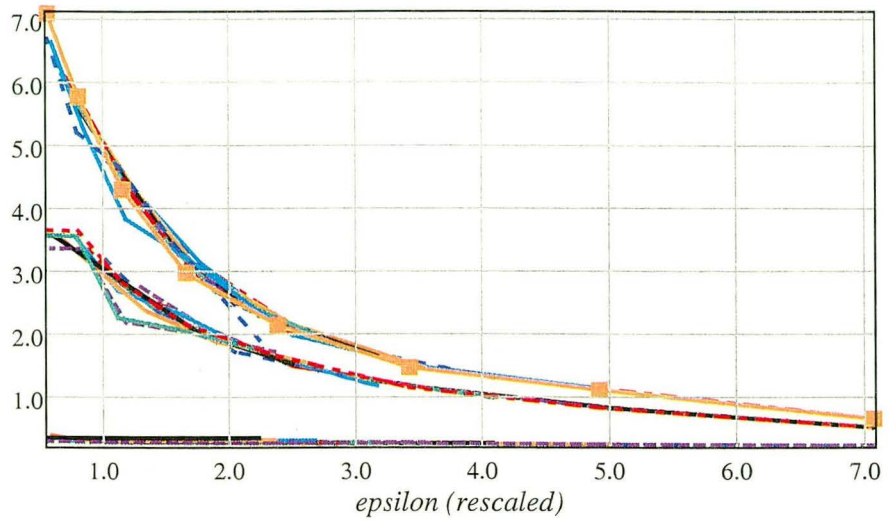


Figure 3.27: Exponential decay times for the autocovariance functions of $\hat{\lambda}_k$. Again, the data points marked by squares are computed from the SDE (3.8).

equations, ranging from the stochastic calculus [RW00b] to a well-developed large deviations theory [FW98] and the methods of modern PDE theory [Eva98]. We do not present these connections here as they are part of on-going work.

Chapter 4

Convergence and Related Problems

Chapter 3 demonstrated the use of random perturbations in numerical computations of statistical averages. The main effect of random perturbations is a reduction in the amount of initialization time required for long time simulations. In a low dimensional family of maps with a simple intermittent behavior, random perturbations also help nudge the system out of the non-expanding region into domains where phase space expansion dominates.

In order to provide a solid theoretical foundation for the use of random perturbations in numerical studies of SRB measures, the following questions must be answered:

1. Let (x_n^ϵ) denote the trajectory of the perturbed system and $g : X \to \mathbb{R}$ an observable. Does $\mathbb{E}\left[g\left(x_n^\epsilon\right)\right]$ converge to $\int g \, d\mu_{\text{SRB}}$ as $n \to \infty$ and $\epsilon \to 0$?

2. What is the effect of random perturbations on the decay rate of correlation functions?

As was pointed out in previous chapters, Kifer's Theorem (see 2.6) provides an affirmative answer to the first question but only when the perturbations have absolutely continuous transition measures. Furthermore, Kifer's result provides no information on the rate of convergence and does not apply when the transitions have singular transition measures.

This chapter provides a simple, alternate proof of a special case of Kifer's Theorem along with a rough error estimate on the convergence rate. The main result is an upper bound in the form of a power law: there exists a constant $\gamma \in (0, 1)$ such that

$$\left| \int g \, d\mu_\epsilon - \int g \, d\mu_{\text{SRB}} \right| = O\left(\epsilon^\gamma\right), \tag{4.1}$$

where μ_ϵ is an invariant measure of the perturbed system and μ_{SRB} is the SRB measure of the unperturbed system. For technical reasons, the proof only works for absolutely continuous transition measures of convolution type, but the result is likely to be true for the more complicated transition measures considered in Chapter 3.

The estimate (4.1) is not likely to be sharp: even in the case of a nonuniformly hyperbolic system like the Lorenz flow, perturbed invariant measures appear to converge with exponent $\gamma \geq 1$ as $\epsilon \to 0$ (see Figure 3.1). The harder problem of determining the effect of random perturbations on the correlation spectrum of hyperbolic systems is left for future work.

4.1 Rate of mixing in Anosov systems

First, we need to review some standard definitions and results. See Bressaud and Liverani [BL02] for details and proofs.

Let X be a compact Riemannian manifold and $F : X \to X$ a topologically transitive Anosov diffeomorphism (see §2.2). For x and x' in X, let $d^s(x, x')$ be the distance between x and x' measured along the stable manifold containing both x and x' (we set $d^s(x, x') = +\infty$ if x and x' do not lie on the same stable manifold). Fix $\delta > 0$ and $\alpha, \beta_s \in (0, 1)$. For $g : X \to \mathbb{R}$, define

$$|g|_s = \sup_{d^s(x,x') \leq \delta} \frac{|g(x) - g(x')|}{d^s(x, x')^{\beta_s}}, \tag{4.2}$$

$$||g||_s = ||g||_\infty + |g|_s, \tag{4.3}$$

$$||g||_u = ||g||_1 + \sup_{d^u(x,x') \leq \delta} \frac{|g(x) - g(x')|}{d^u(x, x')^{\alpha}}. \tag{4.4}$$

(The metric d^u is the exact analog of d^s on unstable manifolds.) Let $C_s(\beta_s)$ denote the set of all measurable functions g such that $||g||_s < \infty$. The following is the main result of [BL02]:

Theorem 4.1. *There exist positive constants C and $\theta < 1$ such that for all $\beta_s \in (0, 1]$, $g \in C_s(\beta_s)$, and $h_1, h_2 \in C^\alpha(X, \mathbb{R})$ with $\int h_1 \, dm = \int h_2 \, dm = 1$,*

$$\left| \int (g \circ F^n) \cdot h_1 \, dm - \int (g \circ F^n) \cdot h_2 \, dm \right| \leq C ||g||_s \max\left(||h_1||_u, ||h_2||_u\right) \cdot \theta^{n\beta_s}. \tag{4.5}$$

This theorem tells us how quickly an absolutely continuous initial distribution relaxes to a statistical steady state:

Corollary 4.2. *The map F possesses a unique SRB measure μ_{SRB}. Furthermore, there exist positive constants C and $\theta_1 < 1$ such that for all $g \in C_s(\beta_s)$ and $h \in C^\alpha(X, \mathbb{R})$,*

$$\left| \int (g \circ F^n)\, h\, dm - \int g\, d\mu_{SRB} \int h\, dm \right| \leq C ||g||_s \cdot ||h||_u \cdot \theta_1^n. \tag{4.6}$$

If we take $h \geq 0$, $\int h\, dm = 1$ in Corollary 4.2, then (4.6) tells us that

$$\text{dist}\,(F_*^n(h\, dm), d\mu_{\text{SRB}}) \leq C||h||_u \theta_1^n$$

where

$$\text{dist}\,(d\mu_1, d\mu_2) = \sup_{||g||_s \leq 1} \left| \int g\, d\mu_1 - \int g\, d\mu_2 \right|.$$

Thus the exponential decay of correlation functions implies exponential relaxation to equilibrium, as claimed above.

4.2 Error estimates

This section contains the proof of (4.1). The basic ingredients are a scaling argument due to Shardlow and Stuart [SS00] and the results stated in the previous section. For simplicity, the discussion is restricted to the simplest case of Anosov diffeomorphisms on the d-torus $\mathbb{T}^d$, though it should not be difficult to extend the argument to Axiom A systems on general manifolds.

Proposition 4.3. *Let $F : \mathbb{T}^d \to \mathbb{T}^d$ be a transitive Anosov diffeomorphism and let μ_{SRB} be the SRB measure of F. Suppose $p_\epsilon(x, \cdot)$ is a family of probability measures satisfying the following conditions:*

1. *The measures $p_\epsilon(x, dx')$ have the form $q_\epsilon(x - x')\, dx'$, that is they are translation-invariant.*
2. *The distributions q_ϵ have standard deviation ϵ.*

Let F_ϵ denote the corresponding random perturbation of F, and let m_0 be an absolutely continuous measure. Set $x_{n+1}^\epsilon = F_\epsilon(x_n^\epsilon)$ with initial condition x_0 drawn from m_0. Then there exist constants $C = C(m_0)$ (depending only on m_0 and F), $\beta \geq \log ||DF||_\infty$, and $r \geq |\sigma(T_F) \setminus \{1\}|$ such that for all Lipschitz-continuous functions $g : X \to \mathbb{R}$, the estimate

$$\left| \mathbb{E}\left[g(x_n^\epsilon)\right] - \int g\, d\mu_{SRB} \right| \leq C\left[Lip(g) + ||g||_s\right] \epsilon^\gamma \tag{4.7}$$

holds with exponent

$$\gamma = \frac{\log r}{\log r - \beta} \tag{4.8}$$

whenever $n \geq n_0 = \frac{\log\epsilon + \beta}{\log r - \beta}$.

Proof. First, define the following Markov chains with initial condition x_0 drawn from m_0:

$$\begin{aligned} x_{n+1} &= F(x_n), \\ x^\epsilon_{n+1} &= F_\epsilon\left(x^\epsilon_n\right). \end{aligned} \tag{4.9}$$

The quantity we wish to estimate, $\left|\mathbb{E}\left[g(x^\epsilon_n)\right] - \int g\, d\mu_{\text{SRB}}\right|$, can be split up in the following way:

$$\left|E\left[g(x^\epsilon_n)\right] - \int g\, d\mu_{\text{SRB}}\right| \leq \left|E\left[g(x^\epsilon_n)\right] - E\left[g(x_n)\right]\right| + \tag{4.10}$$

$$\left|E\left[g(x_n)\right] - \int g\, d\mu_{\text{SRB}}\right|. \tag{4.11}$$

The rest of the proof consists of showing that as $n \to +\infty$,

1. the right hand side of (4.10) grows at most exponentially fast; and
2. the term (4.11) decays at least exponentially fast.

To estimate (4.10), observe that $\left|\mathbb{E}\left[g(x^\epsilon_n)\right] - \mathbb{E}\left[g(x_n)\right]\right| \leq \text{Lip}(g)\mathbb{E}\left[d\left(x^\epsilon_n, x_n\right)\right]$. Choosing β sufficiently large (on the order of $\log ||DF(x)||_\infty$) yields $\mathbb{E}\left[d\left(x^\epsilon_n, x_n\right)\right] \leq C\epsilon \cdot e^{\beta n}$, which implies that

$$\left|\mathbb{E}\left[g(x^\epsilon_n) - \mathbb{E}\left[g(x_n)\right]\right]\right| \leq C \cdot \text{Lip}(g) \cdot \epsilon \cdot e^{\beta n}. \tag{4.12}$$

To estimate (4.11), we can use Corollary 4.2: there exist positive constants C and $r < 1$ such that $C_{gh}(n) \leq C||g||_s \cdot ||h||_u \cdot r^{|n|}$. As the initial measure m_0 is absolutely continuous, it has a density h_0 with respect to the Haar measure m. This means that

$$\begin{aligned} \left|\mathbb{E}\left[g(x_n)\right] - \int g\, d\mu_{\text{SRB}}\right| &= \left|\int g \circ F^n\, dm_0 - \int g\, d\mu_{\text{SRB}}\right| \\ &= \left|\int (g \circ F^n) h_0\, dm - \int g\, d\mu_{\text{SRB}}\right| \\ &\leq C||g||_s \cdot ||h_0||_u \cdot r^n. \end{aligned}$$

It follows that

$$\left|\mathbb{E}\left[g(x^\epsilon_n)\right] - \int g\, d\mu_{\text{SRB}}\right| \leq C\left(\text{Lip}(g) + ||g||_s \cdot ||h_0||_u\right) \cdot \left(\epsilon \cdot e^{\beta n} + r^n\right). \tag{4.13}$$

We need to choose n_0 and γ so that

$$\epsilon \cdot e^{\beta n} \leq \epsilon^{\gamma} \tag{4.14}$$

for all $n \leq n_0 + 1$ and

$$r^n \leq \epsilon^{\gamma} \tag{4.15}$$

for all $n \geq n_0$. It is easy to check that the choice

$$n_0 = \frac{\log \epsilon + \beta}{\log r - \beta} \tag{4.16}$$

$$\gamma = \frac{\log r}{\log r - \beta} \tag{4.17}$$

works. So

$$\left| \mathbb{E}\left[g(x_n^{\epsilon})\right] - \int g \, d\mu_{\text{SRB}} \right| \leq C \left(\text{Lip}(g) + ||g||_s \cdot ||h_0||_u\right) \cdot \epsilon^{\gamma} \tag{4.18}$$

if $n_0 \leq n \leq n_0 + 1$. As n_0 and γ do not depend on m_0, the argument can be repeated with m_0 replaced by the law m_k of x_k^{ϵ} for $k \geq 1$:

$$\left| \mathbb{E}\left[g(x_{n+k}^{\epsilon})\right] - \int g \, d\mu_{\text{SRB}} \right| \leq C \left(\text{Lip}(g) + ||g||_s \cdot ||h_k||_u\right) \cdot \epsilon^{\gamma} \tag{4.19}$$

for $n_0 + k \leq n \leq n_0 + k + 1$. Now

$$\begin{aligned} h_k(x) &= \int q_{\epsilon}(x - F(x'))h_{k-1}(x') \, m\left(dx'\right) \\ &= \int q_{\epsilon}(x - x')h_{k-1}(F^{-1}(x')) \det\left(DF^{-1}(x')\right) \, m\left(dx'\right) \\ &= \int q_{\epsilon}(x')h_{k-1}(F^{-1}(x - x')) \det\left(DF^{-1}(x - x')\right) \, m\left(dx'\right). \end{aligned}$$

The map F behaves like an expanding map along unstable manifolds. Differentiating in the direction $E^u(x)$ shows that there exist constants $a \in (0,1)$ and $b > 0$ such that

$$||h_{k+1}||_u \leq a||h_k||_u + b. \tag{4.20}$$

It follows that

$$\sup_{k \geq 1} ||h_k||_u < +\infty \tag{4.21}$$

and the factor $C||h_k||_u$ in Equation (4.19) can be replaced by a constant depending only on h_0 (or equivalently m_0) and the map F. This yields

$$\left| \mathbb{E}\left[g(x_n^{\epsilon})\right] - \int g \, d\mu_{\text{SRB}} \right| \leq C \left(\text{Lip}(g) + ||g||_s\right) \epsilon^{\gamma} \tag{4.22}$$

whenever $n \geq n_0 = \frac{\log \epsilon + \beta}{\log r - \beta}$. □

Corollary 4.4. *Suppose the randomly perturbed map F_ϵ has a unique invariant measure μ_ϵ. Then under the assumptions of Proposition 4.3,*

$$\sup_{g \in Lip} \left| \int g \, d\mu_\epsilon - \int g \, d\mu_{SRB} \right| = O\left(\epsilon^\gamma\right). \tag{4.23}$$

As mentioned above, the argument should work in the more general setting of Axiom A attractors. It should also work when X is a bounded subset of $\mathbb{R}^d$, although in the Euclidean case one must also take care to ensure that the probability of escape from the basin of attraction is sufficiently small. When the transition measures $p_\epsilon(x, \cdot)$ are not of convolution type, more effort is required to make the argument in the proof of Proposition 4.3 work. Specifically, more care is required in checking that $\sup_k ||h_k||_u < \infty$. Nevertheless the estimate is likely to be true. Also, as pointed out earlier, (4.1) is not likely to be sharp. See Figure 3.1.

Because $r < 1$ and $\beta > 0$, the exponent $\gamma = \frac{\log r}{\log r - \beta}$ generally falls between 0 and 1. Thus the $O(\epsilon^\gamma)$ error can, in principle, be rather large. This rather pessimistic prediction is not reflected in numerical experiments: errors generally decay much more quickly as $\epsilon \to 0$. It is an open problem to develop a perturbation theory which can adequately explain the results of numerical experiments on hyperbolic systems, perhaps along the lines of [CCP97]. It would also be very nice to have some idea of how to handle situations where hyperbolicity is not uniform, especially in systems exhibiting some intermittency. This is left for future work as well.

4.3 Unique ergodicity

Proposition 4.3 says nothing about the unique ergodicity of the Markov chain x_n^ϵ for $\epsilon > 0$. It does say that even if x_n^ϵ possesses several invariant measures, none of them can be more than $O(\epsilon^\gamma)$ away from μ_{SRB} when ϵ is sufficiently small. In some cases, it is possible to establish unique ergodicity directly. One case is when the transition measures $p_\epsilon(x, \cdot)$ is absolutely continuous. Another case is described below:

Theorem 4.5. *Let $F : X \to X$ be a diffeomorphism possessing an Axiom A attractor K. Suppose $F \restriction_K$ is topologically transitive and let μ_{SRB} be the SRB measure of $F \restriction_K$. Suppose $p_\epsilon(x, \cdot)$ is a family of probability measures satisfying the following conditions:*

1. *There exists a C^1 family $Y_\epsilon(x)$ of smooth submanifolds of X of diameter $\leq \epsilon$ such that $p_\epsilon(x,\cdot)$ is absolutely continuous with respect to the Riemann measure of $Y_\epsilon(x)$. The dimension of $Y_\epsilon(x)$ is assumed to be $\dim W^u$. Let us denote the density of $p_\epsilon(x,\cdot)$ on $Y_\epsilon(x)$ by $q_\epsilon(x,\cdot)$.*

2. *The family of manifolds $Y_\epsilon(x)$ are everywhere transverse to the local stable manifolds of F with an uniformly bounded angle.*

3. *The transition densities $q_\epsilon(x,x')$ are uniformly positive and are smooth functions of x and $x' \in Y_\epsilon(x)$ with uniformly bounded derivatives.*

Then the random map F_ϵ has a unique invariant measure for all sufficiently small $\epsilon > 0$.

Proof. The proof relies a coupling argument. The main idea is due to Mattingly [Mat02] and is quite close to the work of Bressaud and Liverani [BL02].

Let us choose two probability measures m_1 and m_2 and fix $\epsilon > 0$. The proof relies on constructing a Markov chain (x_n^1, x_n^2) such that x_n^1 and x_n^2 are two realizations of the Markov chain with the transition measure $p_\epsilon(F(x),\cdot)$, that is at each step we apply the deterministic map F and then perturb the result. The initial conditions x_0^i are drawn from m_i for $i = 1,2$, and the joint process (x_n^1, x_n^2) will have the property that $d\left(x_n^1, x_n^2\right)$ almost surely converges to 0 as $n \to \infty$.

If we can construct a joint process as described above, then we can adapt the basic coupling inequality (see §2.5.1) to this setting. More precisely, fix $h > 0$ and let n_h be the smallest integer such that for all $n \geq n_h$, $d\left(x_n^1, x_n^2\right) \leq h$. Then

$$\begin{aligned}
\left|\mathbb{E}\left[g(x_n^1) - g(x_n^2)\right]\right| &\leq \left|\mathbb{E}\left[g(x_n^1) - g(x_n^2) | n \leq n_h\right]\right| P(n \leq n_h) + \\
&\quad \left|\mathbb{E}\left[g(x_n^1) - g(x_n^2) | n > n_h\right]\right| P(n > n_h) \\
&\leq 2 \cdot ||g||_\infty \cdot P(n \leq n_h) + h.
\end{aligned}$$

As $n \to \infty$, the first term vanishes. Since h was arbitrary, this shows that any two realizations of the Markov chain with the same transition rule $x \mapsto F_\epsilon$ lead to the same steady state distribution, which must be the (unique) invariant measure.

The coupling construction is easiest to describe in the form of an algorithm. Before describing the algorithm, let us select $\delta > 0$ so that the local product $(x, x') \mapsto [x, x']$ makes sense whenever $d(x, x') < \delta$. (Recall that the local product $[\cdot,\cdot]$ maps (x, x') to the unique point in the intersection of the local unstable manifold of x and the local stable manifold of x'. See §2.2.1.)

Algorithm 4.6 (coupling for small ϵ).

1: sample initial points x_0^i from m_i for $i = 1, 2$ independently
2: **for** $n := 1, 2, \ldots$ **do**
3: set $x_n^i := F\left(x_{n-1}^i\right)$ for $i = 1, 2$
4: **if** $d\left(x_n^1, x_n^2\right) \geq \delta$ **then** *(there is no chance for coupling)*
5: set $x_n^i := F_\epsilon\left(x_{n-1}^i\right)$ for $i = 1, 2$ independently
6: **else**
7: let $\Phi : Y_\epsilon\left(x_n^1\right) \to Y_\epsilon\left(x_n^2\right)$ be the stable holonomy map and let $J\Phi$ be its jacobian
8: select sample point x' from $p_\epsilon\left(x_n^1, \cdot\right)$ *(we will try to draw a sample from $q_\epsilon(x_n^2, \cdot)$ by rejection sampling, using $q_\epsilon(x_n^1, \cdot)$ and the holonomy map Φ)*
9: set $c := \sup \left| \frac{q_\epsilon\left(x_n^1, \Phi^{-1}(\cdot)\right)}{J\Phi(\Phi^{-1}(\cdot)) \cdot q_\epsilon(x_n^2, \cdot)} \right|$
10: **if** UNIFORMRANDOM() $\leq c^{-1} \cdot \frac{q_\epsilon\left(x_n^1, x'\right)}{J\Phi(x') \cdot q_\epsilon(x_n^2, \Phi(x'))}$ **then** *(coupled for one step!)*
11: set $x_n^1 := x'$
12: set $x_n^2 := \Phi\left(x'\right)$
13: **else** *(no luck, try again)*
14: set $x_n^1 := x'$
15: set $x_n^2 := F_\epsilon\left(x_{n-1}^2\right)$ independently
16: **end if**
17: **end if**
18: **end for**

To show that this algorithm produces a joint process (x_n^1, x_n^2) with the desired properties, we need to show that as $n \to \infty$, $d\left(x_n^1, x_n^2\right) \to 0$ almost surely. This is implied by the following:

1. Every time x_n^1 and x_n^2 come within a distance of δ of each other, there is a finite probability that they would couple for all future times.

2. The event $\left(d\left(x_n^1, x_n^2\right) \leq \delta\right)$ occurs infinitely often.

The probability of coupling is determined by line 10 of Algorithm 4.6. This is equivalent to the probability of success of rejection sampling (see [Mac98]). This probability is bounded

below by

$$p_{\min}\left(x^1, x^2\right) = \inf_{x'}\left(c^{-1} \cdot \frac{q_\epsilon\left(x^1, x'\right)}{J\Phi(x') \cdot q_\epsilon(x^2, \Phi(x'))}\right),$$
$$c = \sup_{x}\left|\frac{q_\epsilon\left(x^1, \Phi^{-1}(x)\right)}{J\Phi(\Phi^{-1}(x)) \cdot q_\epsilon\left(x^2, x\right)}\right|.$$

The jacobian $J\Phi$ is α-Hölder continuous for some $\alpha > 0$ and the densities vary smoothly with x^1 and x^2. Furthermore, the difference $J\Phi - 1$ is bounded away from 0 and its Hölder norm is proportional to the C^1 distance between the manifolds $Y_\epsilon(x^1)$ and $Y_\epsilon(x^2)$ (see [BP02]). If x^1 and x^2 are coupled at step m, then, they will lie on the same stable manifold. At the next step, the application of the map F will contract the distance along the stable manifold, so that x^1_{n+1} and x^2_{n+1} are closer than x^1_n and x^2_n by a factor of $\beta < 1$. The Hölder continuity of $J\Phi$ and the smoothness of the distributions $q_\epsilon(x, \cdot)$ ensure the existence of a positive constant c such that

$$p_{\min}\left(x^1, x^2\right) \geq e^{-c \cdot d\left(x^1, x^2\right)^\alpha}. \tag{4.24}$$

Thus the probability that x^1 and x^2 remain coupled for all time is bounded below by

$$\prod_{n=0}^{\infty} e^{-a \cdot \beta^{n\alpha} \cdot d\left(x^1_m, x^2_m\right)^\alpha} \geq e^{-c\delta^\alpha/(1-\beta^\alpha)} > 0. \tag{4.25}$$

So each time the processes couple there is a finite probability that they will remain coupled forever.

To see that x^1_n and x^2_n will have infinitely many opportunities for coupling, fix x^1_0 and define

$$\begin{aligned} \tilde{x}^1_{n+1} &= \left[F(\tilde{x}^1_n), x^1_{n+1}\right], \\ \tilde{x}^1_0 &= x^1_0. \end{aligned} \tag{4.26}$$

Thus, $\tilde{x}^1_{n+1}$ is a projection of x^1_{n+1} onto the unstable manifold of $F(x^1_n)$. Let $d_n = d(\tilde{x}^1_n, x^1_n)$. Then

$$\begin{aligned} d_{n+1} &= d(\tilde{x}^1_{n+1}, x^1_{n+1}) \\ &= d\left(\left[F^k(\tilde{x}^1_n), F^k_\epsilon(x^1_n)\right], F^k_\epsilon(x^1_n)\right) \\ &\leq d\left(\left[F^k(\tilde{x}^1_n), F^k_\epsilon(x^1_n)\right], F^k(\tilde{x}^1_n)\right) + d(F^k(\tilde{x}^1_n), F^k(x^1_n)) + d(F^k(x^1_n), F^k_\epsilon(x^1_n)) \\ &\leq C\epsilon^\epsilon + \beta d_n + \epsilon. \end{aligned}$$

By induction,

$$\sup_n \{d_n\} \leq \frac{C\epsilon^\alpha + \epsilon}{1-\beta}. \tag{4.27}$$

The images of a local unstable manifold under the action of the map is dense in the attractor K in the following sense: for every open set U such that $U \cap K \neq \varnothing$, every $\epsilon > 0$, and every $x \in K$, there exists an integer n large enough that $F^n\left(W^u_\epsilon(x)\right) \cap U \neq \varnothing$. To see this, one only needs to check that the density result holds for the corresponding symbolic dynamics, and hence is also true for $F : X \to X$. (The proof is fairly straightforward and is omitted here.)

Let S_n denote the support of the distribution of $\tilde{x}^1_n$. As $\tilde{x}^1_1$ is the projection of x^1_1 onto a the unstable manifold of $F(x^1_0)$ and the transition measure which defines our random perturbations has a uniformly positive density on the manifold $Y_\epsilon(F(x^1_0))$, it follows that the support of the distribution of $\tilde{x}^1_1$ is the image of $Y_\epsilon(F(x^1_0))$ projected onto the unstable manifold of $F(x^1_0)$. Because the manifolds $Y_\epsilon(x)$ are transverse to the stable directions, this projection must contain a small open neighborhood of $F(x^1_0)$ in its unstable manifold. By induction, and by the density of images of local unstable manifolds explained above, there exists an N such that $P\left(\tilde{x}^1_N \in U\right) > 0$. Applying the same construction to x^2_n, we see that for each δ there is an N large enough that $P\left(d\left(\tilde{x}^1_N, \tilde{x}^2_N\right) < \delta/2\right) > 0$. Hence $P\left(d\left(x^1_N, x^2_N\right) < \delta\right) > 0$ if N is large and ϵ is small. By the compactness of K and the continuity of $p_\epsilon(x, \cdot)$, there exists a single N which makes this work. Thus there is a single constant $p'_{\min} > 0$ such that

$$P\left(d\left(x^1_N, x^2_N\right) < \delta\right) \geq p'_{\min} \tag{4.28}$$

for any (fixed) (x^1_0, x^2_0). The rest is straightforward:

$$\begin{aligned} P\left(d\left(x^1_n, x^2_n\right) \geq \delta : \text{almost always}\right) &\leq P\left(d\left(x^1_{nN}, x^2_{nN}\right) \geq \delta : \text{almost always}\right) \\ &= P\left(\cup_{m=0}^\infty \cap_{n \geq m} \left(d\left(x^1_{nN}, x^2_{nN}\right) \geq \delta\right)\right) \\ &\leq \sum_{m=0}^\infty P\left(\cap_{n \geq m} \left(d\left(x^1_{nN}, x^2_{nN}\right) \geq \delta\right)\right). \end{aligned}$$

Since

$$\begin{aligned} &P\left(\cap_{n \geq m} \left(d\left(x^1_{nN}, x^2_{nN}\right) \geq \delta\right)\right) \leq \\ &\qquad P\left(d\left(x^1_m, x^2_m\right) \geq \delta\right) \cdot \prod_{n=m}^\infty P\left(d\left(x^1_{(n+1)N}, x^2_{(n+1)N}\right) \geq \delta | x^1_{nN}, x^2_{nN}\right) \end{aligned} \tag{4.29}$$

and Equation (4.28) implies that

$$P\left(d\left(x^1_{(n+1)N}, x^2_{(n+1)N}\right) \geq \delta | x^1_{nN}, x^2_{nN}\right) \leq 1 - p'_{\min} < 1,$$

we find that

$$P\left(d\left(x^1_n, x^2_n\right) < \delta : \text{infinitely often}\right) = 1.$$

□

This proof also works if the supports $Y_\epsilon(x)$ of the transition measures have dimension strictly larger than $\dim W^u$, but the notation becomes a bit messier. Note that a more detailed analysis of the probability distribution of the coupling time n_h would yield estimates of the decay rate of correlation functions for the perturbed system; this line of attack is part of on-going work. Note also that for technical reasons, Propositions 4.3 and 4.5 apply in different situations. They will both apply as stated only when the convolution-type distributions in Proposition 4.3 are everywhere transverse to the stable subspaces. (It should be possible to replace "everywhere" by "almost everywhere.")

See Baladi [Bal00], Bressaud and Liverani [BL02], and Young [You98] for other uses of the coupling method in the study of hyperbolic systems. David Griffeath's thesis [Gri78] offers a beautifully lucid introduction to the coupling method for Markov chains.

Chapter 5

The Kuramoto-Sivashinsky Equation

Preceding chapters explored the effects of noise in hyperbolic dynamical systems through numerical and analytical techniques. This chapter describes two sets of numerical experiments on the Kuramoto-Sivashinsky equation (KSE):

$$\partial_t u + u\partial_x u + \alpha\partial_x^2 u + \beta\partial_x^4 u = 0, \tag{5.1}$$

where the solution $u = u(t,x)$ is assumed to be periodic in x with period L and α, β are positive constants.

The KSE originally arose as a model of interface phenomena in classical continuum physics (see [KPZ86] and references therein). It also played an important role in the study of infinite dimensional dynamics [Tem88]. This chapter describes the calculation of some statistical properties of the KSE and studies the effects of noise on these quantities.

5.1 Properties of the KSE

In [Tem88], it is shown that the initial value problem (5.1) is well posed for $t \in [0, +\infty)$ with initial conditions in an appropriately chosen Hilbert space $\mathbb{H}$ and that solutions remain bounded for all time. Furthermore, as a dynamical system in the an infinite-dimensional Hilbert space $\mathbb{H}$, Equation (5.1) possesses a finite-dimensional invariant manifold $\Lambda \subset \mathbb{H}$, called the inertial manifold, which attracts all trajectories exponentially fast. The implication of this result is clear: even though the phase space $\mathbb{H}$ has infinitely many dimensions, the dynamics of the KSE can be prescribed by a finite number of variables as

t becomes large. PDEs with inertial manifolds provide a relatively simple setting in which to extend the statistical theory of dynamical systems to infinite-dimensional problems.

Let $e_k(x) = e^{i2\pi kx/L}$ denote the usual Fourier basis, $\mathbb{H}_N$ the span of $(e_{-N}, ..., e_N)$, and $\mathbb{P}_N$ projection onto $\mathbb{H}_N$. Then in addition to the existence of inertial manifolds, one can prove that for all sufficiently large N, there exists a function Φ_N mapping $\mathbb{H}_N$ into its orthogonal complement such that

$$||u(t) - \mathbb{P}_N u(t) - \Phi_N(\mathbb{P}_N u(t))||_{\mathbb{H}} \to 0 \tag{5.2}$$

exponentially fast as $t \to \infty$. That is, asymptotically, we can write

$$u(t,x) \approx \mathbb{P}_N u(t,x) + \Phi_N(\mathbb{P}_N u(t)). \tag{5.3}$$

Furthermore, it is possible to approximate the maps Φ_N numerically [NTW01]. The existence of inertial manifolds implies that any invariant measure μ of (5.1) is supported on a finite dimensional submanifold of $\mathbb{H}$. In the language of optimal prediction (see §1.3), the conditional measures of μ conditioned on $\mathbb{P}_N u$ are trivial if one takes N large enough: they are Dirac δ measures positioned at $\Phi_N(\mathbb{P}_N u)$.

Rewriting (5.1) in Fourier spectral variables yields

$$\frac{d}{dt}\hat{u}_k(t) = -\frac{i\omega_0 k}{2} \sum_{k_1+k_2=k} \hat{u}_{k_1}(t)\hat{u}_{k_2}(t) + \omega_0^2 k^2 \left(\alpha - \beta\omega_0^2 k^2\right) \hat{u}_k(t), \tag{5.4}$$

$$\omega_0 = 2\pi/L.$$

Linearizing about the fixed point $u \equiv 0$ yields

$$\frac{d}{dt}\delta\hat{u}_k(t) = \omega_0^2 k^2 \left(\alpha - \beta\omega_0^2 k^2\right) \delta\hat{u}_k(t). \tag{5.5}$$

Thus, those modes with wave number

$$k < N_{\text{crit}} \stackrel{\text{def}}{=} \frac{1}{\omega_0}\sqrt{\frac{\alpha}{\beta}} \tag{5.6}$$

are linearly unstable, while those modes with wave number $k > N_{\text{crit}}$ are linearly stable. The trivial solution $u \equiv 0$ is thus a hyperbolic fixed point of the KS flow. This linear analysis also shows that the dimension of the inertial manifold is at least N_{crit}, since by definition the inertial manifold must contain all unstable degrees of freedom. In addition, we see that the map $\Phi_N : \mathbb{H}_N \to \mathbb{H}$ exists only if $N > N_{\text{crit}}$.

In the opposite limit $N \ll N_{\text{crit}}$, physical scaling arguments predict that the dynamics of $\hat{u}_N$ can be modeled by a stochastically-forced Burgers equation [For75, KPZ86, Yak81]:

$$\partial_t u + u\partial_x u = \nu \partial_x^2 u + f, \tag{5.7}$$

where ν is an effective viscosity constant and f is a white noise forcing term. In view of these facts, it is natural to ask if one may derive Equation (5.7) from Equation (5.1) using a systematic optimal prediction procedure, such as the one recently suggested by Stinis [Sti03]. In addition, there are other interesting questions regarding the long-term predictability of the KSE:

1. As we increase N past the critical value N_{crit}, how does $(u - \mathbb{P}_N u)$ become more predictable in terms of $\mathbb{P}_N u$?

2. To answer the question above, it is natural to begin by performing numerical experiments. But, how do we tell if the system has "relaxed" to equilibrium in such experiments? What does one do in the presence of multiple attractors (and hence multiple invariant probability measures)?

3. How can we quantify the uncertainly left in $(u - \mathbb{P}_N u)$ given knowledge of $\mathbb{P}_N u$? One candidate is the relative entropy $H(Y|X)$, which measures the amount of information in one random variable Y when one has knowledge of another random variable X [CT91]. The amount of computational resources required to compute $H(u - \mathbb{P}_N u|\mathbb{P}_N u)$ directly is prohibitively large, though.

There is a long history of numerical simulations of the KSE, with a particular emphasis on the statistical properties of solutions and the verification of (5.7) [SKJ+92, Zal89]. However, it is difficult to assess the reliability of these early studies. There is, in contrast, very detailed work on the geometry of the KSE in recent years [JJK01].

Note that the KSE is invariant under translations in the sense that if $u : \mathbb{T} \times [0, +\infty) \to \mathbb{R}$ is a solution, then so is

$$(T_{x_0} u)(x, t) = u(x + x_0, t). \tag{5.8}$$

Furthermore, if $S_t : \mathbb{H} \to \mathbb{H}$ denotes the solution map of the KSE, then $T_{x_0} S_t(u_0) = S_t(T_{x_0} u_0)$. Hence two solutions which are translates of each other remain translates of each other under the KS flow. The unitarity of the translation operator $T_{x_0} : \mathbb{H} \to \mathbb{H}$ and the fact

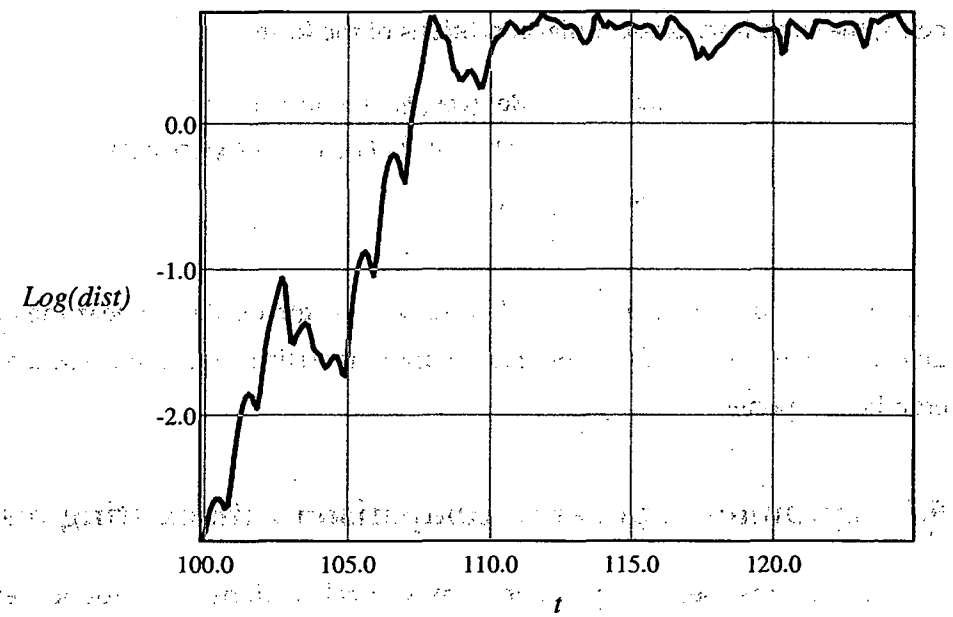


Figure 5.1: The exponential separation of nearby initial conditions in the Kuramoto-Sivashinsky equation, for $\alpha = 1$ and $\beta = 0.085$.

that solutions are bounded for all time means that there are center manifolds associated with the translation invariance.

5.2 Parameters & initial conditions

In studying chaotic dynamical systems, one of the most basic problems is the stability of dynamical features under variations of the system's parameters. For the KSE, if we fix $\alpha = L = 1$ and vary the "hyperviscosity" β, the system will undergo a sequence of bifurcations. (See [JJK01, Sti03] and the references there.) The KSE exhibits very different behavior for different parameters, and not all interesting features of this system are stable under small perturbations of the parameters. For concreteness, the numerical experiments in this chapter focus on the parameter values

$$L = 2\pi, \alpha = 1, \beta = 0.085. \tag{5.9}$$

For this set of parameter values, the KSE exhibits exponential separation of nearby initial conditions (see Figure 5.1). As the KSE may potentially have infinitely ergodic compo-

nents, the experiments use initial conditions of the form

$$\begin{aligned} u_0(x) &= a \cdot \mathrm{Re}\left[(\cos(\theta_0) + i\sin(\theta_0))\, e^{i\omega_0 x}\right], \\ a &= a_0 \cdot (1 + 10^{-6} \cdot \textsc{UniformRandom}()), \\ a_0 &= 10^{-3}, \\ \theta_0 &= \tfrac{\pi}{6}. \end{aligned} \tag{5.10}$$

The initial condition (5.10) is a small long wavelength excitation near the origin and may give us a better chance of obtaining unambiguous statistical results associated with a single ergodic component.

5.3 Lyapunov exponents: comparison with existing results

The first set of experiments involve calculations of Lyapunov exponents estimated using Equation (3.1) and Algorithm 3.1. The exponents were computed using both periodic (gaussian) kicks in asymptotically unstable directions. Figures 5.2 - 5.6 show the results computed by periodic random perturbations in asymptotically unstable directions. In these calculations, the perturbed KS flow is sampled at intervals of $T = 0.01$ for 5×10^5 steps and plotted as a function of the noise variance ϵ. The empirical decay time $\hat{\tau}_{\text{exp}}$ is defined here to be the time required for the correlation function to decay to $\frac{1}{10}$ of its value at $t = 0$, and the empirical estimate $\hat{\tau}_{\text{int}}$ of the integrated autocorrelation times is computed by numerically integrating the normalized autocovariance function from 0 to $\hat{\tau}_{\text{exp}}$ using the trapezoid rule.

As in the case of the Lorenz flow, the exponential decay time decreases with increasing ϵ and $\hat{\tau}_{\text{int}}$ remains more or less static with increasing ϵ (see Figures 5.7 - 5.9 and Figure 5.10). Tables 5.1 - 5.3 show that the standard deviations of the Lyapunov exponents do not decrease with increasing ϵ; this is also similar to the behavior of the Lorenz flow. Note that the algorithm did not produce the Lyapunov exponents in order of magnitude; this phenomenon is explained in [GPL90]. In agreement with Figure 5.1, the Lyapunov exponent data show that this choice of parameters produces (weakly) chaotic dynamics with very strong contractions and only two small positive Lyapunov exponents.

The distribution of Lyapunov exponents may be compared with those computed by Christiansen, Cvitanović, and Putkaradze [CCP97] using more sophisticated periodic orbit techniques. Note that they consider the restriction of the KSE to the space of odd-

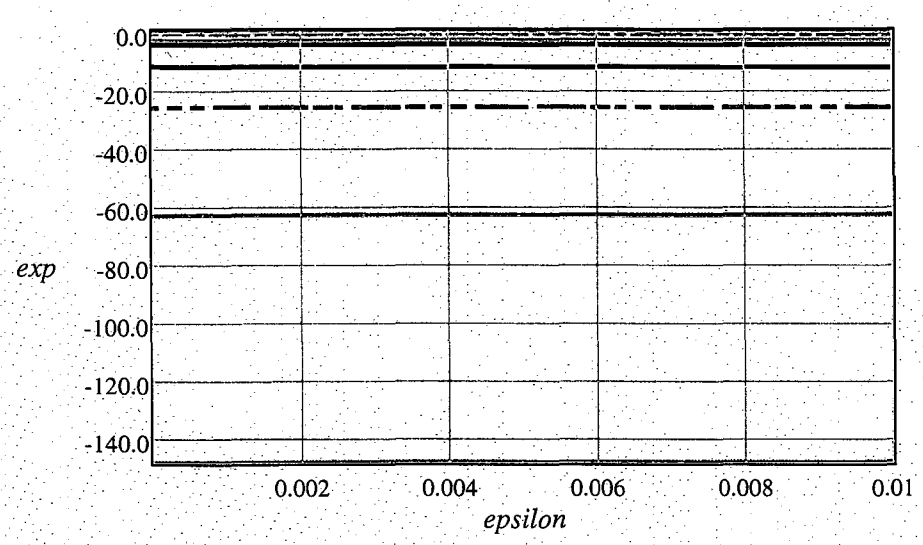


Figure 5.2: Estimated Lyapunov exponents for the KS flow kicked by gaussian noise of variance ϵ in the asymptotically unstable direction.

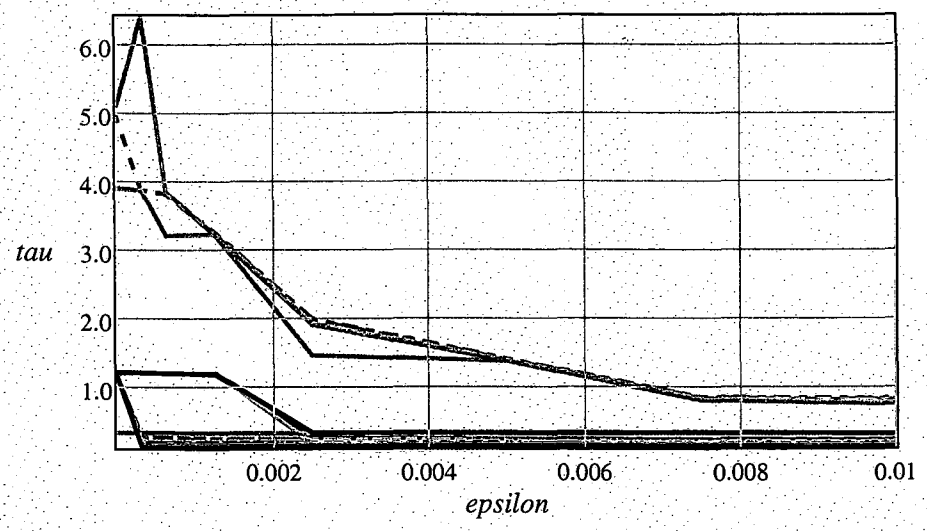


Figure 5.3: Decay times of the KS flow periodically kicked by gaussian noise of variance ϵ in the asymptotically unstable direction.

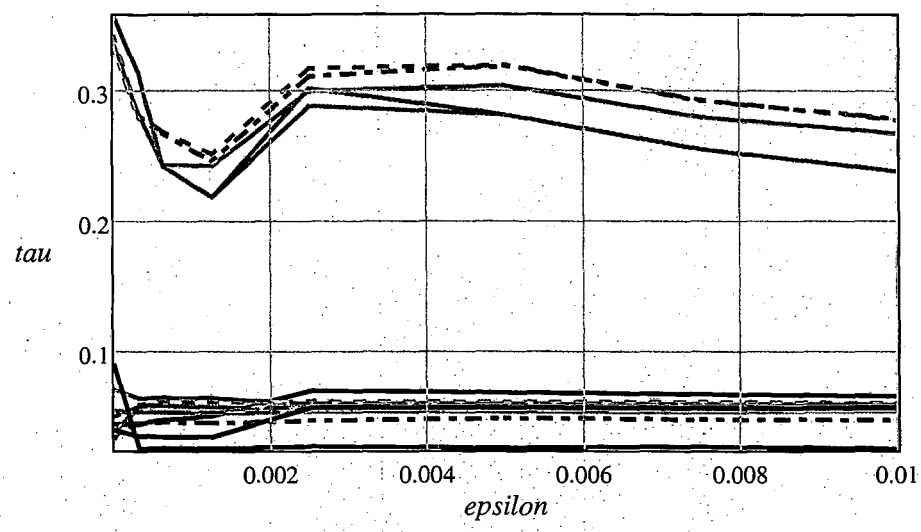


Figure 5.4: Integrated correlation times of the KS flow periodically kicked by gaussian noise of variance ϵ in the asymptotically unstable direction.

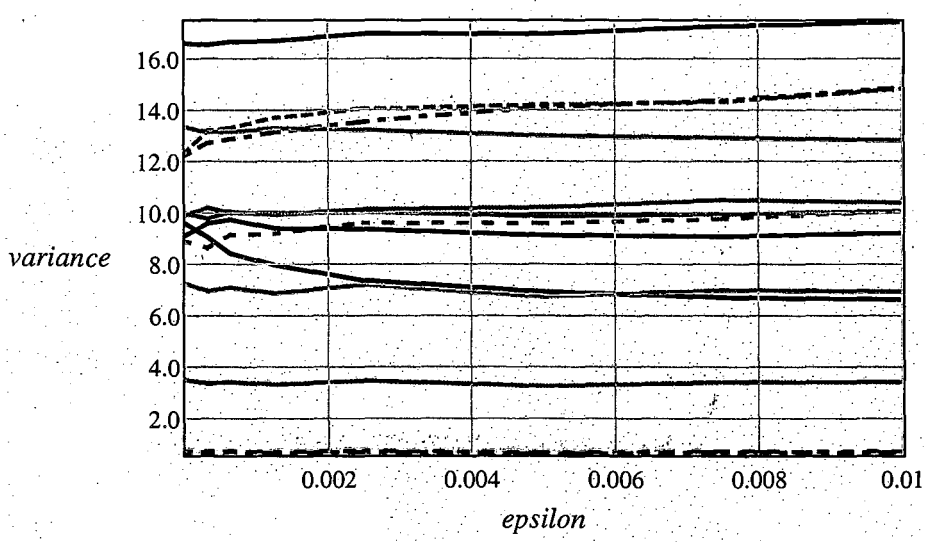


Figure 5.5: Variances of the local expansion rates for the KS flow periodically kicked by gaussian noise of variance ϵ in the asymptotically unstable direction.

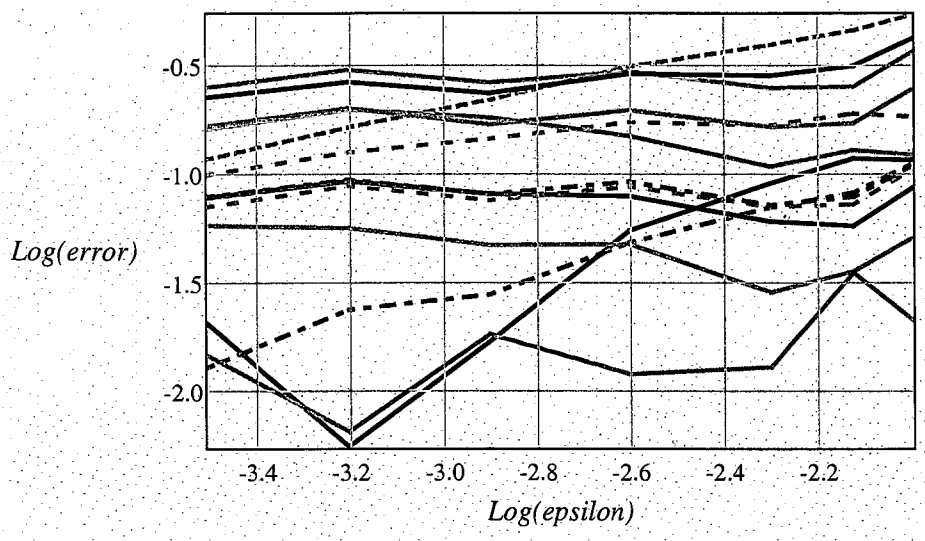


Figure 5.6: Approximate errors for the estimated Lyapunov exponents of the KS flow periodically kicked by gaussian noise of variance ϵ in the asymptotically unstable direction.

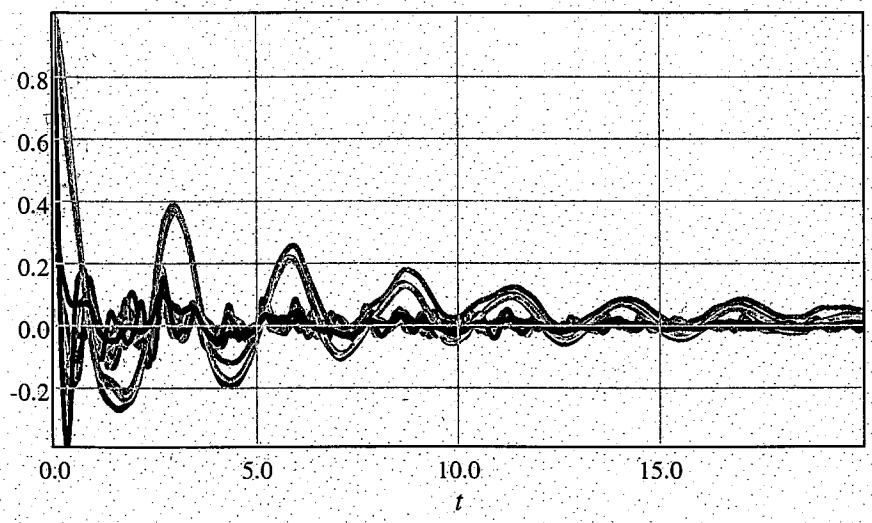


Figure 5.7: Autocovariance functions for the local expansion rates of the unperturbed KS flow.

ϵ	$\hat{\lambda}_1^\epsilon$	τ_{exp}	τ_{int}	$\lvert\hat{\lambda}_1^\epsilon - \hat{\lambda}_1^0\rvert$
0.0000	$0.488 \pm 1.3 \times 10^{-2}$	1.202	0.044	0.000
0.0025	$0.543 \pm 1.4 \times 10^{-2}$	0.260	0.070	0.054
0.0050	$0.579 \pm 1.3 \times 10^{-2}$	0.238	0.068	0.090
0.0075	$0.606 \pm 1.3 \times 10^{-2}$	0.230	0.067	0.117
0.0100	$0.604 \pm 1.3 \times 10^{-2}$	0.217	0.065	0.116

ϵ	$\hat{\lambda}_2^\epsilon$	τ_{exp}	τ_{int}	$\lvert\hat{\lambda}_2^\epsilon - \hat{\lambda}_2^0\rvert$
0.0000	$0.001 \pm 1.0 \times 10^{-2}$	1.159	0.033	0.000
0.0025	$0.175 \pm 1.5 \times 10^{-2}$	0.238	0.061	0.173
0.0050	$0.169 \pm 1.5 \times 10^{-2}$	0.230	0.061	0.167
0.0075	$0.192 \pm 1.5 \times 10^{-2}$	0.234	0.061	0.190
0.0100	$0.185 \pm 1.5 \times 10^{-2}$	0.230	0.059	0.183

ϵ	$\hat{\lambda}_3^\epsilon$	τ_{exp}	τ_{int}	$\lvert\hat{\lambda}_3^\epsilon - \hat{\lambda}_3^0\rvert$
0.0000	$0.00 \pm 1.3 \times 10^{-2}$	1.150	0.038	0.000
0.0025	$-0.04 \pm 1.5 \times 10^{-2}$	0.182	0.046	0.048
0.0050	$-0.07 \pm 1.6 \times 10^{-2}$	0.204	0.048	0.070
0.0075	$-0.08 \pm 1.6 \times 10^{-2}$	0.191	0.047	0.082
0.0100	$-0.11 \pm 1.6 \times 10^{-2}$	0.182	0.047	0.111

ϵ	$\hat{\lambda}_4^\epsilon$	τ_{exp}	τ_{int}	$\lvert\hat{\lambda}_4^\epsilon - \hat{\lambda}_4^0\rvert$
0.0000	$-0.11 \pm 1.5 \times 10^{-2}$	1.155	0.050	0.000
0.0025	$-0.43 \pm 1.7 \times 10^{-2}$	0.238	0.056	0.312
0.0050	$-0.51 \pm 1.7 \times 10^{-2}$	0.234	0.055	0.392
0.0075	$-0.57 \pm 1.7 \times 10^{-2}$	0.225	0.055	0.457
0.0100	$-0.65 \pm 1.7 \times 10^{-2}$	0.221	0.054	0.538

ϵ	$\hat{\lambda}_5^\epsilon$	τ_{exp}	τ_{int}	$\lvert\hat{\lambda}_5^\epsilon - \hat{\lambda}_5^0\rvert$
0.0000	$-2.57 \pm 1.9 \times 10^{-2}$	1.194	0.070	0.000
0.0025	$-2.72 \pm 1.7 \times 10^{-2}$	0.182	0.059	0.149
0.0050	$-2.68 \pm 1.7 \times 10^{-2}$	0.182	0.060	0.108
0.0075	$-2.70 \pm 1.7 \times 10^{-2}$	0.182	0.059	0.129
0.0100	$-2.69 \pm 1.7 \times 10^{-2}$	0.178	0.059	0.122

Table 5.1: Correlation times of the estimated Lyapunov exponents of the KS flow periodically kicked by gaussian noise of variance ϵ in the asymptotically unstable direction. Estimated Lyapunov exponents are stated with their standard deviations.

ϵ	$\hat{\lambda}_6^\epsilon$	τ_{exp}	τ_{int}	$\lvert\hat{\lambda}_6^\epsilon - \hat{\lambda}_6^0\rvert$
0.0000	$-3.56 \pm 1.2 \times 10^{-2}$	1.163	0.038	0.000
0.0025	$-3.57 \pm 1.5 \times 10^{-2}$	0.195	0.057	0.011
0.0050	$-3.55 \pm 1.5 \times 10^{-2}$	0.191	0.056	0.012
0.0075	$-3.52 \pm 1.5 \times 10^{-2}$	0.186	0.056	0.035
0.0100	$-3.54 \pm 1.5 \times 10^{-2}$	0.186	0.055	0.021

ϵ	$\hat{\lambda}_7^\epsilon$	τ_{exp}	τ_{int}	$\lvert\hat{\lambda}_7^\epsilon - \hat{\lambda}_7^0\rvert$
0.0000	$-3.83 \pm 1.4 \times 10^{-2}$	0.338	0.054	0.000
0.0025	$-3.88 \pm 1.4 \times 10^{-2}$	0.325	0.053	0.047
0.0050	$-3.86 \pm 1.4 \times 10^{-2}$	0.321	0.053	0.028
0.0075	$-3.87 \pm 1.4 \times 10^{-2}$	0.325	0.053	0.035
0.0100	$-3.89 \pm 1.4 \times 10^{-2}$	0.312	0.054	0.050

ϵ	$\hat{\lambda}_8^\epsilon$	τ_{exp}	τ_{int}	$\lvert\hat{\lambda}_8^\epsilon - \hat{\lambda}_8^0\rvert$
0.0000	$-4.03 \pm 1.2 \times 10^{-2}$	1.211	0.039	0.000
0.0025	$-4.10 \pm 1.4 \times 10^{-2}$	0.334	0.056	0.079
0.0050	$-4.09 \pm 1.4 \times 10^{-2}$	0.330	0.056	0.060
0.0075	$-4.08 \pm 1.4 \times 10^{-2}$	0.334	0.056	0.057
0.0100	$-4.11 \pm 1.4 \times 10^{-2}$	0.325	0.056	0.086

ϵ	$\hat{\lambda}_9^\epsilon$	τ_{exp}	τ_{int}	$\lvert\hat{\lambda}_9^\epsilon - \hat{\lambda}_9^0\rvert$
0.0000	$-25.7 \pm 9.8 \times 10^{-3}$	4.942	0.341	0.000
0.0025	$-25.7 \pm 9.6 \times 10^{-3}$	1.980	0.317	0.087
0.0050	$-25.7 \pm 9.1 \times 10^{-3}$	1.394	0.319	0.070
0.0075	$-25.7 \pm 9.0 \times 10^{-3}$	0.838	0.293	0.072
0.0100	$-25.6 \pm 8.7 \times 10^{-3}$	0.812	0.277	0.104

ϵ	$\hat{\lambda}_{10}^\epsilon$	τ_{exp}	τ_{int}	$\lvert\hat{\lambda}_{10}^\epsilon - \hat{\lambda}_{10}^0\rvert$
0.0000	$-25.8 \pm 9.4 \times 10^{-3}$	3.899	0.331	0.000
0.0025	$-25.7 \pm 9.2 \times 10^{-3}$	1.993	0.310	0.092
0.0050	$-25.7 \pm 8.8 \times 10^{-3}$	1.402	0.318	0.071
0.0075	$-25.7 \pm 8.8 \times 10^{-3}$	0.838	0.292	0.078
0.0100	$-25.7 \pm 8.5 \times 10^{-3}$	0.820	0.278	0.109

Table 5.2: Correlation times of the estimated Lyapunov exponents of the KS flow periodically kicked by gaussian noise of variance ϵ in the asymptotically unstable direction. Estimated Lyapunov exponents are stated with their standard deviations.

ϵ	$\hat{\lambda}^{\epsilon}_{11}$	τ_{exp}	τ_{int}	$\lvert\hat{\lambda}^{\epsilon}_{11} - \hat{\lambda}^{0}_{11}\rvert$
0.0000	$-63.0 \pm 3.1 \times 10^{-2}$	3.904	0.335	0.000
0.0025	$-62.7 \pm 2.9 \times 10^{-2}$	1.923	0.299	0.296
0.0050	$-62.7 \pm 2.8 \times 10^{-2}$	1.385	0.303	0.249
0.0075	$-62.7 \pm 2.7 \times 10^{-2}$	0.829	0.280	0.252
0.0100	$-62.6 \pm 2.7 \times 10^{-2}$	0.803	0.267	0.364

ϵ	$\hat{\lambda}^{\epsilon}_{12}$	τ_{exp}	τ_{int}	$\lvert\hat{\lambda}^{\epsilon}_{12} - \hat{\lambda}^{0}_{12}\rvert$
0.0000	$-63.0 \pm 3.1 \times 10^{-2}$	3.904	0.335	0.000
0.0025	$-62.7 \pm 2.9 \times 10^{-2}$	1.923	0.299	0.296
0.0050	$-62.7 \pm 2.8 \times 10^{-2}$	1.381	0.303	0.249
0.0075	$-62.7 \pm 2.7 \times 10^{-2}$	0.829	0.280	0.252
0.0100	$-62.6 \pm 2.7 \times 10^{-2}$	0.803	0.267	0.364

ϵ	$\hat{\lambda}^{\epsilon}_{13}$	τ_{exp}	τ_{int}	$\lvert\hat{\lambda}^{\epsilon}_{13} - \hat{\lambda}^{0}_{13}\rvert$
0.0000	$-148.0 \pm 2.2 \times 10^{-2}$	5.107	0.356	0.000
0.0025	$-147.0 \pm 1.9 \times 10^{-2}$	1.906	0.288	0.196
0.0050	$-147.0 \pm 1.9 \times 10^{-2}$	1.376	0.281	0.164
0.0075	$-147.0 \pm 1.8 \times 10^{-2}$	0.803	0.255	0.171
0.0100	$-147.0 \pm 1.8 \times 10^{-2}$	0.751	0.238	0.246

ϵ	$\hat{\lambda}^{\epsilon}_{14}$	τ_{exp}	τ_{int}	$\lvert\hat{\lambda}^{\epsilon}_{14} - \hat{\lambda}^{0}_{14}\rvert$
0.0000	$-148.0 \pm 2.2 \times 10^{-2}$	5.107	0.356	0.000
0.0025	$-147.0 \pm 2.0 \times 10^{-2}$	1.463	0.301	0.196
0.0050	$-147.0 \pm 1.9 \times 10^{-2}$	1.376	0.281	0.164
0.0075	$-147.0 \pm 1.8 \times 10^{-2}$	0.803	0.255	0.171
0.0100	$-147.0 \pm 1.8 \times 10^{-2}$	0.751	0.238	0.246

ϵ	$\hat{\lambda}^{\epsilon}_{15}$	τ_{exp}	τ_{int}	$\lvert\hat{\lambda}^{\epsilon}_{15} - \hat{\lambda}^{0}_{15}\rvert$
0.0000	$-11.4 \pm 2.4 \times 10^{-2}$	1.168	0.088	0.000
0.0025	$-11.7 \pm 1.3 \times 10^{-2}$	0.134	0.027	0.289
0.0050	$-11.7 \pm 1.3 \times 10^{-2}$	0.130	0.026	0.283
0.0075	$-11.7 \pm 1.3 \times 10^{-2}$	0.134	0.026	0.314
0.0100	$-11.8 \pm 1.3 \times 10^{-2}$	0.125	0.025	0.419

Table 5.3: Correlation times of the estimated Lyapunov exponents of the KS flow periodically kicked by gaussian noise of variance ϵ in the asymptotically unstable direction. Estimated Lyapunov exponents are stated with their standard deviations.

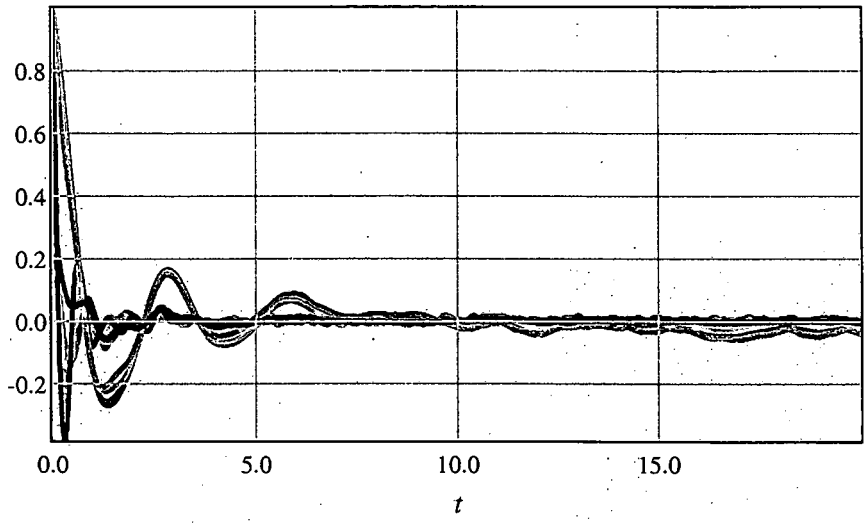


Figure 5.8: Autocovariance functions for the local expansion rates of the KSE. The flow is periodically kicked in asymptotically unstable directions by gaussian random variables of variance $\epsilon = 0.005$.

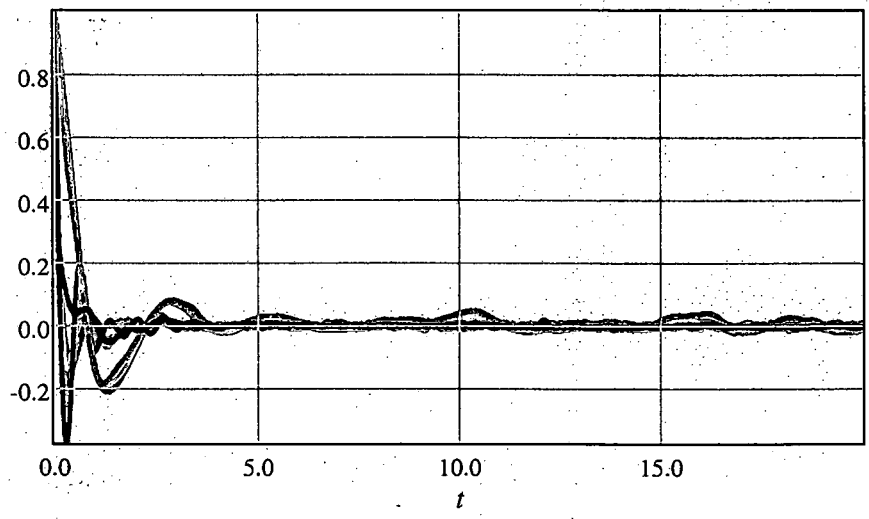


Figure 5.9: Autocovariance functions for the local expansion rates of the KSE. The flow is periodically kicked in asymptotically unstable directions by gaussian random variables of variance $\epsilon = 0.01$.

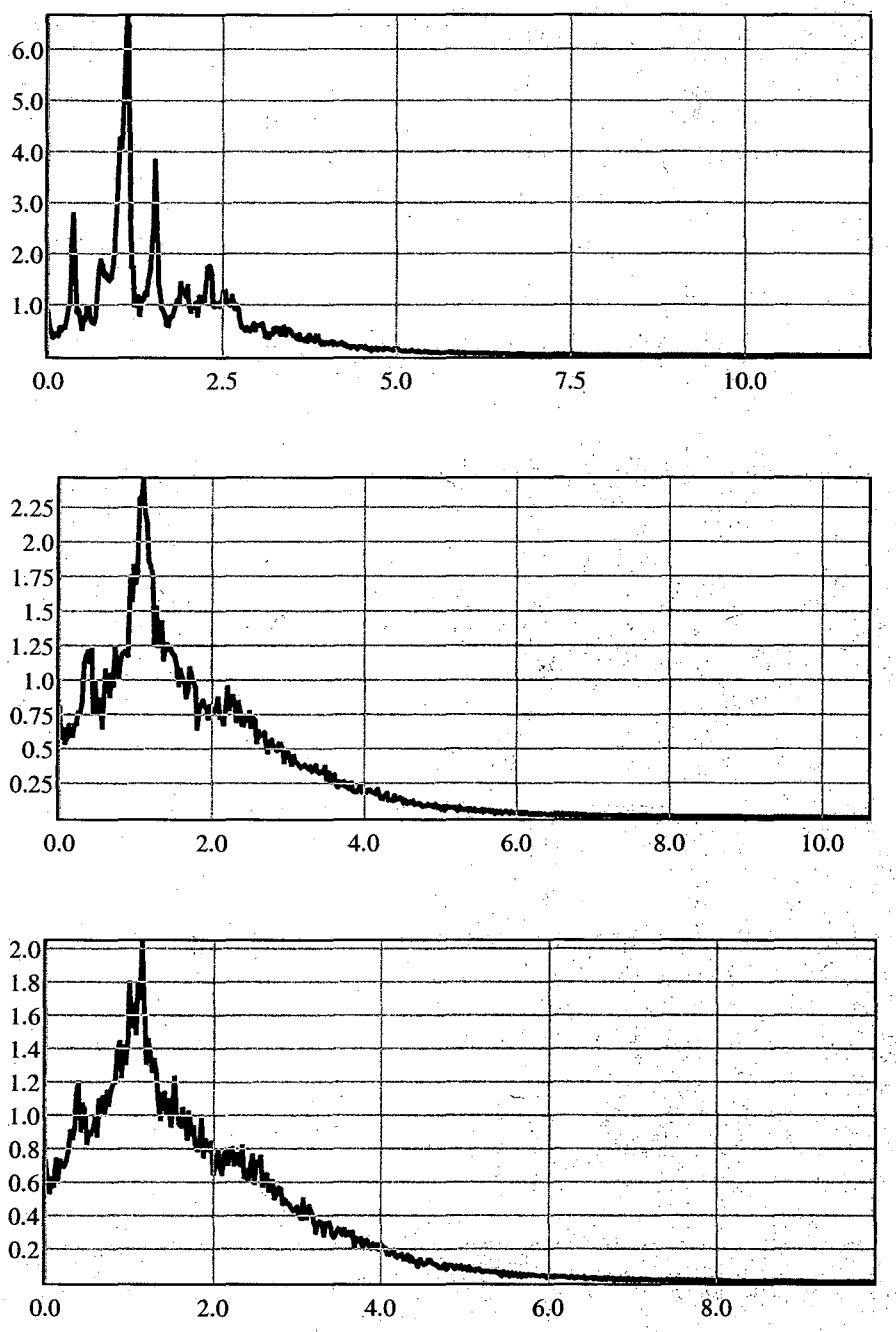


Figure 5.10: Spectral power densities $\hat{C}(f)$ associated with the largest estimated Lyapunov exponent $\hat{\lambda}_1$ of the KS flow, plotted against the frequency f, for $\epsilon \in \{0.0, 0.005, 0.01\}$.

wave-number modes because the relevant cycle expansions converge only for hyperbolic periodic orbits: while the center manifold associated with the direction of the flow can be factored out using a Poincaré section, the center manifold associated with translation invariance may cause problems for cycle expansions.

5.4 Mode-mode mutual information

The second set of experiments concern the statistical dependence between Fourier modes. Specifically, a pair of partitions Ξ_k^r and Ξ_k^i of $\mathbb{R}$ is chosen for each wave numbers k. The partitions allows us to construct coarse-grained random variables ξ_k^r and ξ_k^i whose probability distributions are given by the empirical distribution of $\hat{u}_k$ over the partition $\Xi_k^r \times \Xi_k^i$. That is, ξ_k^r is formally a Ξ_k^r-valued random variable and provides a coarse-grained description of $\mathrm{Re}\,[\hat{u}_k]$.

More precisely, this construction assigns to each set $E \in \Xi_k$ the probability

$$P(E) \stackrel{\text{def}}{=} \int \chi_E(\hat{u}_k)\, d\mu,$$

where μ is an invariant measure of the KSE. This, in turn, allows us to compute the mutual information between the coarse-grained random variables ξ_{k_1} and ξ_{k_2}:

$$I(\xi_{k_1}, \xi_{k_2}) = H(\xi_{k_1}) + H(\xi_{k_2}) - H(\xi_{k_1}, \xi_{k_2}). \tag{5.11}$$

(See §2.3.1 for the definition of the entropy H.) The mutual information measures the amount of statistical dependence between two random variables. It is easy to check that ξ_{k_1} and ξ_{k_2} are independent if and only if $I(\xi_{k_1}, \xi_{k_2}) = 0$. Naturally, the independence of $\hat{u}_{k_1}$ and $\hat{u}_{k_2}$ implies the independence of ξ_{k_1} and ξ_{k_2}, but the converse is false. What is true is that if $I(\xi_{k_1}, \xi_{k_2}) = 0$ for all choices of the partitions $\Xi_k^{(r,i)}$ then $\hat{u}_{k_1}$ and $\hat{u}_{k_2}$ are independent.

The coarse-grained mutual information (5.11) provides a way to quantify the degree of statistical dependence between the Fourier modes of a solution of the KSE. This gives us a starting point for addressing some of the questions raised in §5.1. Because of limits on computational resources, the experiments below use only partitions Ξ_k with $|\Xi_k| = 4$. That is, the histograms used to construct the coarse-grained random variables ξ_k can use only 4 bins per (real, not complex) degree of freedom. The boundaries of the individual bins are determined adaptively by first simulating the flow for $0 \le t \le 100$ to

ϵ	0.00	5.00×10^{-3}	7.50×10^{-3}	1.00×10^{-2}
H_1	9.61×10^{-1}	9.63×10^{-1}	1.02	9.88×10^{-1}
H_2	1.05	1.02	1.00	1.00
H_3	1.12	1.00	1.00	9.98×10^{-1}
H_4	9.98×10^{-1}	1.00	9.85×10^{-1}	1.00
H_5	9.94×10^{-1}	9.96×10^{-1}	1.01	9.88×10^{-1}
H_6	1.02	9.93×10^{-1}	9.96×10^{-1}	9.19×10^{-1}
H_7	1.02	1.00	9.81×10^{-1}	1.00
H_8	1.01	1.02	9.96×10^{-1}	9.80×10^{-1}
H_9	1.08	1.00	9.97×10^{-1}	1.00
H_{10}	1.03	9.91×10^{-1}	9.93×10^{-1}	1.05
H_{11}	1.01	1.00	1.01	1.01
H_{12}	9.94×10^{-1}	1.03	1.01	9.90×10^{-1}
H_{13}	1.03	9.96×10^{-1}	9.87×10^{-1}	1.01
H_{14}	1.06	9.43×10^{-1}	1.01	1.00

Table 5.4: The coarse-grained entropy for each mode of the KSE flow.

determine rough upper and lower bounds on the real and imaginary parts of each mode. The interval is further subdivided to form a partition of $\mathbb{R}$ for each mode.

Tables 5.4 - 5.10 show the results of the computation. Statistical errors in Tables 5.5, 5.7, and 5.9 are computed by the formula

$$\text{error} \stackrel{\text{def}}{=} \sum_k \Delta\hat{p}_k \left|\phi'(\hat{p}_k)\right|, \tag{5.12}$$

$$\phi(p) = -p\log_2(p). \tag{5.13}$$

Here, $\hat{p}_k$ is the empirical probability of the kth bin in the histogram, and $\Delta\hat{p}_k$ denotes the estimated standard deviation of the empirical probability computed from the integrated autocorrelation times of the histogram. Note that the function ϕ is not differentiable at 0 or 1, so this naive error estimate becomes meaningless when numbers very near the end points of $[0, 1]$ arise. This does not happen in these calculations.

If the KSE possesses a unique invariant measure μ, then the translation symmetry of the KSE would imply that μ is also translation invariant. Thus solutions of the KSE with random initial conditions drawn from μ would form a spacetime-stationary stochastic process. This, in turn, would imply (see [Hel91]) that the modes are uncorrelated random variables. Of course, the KSE is unlikely to be uniquely ergodic except possibly at very special parameter values. Nevertheless the Fourier modes would be statistically uncorre-

ϵ	0.00	5.00×10^{-3}	7.50×10^{-3}	1.00×10^{-2}
ΔH_1	2.04×10^{-3}	7.82×10^{-4}	9.34×10^{-4}	9.25×10^{-4}
ΔH_2	1.83×10^{-3}	8.74×10^{-4}	6.82×10^{-4}	7.68×10^{-4}
ΔH_3	2.58×10^{-3}	6.29×10^{-4}	8.29×10^{-4}	8.64×10^{-4}
ΔH_4	1.43×10^{-3}	6.34×10^{-4}	6.08×10^{-4}	6.32×10^{-4}
ΔH_5	1.32×10^{-3}	6.41×10^{-4}	6.90×10^{-4}	7.06×10^{-4}
ΔH_6	1.40×10^{-3}	6.65×10^{-4}	7.07×10^{-4}	9.69×10^{-4}
ΔH_7	1.64×10^{-3}	6.01×10^{-4}	5.61×10^{-4}	7.43×10^{-4}
ΔH_8	1.77×10^{-3}	7.17×10^{-4}	5.11×10^{-4}	6.80×10^{-4}
ΔH_9	1.66×10^{-3}	6.29×10^{-4}	6.26×10^{-4}	6.20×10^{-4}
ΔH_{10}	1.44×10^{-3}	5.53×10^{-4}	5.91×10^{-4}	8.91×10^{-4}
ΔH_{11}	1.65×10^{-3}	7.10×10^{-4}	6.64×10^{-4}	7.61×10^{-4}
ΔH_{12}	1.63×10^{-3}	9.76×10^{-4}	9.52×10^{-4}	8.69×10^{-4}
ΔH_{13}	1.97×10^{-3}	6.84×10^{-4}	6.83×10^{-4}	8.86×10^{-4}
ΔH_{14}	2.08×10^{-3}	6.51×10^{-4}	9.14×10^{-4}	9.11×10^{-4}

Table 5.5: Error estimates for Table 5.4.

	ξ_1	ξ_2	ξ_3
ξ_1	9.61×10^{-1}	0.00	0.00
ξ_2	0.00	1.05	0.00
ξ_3	0.00	0.00	1.12

	ξ_4	ξ_5	ξ_6	ξ_7
ξ_1	0.00	0.00	0.00	0.00
ξ_2	0.00	0.00	0.00	0.00
ξ_3	0.00	0.00	0.00	0.00

	ξ_4	ξ_5	ξ_6	ξ_7
ξ_4	9.98×10^{-1}	0.00	0.00	0.00
ξ_5	0.00	9.94×10^{-1}	0.00	0.00
ξ_6	0.00	0.00	1.02	0.00
ξ_7	0.00	0.00	0.00	1.02

Table 5.6: The coarse-grained mutual information the modes of the unperturbed KS flow.

	$\Delta\xi_1$	$\Delta\xi_2$	$\Delta\xi_3$
$\Delta\xi_1$	1.83×10^{-3}	8.57×10^{-3}	8.83×10^{-3}
$\Delta\xi_2$	8.57×10^{-3}	2.58×10^{-3}	9.35×10^{-3}
$\Delta\xi_3$	8.83×10^{-3}	9.35×10^{-3}	1.43×10^{-3}

	$\Delta\xi_4$	$\Delta\xi_5$	$\Delta\xi_6$	$\Delta\xi_7$
$\Delta\xi_1$	6.37×10^{-3}	6.61×10^{-3}	6.68×10^{-3}	7.24×10^{-3}
$\Delta\xi_2$	7.29×10^{-3}	7.05×10^{-3}	7.55×10^{-3}	8.00×10^{-3}
$\Delta\xi_3$	7.37×10^{-3}	8.30×10^{-3}	8.13×10^{-3}	8.37×10^{-3}

	$\Delta\xi_4$	$\Delta\xi_5$	$\Delta\xi_6$	$\Delta\xi_7$
$\Delta\xi_4$	1.32×10^{-3}	5.38×10^{-3}	5.63×10^{-3}	6.11×10^{-3}
$\Delta\xi_5$	5.38×10^{-3}	1.40×10^{-3}	5.33×10^{-3}	6.27×10^{-3}
$\Delta\xi_6$	5.63×10^{-3}	5.33×10^{-3}	1.64×10^{-3}	6.27×10^{-3}
$\Delta\xi_7$	6.11×10^{-3}	6.27×10^{-3}	6.27×10^{-3}	1.77×10^{-3}

Table 5.7: Error estimates for Table 5.6.

	ξ_1	ξ_2	ξ_3
ξ_1	9.88×10^{-1}	0.00	0.00
ξ_2	0.00	1.00	0.00
ξ_3	0.00	0.00	9.98×10^{-1}

	ξ_4	ξ_5	ξ_6	ξ_7
ξ_1	0.00	0.00	0.00	0.00
ξ_2	0.00	0.00	0.00	0.00
ξ_3	0.00	0.00	0.00	0.00

	ξ_4	ξ_5	ξ_6	ξ_7
ξ_4	1.00	0.00	0.00	0.00
ξ_5	0.00	9.88×10^{-1}	0.00	0.00
ξ_6	0.00	0.00	9.19×10^{-1}	0.00
ξ_7	0.00	0.00	0.00	1.00

Table 5.8: The coarse-grained mutual information the modes of the KS flow at noise level $\epsilon = 0.01$.

	$\Delta\xi_1$	$\Delta\xi_2$	$\Delta\xi_3$
$\Delta\xi_1$	7.68×10^{-4}	3.51×10^{-3}	3.32×10^{-3}
$\Delta\xi_2$	3.51×10^{-3}	8.64×10^{-4}	3.06×10^{-3}
$\Delta\xi_3$	3.32×10^{-3}	3.06×10^{-3}	6.32×10^{-4}

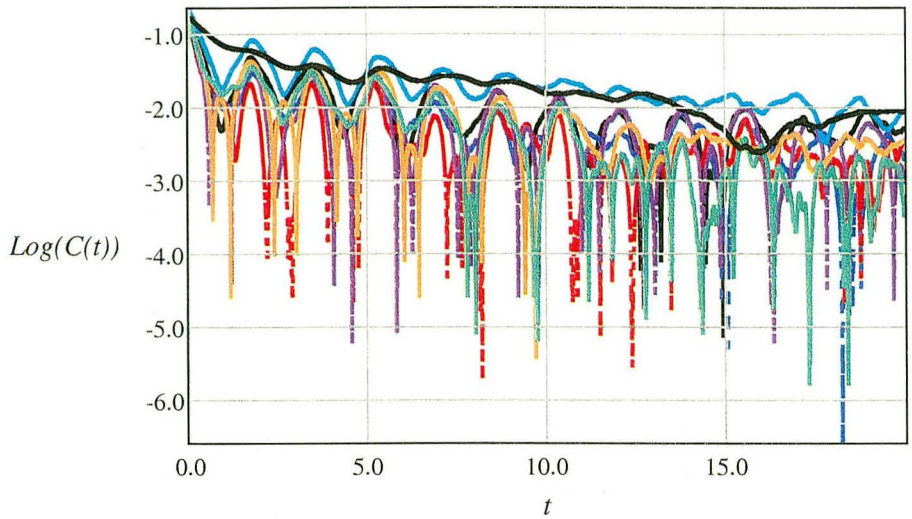
	$\Delta\xi_4$	$\Delta\xi_5$	$\Delta\xi_6$	$\Delta\xi_7$
$\Delta\xi_1$	3.14×10^{-3}	3.45×10^{-3}	3.35×10^{-3}	3.23×10^{-3}
$\Delta\xi_2$	2.86×10^{-3}	3.26×10^{-3}	2.98×10^{-3}	2.99×10^{-3}
$\Delta\xi_3$	2.66×10^{-3}	3.06×10^{-3}	2.89×10^{-3}	2.74×10^{-3}

	$\Delta\xi_4$	$\Delta\xi_5$	$\Delta\xi_6$	$\Delta\xi_7$
$\Delta\xi_4$	7.06×10^{-4}	2.77×10^{-3}	2.63×10^{-3}	2.55×10^{-3}
$\Delta\xi_5$	2.77×10^{-3}	9.69×10^{-4}	3.06×10^{-3}	2.98×10^{-3}
$\Delta\xi_6$	2.63×10^{-3}	3.06×10^{-3}	7.43×10^{-4}	2.80×10^{-3}
$\Delta\xi_7$	2.55×10^{-3}	2.98×10^{-3}	2.80×10^{-3}	6.80×10^{-4}

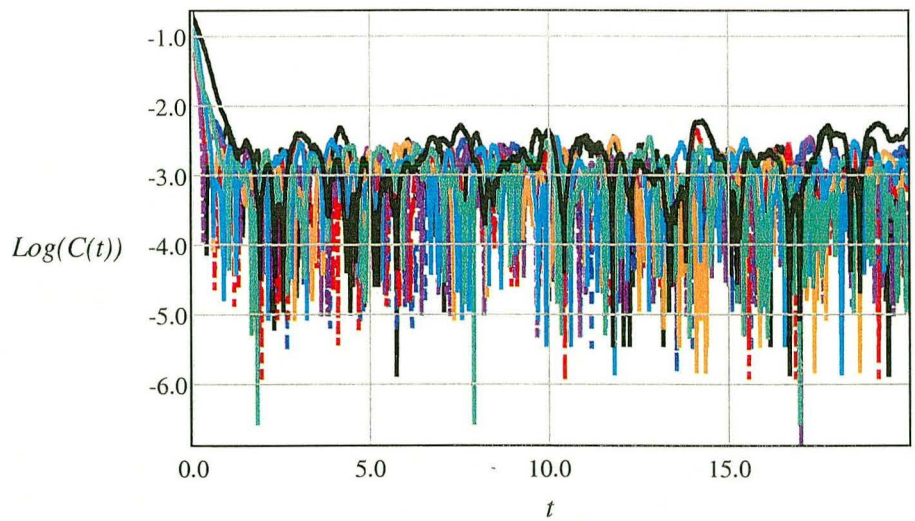
Table 5.9: Error estimates for Table 5.8.

ϵ	0.00	5.00×10^{-3}	1.00×10^{-2}
ξ_1	$-1.53, 1.80 \times 10^{-1}$	$-3.01, -3.03 \times 10^{-1}$	$-2.73, 3.26 \times 10^{-1}$
ξ_2	$-2.77, 1.05 \times 10^{-1}$	$-3.36, -8.57 \times 10^{-2}$	$-3.35, 2.33 \times 10^{-1}$
ξ_3	$-2.36, 2.90 \times 10^{-1}$	$-3.75, 1.15 \times 10^{-1}$	$-3.96, -2.30 \times 10^{-1}$
ξ_4	$-1.72, 1.46 \times 10^{-1}$	$-2.26, -7.20 \times 10^{-2}$	$-2.27, -8.31 \times 10^{-3}$
ξ_5	$-7.94 \times 10^{-1}, 5.19 \times 10^{-2}$	$-1.11, -7.07 \times 10^{-2}$	$-1.08, 1.01 \times 10^{-1}$
ξ_6	$-3.95 \times 10^{-1}, 1.67 \times 10^{-2}$	$-5.10 \times 10^{-1}, 3.39 \times 10^{-2}$	$-5.24 \times 10^{-1}, 9.34 \times 10^{-2}$
ξ_7	$-1.71 \times 10^{-1}, 4.77 \times 10^{-3}$	$-2.59 \times 10^{-1}, -2.73 \times 10^{-3}$	$-2.48 \times 10^{-1}, 1.70 \times 10^{-2}$
ξ_8	$-1.70 \times 10^{-1}, -1.13 \times 10^{-2}$	$-2.48 \times 10^{-1}, -3.02 \times 10^{-3}$	$-2.57 \times 10^{-1}, 2.10 \times 10^{-2}$
ξ_9	$-3.43 \times 10^{-1}, 1.07 \times 10^{-2}$	$-5.68 \times 10^{-1}, 1.56 \times 10^{-2}$	$-6.79 \times 10^{-1}, -9.52 \times 10^{-3}$
ξ_{10}	$-8.08 \times 10^{-1}, -2.25 \times 10^{-2}$	$-1.13, -5.75 \times 10^{-2}$	$-1.07, 4.11 \times 10^{-4}$
ξ_{11}	$-1.73, 7.64 \times 10^{-2}$	$-1.93, 1.02 \times 10^{-1}$	$-2.08, 8.01 \times 10^{-2}$
ξ_{12}	$-3.31, -1.18 \times 10^{-1}$	$-3.48, 3.39 \times 10^{-3}$	$-3.95, -3.43 \times 10^{-1}$
ξ_{13}	$-2.81, 3.16 \times 10^{-1}$	$-3.55, 2.02 \times 10^{-1}$	$-3.70, -3.62 \times 10^{-1}$
ξ_{14}	$-1.51, -5.18 \times 10^{-2}$	$-3.42, -3.51 \times 10^{-1}$	$-3.14, -9.67 \times 10^{-2}$

Table 5.10: Partitions used for this calculation.

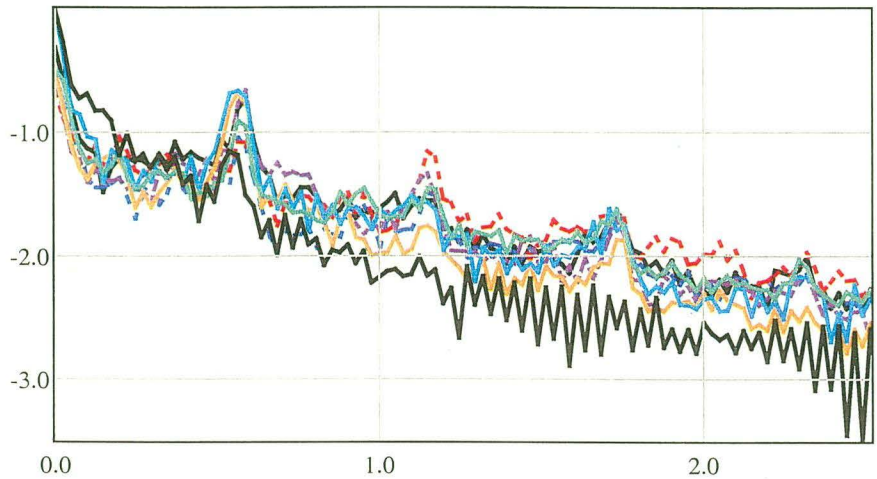


(a) unperturbed

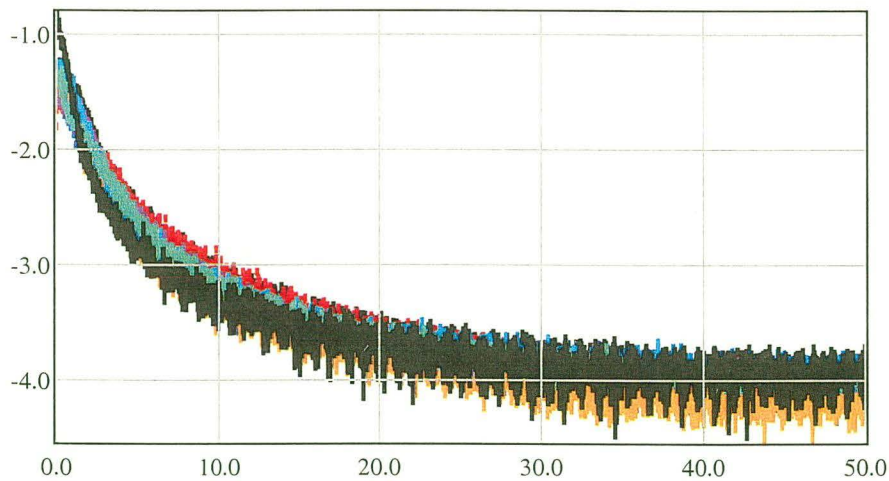


(b) $\epsilon = 0.01$

Figure 5.11: Some autocovariance functions of the histograms used in the mutual information calculation. The latter is quite noisy but does show faster initial decay.



(a) unperturbed



(b) $\epsilon = 0.01$

Figure 5.12: Spectral power densities $\log_{10}(\hat{C}(f))$ for the correlation functions shown in Figure 5.11.

lated if the KS flow were ergodic along the center manifolds defined by orbits of the group action $(T_{x_0} : x_0 \in \mathbb{T})$. The numerical results above suggest that at least one invariant measure μ may be translation-invariant. In particular, Table 5.6 suggests that the modes may be nearly independent. Incidentally, though the simple coarse-graining procedure used here is quite crude, it appears to be fairly effective at detecting the statistical dependence of jointly gaussian random variables, as one can see in simple numerical experiments.

Of course, any conclusions we may draw regarding the statistical independence of the Fourier modes can only be tentative because of the drastic nature of the coarse-graining. As for the questions raised in §5.1, a more refined version of this calculation would let us place a rough lower bound on the (coarse-grained) relative entropy of $u - \mathbb{P}_N u$ with respect to $\mathbb{P}_N u$.

Chapter 6

Future Work

Preceding chapters contain concrete numerical examples and rough analytical results on the effect of random perturbations on hyperbolic dynamics. The main result is that correlation functions can be caused to decay much faster without affecting the expectation values of observables with respect to invariant measures. In order to understand fully the algorithmic usefulness of random perturbations, there is clearly much more to be done in building firmer theoretical foundations and in performing numerical experiments.

The first question which arises from the discussions in preceding chapters is the effect of noise on the correlation spectrum of a dynamical system. As global subdivision techniques and Ulam-type approximations of the transfer operator are unlikely to yield accurate estimates of eigenvalues inside the unit circle (see [BK98]), perhaps the best approach is to directly infer the correlation spectrum from the autocovariance functions [MSW89, Man98, WMS88]. The question of how noise affects the correlation spectrum is also quite interesting from a theoretical point of view. In [BK98], it is suggested that the correlation spectrum of general hyperbolic maps may not be stable under certain types of random perturbations. Nevertheless, it may be that for sufficiently smooth random perturbations one can develop a reasonably complete perturbation theory for the correlation spectrum, perhaps along the lines of [CSP+99]. One may also obtain rough bounds using the coupling method. It is also of considerable practical and theoretical interest to understand whether any stochastic resonances may arise from the types of random perturbations considered here, particularly in the presence of intermittency.

Another way of approaching small-noise problems is via the theory of large deviations, which studies the probability of large deviations from the zero-noise limit. Freidlin

and Wentzell [FW98] have developed a beautiful theory along these lines for stochastic differential equations driven by a small amount of white noise. Their results, when combined with the scaling rule observed in §3.6, may offer another approach to the analysis of Algorithm 3.4.

Algorithmically, the discussion leading to the random perturbation algorithm suggests that it may be possible to develop analogs of Markov chain Monte Carlo methods for SRB measures. However, it is difficult to compute the local expansion coefficients J_F^u. Nevertheless, there may be specific situations where, with appropriate approximations, true Markov chain Monte Carlo methods may be implemented for SRB measures. It is particularly interesting to explore analogs of other standard algorithms of statistical physics, such as multigrid Monte Carlo and Swendsen-Wang type algorithms. It is also interesting to try to develop a numerical coupling algorithm for dynamical systems along the lines of the "coupling from the past" technique of Propp and Wilson [PW96, PW00]. Such an algorithm would provide a way to dynamically estimating the equilibration time of a long simulation.

Much of the work here concerns dissipative chaotic systems. The effects of noise on Hamiltonian (and more generally symplectic) maps and flows are much more difficult to analyze and understand. A particularly simple example which may provide a suitable starting point for numerical and analytical explorations is the standard map [SW01, You98]:

$$S_K(I,\theta) = \begin{bmatrix} I + K\sin\theta \\ \theta + I + K\sin\theta \end{bmatrix} \pmod{2\pi}. \tag{6.1}$$

Another simple example is the cat map (2.3), which is much better understood and for which considerable progress has been made on its quantum statistical mechanics in the semiclassical limit [FW03, Non03]. The effects of noise on classical and quantum Hamiltonian systems present a fascinating set of new challenges.

Finally, as explained in the Introduction, the original motivation for this work was to understand the statistical steady states of dissipative chaotic systems from an algorithmic point of view. In the much more difficult study of transient statistical states, many mathematical, algorithmic, and conceptual problems remain unresolved.

Bibliography

[Bal00] Viviane Baladi, *Positive transfer operators and decay of correlations*, World Scientific, 2000.

[BGS76] Giancarlo Benettin, Luigi Galgani, and Jean-Marie Strelcyn, *Kolmogorov entropy and numerical experiments*, Physical Review A **14** (1976), 2338–2345.

[BK98] M. Blank and G. Keller, *Random perturbations of chaotic dynamical systems: Stability of the spectrum*, Nonlinearity **11** (1998), 1351–1364.

[BKL02] M. Blank, G. Keller, and C. Liverani, *Ruelle-Perron-Frobenius spectrum for Anosov maps*, Nonlinearity **15** (2002), 1905–1973.

[BL02] Xavier Bressaud and Carlangelo Liverani, *Anosov diffeomorphisms and coupling*, Ergodic Theory and Dynamical Systems **22** (2002), 129–152.

[Bow75] Rufus Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Springer-Verlag, 1975.

[BP02] Luis Barreira and Yakov B. Pesin, *Lyapunov exponents and smooth ergodic theory*, American Mathematical Society, c2002.

[CCP97] F. Christiansen, Predrag Cvitanović, and V. Putkaradze, *Spatiotemporal chaos in terms of unstable recurrent patterns*, Nonlinearity **10** (1997), 55–70.

[CHK02] Alexandre Chorin, Ole Hald, and Raz Kupferman, *Optimal prediction with memory*, Physica D **166** (2002), 239–257.

[Cho02] Alexandre Chorin, *Conditional expectations and renormalization*, Multiscale Modeling and Simulation **1** (2002), 105–118.

[CSP+99] Predrag Cvitanović, Niels Søndergaard, Gergely Palla, Gábor Vattay, , and C. P. Dettmann, *Spectrum of stochastic evolution operators: Local matrix representation approach*, Physical Review E **60** (1999), 3936–3941.

[CT91] Thomas Cover and Joy Thomas, *Elements of information theory*, John Wiley and Sons, 1991.

[DH97] Michael Dellnitz and Andreas Hohmann, *A subdivision algorithm for the computation of unstable manifolds and global attractors*, Numerische Mathematik **75** (1997), 293–317.

[DHJR97] Michael Dellnitz, Andreas Hohmann, Oliver Junge, and Martin Rumpf, *Exploring invariant sets and invariant measures*, Chaos **7** (1997), no. 2, 221–228.

[DJ99] Michael Dellnitz and Oliver Junge, *On the approximation of complicated dynamical behavior*, SIAM Journal on Numerical Analysis **36** (1999), no. 2, 491–515 (electronic).

[Dol01] Dmitry Dolgopyat, *Lectures on u-Gibbs states*, Conference on Partially Hyperbolic Systems, Northwestern University, 2001.

[Dor99] J. Robert Dorfman, *An introduction to chaos in nonequilibrium statistical mechanics*, Cambridge University Press, 1999.

[ER85] Jean-Pierre Eckmann and David Ruelle, *Ergodic theory of chaos and strange attractors*, Reviews of Modern Physics **57** (1985), no. 3, part 1, 617–656.

[Eva98] L. Craig Evans, *Partial differential equations*, American Mathematical Society, c1998.

[Fol84] Gerald B. Folland, *Real analysis: Modern techniques and their applications*, John Wiley and Sons, 1984.

[For75] Dieter Forster, *Hydrodynamic fluctuations, broken symmetry, and correlation functions*, W. A. Benjamin, 1975.

[FW98] Mark I. Freidlin and A. D. Wentzell, *Random perturbations of dynamical systems*, Springer-Verlag, 1998.

[FW03] Albert Fannjiang and Lech Wolowski, *Noise induced dissipation in Lebesgue-measure preserving maps on d-dimensional torus*, Journal of Statistical Physics **to appear** (2003).

[Gas98] Pierre Gaspard, *Chaos, scattering and statistical mechanics*, Cambridge University Press, 1998.

[GH83] John Guckenheimer and Philip Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, Springer-Verlag, 1983.

[GPL90] Karlheinz Geist, Ulrich Parlitz, and Werner Lauterborn, *Comparison of different methods for computing Lyapunov exponents*, Progress of Theoretical Physics **83** (1990), 875–893.

[Gri78] David Griffeath, *Coupling methods for Markov processes*, Studies in Probability and Ergodic Theory (Gian-Carlo Rota, ed.), Advances in Mathematics Supplementary Studies, vol. 2, Academic Press, 1978, pp. 1–43.

[Hel91] Carl W. Helstrom, *Probability and stochastic processes for engineers*, second ed., Prentice Hall, Englewood Cliffs, New Jersey, 1991.

[Hun96] Fern Y. Hunt, *Approximating the invariant measures of randomly perturbed dissipative maps*, Journal of Mathematical Analysis and Applications **198** (1996), 534–551.

[JJK01] M. E. Johnson, M. S. Jolly, and I. G. Kevrekidis, *The Oseberg transition: Visualization of global bifurcations for the Kuramoto-Sivashinsky equation*, International Journal of Bifurcations and Chaos in Applied Science and Engineering **11** (2001), 1–18.

[KH95] Anatole B. Katok and Boris Hasselblatt, *Introduction to the modern theory of dynamical systems*, Cambridge University Press, 1995.

[Kif74] Yuri I. Kifer, *Small random perturbations of certain smooth dynamical systems*, Izv. Akad. Nauk SSSR Ser. Mat. **38** (1974), 1091–1115.

[Kif88] Yuri I. Kifer, *Random perturbations of dynamical systems*, Birkhäuser, 1988.

[Kif97] Yuri I. Kifer, *Computations in dynamical systems via random perturbations*, Discrete and Continuous Dynamical Systems **3** (1997), 457–476.

[KP92] Peter E. Kloeden and Eckhard Platen, *Numerical solution of stochastic differential equations*, Springer-Verlag, 1992.

[KPZ86] M. Kardar, G. Parisi, and Y.-C. Zhang, *Dynamic scaling of growing interfaces*, Physical Review Letters **56** (1986), 889–892.

[Lor63] Edward N. Lorenz, *Deterministic nonperiodic flow*, Journal of the Atmospheric Sciences **20** (1963), no. 2, 130–148.

[LSV99] Carlangelo Liverani, Benoît Saussol, and Sandro Vaienti, *A probabilistic approach to intermittency*, Ergodic Theory & Dynamical Systems **19** (1999), 671–685.

[Mac98] David J. C. MacKay, *Introduction to Monte Carlo methods*, Learning in Graphical Models (M. I. Jordan, ed.), NATO Science Series, Kluwer Academic Press, 1998, pp. 175–204.

[Mae99] Christian Maes, *The fluctuation theorem as a Gibbs property*, Journal of Statistical Physics **95** (1999), 367–392.

[Man98] Vladimir Mandelshtam, *Harmonic inversion of time cross-correlation functions: The optimal way to perform quantum or semiclassical dynamics calculation*, Journal of Chemical Physics **108** (1998), 9999–10007.

[Mat02] Jonathan Mattingly, *Exponential convergence for the stochastically forced Navier-Stokes equations and other partially dissipative dynamics*, Communications in Mathematical Physics **230** (2002), 421–462.

[MSW89] George Majda, Walter A. Strauss, and Musheng Wei, *Computation of exponentials in transient data*, IEEE Transactions on Antennas and Propagation **37** (1989), 1284–1290.

[Non03] Stéphane Nonnenmacher, *Spectral properties of noisy classical and quantum propagators*, Nonlinearity **to appear** (2003).

[NTW01] J. Novo, E. S. Titi, and S. Wynne, *Efficient methods using high accuracy approximate inertial manifolds*, Numerische Mathematik **87** (2001), 523–554.

[Ott00] Edward Ott, *Chaos in dynamical systems*, Cambridge University Press, 2000.

[PW96] James Propp and David Wilson, *Exact sampling with coupled Markov chains and applications to statistical mechanics*, Random Structure and Algorithms **9** (1996), 223–252.

[PW00] James Propp and David Wilson, *Coupling from the past: A user's guide*, Microsurveys in Discrete Probability (David Aldous and James Propp, eds.), American Mathematical Society, c2000, pp. 181–192.

[RW00a] L. C. G. Rogers and D. Williams, *Diffusions, Markov processes, and martingales volume 1: Foundations*, Cambridge University Press, 2000.

[RW00b] L. C. G. Rogers and D. Williams, *Diffusions, Markov processes, and martingales volume 2: Itô calculus*, Cambridge University Press, 2000.

[Sha48] Claude Elwood Shannon, *A mathematical theory of communication*, Bell System Technical Journal **27** (1948), 379–423, 623–656.

[Shu87] Michael Shub, *Global stability of dynamical systems*, Springer-Verlag, 1987.

[SKJ+92] K. Sneppen, J. Krug, M. H. Jensen, C. Jayaprakash, and T. Bohr, *Dynamic scaling and crossover analysis for the Kuramoto-Sivashinsky equation*, Physical Review A **46** (1992), R7351–R7354.

[Sok97] Alan D. Sokal, *Monte Carlo methods in statistical mechanics: Foundations and new algorithms*, Functional Integration (Cargèse, 1996), NATO Adv. Sci. Inst. Ser. B Phys., vol. 361, Plenum, 1997, pp. 131–192.

[Spi79] Michael Spivak, *A comprehensive introduction to differential geometry. Vol. I-V*, second ed., Publish or Perish Inc., Wilmington, Del., 1979.

[SS00] T. Shardlow and Andrew M. Stuart, *A perturbation theory for ergodic Markov chains and applications to numerical approximations*, SIAM Journal on Numerical Analysis **37** (2000), 1120–1137.

[Sti03] Panagiotis Stinis, *Optimal prediction for the Kuramoto-Sivashinsky equation*, Ph.D. thesis, Columbia University, 2003.

[SW92] Gerald Jay Sussman and Jack Wisdom, *Chaotic evolution of the solar system*, Science **257** (1992).

[SW01] Gerald Jay Sussman and Jack Wisdom, *Structure and interpretation of classical mechanics*, MIT Press, 2001.

[Tak84] Yōichirō Takahashi, *Entropy functional (free energy) for dynamical systems and their random perturbations*, Stochastic analysis (Katata/Kyoto, 1982), North-Holland Math. Library, vol. 32, North-Holland, Amsterdam, 1984, pp. 437–467.

[Tem88] Roger Temam, *Infinite dimensional dynamical systems in mechanics and physics*, Springer-Verlag, 1988.

[Tuc99] Warwick Tucker, *The Lorenz attractor exists*, C. R. Acad. Sci. Paris **328** (1999), 1197–1202.

[WMS88] M. Wei, G. Majda, and W. Strauss, *Numerical computation of the scattering frequencies for acoustic wave equations*, Journal of Computational Physics **75** (1988), no. 2, 345–358.

[WSSV85] Alan Wolf, Jack B. Swift, Harry L. Swinney, and John A. Vastano, *Determining Lyapunov exponents from a time series*, Physica D **16** (1985), 285–317.

[Yak81] V. Yakhot, *Large-scale properties of unstable systems governed by the Kuramoto-Sivashinsky equation*, Physical Review A **24** (1981), 642–644.

[You95] Lai-Sang Young, *Ergodic theory of differentiable dynamical systems*, Real and Complex Dynamics, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. (Hillerød, 1993), vol. 464, Kluwer Academic Publishers, 1995, pp. 293–336.

[You98] Lai-Sang Young, *Statistical properties of dynamical systems with some hyperbolicity*, Annals of Mathematics **147** (1998), 585–650.

[You02] Lai-Sang Young, *What are SRB measures, and which dynamical systems have them?*, Journal of Statistical Physics **108** (2002), 733–754.

[Zal89] S. Zaleski, *A stochastic model for the large scale dynamics of some fluctuating interfaces*, Physica D **34** (1989), 427–438.

Ernest Orlando Lawrence Berkeley National Laboratory
One Cyclotron Road | Berkeley, California 94720

Prepared for the U.S. Department of Energy under Contract No. DE-AC03-76SF00098