A Signal Processing Framework for the Analysis and Application of Chaotic Systems

Steven H. Isabelle

RLE Technical Report No. 593

May 1995

The Research Laboratory of Electronics
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS 02139-4307

This work was supported in part by the Department of the Navy, Office of the Chief of Naval Research under Grant N00014-93-1-0686 as part of the Advanced Research Projects Agency's RASSP program, and in part by the U.S. Air Force Office of Scientific Research under Grant AFOSR-91-0034-C.

A Signal Processing Framework for the Analysis and Application of Chaotic Systems

by

Steven Hamilton Isabelle

Submitted to the Department of Electrical Engineering and Computer Science
on February 6,1995, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

One contribution of chaos theory to the engineering research community is the notion that complex, erratic behavior in physical systems need not be the result of stochastic phenomena—such phenomena may result from deterministic mechanisms. This idea has been used in the analyses of several engineering systems. Perhaps more interesting are the several proposed engineering applications that take advantage of the structure of signals generated by chaotic systems. In order to take full advantage of the unique properties of chaotic signals in future applications, this structure must be well characterized. This thesis explores two aspects of this issue—the statistical structure of chaotic signals and the to linear distortion of chaotic signals.

In the first portion of the thesis, we concentrate on the time-average behavior of signals generated by chaotic systems with one state variable. Using an analogy between such signals and stationary stochastic processes, we present a framework for analyzing the statistical properties of these chaotic signals. In particular, we provide readily computable analytic expressions for a broad class of statistics of a large class of chaotic signals. We also present a technique for approximating the statistics of certain chaotic signals for which exact results are unavailable. As an example of the utility of these results, we use them to determine the power spectra of chaotic signals and to analyze a model of a switching DC-DC power converter operating in a chaotic regime.

In the second portion of the thesis, we concentrate on chaotic signals that have been linearly filtered. Such signals may arise, for example, when chaotic phenomena are measured through sensors with linear dynamics. We present results relating certain parameters of the original and distorted signals. These results have several useful consequences. For example, they are used to synthesize a family of systems that generate "chaotic white noise" and to deconvolve a chaotic signal from a filtered observation.

Thesis Supervisor: Alan V. Oppenheim
Title: Distinguished Professor of Electrical Engineering

Thesis Supervisor: Gregory W. Wornell
Title: Assistant Professor of Electrical Engineering

Acknowledgments

Sometimes things work out better that you could have planned and I certainly could not have planned for a better thesis committee. I would like to express my most sincere thanks to my thesis supervisors Alan Oppenheim and Gregory Wornell for their guidance and support of this work. They have provided outstanding examples of professional excellence as well as an outstanding environment in which to work and grow. I have learned much from both of them.

I would also like to thank my thesis reader George Verghese for many exciting and challenging discussions. In particular, George showed me that power converters can be chaotic and at the same time that MIT professors can be type B.

Thanks are also due to my graduate student colleagues. I can say without reservation that my MIT experience was much more stimulating with you than it would have been without you. In particular, I'd like to thank Kambiz Zangi, Vince McNeil, Reggie Brothers, Christ Richmond, Andrew Singer and James Preisig for providing, at different times, technical discussion and nontechnical escapes.

I have been fortunate to have a large number of supporters both nearby and far away. I'd like to thank Margo Tyler, Clarence Williams, Sr., Donald Warren, Geoffrey Gillespie, Alvin Fortier, Glen Fuller, Gregory Martin, Peter Barthe, and especially John Anderson and Greg Wornell for their friendship and advice.

I also wish to express my thanks to the AT&T Doctoral Support Program for providing financial support during my graduate studies.

Finally, I would like to thank my parents Leonard and Bettye Isabelle and my siblings Elizabeth and Leonard for their constant love and support. I am truly blessed to have you in my corner.

Contents

Chapter 1

Introduction

Chaotic models have been applied to the description of physical phenomena ranging from turbulence in fluids [1] to population dynamics [2]. Each of these applications, has provided a deterministic explanation for erratic phenomena that have traditionally been described with stochastic models . What is more interesting from an engineering standpoint is that chaotic systems and the signals they generate are potentially applicable in a wide range of engineering settings. In fact, many engineering systems are known to display chaotic behavior [10, 23, 32, 24].

Because of their special characteristics, chaotic systems also have the potential to be useful in certain engineering systems. In this context, rather than using a chaotic system to describe the behavior of an existing system, a chaotic system would be designed to perform a specific function. Some proposed applications of this type include using chaotic systems as random number generators and as waveform generators for communication systems.

Although nonlinear systems are quite common in engineering practice, the typical methods of studying them are often either too unsophisticated or too complex to be generally applicable. For instance, nonlinear systems are often approximated by linear systems, which cannot capture the inherently nonlinear phenomenon of chaos. More sophisticated techniques are often cumbersome or inaccessible. It seems plausible then that chaotic signals and systems would be more widely useful to engineers, both as models used for making inferences about systems and as components for designing systems, if a precise description of their properties and a set of convenient analytical tools were available. Given the scope

of the nonlinear systems theory, this goal may be too ambitious. However, as we shall see, when attention is restricted to a sufficiently tractable class of systems, some sharp results can be obtained.

This thesis presents a framework for the study of two of the salient aspects of systems in two useful classes of chaotic systems. The two classes of systems are quite broad and include systems that been previously used as models in several applications. The two aspects of these systems that we will describe are their statistical and geometric structure. In particular, using an analogy between chaotic signals and stochastic processes, we develop analytic and computational tools that describe can be used to study the statistics certain chaotic signals. We then use these results to analyze one example of a real system exhibits chaotic behavior—a switching power supply. Next we focus on the geometric properties of chaotic signals and specifically how they are altered by linear convolution. Linear convolution is an important operation in many signal processing applications and is hence worthy of study. Moreover, understanding convolutional effects may have implications, for example, when chaotic phenomena are measured using sensors with linear dynamics.

In the remainder of this chapter, we describe in more detail the contents of the thesis.

1.1 Statistics of Chaotic Maps

In the first portion of the thesis, we focus on signals generated by iterating from some initial condition the nonlinear difference equation

$$x[n] = f(x[n-1]), \tag{1.1}$$

where $f(\cdot)$ is a real function of a real variable that is also a chaotic system. Although such one-dimensional systems are in a sense the simplest ones exhibiting chaos, they are important for two reasons: they have many of the same properties as higher dimensional systems, and they have been proposed as models for several engineering systems.

We will see that the signals generated by systems of the form (1.1) are analogous to stationary stochastic processes and as a result, that the notion of statistics is well defined. The specific statistics can be expressed in terms of expected values taken with respect to

a certain measure or probability distribution. Furthermore, using classical results from ergodic theory, it can be shown that statistics of this type can be interpreted physically in terms of the relative frequency of different events.

. Using a statistical approach certain questions about the average operation of a nonlinear system may be addressed. Given that design specifications are often given in terms of averages, this approach is potentially applicable in a wide variety of settings. Also, as we will see, the average properties of a system often have useful physical interpretations.

We use the statistical approach in Chapter 3 and restrict attention to a class of systems of the form (1.1), called Markov maps, that are particularly amenable to analysis. These maps have certain extremely useful qualities and naturally arise several times throughout this thesis.

Among the more interesting properties of Markov maps is that many of their statistics can be determined in closed form. We derive these closed form expressions, which depend on the parameters of the Markov map, for a large class of statistics including all second and higher-order moments of the signal. These expressions provide not only a readily computable means of determining individual statistics of a Markov map, but also some insight into its global statistical structure. For example, we will see that Markov maps generate signals with *rational* power spectra which can be determined in closed form.

In Chapter 4 we consider a larger class of systems, called eventually-expanding maps, which includes Markov maps as a subset along with several maps that have been proposed as models for physical systems. In general, the statistics of an eventually-expanding map (that is not also Markov) are difficult to express in closed form. However, we show how Markov maps can be used to approximate any eventually-expanding map and its statistics arbitrarily well. Again, the results provide both a computational framework for approximating these statistics as well as some insight into their structure.

As an illustration of one application of the results of Chapter 3 and Chapter 4, we present in Chapter 5 an analysis of a switching power converter that is well described by a one-dimensional chaotic system. We derive a map that relates the currents and voltages within the converter at consecutive switching times. We then demonstrate how these sampled-data descriptions can be used to make inferences about the statistics of the converter's

continuous-time waveforms. In particular, we demonstrate that, even in a chaotic operating regime, the continuous-time waveforms have cyclostationary statistics.

1.2 Geometric Changes Due to Measurement Effects

In addition to its statistical structure, a chaotic signal often has a geometric structure that is unique to the system generating it. However, when the signal is filtered by a linear system this geometric structure may potentially be obscured. In Chapters 6 and 7, we present a method of analyzing certain aspects of the effect of linear filtering on two types of chaotic signals: those generated by one-dimensional maps, and those generated by multidimensional, invertible systems. Since, in practice, chaotic signals may potentially be measured through media that are well modeled as linear systems, this scenario is worthy of study.

Chapter 6 concentrates on signals generated by one-dimensional chaotic systems. We first examine the relation between the input and output of a filter driven by a such a signal. As we will see, when the filter has a finite length impulse response, the input and output time series are related by a scalar function, which we term the input-to-output map. We use the properties of this map to draw several conclusions concerning filtered chaos. In particular, we show that the output of certain filters may be governed by a one-dimensional map when the input is governed by a one-dimensional map. For a given chaotic system, we completely characterize this class of filters.

These results suggest several applications. In particular, using the characterization of filtered chaos, we determine an algorithm that deconvolves a chaotic signal from a filtered observation. Some preliminary results on the performance of the algorithm are also presented. As another application, we show how to construct, given an understanding of the statistics of chaotic signals, a family of chaotic systems that generate waveforms that are spectrally white.

In Chapter 7 we extend some of the results derived for one-dimensional noninvertible systems to multidimensional, invertible systems. In particular, we examine filtered signals generated by invertible, multidimensional chaotic systems and show that it is useful to inter-

pret such signals as the output of a composite system that consists of a cascade of a chaotic system, which corresponds to the input signal, and a linear system, which corresponds to the filter. We use this interpretation to relate the properties of filtered signal to those of the unfiltered signal. For example, we will show that an attractor reconstructed from the filter output is related to the attractor of the unfiltered chaotic signal by a nonlinear point transformation and discuss some of the implications of this fact. In particular, we demonstrate how one measure of the geometric structure of a chaotic signal, its so-called fractal dimension, may potentially be changed by filtering, and determine filters that do not alter the fractal The latter result has implications for the design of sensors to measure chaotic signals since accurate sensors should not alter geometric structure.

Many of the results of the thesis depend on classical results from the theory of nonlinear systems. In Chapter 2, we review some of the relevent results.

Chapter 2

Random and Deterministic Dynamics

One contribution of chaos theory to the signal processing community is the notion that complex, erratic behavior in physical systems need not be the result of stochastic phenomena—a deterministic, nonlinear mechanism may cause such behavior. As a consequence, deterministic models may potentially describe these systems more precisely than traditional stochastic models. However, the analysis of such nonlinear models typically requires mathematical tools and concepts that are not used in traditional signal processing tasks. In this chapter, we collect some of the less familiar definitions and results from the theory of nonlinear systems that are most useful in the remainder of the thesis.

We will concentrate on the properties of time series generated by the discrete time difference equation

$$\mathbf{x}[n] = \mathbf{F}(\mathbf{x}[n-1]) \tag{2.1}$$

$$y[n] = g(\mathbf{x}[n]) \tag{2.2}$$

where $\mathbf{x}[n] \in \mathbb{R}^N$ is the state vector of the dynamical system $\mathbf{F} : \mathbb{R}^N \to \mathbb{R}^N$, and $g : \mathbb{R}^N \to \mathbb{R}$ models a scalar observation of the state. A variety of physical systems evolving in discrete time steps with no input are well modeled by systems of this form [3].

The time evolution of the model is computed by iterating $\mathbf{F}$ to generate the set of state

vectors, called the *orbit* associated with $\mathbf{x}[0]$, of the form

$$\{\mathbf{x}[0], \mathbf{F}(\mathbf{x}[0]), \mathbf{F}^2(\mathbf{x}[0]), \mathbf{F}^3(\mathbf{x}[0]), \ldots\}$$

where $\mathbf{F}^n$ denotes the n^{th} iterate of $\mathbf{F}$.

Broadly speaking, the goal of a large body of nonlinear systems theory is to determine the behavior of the sequence $\mathbf{x}[n] = \mathbf{F}^n(\mathbf{x}[0])$ as n gets large. Accordingly, the questions that have been addressed in this respect concern the asymptotic or steady-state behavior of the system (2.1) starting from a given initial condition. Ideally, for each initial condition $\mathbf{x}[0]$ the state sequence could be determined and its features classified. However, this goal is too ambitious in the sense that no general analytical tools are available to provide such information. Nevertheless, some issues related to the long term behavior of the sequence $\mathbf{x}[n]$ can be addressed, at least qualitatively.

Two particularly important issues are what portion of the state space is visited by the state sequence in the long time limit and with what relative frequency subsets of the state space are visited by the state sequence. Significant information about the dynamic behavior of the nonlinear system may be provided by the resolution of these two issues. Indeed, the two approaches to the study of nonlinear systems that are most useful for the purposes of this thesis divide roughly along the lines of these two issues. Loosely speaking, the first approach examines the geometry of orbits generated by (2.1). This is the subject of topological dynamics. Some useful results from topological dynamics are presented in Section 2.1. The second approach examines the average properties of state sequences generated by (2.1). This is the subject of ergodic theory or measurable dynamics. Some results from ergodic theory are presented in Section 2.2.

2.1 Geometry of Solutions

In this section, we concentrate on ideas related to the geometric properties of solutions.

2.1.1 Steady State Behavior: Invariant Sets

As an initial step toward understanding the long term behavior of (2.1) it is convenient to single out some common types of orbits. The simplest orbit consists of only one point. A state vector, $\mathbf{x}$, that is invariant under $\mathbf{F}$—that is, one with the property that $\mathbf{F}(\mathbf{x}) = \mathbf{x}$—is called a *fixed point* of $\mathbf{F}$. Similarly, a fixed point of $\mathbf{F}^k$ is called a periodic point and satisfies $\mathbf{F}^k(\mathbf{x}) = \mathbf{x}$. The period of a periodic point is the smallest integer p such that $\mathbf{F}^p(\mathbf{x}) = \mathbf{x}$. The orbit of a periodic point $\mathbf{x}$ with period p, i.e., the p-element set $V = \{\mathbf{x}, \mathbf{F}(\mathbf{x}), \ldots, \mathbf{F}^{p-1}(\mathbf{x})\}$, is invariant under the action of $\mathbf{F}$ since each set element's image remains in the set.

Sets more complicated than periodic orbits may be invariant under the action of $\mathbf{F}$. In general, an invariant set is any set V with the property that $\mathbf{F}(x) \in V$ whenever $x \in V$. Invariant sets are of interest because they often provide insight into the the steady-state behavior of nonlinear difference equations. In particular, many nonlinear models have the property that state sequences generated from initial conditions in a certain region of the state space approach a well defined, invariant subset of the state space known as an attractor.

Although several definitions of attractor are used in the literature, they commonly define an invariant set to which all nearby orbits converge [4]. One more formal definition, presented here to illustrate the components of a definition of attractor, is a follows: We denote the image of a set V by $\mathbf{F}(V)$, i.e., $\mathbf{F}(V) = \{\mathbf{x} | \mathbf{x} = \mathbf{F}(\mathbf{v}), \mathbf{v} \in V\}$. A set $\Lambda \subset \mathbb{R}^N$ is called an *attractor* for $\mathbf{F}$ if there is a neighborhood N of Λ with closure $\overline{N}$ such that $\mathbf{F}(\overline{N})$ is contained in the interior of $\overline{N}$ and

$$\Lambda = \cap_{n \geq 0} \mathbf{F}^n(\overline{N}). \tag{2.3}$$

The operational definition presented above is made precise by this more technical definition. However, because we will not need a particular definition of attractor for what follows, the operational definition will serve our purposes.

2.1.2 Fractal Dimension of Attractors

Attractors associated with many important nonlinear systems are complex, fractal subsets of state space that have no volume (i.e. they are set of zero Lebesgue measure) and are

correspondingly difficult to precisely describe. One property commonly used to characterize these sets is their fractal dimension. In essence, fractal dimensions quantify the difference in "size" of sets of zero volume.

Although there are many definitions of fractal dimension (see e.g. [5]), we shall concentrate on only two: the Hausdorff dimension and the box dimension. The Hausdorff dimension is often the easier of the two to manipulate analytically, however, there is no way to estimate it from data. Conversely, the box dimension is more applicable to real data but has several analytically inconvenient features. In this section, we will present the definitions of these two notions of dimension.

Before defining a fractal dimension, we examine a shortcoming in one intuitive approach to volume and dimension. Consider a line segment, a plane segment, and a cube in $\mathbb{R}^3$. Although it is perhaps natural to assign these objects dimensions one, two, and three respectively, only the cube has nonzero volume when viewed as a subset of $\mathbb{R}^3$. Thus, the three dimensional Lebesgue measure is not sufficiently sensitive to distinguish between a plane and a line. Part of the usefulness of Hausdorff dimension and its associated measure is that it can distinguish between many sets of Lebesgue measure zero.

From one point of view, the Hausdorff dimension is defined as an byproduct of a generalized notion of volume. Consider the s-dimensional volume of a subset V of $\mathbb{R}^N$, with s nonnegative, defined by the following two step procedure. First, for any small $\delta > 0$ we define

$$\mathcal{H}^s_\delta(V) = \inf\left\{\sum_{i=1}^{\infty} |U_i|^s : V \subset \cup_{i=1}^{\infty} U_i, 0 < |U_i| \leq \delta\right\} \tag{2.4}$$

where $|U_i|$ is the diameter of U_i and each U_i is open. Intuitively, $\mathcal{H}^s_\delta(V)$ is an approximation to the volume of V computed summing the volumes of the small sets U_i that cover V. The term $|U_i|^s$ quantifies the intuitive notion that volumes in s dimensions should depend on length raised to the power of s. In this definition however, s need not be an integer.

As δ decreases, $\mathcal{H}^s_\delta(V)$ must increase and so approaches a limit. The s-dimensional Hausdorff measure of V is defined as this limit, i.e.,

$$\mathcal{H}^s(V) = \lim_{\delta \to 0} \mathcal{H}^s_\delta(V). \tag{2.5}$$

It can be shown that $\mathcal{H}^s(V)$ is a measure and many of its properties follow from the results of classical measure theory. However, the infimum on the right side of (2.4) cannot in general be implemented numerically. Thus, although Hausdorff measure is analytically convenient, it is not useful for the numerical study of attractors.

The Hausdorff dimension is defined by noting that the limit (2.5) exists for any subset V in $\mathbb{R}^n$ but may be 0 or ∞. Intuitively, if a set is measured in too large a dimension, it appears to have no volume; if the dimension is too small, it appears to have infinite volume. The Hausdorff dimension of V is defined as the smallest s such that the measure is 0 or equivalently, the largest s yielding an infinite measure, i.e.,

$$\dim_H V = \inf\{s : \mathcal{H}^s(V) = 0\} = \sup\{s : \mathcal{H}^s(V) = \infty\}$$

The Hausdorff dimension of a set may take a noninteger value, but for well behaved sets its value agrees with intuition. For example, a smooth surface in $\mathbb{R}^3$ has a Hausdorff dimension of two.

Although the Hausdorff dimension has analytically convienent properties due to its definition in terms of a measure, it is difficult to apply to real data. Another fractal dimension, the box dimension, is more easily applied to data. The box dimension is defined as follows. Let $N_\delta(V)$ be the smallest number of cubes of diameter δ that cover V. Then the upper and lower box dimensions of V are defined by

$$\begin{aligned} \underline{\dim}_B V &= \lim_{\delta \to 0} \inf \frac{\log(N_\delta(V))}{-\log \delta} \\ \overline{\dim}_B V &= \lim_{\delta \to 0} \sup \frac{\log(N_\delta(V))}{-\log \delta} \end{aligned}$$

If these are equal we refer to the common value as the box dimension of V. The Hausdorff and box dimensions differ in that the former uses a more general cover in its definition.

It can be shown that the box dimension of a set has many of the properties that are expected from a quantifier of dimension. For example, the box dimension of the unit cube in $\mathbb{R}^N$ can be shown to be N. The box dimension does have some undesirable properties as well. For instance, there are examples of countable sets with nonzero box dimension (see

e.g. [5, 6]).

When we examine the effects of linear convolution on chaotic signals in Chapter 7, we will need to understand how the Hausdorff and box dimensions of a set are altered when the set is nonlinearly transformed. As we will presently see, one property that is shared by both of these dimensions is that they are not altered by invertible transformations that are sufficiently smooth. We now define the specific notion of smoothness that is most useful in the sequel. A function $\mathbf{P} : \Lambda \to \mathbb{R}^{N+M}$ is Lipschitz if it satisfies

$$\|\mathbf{P}(\mathbf{x}) - \mathbf{P}(\mathbf{y})\| \leq c_1 \|\mathbf{x} - \mathbf{y}\|$$

where the norms are defined on the appropriate spaces. Intuitively, Lipschitz functions have a bounded growth rate. A function is bi-Lipschitz if it satisfies

$$c_2 \|\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{P}(\mathbf{x}) - \mathbf{P}(\mathbf{y})\| \leq c_1 \|\mathbf{x} - \mathbf{y}\|.$$

Thus, the growth rate of a bi-Lipschitz function is bounded from below as well as from above.

The significance of bi-Lipschitz transformations in relation to fractal sets is the following theorem.

Theorem 1 *If $\mathbf{P} : \Lambda \to \mathbb{R}^{N+M}$ is a bi-Lipschitz transformation, then the Hausdorff dimensions of Λ and $\mathbf{P}(\Lambda)$ are equal.*

Thus, bi-Lipschitz transformations do not change fractal dimension. It follows that smooth, smoothly invertible coordinate changes cannot change the dimension of a set.

2.1.3 Relations Between System Properties and Time Series Properties

We now turn to properties of time series generated by scalar observations of the state of a nonlinear system. It can be shown that, in this context, the class of nonlinear systems is much too large to work with as a whole; virtually no precise statement can be made about the entire class. In particular, with no restrictions on $\mathbf{F}$ and g, the class of models of the form (2.1) and (2.2) can generate any bounded time series. This section presents two

examples designed to illustrate this point.

If the dimension of the state space is infinite, then a there is a system $\mathbf{F}$ and observation function g that can produce any finite, causal time series. Let the state space, $\mathbf{X}$, be the space of bounded, right sided sequences. If $\mathbf{x}$ is such a sequence, we denote its i-th component by $(\mathbf{x})_i$. In other words, each element, $\mathbf{x} \in \mathbf{X}$ is of the form

$$\mathbf{x} = (\mathbf{x})_0, (\mathbf{x})_1, (\mathbf{x})_2, \ldots.$$

Now suppose $\mathbf{F}$ is the shift operator and g observes the 0-th component of a state vector, i.e.

$$\begin{aligned} (\mathbf{F}(\mathbf{x}))_i &= (\mathbf{x})_{i+1}, \quad i = 0, 1, 2, \ldots \\ g(\mathbf{x}) &= (\mathbf{x})_0. \end{aligned}$$

With these definitions, $\mathbf{F}$ and maps the state space to itself. This system can generate any bounded time series by a proper choice of initial condition. In essence, by choosing the initial condition, the entire time series is chosen. The implication is that simply specifying that a signal was generated deterministically does not significantly limit the possible range of time series behavior.

One might guess that the complex behavior illustrated in the previous example is related to the dimension of the state space. On the contrary, complex behavior is possible even in one dimensional state spaces as the following example shows.

A system that is equivalent to the shift of the previous example can be constructed on a one-dimensional state space. According to a classical result from real analysis, the set of right-sided sequences of real numbers has the same cardinality as the unit interval $I = [0, 1]$ (see, for example, [7]). This means that the points of the interval I can be put into one-to-one correspondence with the points of the space $\mathbf{X}$. Let $\phi : I \to \mathbf{X}$ denote this correspondence. Since ϕ is one-to-one it is invertible. As a consequence, each state sequence generated by $\mathbf{F}$ corresponds to a sequence in I. This corresponding sequence can

be determined by the following map:

$$x[n] = \phi^{-1}(\mathbf{F}(\phi(x[n-1]))) \triangleq f(x[n-1]),$$

where f is a scalar function mapping I to itself. The corresponding output time series is generated by the observation equation

$$y[n] = g(\phi(x)) \triangleq h(x).$$

Thus, the one-dimensional map f along with the observation function h generate the same collection of time series that the infinite dimensional map $\mathbf{F}$ and observation function g do. A similar construction could establish this type of equivalence between any two systems operating on subsets of finite dimensional spaces.

The construction presented in the previous example is not meant to suggest a practical method for designing signal generators. Indeed, the correspondence function ϕ is exceedingly complicated and could never be implemented by any finite precision machine. However, the example does illustrate the point that without narrowing the scope of inquiry, virtually no precise statements can be made about the behavior of nonlinear systems. Since both systems produce the same time series, it follows that joint specifications on the state space, the map and the observation function are necessary to constrain the possibilities. In general, these specifications consist of some restriction on the dimension of the state space, and smoothness conditions on $\mathbf{F}$ and g.

In the remainder of this thesis, we will concentrate on two classes of systems. The first class is that of piecewise continous one-dimensional maps. Such maps appear in models of many engineering systems and physical phenomena [8, 9, 10].The second class consists of differentiable maps from $\mathbb{R}^N$ to itself with a differentiable inverses, i.e., diffeomorphisms of $\mathbb{R}^N$. Such models appear for example in the study of the time sampled evolution of continuous time differential equations [11].

2.1.4 Observability of Nonlinear Systems

In many situations, the entire state of the map $\mathbf{F}$ cannot be measured, either because it is inaccessible or because a full measurement would be prohibitively complex. It is natural to ask then what information about the entire state can be inferred from a scalar measurement like (2.2). For instance, we may be interested in whether the geometry of a system's attractor can be determined from a time series observation. We will see in this section that, loosely speaking, when $\mathbf{F}$ is a diffeomorphism and g is a differentiable observation function, an image of the system's attractor can be recovered from a time series observation to within a smooth, nonlinear coordinate change. An important consequence of this result is that properties of an attractor that are invariant to smooth nonlinear coordinate changes may be measured from a time series observation. We will present an application of this result in Chapter 7 where we explore the effects of linear time invariant filtering on chaotic time series.

One approach to inferring properties of the vector state sequence $\mathbf{x}[n]$ from the scalar observation time series $y[n]$ is to somehow transform $y[n]$ into a sequence of vectors. A widely used transformation of this type, called time delay reconstruction, maps $y[n]$ into the sequence of vectors defined by

$$\hat{\mathbf{x}}[n] = \begin{bmatrix} y[n] \\ y[n-1] \\ \vdots \\ y[n-L+1] \end{bmatrix} \tag{2.6}$$

where L is an integer specifying the dimension of $\hat{\mathbf{x}}[n]$. The vector $\hat{\mathbf{x}}[n]$ is referred to as the reconstructed vector.

A relation between $\mathbf{x}[n]$ and $\hat{\mathbf{x}}[n]$ can be derived by substituting (2.1) and (2.2) into

(2.6) to obtain

$$\hat{\mathbf{x}}[n] = \begin{bmatrix} g(\mathbf{x}[n]) \\ g(\mathbf{F}^{-1}(\mathbf{x}[n])) \\ \vdots \\ g(\mathbf{F}^{-L+1}(\mathbf{x}[n])) \end{bmatrix} \triangleq \Phi_{(\mathbf{F},g,L)}(\mathbf{x}[n]). \tag{2.7}$$

We refer to the map $\Phi_{(\mathbf{F},g,L)}(\mathbf{x}[n])$ as the reconstruction transformation. Thus, the reconstructed vector can be written as a nonlinear transformation of a state vector. As (2.7) makes explicit, the properties of the reconstruction transformation depend on L, g and $\mathbf{F}$. What is remarkable about (2.7) is that, in general, for sufficiently large L and sufficiently smooth $\mathbf{F}$ and g, the reconstruction transformation is smooth and invertible with a smooth inverse. Thus, under certain conditions, the reconstruction transformation is a smooth coordinate change.

The next theorem makes the last statement precise using the topological notion of a generic property, which is defined as follows. Denote by D the set of all pairs $(\mathbf{F}, g)$, where $\mathbf{F}$ is a twice differentiable diffeomorphism mapping a compact subset X of $\mathbb{R}^N$ to itself and g is twice differentiable. Open sets in D may be defined through an appropriate norm. A property is *generic* for systems in D if it is possessed by systems in some subset of D that is open and dense, or a countable intersection of open dense sets. Thus, generic properties are typical in the sense that they are possessed by a large subset of systems in D. Furthermore, if a system $\mathbf{F} \in D$ does not possess a particular generic property, another system arbitrarily close to $\mathbf{F}$ does.

Theorem 2 (Takens [1]) *Let X be a compact subset of $\mathbb{R}^N$. For pairs $(\mathbf{F}, g)$, with $\mathbf{F} : X \to X$ a twice differentiable diffeomorphism and $g : X \to \mathbb{R}$ a twice differentiable function, it is a generic property that the map $\Phi_{(\mathbf{F},g,2N+1)} : X \to \mathbb{R}^{2N+1}$. is a diffeomorphism of X in $\mathbb{R}^{2N+1}$.*

Thus, a large class of systems have the property that their state sequences may be observed (to within a smooth invertible coordinate change) from a smooth scalar measurement of the state. This has a major implication for chaotic time series analysis since some properties of the state vectors are invariant to smooth coordinate changes and hence

may be measured from time series observations. See [1],[12],[13], and [6] for more detailed discussions of this issue.

2.1.5 Multiple Operating Regimes of Nonlinear Systems

Certain a nonlinear systems may operate in one of several qualitatively different regimes, depending on its initial condition,. The simplest example of such behavior is a globally stable system with two distinct attracting fixed points. As this system evolves, one collection of initial conditions approaches one fixed point while another collection of initial conditions approaches the other fixed point. In general, a system with multiple regimes of operation can be analyzed using a "divide and conquer" strategy of examining the system separately in each qualitatively different regime. Returning to the previous example, the system with two fixed points has the property that, with respect to steady state behavior, it may be decomposed into two separate systems, each with a single stable fixed point. These two systems, obtained by restricting the definition of the original system to appropriate regions of the state space, represent two different steady state operating regimes. In essence, this procedure breaks the attractor (the two fixed points) into pieces which cannot be further decomposed.

We will see below that the notion of breaking an attractor in to its smallest pieces provides a helpful way of interpreting certain time series generated by nonlinear systems. Of particular interest in this context are attractors which cannot be decomposed into smaller pieces. A system with an indecomposable attractor is called topologically transitive. More precisely, a map $\mathbf{F} : X \rightarrow X$ is said to be topologically transitive if for any pair of open sets $U, V \in X$ there exists an $n > 0$ such that $\mathbf{F}^n(U) \cap V \neq \emptyset$. It can be shown that, as a consequence of this definition, the attractor of a topologically transitive system cannot be decomposed [4].

Another important property of topologically transitive systems, which has implications for time series analysis, is that their entire attractor geometry can be recovered from time series observations. This result follows from the fact that these systems possess dense orbits [3, 4]. In particular, suppose $\mathbf{F}$ has an attractor X and a state sequence $x[n]$ that is dense in X and a time series observation of the form (2.2) is available. By the embedding theorem,

the vectors $\hat{\mathbf{x}}[n]$ generated from $y[n]$ by time delay reconstruction are in general related to $x[n]$ through a nonlinear coordinate change. Thus, since the set $\{\mathbf{x}[n]\}$ is dense in X and because the the reconstruction transformation is smooth, the geometry of the attractor can be observed from time series given sufficient data.

2.1.6 A Definition of Chaos

Because of the broad scope of nonlinear systems research, no standard definition of chaos exists in the literature. However, many notions of chaos try to capture precisely the property that two state sequences generated from arbitrarily close initial conditions will eventually diverge to very different state values. This notion, called sensitive dependence on initial conditions is defined as follows. A map $\mathbf{F} : X \to X$ has sensitive dependence on initial conditions if there exists a $\delta > 0$ such that for any $x \in X$ and any neighborhood U of x, there exists a $y \in U$ and $n \geq 0$ such that $\|\mathbf{F}^n(x) - \mathbf{F}^n(y)\| > \delta$.

The definition of sensitive dependence on initial conditions does not say that *all* points in a neighborhood of $\mathbf{x}$ generate state sequences that diverge. Some points may, in fact, generate state sequences that converge to the state sequence generated by $\mathbf{x}$. It is this combination of divergence and convergence that results in the complicated structure of the some chaotic attractors.

In the remainder of the thesis, we will concentrate on systems with a strong form of sensitive dependence on initial conditions. These systems have the property that whenever two nearby state sequences diverge, the distance between them increases exponentially, at least in the short term. In Section 2.2.4 we shall describe a commonly used quantifier of this exponential divergence called Lyapunov exponents.

2.2 Statistics of Solutions

Distinct from the geometric approach to nonlinear system analysis, is an approach based on the system's average properties. We will see that many nonlinear systems can be viewed in the same framework as stochastic processes using the tools of classical ergodic theory. In this context, we will refer to the *statistical* properties of chaotic signals even though the

entire signal generation method described by (2.1) and (2.2) is deterministic. An important point, which will be revisited below, is that a statistical approach to the analysis of nonlinear systems need not depend on the assumption of a random initial condition, although this assumption may sometimes prove convenient.

As with stationary stochastic processes, time series derived from state sequences of many chaotic systems have the property that their time averages are equivalent to ensemble average taken with respect to a distribution function. Through this equivalence, the average properties nonlinear system can, at least in principle, be computed precisely without relying on empirical time averages. Thus, certain chaotic systems may be analyzed without resorting to exhaustive simulations. This is a great advantage, for example, for the many chaotic system that are difficult to simulate on digital computers because of finite precision effects.

The mixing of deterministic and stochastic notions is not as odd as it may first appear; certain natural questions concerning deterministic systems, such as what fraction of time a state sequence spends in any given region of the state space, have decidedly probabilistic connotations. Ergodic theory, which concerns among other things the statistical properties of certain nonlinear systems, is the natural framework within which to answer such questions. The key property of the nonlinear system $\mathbf{F}$ of (2.1) that allows ergodic theory to be applied is that there exists a probability distribution μ that is invariant to $\mathbf{F}$; that is, at each time n, the state vector $\mathbf{x}[n]$ is distributed according to μ whenever the initial condition $\mathbf{x}[0]$ is distributed according to μ. Such systems are called measure preserving transformations and μ is called an invariant measure.

Ergodic theory is flexible enough to be applied to both deterministic and stochastic systems primarily because is was developed using highly abstract models for the systems involved. Accordingly, many of its results are qualitative in nature. However, we will see in subsequent chapters that these qualitative results form the basis of a quantitative analysis of a large class of chaotic systems.

The statistical approach models the state space of a nonlinear system as a *measure space*, i.e. a triple $(\Omega, \mathcal{S}, \mu)$ consisting of a set Ω, a σ-ring $\mathcal{S}$ and a probability distribution or *measure* μ defined on $\mathcal{S}$. This is a more detailed model than that used in Section 2.1 in

that the state space is assumed to consist not only of a set of states (Ω), but also a measure (μ) describing the distribution of states, and a collection of measurable sets ($\mathcal{S}$) for which the measure is defined. The measure of a measurable set A is denoted $\mu(A)$.

Because we will always work with measures that are normalized to 1 in what follows (i.e. , $\mu(\Omega) = 1$) the triple $(\Omega, \mathcal{S}, \mu)$ is a probability space. Thus, μ gives the size of any set in $\mathcal{S}$ as a fraction of the size of Ω. Note that this the designation of μ as a probability measure does not mean that the states are "random". However, we will often use the probabilistic terms *event* and *probability* to refer to a particular measurable set and its measure.

We will restrict our attention to maps that have certain properties that allow them to be conveniently analyzed. In particular, the tools of ergodic theory naturally apply to systems (2.1) and (2.2) with the property that sets that map to measurable sets in Ω or $\mathbb{R}$ are themselves measurable. Such $\mathbf{F}$ and g, called measurable transformations, satisfy $\mathbf{F}^{-1}(A) \in \mathcal{S}$ whenever $A \in \mathcal{S}$ and $g^{-1}(E) \in \mathcal{S}$ whenever E is a Borel subset of the real line. It is also useful to restrict attention to systems that that do not map sets of nonzero probability map onto sets of zero probability. Such a map $\mathbf{F}$, so-called a nonsingular transformation, has the property that $\mu(A) = 0$ implies that $\mu(\mathbf{F}^{-1}(A)) = 0$. Finally, to ensure that expected values are well defined in terms of integrals, we often require that g be integrable with respect to μ, that is, that g must satisfy

$$\|g\| \stackrel{\Delta}{=} \int_{\Omega} |g| \, d\mu < \infty.$$

Such a g is said to be μ-integrable with norm $\|g\|$. The collection of such integrable functions is denoted $L^1(\Omega, \mu)$ or simply L^1 hen the measure is obvious from context.

2.2.1 Invariant Measures and the Frobenius–Perron Operator

A central issue in the statistical approach to nonlinear system analysis concerns how a distribution of states evolves as the map $\mathbf{F}$ is repeatedly applied. As mentioned previously, certain distributions have the property that they are invariant to the application of $\mathbf{F}$. Because these invariant distributions are important for what follows, we define them more precisely and describe some of their properties.

Suppose that the measure μ describes a distribution of initial conditions $\mathbf{x}[0]$ and that μ_1 describes the distribution of $\mathbf{x}[1]$, i.e., the points of the form $\mathbf{F}(\mathbf{x}[0])$. Then μ_1 is defined by the relationship

$$\mu_1(A) = \mu(\mathbf{F}^{-1}(A)).$$

for all measurable sets A. More generally, μ_n, the distribution of $\mathbf{x}[n]$, is defined by

$$\mu_n(A) = \mu(\mathbf{F}^{-n}(A)). \tag{2.8}$$

An invariant measure is one satisfying $\mu_n = \mu$ for all $n \geq 0$. Equation (2.8) implies that when μ is invariant with respect to $\mathbf{F}$, it satisfies the relationship,

$$\mu(A) = \mu(\mathbf{F}^{-1}(A)). \tag{2.9}$$

Although all of the systems that we will study in this thesis have invariant measures, it is important to emphasize that, in general, this need not be the case. Examples of systems with no invariant measure can be found in [14].

The evolution of measures described by (2.8) has a probabilistic interpretation. Suppose for the moment that the initial condition $\mathbf{x}[0]$ of the map $\mathbf{F}$ is random and described by the probability distribution μ. Then $\mathbf{x}[n] = \mathbf{F}^n(x[0])$ is random with probability distribution μ_n. When $\mu_n = \mu$ for all n the state vector has the same probability law for all time. We will presently see that this probabilistic analogy allows certain nonlinear systems to be viewed as stochastic processes.

When μ has a density p, the measure of a set is defined by

$$\mu(A) = \int_A p(x)\, dx$$

for all measurable sets A. Since $\mathbf{F}$ is nonsingular, μ_1 also has a density, which we denote by p_1. According to (2.8), the two densities satisfy

$$\int_A p_1(x)\, dx = \int_{\mathbf{F}^{-1}(A)} p(x)\, dx. \tag{2.10}$$

Because (2.10) holds for any probability measure μ defined by a density, it can be shown [15] that (2.10) defines a unique *linear* operator which we denote by $P_{\mathbf{F}}$. The operator $P_{\mathbf{F}}$ maps integrable functions to integrable functions and satisfies

$$\int_A P_{\mathbf{F}} p(x)\, dx = \int_{\mathbf{F}^{-1}(A)} p(x)\, dx. \tag{2.11}$$

With this definition of $P_{\mathbf{F}}$, the initial and time one densities are related by

$$P_{\mathbf{F}} p = p_1.$$

The operator defined in (2.11), commonly called the Frobenius-Perron (FP) operator, compactly describes how densities of states evolve under the application of $\mathbf{F}$. Again, speaking probabilistically, when $\mathbf{x}[0]$ is random and governed by a probability density p, then $\mathbf{x}[1] = \mathbf{F}(\mathbf{x}[0])$ is random and governed by the probability density $P_{\mathbf{F}} p$.

An invariant density, i.e. one associated with a measure satisfying (2.9), satisfies the equation

$$P_{\mathbf{F}} p = p.$$

This is a fixed point equation in the infinite dimensional space of integrable functions and is difficult to solve in general. However, we will see in Chapters 3 that for certain maps the invariant density can be determined exactly. We will also see how the invariant density forms the basis for a numerical approach to the statistical analysis of chaotic systems.

2.2.2 Relating Chaotic behavior and Stochastic processes

In this section, we will explore the connection between nonlinear systems and stochastic processes and in particular, the manner in which their associated time series are generated.

A discrete time stochastic process is an indexed set of random variables. More specifically, suppose $(\Omega, \mathcal{S}, \mu)$ is a measure space. A sample sequence of the stochastic process is determined by choosing a point of $\omega \in \Omega$ randomly according to the probability distribution

μ. Associated with each ω is a sequence

$$f_\omega[0], f_\omega[1], f_\omega[2], \ldots . \tag{2.12}$$

The stochastic process is comprised of sequences of the form (2.12). For fixed k, $f_\omega[k]$ may be viewed as a random variable. From this point of view, properties of the stochastic such as stationarity, independence and correlation are relation between the random variables $f_\omega[k]$. For example, one implication of stationarity is that $E\{f_\omega[k]\}$ is independent of k.

Once the random experiment is performed, ω is fixed and the sequence (2.12) is determined for all k. Thus, the properties of any particular sequence can be studied outside of any probabilistic framework. One subset of the theory of stochastic processes seeks to relate the properties of individual sequences—typically a time average—to those of the ensemble—typically an average over Ω. As we shall see in the next section, such an approach can be applied to chaotic signals as well.

A chaotic sequence with random initial condition is a stochastic process whenever the nonlinear system is a measurable transformation. Each member of the ensemble of sequences generated by (2.1) and (2.2) is of the form

$$g(\mathbf{x}[0]), g(\mathbf{F}(\mathbf{x}[0])), g(\mathbf{F}^2(\mathbf{x}[0])), \ldots . \tag{2.13}$$

A relation between the sequences of (2.12) and (2.13) is determined by the correspondences $\omega = \mathbf{x}[0]$ and

$$f_\omega[k] = g(\mathbf{F}^k(\mathbf{x}[0])).$$

Again, as above, once the initial condition is chosen, the chaotic sequence is determined for all time.

The primary feature distinguishing chaotic sequences generated by smooth nonlinear systems from more general stochastic processes is that for different values of k, the functions $g \circ \mathbf{F}^k$ of (2.13) are necessarily closely related to one another because of the smoothness of $\mathbf{F}$ and g. On the other hand, more general measure preserving transformations need have no structure other than measurability and nonsingularity.

It can be shown that if $\mathbf{F}$ is a measure preserving transformation and the initial condition $\mathbf{x}[0]$ is distributed according to the invariant measure, the resulting process, (2.13), is stationary. In fact, it can be shown that every stationary process can be modeled by an observation of a measure preserving transformation of the form (2.13) [16]. Again, the general stationary stochastic processes and ensembles of chaotic sequences differ primarily because of the smoothness of $\mathbf{F}$ and g.

2.2.3 Relating Time and Ensemble Averages

Given the close relationship between chaotic time series and stochastic processes, it is not surprising that similar analysis techniques apply to both areas. In particular, time averages of chaotic signals can be related to ensemble averages taken with respect to an invariant measure.

Before proceeding to relate time averages to ensemble averages, the existence of time averages must be established. The existence of the limit

$$\bar{g}(\mathbf{x}[0]) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g(\mathbf{F}^k(\mathbf{x}[0])). \tag{2.14}$$

for arbitrary measurable functions g is a classical problem of ergodic theory. The Birkhoff ergodic theorem asserts that the average (2.14) converges when $\mathbf{F}$ is measure preserving.

Theorem 3 (see e.g. [16]) *Let $(\Omega, \mathcal{S}, \mu)$ be a probability space, $\mathbf{F} : \Omega \to \Omega$ a measure preserving transformation, and $g \in L^1(\Omega, \mu)$. Then*

1. $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g(\mathbf{F}^k(\mathbf{x})) = \bar{g}(\mathbf{x})$ *exists almost everywhere;*
2. $\bar{g}(\mathbf{F}(\mathbf{x})) = \bar{g}(\mathbf{x})$ *almost everywhere;*
3. $\bar{g}(\mathbf{x}) \in L^1$, *and* $\|\bar{g}\| \leq \|g\|$;
4. *if* $A \in \mathcal{S}$ *with* $\mathbf{F}^{-1}(A) = A$, *then* $\int_A g \, d\mu = \int_A \bar{g} \, d\mu$;
5. $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g(\mathbf{F}^k(\mathbf{x})) \to \bar{g}(\mathbf{x})$ *in* L^1.

Note that the time average $\bar{g}(\mathbf{x})$ may depend on the initial condition of $\mathbf{F}$. When $\bar{g}(\mathbf{x})$ is independent of $\mathbf{x}$, i.e., when its value is constant, $\mathbf{F}$ is referred to as an ergodic transformation.

We will define ergodicity more formally below.

As discussed in Section 2.1.5, a nonlinear system may potentially operate in one of several different regimes depending on its initial condition. We saw that this could occur when the attractor of the system is decomposable. The notion of indecomposable attractors lead to the definition of topological transitivity. The analogous notion in ergodic theory is metric transitivity or ergodicity. Ergodic transformations are those for which the invariant measure is indecomposable. The property of ergodicity also isolates the important class of systems for which the time averages (2.14) are independent of initial condition.

In order to formally define ergodicity, we first require a definition of invariant set which takes into account the measure space setting of the current discussion. A set $B \in \mathcal{S}$ is called invariant if $\mu(f^{-1}(B)\Delta B) = 0$ where Δ is the symmetric difference $A\Delta B = (A-B)\cap(B-A)$. Note that this definition of invariant set differs from that given in Section 2.1.1. A measure preserving transformation is ergodic or metrically transitive if the only invariant sets have measure either 0 or 1. A consequence of the definition of ergodicity is that an invariant measure, μ, cannot be decomposed into smaller nontrivial measures. In other words, if $\mathbf{F}$ is ergodic and $\mu = \mu_a/2 + \mu_b/2$ with μ_a and μ_b defined on disjoint invariant sets, then either $\mu_a = 0$ or $\mu_b = 0$.

When $\mathbf{F}$ is ergodic, it can be shown that the time average $\bar{g}(\mathbf{x})$ is constant almost everywhere, i.e. it is almost everywhere independent of the initial condition. It follows from the Birkhoff ergodic theorem that for ergodic transformations,

$$\bar{g} = \int_\Omega g \, d\mu. \tag{2.15}$$

Thus time and ensemble averages are equal only for ergodic systems.

Birkhoff's theorem implies a special role for integrals with respect to the invariant density of a system as in (2.15). We will often refer to such integrals as expected values. Using this terminology, it is now possible to refer to many statistics of deterministic systems in familiar terms. For example, the autocorrelation sequence of a chaotic signal is of the form

$$R_{xx}[k] = E\{x[n]x[n+k]\}$$

where E denotes expectation. Similarly, a chaotic signal's power spectrum is the Fourier transform of its autocorrelation sequence.

2.2.4 Quantifying Sensitivity to Initial Conditions

In addition to arithmetic averages of the form (2.14), geometric average also arise occasionally in the study of stochastic processes. In the case of matrix functions of a stochastic process, these averages are of the form

$$\Lambda(\mathbf{x}[0]) = \lim_{n\to\infty} (\mathbf{M}(n, \mathbf{x}[0])^T \mathbf{M}(n, \mathbf{x}[0]))^{\frac{1}{2n}}, \tag{2.16}$$

where

$$\mathbf{M}(n, \mathbf{x}[0]) = \mathbf{M}(\mathbf{x}[n-1])\mathbf{M}(\mathbf{x}[n-2]) \cdots \mathbf{M}(\mathbf{x}[0]).$$

The Multiplicative Ergodic Theorem, due to Oseledec [17], asserts that under some mild restrictions, geometric averages of the form (2.16) exist.

The Multiplicative Ergodic Theorem (MET) is phrased in terms of general measure preserving transformations. However, when $\mathbf{F}$ is also differentiable, the MET has an interpretation in terms of quantifying sensitivity to initial conditions. Specifically, consider a nominal state state sequence generated from an initial condition $\mathbf{x}[0]$ and a state sequence generated from the perturbed initial condition $\mathbf{x}[0]+\Delta$. We wish to quantify how these two state sequences diverge from or converge to one another. To this end, define the difference between state values

$$d(n, \Delta, \mathbf{x}[0]) = \mathbf{F}^n(\mathbf{x}[0] + \Delta) - \mathbf{F}^n(\mathbf{x}[0]). \tag{2.17}$$

By linearizing $d(n, \Delta, \mathbf{x}[0])$ in its second argument, first order information concerning its may be obtained for the case of small perturbations Δ . Differentiating (2.17), we obtain that the linearized difference is of the form

$$\tilde{d}(n, \Delta, \mathbf{x}[0]) = D\mathbf{F}^n(\mathbf{x}[0])\Delta, \tag{2.18}$$

where $D\mathbf{F}$ is the Jacobian of $\mathbf{F}$. Using the chain rule , (2.18) can be written in the form

$$\tilde{d}(k, \Delta, \mathbf{x}[0]) = D\mathbf{F}(\mathbf{F}^{n-1}(\mathbf{x}[0]))D\mathbf{F}(\mathbf{F}^{n-2}(\mathbf{x}[0])) \cdots D\mathbf{F}(\mathbf{x}[0])\Delta. \tag{2.19}$$

Setting $\mathbf{M}(\mathbf{x}) = D\mathbf{F}(\mathbf{x})$, we see that the conclusions of the the MET hold for the product (2.19).

In particular, the MET asserts that there are numbers $\alpha_i(\mathbf{x})$, commonly called the Lyapunov exponents of $\mathbf{F}$, such that for large n

$$\|\tilde{d}(n, \mathbf{x}[0], \Delta)\| = \|D\mathbf{F}^n(\mathbf{x}[0])\Delta\| \approx e^{\alpha_i(\mathbf{x})n} \tag{2.20}$$

It follows from (2.20) that when $\alpha_1 > 0$, the norm of the linearized difference grows exponentially for large n. The implication is that state sequences generated from nearly identical initial conditions may diverge exponentially.

By definition, the expression for the linearized difference $\tilde{d}(n, \mathbf{x}[0], \Delta)$ holds only for small Δ, so the exponential divergence of state sequences will occur only until the linear approximation ceases to be accurate. It should not be surprising that systems with positive Lyapunov exponents have sensitive dependence on to initial conditions [3]. The Lyapunov exponents provide a quantitative measure of this sensitivity by providing rates of divergence.

Because they address very different types of questions, the geometric approach and the statistical approach to nonlinear systems analysis are complementary. In subsequent sections we shall use both as a framework for a quantitative study of the statistical behavior of a specific class of chaotic systems as well as an investigation into the effects of linear distortions on chaotic signals.

Chapter 3

Statistical Properties of Markov Maps

In this chapter and the next, we consider signals generated by what are arguably the simplest chaotic systems – those with only one state variable. Such signals are generated by the recursion

$$x[n] = f(x[n-1]), \tag{3.1}$$

where f maps scalars to scalars.

As demonstrated by the examples of Chapter 2, restricting attention one-dimensional systems may be no restriction at all; one-dimensional systems can generate the same signals that higher dimensional systems can. A less general approach allows some sharp statements to be made. In particular, we shall concentrate on a class of one-dimensional systems that are piecewise-smooth. The detailed definition of the class will be deferred until a subsequent section. For now, it suffices to say that the class includes systems that have been proposed as models for a variety of engineered systems and physical phenomena.

This chapter and Chapter 4 have the specific goal of determining the statistics of certain one-dimensional systems. We consider a class of statistics broad enough to include those of interest to signal processors – for example the autocorrelation and all higher order moments of the process generated by the system. The chapters differ in the class of systems considered. In this chapter, we consider a class of piecewise-linear systems that can be analyzed

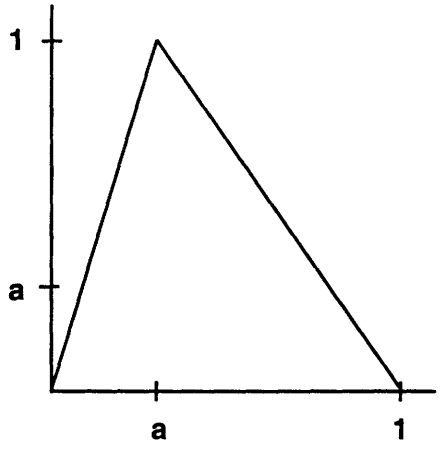


Figure 3-1: The Tent Map of Eq. (3.2)

exactly. In Chapter 4, we consider a much larger class of piecewise-continuous systems that admit an approximate analysis.

We begin this chapter with a motivating example. The remainer of the chapter factors into three parts. First, we define the class of maps and the class of statistics that will be studied. Next, we determining analytical expressions for the statistics. Finally, we make some general observations concerning the structure of the statistics.

3.1 A Motivating Example

In this section we motivate the results in the remainder of the chapter by deriving an expression for the power spectrum of a simple chaotic map. The result presented here is not new—it was first reported in [18]. However, our derivation has two advantages over that of [18] : it is substantially simpler and, more importantly, it can be generalized to apply to a much larger class of maps.

We wish to determine the power spectrum of the stationary process generated by the map

$$f(x) = \begin{cases} \frac{x}{a} & 0 \leq x \leq a \\ \frac{1-x}{1-a} & a \leq x \leq 1 \end{cases} \tag{3.2}$$

by computing the Fourier transform of its autocorrelation sequence. The map f generates time series of the form $x[n] = f(x[n-1])$ so the k-th autocorrelation lag of the process is

of the form

$$\begin{aligned} R_{xx}[k] &= E\{x[n]x[n+k]\} \\ &= E\{x[n]f^k(x[n])\}. \end{aligned} \tag{3.3}$$

The expected value (3.3) is taken with respect to the marginal density of $x[n]$ or more precisely, the invariant density of the map f.

The invariant density must be determined in order to compute the expected value (3.3). According to the discussion of Section 2.2.1, the invariant density of f is a fixed point its Frobenius–Perron operator. Although this fixed point is difficult to determine for general maps, in this special case a solution straightforward. The FP operator corresponding to f relates the probability density of x to that of $y = f(x)$ and has the form

$$P_f p(x) = ap(ax) + (1-a)p(1-(1-a)x). \tag{3.4}$$

It follows that invariant density satisfies

$$p(x) = ap(ax) + (1-a)p(1-(1-a)x). \tag{3.5}$$

A solution is the uniform density

$$p(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.6}$$

as can be verified by substitution into (3.5). We shall take the expected value with respect to this density.

Using (3.6), we can relate consecutive autocorrelation values through the following series of manipulations:

$$
\begin{aligned}
R_{xx}[k] &= E(x[n]x[n+k]) && (3.7)\\
&= \int_{[0,1]} x f^k(x)dx \\
&= \int_{[0,a]} x f^k(x)dx + \int_{[a,1]} x f^k(x)dx \\
&= a^2 \int_{[0,1]} x f^k(x/a)dx - (1-a)^2 \int_{[0,1]} x f^k(1-(1-a)x)dx \\
&\quad + \int_{[0,1]} (1-a) f^k(1-(1-a)x)dx \\
&= a^2 \int_{[0,1]} x f^{k-1}(x)dx - (1-a)^2 \int_{[0,1]} x f^{k-1}(1-(1-a)x)dx \\
&\quad + \int_{[0,1]} (1-a) f^{k-1}(x)dx \\
&= (2a-1) \int_{[0,1]} x f^k(x)dx + \int_{[0,1]} (1-a) f^{k-1}(x)dx \\
&= (2a-1)R_{xx}[k-1] + (1-a)E(x). && (3.8)
\end{aligned}
$$

The second equality is a result of the definition of the expected value and the uniform invariant density; the third is obvious; the fourth is a result of the substitution $y = \frac{x}{a}$ in the first term and $y = \frac{1-x}{1-a}$ in the second; the fifth follows from the the definition of the map f; collecting terms yields the sixth ; and the seventh is an application of the definition of the autocorrelation.

The derivation of (3.8) is valid for $k > 0$. For $k < 0$, we have $R_{xx}[k] = R_{xx}[-k]$ since the autocorrelation sequence is conjugate symmetric and $x[n]$ is real. Using in (3.8) the quantities $R_{xx}[0] = E(x^2) = 1/3$ and $E(x) = 1/2$, which both follow from the form of the invariant density (3.6), we obtain that the Fourier transform $R_{xx}[k]$ is of the form

$$
S_{xx}(e^{j\omega}) = \frac{1}{12(1-(2a-1)z)(1-(2a-1)z^{-1})}\bigg|_{z=e^{j\omega}} + \pi\delta(\omega).
$$

Thus, the power spectrum associated with f is that of a first order autoregressive process with mean $1/2$.

The technique presented in the sequel is more general than the example of this section—

it computes a larger class of statistics for a larger family of maps. The present example does, however, illustrate the approach we will take in the sequel. In particular, the derivation relied on the explicit form of f and its piecewise linearity. In the next section, a class of piecewise-linear systems called Markov maps, unique because their statistics can be determined in closed form, is introduced and analyzed. As the Markov maps include the tent maps presented in this section, the results of this section will be special cases of more general results.

3.2 Piecewise-Linear Markov Maps

The main objects of study in this chapter are the eventually-expanding, piecewise-linear Markov maps. These maps, which are defined below, were first introduced because their invariant densities exist and can be exactly determined [19, 20]. Very little subsequent work has been directed toward understanding their more general statistical properties. The following sections explore these more general properties.

Eventually-expanding, piecewise-linear, Markov maps are one-dimensional chaotic systems which are amenable to analysis. They are defined as follows.

Definition 1 *A map $f : [0,1] \to [0,1]$ is an eventually-expanding, piecewise-linear, Markov map if and only if*

1. *There is a set of* partition points $0 = a_0 < a_1 < \cdots < a_N = 1$ *such that restricted to each of the intervals (a_{i-1}, a_i), the map f is affine, i.e. $f(x)|_{(a_{i-1},a_i)} = s_i x + b_i$.*

2. *For each i, $f(a_i) = a_j$ for some j.*

3. *There is an integer $k > 0$ such that $\inf_{x\in[0,1]} |\frac{d}{dx} f^k(x)| > 1$.*

Thus, an eventually-expanding, piecewise-linear Markov map consists of a finite number of affine segments; maps partition points to partition points; and is strictly expanding after a finite number of iterations. For example, the tent map (3.2) is an eventually-expanding, piecewise-linear, Markov map. In fact, all of the maps we will examine in this chapter are eventually-expanding, piecewise-linear Markov maps. Rather than repeatedly using this

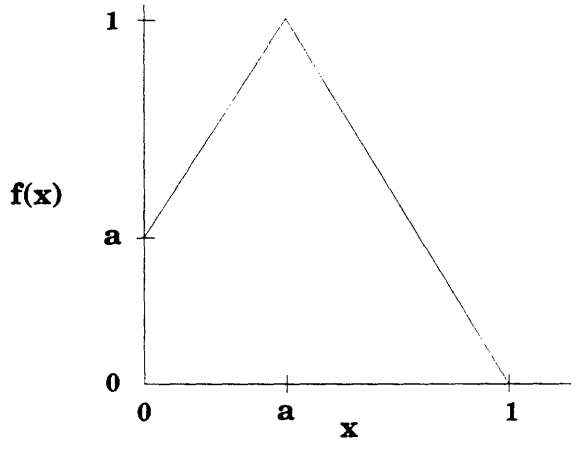


Figure 3-2: An Example of a Piecewise-Linear Markov Map with Two Partition Elements

cumbersome but precise name, the abbreviated name "Markov maps" will be used when there is no possible ambiguity.

It is appropriate here to introduce some notation that is useful in the subsequent development. The interval $[a_{i-1}, a_i]$ is called the i-th partition element and denoted by I_i. By definition, Markov maps take partition points to partition points and hence map partition elements onto unions of partition elements. We denote by $\mathcal{I}_i$ the set of indices of partition elements in the image of I_i. With this notation, the image may be expressed in the form

$$f(I_i) = \cup_{j \in \mathcal{I}_i} I_j.$$

Figure 3-2 shows the graph of the simple Markov map

$$f(x) = \begin{cases} \frac{1-a}{a}x + a & 0 \leq x \leq a \\ \frac{1}{1-a}(1-x) & a < x \leq 1 \end{cases} \tag{3.9}$$

This map has partition elements $I_i = [0, a]$ and $I_2 = [a, 1]$. The index sets associated with the partition elements are $\mathcal{I}_1 = \{2\}$ and $\mathcal{I}_2 = \{1, 2\}$.

More generally, the dependence of a Markov map on its parameters is made explicit in

the following functional form:

$$f(x) = \sum_{i=1}^{N} (s_i x + b_i)\chi_i(x), \tag{3.10}$$

where the $\chi_i(x)$ is the indicator function

$$\chi_i(x) = \begin{cases} 1 & \text{if } x \in I_i \\ 0 & \text{otherwise} \end{cases} . \tag{3.11}$$

As a result of the eventually-expanding property, the slopes satisfy $|s_i| > 0$. It follows that Markov maps are nonsingular transformations (see Section 2.2).

As mentioned in Section 2.2.1, analyzing a chaotic systems using a statistical approach requires it to have an invariant measure. All Markov maps have invariant measures (see [19] for a proof). The next section shows how these measures can be explicitly determined[1].

3.3 Computing the Statistics of Markov Maps

An important property of Markov maps is that many of their statistics can be calculated in closed form. Section 3.1 presented an example calculation for the special case of the autocorrelation of the tent map. This section, introduces a broad class of statistics that can be solved for in closed form.

We will derive closed form expressions for statistics of the form

$$R_{f;h_0,h_1,\ldots,h_r}[k_1,\ldots,k_r] = E\{h_0(x[n])h_1(x[n+k_1])\cdots h_r(x[n+k_r])\}, \tag{3.12}$$

when $x[n]$ is a time series generated by a Markov map f. These statistics, which we call correlation statistics, include the autocorrelation and all higher order moments of the process generated by f. The broad definition of correlation statistics makes them potentially applicable in a number of chaotic data analysis and synthesis problems. Two applications

[1] An invariant measure of a Markov map is smooth with respect to Lebesgue measure. Coupled with the piecewise-linearity of the map, this implies that, formally, the rules of classical calculus can be used rather than those of measure theory for our statistical calculations in the following sections.

are presented in this thesis: Chapter 5 presents a statistical analysis of a switching power converter; Chapter 6 presents a method of synthesizing Markov maps with white power spectra.

The statistics of a chaotic system depend on its Frobenius–Perron operator. One connection between statistics and the FP operator is that the expectation of (3.12) is taken with respect to an invariant density that is in turn a fixed point of the FP operator. Invariant density aside, the FP operator is involved in the computation of correlation statistics at a more fundamental level. In particular, it can be shown that the FP operator satisfies the relationship

$$\int_{[0,1]} h_1(x) P_f h_2(x) dx = \int_{[0,1]} h_1(f(x)) h_2(x) dx \tag{3.13}$$

for all integrable h_i. The correlation statistic of (3.12) can be written in the form

$$R_{f;h_0,h_1,\ldots,h_r}[k_1,\ldots,k_r] = \int_{[0,1]} h_0(x) h_1(f^{k_1}(x)) \cdots h_r(f^{k_r}(x)) p(x) dx \tag{3.14}$$

where $p(x)$ is the invariant density of the map f. Repeated application (3.13) to (3.14) yields

$$R_{f;h_0,h_1,\ldots,h_r}[k_1,\ldots,k_r] = \int_{[0,1]} h_r(x) (P_f^{k_r-k_{r-1}}(h_{r-1}(x) \cdots P_f^{k_2-k_1}(h_1(x) P_f^{k_1}(h_0(x)p(x))) \cdots) dx. \tag{3.15}$$

Equation (3.15) explicitly displays the relationship between the FP operator of a map and its correlation statistics. This relationship is valid not only for Markov maps but for all nonsingular one-dimensional maps—a fact we shall make use in Chapter 4. For Markov maps we will show below that (3.15) suggests a strategy for computing the integral in closed form. As a first step toward this end, the invariant density must be determined.

3.3.1 Invariant Densities of Markov Maps

Markov maps have invariant densities that can be determined in closed form. As the invariant density must be known in order to determine the correlation statistics (3.15), this result is of significant interest. The derivation of the closed form solution for the invariant

density of a Markov map was first reported in [20] and later in [19]. We include it here for primarily for completeness and to lead into the discussion of more general statistics.

The FP operator for nonsingular, one-dimensional maps of the form

$$f(x) = \sum_{i=1}^{N} f_i(x)\chi_i(x)$$

is (see [15])

$$(P_f h)(x) = \sum_{i=1}^{N} \frac{h(f_i^{-1}(x))\chi_{f_i(I_i)}(x)}{|f'(f_i^{-1}(x))|}. \tag{3.16}$$

The FP operator for a Markov map follows by substituting (3.10) into (3.16) to obtain

$$(P_f h)(x) = \sum_{i=1}^{N} \frac{h(\frac{x-b_i}{s_i})\chi_{f_i(I_i)}(x)}{|s_i|}. \tag{3.17}$$

Our goal is to determine invariant density of a Markov map, or phrased differently, to find a fixed point of (3.17). A fact that is instrumental in finding such a fixed point is that the invariant densities of a Markov map is constant on each partition element [19]. More specifically, $p(x)$ is of the form

$$p(x) = \sum_{i=1}^{N} p_i \chi_i(x),$$

where $p_i \geq 0$ is a constant. It is thus natural to examine the structure of the FP operator acting on functions that are constant on the partition elements.

We proceed by examining the FP operator acting the indicator functions χ_i. Its action on linear combinations of indicator functions follows from its linearity. Substituting $h(x) = \chi_j(x)$ into (3.16) and simplifying yields

$$P_f \chi_j(x) = \frac{1}{|s_j|} \sum_{i \in \mathcal{I}_j} \chi_i(x). \tag{3.18}$$

Thus, the image of the indicator function of a partition element is a sum of indicator functions of partition elements. As a consequence of its linearity and (3.18), the FP operator maps linear combinations of the indicator functions $\{\chi_i\}_{i=1}^{N}$ to linear combinations of indicator functions.

In what follows, it is convenient to denote by $\mathcal{P}_0$ the N dimensional space spanned by $\{\chi_i\}_{i=1}^N$. Each point in a $\mathcal{P}_0$ is of the form

$$h(x) = \sum_{i=1}^{N} h_i \chi_i(x). \tag{3.19}$$

where h_i is a constant. Each point of $\mathcal{P}_0$ can thus be uniquely represented by the N-tuple $\bar{h} = [h_1, \ldots, h_N]^T$. So for example, the characteristic functions χ_1 and χ_2 correspond to the N-tuples $\bar{\chi}_1 = [1, 0, \ldots, 0]^T$ and $\bar{\chi}_2 = [0, 1, 0, \ldots, 0]^T$ respectively.

By the previous discussion, a Markov map's FP operator maps $\mathcal{P}_0$ to itself and thus, when restricted to $\mathcal{P}_0$, can be represented by an $N \times N$ matrix. We denote this matrix by $\bar{P}_0$. With respect to the vectors $\{\chi_i\}_{i=1}^N$, the i-th column of the matrix is the coordinate vector of the image of $\chi_i(x)$ under P_f as in (3.18). More specifically, the elements of $\bar{P}_0$ are

$$[\bar{P}_0]_{i,j} = \begin{cases} \frac{1}{|s_j|} & \text{if } i \in \mathcal{I}_j \\ 0 & \text{otherwise} \end{cases}, \tag{3.20}$$

where s_j is the slope of the j-th affine segment of f (see (3.10)).

The matrix $\bar{P}_0$ is the basis for an analytic solution for the invariant density of a map f. An invariant density is a solution to the fixed point equation

$$P_f p(x) = p(x) \tag{3.21}$$

Suppose $p(x)$ is piecewise-constant on the partition $\{I_i\}_{i=1}^N$. Then (3.21) can be expressed in terms of $\bar{P}_0$ as

$$\bar{P}_0 \bar{p} = \bar{p}, \tag{3.22}$$

where $\bar{p}$ is the coordinate vector of the invariant density $p(x)$. In other words, according to (3.22), the coordinate vector of an invariant density is the eigenvector of $\bar{P}_0$ corresponding to the eigenvalue 1.

That such an eigenvector/eigenvalue pair always exists has been shown by Friedman and Boyarsky [21]. Specifically, they have shown:

Lemma 1 *The matrix $\bar{P}_0$ is diagonally similar to a column stochastic matrix, i.e.*

$$T = D^{-1}\bar{P}_0 D, \tag{3.23}$$

where, D is diagonal and has positive entries. The matrix T has positive elements and each of its columns sums to unity.

Similar matrices have the same eigenvalues. It follows that $\bar{P}_0$ has the same eigenvalues as some stochastic matrix. Frobenius's theorem [22] asserts that all stochastic matrices, and hence $\bar{P}_0$, have an eigenvalue equal to unity and further that this is an eigenvalue of maximum magnitude. Because the elements of $\bar{P}_0$ are nonnegative, the eigenvector corresponding to the eigenvalue 1 has positive components. The invariant density resulting from this eigenvector is thus nonnegative in accordance with intuition.

Suppose that the solution to the fixed point equation (3.22) is $\bar{p} = [p_1, \ldots, p_N]^T$. An invariant density of f is of the form

$$p(x) = \frac{1}{\sum_{i=1}^{N} p_i |I_i|} \sum_{i=1}^{N} p_i \chi_i(x) \tag{3.24}$$

where $|I_i| = a_i - a_{i-1}$ is the length of the i-th partition element and the leading constant ensures that the density has an integral equal to one.

Example

Consider again the piecewise-linear Markov map of (3.9) shown in Figure 3-2. When restricted to the two-dimensional space of piecewise-constant functions, the FP operator associated with f is represented by the 2×2 matrix

$$\bar{P}_0 = \begin{bmatrix} 0 & 1-a \\ \frac{a}{1-a} & 1-a \end{bmatrix}.$$

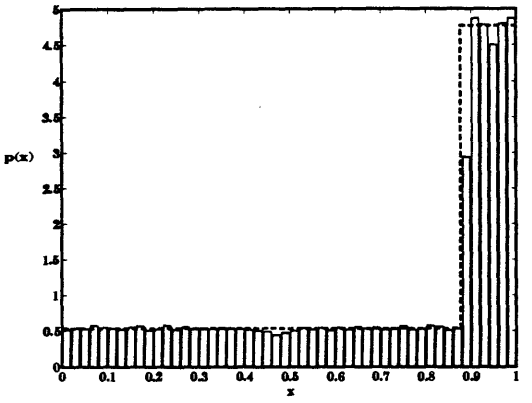


Figure 3-3: A Comparison of the Analytic and Empirical Density for the Markov Map of the Example

The eigenvector associated with the eigenvalue 1 is $p = [(1-a), 1]^T$. Referring to (3.24), the invariant density is the piecewise-constant function

$$p(x) = \begin{cases} \frac{1}{1+a} & 0 \leq x \leq a \\ \frac{1}{1-a^2} & a \leq x \leq 1 \end{cases}.$$

In Figure 3-3 this density is compared with an empirically estimated density computed via a histogram of 50000 points generated by the map f with parameter $a = 8/9$.

3.3.2 Matrix Representations for the FP operator

In light of the relationship (3.15) between the FP operator and correlation statistics, a compact representation of the correlation statistic relies on a compact representation of the FP operator. By definition, the FP operator is an infinite dimensional linear operator since it maps the infinite dimensional space of integrable functions to itself. However, as we have seen above, when restricted to a space of piecewise-constant functions, a Markov map's FP operator has representation as a finite dimensional matrix. In this section, we show that the FP operator may be represented by finite dimensional matrices on larger spaces.

The results of the previous section, which held for the piecewise-constant functions

(3.19), generalize in a straightforward way to piecewise-polynomial functions of the form

$$h(x) = \sum_{i=1}^{N} \sum_{j=0}^{k} a_{ij} x^j \chi_i(x). \tag{3.25}$$

It is convenient to introduce the notation

$$\{\theta_1, \ldots, \theta_{N(k+1)}\} \stackrel{\Delta}{=} \{\chi_1, \ldots, \chi_N, x\chi_1, \ldots, x\chi_N, \ldots, x^{k-1}\chi_1, \ldots, x^{k-1}\chi_N\}. \tag{3.26}$$

so that $h(x)$ can be written in the form

$$h(x) = \sum_{i=1}^{N(k+1)} h_i \theta_i(x). \tag{3.27}$$

With this notation, each piecewise-polynomial of degree k is uniquely specified by the $N(k+1)$-tuple $\bar{h} = [h_1, \ldots, h_{N(k+1)}]^T$ which we refer to as the coordinate vector of $h(x)$. Keep in mind that the piecewise-polynomial (3.27) depends on the partition elements I_i (and hence the map f) through the characteristic functions $\chi_i(x)$. This dependence will be mentioned only when the partition is not clear from the context of the discussion.

Proceeding along the same lines as the discussion of Section 3.3.1, we consider first the FP operator acting on functions of the form $h(x) = x^k \chi_i(x)$. Substituting $h(x)$ into (3.17) yields

$$(P_f x^k \chi_j)(x) = \sum_{i=1}^{N} \frac{(\frac{x-b_i}{s_i})^k \chi_i(\frac{x-b_i}{s_i}) \chi_{f_i(I_i)}(x)}{|s_i|}. \tag{3.28}$$

Expression (3.28) simplifies substantially upon taking advantage of f's Markov property. After some straightforward manipulations we obtain

$$P_f x^k \chi_j(x) = (\frac{x - b_j}{s_j})^k \frac{1}{|s_j|} \sum_{i \in \mathcal{I}_j} \chi_i(x), \tag{3.29}$$

where s_j and b_j are the parameters of f (see (3.10)). Thus the image of $h(x)$ is a piecewise-polynomial. By the linearity of the FP operator, it follows that the image of a piecewise-polynomial of degree k is a piecewise polynomial of degree k.

Denote by $\mathcal{P}_k$ the space of piecewise-polynomials of degree k. Because $\mathcal{P}_k$ is spanned by

the linearly independent functions $\{\theta_i\}_{i=1}^{N(k+1)}$, its dimension is $N(k+1)$. By the previous discussion, P_f maps the finite dimensional space $\mathcal{P}_k$ to itself and hence, its restriction to $\mathcal{P}_k$ can be represented by an $N(k+1) \times N(k+1)$ matrix. We denote this matrix by $\bar{P}_k$. With respect to the set of functions $\{\theta_i\}_{i=1}^{N(k+1)}$, the i-th column of the matrix $\bar{P}_k$ is the coordinate vector of the image of θ_i under P_f which is given by (3.29). More specifically, expanding the first factor on the right hand side of (3.29) using the binomial theorem yields that $\bar{P}$ is of the form

$$\bar{P}_k = \begin{bmatrix} P_{00} & P_{01} & \cdots & \cdots & P_{0k} \\ 0 & P_{11} & P_{12} & \cdots & P_{1k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & P_{kk} \end{bmatrix}, \tag{3.30}$$

where each block of $\bar{P}_k$ is an $N \times N$ matrix of the form

$$P_{ij} = \begin{cases} \begin{pmatrix} j \\ i \end{pmatrix} \bar{P}_0 B^{j-i} S^j & j \geq i \\ 0 & \text{otherwise} \end{cases}. \tag{3.31}$$

The matrices B and S are diagonal with elements

$$B = \begin{bmatrix} -b_1 & 0 & \cdots & \cdots & 0 \\ 0 & -b_2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & -b_p \end{bmatrix}. \tag{3.32}$$

and

$$S = \begin{bmatrix} \frac{1}{s_1} & 0 & \cdots & \cdots & 0 \\ 0 & \frac{1}{s_2} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \frac{1}{s_p} \end{bmatrix}. \tag{3.33}$$

and $\bar{P}_0$ is the $N \times N$ matrix with elements defined in (3.20).

In summary, a Markov map's FP operator on the space $\mathcal{P}_k$ may be represented by the

$N(k+1) \times N(k+1)$ dimensional matrix defined by (3.30), (3.32), (3.33) . The matrix is quite structured and easy to compute directly from the map's parameters.

Since an invariant density is an eigenvector of the FP operator associated with a unity eigenvalue, the eigenstructure of the matrix $\bar{P}_k$ also provides some information about the form of the invariant densities of Markov maps. One consequence of the block triangular structure of $\bar{P}_k$ is that its eigenvalues are the eigenvalues of the submatrices $P_{ii} = \bar{P}_0 S^i$ on its diagonal. For an expanding map, i.e. one with a slope that is strictly greater than one, the the elements of the diagonal matrix S are strictly less than one. It follows that the eigenvalues of P_{ii}, which we denote by $\{\lambda_{ij}\}_{j=1}^{N}$, satisfy

$$\begin{aligned} |\lambda_{ij}| &\leq \|P_{ii}\| \\ &\leq \|\bar{P}_0\|\|S\|^i \\ &< 1. \end{aligned}$$

when $i > 0$ When f is eventually-expanding, a similar argument applied to P_{ii}^m for sufficiently large m shows that its eigenvalues are strictly less than one in magnitude when $i > 0$. It follows that the only piecewise-polynomial invariant densities of a Markov map are associated with $\bar{P}_0$, or, phrased differently, that all piecewise-polynomial invariant densities of a Markov map are piecewise-constant.

Example

In Section 3.4, expressions for the power spectrum of time series generated by Markov maps will be derived. As a first step in this process, the second order correlation $E\{x[n]x[n+k]\}$ must be computed, which in turn requires the FP operator on the space of piecewise-linear functions. For the map pictured in Figure 3-2 the matrix representation of the FP operator on $\mathcal{P}_1$ is

$$\bar{P}_1 = \begin{bmatrix} 0 & 1-a & 0 & (1-a) \\ \frac{a}{1-a} & 1-a & -\frac{a^3}{(1-a)^2} & (1-a) \\ 0 & 0 & 0 & -(1-a)^2 \\ 0 & 0 & \frac{a^2}{(1-a)^2} & -(1-a)^2 \end{bmatrix}.$$

Here the block upper triangular form mentioned above is clear.

3.3.3 Analytical Expressions for Correlation Type Statistics

The matrix representation for the FP operator developed in the previous section is the central tool needed to determine closed form expressions for the correlation statistics (3.12). Once the appropriate representation for the FP operator has been determined and the invariant density has been computed, what remains primarily a matter of notation.

Consider first the integrand of (3.15). When each h_i is a piecewise-polynomial, straightforward application of the definitions of the previous sections yields that the integrand

$$h_r(x)(P_f^{k_{r-1}-k_{r-2}}(h_{r-1}(x)\cdots P_f^{k_2-k_1}(h_2(x)P_f^{k_1}(h_1(x)p(x)))\cdots) \tag{3.34}$$

has coordinate vector of the form

$$\bar{h}_r \odot (\bar{P}^{k_{r-1}}(\bar{h}_{r-1} \odot \bar{P}^{k_{r-2}}(\cdots \bar{P}^{k_1}(\bar{h}_1 \odot \bar{p})\cdots), \tag{3.35}$$

where $\bar{P}$ is the matrix representation of the FP operator acting on the space of piecewise-polynomials of appropriate dimension, and the notation $\odot$ denotes multiplication of piecewise-polynomials. More specifically, if $\bar{h}_1$ and $\bar{h}_2$ are coordinate vectors of piecewise-polynomials, the vector product product $\bar{h}_3 = \bar{h}_1 \odot \bar{h}_2$ is defined as the coordinate vector of the product $h_3(x) = h_1(x)h_2(x)$.

At first glance, (3.35) appears to be a complex way to express the integrand. However, it provides a direct computational approach using only straightforward matrix–vector operations. Furthermore, it is often the case that the individual statistics are of less interest than some transformation of a sequence of statistics. This is the case for instance when we are interested in the power spectrum of a process generated by a Markov map. In many situations of this type, (3.35) is more easily manipulated than (3.34).

What remains is to evaluate the integral on the right hand side of (3.15). The integrand (3.34) can be viewed as a product of two piecewise-polynomials—$h_r(x)$ and the parenthesized term on the right side of (3.35). The integral is a bilinear form in these two functions and can be implemented by a finite dimensional matrix bilinear form since each function is

an element of a finite dimensional vector space. Thus, the integral itself can be computed in terms of the coordinate vectors. Specifically, suppose two piecewise-polynomials $g_1, g_2 \in \mathcal{P}_k$ have coordinate vectors $\bar{g}_1, \bar{g}_2$, then

$$\int_{[0,1]} g_1(x) g_2(x) dx = \bar{g}_1^T M_k \bar{g}_2. \tag{3.36}$$

where

$$[M_k]_{ij} = \int_{[0,1]} \theta_i(x) \theta_j(x) dx \tag{3.37}$$

and θ_i, θ_j are are defined in (3.26). Thus, M is a $N(k+1) \times N(k+1)$ matrix. Note that in each row and column there are at most k nonzero entries because the support of only k of the θ_i's coincide and the rest are disjoint. The sparse structure of M may be of use in computationally efficient implementations of (3.36).

Example: Returning again to the map pictured in Figure 3-2, the matrix M_1 is of the form

$$M_1 = \begin{bmatrix} a & 0 & \frac{a^2}{2} & 0 \\ 0 & 1-a & 0 & \frac{1-a^2}{2} \\ \frac{a^2}{2} & 0 & \frac{a^3}{3} & 0 \\ 0 & \frac{1-a^2}{2} & 0 & \frac{1-a^3}{3} \end{bmatrix}.$$

Given this matrix representation of the integral and the matrix representation of the FP operator on $\mathcal{P}_k$ we can obtain a matrix representation of correlation type statistics as follows. Let $\bar{g}_2$ be the parenthesized term of (3.35). Then

$$R_{f;h_0,\ldots,h_r}[k_1, \ldots, k_r] = \bar{h}_r M \bar{g}_2 \tag{3.38}$$

where M is the matrix with elements given by (3.37) with the appropriate basis elements.

Equations (3.38) and the relationship expressed in (3.34) and (3.35) together provide closed form analytical expression for a general class of correlation statistics. This closed form result not only provides a method to compute the numerical values of correlation statistics, but also provides some insight into the overall statistical structure of Markov maps. In the next section we consider one aspect of this structure, namely, the form of the

power spectra associated with Markov maps.

3.4 Spectra of Markov Maps

In this section, the techniques developed above will be used to find the power spectra of time series generated by Markov maps. We will see that such time series have rational spectra with readily computable parameters. This is a rather remarkable property of Markov maps—the results of Chapter 4 show that more general maps need not have rational spectra. As in the example of Section 3.1 our approach is to compute the autocorrelation sequence of a map; its Fourier transform is the desired power spectrum.

The correlation values to be computed are

$$\begin{aligned} R_{f;x,x}[k] &= E\{xf^k(x)\} \\ &= \int_{[0,1]} xf^k(x)p(x)dx \end{aligned}$$

where p is the invariant density associated with f and m_x is the expected value of x. Let $h_1(x) = x$ and $h_2(x) = xp(x)$. Note that because the invariant density of any Markov map is piecewise-constant (see Section 3.3.1), the product $xp(x)$ is piecewise-linear, as is $h_1(x)$. It follows that $\bar{P}_1$, the matrix representation of the FP operator on $\mathcal{P}_1$, and M_1 must be determined in order to compute the autocorrelation sequence. Using (3.38) and (3.35), the correlation sequence can be expressed in the form

$$R_{xx}[k] = \bar{h}_1^T M_1 \bar{P}_1^{|k|} \bar{h}_2, \tag{3.39}$$

where each term in (3.39) may be determined as in the preceding sections.

The power spectrum associated with f is obtained by taking the Fourier transform of (3.39)

$$\begin{aligned} S_{xx}(e^{j\omega}) &= \sum_{n=-\infty}^{\infty} R_{xx}[n]e^{-j\omega n} \\ &= \bar{h}_1^T M(\sum_{n=-\infty}^{\infty} \bar{P}_1^{|n|} e^{-j\omega n})\bar{h}_2 \end{aligned} \tag{3.40}$$

The only issue that needs to be resolved is whether the infinite sum in (3.40) converges. Because $\bar{P}_1$ has at least one eigenvalues of unity magnitude (corresponding to the invariant density), a complication arises when performing the sum (3.40). Any eigenvalues of $\bar{P}_1$ with unit magnitude correspond to impulses in the Fourier transform. In particular, because a Markov map on the unit interval generates time series with nonzero mean, its spectrum must have an impulse at zero frequency. These impulsive components may be isolated however.

The eigenvalues of unit magnitude can be isolated by writing $\bar{P}_1$ in Jordan form:

$$\bar{P}_1 = E^{-1} J E$$

where E is a matrix of generalized eigenvectors. Writing J in the form

$$J = \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix},$$

where J_1 consists of Jordan blocks with eigenvalues of unit magnitude and J_2 consist of Jordan blocks with magnitude strictly less than unity. By lemmas 1 and and the fact that all eigenvalues of unit magnitude are associated with $\bar{P}_0$ (see Section 3.3.2), there are no other eigenvalues. Now, $\bar{P}_1$ can be expressed as

$$\bar{P}_1 = \Gamma_1 + \Gamma_2$$

where Γ_1 and Γ_2 are defined by

$$\begin{aligned} \Gamma_1 &= E^{-1} \begin{bmatrix} J_1 & 0 \\ 0 & 0 \end{bmatrix} E, \\ \Gamma_2 &= E^{-1} \begin{bmatrix} 0 & 0 \\ 0 & J_2 \end{bmatrix} E \end{aligned}$$

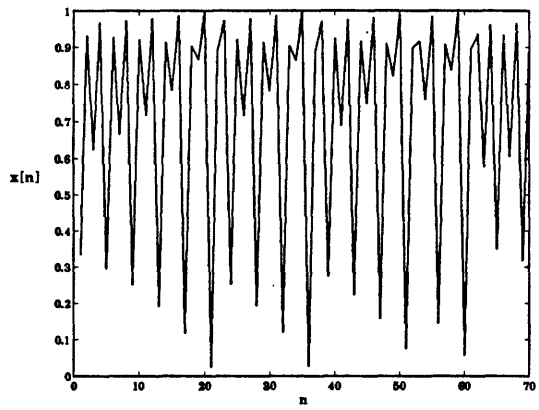


Figure 3-4: A Sample Time Series Generated by the Markov Maps of the Example

It can be readily verified that $\bar{P}_1^k = \Gamma_1^k + \Gamma_2^k$ so that the sum (3.40) is of the form

$$S_{xx}(e^{j\omega}) = \bar{h}_1^T M(I - \Gamma_2 e^{-j\omega})^{-1}(I - \Gamma_2^2)(I - \Gamma_2 e^{j\omega})^{-1}\bar{h}_2 + \sum_{i=1}^{p} C_i \delta(\omega - \omega_i), \qquad (3.41)$$

where C_i and ω_i depend on Γ_1 and p is no larger than the dimension of the J_1. Since the process generated by f has a nonzero mean, $p \geq 1$ and $\omega_i = 0$ for some i.

Examining (3.41) more closely, we may infer that the spectrum of a Markov map is a linear combination of a rational function and impulses. The poles of the rational portion of the spectrum correspond to the eigenvalues of the matrix Γ_2, i.e. the eigenvalues of $\bar{P}_1$ with magnitude less than one. The zeros of $S_{xx}(e^{j\omega})$ depend on the vectors $\bar{h}_1$ $\bar{h}_2$ and the matrix M.

Example

Consider again the Markov map

$$f(x) = \begin{cases} \frac{1-a}{a}x + a & 0 \leq x \leq a \\ \frac{1}{1-a}(1-x) & a < x \leq 1 \end{cases}$$

with parameter $a = \frac{8}{9}$. A time series generated by iterating f from the initial condition $x[0] = 1/3$ is shown Figure 3-4.

The matrices M and $\bar{P}_1$ and the invariant density $p(x)$ have been computed in previous

examples and are evaluated here for the parameter value $a = \frac{8}{9}$:

$$\bar{P}_1 = \begin{bmatrix} 0 & \frac{1}{9} & 0 & \frac{1}{9} \\ 8 & \frac{1}{9} & -\frac{512}{9} & -\frac{1}{9} \\ 0 & 0 & 0 & -\frac{1}{81} \\ 0 & 0 & 64 & -\frac{1}{81} \end{bmatrix} ..$$

$$M = \begin{bmatrix} \frac{8}{9} & 0 & \frac{32}{81} & 0 \\ 0 & \frac{1}{9} & 0 & \frac{17}{162} \\ \frac{32}{81} & 0 & \frac{512}{2187} & 0 \\ 0 & \frac{17}{162} & 0 & \frac{217}{2187} \end{bmatrix} ..$$

$$p(x) = \begin{cases} \frac{9}{17} & 0 \le x \le \frac{8}{9} \\ \frac{81}{17} & \frac{8}{9} \le x \le 1 \end{cases} .$$

The expected value is $m_x = \frac{217}{306}$. Using (3.41) we obtain that the rational part of the z transform of the correlation sequence is of the form

$$S_{xx}(z) = -\frac{42632}{459} \frac{36z^{-1} - 145 + 36z}{(9+8z)(9+8z^{-1})(64z^2 + z + 81)(64z^{-2} + z^{-1} + 81)}. \tag{3.42}$$

Both the numerator and denominator polynomials are symmetric in z and z^{-1} which is consistent with the fact that any rational power spectrum must have a factorization

$$S_{xx}(z) = \tilde{S}(z)\tilde{S}(z^{-1}),$$

for some rational function $\tilde{S}$. Thus, a time series generated by the map f has a three pole spectrum.

The power spectrum $S_{xx}(e^{j\omega})$ is plotted in Figure 3-5 along with an empirical spectrum computed by periodogram averaging with a window length of 128 on a time series of length 50000. The solid line is the plot of the analytical expression and the dotted line is the empirically computed spectral estimate. The agreement between the analytical and the empirical results is good.

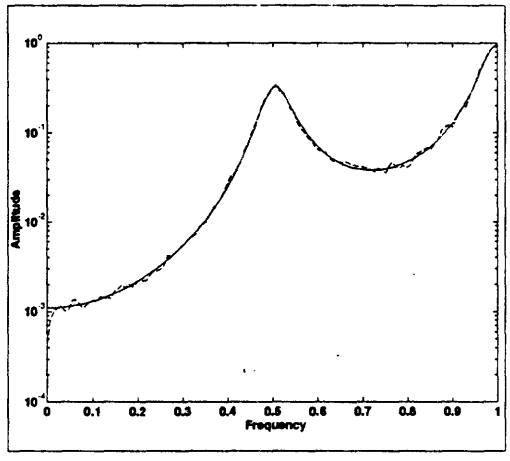


Figure 3-5: Comparison of Analytically Computed Spectrum and Empirical Spectrum

We have seen that corresponding to each Markov map is a rational power spectrum. It is natural to ask whether the converse is true. That is, given an arbitrary rational system function S_{xx} that is positive on the unit circle, is there a Markov map with S_{xx} as its power spectrum? This question appears to be difficult to answer. However, the following discussion may provide some insight into the question itself.

As we have seen, the spectrum of a Markov map depends FP operator, and in particular its eigenvalues, in a fundamental way. Because the matrix $\bar{P}_1$ is highly structured, its eigenvalues may not take on arbitrary values. Specifically, $\bar{P}_1$ is block upper triangular and each of its diagonal blocks is a product of the matrix $\bar{P}_0$ with a diagonal matrix. As discussed above $\bar{P}_0$ is similar to a stochastic matrix. It follows that some insight into the poles of the spectrum can be gained by looking at eigenvalue structure of stochastic matrices. Indeed, it is known that there need not exist a stochastic matrix with an arbitrary set of eigenvalues (see [22]). Moreover, the set of possible eigenvalues of a stochastic matrix is highly constrained. As a consequence, Markov maps with arbitrary rational spectra need not exist. While these remarks do not answer the question posed above, they do point out the fundamental difficulties in finding such an answer.

Chapter 4

Maps of the Interval: NonMarkov Maps

The Markov maps introduced in Chapter 3 have several properties made them worthy of further study. Perhaps as important as their analytic tractability and easily described statistical structure is their ability to serve as useful models for much larger class of systems. In particular, in this chapter, we show that a much larger class of maps—the eventually-expanding maps—can be analyzed indirectly by applying the techniques developed for Markov maps. Indeed, every map in this larger class can be approximated arbitrarily well by a Markov map. Of course, in order to make the previous statement meaningful eventually-expanding maps must be defined and the sense of approximation must be described more precisely. This chapter addresses these two issues.

The relationship between eventually-expanding maps and Markov maps has other implications. For instance, we will show that the power spectra of eventually-expanding maps are similar to those of Markov maps. More broadly, the goal of this chapter is to determine the structure of correlation statistics of the form

$$R_{f;h_0,h_1,\ldots,h_r}[k_1,\ldots,k_r] = \int_{[0,1]} h_0(x)h_1(f^{k_1}(x))\cdots h_r(f^{k_r}(x))p(x)dx \qquad (4.1)$$

for the larger class of eventually-expanding maps.

4.1 Eventually-Expanding Maps

Many examples of maps occurring in practice are neither Markov nor piecewise linear [10, 23, 24]. In this section, we define a broader class of maps, called eventually-expanding maps, that retains some of the analytic tractability of Markov maps. Unlike the Markov maps considered in the previous chapter, we will not be able to determine closed form expressions for the statistics of eventually-expanding maps. However, we will be able to approximate them as closely as desired.

With little loss in generality, we will consider only maps on the unit interval. This restriction is not as stringent as it may appear, since if a time series is generated by $x[n] = g(x[n-1])$ with g a map on an interval $[a, b]$, the time series $z[n] = x[n]/(b-a) - a$ satisfies

$$z_n = \frac{1}{b-a}(g((b-a)z_{n-1} + a) - a) \stackrel{\Delta}{=} f(z_{n-1}),$$

and f is a map on the unit interval. Of course, $x[n]$ can be recovered from $z[n]$ and so the properties of x_n can be inferred from those of z_n.

Eventually-expanding maps possess two of the properties of the Markov maps we studied in Chapter 3, namely, the eventually-expanding property and piecewise continuity. Their precise definition is as follows:

Definition 2 *A nonsingular map $f : [0,1] \to [0,1]$ is called eventually-expanding if*

1. *There exists a partition $0 = a_0 < a_1 < \cdots a_N = 1$ of the unit interval such that for each i the restriction of f to the open interval $I_i = (a_{i-1}, a_i)$ is monotonic, continuous and differentiable.*

2. *The function $\frac{1}{|f'(x)|}$ is a function of bounded variation.*

3. *There is a number $\lambda > 1$ and an integer m such that $|\frac{d}{dx} f^m(x)| \geq \lambda$ wherever the derivative exists.*

The bounded variation condition in the definition is essentially a smoothness requirement on the derivative of the map. Among the maps that satisfy this condition are Markov maps, maps with bounded slope, and certain maps with a finite number of points of infinite slope.

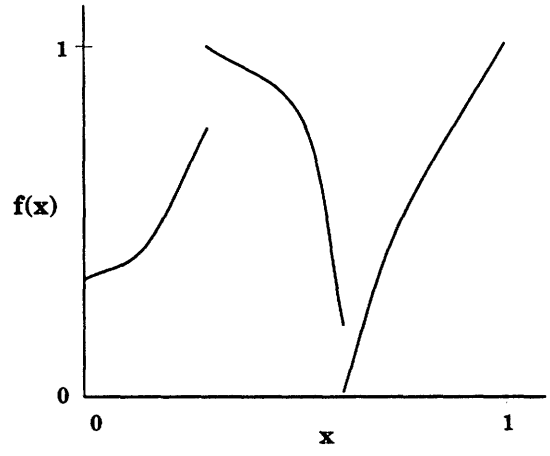


Figure 4-1: The Graph of an Eventually-Expanding Map

An eventually-expanding map can be expressed in the form

$$f(x) = \sum_{i=1}^{N} f_i(x)\chi_i(x), \tag{4.2}$$

where each f_i is monotonic and continuous on the interval I_i and χ_i is the indicator function of the i-th interval. An example of an eventually-expanding map is shown in Figure 4-1. Note in particular, that an eventually-expanding map need not be continuous. Eventually-expanding maps differ from Markov maps in that they need neither be piecewise linear nor map partition points to partition points.

The broad statistical properties of eventually-expanding maps have been well studied [25]. For example, existing results establish conditions for ergodicity and the applicability of central limit theorems. Less attention has been paid to statistics of time series generated by these maps. In this chapter we will extend some of the results of Chapter 3 to eventually-expanding maps. As an intermediate step, the next section presents some previously reported results establishing the existence of invariant densities and the form of the FP operator for eventually-expanding maps.

4.1.1 FP operator and Invariant Densities for Eventually-Expanding Maps

As was the case for Markov maps, the Frobenius–Perron operator plays a central role in understanding the statistical properties of time series generated by eventually-expanding

maps. The FP operator associated with an eventually-expanding map can be derived in the same manner as for a Markov map. In fact, the expression (3.16) for the FP operator of a one-dimensional map is sufficiently general to include all eventually-expanding maps. It is repeated here for convenience:

$$(P_f h)(x) = \sum_{i=1}^{N} \frac{h(f_i^{-1}(x))\chi_{f_i(I_i)}(x)}{|f'(f_i^{-1}(x))|}.$$

As for all nonsingular maps, the invariant density of an eventually-expanding map is a fixed point of the FP operator; a map has an invariant density only when the fixed point exists. In Section 3.3.1, we reasoned that, for Markov maps, such a fixed point exists because of properties of a matrix representation of the FP operator. For more general eventually-expanding maps, the existence of a fixed point and an invariant density must be verified by other means since, in general, there is no matrix representation of the FP operator. The following theorem due to Wong [26] ensures that an eventually-expanding map has an invariant density:

Theorem 4 *Let $f : [0,1] \to [0,1]$ be an eventually-expanding map. Then there exists a density p of bounded variation such that $P_f p = p$, i.e. p is an invariant density. If f has N partition points, then f has at most N such densities.*

Once the existence of an invariant density has been established, the integral on the right side of 4.1 (and hence any correlation statistic we seek) is well defined.

Theorem 4 does not rule out the possibility that more than one absolutely continuous invariant density exists. Certain auxiliary conditions on f guarantee the uniqueness of the invariant density. For example, it can be shown (see [26]) that any map with two partition elements has a unique invariant density. More generally, when f has multiple invariant densities, each is supported on disjoint subsets of the unit interval consisting of finite unions of intervals. When restricted to one of these subsets, f is an ergodic map, i.e., it has a unique invariant density. In this manner, an eventually-expanding map may be decomposed into a finite number ergodic components. Using this decomposition, maps with multiple invariant densities are studied by examining each ergodic component[1]. For

[1]See, for example, [25] for a more detailed treatment of these issues.

this reason, we will restrict attention to ergodic maps.

Now that eventually-expanding maps have been defined, and some of their properties described, we turn to the problem of approximating their statistics.

4.2 Approximate Statistics for Eventually-Expanding Maps

The statistics of a general eventually-expanding map, unlike those of a Markov map, seem to have no simple closed form expressions. Nevertheless, eventually-expanding maps can be analyzed approximately using techniques based on Markov maps. In particular, any eventually-expanding map, along with its correlation statistics, can be approximated to any desired accuracy by a Markov map. Furthermore, the statistics of the approximating Markov map may be computed using the results of Chapter 3. In this manner, the approximate value of statistics of an eventually-expanding map may be determined in closed form.

In this section, we use a particular method of approximating eventually-expanding maps by Markov maps. Although this approximation technique has been previously reported [27], its application to the problem of computing correlation statistics is new. The next subsection presents a description of the approximation method. Subsection 4.2.2 shows how the Markov approximation can be used to compute approximate correlation statistics of an eventually-expanding map.

4.2.1 Markov Approximations to Eventually-Expanding Maps

This section presents a sequence of piecewise linear Markov maps that approximates an arbitrary eventually-expanding map. These Markov approximations form the basis of a technique approximating the statistics of eventually-expanding maps. As a preliminary step, we consider the special case of approximating eventually-expanding Markov maps that are not piecewise linear. This special case provides some insight into the general case.

Suppose f is Markov map that is not piecewise linear. Then, as discussed in Chapter 3, f takes partition points to partition points. That is, there exists a set of partition points $Q = \{a_i\}_{i=0}^{N}$ such that f is monotone on the interval (a_{i-1}, a_i) and $f(Q) \subset Q$. Suppose we denote by f_0 the piecewise linear approximation of f with respect to Q. More precisely, f_0

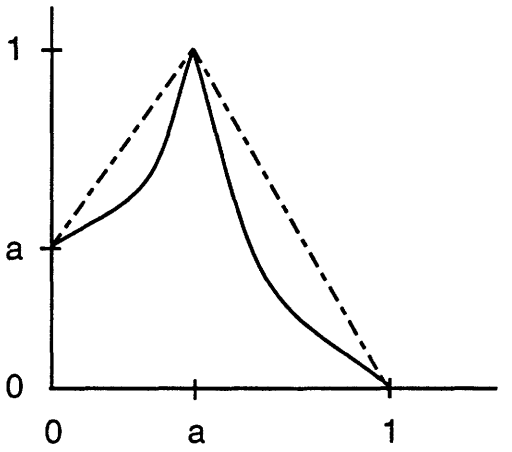


Figure 4-2: A Markov Map and its First Piecewise Linear Approximation

is defined at partition points by

$$\begin{aligned} f_0(a_i^-) &= f(a_i^-) \\ f_0(a_i^+) &= f(a_i^+), \end{aligned}$$

where we define the left and right limiting values since f may be discontinuous. Between partition points , f_0 is defined by linear interpolation, i.e.

$$f_0(x) = \frac{x - a_{i-1}}{a_i - a_{i-1}} f(a_i^-) + \frac{a_i - x}{a_i - a_{i-1}} f(a_{i-1}^+). \tag{4.3}$$

By definition, the map f_0 satisfies $f_0(Q) \subset Q$ and is thus a piecewise linear Markov map. It can be shown (using for example the mean value theorem) that f_0 is also eventually-expanding and hence has a smooth invariant density. An example of a Markov map and its piecewise linear approximation are shown in Figure 4-2.

The map f_0 is one member of a family of piecewise linear Markov maps that approximates f. The family can be determined as follows. Note that if f is a Markov map with respect to the partition points Q, it is also a Markov map with respect to the partition points $Q_1 = Q \cup f^{-1}(Q)$ since $f(Q_1) = f(Q) \cup Q \subset Q \subset Q_1$. We denote by f_1 the piecewise linear approximation of f with respect to Q_1. Then f_1 is a piecewise linear Markov map with respect to Q_1 as well. More generally, we define the partition points $Q_i = Q_{i-1} \cup f^{-1}(Q_{i-1})$.

We denote by f_i the piecewise linear version of f with respect to the partition points Q_i. For each i, the map f_i is a piecewise linear Markov map. This sequence of piecewise linear maps provides an increasingly good approximation to f as i gets large. In fact, using the eventually-expanding property of f, it can be shown [25] that f_i converges to uniformly to f. For the Markov map of Figure 4-2, two subsequent steps of this process are shown in Figure 4-3.

We will show in the next section that not only does the map converge, but, more importantly, the statistics of f_i approach those of f for sufficiently large i. We now generalize this approach to all eventually-expanding maps.

The approximation technique described above, which applies to Markov maps that are not piecewise linear, had two steps. First, a sequence of increasingly fine sets of partition points was defined. Then, a piecewise linear Markov map was defined using each set of partition points. These two steps can be generalized to arbitrary eventually-expanding maps as follows. First, suppose f is an eventually-expanding map with partition points Q. As above, we define the sequence of sets of partition points $Q_i = Q_{i-1} \cup f^{-1}(Q_{i-1})$. Since f need not be Markov, defining the piecewise linear approximations with respect to Q_i requires some care to assure both the eventually-expanding and Markov properties. As above, we define the approximating map at the partition points linearly interpolate. One approach is the following:

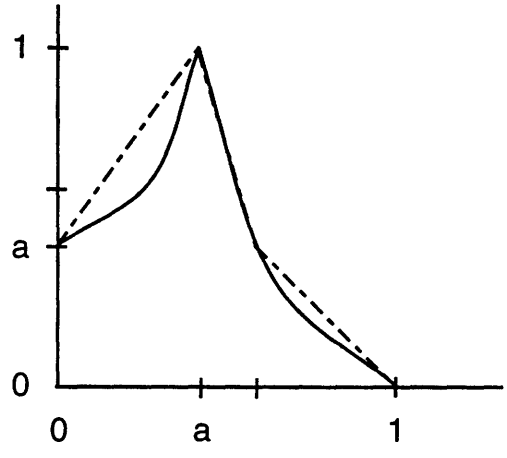
1. For each partition point $q \in Q$ such that f is increasing at q^+, set $f_i(q^+)$ equal to the maximum element of Q_i that is no greater than $f(q^+)$, i.e.

$$f_i(q^+) = \max\{v \in Q_i | v \leq f(q^+)\}.$$

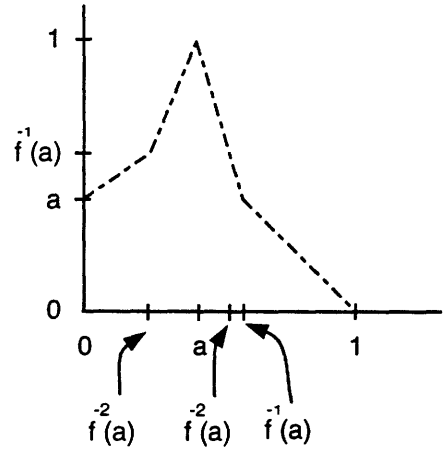
2. For each partition point $q \in Q$ such that f is decreasing at q^+, set $f_k(q^+)$ equal to the minimum element of Q_i that is no less than $f(q^+)$, i.e.

$$f_k(q^+) = \min\{v \in Q_i | v \geq f(q^+)\}.$$

3. For each partition point $q \in Q$ such that f is increasing at q^-, set $f_k(q^-)$ equal to the



(a) The Second Markov Approximation $f_2(x)$



(b) The Third Markov Approximation $f_3(x)$

Figure 4-3: Two Members of a Sequence of Piecewise Linear Markov Approximations to a General Eventually-Expanding Map

minimum element of Q_i that is no less than $f(q^-)$, i.e.

$$f_k(q^-) = \min\{v \in Q_i | v \geq f(q^-)\}.$$

4. For each partition point $q \in Q$ such that f is decreasing at q^-, set $f_k(q^-)$ equal to the maximum element of Q_i that is no greater than $f(q^-)$, i.e.

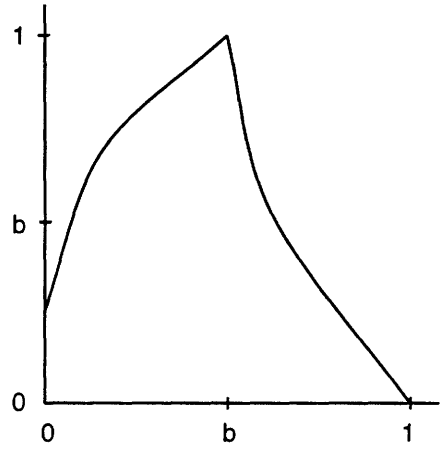
$$f_k(q^-) = \max\{v \in Q_i | v \leq f(q^-)\}.$$

The map f_i is defined at all other points by linear interpolation in the same manner as in (4.3). By construction, the function f_i satisfies $f(Q_i) \subset Q_i$ and is thus not only piecewise linear but also Markov. It has also been constructed to be eventually-expanding and piecewise linear. A few steps in this process applied to a simple non-Markov map are shown in Figure 4-4. Although other sequences of Markov approximations may potentially be applicable in the current context, the method described above will be used exclusively. We therefore refer to the sequence f_i as the sequence of Markov approximations to f with no possibility of ambiguity. That the sequence of approximations has the desired properties is verified in the next section.

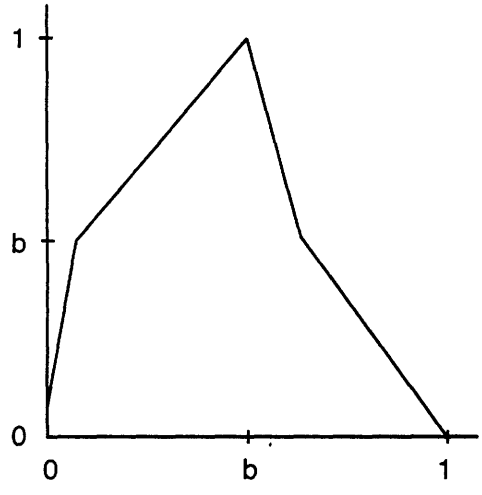
4.2.2 Properties of Markov Approximations

In this section, we show that the sequence of Markov approximations can be used to approximate the correlation statistics of an eventually-expanding map. Previously reported results show that the sequence of Markov approximations to a map f converges uniformly to f [27]. However, uniform convergence is not sufficient to guarantee that the statistics of the Markov approximations converge to the statistics of f. The main result of this section is that correlation statistics of the approximate maps do indeed converge to the appropriate statistics of f. This result represents a substantial generalization of previously reported results, which considered only simple statistics of the form $E(h(x))$ (see e.g. [28]).

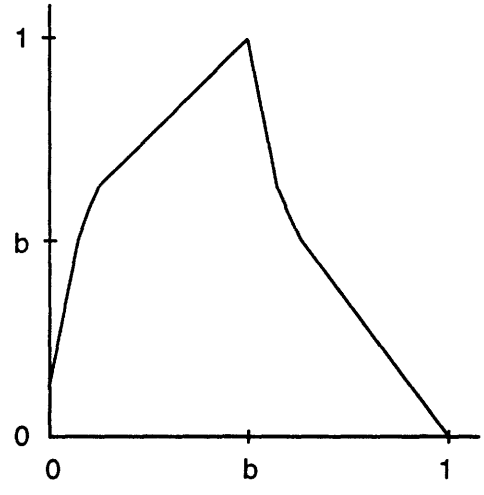
A sequence of Markov approximations to an eventually-expanding map generates a sequence of invariant densities that in turn dictates the manner in which the statistics



(a) A NonMarkov Eventually-Expanding Map



(b) The First Markov Approximation



(c) The Second Markov Approximation

Figure 4-4: An Example of Markov Approximations to a NonMmarkov Map

of the approximations converge. In this context, the notion of weak convergence [15] is particularly useful. More specifically, a sequence of functions $p_i \in L_1$ converges weakly to $p \in L_1$ if

$$\lim_{i\to\infty} \int_{[0,1]} h(x)p_i(x)dx = \int_{[0,1]} h(x)p(x)dx \tag{4.4}$$

for all bounded measurable functions h. Weak convergence has an interpretation in terms of statistics of the form $E(h(x))$. In particular, by construction, each f_i is eventually-expanding and hence has an invariant density p_i. It follows that the following sequence of statistics is well defined:

$$E_i\{h(x)\} = \int_{[0,1]} h(x)p_i(x)dx,$$

where $h(x) \in L_1$. For f_i to be a useful approximation to f, we require that

$$\begin{aligned} \lim_{i\to\infty} E_i\{h(x)\} &= \int_{[0,1]} h(x)p(x)dx \\ &= E\{h(x)\}, \end{aligned}$$

for every $h \in L_1$. Phrased differently, the sequence of invariant densities of f_i must converge weakly to the invariant density of f.

According to the following theorem due to Gora [27], the sequence of invariant densities associated with a sequence of Markov approximations converges weakly.

Theorem 5 *Suppose f is an eventually-expanding map with invariant density p, and f_i is the sequence of Markov approximation to f. If f_i has invariant density p_i, then $f_i \to f$ uniformly and $p_i \to p$ weakly.*

In light of the previous discussion, this theorem may be interpreted as an approximation result. Namely, f_i approximates the statistical properties of f to the extent that expected values computed with respect to the invariant densities of f_i converge to expected values with respect to the invariant density of f.

Theorem 5 guarantees that $E\{h(x)\}$ can be well approximated by $E_i\{h(x)\}$ for sufficiently large i. This fact suggests a a three step method for approximating statistics of the form $E\{h(x)\}$ for eventually-expanding maps. First, choose i sufficiently large so that the Markov approximation f_i provides a good approximation to f. Next, use the tech-

niques of Chapter 3 to determine the invariant density p_i. Finally, compute the integral $E_i\{h(x)\} = \int_{[0,1]} h(x)p_i(x)dx$. This procedure is straightforward to apply and requires only that the invariant density of f_i be determined. It has been applied to the problem of approximating the Lyapunov exponent of a one-dimensional map [28].

We will show in the sequel that the approximation technique outlined above can be applied to compute correlation statistics. In particular, we will verify that correlation statistics of the sequence of approximations satisfy

$$\lim_{i\to\infty} R_{f_i,h_0,\ldots,h_l}[k_1,\ldots,k_r] = R_{f,h_0,\ldots,h_l}[k_1,\ldots,k_r],$$

or, in more detail,

$$\lim_{i\to\infty} \int_{[0,1]} h_0(x)h_1(f_i^{k_1}(x))\cdots h_r(f_i^{k_r}(x))p_i(x)dx = \int_{[0,1]} h_0(x)h_1(f^{k_1}(x))\cdots h_r(f^{k_r}(x))p(x)dx. \tag{4.5}$$

Comparing (4.5) and (4.4) we see that weak convergence of p_i does not immediately guarantee the convergence of the quantities (4.5)—both the invariant density and the remainder of the integrand of (4.5) depend on the index i. In order to verify that correlation statistics converge to the correct values, we introduce a stronger notion of convergence.

We define *statistical convergence* to occur when a sequence of maps and its correlation statistics converge to the desired values.

Definition 3 *Suppose an eventually-expanding map f has a unique invariant density p. A sequence of maps $\{f_i\}$ statistically converges to f if each f_i has a unique invariant density p_i and*

$$R_{f_n,h_0,\ldots,h_l}[k_1,\ldots,k_r] \stackrel{n\to\infty}{\to} R_{f,h_0,\ldots,h_l}[k_1,\ldots,k_r]$$

for any continuous h_i and all nonnegative integers k_i and r.

Note that statistical convergence depends directly on both the sequence of maps f_i and the sequence of invariant densities.

The following theorem, which is proved in Appendix A, asserts that the sequence of Markov approximations to an eventually-expanding map statistically converges.

Theorem 6 *Suppose f is an eventually-expanding map. Then sequence of piecewise linear Markov approximations statistically converges to f.*

Theorem 6 implies that the three step approach described above may be used to approximate the correlation statistics of any eventually-expanding map.

Theorem 6 states that an eventually-expanding map is the limit of a sequence of Markov maps in a very specific sense. We will see in the next section that an alternative interpretation of the theorem is that the FP operator of a eventually-expanding map is the limit of a sequence sequence of FP operators associated with the Markov approximations. We now explore this interpretation and some of its consequences.

4.3 Power Spectra of Eventually-Expanding Maps

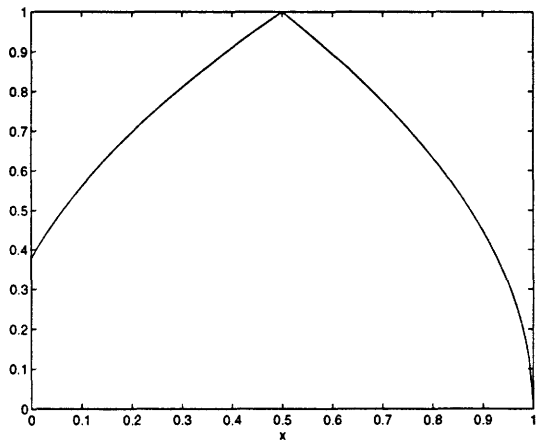
In this section, we consider some properties of the FP operator and the insight it provides into the broad structure of statistics of eventually-expanding maps. For concreteness, we concentrate on the autocorrelation sequence and power spectra associated with these maps. We begin by presenting a concrete example illustrating how the techniques of the previous section can be used to approximate the power spectrum of a non-Markov map.

Consider the map

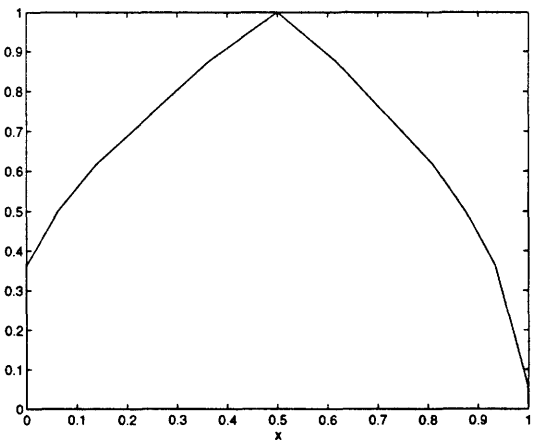
$$f(x) = \begin{cases} (\frac{12}{7}x + \frac{1}{7})^{\frac{1}{2}} & x \leq \frac{1}{2} \\ (2(1-x))^{\frac{1}{2}} & x > \frac{1}{2} \end{cases} \tag{4.6}$$

with graph as in Figure 4-5 (a). We shall approximate this map by f_3, the third in the sequence of Markov approximations to f. By directly applying the method described above, the approximate map is determined to be of the form shown in Table 4.1 The graph of $f(x)$ is shown in Figure 4-5 (b).

Once an approximating map is chosen, the techniques of Chapter 3 may be used to determine a closed form expression for the power spectrum. The analytical power spectrum computed from the approximate map is shown in Figure 4-6 along with an empirically determined power spectrum computed from a 50000 sample time series by periodogram averaging with a length 256 window. It is evident that the agreement between analytical and empirical spectra is good. There are, however, no results available to bound the ap-



(a) The Map of Equation (4.6)



(b) The Third Piecewise Linear Markov Approximation of Equation (4.6)

Figure 4-5: An Example of Approximating an Eventually-Expanding Map with a Piecewise-linear Markov Map

$f_3(x)$	Range
$\frac{35}{16}x + \frac{93}{256}$	$0 < x \leq \frac{1}{16}$
$\frac{1536}{1001}x + \frac{809}{2002}$	$\frac{1}{16} < x \leq \frac{9101}{65536}$
$\frac{1536}{1337}x + \frac{4895}{10696}$	$\frac{9101}{65536} < x \leq \frac{93}{256}$
$\frac{6144}{6713}x + \frac{29135}{53704}$	$\frac{93}{256} < x \leq \frac{521901}{1048576}$
$\frac{2048}{2387}x + \frac{1363}{2387}$	$\frac{521901}{1048576} < x \leq \frac{1}{2}$
$-\frac{1024}{1023}x + \frac{1535}{1023}$	$\frac{1}{2} < x \leq \frac{263167}{524288}$
$-\frac{1024}{959}x + \frac{11769}{7672}$	$\frac{263167}{524288} < x \leq \frac{79}{128}$
$-\frac{256}{191}x + \frac{2601}{1528}$	$\frac{79}{128} < x \leq \frac{26527}{32768}$
$-\frac{256}{143}x + \frac{591}{286}$	$\frac{26527}{32768} < x \leq \frac{7}{8}$
$-\frac{512}{221}x + \frac{1117}{442}$	$\frac{7}{8} < x \leq \frac{122423}{13172}$
$-\frac{512}{109}x + \frac{8285}{1744}$	$\frac{122423}{13172} < x \leq \frac{511}{512}$
$-32x + 32$	$\frac{511}{512} < x \leq 1$

Table 4.1: The Definition of the Third Markov Approximation to the Map of Equation (4.6)

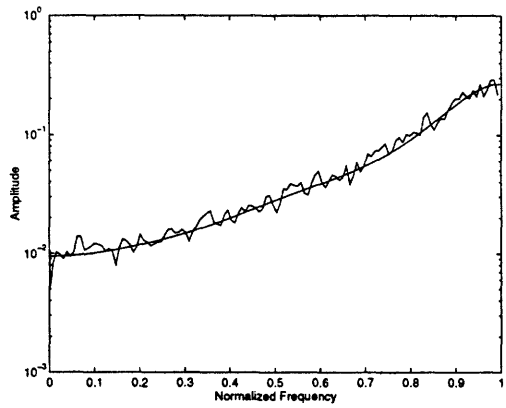


Figure 4-6: A Comparison of the Analytically Computed Spectrum of the Map of Figure 4-5 (b) and an the Empirically Spectrum of the Map of Figure 4-5 (a)

proximation error. Thus, choosing the order of the approximation becomes a matter of trial and error. More work is necessary in order to address this clearly important issue.

The results of the previous section guarantee that any eventually-expanding map is the statistical limit of a sequence of Markov maps; an eventually-expanding map's statistics are thus close to the statistics of a Markov map. It follows that the approach used in the previous example may be applied to compute any correlation statistic of an eventually-expanding map. Beyond its practical use in computing numerical values of statistics, this approximation result also illuminates some of the more general statistical characteristics of eventually-expanding maps. In particular, the approximation procedure outlined in Section 4.2.1 may be viewed as a method of approximating the FP operator of an eventually-expanding map f by that of a Markov map; once a Markov approximation is chosen its FP operator is used to compute the statistics of f. Given the close relationship of the FP operators of eventually-expanding maps and Markov maps, it should not be surprising that their statistical properties are similar.

The structural and computation results presented above hold for the entire class of eventually-expanding maps. We have not shown how eventually-expanding maps may arise in engineering practice. The next chapter presents a map that models the sampled inductor current in a switching power supply and shows how this model can be used to compute the statistics of the continuous inductor current.

Chapter 5

Chaos in a Switching Power Converter

The previous chapters have presented techniques with which to analyze a large class of one-dimensional systems. In this chapter we present an example of an engineering system that is well described by such a chaotic system. In particular, we analyze the operation of an idealized model of a switching DC-DC power converter shown in Figure 5-1. We shall see that in certain of its regimes of operation, the converter is quite well described by an eventually expanding one-dimensional map of the type analyzed in Chapter 4. Once this dynamical model is established, the techniques previously described may be used to determine its statistical properties. Using this approach, certain time-average properties of the converter may be determined.

Despite the variety of types of switching power converter implementations, e.g., direct and indirect converter topologies operating in continuous or discontinuous conduction mode, the same basic approach has typically been applied to modeling them (see e.g. [32, 23] and the references therein). In particular, because the power converters of interest are periodically switched systems, they give rise to different models corresponding to each state of the switch. By incorporating knowledge of the switching strategy, an overall solution can be constructed. We shall use a similar approach, however, our results differ in two ways from those previously reported. In particular, unlike the results of Tse [32], which apply only

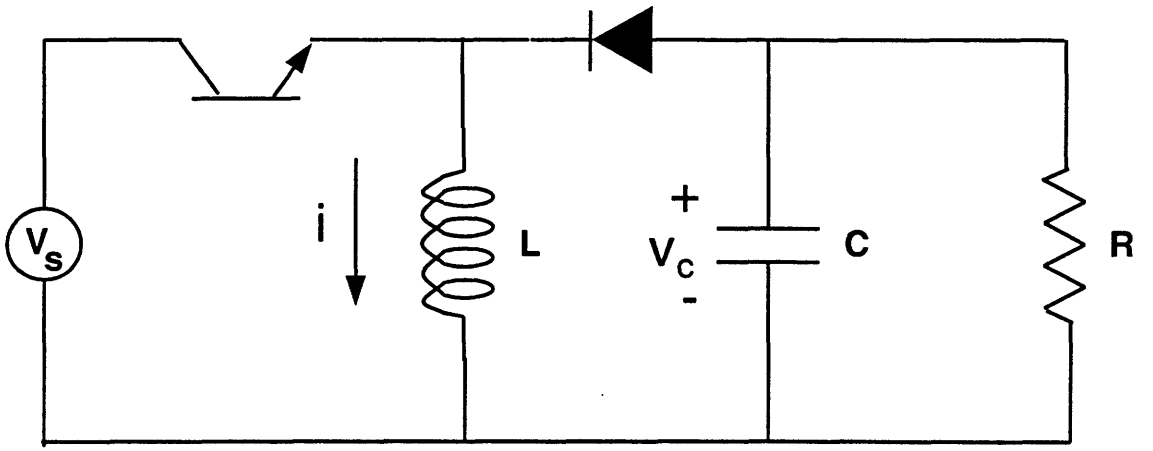


Figure 5-1: The Circuit Model of a Buck/Boost Switching Power Converter

to converters operating in discontinuous conduction mode that can be exactly described by a one-dimensional model, we will consider simple converters operating in continuous conduction mode, which are inherently two-dimensional systems. Furthermore, our model describes the evolution of the sampled inductor current and sampled capacitor voltage from one switching time to the next, in synchrony with the underlying switching clock . In distinction to the model proposed by Deane [23], which models an asynchronous current and voltage evolution, our synchronous model allows a straightforward analysis of the statistics of the converter waveforms.

The chapter is organized as follows. Section 5.1 derives a map that describes the cycle-to-cycle operation of the converter under current-mode, control with no approximations aside from the assumption of ideal circuit elements. This model has two state variables, as might be expected since the circuit has two energy storing elements. However, under the reasonable assumption of small output voltage ripple, the converter can be well described by a one-dimensional model. Section 5.2 presents this simplified model and describes some of its properties. Finally, Section 5.3 presents an analysis of the converter.

5.1 Derivation of the Map

We shall analyze the buck/boost converter of Figure 5-1. The circuit is assumed to operate in continuous conduction mode, that is, the inductor current i is always strictly positive.

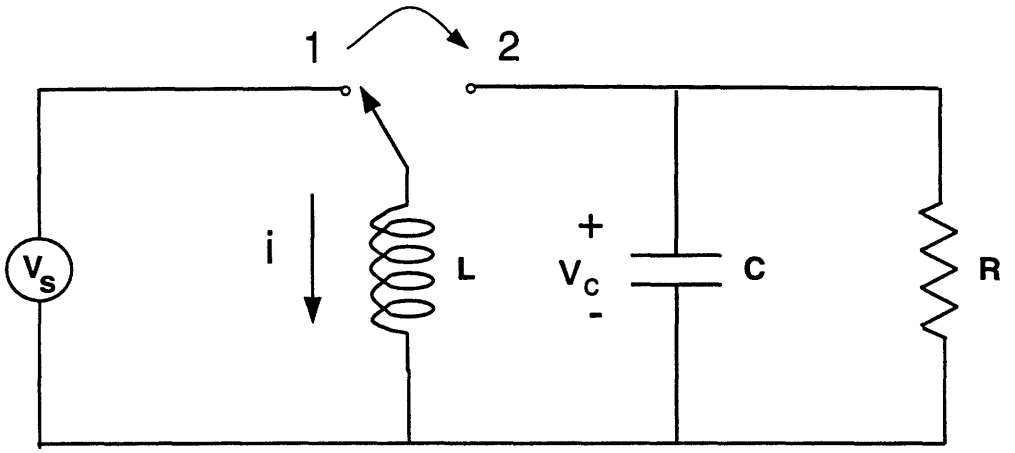


Figure 5-2: An Equivalent Buck/Boost Converter with and Idealized Switch

Furthermore, the transistor and diode are assumed to operate as ideal switches, in other words, i.e., they are either fully conducting (on) or fully non-conducting (off). With these assumptions, the transistor and diode act as single-pole double-throw switch in that only one of the transistor/diode pair will conduct at any time. A circuit with an explicitly represented switch is shown in Figure 5-2.

With the switch in position 1, the inductor current increases linearly and the load current equals the negative capacitor current. With the switch in position 2, the load current is the sum of the inductor and capacitor currents. The load current can be regulated by controlling the switching between position 1 and position 2. Thus, the converter's operation depends on both the topology of the circuit and switching strategy employed. The model developed below describes the behavior of a converter operating under current-mode control. Using this control strategy, the switch is thrown into position 1 periodically with period T. The switch is put into position 2 only when the inductor current reaches a peak value, which we denote I_p. The switch is thrown back into position 1 when the next switching time arrives. Thus, the peak inductor current, I_p, is the control parameter.

We will see below that when the converter operates under current-mode control, its voltage and current waveforms need not be periodic. In fact, the switch may stay in position 1 through multiple periods while, it can remain in position 2 for at most one period. This is in contrast to typical duty ratio control schemes (see, e.g., [33]) in which the switch assumes each position during every cycle.

With the switch fixed in either position, the resulting circuit is linear and time invariant, and can be analyzed using standard techniques from circuit theory. The details of the switching strategy dictate how these two solutions are pieced together into a complete model of the converter. We begin by solving the circuit for each switch position.

With the switch is in position 1, that is, with the transistor fully conducting, the model behaves like two uncoupled first order circuits described by the equations

$$\frac{d}{dt}\begin{bmatrix} i \\ v_c \end{bmatrix} = \begin{bmatrix} \frac{V_s}{L} i \\ \frac{v_c}{RC} \end{bmatrix} \tag{5.1}$$

The solutions of these two equations are

$$\begin{bmatrix} i(t) \\ v_c(t) \end{bmatrix} = \begin{bmatrix} i(0) + \frac{V_s}{L} \\ v_c(0) e^{-\frac{t}{RC}} \end{bmatrix}. \tag{5.2}$$

When the switch is in position 2, that is when the diode is conducting, the model behaves like an undriven RLC circuit governed by the state equations

$$\begin{aligned} \frac{d}{dt}\begin{bmatrix} i \\ v_c \end{bmatrix} &= \begin{bmatrix} 0 & \frac{1}{L} \\ -\frac{1}{C} & -\frac{1}{RC} \end{bmatrix} \begin{bmatrix} i \\ v_c \end{bmatrix} \\ &= \mathbf{A} \begin{bmatrix} i \\ v_c \end{bmatrix}. \end{aligned}$$

The solution to these state equations is

$$\begin{bmatrix} i(t) \\ v_c(t) \end{bmatrix} = e^{\mathbf{A}t} \begin{bmatrix} i(0) \\ v_c(0) \end{bmatrix}. \tag{5.3}$$

For convenience, we will model the evolution of the inductor current and capacitor voltage sampled at consecutive switching times. As we will see, the continuous voltage and current waveforms can be recovered from these sampled waveforms.

We first consider how the inductor current evolves from time nT to time $(n+1)T$. When the inductor current at the beginning of the switching interval satisfies $i(nT) < I_p - \frac{V_s}{L}T$,

the current $i(t)$ does not reach the peak value during the interval $nT \leq t \leq (n+1)T$. The transistor therefore conducts throughout the period and, according to (5.2), the state variables at times $t = (n+1)T$ and $t = nT$ are related by

$$\begin{bmatrix} i((n+1)T) \\ v_c((n+1)T) \end{bmatrix} = \begin{bmatrix} i(nT) + \frac{V_s}{L}T \\ v_c(nT)e^{-\frac{T}{RC}} \end{bmatrix}$$

On the other hand, if $i(nT) \geq I_p - \frac{V_s}{L}T$ the inductor current will reach its peak value sometime in the interval $nT \leq t \leq (n+1)T$, at which time the transistor will cut off. Referring again to Equation (5.2), the peak current is reached at time $t' = nT + \frac{L}{V_s}(I_p - i(nT))$. At this time, the inductor current has value I_p and the capacitor voltage has value

$$v_c(t') = v_c(nT)e^{-\frac{t'-nT}{RC}}.$$

At time t' the transistor switches off and the circuit is governed by (5.3). Thus, at time $t = (n+1)T$ the state variables have values

$$\begin{bmatrix} i((n+1)T) \\ v_c((n+1)T) \end{bmatrix} = e^{\mathbf{A}((n+1)T-t')} \begin{bmatrix} I_p \\ v_c(nT)e^{\frac{1}{RC}(t'-nT)} \end{bmatrix}$$

By substituting the expression for t', the state variables at any switching time can be expressed in terms of their values at the previous switching time. In particular, the state variables evolve according to

$$\begin{bmatrix} i((n+1)T) \\ v_c((n+1)T) \end{bmatrix} = \begin{cases} \begin{bmatrix} i(nT) + \frac{V_s}{L}T \\ v_c(nT)e^{-\frac{T}{RC}} \end{bmatrix} & i(nT) < I_p - \frac{V_s}{L}T \\ e^{\mathbf{A}(T-\frac{L}{V_s}(I_p-i(nT)))} \begin{bmatrix} I_p \\ v_c(nT)e^{\frac{1}{RC}(\frac{L}{V_s}(I_p-i(nT)))} \end{bmatrix} & i(nT) \geq I_p - \frac{V_s}{L}T \end{cases} \tag{5.4}$$

The sampled-data state equations (5.4) are nonlinear. As such, their behavior is difficult to determine directly from their form. However, computer simulations of (5.4) for different values of the circuit parameters suggest that sampled waveforms in the converter may

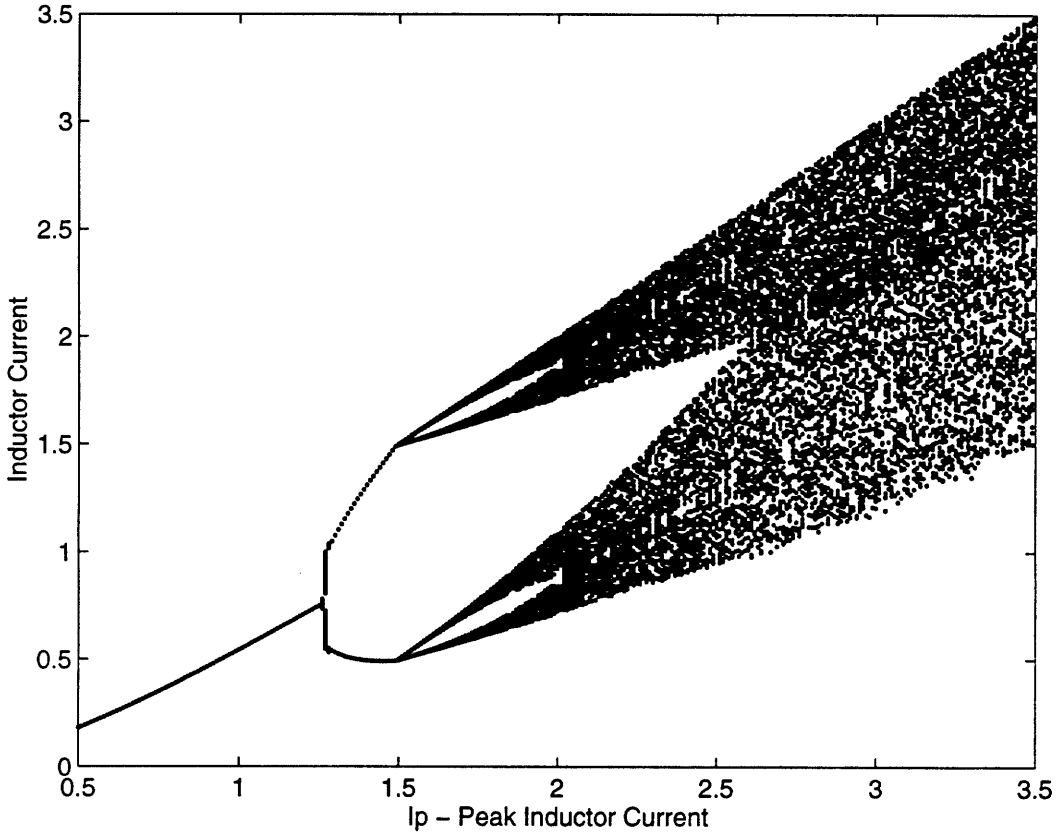


Figure 5-3: Bifurcation Diagram of the Two-Dimensional Model as Computed by Sweeping the Peak Current

display three types of steady state behavior: constant, periodic, or aperiodic.

When the sampled-data state variables are constant or periodic, the continuous state variable are periodic. Traditionally, switching power converters have been designed to operate in this periodic regime. As a consequence, the analysis and design techniques developed for these converters are typically applicable only in a periodic regime of operation [33].

The bifurcation diagram of Figure 5-3 shows all three types of behavior. For all of the simulations used to construct the figure, the circuit parameters in the model were fixed at $V_s = 10V$, $T = 100\mu s$, $L = 1mH$, $C = 100\mu F$, and $R = 20\Omega$. The diagram was generated by sweeping the peak current from $I_p = 0.5A$ to $I_p = 3.5A$ in increments of $.01A$. For each value of I_p the model (5.4) was iterated through 1000 switching periods and the final 100 values of the inductor current were plotted on the vertical axis. By discarding the initial 900

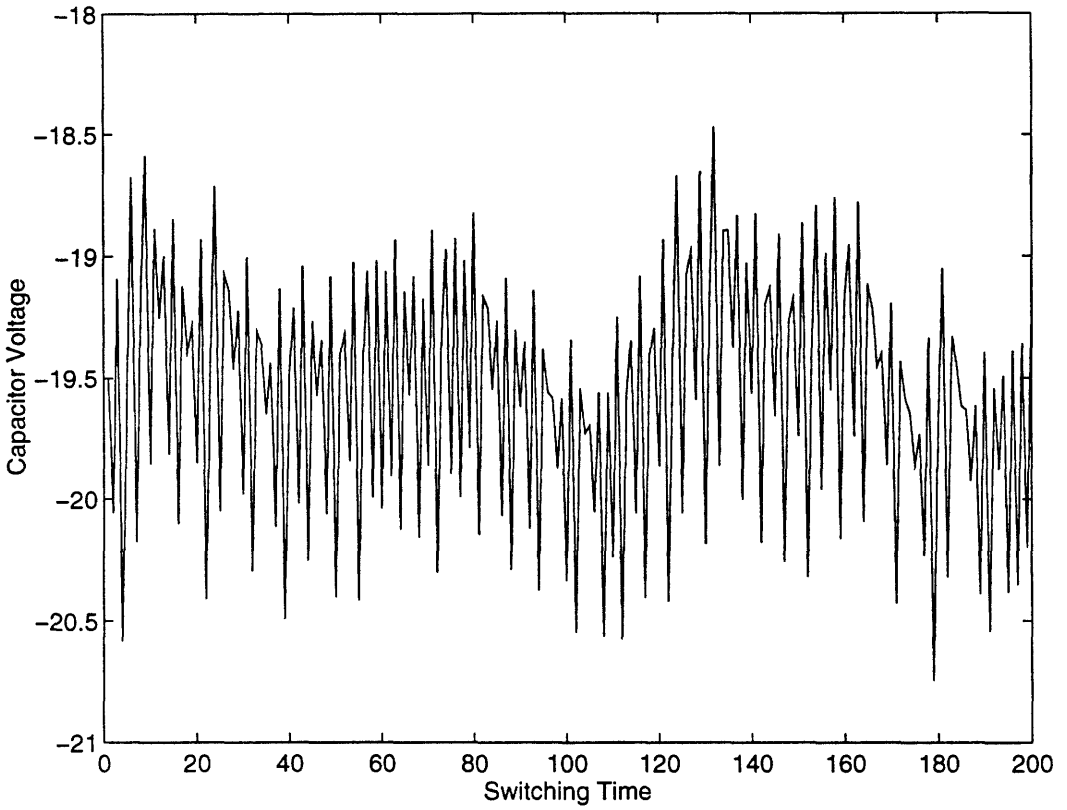


Figure 5-4: Sampled Capacitor Voltage Computed From Equation (5.4)

points, transient behavior was ignored. We can see for instance that the sampled current settles into a fixed point for I_p less than about $1.25A$. The sampled current is periodic with period 2 for I_p in the range of about $1.3A$ to $1.5A$. As the peak current increases, we see behavior that appears to be chaotic.

An example of the sampled capacitor voltage is shown in Figure 5-4 for the same circuit parameters as above and a peak inductor current of $I_p = 3.5A$. Although the sampled output voltage does not appear to be periodic, it does have a manageable ripple. This suggests the possibility that these power converters may be practically operated in a chaotic regime. In fact, the broadened power spectrum of the waveforms in the converter may have certain benefits.

Figure 5-5 shows the relation between consecutive values of the sampled inductor current for the same circuit setup as above. Again, the peak current I_p is set to $3.5A$ and the data were generated by iterating (5.4). It is evident that there is a strong relation

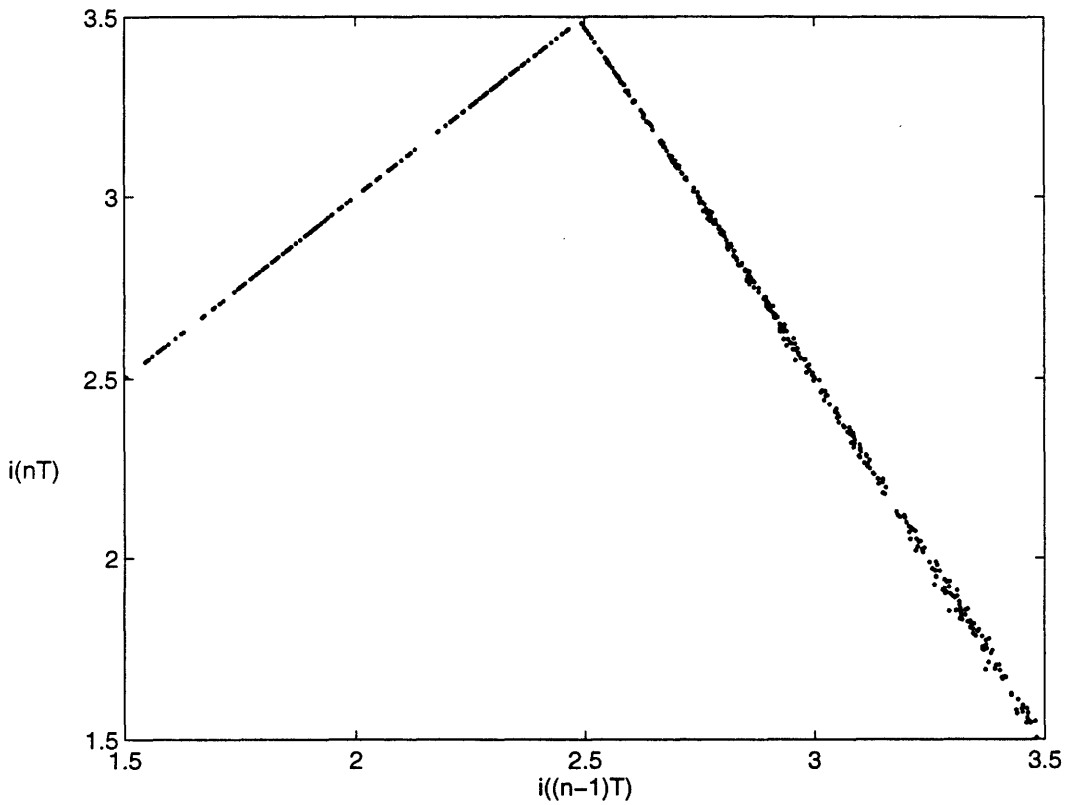


Figure 5-5: Relation Between Consecutive Samples of the Sampled Inductor Current

between consecutive samples of the current that would appear to be well modeled by a one-dimensional map. In the next section, we determine the form of a one-dimensional map that approximates Figure 5-5.

5.2 The One-Dimensional Model

Consider the operation of the converter when the transistor is not conducting. When the output voltage ripple is small, the capacitor voltage $v_c(t)$ is nearly constant, say $\bar{v}_c$. In our specific situation, such an approximation could be expected to hold when $\frac{T}{RC}$ and $\frac{T}{\sqrt{LC}}$ are small.

Reexamining (5.4) under the assumption that $\frac{T}{RC}$ is negligible and that T is small, we make the following approximations:

$$e^{\mathbf{A}(T-\frac{L}{V_s}(I_p-i(nT)))} \approx \mathbf{I} + \mathbf{A}(T - \frac{L}{V_s}(I_p - i(nT)))$$

$$e^{\frac{1}{RC}(\frac{L}{V_s}(I_p-i(nT)))} \approx 1$$

$$e^{\frac{T}{RC}} \approx 1$$

With these approximations, the component of Equation (5.4) describing the evolution of the inductor current reads:

$$i((n+1)T) = \begin{cases} i(nT) + \frac{V_s}{L}T & i(nT) \leq I_p - \frac{V_s}{L}T \\ I_p - \frac{\bar{v}_c}{L}(T - \frac{L}{V_s}(I_p) - i(nT) & i(nT) > I_p - \frac{V_s}{L}T \end{cases} \tag{5.5}$$

Intuitively speaking, Equation (5.5) reflects the fact that with the transistor conducting, the inductor current increases linearly from its initial value with slope $\frac{V_s}{L}$. When the peak current is attained during a switching period, it occurs at time t', subsequent to which the inductor current decreases linearly with slope $\frac{\bar{v}_c}{L}$ until the next switching time, which occurs $T - \frac{L}{V_s}(I_p - i(nT))$ seconds after the peak current is achieved.

It is convenient to isolate some of the important quantities associated with the operation of the circuit through the following definitions:

- The maximum inductor current increase per period when the transistor is conducting is defined as $m_1 = \frac{V_s T}{L}$.
- The maximum current decrease per period is defined as $m_2 = \frac{\bar{v}_c T}{L}$.
- The ratio of increase to decrease is defined $\alpha = \frac{m_1}{m_2}$.
- The scaled peak current is defined $\beta = \frac{I_p}{m_1}$.

It is also convenient for what follows to rewrite the dynamics for the scaled inductor current $y_n = \frac{i(nT)}{m_1}$. Substituting the variables defined above into Equation (5.5) yields that the scaled inductor current evolves according to

$$y_{n+1} = f(y_n), \tag{5.6}$$

where f is of the form

$$f(y) = \begin{cases} y+1 & y < \beta - 1 \\ \beta - \frac{1}{\alpha}(1 - \beta + y) & y > \beta - 1 \end{cases}. \tag{5.7}$$

The graph of the function $f(y)$ is shown in Figure 5-6. Its maximum value at occurs at $y = \beta - 1$ where $f(\beta - 1) = \beta$. According to this model, the sampled inductor current may become negative for certain combinations of α and β. Recall however that the model was derived assuming that the converter operates in continuous conduction mode, i.e., that the inductor current is always positive. The model may be restricted to parameter values such that positive inductor currents will always map to positive currents as follows. From the graph of Figure 5-6 it can be seen that a positive value of the current may map to a negative value only when $f(\beta) < 0$. Since $f(\beta) = \beta - \frac{1}{\alpha}$, we may eliminate the possibility of negative currents by considering only parameter satisfying $\beta > \frac{1}{\alpha}$.

The map f has a fixed point at $y = \frac{\alpha}{\alpha - 1}(\beta - \frac{1-\beta}{\alpha})$ The slope of f at this fixed point is $\frac{df}{dx} = -\frac{1}{\alpha}$ and thus the fixed point is stable when $\frac{1}{\alpha} < 1$. Referring back to the definition of

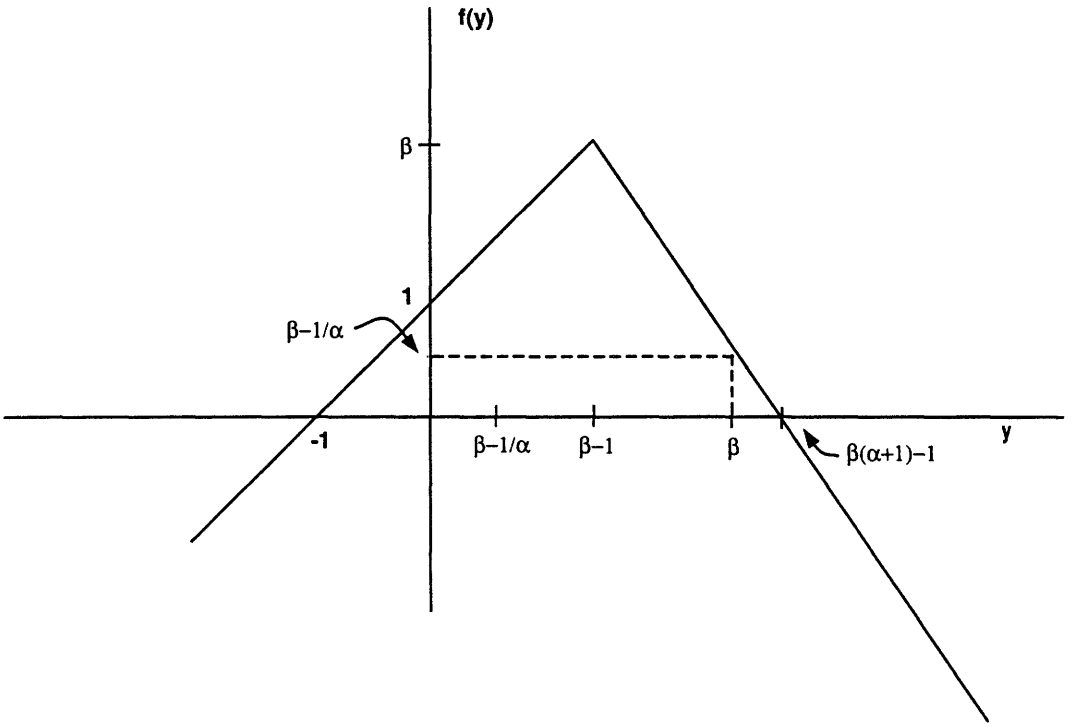


Figure 5-6: A Graph of the Scaled Inductor Current Map (5.7) for $\frac{1}{\alpha} > 1$

α, this condition may be interpreted as saying that the inductor current may converge to a stable fixed point when the maximum inductor current decrease during a switching period is greater than the maximum inductor current increase during a switching period. When $\frac{1}{\alpha} > 1$, the fixed point is unstable. We will see below that in this case the sampled inductor current may be chaotic.

Recall that the model f is valid only for parameter values satisfying $\beta > \frac{1}{\alpha}$. Under this condition, the map of Figure 5-6 has several features that make its steady state behavior more transparent. First, any initial condition in the interval $I_1 = [\beta - \frac{1}{\alpha}, \beta]$ stays in I_1 under iteration of f. Second, any initial condition in the interval $I_2 = [0, \beta - \frac{1}{\alpha})$ eventually leaves I_2 since all values of $y \in I_2$ map to $f(y) = y + 1$. Finally, initial conditions in the interval $I_3 = (\beta, \beta(\alpha + 1) - 1]$ leave I_3 in one step and enter I_2. These three features of the map together imply that for all positive initial inductor current values, the sampled inductor current eventually ends up taking values only in the interval I_1 after a sufficiently large number of iterations of f. Thus, positive initial conditions outside of I_1 may start out

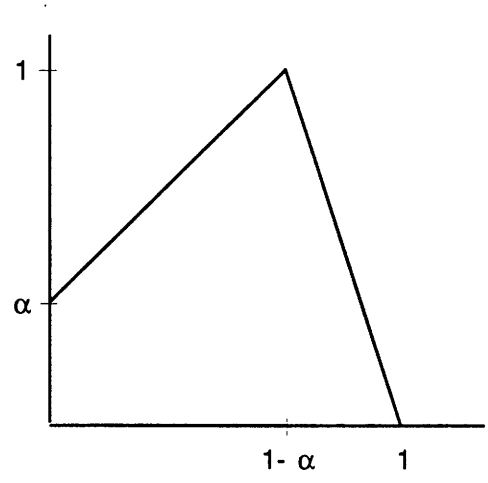


Figure 5-7: Graph of Modified Inductor Current Dynamics (5.10)

with some transient behavior, but the only interesting steady state behavior occurs in the interval I_1. We are only be interested in steady state behavior and so will concentrate on f restricted to the interval I_1.

In order to isolate the behavior of the map on I_1, we consider the scaled and shifted variable

$$z_n = \frac{y_n - (\beta - \frac{1}{\alpha})}{\frac{1}{\alpha}}. \tag{5.8}$$

Substituting (5.8) into (5.6) and (5.7), the dynamics governing the evolution of z are found to be

$$z_{n+1} = g(z_n), \tag{5.9}$$

where g is of the form

$$g(z) = \begin{cases} z + \alpha & z < 1 - \alpha \\ \frac{1}{\alpha}(1 - z) & z \geq 1 - \alpha \end{cases}. \tag{5.10}$$

The graph of g is shown in Figure 5-7. Comparing Figure 5-7 with the scatter diagram of sampled inductor currents of Figure 5-5, a strong qualitative similarity can be seen. Both plots can be brought into close quantitative agreement by accounting for the scale factor and translation inherent in the definition of z_n.

In particular, the unscaled sampled inductor current $i(nT)$ can be determined from z_n

by reversing the previously described transformations. Specifically,

$$i(nT) = I_p - m_2 + m_2 z_n.$$

It follows that properties of the sampled inductor current may be determined from those of of z_n and g, since $i(nT)$ is simply a scaled and shifted version of z_n.

The bifurcation diagram shown in Figure 5-8 gives an indication of the range of behavior displayed by the one-dimensional model of (5.9) and (5.10) as the value of $\frac{1}{\alpha}$ is varied in the range $0.8 \leq \frac{1}{\alpha} \leq 3$. By comparing Figure 5-8 with the bifurcation diagram of the more complex model of Figure 5-3, we can get a sense of how well the one-state-variable map mimics the behavior of the two-state-variable model. Although the diagrams plot differently scaled variables, their qualitative properties are similar. Namely, for small values of the bifurcation parameter, fixed point behavior is displayed by both models. For large parameter values, the behavior appears to be chaotic, and both models appear to take all values in a certain interval. The primary qualitative difference between the two diagrams occurs in the transition from fixed point behavior to chaotic behavior. In the case of the more accurate two-state-variable model, we see a period doubling before chaos develops. The one-state-variable model transitions directly from fixed point behavior to chaos. Since the primary interest of this section is chaotic behavior in the power converter, this discrepancy may not be of great significance.

5.3 Some Properties of the One-State-Variable Model

The derivative of the map g of (5.10) has one discontinuity at $z = \beta - 1$. The slope takes on only two different values: it is either 1 or $\frac{1}{\alpha}$. When $\frac{1}{\alpha} > 1$, the map is eventually expanding and the theory of eventually expanding maps can be applied to the analysis of its properties. In the context of a specific map, namely g, the general and abstract results of the theory may be given physical interpretations. We shall see in this section that the map g has no stable periodic orbits and that it is ergodic and relate these properties to the operation of the power converter. We shall also see that when the parameter takes the value $\alpha = \frac{1}{2}$, the map g is a piecewise linear Markov map and its operation has a straightforward

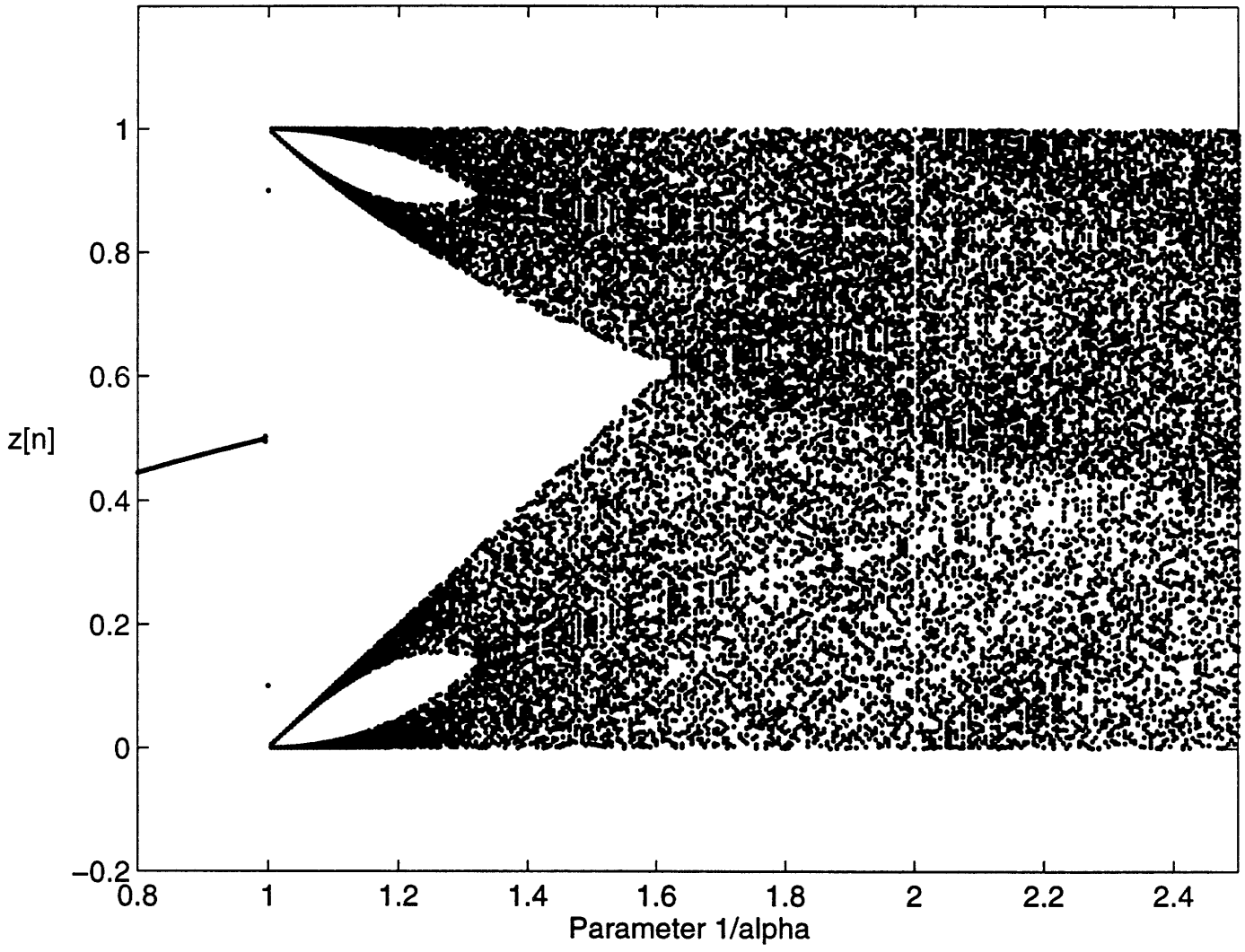


Figure 5-8: Bifurcation Diagram Associated With the One-Dimensional Scaled Inductor Current Model

interpretation.

Intuitively speaking, eventually expanding maps have no stable periodic orbits because of sensitive dependence on initial conditions; points near periodic point generate orbits that diverge from periodicity. More precisely, consider a periodic point z with period p. This means that p is the smallest positive integer such that $g^p(z) = z$. If z is a stable periodic point then the magnitude of the derivative of the single period map must be strictly less than one, i.e. $|\frac{d}{du}g^p(u)|_{u=z}| < 1$. However, since g is eventually expanding there is an integer m such that $|\frac{d}{du}g^m(u)|_{u=z}| > 1$. If we choose some iterate of the period p map such that $kp > m$, stability would imply that $|\frac{d}{du}g^{kp}(u)|_{u=z}| < 1$ while the eventually expanding property requires that $|\frac{d}{du}g^{kp}(u)|_{u=z}| > 1$. Thus, stable periodic orbits are incompatible with the eventually expanding property.

Because the converter is a nonlinear circuit, there is no *a priori* reason to believe that its qualitative behavior should be independent of the initial current in the inductor, or insensitive to small inductor current perturbations caused by external influences. However, to the extent that the one-state-variable model is an accurate model of the operation of the converter, a converter designed to operate in a chaotic regime may not slip into periodic behavior.

The previous argument does not rule out the possibility that the same circuit may display qualitatively different types of chaotic behavior, depending on its initial state. The scenario in which the same circuit operates in several different regimes may be inconvenient in practical applications. It turns out that when g is eventually expanding, it can have only one regime of operation. More specifically, Theorem 4 cited above guarantees that g is ergodic in the sense that it has only one absolutely continuous invariant density. This implies that almost all initial conditions lead to behavior that is qualitatively the same. In particular, the average properties of the sampled inductor current in the steady state are independent of initial condition.

When $\alpha = 2$, the map g is of the form

$$g(z) = \begin{cases} z + \frac{1}{2} & 0 \le z < \frac{1}{2} \\ 2(1-z) & \frac{1}{2} \le z \le 1 \end{cases}, \tag{5.11}$$

and is a member of the family of Markov maps described in the example of Section 3.2. Some insight into its operation may be gained by expressing the initial condition in binary form, i.e. $z = .b_1b_2b_3\ldots$. To remove any ambiguity in the binary expansion, we introduce the convention that when a number has two binary expansions, we choose the terminating one. With this notation, the conditions $0 \leq z < \frac{1}{2}$ and $\frac{1}{2} \leq z \leq 1$ correspond to the conditions $b_1 = 0$ and $b_1 = 1$ respectively. In terms of the binary expansions, the operation of g can be expressed in the form

$$g(.b_1b_2b_3\ldots) = \begin{cases} .1b_2b_3\ldots & \text{for } b_1 = 0 \\ .\bar{b}_2\bar{b}_3\bar{b}_4\ldots & \text{for } b_1 = 1 \end{cases}, \tag{5.12}$$

where $\bar{b}_i = 1 - b_i$.

Equation 5.12 may be used to get an idea of the relative frequency with which z_n takes on certain values. The binary expansion of $g(z)$ will have a 1 as its most significant bit whenever the first bit of z is 0 or the second bit of z is 0. Thus, if the binary expansion of z_0 contains roughly equal proportions of 1's and 0's, we expect that the binary expansion of $z[n] = g^n(z[0])$ will have 1 as its most significant bit about twice as frequently as 0. This intuition may be verified using the techniques of Chapter 3 to compute the the invariant density, $p(z)$, associated with g. The density is of the form

$$p(z) = \begin{cases} \frac{2}{3} & 0 \leq z < \frac{1}{2} \\ \frac{4}{3} & \frac{1}{2} \leq z \leq 1 \end{cases}.$$

Thus, $z[n]$ takes values in the interval $[\frac{1}{2}, 1]$ with about twice the frequency with which it takes values in the interval $[\frac{1}{2}, 1]$.

5.3.1 Statistics of the Continuous-Time Current

In previous sections we have determined a model for the evolution of the inductor current at switching times and shown how its properties may provide some insight into the converter's operation. However, properties of the continuous time inductor current are not immediately evident from the previous discussion. For example, even though the sampled

inductor current is described by an eventually expanding map with statistics that may be approximated using the previously presented tools, we will see that the statistics of the continuous time inductor current are periodically-time-varying, or cyclostationary. The periodicity stems from the periodically switched nature of the current-mode control scheme. This result is potentially of more than academic interest since a statistical characterization of the continuous inductor current in the converter may provided some insight into energy radiated from such devices.

This section determines the autocorrelation of the continuous time inductor current. Although, we will concentrate on the autocorrelation, the technique presents may be more applicable to a broader class of statistics. Our task is accomplished in two steps. First, we relate the continuous inductor to the sampled inductor current. Next, using this relation, we relate the statistics of the continuous inductor current to those of the sampled inductor current.

5.3.2 Continuous Inductor Currents

We begin by determining a compact representation for the continuous time inductor current in terms of the sampled inductor current. The time instant at which the inductor current takes its maximum value during each switching interval figures prominently in this representation. As discussed in the derivation of the sampled current map, there are two possible inductor current behaviors for t between nT and $(n+1)T$: either the inductor current increases linearly with slope m_1 for the entire switching interval or the peak current is reached at some instant during the interval, at which time the inductor current decays linearly with slope $-m_2$. In the first case, the maximum current occurs when $t = (n+1)T$. In the second case, the maximum is reached at time $t = n + \frac{i_p - i(nT)}{m_1}$. We define t^* to be the time instant of maximum current, i.e,

$$t^*(i) = \min(T, \frac{I_p - i}{m_1}).$$

With this notation, the continuous time inductor current is of the form

$$i(t) = \begin{cases} i(nT) + m_1 t & nT \le t \le nT + t^*(i(nT)) \\ I_p - m_2(t - t^*(i(nT))) & nT + t^*(i(nT)) \le t \le (n+1)T \end{cases} . \tag{5.13}$$

Equation (5.13) shows that the continuous current is exactly determined by its samples. Thus, no information is lost by considering the sampled current.

In order to simplify notation, define the shifted inductor current pulse by

$$p(i,t) = \begin{cases} i + m_1 t & 0 \le t \le t^*(i) \\ i_p - m_2(t - t^*(i)) & t^*(i) \le t \le T \\ 0 & \text{otherwise} \end{cases} . \tag{5.14}$$

Thus, for fixed i, the graph of $p(i,t)$ is the continuous time current for times between 0 and T given that the current at time 0 is i. Alternatively, for fixed t, $p(i,t)$ as a function of i is the inductor currents value at a fixed time t between 0 and T for all initial inductor currents i.

Substituting Equation (5.14) into Equation (5.13), the inductor current at any positive time may be expressed as

$$i(t) = \sum_{n=0}^{\infty} p(i(nT), t - n).$$

We shall see below how this expression can be used to compute the autocorrelation of the continuous time inductor current.

Autocorrelations of the Continuous Time Current

The autocorrelation of the continuous time current is of the form

$$\begin{aligned} R_{ii}(t,s) &= E(i(t), i(s)) \\ &= E(\sum_{n=0}^{\infty} p(i(nT), t - nT) \sum_{m=0}^{\infty} p(i(mT), s - mT)). \end{aligned}$$

Because $p(i,t) = 0$ when t is not in the interval $[0,T]$, the above expression simplifies considerably. Specifically, let we break the times t and s into an integer multiple of the

switching period and a fractional part, i.e. let $\gamma = t \bmod T$, $\nu = s \bmod 1$, $l = \frac{(t-\gamma)}{T}$, and $k = \frac{(s-\nu)}{T}$ so that $t = lT + \gamma$, $s = kT + \nu$ and with l and k integers and $0 \leq \gamma, \nu \leq T$. With this notations, the autocorrelation may be expressed as

$$R_{ii}(t,s) = E(\sum_{n=0}^{\infty} p(i(nT), lT + \gamma - nT) \sum_{m=0}^{\infty} p(i(mT), kT + \nu - mT))$$

From the definition of $p(\cdot,\cdot)$ (Equation (5.14)) and fact that $0 \leq \nu, \gamma \leq T$ we see that $p(i, kT + \nu - mT)$ is nonzero only when $k = m$. Thus, the above expression reduces to the simpler

$$\begin{aligned} R_{ii}(t,s) &= E(p(i(lT),\gamma)p(i(kT),\nu)) && (5.15) \\ &= E(p(i(lT), t \bmod T)p(i(kT), s \bmod T)) && (5.16) \end{aligned}$$

where k and l have are as previously defined. Also, assuming $s \geq t$ implies $k \geq l$, we have

$$R_{ii}(t,s) = E(p(i(lT), t \bmod T)p(f^{k-l}(i(lT)), s \bmod T))$$

so that the expectation is a correlation statistic of the form discussed earlier. The case $t \geq s$ may be handled similarly.

Equation (5.15) makes clear the relation between the statistics of sampled and continuous inductor currents; The autocorrelation is a correlation statistic of the sampled inductor current. It follows that the techniques of Chapter 4 can be used to compute numerical values of the autocorrelations. Similar expressions can be derived for the average current, and the technique seems to be applicable in a broader settings.

Since $R_{ii}(t,s)$ depends on both t and s, the continuous current is not a stationary process. However, the autocorrelation does satisfy

$$R_{ii}(t + nT, s + nT) = R_{ii}(t,s),$$

any integer n. Thus, the continuous time inductor current is wide sense cyclostationary. This periodicity is a consequence of the switched nature of the circuit.

Chapter 6

Observation Effects : Maps of the Interval

This chapter and the next consider the effects observing chaotic signals through linear time invariant filters. As chaotic signals have been proposed as models for several physical phenomena [2, 8] as well as for use as signaling waveforms in communication systems [34], filtered chaotic signals may arise, for example, when chaotic phenomena are measured through sensors with linear dynamics [35, 36], when chaotic signaling waveforms pass through a bandlimited channel, and in certain models of physical systems [37].

The means by which a signal is observed may alter the results of the most common techniques of chaotic time series analysis. For example, estimates of the fractal dimension and Lyapunov exponents computed from empirical data are often used as diagnostic quantities to infer the presence of chaos; a low fractal dimension and a positive Lyapunov exponent indicate chaos. Also, fractal dimension can potentially be used as a signature quantity for detecting and classifying chaotic signals. When the sensors used to measure the data are not ideal in the sense that they contain dynamics of their own, there are two components affecting the results of the data analysis: the underlying chaotic signal, and the sensor itself. In this chapter, we show that, in general, these two components interact in a nontrivial manner and that the properties of the measured signal may differ from those of the underlying signal being measured.

We will see in Chapter 7 that the behavior under filtering of chaotic signals generated by invertible chaotic systems differs from that of signals generated by noninvertible systems. The focus of this chapter is on time series generated by filtered one-dimensional maps, which are inherently noninvertible.

As we have seen in Chapters 2,3, and 4 the ensemble of time series generated by an eventually-expanding map with random initial condition may be viewed as a stochastic process. Furthermore, when the initial condition is chosen according to the invariant density the resulting process is stationary. It follows that general results from the theory of linear systems driven by stochastic processes (see e.g. [38, 39, 40]) may be applied to the problem of determining the statistical relationship between the filter input and output for filtered chaos. However, we shall be interested in chaotic signals, which have extra structure that is not accounted for in the classical theory of stochastic processes. In particular, classical results give no insight into how the geometric structure of a chaotic signal manifests itself at the filter output.

Empirical evidence indicates that the geometric structure of the input can be obscured by linear filtering [37, 41]. In order to understand this phenomenon, this chapter considers two interpretations of the filtered time series—it may be viewed either as the output of a filter driven by a chaotic signal, or as an observation of the state of a large nonlinear system that is, loosely speaking, the cascade of a chaotic system and a linear system. These two views correspond roughly to input-output descriptions and state space descriptions of the system respectively.

We begin in Section 6.1 by presenting several examples illustrating some of what may occur when chaotic signals are filtered. Next, we consider in Section 6.2, the input-output description of filtered chaos. We will see that it is often the case that the filter's input is related to its output by a simple nonlinear map. Furthermore, in certain circumstances, the filter output is also governed by a chaotic map. We determine the conditions under which this occurs. These conditions are applied to synthesize a family of chaotic white noise generators that may have potential applications in certain signal generation contexts.

Section 6.3 relates the statistical properties of the filter input to those of the output. The statistical dependencies inherent in chaotic signals allow for some unusual developments.

We compare chaotic signals to IID non-Gaussian processes, and prove some results leading to a deconvolution algorithm.

Finally, Section 6.5 views the filtered chaotic signal as an observation of the state of the larger chaotic system. Through an example, we show that, as a direct consequence of the noninvertibility of the map, the state of this larger system may evolve on a fractal subset of the state space. Furthermore, the dimension of this fractal depends on the parameters of both the map and the filter.

6.1 Some Examples of Filtered Maps

This section presents three examples designed to illustrate some of the possible characteristics of filtered chaotic signals. The basic situation considered in this chapter is as follows. A signal $x[n]$, generated by a chaotic map f, drives a linear time invariant filter with impulse response $b[n]$. The filter output $y[n]$ is completely determined by the following equations:

$$\begin{aligned} x[n] &= f(x[n-1]) \\ y[n] &= b[n] * x[n]. \end{aligned}$$

As in the previous chapters, we restrict our study to eventually-expanding maps. Even for this restricted class, however, the filter output may display a wide range of behaviors. Also, as might be expected, the properties of the output signal depend on both the filter impulse response and the map generating the input.

Example 1:

Filtering a chaotic time series need not obscure the entire structure of the input. Suppose, for example, that the FIR filter with system function $H(z) = 1 - .9z^{-1}$ is driven by a typical time series generated by the simple Markov map f shown in Figure 6-1 and of the form

$$f(x) = \begin{cases} \frac{3}{7}x + \frac{7}{10} & 0 \le x \le \frac{7}{10} \\ \frac{10}{3}(1-x) & \frac{7}{10} < x \le 1 \end{cases} . \tag{6.1}$$

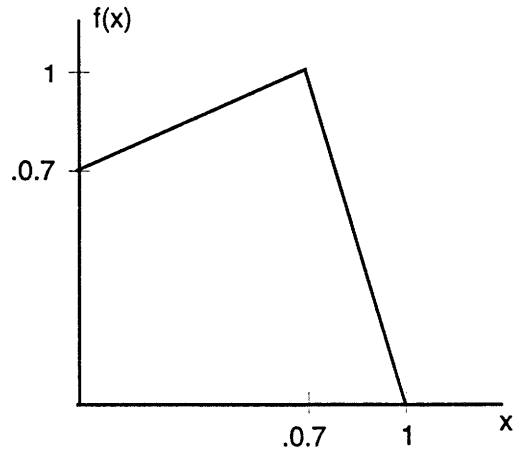


Figure 6-1: The Graph of the Markov Map of (6.1)

The output of the filter satisfies

$$y[n] = x[n] - .9x[n-1]. \tag{6.2}$$

Since the values of $x[n]$ are deterministically related, the filter output at time n may be expressed as a function of the previous input by substituting (6.1) into (6.2) to obtain

$$y[n] = f(x[n-1]) - .9x[n-1].$$

In order to compactly describe the effect of the filter, we define the function $\phi(x) = f(x) - .9x$ so that $y[n] = \phi(x[n-1])$. The function ϕ, shown in Figure 6-2, is one to one with a continuous inverse. Thus, the time series $x[n]$ and $y[n]$ are related in a straightforward manner. Furthermore, by mapping values of $y[n-1]$ back to values of $x[n-2]$ through the inverse of ϕ, followed by mapping $x[n-2]$ to $x[n-1]$ through f, and finally mapping $x[n-1]$ to $y[n]$ through ϕ, it is evident that $y[n]$ is completely determined by $y[n-1]$. This series of operations is shown in Figure 6-3. More precisely, the consecutive values of the sequence $y[n]$ are related by a map of the form

$$y[n] = \phi(f(\phi^{-1}(y[n-1]))) \triangleq g(y[n-1]). \tag{6.3}$$

The function relating $y[n]$ and $y[n-1]$ is shown in Figure 6-4.

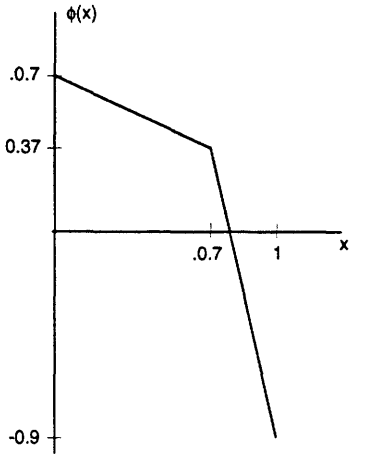


Figure 6-2: The Function $\phi(x)$ Relating the Filter's Input to its Output

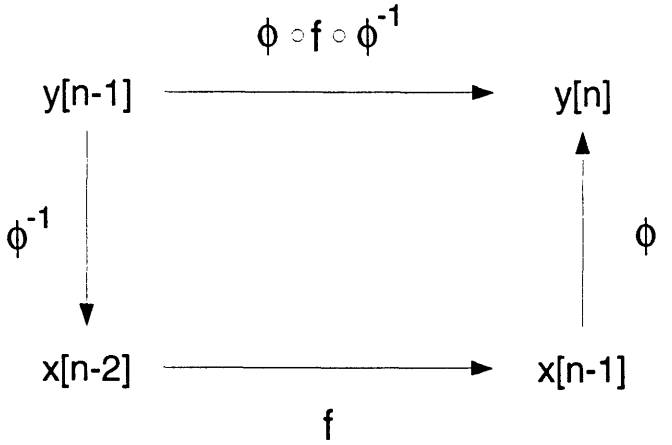


Figure 6-3: The Evolution of the Filter Output

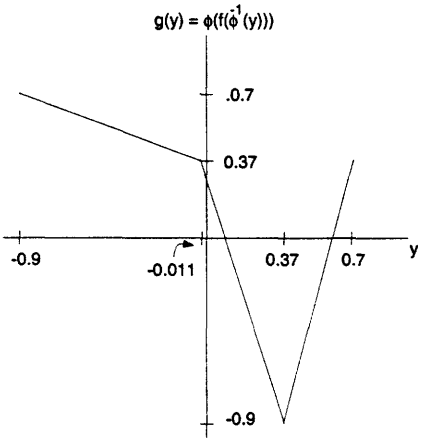


Figure 6-4: The Map Relating Consecutive Filter Outputs

In this case, the filter input and output are both determined by (different) one-dimensional piecewise smooth maps. The structure of the filter output is particularly easy to describe because the function ϕ, which depends on the filter and the map f, is piecewise smooth and invertible. As we will see in Section 6.2, the situation in which both a filter's input and output are governed by a one-dimensional map is not uncommon. The next example shows that other types of behavior are also possible.

Example 2:

Filtering may obscure some of the structure of a chaotic signal. Consider for example, time series generated by the map f define by (6.1) filtered by the single pole filter

$$y[n] = ay[n-1] + x[n].$$

Plots of $y[n-1]$ vs. $y[n]$ for filter parameters $a = .3$ and $a = .7$ are shown in Figures 6-5 (a) and (b) respectively. These figures suggest that (for these values of the filter parameter) the relationship between consecutive samples of the output is multivalued. Indeed this is the case—the output is not determined by a piecewise smooth one-dimensional map. In fact, the set of values of the pair $(y[n-1], y[n])$ appears to have a complicated structure. In Section 6.5 we will present an example showing that the output of a filter driven by a one-dimensional map may actually be a fractal.

One might guess that the fundamental difference between the two previous examples is that the first deals with FIR filters while the second deals with IIR filters. However, as the next example shows, IIR filters and FIR filters may have fundamentally equivalent effects on chaotic signals.

Example 3:

Consider the system depicted in the block diagram of Figure 6-6, driven by a time series generated by the map of (6.1). From the standpoint of input/output behavior, the cascade of two subsystems is equivalent to a two tap FIR system. However, we consider each subsystem separately. As was determined in Example 1, the output of the first subsystem

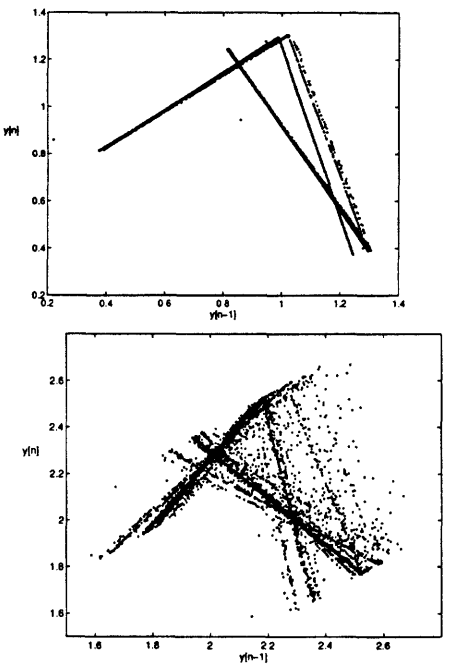


Figure 6-5: Plots of Consecutive Filter Outputs for the Filters of Example 2

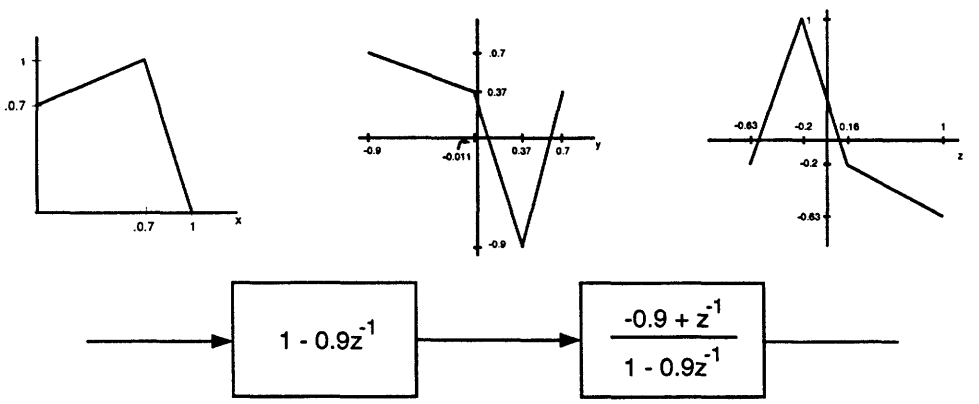


Figure 6-6: The Block Diagram of the Construction Described in Example 3

$y[n]$ obeys the one-dimensional recursion of (6.3). It follows that the second subsystem of Figure 6-6 can be considered independently as an IIR filter driven by a time series from the chaotic map

$$y[n] = g(y[n-1]),$$

where g is defined in Example 1. The output of the second subsystem satisfies the relationship

$$z[n] = g_1(z[n-1]),$$

where g_1 is as shown in Figure 6-6. Thus, a one-dimensional map may maintain some of its one-dimensional character even when filtered by an IIR filter.

As the previous examples illustrate, a linear filter may alter a chaotic signal in a variety of ways. The remainder of this chapter explores conditions under which these different types of effects occurs.

6.2 Filtered Maps: FIR filters

To begin, let us restrict our attention to filters with finite impulse responses. As we will show, a functional relationship can always be derived relating a single sample of the input to a corresponding output sample. The properties of this function determine the properties of the output.

6.2.1 The Input-to-Output Map

The input signals considered here are generated by the scalar recursion

$$x[n] = f(x[n-1]). \tag{6.4}$$

Since the time series $x[n]$ will be the input to a filter, we refer to f as the *input map.* When a time series generated by f drives a finite impulse response filter, the filter input and output are related in a particularly straightforward way. Specifically, suppose the filter

output satisfies the difference equation

$$y[n] = b[0]x[n] + b[1]x[n-1] + \cdots + b[N-1]x[n-N+1]$$

where the sequence $b[n]$ is the finite length impulse response of the filter. Since the values of the sequence $x[n]$ are related deterministically by (6.4), the filter output also satisfies

$$y[n] = b[0]f^{N-1}(x[n-N+1]) + b[1]f^{N-2}(x[n-N+1]) + \cdots + b[N-1]x[n-N+1]. \quad (6.5)$$

Thus, the n-th sample of the filter output depends only on the $(n-N+1)$-th sample of the input.

When FIR filters are driven by chaotic maps, the filter output at any time is a function of one input sample. We refer to this function as the *input-to-output map* and denote it by $\phi_b(x)$ (the subscript "b" is included to emphasize the dependence of the map on the filter coefficients). With this definition, the filter input and output are related by the expression

$$y[n] = \phi_b(x[n-N+1]) \quad (6.6)$$

where N is the number of filter taps and $\phi_b(x)$ is defined as the right side of (6.5), i.e.,

$$\phi_b(x) = b_0 f^{N-1}(x) + b_1 f^{N-2}(x) + \cdots + b_{N-1}x.$$

The input to output map $\phi_b(x)$ is a linear combination of iterates of the input map f. Since f along with each of its iterates, is piecewise continuous, the input to output map ϕ_b is piecewise continuous as well. Its other characteristics depend rather strongly on the input map and the impulse response of the filter.

6.2.2 Filters Preserving Determinism

Expression (6.6) allows us to determine how some of the analytical and statistical properties of the output of the filter depend on its input and impulse response. We begin by referring back to the examples. As illustrated in Example 1, the filter output may, for a certain input maps and impulse responses, be described by a one-dimensional piecewise smooth map. In

this case, the filter output may be determined by iterating the map describing the output from an appropriate initial condition. To emphasize that the value of the filter output at any time is uniquely determined by a single previous value of the state, we say that such filters preserve one-state determinism.

The behavior illustrated in Example 1 is not contrived—one-state determinism may be preserved by more general filters, each filter inducing a different map describing the filter output. Indeed, any filter $b[n]$ for which ϕ_b is invertible generates an output governed by

$$\begin{aligned} y[n] &= \phi_b(f(\phi_b^{-1}(y[n-1]))) \\ &= g_b(y[n-1]) \end{aligned} \tag{6.7}$$

where the innermost function maps an output value $y[n-1]$ to an input value $x[n-N]$, the middle function maps the input values $x[n-N]$ to $x[n-N+1]$, and the outermost function maps the input value $x[n-N+1]$ to output value $y[n]$. This sequence of operations is equivalent to those shown in Figure 6-3. Since it describes the evolution of the filter output, we will refer to $g_b(y)$ as the *output map*. Compositions of piecewise continuous functions are also piecewise continuous. It follows that g_b is piecewise continuous when f is.

Although all input map/ FIR filter pairs generate an input-to-output map, an output map exists only in certain cases. When the input map is eventually-expanding and the input-to-output map is invertible, then the filter output is governed by an eventually expanding map. We turn now to the converse question: if the output of a filter has one-state determinism, must its input-to-output map be invertible? We will see that the answer to this question is a qualified yes.

A Characterization Theorem

It is conceivable that for certain non-invertible input-to-output maps ϕ_b, the output still obeys a one-dimensional recursion, albeit different from (6.7) in form. For instance, when the filter is simply a delay, its impulse response is $b[n] = \delta[n-n_0]$. The input-to-output map is $\phi_b = f^{n_0}$, which is certainly not invertible when f is an eventually-expanding map.

However, the filter output $y[n]$ is simply a shifted version of the input and must therefore be governed by f. Furthermore, if the filter $b[n]$ preserves one-state determinism, any shifted version $b[n - n_0]$ will preserve one-state determinism although the input to output map of the shifted version will not in general be invertible. The following theorem, which is proved in Appendix B, asserts that for eventually-expanding input maps, these are in essence the only non-invertible input-to-output maps preserving one-state determinism.

Theorem 7 *Suppose that an eventually-expanding map f is the input map to a nonzero FIR filter $b[n]$ and that ϕ_b is the associated input-to-output map. Then the filter output satisfies $y[n] = g(y[n-1])$ with g eventually-expanding if and only if $\phi_b(x)$ has the form*

$$\phi_b = \phi_1 \circ f^k$$

for some invertible ϕ_1 and some integer $k \geq 0$.

Remark 1: Theorem 7 may be viewed as a characterization of all filters that preserve one-state determinism for a fixed input map. Specifically, when f is the input map, the set of all such FIR filters is precisely the set of filters for which ϕ_b is invertible.

Remark 2: From another point of view, Theorem 7 suggests that generally, filtered chaos is difficult to distinguish using naive methods. Suppose that a signal is observed through a sensor which is well modeled as an FIR filter. Even when the observed signal displays one-state determinism, we may not be able to infer the form of the unfiltered input map. In other words, several input maps may generate the same output map. If neither the filter nor the input is specified, there is an inherent ambiguity. In short, one-state determinism at the output is neither a sufficient criterion to determine that a chaotic signal has been filtered, nor a sufficient criterion to recover the filter input. Furthermore, the absence of one-state determinism at the output does not preclude determinism at the input. This implies that simple scatter plots of physical data may be inadequate to discern a filtered chaotic signal.

Several groups have proposed techniques for detecting filtered chaos and inverting the filter based only on the filter output [35, 36, 42]. These techniques are based primarily on heuristic analyses. The results of this section indicate that any method for detecting

filtered chaos and recovering the input that is based only on the filter output cannot be consistently successful. However, as we will see in Section 6.3, the ambiguity can be precisely characterized.

Application: Synthesis of Chaotic White Noise

The results of the previous section suggest a procedure to generate a family of eventually-expanding maps that generate time series with flat power spectra. Such maps may be potentially useful, for example, as signaling waveforms in some communication systems [34]. The only previously reported examples of maps that generate spectrally white chaotic signals seem to be the symmetric tent map and the symmetric logistic map [18]. The family derived below depends on a parameter and hence contains a continuum of maps. We will refer to such maps as chaotic white noise generators.

Consider the family of tent maps with parameter a given by

$$f(x) = \begin{cases} \frac{x}{a} & 0 \leq x \leq a \\ \frac{1-x}{1-a} & a < x \leq 1. \end{cases} \tag{6.8}$$

As was shown in Section 3.1, the process associated with f has a power spectrum of the form

$$S_{xx}(z) = \frac{1}{|1-(2a-1)z^{-1}|^2}.$$

If the whitening filters for this spectrum preserve one-state determinism, the output map of the whitening filter will generate time series with a white spectrum.

There are two FIR whitening filters corresponding to the minimum phase and maximum phase spectral factors of $S_{xx}(z)$. Specifically, the filters with system functions

$$\begin{aligned} B_1(z) &= 1-(2a-1)z^{-1} \\ B_2(z) &= 1-(2a-1)^{-1}z^{-1} \end{aligned}$$

both perform the whitening operation. Since these are FIR filters, the previous results of this section apply. Suppose we choose $B_2(z)$ as our whitening filter. The input-to-output

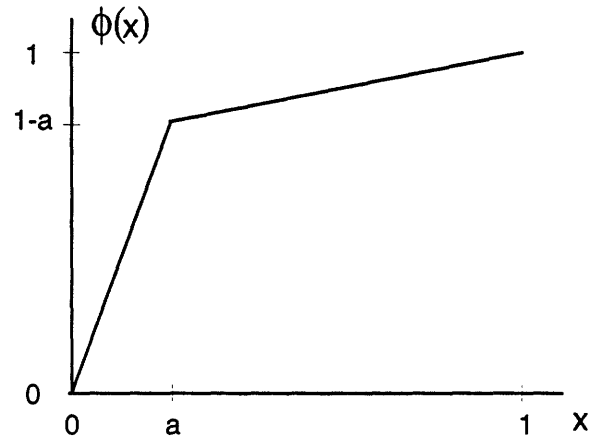


Figure 6-7: Input-to-Output Map the Whitening Filter

map is

$$\phi_{b_2}(x) = f(x) + (2a-1)^{-1}x.$$

By substituting in the explicit form of f from (6.8), we obtain the following functional form for ϕ_{b_2},

$$\phi_{b_2}(x) = \begin{cases} \frac{a-1}{a(2a-1)}x & 0 \le x < a \\ \frac{1}{1-a} + \frac{a}{(a-1)(2a-1)}x & a \le x \le 1 \end{cases}.$$

Figure 6-7 shows the normalized input-to-output map $(1-2a)\phi_{b_2}(x)$ corresponding to the scaled whitening filter $(1-2a)B_2(z)$, which generates a filter output taking values only between 0 and 1. We infer from this plot that ϕ_{b_2} is invertible for every value of a between 0 and 1; hence, the filter output is governed by a piecewise linear output map.

Since the input-to-output map is invertible and we have an explicit expression for it, the output map exists and can be determined by a direct application of (6.7). Through straightforward but somewhat tedious calculations the output map can be shown to have the graph shown in Figure 6-8. For parameter values a between 0 and 1, the above construction yields a map which generates spectrally white time series.

Although each member of the family of white noise generators constructed above has the same spectral properties, the time series will in general have different characteristics. Examples of time series generated using several different parameter values are shown in Figure 6-9 (a)–(d).

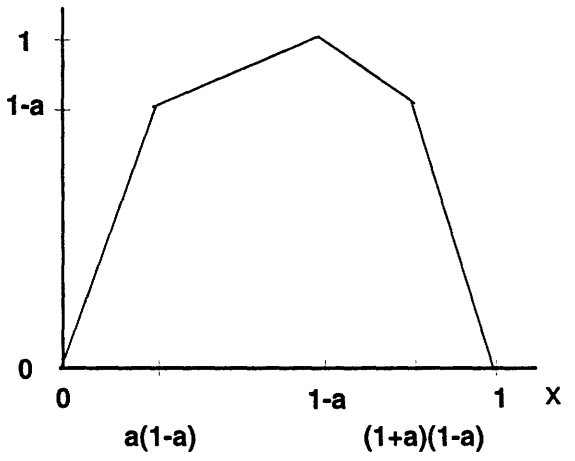


Figure 6-8: A Family of Maps Generating Chaotic White Noise

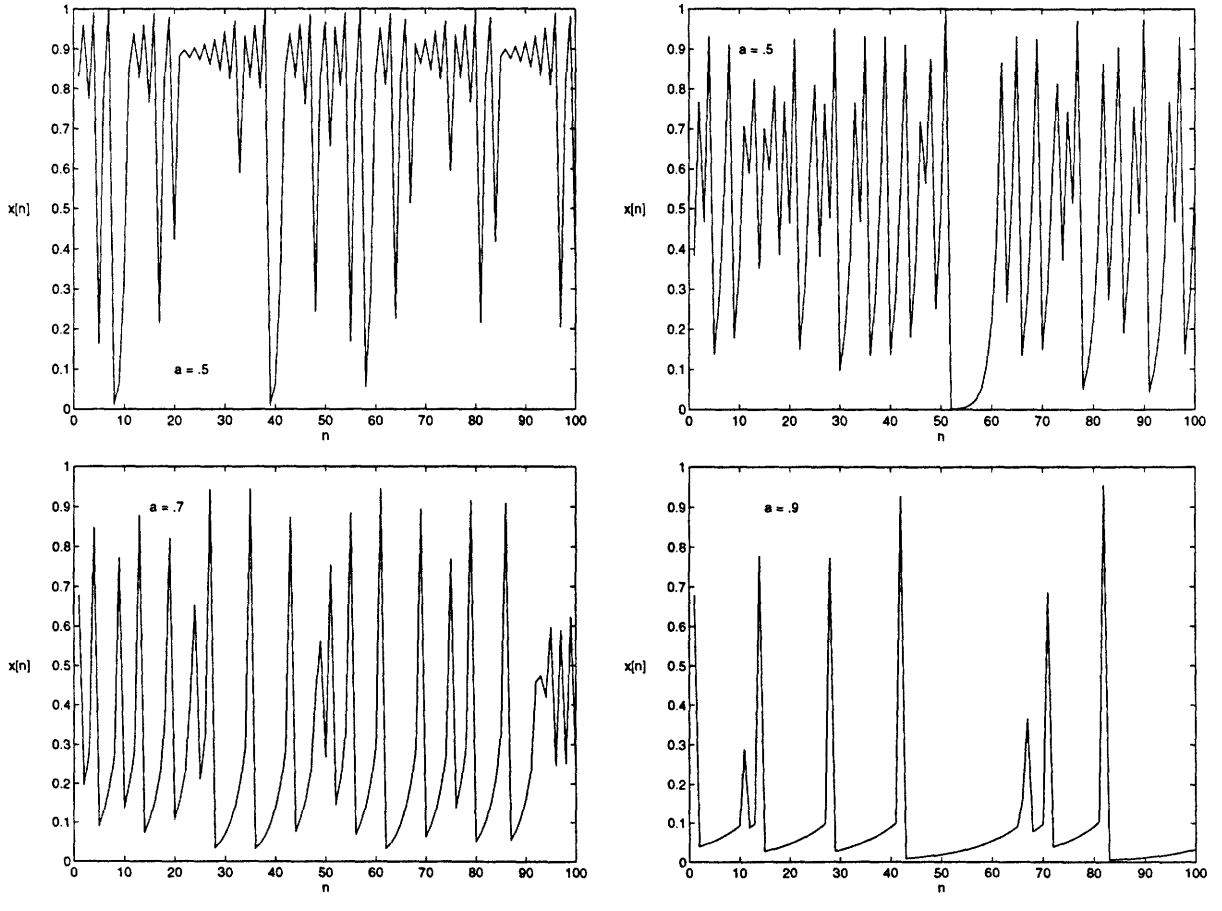


Figure 6-9: Examples of Chaotic White Noise Time Series

In principle, the technique of applying a whitening filter applies to maps more general than tent maps. In particular, as we have seen in Chapter 3, all piecewise linear Markov maps have rational spectra and hence have rational whitening filters. However, it is important to emphasize that many piecewise linear Markov maps have no whitening filter that preserve one-state determinism. An interesting area of future work is the characterization of all piecewise linear Markov maps that generate chaotic white noise.

6.3 FIR Filtered Maps: Statistical Properties

The results of Section 6.2.2 give conditions under which the output of an FIR filter can be generated by one-dimensional maps when the input is generated by a one-dimensional map. This result relates the deterministic properties of the filter input and output. It does not however provide much insight into the relationship between the statistical properties of the input and output maps. In this section, we explore this relationship.

6.3.1 Chaos vs. Statistical Independence

As might be expected, the probabilistic properties of filtered chaotic signals are quite different from those of filtered sequences of IID random variables. However, because the properties of filtered IID sequences are well understood, they provide a counterpoint to filtered chaos. For example, when finite variance, non-Gaussian IID sequences drive a non-trivial linear filter, the marginal pdf of the output must differ from the marginal pdf of the input [43]. This fact has practical implications; for example, it forms the basis for a large number of deconvolution algorithms. The most common techniques choose an inverse filter that forces the output pdf to match the input pdf (see e.g. [43]). As the following example shows, the marginal pdf of a chaotic signal can sometimes remain unaffected by filtering. This implies that chaotic signals generated by one-dimensional maps cannot be deconvolved using traditional techniques.

Example: According to the results of the Section 6.2.1, the output of an FIR filter driven by a one-dimensional map is related to an input sample by the piecewise monotonic input-to-output map ϕ_b. The existence of such a mapping allows us to directly determine some

of the statistical properties of the output.

Suppose the input map $f(x)$ is the skewed tent map given by

$$f(x) = \begin{cases} 3x & 0 \leq x \leq \frac{1}{3} \\ \frac{3(1-x)}{2} & \frac{1}{3} < x \leq 1 \end{cases} . \tag{6.9}$$

The map (6.9), like all tent maps, has a uniform invariant density (see Section 3.1). Suppose that a time series $x[n]$ generated by f drives the filter

$$y[n] = .6x[n] - .2x[n-1] + .6x[n-2].$$

It follows that the input-to-output map $\phi(x)$ is of the form

$$\phi(x) = .6f^2(x) - .2f(x) + .6x.$$

Straightforward calculations yield the following expression for $\phi(x)$:

$$\phi(x) = \begin{cases} \frac{27}{5}x & 0 \leq x \leq \frac{1}{9} \\ \frac{9}{10} - \frac{27}{10}x & \frac{1}{9} < x \leq \frac{1}{3} \\ \frac{9}{4}x - \frac{3}{4} & \frac{1}{3} < x \leq \frac{7}{9} \\ \frac{12}{5} - \frac{9}{5}x & \frac{7}{9} < x \leq 1 \end{cases} \tag{6.10}$$

The graph of ϕ is shown in Figure 6-10.

If the initial condition of f is chosen according to the uniform invariant density, it is straightforward to show that the marginal density of $x[n]$ is uniform for all positive n. In particular, since the filter output is related to the input through the input to output map of (6.10), its marginal density can be determined using standard techniques for determining the density of a function of a random variable [40, 44]. Specifically, the density, p_y, of the filter output satisfies

$$p_y(y) = \begin{cases} \frac{5}{27}p_x(\frac{5y}{27}) + \frac{10}{27}p_x(\frac{9-10y}{27}) + \frac{12}{27}p_x(\frac{4y+3}{9}) & 0 < y < .6 \\ \frac{4}{9}p_x(\frac{4y+3}{9}) + \frac{5}{9}p_x(\frac{12-5y}{9}) & .6 < y < 1 \end{cases}$$

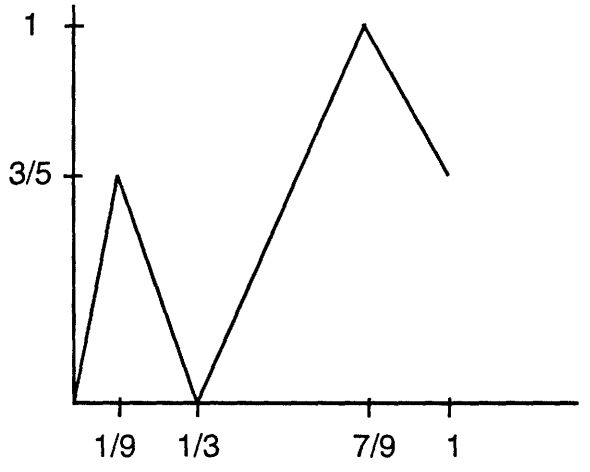


Figure 6-10: The Input-to-Output Map (6.10)

Since $p_x = 1$ on $[0, 1]$ it follows that $p_y = p_x$ and so the pdf is unchanged by filtering.

Remark 1: The example is intended to illustrate the possibility that the marginal pdf of the input may be unaffected by filtering. This phenomenon is possible only because of the strong dependence between the samples of the filter input; when there is no dependence, the input is an IID random process and the marginal density must be affected by filtering. Thus, unlike IID non-Gaussian processes, chaotic processes may not be uniquely recovered after filtering given only the marginal pdf.

Remark 2: Although the marginal densities of the input and output of the filter are identical in this example, the filter does not preserve one-state determinism since ϕ is not invertible on $[0, 1]$. Thus, the filter output $y[n]$ need not be governed by a one-dimensional map. As will be seen below, nontrivial filters must alter either the marginal pdf or the one-state determinism of the input. This observation leads to a technique for deconvolving chaotic signals.

6.3.2 Effects of FIR filtering on Probabilistic Properties

The arguments of the previous example carry over to the general case. Suppose that the input map f has invariant density p and its initial condition is chosen randomly according to p. The results of Chapter 2 show that the input sequence $x[n]$ is a stationary stochastic

process with marginal density p. Since the output sequence $y[n]$ is a filtered version of $x[n]$, it is also stationary. We denote its (time invariant) marginal pdf by p_y. By the results of the previous section, the input and output of the filter are related by the piecewise smooth input to output map ϕ, i.e.

$$y[n] = \phi_b(x[n]).$$

As above, the pdf of $y[n]$ may be readily determined using standard results from probability theory [40, 44]. In particular, when ϕ_b is invertible the pdf of the filter output may be calculated by

$$p_y(y) = p_x(\phi_b^{-1}(y))|\frac{d}{dy}\phi_b^{-1}(y)| \tag{6.11}$$

whenever the derivative on the right hand side exists. When the input-to-output transformation is not invertible, the situation is only slightly more complex. in this case, the relation between the input and output is

$$p_y(y) = \sum_{x \in \phi_b^{-1}(y)} p_x(x)/|\frac{d}{dx}\phi_b(x)|. \tag{6.12}$$

In this noninvertible case, the value of the output pdf at a point y is a weighted sum of the values of the input pdf at each point mapping to y through the transformation ϕ_b. The weight of each term in the sum is the reciprocal of the slope of ϕ_b at that point[1].

The relationship of (6.12) may be viewed as an operator mapping an input pdf to an output pdf. From this viewpoint, it may be interpreted as the Frobenius–Perron operator of the nonlinear mapping ϕ_b. When the stationary density of f is also a stationary density of ϕ_b, the filter input and output will have the same marginal pdf. In other words, the marginal pdf of the input will appear at the output for filters satisfying

$$p_x(y) = \sum_{x \in \phi_b^{-1}(y)} p_x(x)/|\frac{d}{dx}\phi_b(x)|. \tag{6.13}$$

[1]Since $\phi_b(x)$ is piecewise monotonic, so is its inverse. It follows that the derivative exists at all but a finite number of points. These points do not affect the density.

Remark : In Section 6.2.2 we completely characterized the set of filters preserving one-state determinism. The reasoning of Remark 2 of that section can be used to infer that an inherent ambiguity exists in any scheme to recover the filter input that is based strictly on an observation of its output. This follows because several filters may preserve one-state determinism. Equation (6.13) provides an analogous result for filters that preserve the marginal probability density of the input. Since we can construct nontrivial filters preserving the marginal pdf of a chaotic signal, any scheme to recover the chaotic input to a filter based solely on its marginal pdf will also have an ambiguity.

6.4 Deconvolution of Chaotic Signals

In light of the last remark, it is natural question to ask what amount of *a priori* information about the filter input is necessary to recover it based only on the filter output? This section addresses this question for certain input maps and filters.

6.4.1 Deconvolution Theorems

The results of the previous section show that individually, neither one-state determinism nor the marginal pdf is sufficient to recover the filter input from its output. However, the next lemma shows that together, one-state determinism and marginal pdf impose stronger constraints on the filter input and output pair when the input map is continuous. In particular, the next theorem, which is proved in Appendix B, asserts that for continuous input maps, the only FIR filters preserving both one-state determinism and marginal pdf are simple delays.

Theorem 8 *Consider a continuous eventually-expanding input map and suppose that a nonzero FIR filter $b[n]$ preserves one-state determinism. Then $b[n]$ preserves the marginal pdf of the input if and only if $b[n] = \delta[n - n_0]$ for some integer n_0.*

Although one hypothesis of Theorem 8 is that the input map is continuous, the results do not appear to depend fundamentally on continuity. It seems reasonable that the proof of Theorem 8 can be extended to include discontinuous input maps.

Suppose that an FIR filter $b[n]$ preserves one-state determinism and has the further property that the input and output maps are identical. Then the marginal pdf of the input is preserved. According to Theorem 8 the filter is necessarily a delayed impulse. Thus Theorem 8 has the following immediate corollary.

Corollary 1 *Consider an FIR filter $b[n]$ which preserves the one-state determinism of a continuous eventually-expanding input map. Suppose that the input map and the output map are identical. Then the filter is a simple delay, i.e. $b[n] = \delta[n - n_0]$.*

Note that the previous theorem and its corollary are stated without direct reference to the input-to-output map. This leads us to conjecture that Corollary 1 may apply to IIR filters as well. In the case of IIR filters, the input-to-output transformation may be more difficult to determine and the techniques used to prove Theorems 7 and 8 (and hence Corollary 1) may not apply. However, since Corollary 1 applies to all finite length filters, it seems plausible that it will hold also for filters with infinite extent impulse responses. More specifically, we conjecture that when $b[n]$ is a stable filter that preserves the one-state determinism of an eventually-expanding input map and the input map and the output map are identical, then the filter is a simple delay, i.e. $b[n] = \delta[n - n_0]$.

6.4.2 Deconvolution by Dynamics Matching

This section describes a technique for recovering a chaotic signal from a filtered observation. The approach is based on the conjecture of Section 6.4.1 and assumes that the input map is known. The results presented here are of a rather preliminary nature and are included primarily to suggest the viability of our approach.

Again, we consider a filter driven by a chaotic signal generated by a one-dimensional map. As usual, the input map, the input time series and the filter impulse response are denoted by f, $x[n]$ and $b[n]$ respectively. The filter output is $y[n] = b[n] * x[n]$. We wish to recover $x[n]$ based only on an observation of the filter output and the knowledge of the input map f.

The approach involves choosing an inverse filter with impulse response $c[n]$ such that

its output

$$\begin{aligned} z[n] &= c[n] * y[n] \\ &= c[n] * b[n] * x[n] \end{aligned} \tag{6.14}$$

satisfies $z[n] = f(z[n-1])$. Theorem 8, Corollary 1 and our running conjecture suggest that when such a $c[n]$ can be found, the impulse response of the cascade, $c[n] * b[n]$, must be a simple delay. Because the inverse filter $c[n]$ is chosen so that its output follows the same dynamics as the input, we refer to this approach as *dynamics matching.*

A straightforward algorithm for determining a dynamics matching inverse filter is based on an objective function that is minimized by the desired filter. A dynamics matching filter will necessarily minimize the objective function

$$J = \sum_{n=1}^{\infty} (z[n] - f(z[n-1]))^2,$$

since in this case, $J = 0$. This suggests choosing $c[n]$ to minimize J. One difficulty with this approach occurs when the map $f(x)$ has a fixed point at zero, i.e. $f(0) = 0$. In this case, the trivial filter $c[n] = 0$ for all n minimizes J. This trivial minimizing solution can be eliminated by requiring the filter output power to equal the input power. Since the power in the input and power in the output are their respective variances, this corresponds to adding the constraint $\sigma_z^2 = \sigma_x^2$.

The approach we will pursue is to solve the following optimization problem

$$\begin{aligned} &\text{minimize } J = \textstyle\sum_{n=1}^{L-1} (z[n] - f(z[n-1]))^2 \\ &\text{subject to } \tfrac{1}{L} \textstyle\sum_{n=0}^{L-1} (z[n] - \bar{z})^2 = \sigma_x^2 \end{aligned} \tag{6.15}$$

where L is the length of the filtered data record, the variance of the $z[n]$ is estimated from the data by the sample variance, and $\bar{z}$ is the sample mean

$$\bar{z} = \frac{1}{L} \sum_{n=0}^{L-1} z[n].$$

The objective function J depends on the filter parameters through (6.14). The value of σ_x^2

may be determined directly from the map f using the techniques of Chapters 3 and 4. Of course, since we are optimizing with respect to the impulse response parameters we will necessarily be restricted to FIR filters $c[n]$. However, this restriction does not appear to be serious since we are free to choose arbitrarily long filters.

In order to determine whether optimization of the constrained objective function (6.15) is a potentially useful method for deconvolution, a set of computer simulations has been performed. In each experiment, the optimization (6.15) has been performed on a time series generated by filtering a sequence generated from an eventually-expanding map. The one-dimensional map used in each experiment is of the form

$$f(x) = \begin{cases} \frac{7}{3}x + \frac{3}{10} & 0 \leq x < \frac{3}{10} \\ -\frac{4}{3}x + \frac{7}{5} & \frac{3}{10} \leq x < \frac{6}{10} \\ -\frac{3}{2}x + \frac{3}{2} & \frac{6}{10} \leq x \leq 1 \end{cases} .$$

This map was chosen because of its relatively simple form and because its power spectrum is not white. Its associated time series are therefore not signals for which a traditional whitening approach to parameter estimation would work.

The following three filters were used to generate the data for the experiments:

- The linear phase filter

$$B_1(z) = 0.0565 - 0.1885z^{-1} + 0.5100z^{-2} - 0.1885z^{-3} + 0.0565z^{-4}.$$

- The all pole filter

$$B_2(z) = \frac{1}{A(z)}.$$

- The all pass filter

$$B_3(z) = \frac{A(z^{-1})}{A(z)}.$$

The polynomial $A(z)$ used in the definition of $B_2(z)$ and $B_3(z)$ is defined to be

$$A(z) = 1 + 0.78z^{-1} + 0.58z^{-2} - 0.04z^{-3} - 0.11z^{-4} - .0882z^{-5}.$$

Filter	Initial Value	Final Value
H_1	452	0.025
H_2	1630	0.12
H_3	84.3	0.038

Table 6.1: Initial and Optimized Values of the Objective Function J

and both $B_2(z)$ and $B_3(z)$ are causal filters. Neither B_1 nor B_3 has a stable causal inverse; they were chosen to evaluate the ability of the objective function (6.15) to approximately invert non-minimum-phase systems . On the other hand, $B_2(z)$ has a causal FIR inverse. It was chosen to evaluate how the objective function performs under ideal conditions—as mentioned above, we are restricted in practice to FIR filters $c[n]$.

For each filter, 50 experiments were performed. Each time series was generated with a random, uniformly distributed initial condition between 0 and 1. For each time series, the optimization was performed using the optimization routines of the commercial software package Matlab. In each case, the length of the filter $c[n]$ is fixed somewhat arbitrarily at 19 taps. Two measures of the effectiveness of the algorithm have been computed. First, Table 6.1 presents a measure of how well the dynamics are matched before and after the optimization is performed. More specifically, it shows the value of the objective function J before the optimization and the value after optimization. Second, Table 6.2 displays a measure of how well the filter $c[n]$ inverts the effect of the B_i. The measure has the form

$$\rho_i(c[n]) = \frac{\sum_n |(h_i * c)[n]| - \max_n(|(h_i * c)[n]|)}{\sum_n |(h_i * c)[n]|}.$$

The quantity ρ_i takes values between 0 and 1 and is a measure of concentration of the composite impulse response. Loosely speaking, smaller values of ρ_i indicate a more impulse-like sequence with most of the cumulative magnitude of its values concentrated in just one of its samples. Larger values of ρ_i indicate that the impulse response is spread out. Note that ρ_i is invariant to scaling of $c[n]$. The table presents the value of ρ_i for the initial and optimized values of $c[n]$. The numbers presented in both tables are averages over all of the experiments. An example of the result of the technique is shown in Figure 6-11. The upper

Filter	Initial Value	Final Value
H_1	0.49	0.017
H_2	0.66	0.033
H_3	0.81	0.041

Table 6.2: Initial and Optimized Values of the Dispersion Measure ρ_i

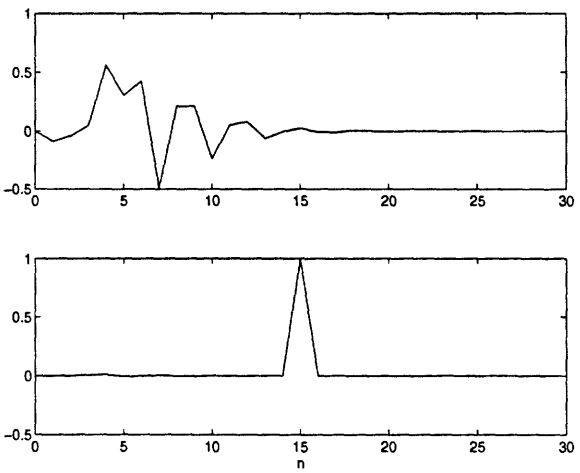


Figure 6-11: Example of an Equalized Impulse Response

plot is a graph of the impulse response associated with the system function $H_3(z)$. The lower plot is a graph of the overall impulse response $h_3[n] * c[n]$ after the optimization has been performed. This is the result of a single experiment and does not represent an average value.

For each experiment the optimization routine was initialized with a filter consisting of all zeros with the exception of the tenth tap, which was set to $\frac{\sigma_x}{\sigma_y}$. Using this initialization, the optimization routine did not always converge to a good inverse filter; that is, for some experiments, the value of the objective J was large even after the optimization was performed. For the filter H_1, the optimization routine converged to a "good" inverse filter in 49 out of 50 runs; for H_2 44 out of 50 runs gave good results; and for H_3 all 50 runs gave good results.

These computational experiments are not meant to be an exhaustive performance evaluation. Rather, they are meant to indicate the potential of the objective function and

constraint of (6.15) to determine an approximate inverse filter and recover a filtered chaotic signal. There are several issues raised by these preliminary results that seem worthy of further study.

- It is natural to ask whether the objective function has a unique minimum value. The lack of consistent convergence cited above suggests that the constrained optimization problem (6.15) has local minima. Further study of this issue may provide some insight into how the result of any optimization procedure depends on its initialization.
- Since the practical implementation of this technique attempts to invert one filter with a finite length filter we may ask how the performance of the technique depends on the filter length.
- It is possible that the performance of the technique may depend on both the filter and the chaotic map. These effects should be studied.
- The experiments above were performed in a noise free setting. The effects of noise on the objective function deserves further study.
- The possibility exists that an efficient optimization, taking advantage of the specific structure of the objective function, may exist. This issue should also be explored.

6.5 Fractal State Trajectories

When a filter preserves one-state determinism, the graph of the output map is well estimated from the filter output $y[n]$ by a scatter plot, i.e. a plot of $y[n-1]$ versus $y[n]$. The set of points in the scatter plot is necessarily "well behaved" since it is the graph of a piecewise smooth function. However, when one-state determinism is not preserved, the return map may have a very complex structure. Example 2 of Section 6.1 suggests that an IIR filter may generate a return map with a fractal structure with the fractal dimension depending on the filter impulse response. This is a rather remarkable possibility since the input map does not, in general, have any fractal structure. In order to gain some insight into how the fractal return map arises, this section examines an example admitting a fairly complete analysis.

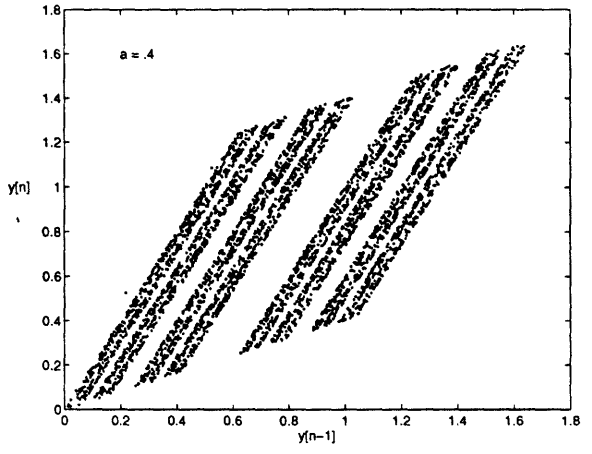


Figure 6-12: Return Map Showing Fractal Structure in the Output of a Single-Pole System Driven by a Chaotic Signal

Consider an input map f of the form

$$f(x) = 2x \bmod 1 = \left\{ \begin{array}{ll} 2x & 0 \leq x < \frac{1}{2} \\ 2x-1 & \frac{1}{2} \leq x < 1 \end{array} \right\}.$$

Suppose that time series generated by f drives the stable causal single pole AR filter with system function $B(z) = \frac{1}{1-az^{-1}}$. The filter output satisfies $y[n] = ay[n-1] + x[n]$. When $y[n-1]$ is plotted vs. $y[n]$ as in Figure 6-12, the result has a Cantor set cross section, as we will now show.

The approach we take to analyzing this set is to examine the state space of composite system consisting of both the chaotic map and the filter. We will then infer the properties of the filter output from the properties of the state space of the composite system. We will see in Chapter 7 that this approach is also useful for understanding the effects of filtering on invertible chaotic systems.

Consider the composite system

$$\left[\begin{array}{c} x[n] \\ y[n] \end{array} \right] = \left[\begin{array}{c} f(x[n-1]) \\ ay[n-1] + f(x[n-1]) \end{array} \right],$$

describing the states of the chaotic system and the filter. We wish to determine the set of

values taken by the state vector in the steady state. We denote this set by Γ. The set Γ is a subset of the plane. Its points can be determined by first calculating the set of values that the filter state can possibly take for a fixed value of the input, then taking the union of these sets over all input values. We detail this procedure below.

We wish to find a relationship between the state of the filter and the value of the filter input. The filter state at time n_0 is

$$y[n_0] = \sum_{k=0}^{\infty} a^k x[n_0 - k], \tag{6.16}$$

where transient behavior is ignored by assuming that the system has been running for a sufficiently long time. We now fix the value of the input at time n_0 to $x[n_0] = x^*$ and determine the possible values of the output. Since the chaotic map is noninvertible, many different input sequences satisfy $x[n_0] = x^*$. It follows that $y[n_0]$ may take many values. We will see below that the multivalued nature of the map from $x[n]$ to $y[n]$ induces the fractal structure apparent in Figure 6-12.

If f were invertible, (6.16) could be rewritten in the form

$$y[n] = \sum_{k=0}^{\infty} a^k f^{-k}(x[n]). \tag{6.17}$$

However, since f is two-to-one, it follows that f^{-1} has two branches, f^{-2} has four branches, and f^{-N} has 2^N branches. Suppose we wish to construct a sequence $x[n]$ such that $x[n_0] = x^*$. The value of $x[n_0 - 1]$ is one of two possibilities given by one of the two branches of f^{-1}. Similarly, given the value at time $n_0 - 1$, the value of the sequence at time $n_0 - 2$ is given by one of the two branches of f^{-1}. Continuing in this manner, we can construct a sequence satisfying $x[n_0] = x^*$ by making a series of binary decisions. Thus, each chaotic time series generated by f with value $x[n_0] = x^*$ can be associated with a binary fraction $c = .c_1c_2c_3\ldots$ where c_k is either zero or one depending on whether the left or right branch of the inverse function was taken at time $n_0 - k$.

Given a sequence of binary decisions c, the inverse function is well defined and has the

form

$$f_c^{-k}(x) = \frac{x}{2^k} + b_c[k]$$

where the offset parameter $b_c[k]$ satisfies

$$b_c[k] = \begin{cases} \frac{1}{2}b_c[k-1] & \text{if } c_k = 0 \\ \frac{1}{2}b_c[k-1] + \frac{1}{2} & \text{if } c_k = 1 \end{cases}.$$

From the definition of c it follows that

$$b_c[k] = \sum_{i=1}^{k} \frac{c_{i-1}}{2^i}. \tag{6.18}$$

Substituting (6.18) into (6.17) we obtain that the filter outputs associated with $x[n] = x^*$ are given by

$$\begin{aligned} y[n] &= \sum_{k=0}^{\infty} a^k \left(\frac{x^*}{2^k} + b_c[k]\right) \\ &= x^* \frac{a}{2-a} + \sum_{k=0}^{\infty} a^k \sum_{i=1}^{k} \frac{c_{i-1}}{2^i} \\ &= x^* \frac{a}{2-a} + \sum_{i=1}^{\infty} \frac{c_{i-1}}{2^i} \sum_{k=i-1}^{\infty} a^k \\ &= x^* \frac{a}{2-a} + \frac{1}{a(1-a)} \sum_{i=1}^{\infty} c_{i-1} \frac{a^i}{2^i}, \end{aligned} \tag{6.19}$$

where the first equality follows from the definition of $y[n]$ and the inverse function; the second equality follows from the definition of $b_c[k]$; the third and fourth equalities result from rearranging the innermost summations and simplifying sums of exponentials. The structure of the set of state values may be determined by examining how $y[n]$ depends on the binary sequence c. In fact, each value of c corresponds to a different sequence $.c_1c_2\ldots$ so the final summation (6.19) can be interpreted as a fraction in base $\frac{a}{2}$.

With $y[n]$ expressed in this form we proceed with the calculation of the box dimension (see Section 2.1.2) of the set Γ. First note that the $y[n]$ depends in an affine way on x^*. Each value of c provides a different offset term for the affine function. Thus, we expect the set Γ to have the form of a parallel set of lines. Suppose first that a is positive and choose

boxes of length $\delta_n = \frac{a}{2}^n$. For fixed x^*, all $y[n]$ corresponding to bit sequences c agreeing in the first n bits will fall within one δ_n box. There are 2^n such sequences. Along the x axis, $\frac{1}{\delta_n}$ boxes are required. It follows that $N_{\delta_n} = 2^n \times \delta_n^{-1}$. The box dimension is calculated as

$$\begin{aligned} \dim_B \Gamma &= \lim_{n\to\infty} \frac{N_{\delta_n}}{-\log \delta_n} \\ &= \lim_{n\to\infty} \frac{\log 2^n \times \delta_n^{-1}}{-\log \delta_n} \\ &= 1 + \frac{\log 2}{\log 2 - \log a} \end{aligned}$$

A similar argument applies to the case a negative. The final result is

$$\dim_B \Gamma = 1 + \frac{\log 2}{\log 2 - \log |a|} \tag{6.20}$$

According to (6.20), the dimension of the occupied set in the state space increases continuously from a minimum of $\dim_B = 1$ when the filter pole is at the origin ($a = 0$) to a maximum of $\dim_B = 2$ when the pole approaches the unit circle. This is in agreement with empirical observations.

The fractal dimension of the filter output's return map can be obtained from the above results. We wish to examine points of the form $[y[n-1], y[n]]$. From the definition of the filter, we obtain the relationship $y[n-1] = \frac{y[n]-x[n]}{a}$. Thus, the points of Γ are related to the points of the return map by the linear transformation

$$\begin{bmatrix} y[n-1] \\ y[n] \end{bmatrix} = \begin{bmatrix} \frac{-1}{a} & \frac{1}{a} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x[n] \\ y[n] \end{bmatrix}.$$

The box dimension of a fractal set is not changed by linear transformations so (6.20) holds for the return map as well.

Remark 1: The fractal structure in the return map is linked to the noninvertibility of the chaotic map f through the choice of binary sequence c.

Remark 2: The approach used to determine the fractal dimension of the return map hinged on associating a unique binary sequence with each time series taking a fixed value at time n. The idea of associating sequence taking a finite number of values (in this case zero and one) with a chaotic sequence is referred to as symbolic dynamics. All eventually-expanding maps may be described by symbolic dynamics [45, 19] so it is possible that the technique presented here may be generalized to handle a larger class of chaotic maps and filters.

Chapter 7

Observation Effects: Invertible Maps

Although the one-dimensional maps on which we have focused until this point form a large and useful class of systems generating chaotic signals, they are inherently noninvertible. Many interesting chaotic systems, however, are invertible. Invertible chaotic maps may arise, for instance, when studying time sampled ordinary differential equations. This type of map often occurs when analyzing physical systems and is well known to be invertible [3]. In this chapter, we extend some of the results of Chapter 6 concerning observation and measurement effects to a class of invertible maps.

The focus of this chapter is on chaotic systems that are dissipative, finite dimensional diffeomorphisms observed through sensors with linear time invariant dynamics. We concentrate on dissipative diffeomorphisms because they model a variety of physical systems [11] and are analytically tractable. We focus on linear time invariant sensor dynamics for similar reasons. Linear time invariant systems are a well defined analytically tractable class that reasonably model a variety of physical sensors. In contrast to the chaotic systems examined in the Chapters 3, 4, 5, and 6, chaotic diffeomorphisms are invertible and have more than one state variable. However, the effects of filtering on diffeomorphisms can be precisely characterized using tools that are similar to those of Chapter 6.

The chapter is organized as follows. Mathematical representations of filtered chaos are

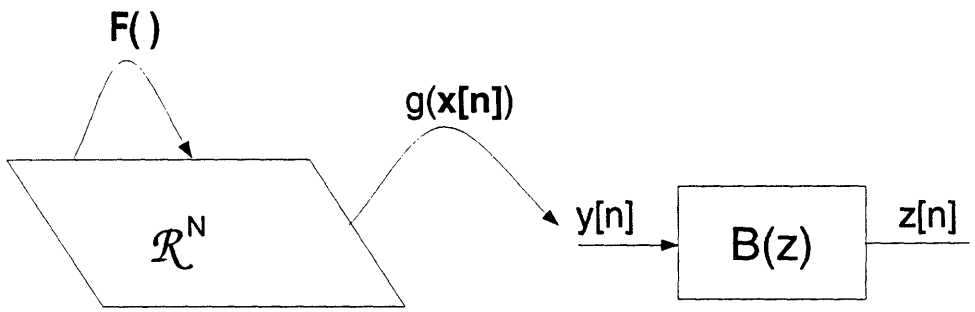


Figure 7-1: A Schematic Representation of a Filtered Chaotic Signal

discussed in Section 7.1. These representations lead to an interpretation of filtered chaos as a transformation of the underlying chaotic system that is used in subsequent sections to determine the structure of the filtered signal. The effects of filtering on Lyapunov exponents is detailed in Section 7.4.

7.1 Filtered Chaos: Representations

In order to begin the study of the effects of linear time invariant sensors dynamics on chaotic signals, it is natural to ask how the properties of the measured signal are related to those of the underlying system. The answer to this question is potentially important, for example, in interpreting chaotic data measured from some physical system. The specific situation that we will analyse is illustrated schematically in Figure 7-1.

The chaotic dynamical system $\mathbf{F}$ may model for example some chaotic phenomena that we would like to measure. Its state at time n is denoted $\mathbf{x}[n]$, and, unlike the one dimensional maps we have seen before, is assumed to be a vector in the space $\mathbb{R}^N$ with $N > 1$. The signal $y[n]$ that drives the filter is generated by a scalar observation of the state $\mathbf{x}[n]$ through the observation function g. The observation function maps state vectors to scalars and may be interpreted as a model of a scalar measurement of the state of the chaotic system. Finally, the linear filter with system function $B(z)$ models the dynamics of the sensor.

We assume that the system $\mathbf{F}$ is operating in the steady state and thus that the state vector is confined to some attractor $\Lambda \subset \mathbb{R}^N$. Because of the interpretation of the filtered chaotic signal as a measurement of some underlying chaotic phenomenon, we will often refer to $\mathbf{F}$ is as the underlying chaotic system and Λ as the underlying attractor.

The state sequence of the $\mathbf{F}$ and the scalar observation of the state are generated as follows:

$$\mathbf{x}[n] = \mathbf{F}(\mathbf{x}[n-1]) \tag{7.1}$$

$$y[n] = g(\mathbf{x}[n]) \tag{7.2}$$

The filter output is denoted by $z[n]$ and is generated by the convolution

$$z[n] = b[n] * y[n],$$

where $b[n]$ is the impulse response of the filter. In this chapter, we restrict attention on smooth, invertible dynamical systems $\mathbf{F}$ with smooth observations functions g. In particular, we assume that the chaotic map $\mathbf{F}$ is a diffeomorphism and that the observation function g is continuous with bounded partial derivatives. Furthermore, we consider only stable, causal filters. These technical assumptions allow a fairly precise analysis of the properties of the filter output. They are also consistent with a large class of physical situations and as such are not overly restrictive.

Our goal is to determine conditions under which inferences about the structure of the underlying system and its attractor can be made based on the filter's output. We will see that, in this regard, embedding theorems (see Chapter 2) are indirectly important. In particular, we will often perform time delay reconstruction using the filter's output. An embedding theorem is then used to infer the properties of what we have reconstructed.

There are two equivalent models of a filtered chaotic signal that will be useful in what follows. They reflect different interpretations of the filter output. We will see that the filter output may be viewed either as a modified scalar observation of the state of the system $\mathbf{F}$ or as an observation of a modified dynamical system. Each interpretation provides a separate insight into the effect of filtering on chaotic signals.

The interpretation of filtered chaos as a modified scalar observation of the chaotic system has been proposed in a less general form in the economics literature [46]. This interpretation has the advantage that no complexity is added to the underlying dynamics. Its drawback is that, in general, the modified observation function may have an analytically inconvenient

form. We circumvent this drawback by introducing a useful model that expresses filtered chaos as the output of a larger dynamical system. This modified system is the cascade of the chaotic system and the linear filter. Although the modified system has a larger state space than that of $\mathbf{F}$, this increased complexity is offset by increased tractability. We will see that the properties of the filter output may be related to those of this larger system in a straightforward manner. We now show how these interpretations arise.

7.1.1 Scalar Observation Function

For invertible chaotic system, the filter output can always be expressed as a complicated observation of the state. This fact seems to have been first noted by Brock [46] in the context of economic time series analysis. The filtered measurement $z[n]$ is related to the chaotic driving signal through convolution sum

$$z[n] = \sum_{i=0}^{\infty} b[i]y[n-i]. \tag{7.3}$$

Substituting Equations (7.1) and (7.2) into Equation (7.3) and taking advantage of the invertibility of $\mathbf{F}$, the filter output at each time can be expressed as a scalar observation of the state $\mathbf{x}[n]$ as follows :

$$\begin{aligned} z[n] &= \sum_{i=0}^{\infty} b[i]g(\mathbf{F}^{-i}(\mathbf{x}[n])) \\ &= \tilde{g}(\mathbf{x}[n]). \end{aligned} \tag{7.4}$$

When the impulse response represents a bounded–input–bounded–output stable linear system, the sum (7.3) converges for all $\mathbf{x}$ on the attractor of the nonlinear system. This interpretation is shown schematically in Figure 7-2.

Recall from Chapter 6 that when the noninvertible maps of the interval are filtered, the filter output can be expressed as a function of a single input sample when the filter has a finite impulse response length. When the filter has an infinite impulse response length, such a representation may or may not exist. In contrast, Equation (7.4) shows that when the underlying dynamics are invertible, the filter output can always be expressed as a function

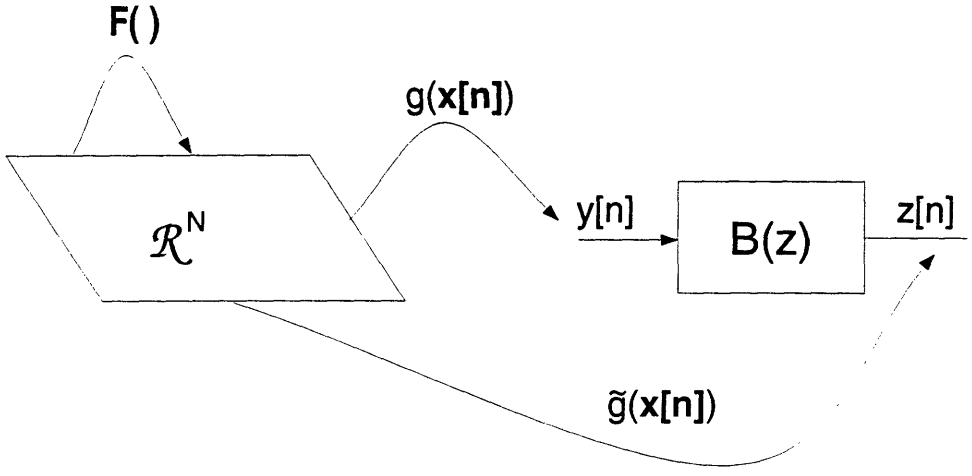


Figure 7-2: A Schematic Representation of Filtered Chaos Interpreted as a Complex Observation Function

.

of the current state of the underlying system regardless of the impulse response length.

The representation of filtered chaos as a scalar observation of the chaotic state can be used to show that in general, finite impulse response filters do not affect the fractal dimension. The following theorem makes the previous statement more precise.

Theorem 9 *It is a generic property that the fractal dimension (Hausdorff or Box dimension) of an attractor reconstructed from output of an FIR filter by time delay embedding equals the fractal dimension of the attractor of* **F**.

Theorem 9 is similar to a theorem of Brock [46]. Our result differs in approach and generality—Brock uses a different technique to prove his result, which holds only for box dimension.

We shall see below that a different approach is necessary in order to describe the effects filtering chaotic signals with infinite impulse response filters. In fact, although the observation function $\tilde{g}$ can be determined for general impulse responses, it may not provide much information about the relationship of the filter output to the underlying chaotic system. This is because infinite impulse response filters may give rise to observation functions $\tilde{g}$ that may not be differentiable. In this case, the hypotheses of the embedding theorem are not satisfied and, as a consequence the properties of the underlying attractor need not be

related to those of the attractor reconstructed from the filtered time series. In order to understand the effects of IIR filtering on chaotic signals, we take a different approach.

7.1.2 The Composite System

In order to understand the properties of the filter output, it is convenient to restrict attention to filters with strictly proper rational system functions. Since many useful non-rational system functions are well approximated by rational functions, this restriction does not appear to be too severe. When the filter has a rational system function, it has a linear state space representation and the filter output can be viewed as the output of a cascade of two systems—one generating the underlying chaotic signal and one corresponding to the filter dynamics. Specifically, the filter output $z[n]$ can be expressed as the output of the linear dynamical system, i.e.,

$$\begin{aligned} \mathbf{w}[n] &= \mathbf{A}\mathbf{w}[n-1] + \mathbf{b}g(\mathbf{x}[n]) \\ z[n] &= \mathbf{c}^T\mathbf{w}[n] \end{aligned}$$

where $\mathbf{w}[n]$ is the state of the filter at time n, $g(\mathbf{x}[n])$ is the chaotic input to the filter, and the matrix $\mathbf{A}$, and vectors $\mathbf{b}$, and $\mathbf{c}$ are a realization of the LTI system. The poles of the filter correspond the eigenvalues of the matrix $\mathbf{A}$ and the zeros of the filter are determined by the vectors $\mathbf{b}$, and $\mathbf{c}$. Many such realizations exist [47] but because we are primarily concerned with input-ouput behavior, we exclude those that have poles and zeros that cancel we consider only minimal realizations of the filter.

Since the underlying chaotic system and the filter dynamics are coupled, the filter output can be expressed as an observation of the state of a nonlinear dynamical system with state equations :

$$\begin{aligned} \mathbf{x}[n] &= \mathbf{F}(\mathbf{x}[n-1]) \\ \mathbf{w}[n] &= \mathbf{A}\mathbf{w}[n-1] + \mathbf{b}g(\mathbf{F}(\mathbf{x}[n-1])). \end{aligned} \tag{7.5}$$

The filter output is given by

$$z[n] = \mathbf{c}^T \mathbf{w}[n]. \tag{7.6}$$

In what follows, it is convenient to write Equation (7.5) in the more compact form

$$\begin{bmatrix} \mathbf{x}[n] \\ \mathbf{w}[n] \end{bmatrix} = \mathbf{s}[n] = \mathbf{Q}(\mathbf{s}[n-1]) \tag{7.7}$$

where the $\mathbf{Q}$ is defined by

$$\mathbf{Q}(\mathbf{s}) = \mathbf{Q}(\begin{bmatrix} \mathbf{x} \\ \mathbf{w} \end{bmatrix}) = \begin{bmatrix} \mathbf{F}(\mathbf{x}) \\ \mathbf{A}\mathbf{w} + \mathbf{b}g(\mathbf{F}(\mathbf{x})) \end{bmatrix}.$$

We refer to $\mathbf{Q}$ as the composite dynamics or the composite system and the vector $\mathbf{s}[n]$ as as the composite state. The composite system has dimension $M+N$ where N is the dimension of the state space of the underlying chaotic system $\mathbf{F}$ and M is the degree of the filter. Note that the composite system approach applies to both finite impulse response and infinite impulse response filters.

The filter output can be determined from the state of the composite system by the observation function

$$z[n] = \tilde{\mathbf{c}}^T \mathbf{s}[n] = \tilde{g}_1(\mathbf{s}[n])$$

where $\tilde{\mathbf{c}}$ is the $N+M$ dimensional vector

$$\tilde{\mathbf{c}} = \begin{bmatrix} \mathbf{0} \\ \mathbf{c} \end{bmatrix}$$

and $\mathbf{c}$ is the observation vector of the linear system in Equation (7.6). Since the observation function $\tilde{g}_1$ for the composite system is a linear function of the state, it is continuous and differentiable. The overall situation is illustrated in Figure 7-3.

The motivation for introducing the composite system is that in many circumstances its properties may be inferred from a smooth observation of its state. This is a direct consequence of the embedding theorem. Since the filter output is a smooth observation of

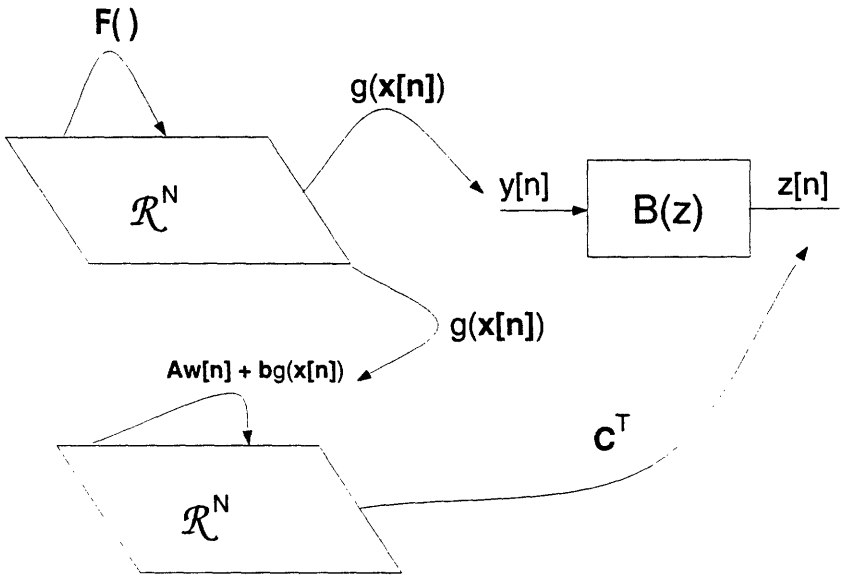


Figure 7-3: A Schematic Representation of Filtered Chaos Interpreted as the Output of a Composite System

the composite system state, an alternative view is that the properties of the filter output may be determined by understanding the properties of the composite system. It is the latter viewpoint that is most useful in the current context. In essence, the embedding theorem allows us to determine the properties of filtered chaos indirectly by determining the properties of the composite system.

Note that the composite system is differentiable because $\mathbf{F}$ and g are. Furthermore, when A is invertible, the transformation $\mathbf{Q}$ is invertible with inverse given by

$$\mathbf{Q}^{-1}(\mathbf{s}) = \mathbf{Q}^{-1}\left(\begin{bmatrix} \mathbf{x} \\ \mathbf{w} \end{bmatrix}\right) = \begin{bmatrix} \mathbf{F}^{-1}(\mathbf{x}) \\ A^{-1}(\mathbf{w} - \mathbf{b}g(\mathbf{x})) \end{bmatrix}.$$

Since $\mathbf{F}$ is a diffeomorphism and g is continuous and differentiable, $\mathbf{Q}$ has a continuous inverse with continuous derivative. Thus, when A is invertible, $\mathbf{Q}$ is a diffeomorphism and the composite dynamics and observation satisfy the hypotheses of the embedding theorem. We thus expect that many properties of the composite system may be determined from the filter output.

When A is not invertible, as for example when it represents filter with a finite impulse

response, the composite system $\mathbf{Q}$ may not be a diffeomorphism. It follows that the hypotheses of the embedding theorem are not met. However, in this case, Theorem 9 of Section 7.1.1 is applicable. Thus the combination of the composite system and the scalar observation representations cover both infinite and finite impulse response length filters.

As shown in Figure 7-3, the composite system consists of a nonlinear component and a linear component coupled through the observation function $g(\cdot)$. The state of the nonlinear system evolves independently of the state of the linear system. We thus expect that some of the properties of the underlying system are preserved by filtering. We shall see below that for Lyapunov exponents, this is the case. On the other hand, the extra state variables of the composite dynamics may influence the behavior of the overall system. We will next examine the relationship between the properties of the composite system $\mathbf{Q}$ and those of the underlying system $\mathbf{F}$.

7.2 Filtered Chaos: The State Space Transformation

The state of the composite system consists of two components: one due to the underlying nonlinear system and one due to the linear filter. In this section we develop the relationship between these two components. We shall see that to each point in the state space of $\mathbf{F}$ there corresponds a point in the state space of the composite system $\mathbf{Q}$. The correspondence between these points is determined by a transformation depending on both the filter and $\mathbf{F}$. The properties of this transformation will allow us to infer, among other things, the form of the attractor of the composite system and the fact that filtering cannot completely obscure chaotic structure in the filter input. We now turn to the problem of deriving this transformation.

Referring to Equation (7.7), we see that the second component of the composite state, $\mathbf{s}[n]$, is the state of a linear time invariant system. Its value at time n can be computed in terms of its initial value and all previous inputs through a convolution sum. More precisely, the state of the composite system at each time is

$$\mathbf{s}[n] = \begin{bmatrix} \mathbf{x}[n] \\ \mathbf{w}[n] \end{bmatrix} = \begin{bmatrix} \mathbf{x}[n] \\ \mathbf{A}^n\mathbf{w}[0] + \sum_{k=0}^{n} \mathbf{A}^k \mathbf{b} g(\mathbf{F}^{-k}(\mathbf{x}[n])), \end{bmatrix}. \tag{7.8}$$

where we have used the relationship $x[n-k] = F^{-k}(x[n])$. Notice that the filter state, $\mathbf{w}[n]$, depends on the filter's initial condition. Although this fact is an elementary result from linear systems theory, it has an important implication when the filter is stable. In this case, the term $\mathbf{A}^n\mathbf{w}[n]$ approaches zero as n gets large. Thus, in the long time limit the initial state of the filter is negligible and the filter state depends only on its input. Because the underlying chaotic system $\mathbf{F}$ is invertible, the entire input state sequence is determined by its current value and thus, in the steady state, the filter's state depends only on the current state of the underlying dynamics. In what follows, we assume that the filter has reached the steady state and consequently that its the effect of its initial condition is negligible.

To simplify the notation, we rewrite (7.8), which describes the state of the composite system at time n, in the form

$$\mathbf{s}[n] = \begin{bmatrix} \mathbf{x}[n] \\ \mathbf{T}(\mathbf{x}[n]) \end{bmatrix}, \tag{7.9}$$

where $\mathbf{T}(\mathbf{x}) = \sum_{k=0}^{\infty} \mathbf{A}^k \mathbf{b} g(\mathbf{F}^{-k}(\mathbf{x}))$. and the infinite upper limit in the sum defining $\mathbf{T}(x)$ is a consequence of the assumption of steady state operation Equation (7.9) shows that the state of the composite system operating in the steady state is a function only of the current state of the underlying chaotic system. We show next that this relationship between the composite state vector $\mathbf{s}[n]$ and the underlying state vector $\mathbf{x}[n]$ is a useful tool for examining the effects of filtering on chaotic signals.

Equation (7.9) implies that for any value of the state of the underlying chaotic system, the filter state is uniquely determined. Suppose we define the transformation $\mathbf{P}(\mathbf{x})$ by

$$\mathbf{P}(\mathbf{x}) = \begin{bmatrix} \mathbf{x} \\ \mathbf{T}(\mathbf{x}) \end{bmatrix}, \tag{7.10}$$

Then $\mathbf{P}$ maps the N dimensional state space of the underlying chaotic system, $\mathbf{F}$, to the $N + M$ dimensional state space of the composite system $\mathbf{Q}$. We call the map $\mathbf{P}$ the state space transformation associated with $\mathbf{F}$ and the filter.

The state space transformation, $\mathbf{P}$, relates points in the state space of the underlying chaotic system to points in the composite state space in the steady state. Strictly speaking, the map $\mathbf{P}$ should be viewed as a map from the attractor of the underlying system to

the composite state space since in the steady state, the state of the underlying system is confined to the attractor Λ. On the other hand, the primary assumptions used to derive the state space transformation were that the initial condition of the filter is negligible, and the sum $\mathbf{T}(\mathbf{x}) = \sum_{k=0}^{\infty} \mathbf{A}^k \mathbf{b} g(\mathbf{F}^{-k}(\mathbf{x}))$ converges. Thus, the state space transformation is valid for all $\mathbf{x}$ for which these two assumptions hold. This set may be larger than the attractor.

In the next section, we examine how the properties of the state space transformation influence the properties of the composite system.

7.3 Fractal Dimension of the Composite Attractor

When the attractor Λ of the underlying chaotic system $\mathbf{F}$ is a fractal set, the attractor Γ of the composite system $\mathbf{Q}$ is the image of a fractal set through the state space transformation. Thus, the problem of determining the fractal properties of the composite attractor is related to the problem of determining how fractal sets behave under transformations. The problem of determining the fractal properties of the composite attractor is complicated by the fact that, although the state space transformation is continuous[1], continuous transformations may change the characteristics of a fractal set. However, the state space transformation, $\mathbf{P}$, is not an arbitrary continuous transformation. Its special form allows us to determine a condition, which depends on the filter, under which the dimension of the composite attractor is exactly equal to that of the underlying attractor. This result is potentially important when performing time series analysis with measured chaotic signals since as we shall see, it implies that some filters provide a more faithful measurement of chaos than others.

7.3.1 Weak Filters and Fractal Dimension

We will now determine conditions under which filtering does not change the fractal dimension of a chaotic attractor. Motivated by the discussion of fractal sets Chapter 2 and in particular Theorem 1, we make the following definition.

Definition 4 *A causal, filter with rational system function is a called weak with respect to the pair* $\mathbf{F}, g$ *if the state space transformation it induces is bi-Lipschitz.*

[1] Continuity of the state space transformation is established in Appendix C

We will refer to filters that are weak with respect to pair $(\mathbf{F}, g)$ as weak filters when $(\mathbf{F}, g)$ is obvious from context. Combining Theorem 1 with the definition of a weak filter immediately yields the following:

Theorem 10 *The attractor of a composite system,* $\mathbf{Q}$ *induced by a weak filter has fractal dimension that is equal to the attractor dimension of the underlying system* $\mathbf{F}$.

In other words, Theorem 10 says that weak filters do not change fractal dimension.

Weak filters were defined above in an indirect manner. The following theorem, which is proved in Appendix C, directly provides a characterization of a subclass of weak filters in terms of state space filter representations.

Theorem 11 *Suppose that* $\mathbf{F}$ *is a dissipative diffeomorphism with Lyapunov exponents* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ *and* $\mathbf{A}$ *has eigenvalues* $\gamma_1, \ldots, \gamma_M$ *ordered such that* $|\gamma_1| \geq |\gamma_2| \geq \cdots |\gamma_M|$. *Then if* $\log(|\gamma_1|) < \lambda_N$ *the filter is weak.*

Theorem 11 provides an easily verified sufficient condition for a filter to be weak. A necessary and sufficient condition would provide a complete characterization of the class of weak filters; however, such a condition appears to be more difficult to determine. In any event, the sufficient condition of the theorem does have a potentially important consequence. In particular, the fractal dimension of the composite attractor and the fractal dimension of the underlying attractor will be equal when the filter is weak with respect to the underlying dynamics. Since the composite attractor represents the measurement through a linear time invariant system, Theorem 11 implies that if the fractal dimension of a measured chaotic signal must be determined precisely, the sensor should be designed to be weak with respect to the dynamics being measured.

Theorem 11 has an intuitive explanation based on the interpretation of Lyapunov exponents as a measure of sensitive dependence on initial conditions. It can be shown (Yamamoto [48]) that the Lyapunov exponents of a linear system with dynamics characterized by the matrix $\mathbf{A}$ are given by $\{\log(|\gamma_i|)\}$ where $\{\gamma_i\}_{i=1}^{M}$ is the set of eigenvalues of $\mathbf{A}$. It follows that stable linear systems have negative Lyapunov exponents. The state space transformation

has the form

$$\mathbf{P}(\mathbf{x}) = \begin{bmatrix} \mathbf{x} \\ \sum_{k=0}^{\infty} A^k \mathbf{b} g(\mathbf{F}^{-k}(\mathbf{x})) \end{bmatrix}$$

and so depends on iterates of the inverse system $\mathbf{F}^{-1}$. According to the results of Ruelle [49], the Lyapunov exponents of the inverse system are simply related to those of the forward system. Specifically, Lyapunov exponents of $\mathbf{F}^{-1}$ are the simply those of $\mathbf{F}$ but with opposite signs. In other words, if the Lyapunov exponents of $\mathbf{F}$ are ordered such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ then $\mathbf{F}^{-1}$ has exponents $-\lambda_N \geq -\lambda_{N-1} \geq \cdots \geq -\lambda_1$.

With this in mind, the condition $\log(|\gamma_1|) < \lambda_N$ can be seen to have the following explanation. Since the filter is stable, it tends to contract distances in the composite state space at a rate that is at least $\log(|\gamma_1|)$ per iterate. The inverse system $\mathbf{F}^{-1}$ can expand distances in the composite state space at an asymptotic rate no greater than $|\lambda_N|$. The condition $\log(|\gamma_1|) < |\lambda_N|$ implies that when the rate of contraction due to the filter dominates the rate of expansion due to the inverse chaotic system the filter is weak. In terms of the forward system, we see that the filter response must decay faster than the rate defined by the minimum Lyapunov exponent of $\mathbf{F}$. When this is the case, the state space transformation is differentiable and no change in fractal dimension occurs.

7.3.2 Potential Increase in Fractal Dimension

Until now, we have concentrated on determining filters that do not change the fractal dimension of the attractors we wish to measure. Neither finite impulse response length filters nor weak infinite impulse response filters will affect the dimension. One may ask what happens when an infinite impulse response length filter is not weak. In this section, we present some empirical evidence that suggests that fractal dimensions may potentially be changed by filters that are not weak.

It is difficult to quantify the effect of filtering on fractal dimension when the filter is not weak. There is however empirical evidence that filtering may cause the composite attractor (and hence the reconstructed attractor) to have higher fractal dimension than the underlying attractor [36, 42, 50, 35, 51, 41]. Figure 7-4 illustrates this phenomenon with a plot of estimated box dimension versus filter parameter for an attractor reconstructed from

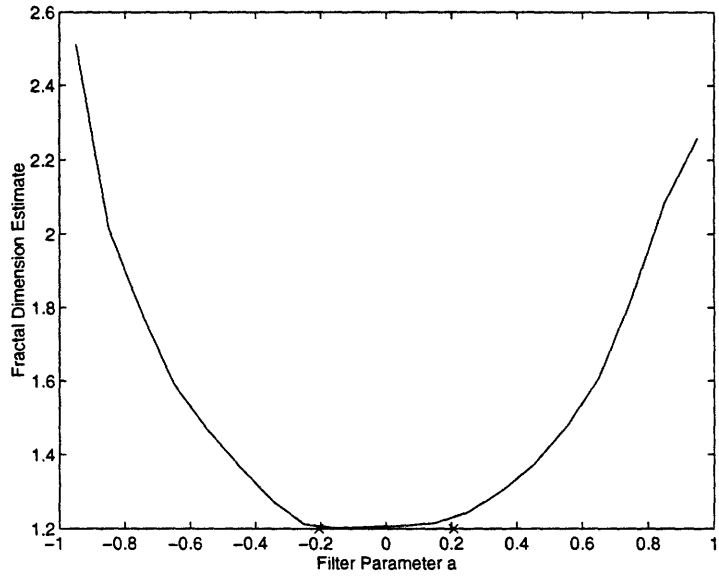


Figure 7-4: Empirically Observed Increase in Fractal Dimension Due to Filtering

the output of a single pole filter driven by a chaotic signal. The chaotic signal is generated by the Henon map

$$\begin{bmatrix} x[n] \\ y[n] \end{bmatrix} = \begin{bmatrix} 1 + y[n-1] - \alpha x^2[n-1] \\ \beta x[n-1] \end{bmatrix}. \tag{7.11}$$

The filter output is generated by the one dimensional linear state space system

$$z[n] = az[n-1] + x[n].$$

The fractal dimension computed is the capacity dimension and is computed from a data set of 10000 state vectors of the composite system using a standard algorithm []. The fractal dimension was computed at 20 values of the filter parameter between $a = -0.95$ and $a = 0.95$. The actual dimension of the underlying attractor is approximately 1.2.

There are several interesting features apparent in the graph of Figure 7-4. First, the fractal dimension seems to show a negligible increase due to filtering for values of the filter parameter with absolute value less than about 0.2. The Lyapunov exponents of the Henon map are approximately log(1.52) and log(0.198). This would appear to be consistent with the analysis of the previous section since in the range $-0.198 < a < 0.198$ the filter is

weak and we expect no increase in dimension due to filtering. Second, the fractal dimension plot is nearly symmetric about the origin. This may reflect a relation between the filter parameter and the fractal dimension of the composite attractor that depends on the radius of the filter pole and not its phase. Finally, the observed fractal dimension seems to increase as the pole radius increases. This may reflect the fact the as the pole radius increases, the filter is farther away from being a weak filter. Perhaps filters with strong dynamics are able to effect a stronger change in the fractal dimension of the composite attractor.

FIR vs. IIR Filtering

The above results may be summarized as follows. Weak filters cannot increase fractal dimension while other filters may. All finite impulse response filters are weak, while some infinite impulse response filters are not. The empirical results of the previous section imply that IIR filtering may cause an increase in the dimension of the attractor observed at the filter output. In light of the fact that any IIR filter with a rational system function can be approximated arbitrarily well by an FIR filter of sufficient length, it is not clear why the two types of filters should have fundamentally different effects on chaotic signals.

This issue has not been fully resolved and deserves further study. However the following heuristic discussion indicates a fundamental difference between FIR and IIR filtering that may explain the different effects on chaotic signals. Fractal sets display scaling behavior at all length scales. Stable, rational IIR filters have exponentially decaying impulse responses that also display scaling behavior at all scales. It seems that this infinite scaling property is what allows IIR filters to distort the fractal structure of chaotic signals.

The impulse response values of FIR filters must show scaling behavior over at most a range of scales—no matter how long an FIR filter is, there is an impulse response value that is largest in magnitude and one that is smallest. It seems that "scale limited" nature of FIR filters is what limits their ability to distort the fractal structure of a chaotic signal. On the other hand, since all numerical techniques for estimating fractal dimensions from data operate only over a finite range of scales, FIR filters may have an effect on numerically computed dimension estimates.

7.4 Lyapunov Exponents for the Composite System

In this section we explore the effects of filtering on Lyapunov exponents. A defining property of chaotic system is sensitive dependence on initial conditions, which is characterized by the Lyapunov exponents of the system. As mentioned in Chapter 2, Lyapunov exponents may be interpreted as the divergence rate of state sequences started from initial conditions that are slightly perturbed from one another. Because the composite system consists of a nonlinear chaotic system driving a linear system the effect of perturbations to each subsystem may be examined separately. Changes to the state of the nonlinear system affect the state of the linear system. However, because there is no feedback from the filter output to the nonlinear subsystem, a perturbation to the state of the linear subsystem has no effect on the state of the nonlinear subsystem. It seems plausible therefore that the Lyapunov exponent of the composite system can be broken into two groups, one due to the linear system and one due to the combination of the nonlinear and linear systems.

We will see in this section that the group of exponents due to the linear system can be computed directly by computing the magnitudes of the eigenvalues of the linear system. However it is not clear that the second group of exponents are simply related to the Lyapunov exponents of the system $\mathbf{F}$. We will explore these issues below.

The Lyapunov exponents are computed from the derivative matrix of the dynamics as described in Chapter 2. The derivative matrix of the composite system has the form

$$D\mathbf{Q}(\mathbf{s}) = \begin{bmatrix} D\mathbf{F}(\mathbf{x}) & \mathbf{0} \\ \mathbf{b}dg(\mathbf{F}(\mathbf{x}))D\mathbf{F}(\mathbf{x}) & \mathbf{A} \end{bmatrix}. \tag{7.12}$$

The block triangular form of the right side of Equation (7.12) is a consequence of the manner in which the linear and nonlinear systems are coupled; the zero matrix shows that perturbations to the state of the linear system do not affect the state of the underlying nonlinear system.

The Lyapunov exponents depend on the behavior of the product

$$(D\mathbf{Q}^n)(\mathbf{s}) = \prod_{n}^{1} D\mathbf{Q}(\mathbf{s}[i]) \tag{7.13}$$

where $\mathbf{s}[i] = \mathbf{Q}^i(\mathbf{s})$ is the state value at time i. Referring to Equation (7.12), we can derive an expression for the matrix product in terms of the component matrices of $D\mathbf{Q}(\mathbf{s})$. The product has the form

$$(D\mathbf{Q}^n)(\mathbf{s}) = \begin{bmatrix} (D\mathbf{F}^n)(\mathbf{x}) & \mathbf{0} \\ \sum_{j=0}^{\infty} A^{n-j}\mathbf{b}dg(\mathbf{F}(\mathbf{x}))(D\mathbf{F}^j)(\mathbf{x}) & \mathbf{A}^{n+1} \end{bmatrix}, \tag{7.14}$$

and is lower triangular since each term of the product is.

7.4.1 Exponents Due to the LTI Subsystem

In order to use expression (7.14) to determine the Lyapunov exponents for the composite system, we recall that the Lyapunov exponents satisfy the following equations

$$\lambda_i = \lim_{n\to\infty} \frac{1}{n} \log(\|(D\mathbf{Q}^n)(\mathbf{s})\mathbf{u}_i\|),$$

where $\mathbf{u}_i$ is a vector related to the i^{th} exponent. Consider first vectors of the form

$$\mathbf{u} = \begin{bmatrix} \mathbf{0} \\ \tilde{\mathbf{u}} \end{bmatrix}.$$

In this case, the product $(D\mathbf{Q}^n)(\mathbf{s})\mathbf{u}$ is equal to $A^k\tilde{\mathbf{u}}$. If $\{\gamma_i\}_{i=1}^M$ are the eigenvalues of A, a theorem of Yamamoto [48] can be invoked to show that the Lyapunov exponents in this case are simply the magnitudes of the eigenvalues of A. More specifically,

$$\lim_{k\to\infty} \frac{1}{k} \log(\|A^k\tilde{\mathbf{u}}_i\|) = \log(|\gamma_i|)$$

where $\tilde{\mathbf{u}}_i$ is in the span of the generalized eigenvectors corresponding to the eigenvalue γ_i. Thus, the heuristic reasoning outlined above holds for the linear portion of the composite system.

7.4.2 Exponents Due to the Nonlinear Subsystem

The negative the Lyapunov exponents corresponding to the underlying nonlinear system $\mathbf{F}$ are a bit more difficult to determine. However, the positive Lyapunov exponents can be determined in a manner analogous to the above. We consider vectors of the form

$$\mathbf{v} = \begin{bmatrix} \tilde{\mathbf{v}} \\ \mathbf{0} \end{bmatrix}.$$

In this case, the product of interest is of the form

$$(D\mathbf{Q}^n)(\mathbf{s})\mathbf{v} = \begin{bmatrix} (D\mathbf{F}^n)(\mathbf{x})\tilde{\mathbf{v}} \\ \sum_{j=0}^{\infty} A^{n-j}\mathbf{b}dg(\mathbf{F}(\mathbf{x}))(D\mathbf{F}^j)(\mathbf{x})\tilde{\mathbf{v}} \end{bmatrix}.$$

When $\tilde{\mathbf{v}}$ corresponds to an unstable direction of $\mathbf{F}(x)$ with Lyapunov exponent λ_i, straightforward but tedious manipulations show that

$$\lim_{n\to\infty} \frac{1}{n} \log(\|(D\mathbf{Q}^n)(\mathbf{s})\mathbf{v}\|) = \lambda_i. \tag{7.15}$$

A intuitive but less precise argument that Equation (7.15) is plausible for positive Lyapunov exponents is the following. Because of the form of the vector $\mathbf{v}$, the norm of the product matrix satisfies

$$\|(D\mathbf{Q}^n)(\mathbf{s})\mathbf{v}\| \geq \|(D\mathbf{F}^n)(\mathbf{x})\tilde{\mathbf{v}}\|.$$

On the other hand, the norm of the term

$$\sum_{j=0}^{\infty} A^{n-j}\mathbf{b}dg(\mathbf{F}(\mathbf{x}))(D\mathbf{F}^j)(\mathbf{x})\tilde{\mathbf{v}}$$

can be bounded by the sum of a constant and a term that is strictly less than a fixed multiple of $\|(D\mathbf{F}^n)(\mathbf{x})\tilde{\mathbf{v}}\|$. Since the vector $\tilde{\mathbf{v}}$ is assumed to correspond to a positive Lyapunov exponent, the limiting growth rate of the overall vector norm must be equal to λ_i.

In order to determine if the negative Lyapunov exponents of the system $\mathbf{F}$ contribute to the Lyapunov exponents of the composite system, a more delicate argument seems to be

necessary. However, empirically, it appears that the negative Lyapunov exponents of the nonlinear system are preserved. Two example are shown in Figures 7-5 and 7-6. Both

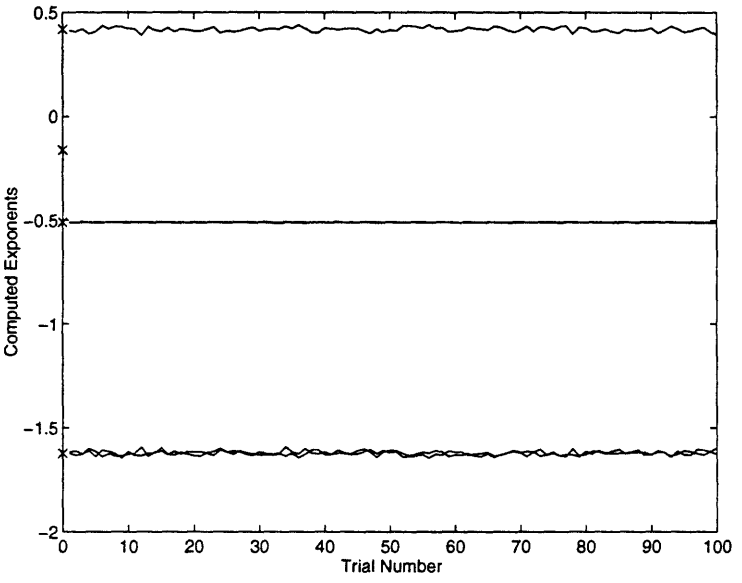


Figure 7-5: Empirically Determined Lyapunov Exponents for the Composite System Henon Map

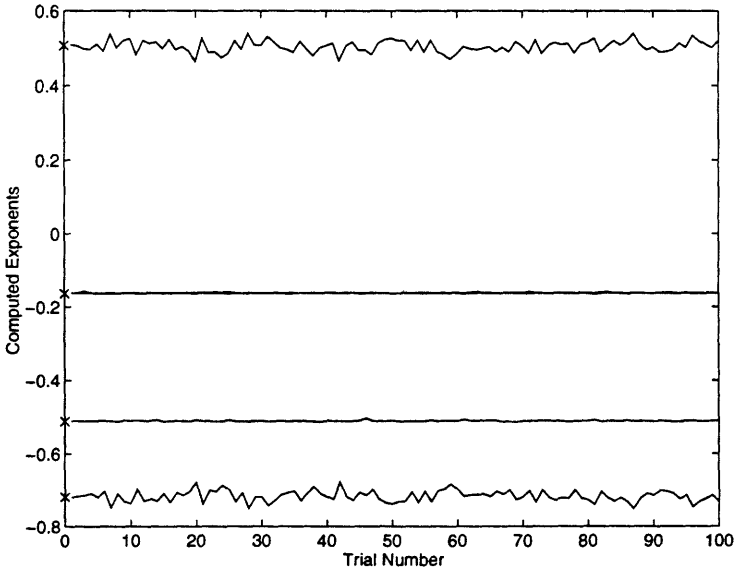


Figure 7-6: Empirically Determined Lyapunov Exponents for the Composite System: Ikeda Map

figures show Lyapunov exponents computed using the numerical algorithm suggested in

[52]. For these examples, the composite systems are formed from the cascade of an AR filter (poles at $z = 0.6, 0.85e^{\frac{5\pi}{8}}, 0.85e^{-\frac{5\pi}{8}}$) with a Henon map (Figure 7-5) or an Ikeda map (Figure 7-6). The equations defining the Henon maps are given by (7.11) while the Ikeda map is defined by

$$\begin{bmatrix} x[n] \\ y[n] \end{bmatrix} = \begin{bmatrix} 1 + u(x[n-1]\cos(t) - y[n-1]\sin(t)) \\ u(x[n-1]\sin(t) + y[n-1]\cos(t)) \end{bmatrix},$$

where $t = 0.4 - \frac{6}{(1+x^2+y^2)}$. Since the Lyapunov exponents are computed numerically from a finite number of iterates of the dynamics, their estimated values depend on the initial condition on the system. Figures 7-5 and 7-6 show the empirically determined Lyapunov exponents for one hundred random initial conditions. These figure seem to indicate that, at least for these two maps, the negative Lyapunov exponents of the nonlinear system are not affected by filtering. Indeed, extensive numerical simulations with these two chaotic systems and many linear filters leads us to conjecture that the negative Lyapunov exponents of general diffeomorphisms are unaffected by filtering.

Chapter 8

Contributions and Future Work

The previous chapters presented a study of some of the geometric and statistical properties of chaotic signals. This study was partially motivated by the need develop computational tools address time series analysis problems involving such signals. In this chapter, we review the major contributions of the thesis and suggest directions for future work.

8.1 Contributions

With regard to statistical properties, we defined the notion of statistics of chaotic signals by drawing on an analogy between them stochastic processes. Concentrating on the class of Markov maps in Chapter 3, we introduced a broad class of statistics and developed closed form expressions for them. These expressions, which are easily computed using standard vector and matrix operations, provided some additional insight into the statistical structure of Markov maps. In particular, we found the their power spectra were rational functions of frequency.

Markov maps are closely associated with the much larger class of eventually-expanding maps as we demonstrated in Chapter 4. We used this association to extend the computational techniques developed for Markov maps to the problem of approximating the statistics of eventually-expanding maps. Again, as with Markov maps, we inferred from this computational result some deeper properties of the statistics of the class.

From a practical standpoint, both the computational techniques for Markov maps and

for eventually-expanding maps are general enough to be widely applied. We used this generality in Chapter 5, where, after deriving a map describing the operation of a chaotic power converter, we suggested a method for determining the statistics of its voltage and current waveforms.

With regard to geometric properties, we presented in Chapter 6 a technique for analyzing the effects of linear convolution on chaotic signals generated by one-dimensional maps. Using this technique, we characterized the entire class of filters that does not alter the geometric structure of the signal. These results pointed to a fundamental ambiguity in the output of a filter driven by a one-dimensional chaotic signal; namely, we showed that there is no way to detect that an observed signal is the result of nontrivial filtering applied to a chaotic signal. We also showed through an example that a filtered one-dimensional chaotic system may potentially display a geometric complexity that is absent in the unfiltered signal.

As an example of the practical application of these results, we used the knowledge of the spectral structure of a particular Markov map and its response to filtering to synthesize a family of "chaotic white noise generators". We also suggested a strategy for recovering a filter's chaotic input given knowledge of the system from which it was generated.

Finally, in Chapter 7 we explored the effects of filtering on multidimensional, invertible chaotic systems. In this case, we determined conditions on the filter under which the geometric structure of the signal would not be obscured and discussed the implications of this result for the design of sensors used to measure chaotic signals.

8.2 Future Work

We conclude the thesis by listing several directions for potentially important extensions of the work we have presented.

1. **Realizing Prescribed Statistics with Chaotic Maps:** The results of Chapter 3 show that all Markov maps have rational spectra. It is natural to ask what subset of rational spectra can be realized by Markov maps. As mentioned in Chapter 3, the solution to this problem seems to be related to the properties of stochastic matrices. Perhaps more interesting from the point of view of applications is the problem of

synthesizing an eventually expanding map that realizes or approximates a prescribed power spectra. The solution to this problem is of potential use in certain signal generation applications.

One approach to solving the approximation problem is based on the fact there are a *finite* number of Markov maps corresponding to any fixed partition of the unit interval. One method of approximation is to first choose a fine partition of the interval and then compute the spectra of all possible Markov maps on this partition. The best match could then be selected. Although this "brute force" approach seems to be impractical—the number of possible Markov maps increases exponentially with the number of partition elements—and provides no insight into any structure that the solution may have, it nevertheless demonstrates that, given sufficient computational power, a solution could be determined.

2. **Approximation Error Bounds:** The results of Chapter 4 show that any eventually-expanding map is the statistical limit of a sequence of Markov maps. This result does not, however, indicate the approximation error introduced when a fixed Markov map is used to approximate an eventually-expanding map. Such bounds, which would be of interest in any application using the Markov approximation scheme, would seem to depend on the Frobenius–Perron operator. In particular, because of the close relation between the statistics and the FP operator of a chaotic system, one avenue to pursue in determining such bounds involves quantifying the difference in the FP operators of the chaotic system and its Markov approximation.

3. **More efficient Approximations of Eventually-Expanding Maps:** Although we suggested in Chapter 4 that many different Markov approximations of a eventually-expanding map were possible, we focused on an approximation of a specific form in our discussion. This raises the question as to whether another approximation exists that is optimal in some sense. Of course, to make this problem well-posed an appropriate optimality criterion must be defined. For example, given an eventually expanding map and set of statistics that we desire to compute, we would like a Markov map that "best" approximates the these statistics for a fixed number of partition elements. It

seems plausible that results concerning approximation error bounds could be useful in this context.

4. **An Extended Analysis of a Switching Power Converter:** One shortcoming of the results of Chapter 5 is that they do not provide any methods to assist in the design of power supplies that operate chaotically. Extensions of this work should address the issue of developing design rules and design techniques. Furthermore, future work should attempt to more precisely determine the advantages and disadvantages of operating power converters in the chaotic regime.

5. **Efficient Deconvolution Techniques:** No attempt was made in Chapter 6 to develop a computationally efficient scheme to perform the deconvolution algorithm that we presented there. Computationally efficient methods of performing this deconvolution could potentially make certain applications of chaotic signal practical. For example, since the our deconvolution technique provides an estimate of the inverse filter response, chaotic signals may perhaps be useful as training sequences for communication channel equalizers.

6. **Bounds on the Increase in Fractal Dimension Caused by Filtering:** One of the weaknesses of the analysis of Chapter 7 is that we have only established conditions under which the fractal dimension of a chaotic signal will not increase. An interesting issue is whether an increase in fractal dimension can be established and whether a bound on the increase, as a function of filter parameters, can be determined.

Appendix A

A Proof of the Statistical Convergence of a Sequence of Markov Approximations to an Eventually Expanding Map

By constructing a special sequence of Markov maps, we showed in Chapter 4 that the statistics of any eventually-expanding map can be approximated by those of a piecewise-linear Markov map. This key result is a consequence of the assertion that the sequence statistically convergences to the desired eventually-expanding map. In this appendix, we show that this assertion is indeed true. More specifically, we prove the following theorem.

Theorem 12 *Suppose f is an eventually expanding map. Then the sequence of piecewise linear Markov approximations f_n statistically converges to f.*

Proof: By the definition of statistical convergence, we must show that

$$R_{f_n,h_0,\ldots,h_l}[k_1,\ldots,k_r] \stackrel{n\to\infty}{\to} R_{f,h_0,\ldots,h_l}[k_1,\ldots,k_r]$$

for any continuous h_i and all nonnegative integers k_i and r. We proceed by first assuming that $f_n^k \to f$ uniformly for all finite k. We will then show that this assumption is indeed

true.

Under our assumptions, the theorem follows from the continuity of the integral and the weak convergence of the invariant densities (see Theorem 5). More specifically, since each h_i is continuous on $[0,1]$ it is uniformly continuous there. The difference between the statistics is

$$\begin{aligned}
&R_{f_n,h_0,\ldots,h_l}[k_1,\ldots,k_r] - R_{f,h_0,\ldots,h_l}[k_1,\ldots,k_r] = \\
&\quad \int h_0(x)h_1(f_n^{k_1}(x))\ldots h_l(f_n^{k_l}(x))p_n(x)dx - \\
&\quad \int h_0(x)h_1(f^{k_1}(x))\ldots h_l(f^{k_l}(x))p(x)dx \\
&= \int h_0(x)h_1(f^{k_1}(x))\ldots h_l(f^{k_l}(x))(p_n(x)-p(x))dx + \\
&\quad \int (h_0(x)h_1(f_n^{k_1}(x))\ldots h_l(f_n^{k_l}(x)) - h_0(x)h_1(f^{k_1}(x))\ldots h_l(f^{k_l}(x)))p_n(x)dx
\end{aligned}$$

The first term goes to zero by the weak convergence of p_n to p. The second term goes to zero by the uniform continuity of each h_i and the uniform convergence of $f_n^{k_i}$ to f^{k_i}.

The proof of the theorem is completed by establishing that the uniform convergence assumption made above always holds. This is the subject of the following lemma.

Lemma 2 *The sequence of Markov approximations has the property that* $f_i^j \overset{n\to\infty}{\to} f^j$ *uniformly for each nonnegative* j.

Proof: When f is continuous, the uniform convergence of f_n to f implies immediately that f_n^j converges uniformly to f^j. When f is not continuous, the proof is a bit more involved.

As in Chapter 4, we will need the set of partition points $Q_0 = \{a_j\}_{j=1}^N$ and $Q_i = Q_{i-1} \cup Q_i = \{a_j^{(i)}\}$, and the associated partitions $I_j^{(i)} = [a_{j-1}^{(i)}, a_j^{(i)}]$. It can be verified that by construction the maps f and f_n satisfy

$$f(I_j^{(n)}) \subset f_n(I_j^{(n)}) \subset I_{j_1}^{(n-1)} \tag{A.1}$$

for some j_1, which depends on j. Furthermore, it can be shown (see e.g. [25]) that the partition elements satisfy

$$\max_j |I_j^{(n)}| < d^n \tag{A.2}$$

for some $d < 1$. From (A.1) and (A.2) it follows that

$$|f^k(x) - f_n^k(x)| < Cd^{n-k} \tag{A.3}$$

for all $x \in [0,1]$, for $n > k$, and for some constant C. Uniform convergence of f_n^k to f^k follows from (A.3). ■

Appendix B

Characterizations Theorems for Filtered Eventually Expanding Maps

The deconvolution results of Chapter 6 depend on two theorems that characterize the effects of filtering on signals generated by eventually expanding maps. These theorems and their proofs are presented in this chapter.

B.1 Filters that Preserve One-State Determinism

Theorem 13 *Suppose that an eventually-expanding map f is the input map to a nonzero, invertible Mth order FIR filter $b[n]$ and that ϕ_b is the associated input-to-output map. Then the filter output satisfies $y[n] = g(y[n-1])$ with g eventually-expanding if and only if $\phi_b(x)$ has the form*

$$\phi_b = \phi_1 \circ f^k$$

for some invertible ϕ_1 and some integer $k \geq 0$.

Proof: $\Rightarrow$ Suppose $\phi_b = \phi_1 \circ f^k$ where ϕ_1^{-1} exists. Then $y[n] = \phi_1(f(\phi_1^{-1}y[n-1])) = g(y[n-1])$.

$\Leftarrow$ We present a proof by contradiction. Suppose $y_n = g(y_{n-1})$ but ϕ_b is not of the form $\phi_1 \circ f^k$, with ϕ_1 invertible. We will show that in this case, x_n could not have been generated by an eventually-expanding map.

We consider first the case where $\phi_b = \phi_1$. Since ϕ_1 is not invertible, we can find two sequences $x[n]$ and $x'[n]$ such that for some L

$$x[L] \neq x'[L]$$

but

$$\phi_1(x[L]) = \phi_1(x'[L]). \tag{B.1}$$

The resulting filtered sequences are of the form

$$y[n] = b[n] * x[n] \tag{B.2}$$

$$y'[n] = b[n] * x'[n]. \tag{B.3}$$

By (B.1), $y[L] = y'[L]$. By our assumption that the filter output is governed by a one-dimensional map, the filtered sequences satisfy

$$y[n] = y'[n], \qquad n \geq L. \tag{B.4}$$

Also, the two input sequences satisfy $f(x[n]) \neq f(x'[n])$, for if not

$$\phi_1(x_L) - \phi_1(x'_L) = b[M](x[L] - x'[L]) \neq 0$$

in contradiction to (B.1. Repeatedly applying the above reasoning, we find that $f^k(x_L) \neq f^k(x'_L)$ for any $k \geq 0$. Thus the sequences $x[n]$ and $x'[n]$ differ for all $n \geq 0$.

Let the stable inverse of $b[n]$ be denoted $b^{-1}[n]$, and let $e[n] = x[n] - x[n]'$. It follow from (B.2) and (B.3) that

$$e[n] = b^{-1}[n] * (y[n] - y'[n]).$$

Since $b[n]$ is FIR, the magnitude of $b^{-1}[n]$ must decay exponentially. Since both $y[n]$ and $y'[n]$ are bounded and

$$y[n] - y'[n] = 0, \qquad n \geq L.$$

it follows that the magnitude of e_n can bounded by

$$|e_n| \leq MC\gamma^n$$

for some constant M and some $\gamma < 1$.

Thus, the difference between the two input sequences must decay at an exponential rate. On the other hand, we know that because $f(\cdot)$ is eventually-expanding, two points that approach each other must diverge. Our hypothesis is contradicted and the theorem is proved.

A similar argument applied to a shifted input sequence established the theorem for ϕ_b of the form $\phi_b = \phi^1 \circ f^k$ ■

B.1.1 Altering the Marginal PDF

Theorem 14 *Consider a continuous eventually-expanding input map and suppose that a nonzero FIR filter $b[n]$ preserves one-state determinism. Then $b[n]$ preserves the marginal pdf of the input if and only if $b[n] = \delta[n - n_0]$ for some integer n_0.*

Proof : $\Leftarrow$ Suppose that $b[n] = \delta[n - n_0]$. Then clearly the output map equals the input map and the two invariant densities coincide.

$\Rightarrow$ Suppose that the marginal pdf of the output equals that of the input.

When f is continuous, each of its iterates f^k is continuous as well. Since the input-to-output transformation is a finite linear combination of iterates of f, ϕ_b is continuous as well.

Since, by assumption, the output has one step determinism, Theorem 7 asserts that the input-to-output transformation is of the form $\phi_b = \phi_1 \circ f^k$ for some nonnegative integer k where ϕ_1 is invertible. Suppose first that $\phi_b = \phi_1$. Then according to Equation (6.11), the

common pdf of input and output satisfies

$$p_x(x) = p_x(\phi_b^{-1}(x))|\frac{d}{dx}\phi_b^{-1}(x)|. \tag{B.5}$$

Because the map ϕ_b^{-1} is continuous and invertible its slope has constant sign. Suppose, without loss of generality, that its slope is positive. Integrating both sides (B.5) from $-\infty$ to x we obtain

$$F(x) = F(\phi^{-1}(x)) \tag{B.6}$$

for each x, where F is the distribution function associated with p_x. The distribution function F is continuous, monotonically nondecreasing and takes valued between 0 and 1. Applying F^{-1} to both sides of Equation (B.6) yields $\phi_1^{-1}(x) = x$. This of course implies that $\phi_1(x) = x$ for all x in the support of the invariant density. By the definition of ϕ_1 we have that $b[n] = \delta[n]$.

A similar argument applies to the more general form of ϕ_b. ■

Appendix C

Results on Filtered Diffeomorphisms

Many of the results of Chapter 7 concerning filtered chaos depended on properties of the state space transformation induced by filtering. In this appendix, we show that the state space tranformation is continuous and prove a theorem that characterizes weak filters.

C.1 Continuity of the State Space Transformation

The next lemma shows that the state space transformation $\mathbf{P}$ (7.10)is continuous from the attractor Λ to composite state space $\mathbb{R}^{N+M}$.

Lemma 3 *The state space transformation P is continuous from Λ to $\mathbb{R}^{N+M}$.*

Proof: Since $\|\mathbf{P}(\mathbf{x}) - \mathbf{P}(\mathbf{y})\|_2^{N+M} = \|\mathbf{x} - \mathbf{y}\|_2^N + \|\mathbf{T}(\mathbf{x}) - \mathbf{T}(\mathbf{y})\|_2^M$, we need only establish the continuity of $\mathbf{T}(\mathbf{x})$. Recall that

$$\mathbf{T}(\mathbf{x}) = \sum_{k=0}^{\infty} A^k \mathbf{b} g(\mathbf{F}^{-k}(\mathbf{x}[n-1])).$$

To show continuity, we note that $g(\mathbf{x})$ is continuous and for any finite value of k the function $\mathbf{F}^{-k}$ is continuous by the hypothesis that $\mathbf{F}$ is a diffeomorphism . Rewriting T in the form

$$\mathbf{T}(\mathbf{x}) = \sum_{k=0}^{L} A^k \mathbf{b} g(\mathbf{F}^{-k}(\mathbf{x})) + \sum_{k=L+1}^{\infty} A^k \mathbf{b} g(\mathbf{F}^{-k}(\mathbf{x})), \tag{C.1}$$

and noting that the first term on the right hand side of (C.1) is a finite combination of continuous functions, we see that the continuity of $\mathbf{T}(\mathbf{x})$ follows from the continuity of the second term of (C.1).

We will show that for each $\epsilon > 0$ there is a $\delta > 0$ such that $\|\mathbf{T}(\mathbf{x}) - \mathbf{T}(\mathbf{y})\| < \epsilon$ whenever $\|\mathbf{x} - \mathbf{y}\| < \delta$ for all $\mathbf{x}, \mathbf{y} \in \Lambda$.

$$\begin{aligned} \|\mathbf{T}(\mathbf{x}) - \mathbf{T}(\mathbf{y})\| &= \|\sum_{k=0}^{L} A^k \mathbf{b}(g(\mathbf{F}^{-k}(\mathbf{x})) - g(\mathbf{F}^{-k}(\mathbf{y}))) + \sum_{k=L+1}^{\infty} A^k \mathbf{b}(g(\mathbf{F}^{-k}(\mathbf{x})) - g(\mathbf{F}^{-k}(\mathbf{y})))\| \\ &\leq \|\|\sum_{k=0}^{N} A^k \mathbf{b}(g(\mathbf{F}^{-k}(\mathbf{x})) - g(\mathbf{F}^{-k}(\mathbf{y})))\| + \|\sum_{k=N+1}^{\infty} A^k \mathbf{b}(g(\mathbf{F}^{-k}(\mathbf{x})) - g(\mathbf{F}^{-k}(\mathbf{y})))\|. \end{aligned}$$

Since the system is stable and the function g is bounded, the integer L can always be chosen to be large that the second term of the right side of the inequality is less than $\epsilon/2$. Next choose δ such that the first term of the inequality is less that $\epsilon/2$ whenever $\|\mathbf{x} - \mathbf{y}\| \leq \delta$. This is always possible since the first term is continuous. ■

C.2 Proof of weak filter lemma

Chapter 7 presented a characterization of weak fitlers in terms of the parameters of the chaotic diffeomorphism and the filter. We now present a proof of this condition.

Theorem 15 *Suppose that* $\mathbf{F}$ *is a dissipative diffeomorphism with Lyapunov exponents* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ *and* $\mathbf{A}$ *has eigenvalues* $\gamma_1, \ldots, \gamma_M$ *ordered such that* $|\gamma_1| \geq |\gamma_2| \geq \cdots |\gamma_M|$. *Then if* $\log(|\gamma_1|) < \lambda_N$ *the filter is weak.*

Proof: We show that the state space transformation is bi-Lipschitz whenever the condition of the theorem is met.

Recall that a transformation of a compact set with bounded Jacobian is Lipschitz. Using to the expression for the state space transformation $\mathbf{P}$ of (7.10) and the chain rule [54], we obtain that the Jacobian has the form

$$D\mathbf{P}(\mathbf{x}) = \begin{bmatrix} \mathbf{I} \\ \sum_{k=0}^{\infty} A^k b Dg(\mathbf{F}^{-k}(\mathbf{x}))(D\mathbf{F}^{-k})(x) \end{bmatrix}, \tag{C.2}$$

where $\mathbf{I}$ is the $N \times N$ identity matrix. It follows that the Jacobian is finite whenever the sum $\sum_{k=0}^{\infty} A^k b Dg(\mathbf{F}^{-k}(\mathbf{x}))(D\mathbf{F}^{-k})(x)$ converges. The convergence of the sum depends on both the linear dynamics through the matrix A and the nonlinear dynamics through the matrix $(D\mathbf{F}^{-k})(x)$.

For large k, the Jacobian of the nonlinear dynamics $(D\mathbf{F}^{-k})(x)$ may have norm which increases exponentially with k. This is a consequence of the sensitive dependence on initial conditions displayed by the dynamical system $\mathbf{F}$, or equivalently, the fact that $\mathbf{F}$ has positive Lyapunov exponents (see Chapter 2). On the other hand, for large k the norm of the matrix A^k decreases exponentially when the filter is stable is a stable system.

A condition that guarantees the boundedness of the Jacobian of Equation (C.2) can be determined as follows. Suppose we simply examine the norm of the Jacobian matrix of $\mathbf{T}(x)$. Using the elementary properties of matrix norms, we arrive at the following series of inequalities:

$$\begin{aligned} \|\mathbf{T}(x)\| &= \|\sum_{k=0}^{\infty} A^k (Dg)(\mathbf{F}^{-k}(\mathbf{x}))(D\mathbf{F}^{-k})(x)\| \\ &\leq \sum_{k=0}^{\infty} \|A^k b Dg(\mathbf{F}^{-k}(\mathbf{x}))(D\mathbf{F}^{-k})(x)\| \\ &\leq \sum_{k=0}^{\infty} \|A^k\|\|b\|\|Dg(\mathbf{F}^{-k}(\mathbf{x}))\|\|(D\mathbf{F}^{-k})(x)\| \end{aligned} \tag{C.3}$$

The convergence of the above expression can be determined by noting that the matrix norm $\|(D\mathbf{F}^{-k})(x)\|$ is exactly the expression that occurs in the definition of the Lyapunov exponents for the system $\mathbf{F}^{-1}$.

According to the results of [49], the Lyapunov exponents of the inverse system are simply related to those of the forward system. Specifically, Lyapunov exponents of $\mathbf{F}^{-1}$ are the

simply those of $\mathbf{F}$ but with opposite signs, i.e. if the Lyapunov exponents of $\mathbf{F}$ are ordered such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ then $\mathbf{F}^{-1}$ has exponents $-\lambda_N \geq -\lambda_{N-1} \geq \cdots \geq -\lambda_1$. Intuitively, the change in sign reflects the fact that trajectories which approach one another under the forward system diverge from one another under the inverse system. The opposite is true of course for trajectories which diverge under the forward system. Since we are interested in dissipative systems, the value of λ_N must be negative.

The norm of the Jacobian matrix $\|(D\mathbf{F}^{-k})(x)\|$ for large k grows at a rate determined by the largest Lyapunov of the inverse system $\mathbf{F}^{-1}$, and satifies the bound $\|(D\mathbf{F}^{-k})(x)\| \leq Ce^{(|\lambda_N|+\epsilon)k}$ for some small $\epsilon > 0$ and an appropriate constant C. Furthermore, the norm of the matrix A^k for large k decays at a rate determined by its eigenvalue of largest magnitude. More specifically, if the eigenvalues of A are ordered as $|\gamma_1| \geq |\gamma_2| \geq \cdots |\gamma_M|$ then for large k we have the inequality $\|A^k\| \leq (|\gamma_1| + \delta)^k$ where $\delta > 0$ is arbitrarily small.

The condition for the convergence of the upper bound (C.3)can be determined as follows

$$\begin{aligned}\sum_{k=0}^{\infty} \|A^k\|\|b\|\|Dg(\mathbf{F}^{-k}(\mathbf{x}))\|\|(D\mathbf{F}^{-k})(x)\| \;\leq\; & \sum_{k=0}^{N} \|A^k\|\|b\|\|Dg(\mathbf{F}^{-k}(\mathbf{x}))\|\|(D\mathbf{F}^{-k})(x)\| \\ & + K \sum_{k=N+1}^{\infty} (|\gamma_1| + \delta)^k \|e^{(|\lambda_N|+\epsilon)k},\end{aligned}$$

where K is a bound on the value of $\|b\|\|Dg(\mathbf{F}^{-k}(\mathbf{x}))\|$. The first term on the right is finite. The second term is also finite under certain conditions on the filter. Specifically, since δ and ϵ can be made arbitrarily small by appropriate choice of k, we conclude that the Jacobian is bounded when $\gamma_1 e^{|\lambda_N|} < 1$, or alternatively, when $\log(|\gamma_1|) < \lambda_N$.

When this condition is met, the transformation $\mathbf{P}$ is bi-Lipschitz and the filter is weak.

■

Bibliography

[1] F. Takens, "Detecting strange attractors in turbulence," in *Lecture Notes in Mathematics 898* (D. Rand and L. S. Young, eds.), p. 366, New York, NY: Springer-Verlag, 1981.

[2] J. Guckenheimer, G. Oster, and A. Ipaktchi, "The dynamics of density dependent population models," *J. Math. Bio.*, vol. 4, no. 2, pp. 101–147, 1977.

[3] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields.* New York: Springer–Verlag, 1983.

[4] R. Devaney, *An Introduction to Chaotic Dynamical Systems.* Reading, Massachusetts: Addison–Wesley, 1989.

[5] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications.* New York, NY: John Wiley and Sons, 1990.

[6] R. Mañe, "On the dimension of the compact invariant sets of certain nonlinear maps," in *Lecture Notes in Mathematics 898* (D. Rand and L. S. Young, eds.), New York, NY: Springer-Verlag, 1981.

[7] I. P. Natanson, *Theory of Functions of a Real Variable.* New York: Frederick Ungar Publishing, 1961.

[8] A. V. Holden, *Chaos.* Princeton University Press, 1986.

[9] J. Baillieul, R. W. Brockett, and R. B. Washburn, "Chaotic motion in nonlinear feedback systems," *IEEE Trans. Circuits and Systems*, vol. 27, pp. 990–997, 1980.

[10] J. H. B. Deane and D. J. Jeffries, "Periodic, intermittent and chaotic states of clocked autonomous digital systems," *Int. J. Cir. Th. App.*, vol. 18, pp. 175–187, 1990.

[11] S. Smale, *The Mathematics of Time.* New York: Springer-Verlag, 1980.

[12] D. Aeyels, "Generic observability of differentiable systems," *SIAM J. Cont. Optim.*, vol. 19, no. 5, pp. 595–603, 1981.

[13] D. Aeyels, "Global observability of morse–smale vectorfields," *J. Diff. Eq.*, vol. 45, pp. 1–15, 1982.

[14] R. Mane, *Ergodic Theory and Differentiable Dynamics.* Berlin: Springer–Verlag, 1987.

[15] A. Lasota and M. Mackey, *Probabilistic Properties of Deterministic Systems.* Cambridge: Cambridge University Press, 1985.

[16] K. Peterson, *Ergodic Theory.* Cambridge: Cambridge University Press, 1983.

[17] V. I. Oseledec, "A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems," *Trudy Mosk. Mat. Obsc. (Moscow Math. Soc.)*, vol. 19, p. 179, 1968.

[18] H. Sakai and H. Tokumaru, "Autocorrelations of a certain chaos," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 28, no. 5, pp. 588–590, 1990.

[19] A. Boyarsky and M. Scarowsky, "On a class of transformations which have unique absolutely continuous invariant measures," *Trans. Am. Math. Soc.*, vol. 255, pp. 243–262, 1979.

[20] A. A. Kosyakin and E. A. Sandler, "Ergodic properties of a class of piecewise-smooth transformations of a segment," *Izv. VUZ Mat.*, vol. 118, pp. 32–40, 1972. English translation available from the British Library Service.

[21] N. Friedman and A. Boyarsky, "Matrices and eigenfunctions induced by markov maps," *Lin. Alg. App.*, vol. 39, pp. 141–147, 1981.

[22] H. Minc, *Nonnegative Matrices.* New York, NY: John Wiley and Sons, 1988.

[23] J. H. B. Deane and D. C. Hamill, "Chaotic behavior in a current-mode controlled dc-dc converter," *Electronics Letters*, vol. 27, pp. 1172–1173, 1991.

[24] Y. S. Tang, A. I. Mees, and L. O. Chua, "Synchronisation and chaos," *IEEE Trans. Circuits and Systems*, vol. CAS-30, no. 9, pp. 620–626, 1983.

[25] F. Hofbauer and G. Keller, "Ergodic properties of invariant measures for piecewise monotonic transformations," *Math. Z.*, vol. 180, pp. 119–140, 1982.

[26] S. Wong, "Some metric properties of mappings of the interval," *Trans. Am. Math. Soc.*, vol. 246, pp. 493–500, 1978.

[27] P. Gora and A. Boyarsky, "Compactness of invariant densities for families of expanding, piecewise monotonic transformations," *Can. J. Math*, vol. 41, no. 5, pp. 855–869, 1989.

[28] A. Boyarsky, "A matrix method for estimating the lyapunov exponent of one dimensional systems," *J. Stat. Phys.*, vol. 50, no. 1/2, pp. 213–229, 1988.

[29] E. Kreyszig, *Introductory Functional Analysis and Applications*. New York: John Wiley and Sons, 1989.

[30] N. Dunford and J. T. Schwartz, *Linear Operators Part I:*. New York: John Wiley and Sons, 1957.

[31] T. Li, "Finite approximation for the frobenius–perron operator: A solution to ulam's conjecture," *J. Approx. Th.*, vol. 17, pp. 177–186, 1976.

[32] C. K. Tse, "Flip bifercation and chaos in three-state boost switching regulators," *IEEE Trans. Circuits and Systems*, vol. 41, no. 1, pp. 16–23, 1994.

[33] J. G. Kassakian, M. F. Schlecht, and G. C. Verghese, *Principles of Power Electronics*. Reading, Mass.: Addison–Wesley, 1991.

[34] H. C. Papadopoulos and G. W. Wornell, "Maximum likelihood estimation of a class of chaotic signals," *IEEE Trans. Inform. Theory*, p. to appear, 1994.

[35] F. Mitschke, M. Moller, and W. Lange, "Measuring filtered chaotic signals," *Phys. Rev. A*, vol. 37, no. 11, pp. 4518–4521, 1988.

[36] R. Badii, G. Broggi, B. Derighetti, M. Ravani, S. Cilberto, A. Politi, and M. A. Rubio, "Dimension increase in filtered chaotic signals," *Phys. Rev. Lett.*, vol. 60, no. 11, pp. 979–982, 1988.

[37] J. Scargle, "Studies in astonomical time series analysis. iv. modeling chaotic and random processes with linear filters," *The Astrophysical Journal*, vol. 359, no. 2, pp. 469–482, 1990.

[38] H. L. V. Trees, *Detection, Estimation and Modulation Theory.* New York: John Wiley and Sons, 1968.

[39] E. Wong, *Introduction to Random Processes.* New York: Springer–Verlag, 1983.

[40] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* New York, NY: McGraw-Hill, 1984.

[41] J. Theiler and S. Eubanks, "Don't bleach chaotic time series," Tech. Rep. No. LA-UR-92-1575, Los Alamos National Laboratory, june 1992.

[42] A. Chennaoui, J. Liebler, and H. G. Schuster, "The mechanism of the increase of the generalized dimension of a filtered chaotic signal," *J. Stat. Phys.*, vol. 59, no. 5/6, pp. 1311–1328, 1990.

[43] D. Donoho, "On minimum entropy deconvolution," in *Applied Time Series Analysis II* (D. Findley, ed.), Academic Press, 1981.

[44] A. W. Drake, *Fundamentals of Applied Probability Theory.* New York: McGraw–Hill, 1967.

[45] R. Bowen, "Bernoulli maps of the interval," *Israel J. Math.*, vol. 28, no. 1–2, pp. 161–168, 1977.

[46] W. A. Brock, "Distinguishing random and deterministic systems: Abridged version," *Journal of Economic Theory*, vol. 40, pp. 168–195, 1986.

[47] W. L. Brogan, *Modern Control Theory.* Englewood Cliffs, New Jersey: Prentice–Hall, 1985.

[48] T. Yamamoto, "On the extreme values of the roots of matrices," *J. Math. Soc. Japan*, vol. 19, no. 2, pp. 173–178, 1967.

[49] D. Ruelle, "Ergodic theory of differentiable dynamical systems," *Publ. Math. IHES*, vol. 50, pp. 285–306, 1979.

[50] A. Chennaoui, K. Pawelzik, W. Liebert, H. G. Schuster, and G. Pfister, "Attractor reconstruction from filtered chaotic time series," *Phys. Rev. A*, vol. 41, no. 8, pp. 4151–4159, 1990.

[51] F. Mitschke, "Acausal filters for chaotic signals," *Phys. Rev. A*, vol. 41, no. 2, pp. 1169–1171, 1990.

[52] J. P. Eckmann and D. Ruelle, "Ergodic theory of chaos and strange attractors," *Rev. Mod. Phys.*, vol. 57, no. 3, pp. 617–656, 1988.

[53] W. Rudin, *Principles of Mathematical Analysis*. New York: McGraw-Hill, 1976.

[54] K. Hoffman, *Analysis in Euclidean Space*. Englewood Cliffs, NJ: Prentice Hall, 1975.