

Lacunarity and Chaos in Nature

Leonard Allen Smith

Thesis

**Submitted in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy in the
Graduate School of Arts and Sciences**

Columbia University

1987

© 1987

Leonard Allen Smith

ALL RIGHTS RESERVED

ABSTRACT

Lacunarity and Chaos in Nature

Leonard Allen Smith

The general solution of self-similar scaling equations is seen to contain an oscillatory component. The lacunarity of a self-similar set is reflected in this oscillatory component of the set's scaling statistics. Strange attractors found in nonlinear dynamical systems often possess self-similar structure. After introducing some new results regarding a strange attractor of astrophysical interest and considering the effects of using digital computers to study chaos, we demonstrate that the standard method of characterizing the global structure of these sets is vulnerable to errors arising from a misinterpretation of the oscillatory component. A new method of describing strange attractors by a spectrum of local scaling exponents computed about unstable periodic orbits is proposed. In addition to being computationally tractable, this approach utilizes the lacunarity oscillation to compute an accurate estimate of the residual error.

The lacunarity oscillation exists both for self-similar physical distributions and strange attractors. The role of this oscillation in the structure functions of fully developed fluid turbulence is exhibited through the β model. The physical structure of the observable matter distribution in the universe is also considered.

A new example of Hamiltonian chaos which is of geophysical interest is introduced. Specifically, a model is developed to simulate the motion of slowly settling tracers in a time-dependent flow. Here the time development of the length of a material line of

tracer particles is shown to be quite complicated even when the background flow field is laminar. A class of two-dimensional maps that simulate this type of flow is developed. Finally, the concept of a strange accumulator is introduced to describe the strong inhomogeneities observed to form in the probability distribution of a single chaotic trajectory.

Acknowledgments

First it is a pleasure to thank Ed Spiegel for his guidance throughout this project and for infecting me with an idea of what it means to be serious. Although the New York Convective Collective has dissipated, the ideas still being thrown about in the attic of Pupin have been a major source of inspiration for me. I am particularly indebted to P. Cvitanovic, J.-D. Fournier, I. Proccacia, K. Prendergast, C. Sparrow, M. Tabor and A. Wolf.

I am happy to acknowledge the financial support provided by Jim Hansen and the Goddard Institute for Space Sciences via NASA Cooperative agreement #NCC 5-29. In addition to financial aid, the Institute provided generous computer resources without which the studies reported here would not have been possible. I would like to thank Nadim Habra and the systems group at GISS for their patience and understanding that chaos and virtual machines sometimes interact in a violent manner. I owe a great debt to Rene Galarza for letting one more plot through, several times.

For allowing and encouraging me to roam this far from home I thank my parents; for various absurd notions and cards I thank my sister, Wendy Faheysmith. I am grateful to Bryan and Terrie Katz for a chance to really play with Squiggles. For their ever present friendship I thank Pat and Tip LaPorte, with a special thanks to Tip for reminding me that things are not always what they seem.

For the personal support I have received in New York I owe thanks to far too many people to name here. For sanity and focus I wish to

thank Kazumi Tabata and Stuart Chassen. My first years in New York were made livable through the craziness and caring of Dan Johnson, Walter Robinson and Audrey Wolf.

My successful completion of this project was possible only with the love, friendship and support of Bonnie Scarborough, Dave DeBeer and Richard Seager. This debt I hope to partially repay over the next two years, or so.

Preface and Overview

The study of nonlinear dynamical systems has provided a deterministic framework for the analysis of physical systems which behave in an irregular, seemingly stochastic manner. A major goal of this research is to discover, from an observational record, a system of equations governing the process, or failing this, a model of sufficient sophistication to be of use in understanding the character of the system and its behavior. The initial steps in this direction require a method of characterizing the complexity of a system's dynamics in its phase space. The recent development of a procedure for reconstructing this trajectory from an observed signal has inspired a great deal of interest interpreting the nature of these orbits. For ergodic systems, a trajectory may explore all energetically accessible regions of phase space, while in many interesting physical systems the trajectory appears to be restricted to a small subset of this region. The complexity of a system's dynamics are reflected in the properties of this trajectory which is taken to approximate the set of all accessible states. The focus of this thesis is on methods of characterizing the resulting set. This is done, for the most part, through the calculation of scaling exponents (generalized dimensions). These exponents provide estimates of the number of excited degrees of freedom in the system and thus a lower bound on the complexity of a model.

This thesis is divided into five chapters. Chapter 1 is titled "Concepts of Dimension and their relevance to Mechanics" and introduces the concepts of self similarity through a new analysis of

an old strange attractor (Moore and Spiegel, 1966). The notion of the scaling exponents of a distribution is discussed and the existence of an oscillatory component of the scaling functions of self-similar sets is established. It is the existence of this oscillatory function which provides, for the first time, a method of reliably evaluating the accuracy of numerically determined scaling exponents. In Chapter 2 the question of determining scaling exponents from data is considered. Two common methods are reviewed and their shortcomings are demonstrated. Specifically, we consider the methods of Broomhead and King (1986) and Grassberger and Procaccia (1983). It is shown by example that the former method fails to converge in test cases where the full structure of the system generating the data is known. Using the second method, the oscillatory component of the correlation integral in some simple fractal sets is examined. When the magnitude of this component is large, the method may converge (locally) yet provide a poor estimate of the correlation exponent. A fundamental issue for all methods of this type (which consider the large scale structure of a set) involves the quantity of data required for a reliable result. In the fourth section of Chapter 2 an absolute lower bound on the amount of optimally distributed data required as a function of the dimension of the set is established. Further restrictions implied by the use of a single time series are also discussed.

The local scaling of fractal sets is investigated in Chapter 3. Using the periodicity of the oscillatory component of scaling functions, a new method of characterizing this scaling is developed

which provides an estimate of the accuracy of the results obtained.

The oscillation in the scaling of physical systems is considered in Chapter 4. In fluid turbulence, where recent experiments suggest the presence of an oscillatory term, a self similar model of the turbulent cascade is generalized to demonstrate the observable manifestations of this phenomena. The presence of an oscillatory component is inherent to the geometry of self-similar distributions, whether the distribution arises from the dynamics of a system in its phase space or a physical distribution in real space. One of the oldest systems to be considered self-similar is the distribution of matter in the Universe; this distribution is considered in the second part of Chapter 4. No oscillatory component is evident in the data set examined.

In the final chapter, Chapter 5, a model for the motion of a particle in a two dimensional time dependent flow is developed. The chaos described there is of a different sort than that observed in dissipative systems. The development of self-similar structures in the flow indicates that flow visualization techniques should be interpreted cautiously when non-ideal tracers are used. Considering the model as a dynamical system, it is shown that the self-similar structure of the probability distribution of (finite length) trajectories can also bias measurement of the scaling exponents.

Overall, the research reported here is a study of an oscillatory component in what has often been considered a line. This phenomenon holds deep implications for existing methods of measurement of

scaling sets, and also suggests new approaches which circumvent the shortcomings of existing methods.

TABLE OF CONTENTS

	page
Acknowledgments	i
Preface and Overview	iii
Table of Contents	vii
Chapter 1: Concepts of dimension and Their Relevance to Dynamics	1
Section 1: Two Concepts of Dimension	5
Section 2: Deterministic Dynamical Systems	13
Section 3: Probability Distributions on Fractals	25
Section 4: The Oscillatory Component of Scaling	32
Figures	35
Appendix 1: Apparent Chaos in Digital Computers	46
Chapter 2: Analysis of Dimension in Nature	60
Section 1: The Broomhead-King Procedure	65
Section 2: The Grassberger-Procaccia Algorithm	70
Section 3: Evaluating the Correlation Integral of Simple Fractal Sets	73
Section 4: Scaling Limits in Nonlacunar Sets	77
Section 5: Edge effects as Strict Limits to Dimensional Calculations	84
Appendix 2: Explicit construction of the oscillation in Cantor Sets.....	89
Figures and Tables	92
Chapter 3: The Oscillatory Component of Scaling	115
Section 1: The Local Correlation Integral	117
Section 2: Intrinsic Oscillations in the Statistics on Fractal Sets	121
Section 3: An Oscillatory Component of Scaling in Inhomogeneous Fractals	125
Section 4: Observing the Oscillatory Component in the Scaling of a Strange Attractor	129
Section 5: Conclusions	137
Figures	138

Chapter 4: Lacunarity in Nature:.....	158
Part I: Lacunarity in Fluid Turbulence	160
Section 1: Characteristics of Turbulence	160
Section 2: The β Model	166
Section 3: β Model Oscillations	169
Part II: Cosmic Lacunae	173
Section 4: The Fournier Universe	174
Section 5: A Geometric Construction	178
Section 6: Projection Effects	180
Section 7: Observations	181
Figures	183
Chapter 5: Particulate Dispersal in a Time-Dependent Flow	189
Section 1: Steady Background Flow	191
Section 2: Phase Wrapping	197
Section 3: Experimental Observations	199
Section 4: Time Dependent Flows	200
Section 5: Conclusions	209
Figures	211
Appendix 3: A Class of Iterative Maps	233
Appendix 4: Strange Accumulators	242
References	251

CHAPTER I Concepts of Dimension and Their Relevance to Dynamics

Introduction

The laws which govern many physical processes may be expressed through systems of nonlinear equations. In first attempts to solve such systems, the nonlinearity is often assumed to be small and is suppressed or ignored entirely. Study of the resulting linear equations has established many time-honored methods of analysis, the failure of which was once taken to imply pathological behavior such as random noise disturbing the system. This is no longer the case. Beginning with Lorenz's numerical experiments in 1963, it has been recognised that signals which appear noisy and intermittent may be generated by completely deterministic nonlinear processes. Instead of requiring an external stochastic source term, the nonlinearities of these systems rapidly (exponentially) amplify small differences in the initial conditions. Deterministic systems with sustained dynamics that show sensitive dependence on initial condition¹ are called chaotic; their trajectories in phase space are said to display deterministic chaos in contrast to truly random or stochastic processes. Identifying a complicated physical system as chaotic provides insight into the processes that govern the system. Attempts to quantify these ideas have lead to a variety of scaling exponents (dimension(s), generalized dimensions, Lyapunov exponents, and so on), which are introduced in §§2 and 4 of this chapter. First we

¹ This phrase, coined by Ruelle (1980), describes the main feature of generic chaos.

give an idea of the range of physical phenomena to which these ideas are relevant.

The lack of linearity normally implies the loss of superposition; nonlinearity couples modes which would act independently in the corresponding linear system and it promotes interaction between disturbances of all scales. As noted by Lorenz (1986), atmospheric motions (weather) and climate models provide examples of this coupling where small-scale effects play a crucial role in large-scale dynamics. Often a self-organization is imposed via this coupling, and the properties which characterize the system are invariant under certain magnifications (changes in scale). This scaling behavior is observed in physical distributions. Typical examples include diffusion limited aggregates (Witten and Sander, 1981, 1983) and the distribution of visible matter in the Universe (de Vaucouleurs, 1971, Mandelbrot, 1972), coastlines (Richardson, 1961) and the fracture surfaces of metals (Mandelbrot et al., 1984). Mathematical structures which display scaling behavior include strange attractors (e.g. Hénon, 1976; Ruelle, 1980; Spiegel, 1986), strange accumulators (e.g. Smith and Spiegel, 1987), the boundaries of the basins of attraction (Grebogi et al., 1984) and fat fractals (Umberger and Farmer, 1985, Grebogi et al., 1985). As a result of the reconstruction theorem due to Takens (1981), the phase space structure of a host of physical systems has been elucidated in terms of "simple" nonlinear systems. These systems include particular behaviors of Josephson junctions (Iansiti, et al., 1985), lasers (Albano, et al., 1985) and Rayleigh-Bernard convection in a variety of fluids: mercury (Libchaber et al.,

1982), liquid Helium (Libchaber et al., 1982) and water (Giglio et al., 1984). Even some of the behavior of a "leaky" faucet is found to be in quantitative agreement with a simple mapping (Martien et al., 1985).

Since the particular subject of inquiry may be drawn from diverse categories, the entity (set, surface, boundary, trajectory, ...) to be investigated will be referred to simply as an object or set. This is natural since, in all computational realizations, the entity will be represented as a denumerable collection of distinct points specified with finite accuracy. Similar techniques may be used to analyze objects from each class, although the class of the object often provides information useful in analyzing its structure. For example, a static physical boundary may be embedded in three dimensional space. Similarly, the solution of a set of ordinary differential equations must lie in a space of dimension less than or equal to the degree of the system.

Why should sets and their scaling parameters be of interest to the physicist? A precise knowledge of the geometric and probabilistic structure of dynamical systems provides detailed limits on the predictability and dynamics of the system. Even a rough knowledge of the simplest geometrical quantities can provide insight into the nature of a system. Consider the power spectrum shown in Figure 1.1 (page 35). Traditionally, such "broadband noise" is interpreted as an indication that the system possesses many active degrees of freedom; if so, a complete description of the system would necessarily be complex. The power spectrum alone cannot

differentiate non-deterministic noise from highly structured chaos, but scaling exponents distinguish between the two and quantify the latter. In many cases, the state of the system may be completely described by only a few independent coordinates (in the case of Figure 1.1, three; see §1.3). Equivalently, the system's dynamics are restricted to a subspace of all possible states and this subspace may be embedded in a Euclidian space of small dimension. This means that after transients, the dynamics may be described by a model with only a few degrees of freedom (relative to the dimension of the full phase space of the system). A standard example from chemistry is the Belousov-Zhabotinsky experiment which involves chemical reactions between 30 chemical species under a variety of interaction parameters (Zaiken and Zhabotinsky, 1970). Parameter regimes exist where, for a variety of initial conditions, this system exhibits deterministic chaos; the asymptotic motion in phase space lies in a low dimensional subspace and may be completely described by many fewer variables than the number required to specify an arbitrary initial condition in the original 30+ dimensional phase space. Two examples from fluid dynamics, where the phase space is infinite dimensional, are Rayleigh-Bernard convection and Taylor-Couette flow (Branstater et al., 1983). Again there exists a parameter range where the system trajectory is restricted through self-organization to motion within a subspace of small dimension. Swinney and coworkers have described the transition of Taylor-Couette flow from a (zero dimensional) fixed point behavior up through chaotic behavior. In these cases, a knowledge of the geometry of the system provides a quantitative

measure related to the fundamental complexity of the problem.

The development of a formalism with which to describe nonlinear systems has a colorful, multiply connected history and is still incomplete. A survey might start with Poincaré's work on the three body problem at the turn of the century. For the most part, these ideas were developed by mathematicians during the first half of this century. They were advanced further in the work of probability theorists, meteorologists, astrophysicists, psychologists and computer scientists interested in pattern recognition during the sixties and early seventies, to be re-embraced by physicists, employed by a diverse cross-section of scientists including chemists, ecologists, population biologists and neurologists, and applied to a still increasing variety of systems today.

With no attempt to disentangle this development, the current formalism will be introduced in parallel with several ideas from the theory of dynamical systems which provide physical motivation. The discussion is presented with the bias found in the work of Farmer and Procaccia. It begins with a discussion of the meaning of dimension in §1.1 where two fundamental notions of dimension are developed and related to several of the "dimensions" commonly referenced in the literature. After the relevance of these ideas to physical models is presented in §1.3, the generalized dimensions of a probability distribution are defined.

SECTION 1.1 Two Concepts of Dimension

There are many different conceptions of dimension. For the sets discussed by Euclid, the different definitions based on each of these concepts often coincide in the same, integer value; while for many physical boundaries and most of the objects of nonlinear systems, this is not true. Two ideas form the basis for the majority of definitions of dimension to be dealt with here. The first, formulated by Poincaré, is related to the notion of the number of "distinct directions" on the set; the second, credited to Hausdorff, is derived by examining how the "volume" of an object changes as it is measured with increasing precision.

First, consider what is now called the topological dimension, based in the ideas of Poincaré and later supplied with a firm mathematical foundation by Brouwer, Menger and Urysohn. Our discussion draws heavily from the complete development presented in Hurewicz and Wallman (1941). Poincaré suggested that the intuitive notion of the dimension of a given set or space is formed from consideration of the properties (dimension) of the simplest space capable of isolating a region in the original space. If a point has zero dimension, the dimension of higher continua may be established inductively:

I shall base the determination of the number of dimensions on the notion of cuts. Let us consider first of all a closed curve, that is, a continuum of one dimension. If on this curve we take any two points through which we shall not permit ourselves to pass, the curve will be cut into two parts, and it will become impossible to go from one to the other, still remaining on the curve, but not passing through the excluded

points. Let us consider, on the other hand, a closed surface which forms a continuum of two dimensions. It will be possible to take on this surface one, two, or any number of excluded points whatever. The surface will not be divided into two parts because of this; it will be possible to go from one point to another on this surface without encountering any obstacle because it will always be possible to go around the excluded points...

We now know what a continuum of n dimensions is. A continuum has n dimensions when it is possible to divide it into many regions by means of one or more cuts which are themselves continua of $n-1$ dimensions. The continuum of n dimensions is thus defined by the continuum of $n-1$ dimensions. This is a definition by recurrence.²

In the formulation of Menger, this idea is capsulized in the definition of the topological dimension, D_T . Specifically, the topological dimension of the empty set is -1 ; for any nonempty set, D_T is equal to the least integer n such that every member of the set has arbitrarily small neighborhoods, the boundaries of which have dimension less than n .

It follows that if every point in the set has arbitrarily small neighborhoods with empty boundaries, the set has $D_T = 0$. This is clearly the case for any finite set of points and it is straightforward to show that D_T is equal to zero for any countable set as well as the set of all rational (or irrational) numbers on the unit interval. Furthermore it may be proven (Hurewicz and Wallman, Chapter 4) that the Euclidian N space, E^N , has $D_T = N$.

The utility of the topological dimension is limited by the difficulty of its determination from experimental data (Farmer, 1982). For the majority of systems found in texts on classical

² From Poincaré (1963) [originally 1913] page 29; a similar discussion may be found in Poincaré (1982) [originally 1913] page 256.

mechanics, however, this concept is sufficient and may be directly related to the number of degrees of freedom of the system. This is shown in the following section. Before doing so, we consider a second notion of dimension, the capacity.

The general trend is to establish by analogy a definition which admits real values equal to the topological dimension for Euclidean objects but is sensitive to the scaling behavior of more general sets. The mathematical framework here is found in the arguments of Hausdorff (1919) and Besicovitch (1953; and references therein).

Consider a geometrical object embedded in E^N , such as a rectangular solid in a 3-dimensional space. Put $M(r)$ equal to the minimum number of N -dimensional spheres (N -spheres) of radius r required to cover the entire set. When r is decreased by a factor of two, $M(r)$ increases by a factor of 2^N (e.g. 8 for the rectangular solid). The capacity dimension d_c is then defined as:

$$d_c = \lim_{r \rightarrow 0} \frac{\log M(r)}{\log r} \quad (1.1)$$

Here d_c is not restricted to integer values. There is, of course, no restriction to noninteger values; a set that has no topological resemblance to a line may have d_c equal to one, an example is the Fournier Universe discussed in Chapter 4.

To demonstrate that the capacity and the topological dimension do not always coincide, consider the "middle thirds" Cantor set. This set is perhaps the simplest and is certainly the most common example for which these two notions of dimensions differ; its construction is shown schematically in Figure 1.2. Starting with the interval

$0 \leq x \leq 1$, delete the middle third, $1/3 < x < 2/3$. Now repeat the process on each of the two remaining subintervals. And so on. This simple destruction algorithm may be used to generate members of a special class of Cantor sets (see §2.3). For the middle thirds Cantor set, a point is uniquely labeled by the record of which of the retained subinterval it was located in at each generation. At each iteration of the process the point in question will lie either on the right section (R) or the left section (L) of the subinterval. Recording these symbols in order identifies the point with a symbolic sequence; for example the point marked by the arrow in Figure 1.2 b is represented by the sequence LRLLLLLL... This symbolic representation will prove useful when we consider the correlation integral in Chapter 2. Through the correspondence $R \rightarrow 1$, $L \rightarrow 0$, a member of the set may be identified with each binary fraction between zero and one inclusive, therefore the set contains an uncountable infinity of points.

In order to determine the topological dimension of this set, note that the location of any member of the set can be expressed as

$$x = \sum_{n=1}^{\infty} \frac{a_n}{3^n} \quad a_n = 0 \text{ or } 2.$$

Thus there exist arbitrarily small neighborhoods about each point with empty boundaries and the set has $D_T = 0$. As shown below, it has a nonzero capacity dimension.

Every Cantor set may be decomposed into an integer number, ρ , of identical subsets A_i each of which reproduces the structure of the entire set when magnified by a factor $1/\sigma$. Any set with this

property is called strictly self-similar with similarity ratio σ and structure ratio ρ . If only general properties of a distribution are reproduced, it is called statistically self-similar or simply, self-similar. The middle thirds Cantor set is strictly self-similar with $\rho = 2$ and $\sigma = 1/3$.

The capacity dimension of a strictly self-similar set is nonzero. If m spheres of radius ℓ are required to cover the set, $m\rho^N$ spheres of radius $\ell\sigma^{-N}$ will be required. This is true for any integer N . Thus

$$M(\ell/\sigma) = \rho M(\ell) \quad (1.2)$$

with

$$M(\ell) \sim \ell^{d_c} \quad (1.3)$$

implies

$$d_c = - \frac{\log \rho}{\log \sigma} \quad (1.4)$$

For the Cantor set above with $\rho = 2$ and $\sigma = 1/3$

$$d_c = \frac{\log(2)}{\log(3)} = 0.63093\dots$$

Many of the statistics of self-similar sets obey an equation of the form

$$C(\ell/\sigma) = \rho C(\ell) \quad (1.5)$$

the general solution of which is (Smith et al, 1986)

$$C(\ell) = \chi(\ell) \ell^d \quad (1.6)$$

with

$$d = - \frac{\log \rho}{\log \sigma} \quad (1.7)$$

where $\chi(\ell)$ is a log-periodic function³ of ℓ . The importance of this oscillatory component of scaling is discussed in Chapters 2, 3, and 4. The power law component is often expressed as

$$C(\ell) \sim \ell^d \quad (1.8)$$

where the symbol " $\sim$ " is read "scales as". Any quantity d that satisfies an equation of the form 1.8 is called a scaling exponent. The capacity dimension is the first of several scaling exponents which will be discussed in this thesis.

The fractal dimension, d_f , is a slight variation on the capacity. Instead of the minimal number of hyperspheres to cover the set, consider the coarse grained image of the set produced by a fixed, space-filling square mesh of grid size ϵ . Denote the number of hyper-boxes containing at least one member of the set as $M(\epsilon)$. The fractal dimension is then defined (Mandelbrot, 1972) as

$$d_f = \lim_{\epsilon \rightarrow 0} \frac{\log M(\epsilon)}{\log \epsilon} \quad (1.9)$$

For most physical examples, the fractal dimension and capacity dimension are thought to coincide. Questions regarding the effect of covering a set with boxes as opposed to spheres have been addressed by Cawley et al. (1985 preprint). The formal basis for both of these concepts is the Hausdorff-Besicovitch dimension, d_{HB} ; the definition of which requires the introduction of several additional concepts in order to establish a more rigorous notion of covering a set.

Let X be a set in an N dimensional Euclidean space with metric d . Define $\delta(X)$, the diameter of X , as the maximum separation of any two

³Alternatively, if χ is considered a function of $\log(\ell)$, χ is periodic.

members of the set.

$$\delta(X) = \sup d(x,y) \quad (1.10)$$

where the supremum is taken over all pairs of the elements of X . Decompose the set X into a countable number of subsets A_i each of which has diameter less than ϵ . For any real number p such that $0 \leq p \leq \infty$, define

$$m_p^\epsilon = \inf \sum_{i=1}^{\infty} [\delta(A_i)]^p . \quad (1.11)$$

Here the infimum (minimum) is taken over all possible decompositions. This defines the minimal covering.

The p -dimensional measure of X , denoted $m_p(X)$ is given by

$$m_p(X) = \sup_{\epsilon > 0} m_p^\epsilon(X) . \quad (1.12)$$

The Hausdorff-Besicovitch dimension of X is then defined as the supremum of all real numbers p such that $m_p > 0$.

While special cases may be constructed where the capacity dimension, fractal dimension and Hausdorff-Besicovitch dimension differ, it is not clear whether this will occur in any system of physical interest. Once self-similar behavior is assumed, many techniques are available to quantify it. Since each method yields a somewhat different quantity, a large number of "dimensions" have been defined. In most cases they are related to the three above.

The term fractal was originally introduced to describe sets whose Hausdorff dimensions exceed their topological dimensions (Mandelbrot, 1972). In this thesis the term is used for any set which scales in a self-similar manner. Fractals are then classified into groups: the

middle-thirds Cantor set is called a homogeneous fractal. The meaning of this qualification is made clear in §1.4. We find this preferable to the suggestion of Taylor (1987) to call only homogenous fractals "fractal."

In this section the two most common notions of dimension have been described. Before generalizing these geometric quantities to include probability distributions, we will demonstrate the physical relevance of these notions of dimension.

Section 1.2 Deterministic Dynamical Systems

A dynamical system may be defined as a collection of rules which prescribes the manner in which any initial point in phase space evolves as a function of time. Here the phase space of a system is an Euclidean N -space, E^N , where N corresponds to the number of independent real numbers required to completely specify an arbitrary initial condition (state) of the system. The phase space dimension of the system is equal to N and to every possible state of the system there corresponds an N -vector $\mathbf{x}$ in this space:

$$\mathbf{x} \equiv (x_1, x_2, \dots, x_N) .$$

As the system evolves in time, a trajectory is defined by the sequence of states through which it passes. The focus here will be on deterministic systems; a system is deterministic only if a complete knowledge of the future (and past) state of a given realization is encoded in its current state; given the rules governing the system, a single space-time point defines the entire

trajectory uniquely. This excludes any random or stochastic elements from the system. A chaotic system is completely deterministic in this sense.

Note that almost all "random number generators" available on digital computers are deterministic dynamical systems. Whether the distinction between random and deterministic generators is important depends on the application; numbers from deterministic generators will be called pseudo-random or p-random. An early p-random number generator suggested by Ulam and von Neumann (1947) is a special case of what is now one of the most studied chaotic systems - the logistic map (see §3.3). Its utility as a p-random number generator is still a subject of interest (Arneodo and Sornette, 1984).

The rules governing dynamical systems of physical interest are often described by difference or differential equations, the choice determined by whether it is more natural/convenient to treat time as a continuous or discrete variable (cellular automata (Wolfram, 1985) may be considered as an additional example). There has been some striking progress in the modeling and interpretation of physical systems through the use of nonlinear dynamical systems in this way. Qualitative behavior such as period doubling and quantitative predictions such as the values of "universal" numbers have been made in systems ranging from dripping faucets (Crutchfield et al., 1986) to the behavior of convection in liquid helium (Libchaber and Maurer, 1982). In addition to fixed points, limit cycles and chaos, seemingly chaotic behavior which slowly dies away has been observed in the work of some authors (e.g. Spiegel, 1986); this transitory

chaos has been called pandemonium.

Hamiltonian systems form an important special case of dynamical systems. Introductions to the exploration of chaos in Hamiltonian systems are given by Hénon (1983) and Berry (1978) and in the textbook of Lichtenberg and Lieberman (1983). This brief introduction demonstrates the fundamental role played by dimension in Hamiltonian dynamics and serves to preface the discussion of a system of this type given in chapter 5.

Consider a deterministic dynamical system with an N dimensional phase space. The system is Hamiltonian if it is possible to write the N variables x_i as n pairs of conjugate variables

$$q_i, p_i \quad i=1,2,\dots,n; \quad n = \begin{cases} N/2 & \text{for } N \text{ even} \\ (N+1)/2 & \text{for } N \text{ odd} \end{cases}$$

in such a way that the dynamics of the system are described by a single function $H(q_i, p_i)$ through the relations

$$H(p_i, q_i) = 0 \quad \dot{p} = - \frac{\partial H}{\partial q} \quad \dot{q} = \frac{\partial H}{\partial p} \quad (1.13)$$

In two-dimensional fluid dynamics, a stream flow is one for which there exists a function $\Psi(x, y, t)$ such that the velocity field determined by

$$\dot{x} = \frac{\partial \Psi(x, y, t)}{\partial y} \quad \dot{y} = - \frac{\partial \Psi(x, y, t)}{\partial x} \quad (1.14)$$

All such systems have Hamiltonian structure. The motion of tracers in this type of flow has been considered by Smith and Spiegel (1985;

also Smith, 1984 and Chapter 5). An experimental realization of Lagrangian turbulence has been analyzed and modeled by Chaiken et al. (1986, 1987).

A similar Hamiltonian structure exists for dynamical systems composed of interacting point vortices in two-dimensional flows. Vortex dynamics may be observed in superfluid helium (see Roberts and Donnelly, 1974). The appearance of chaotic motion in these systems has been studied by Aref (1983).

The Hamiltonian structure of a system imposes constraints on the types of motion it may exhibit. Such systems are deterministic (one to one), and the Liouville theorem (eg. Lichtenberg and Lieberman, 1983) guarantees that the N-volume of an ensemble of initial conditions will be preserved under the flow. By requiring the conservation of N-volume, Liouville's theorem excludes the existence of any attracting sets in the full phase space of a Hamiltonian system.

A system is integrable if there exists a canonical transformation to new variables P_i, Q_i such that

$$H_0(P_i) = 0, \quad \dot{P}_i = - \frac{dH_0}{dQ_i} = 0, \quad \dot{Q}_i = \frac{dH_0}{dP_i} \equiv \omega_i(P_1, \dots, P_n) . \quad (1.15)$$

Hence,

$$P_i(t) = P_i(t=0) , \quad Q_i(t) = \int_{t_0}^t \omega_i(P_1, \dots, P_n) dt \quad (1.16)$$

and the system has been reduced to quadratures.

In the majority of pedagogical examples, a trajectory does not wander throughout phase space but is confined to a lower dimensional

surface by constants (integrals) of the motion, such as the total energy, the angular momentum and so on. If a sufficient number ($N/2$) of these integrals exist, the motion may be placed in the form of Equations 1.15. In this case, the motion of each (P_i, Q_i) pair is independent; each pair represents one degree of freedom of the system. If the solution is bound to a restricted region of phase space, the Q_i must appear as arguments of cyclic functions and the system evolves on an n -dimensional torus (an n -torus). Both the topological dimension and the Hausdorff-Besicovitch dimension of an n -torus are equal to n (as are d_c , d_f , and so on). For a typical integrable system, the topological dimension of the complete trajectory is equal to the number of degrees of freedom (over-integrable systems are an exception; see Henon, 1983). It is in this way that the topological dimension of an orbit is related to the properties of the system. While Liouville's theorem requires every ensemble to maintain an N -dimensional volume, the individual trajectories of an integrable system are each confined to a manifold whose topological dimension is less than (in general one half) the phase space dimension of the system.

This idea is demonstrated by two uncoupled pendula. The phase space of this system is four dimensional (angular position and velocity of each pendulum); trajectories are restricted by the total energy to a three dimensional surface. As long as they remain uncoupled, the energy of each pendulum is conserved individually and the motion is on a two-torus.

Due to the difficulty of visualizing objects which sit in spaces

of dimension greater than or equal to three, phase space trajectories are usually viewed by taking a surface of section. The method, credited to Poincaré, entails following a trajectory in N dimensions and recording the locations at which it passes through a given $N-1$ dimensional surface (in a given direction). For a system constrained to a three dimensional surface, this would correspond to recording the points at which the trajectory pierces a chosen plane. Many of the properties of the full flow may be determined from such a surface of section, for example the capacity and topological dimension of the section of an orbit should be one unit less than the corresponding value of the flow. For example, a Poincaré section of a 2-torus is a circle. Integrable Hamiltonian systems may be perturbed in two distinct ways. Consider first a perturbation, H' , which does not destroy the Hamiltonian structure of the original integrable system, H_0 . That is one which may be written

$$H = H_0 + \epsilon H' \quad \dot{q} = \frac{dH}{dp} \quad \dot{p} = - \frac{dH}{dq} \quad (1.17)$$

At present, there is no method to determine a priori whether or not a given system is integrable. The current search for a general method to answer this question is discussed in its historical context by Tabor (1984). A complementary approach to the analytic search for integrals of the motion is provided by numerical examination of individual trajectories. The system is assumed to be integrable over any range of initial conditions where the computed trajectories appear to lie on tori. Trajectories which appear not to lie on any torus and seem to display sensitive dependence on initial condition are called chaotic. Computer surveys by Hénon and Heiles (1964) and

Walker and Ford (1969) revealed a complex, self-similar interweaving of chaotic and integrable trajectories. While the intricacy of this structure was envisioned by Poincaré; the question as to which type of motion prevailed in the limit of small but finite perturbation was not resolved until the 1960's (See Moser, 1973). A simple perturbed Hamiltonian system is discussed in Chapter 5.

Other physically motivated modifications of system 1.15 add terms which destroy the Hamiltonian structure. These systems are often referred to as dissipative regardless of whether or not they were derived from a Hamiltonian formulation. For such systems there is no Liouville theorem; the volume of an ensemble of initial conditions may contract as the ensemble evolves under the flow and the asymptotic motion of the system is generally governed by an attractor. Consider the system of two uncoupled pendula when both pendula are damped. For almost all initial conditions, the asymptotic state of the system is motionless with both pendula at the orientation of the downward plumb line. By almost all we mean that the set of all initial conditions which behave otherwise has Lebesgue measure zero (zero N -volume in a phase space of dimension N). Equivalently, this set has dimension less than the set of all initial conditions which approach the attractor, and a randomly selected initial condition will belong to the basin of the attractor with probability one. In this particular case, the attractor is a set of zero dimension (a point) and the basin of attraction consists of the entire four-dimensional phase space with the exception of the three-dimensional manifolds on which the kinetic energy of either (or both)

pendulum is such that the pendulum comes to rest at the unstable equilibrium point. Similarly, in the case of a simple limit cycle (damping of only one of the pendula) the attractor would be one-dimensional. Again, knowledge of the dimension of the attractor yields insight into the dynamics of the system.

In addition to limit cycles, attractors often are observed where the large time motion is neither periodic or quasiperiodic (motion on an n -torus), but which is irregular. A severe truncation of the equations describing convection, originally proposed by Lorenz (1963) to demonstrate that the difficulty of weather prediction might be intrinsic to the system, has produced the paradigm of strange attractors⁵. The Lorenz equations have been subjected to intense study which is surveyed and extended by Sparrow (1970). Similar "irregular behavior" was observed by Moore and Spiegel (1966) in a study of stellar pulsations. Here we review the tools of nonlinear dynamics through an examination of the attractors of this system. We note in passing that the phase space trajectories of both of these systems are particular cases of the motion of a particle in a time-dependent potential, where the potential is described by polynomial of the position (see Marzec and Spiegel, 1980).

The original Moore-Spiegel model is a simple thermo-mechanical oscillator; a fluid parcel oscillating vertically in a temperature-stratified background atmosphere. Physically, the restoring force for the oscillations may be, say, differential rotation; while the

⁵This term was first coined by Ruelle and Takens; we maintain the original definition as applying to any attractor the time dependence on which is not periodic or quasiperiodic.

thermal interaction of the parcel with the surrounding fluid leads to the asymmetries in the energy flow which may amplify the oscillations. All displacements are assumed to be small when compared to a scale height of the background atmosphere, implying a Boussinesq equation of state (the density variation of both the parcel and the ambient fluid is linearly proportional to the temperature deviation from some standard temperature). With the additional assumption that the thermal interaction is described by the Newton law of cooling (valid in the case of an optically thin parcel, see Spiegel, 1957), the system may be described by three time scales: the gravitational free fall time as modified by buoyancy, the cooling time due to thermal/radiative interactions, and the time constant of the restoring force.

For a parcel in a linear temperature field, small oscillations either grow or decay depending whether or not the stratification is stable. A nonlinearity is introduced through the consideration of a locally unstable temperature profile which varies quadratically with the vertical displacement, z . Such a profile mimics an unstable layer sandwiched between two stable layers. Measuring the time in units of the cooling time and distance in units of the half height of the unstable layer leads to the equation of motion of the parcel (Moore and Spiegel, 1966)

$$\frac{d^3 z}{dt^3} + \frac{d^2 z}{dt^2} + (T - R + Rz^2) \frac{dz}{dt} + Tz = 0, \quad (1.18)$$

where T and R are positive real numbers. Physically, T represents the square of the quotient of the thermal relaxation time divided by

the free period of the oscillation while R is the square of the ratio of the relaxation time to the free fall time. For $R \gg T$ or $R \ll T$, numerical simulations of system 1.18 display, after a transient, periodic behavior; alternatively, in a band of R , T values near $R = T$, aperiodic motion is observed in a restricted region of phase space. A limit cycle behavior of the system is revealed in Figure 1.3a, where $z(t)$ is plotted for the parameter values $R = 100$ and $T = 10$ (hereafter Case I); the motion lies on a periodic attractor and small perturbations relax back to this attractor. Two examples of aperiodic motion are found for $T = 26$ and $R = 100$ (Case II), and for $T = 36$ and $R = 100$ (Case III). Figures 1.3 b,c are time series of z in these cases. Case II is the source of the power spectrum of Figure 1.1. More detailed spectra of Cases I and II are provided in Figure 1.4. Periodic orbits exist for all three sets of parameter values (Baker, Moore and Spiegel, 1966, 1971). In Cases II and III there are several periodic orbits, all of which are unstable. In the Case II time series, the system trajectory is observed to wander between the neighborhoods of different periodic orbits in what appears to be an erratic manner. Understanding these unstable periodic orbits is central to describing the structure of strange attractors.

A stereogram of the phase space trajectory of Case I is shown in Figure 1.5. The complications of the chaotic trajectories reduces the quality of their stereograms; perspective projections of these trajectories from several viewpoints are shown in Figures 1.6 and 1.7.

A Poincaré section of a trajectory of Case I reveals a collection of points scattered about several specific locations (the number depends on the section taken) which are visited in a well defined order. The scatter is consistent with the numerical accuracy of the calculation. In contrast, a section from Case III reveals no simple geometry or ordered motion. Figure 1.8 is a section showing the velocity and acceleration at successive downward passes through the $z=0$ plane, while a blow up of this section is given in Figure 1.9. The section shown in Figure 1.8 and 1.9 corresponds to structure in the lower left branch of the attractor as viewed in Figure 1.7 a. Evidently, there is no simple description of the system in this parameter regime.

Long integrations of the Lorenz or Moore-Spiegel equations, and observations of the systems they are designed to represent, indicate that this complex motion is not transitory. Numerically, these systems display sensitive dependence on initial condition and they exhibit period doubling bifurcations as the parameters are varied (see Cvitanović 1984, introduction). The evolution of nonlinear ODE's which make up these systems requires a time consuming integration of the system. In an effort to make the structure of the attractors of such systems more accessible, Hénon (1976, see also Hénon and Pomeau, 1977) developed a two-dimensional iterative mapping to simulate the Poincaré sections directly. Designed to represent dissipative systems in general and the Lorenz system in particular,⁶

⁶ A map designed to simulate Moore-Spiegel type systems has been developed by Arneodo et al. (1985).

Hénon determined the most general quadratic map with constant Jacobian to be:

$$\begin{aligned}x_{i+1} &= 1 - ax_i^2 + y \\ y_{i+1} &= bx_i\end{aligned}\tag{1.19}$$

This map has been discussed extensively in the literature. A review is given by Hénon (1980). When the absolute value of the Jacobian of a map (here equal to $|b|$) is less than one, the map is area contracting. In this case, the area of an ensemble of initial conditions will be reduced by a factor of $|b|$ after each iteration of the mapping. By demonstrating a region of the plane which is mapped into its interior, Hénon proved this mapping possessed a trapping region; such a region must contain an attractor. As the value of parameter a is increased, this mapping also follows a period doubling cascade to chaos (Derrida et al., 1978). For the canonical parameters ($a = 1.4$, $b = 0.3$) the mapping appears to have sensitive dependence on initial condition.

Figure 1.10 shows 2^{20} iterations of the map starting from a 64 bit approximation of the fixed point. It appears that this attractor is not periodic, an assumption given empirical support by the self-similar structure observed when the circled region is magnified. It is generally believed (although not proven) that the asymptotic motion described by this system has this type of structure at all scales. This set is the Hénon attractor. Numerical evidence indicates that for this set (Russel et al., 1980)

$$d_c = 1.261 \pm 0.003$$

In addition to the lack of strict self-similarity, two additional

quantitative features distinguish this set and those of Figures 1.5 and 1.6 from the middle thirds Cantor set. The first is a richer mathematical structure due to the nonuniform distribution of points on any coarse grained representation of the attractors. The second is a physical limitation of scaling properties due to the nature of the underlying physical (or computational) system. Physical constraints and the relevant probability theory are discussed in the next section where a few more scaling exponents are introduced. Further discussion of these features is presented in §3.1 where the oscillatory component of the scaling is utilized to quantify these properties with precision.

As an important aside we note that the orbit displayed in Figure 1.10 must evolve onto a periodic orbit. All the calculations were done on a digital computer capable of representing only a finite number of distinct values. Thus every deterministic trajectory must eventually close upon itself. Furthermore, there cannot be true sensitive dependence on initial conditions; if the initial condition is specified to an accuracy beyond the resolution of the numerical grid the future trajectories are identical. The theorems of dynamical systems were developed for continuous systems which are not accurately represented by discrete computational grids. The effects of treating finite separations as infinitesimal in the determination of Lyapunov exponents from experimental data has been studied by Alan Wolf (personal communication). The ultimate periodicity of orbits and apparent chaos in digital computers are explored in Appendix 1.

SECTION 1.3 Probability Distributions on Fractals

The mathematical definitions of dimension in the previous section were given in the limit of small scales. In physical realizations of scaling there exists an inner cutoff below which the scaling relations do not hold. The early 19th century wood print, *The Great Wave off Kanagawa* by Hokusai (see Mayor, 1985), often cited as representative of self-similar cascades in nature, also captures the fundamental limits to any realization of this self-similarity. The great wave is composed of lesser waves, each of which in turn has a finer structure similar in form to the first: hence, self-similarity. After several generations the cascade stops, there is a well-defined smallest wave size, as there is a largest; no wave exceeds Fuji. The concepts of an inner and outer cutoff of the scaling region are fundamental in all physical realizations of self-similarity. The terms self-similar and scaling are nonetheless used to describe physical and numerical distributions where the inner cut off prohibits complete reproduction under magnification. Implicit in doing so is the assumption that there exists a "large" range of scales over which scaling holds approximately. The question of quantifying these properties in empirical data is taken up in Chapter 2.

Even in the case of noise-free data, the inner cutoff may not be observable for at least two reasons. The limited accuracy of any measuring process imposes a partition upon the (assumed continuous) variables which are recorded. If this threshold is above the inner cutoff, the latter will not be observed. Even when the threshold is

sufficiently low as to make the cutoff observable in principle, the quantity of data required to detect it may exceed the capacity of the analysis. Since the amount of data required to determine a surface grows exponentially with the dimension of the set, the effect may pose a real, if technology dependent, constraint.

Consider an experiment in which the finest partition available fails to resolve the inner cutoff. From a histogram of the data, often computed as the experiment proceeds, the relative probability of the i^{th} partition, p_i , may be computed. The Hausdorff dimension (equivalently d_c and d_f) is not sensitive to the magnitudes of the p_i , but only to the number that are nonzero. Assuming this probability distribution converges (i.e. there exists an invariant measure underlying the system) one may compute d_I , the information dimension (Farmer 1982)

$$d_I = \lim_{\epsilon \rightarrow 0} \frac{I'(\epsilon)}{|\log \epsilon|} \quad (1.20)$$

where

$$I'(\epsilon) = \sum p_i \log p_i \quad (1.21)$$

where the sum is over all boxes with nonzero probability. $I'(\epsilon)$ represents the amount of information gained, on average, by an isolated measurement of the system using instruments of resolution ϵ .

The information dimension is just one of an infinite number of generalized dimensions originally defined in the context of information theory by Balatoni and Renyi (1956). General introductions to information theory are Shanon and Weaver (1964) and Renyi (1970, Appendix A); discussions in the context of dynamical

systems are available in Shaw (1981) and Farmer (1982, 1983, 1984). The mathematical foundations of the material presented in this section were laid by Renyi based on the ideas of Boltzmann, Kolmogorov and Shannon. Loosely speaking the amount of information in a signal (measurement) is related to the minimum number of bits required to identify that measurement from the set of all possible measurements. Following Renyi (1970), consider a random variable ξ which assumes denumerably many values x_k each with probability p_k . That is

$$p_k \equiv \text{Prob}(\xi = x_k) \quad k = 1, 2, 3, \dots \quad (1.22)$$

and define the information of order α in the value of ξ , $I_\alpha(\xi)$, as

$$I_\alpha(\xi) = \frac{1}{(1 - \alpha)} \log \left(\sum_{k=1}^{\infty} p_k^\alpha \right) \quad \alpha \neq 1$$

$$I_1(\xi) = \sum_{k=1}^{\infty} p_k \log \frac{1}{p_k} \quad (1.23)$$

The ξ dependence of I_α is implicit in the determination of the p_k . If the distribution of ξ were continuous, a single irrational value of ξ would contain an infinite amount of information. With the physically motivated argument that ξ is known only with finite precision and on the assumption that the observed values converge to the continuous values as the precision is increased, Renyi introduces the partitioned variable

$$\xi_N = \frac{[\xi N]}{N} \quad (1.24)$$

where $[x]$ denotes the greatest integer less than or equal to x . For a continuously distributed ξ , the information tends to infinity as $N \rightarrow \infty$ but the quantity

$$d_\alpha(\xi) = \lim_{N \rightarrow \infty} \frac{I_\alpha(\xi_N)}{\log N} \quad (1.25)$$

exists in many cases and is called the dimension of order α of the ξ . Renyi extends the definition to m -dimensional random vectors ($m = 1, 2, 3, \dots$), proving that for an absolutely continuous distribution the d_α are equal and correspond to the geometric notion of dimension (that is $d_\alpha = m$) and suggesting the application of the ideas of Hausdorff.

In the case $\alpha = 0$

$$I_0 = \log \sum_{k=1}^{\infty} (p_k)^0 = \log M(\epsilon)$$

where $M(\epsilon)$ is the number of hypercubes of side $\epsilon = 1/N$ which have nonzero probability. Thus

$$d_0 = \lim_{N \rightarrow \infty} \frac{\log I_0(\xi)}{\log N} = \lim_{\epsilon \rightarrow 0} \frac{\log M(\epsilon)}{|\log \epsilon|}$$

and d_0 is equivalent to the fractal dimension. Similarly d_1 is the information dimension

$$d_1 = \lim_{N \rightarrow \infty} \frac{\sum p_k \log p_k}{\log N} = d_I$$

This case has been studied extensively by Farmer in the context of dynamical systems.

Since

$$\begin{aligned} p_i^{\alpha'} &< p_i^{\alpha} & \text{for } \alpha < \alpha', \\ I_{\alpha'}(\xi_N) &\geq I_{\alpha}(\xi_N) & \text{for } \alpha < \alpha' \end{aligned} \quad (1.26)$$

which implies

$$d_{\alpha'} \geq d_{\alpha} \quad \text{for all } \alpha < \alpha'. \quad (1.27)$$

Any set for which $d_0 > D_T$ and $d_{\alpha} = d_0$ for all α will be referred to as a homogeneous fractal. The middle thirds Cantor set is of this type. Note that d_{α} may depend on the geometry of the set, the metric, and the probability distribution. As pointed out by Renyi and emphasized by Ott et al. (1984), the d_{α} are not invariant to smooth changes in coordinate.

A similar formalism was described by Hentschel and Procaccia (1983) and applied to several problems of nonlinear dynamical systems theory. At this point we adopt their notation letting $\alpha \rightarrow q$ and $N \rightarrow 1/\epsilon$. Here q is considered a continuous real variable ($-\infty < q < \infty$) (see also Grassberger and Procaccia, 1984).

Recently, Halsey et al. (1986) and Kadanoff (1987) have introduced a new formalism which relates the d_q to a smooth function, $f(\alpha)$, by a Legendre transformation. This function characterizes the strengths of the singularities of the set. Halsey et al. state the following relations (their equations 1.10 and 1.11)

$$d_q = \frac{1}{q-1} [q\alpha(q) - f(\alpha(q))] \quad (1.28)$$

$$\alpha(q) = \frac{d}{dq} [(q-1) d_q] \quad (1.29)$$

$f(\alpha)$ is related to the dimension of the set upon which singularities of strength α lie. Having calculated the d_q with equation (1.25), one may compute $\alpha(q)$ with (1.29)⁷ and then $f(\alpha)$ via (1.28). In cases where it can be determined, $f(\alpha)$ yields a new way to visualize the information in d_q . A comparison with experiment is provided in Libchaber (1986). Note that for homogeneous fractals all the d_q are equal and so

$$\alpha(q) = d_q = d_0 \quad (1.30)$$

Furthermore,

$$f(\alpha) = \alpha(q) = \text{constant}.$$

We shall consider the effects of sectioning and projecting on $f(\alpha)$ in Chapters 3 and 4, respectively. Calculation of d_q for large negative q from experimental data is difficult because these scaling exponents emphasize the structure of the lowest probability singularities of the set. A method which provides a useful lower bound on d_∞ is developed in Chapter 3. A dimension function has also been introduced by Badii and Politi (1985) who have directly extended the work of Renyi. A somewhat more complicated function representing the scaling structure of a set has been introduced by Feigenbaum (1978, 1979).

⁷Note that in general this requires differentiating a numerically derived quantity.

Conclusions

As evidenced by the dimension functions above, the structure of sets that scale is rich and a useful description requires the consideration of more than a single scaling exponent. The general solution to the scaling equations can be used to obtain additional information regarding a set beyond its dimension (Smith et al, 1986). When determining scaling exponents, we often write the general solution

$$C(\ell) = \chi(\ell) \ell^d \quad (1.6)$$

in the form

$$\log C(\ell) = d \log \ell + \log \chi(\ell)$$

and let

$$\psi(\ell) = \log \chi(\ell)$$

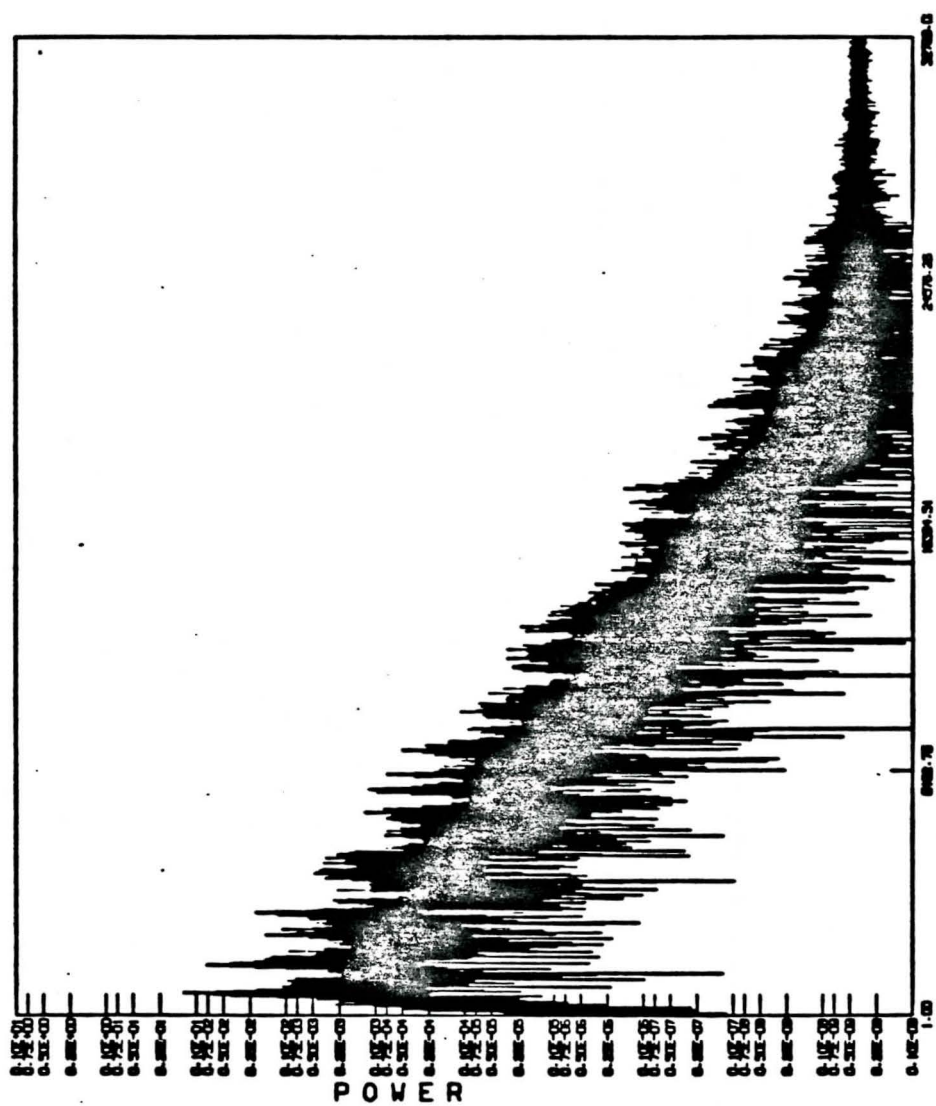
Since $f(\alpha)$ characterizes the singularities of a set, the $\psi(\ell)$ functions describe additional structure to which $f(\alpha)$ is insensitive. Once we have presented a method for calculating the correlation integral, we will examine the structure of this $\psi(\ell)$ function for several fractals.

The function $\chi(\ell)$ (or $\psi(\ell)$) characterizes a textural quality of the underlying distribution. Its precise interpretation depends on the statistical quantity in question, for the minimal cover (and also, though less strikingly, for the correlation integral) $\chi(\ell)$ describes the property of the set called lacunarity (Mandelbrot, 1973). The qualitative character implied by this term is difficult

to capture in a single parameter. The examination of several functionals of $\chi(\ell)$ indicates that it is preferable to retain the information intact in one period of χ itself. Lacunarity describes the lulls in a distribution. The periodicity of the oscillation reveals information as to how different scales are coupled, and as such provides valuable insight into such problems as the spacial intermittency of fluid turbulence. This is the topic of Chapter 4. First, we will discuss the problem of determining the scaling exponents of a distribution. We will see that the ψ functions plays a role here also.

Figure Captions

Figure	Caption
1.1	Power spectrum of a chaotic time series. This figure shows the case of the Moore-Spiegel system (introduced in Section 2) with parameters $T = 26.0$ $R = 100.0$.
1.2	The first four steps in the construction of the middle thirds Cantor set by repeated application of the destruction algorithm. The arrow denotes the point LRRLLLL... .
1.3	Time series of z from the Moore-Spiegel system for parameter values $R = 100$ with a) $T = 10$ (Case I); b) $T = 26$ (Case II); and c) $T = 36$ (Case III). Data were generated using a 6 th order Runge-Kutta routine. The series is sampled with a time-step $\Delta t = 0.05$, individual data points are noted by '+' on the interval $0 \leq t \leq 64$.
1.4	Details in the low frequency power spectra computed from the time series of $z(t)$ for a) Case I and b) Case II.
1.5	Stereograph of the attractor for Case I.
1.6	A perspective view of the attractor for Case II from the point $z = 27.0$, $\dot{z} = 22.0$, $\ddot{z} = 0.0$
1.7	Two perspective views of the attractor for Case III $T = 36.0$ from the point (a) $(z = 30.0, \dot{z} = 5.0, \ddot{z} = 0.0)$. (b) $(z = 27.0, \dot{z} = 22.0, \ddot{z} = 0.0)$.
1.8	Poincaré section of the Moore-Spiegel attractor at $z = 0$ for $T = 36$ $R = 100$. This section corresponds to the lower left region of Figure 1.7a.
1.9	A close-up revealing the many leveled structure of the section shown in Figure 8
1.10	2^{20} iterations of the Hénon mapping (Equations 1.19). The region shown is $-1.5 < x < 1.5$, $-1.5 < y < 1.5$.



MODE

Figure 1

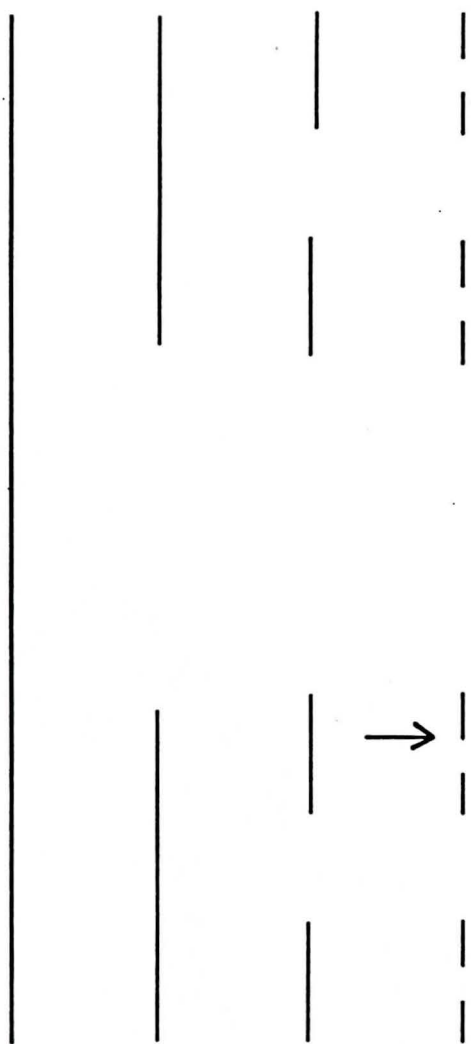


Figure 2

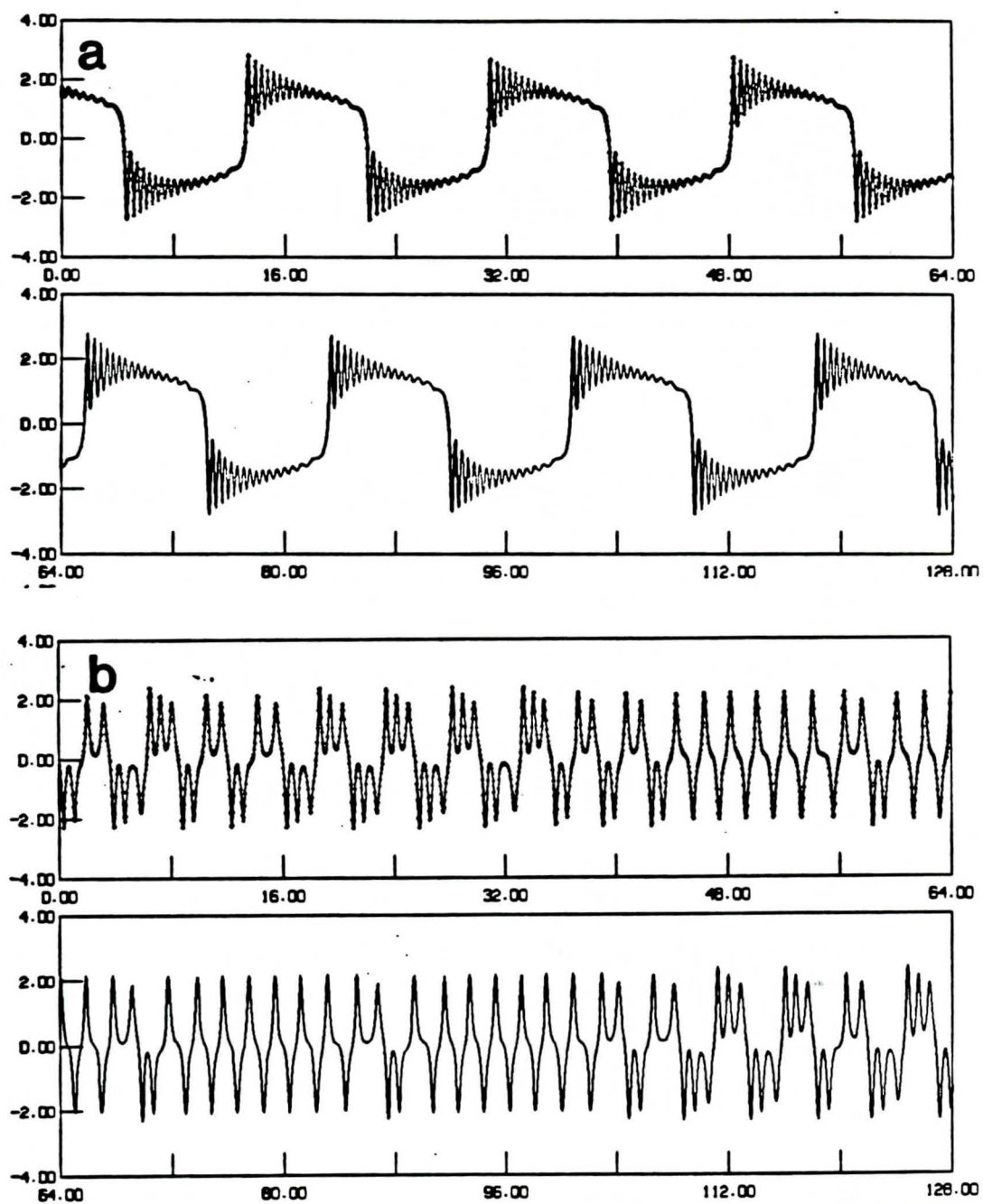
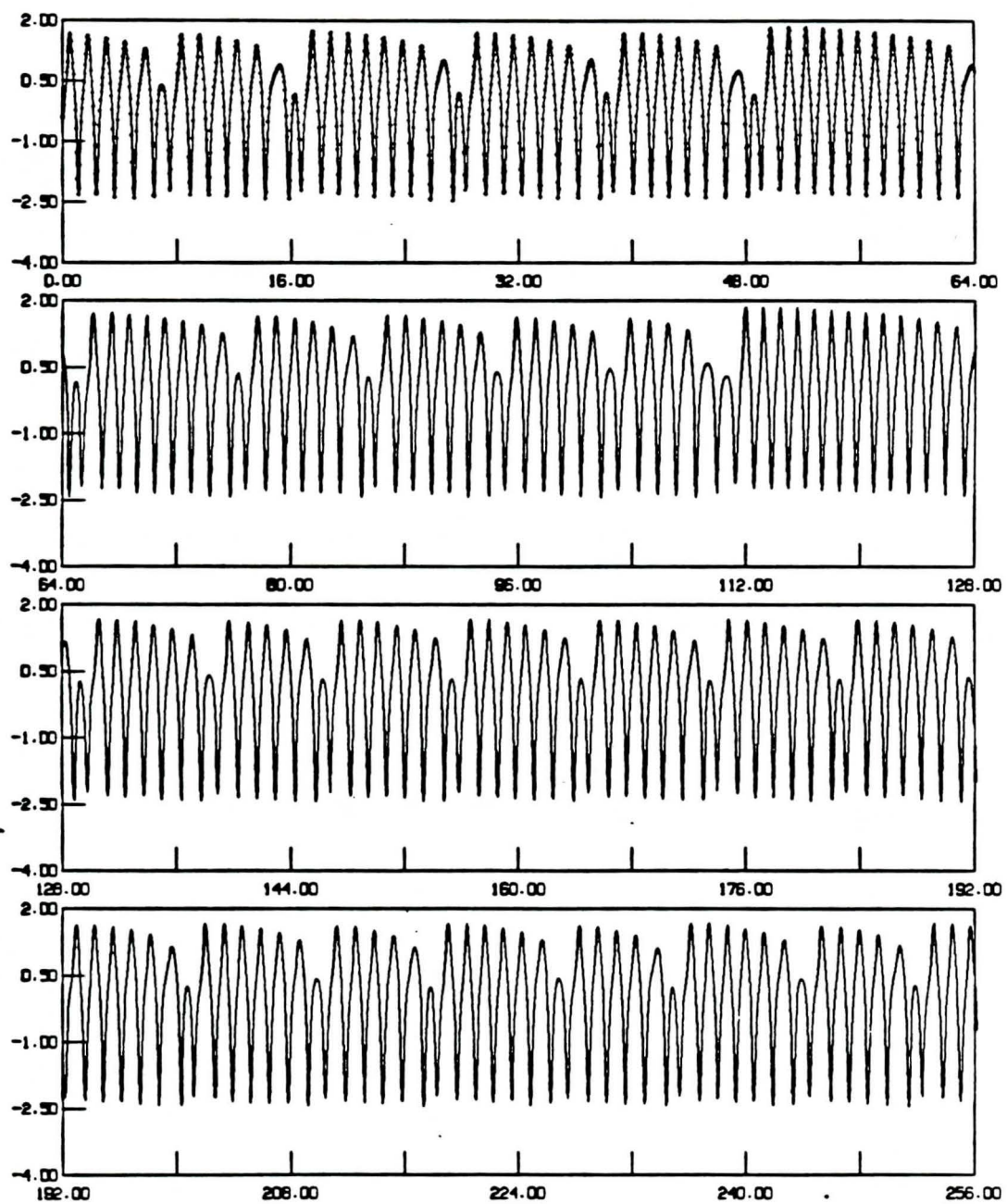


Figure 3

Figure 3



C

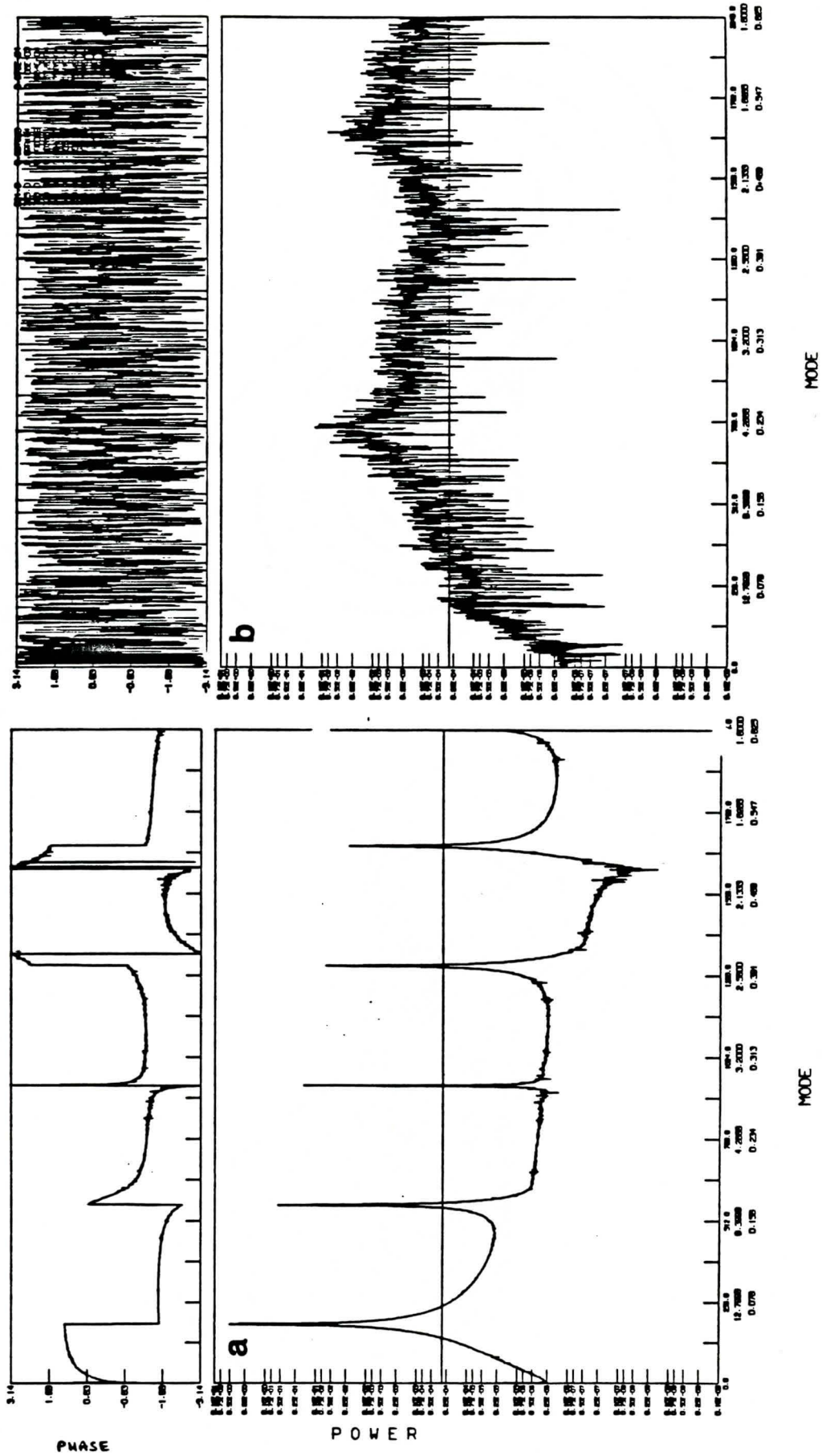


Figure 4

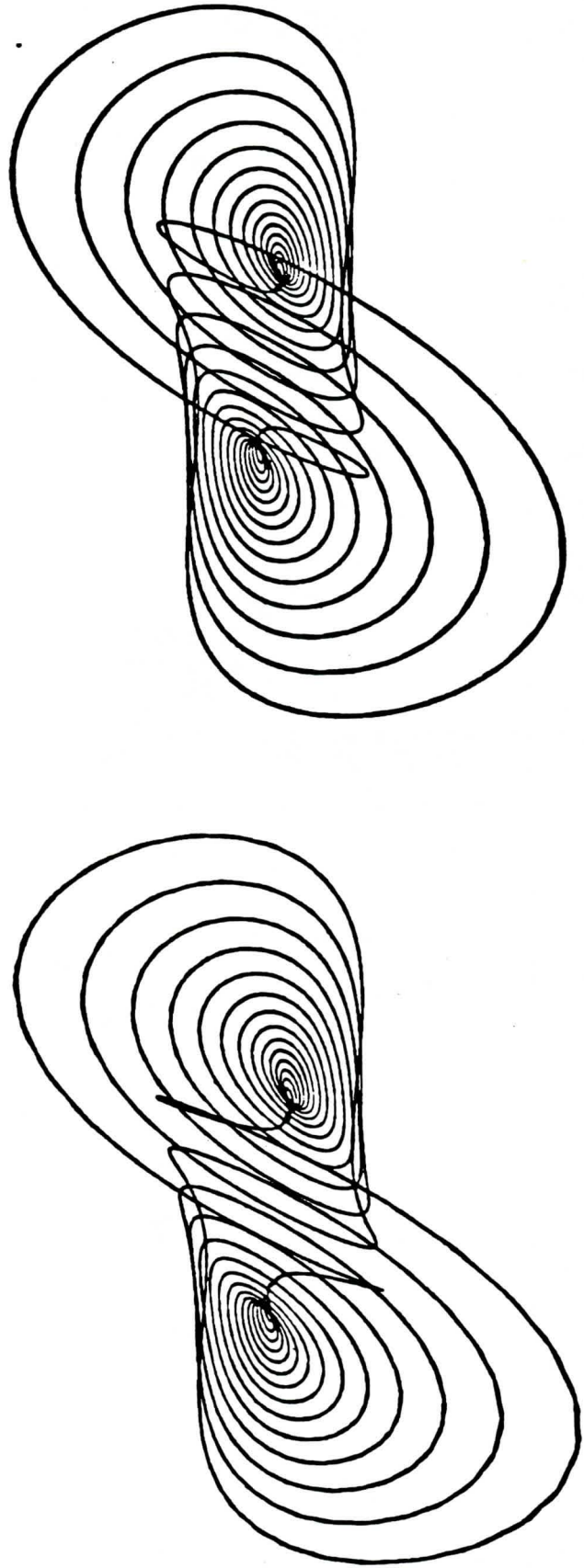


Figure 5

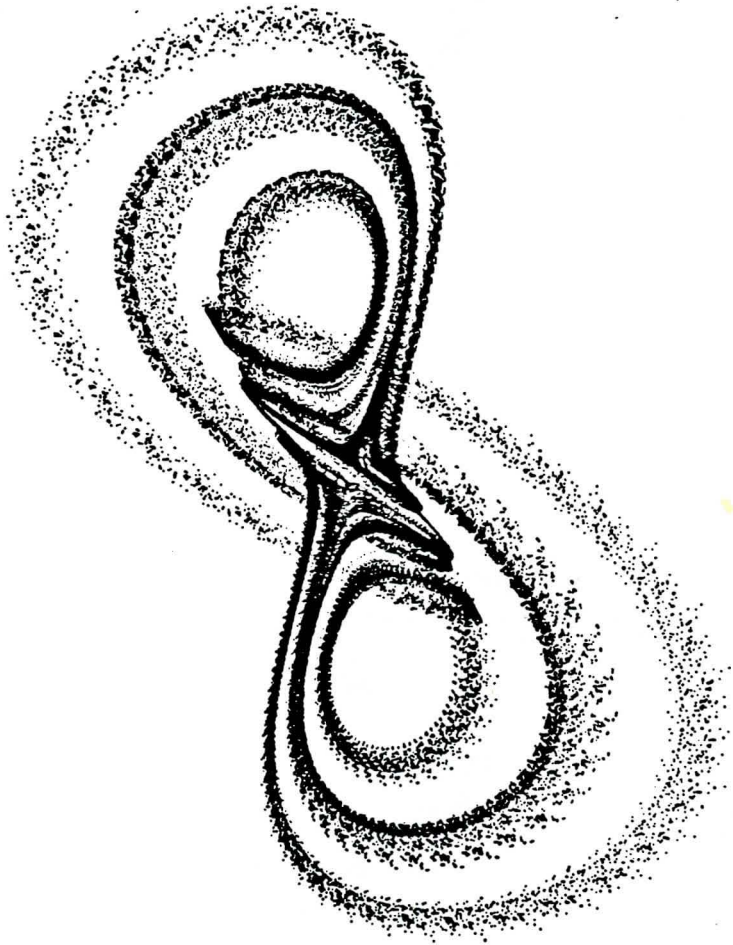


Figure 6

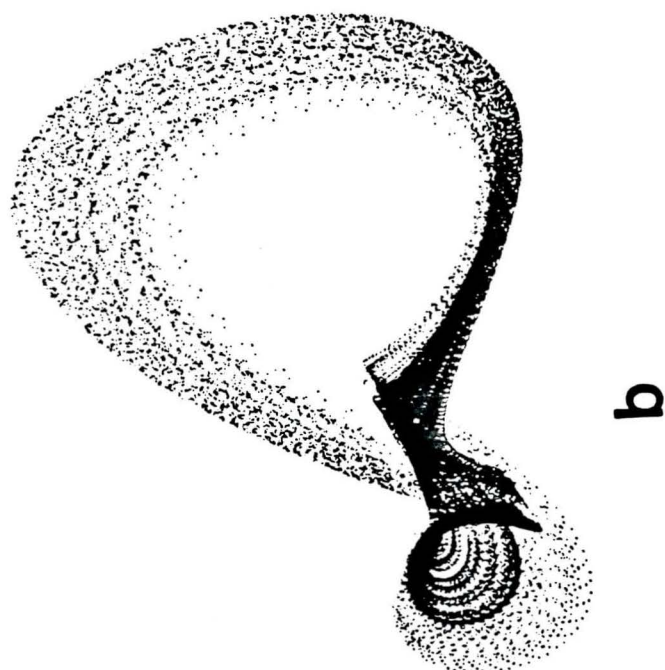
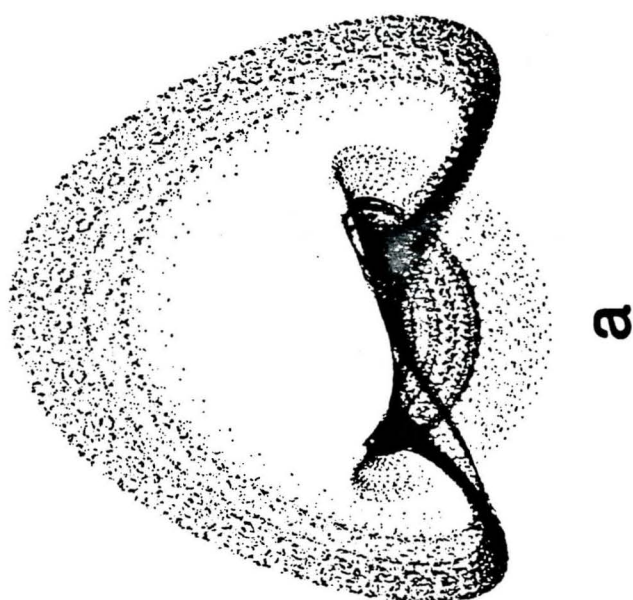
**b****a**

Figure 7

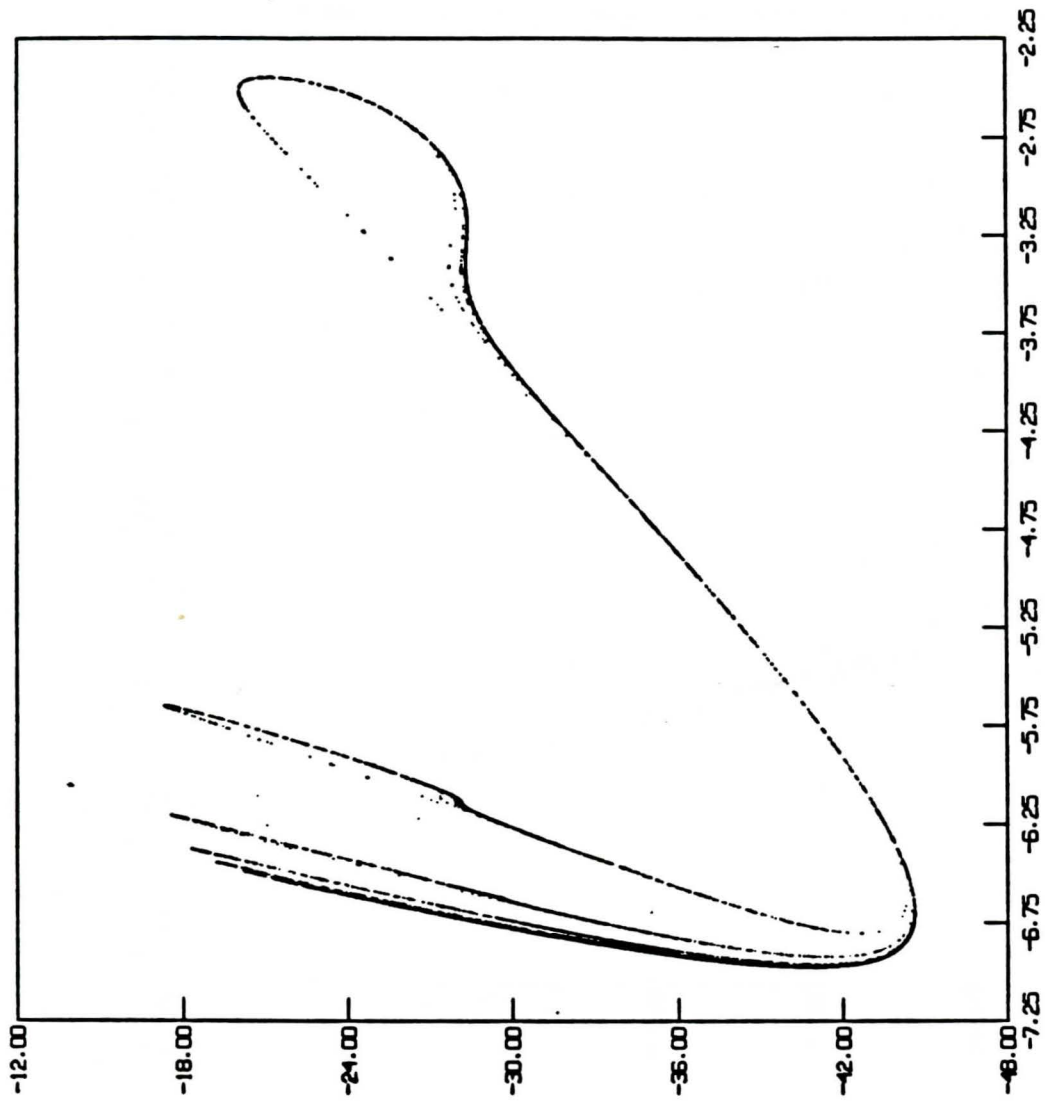


Figure 8

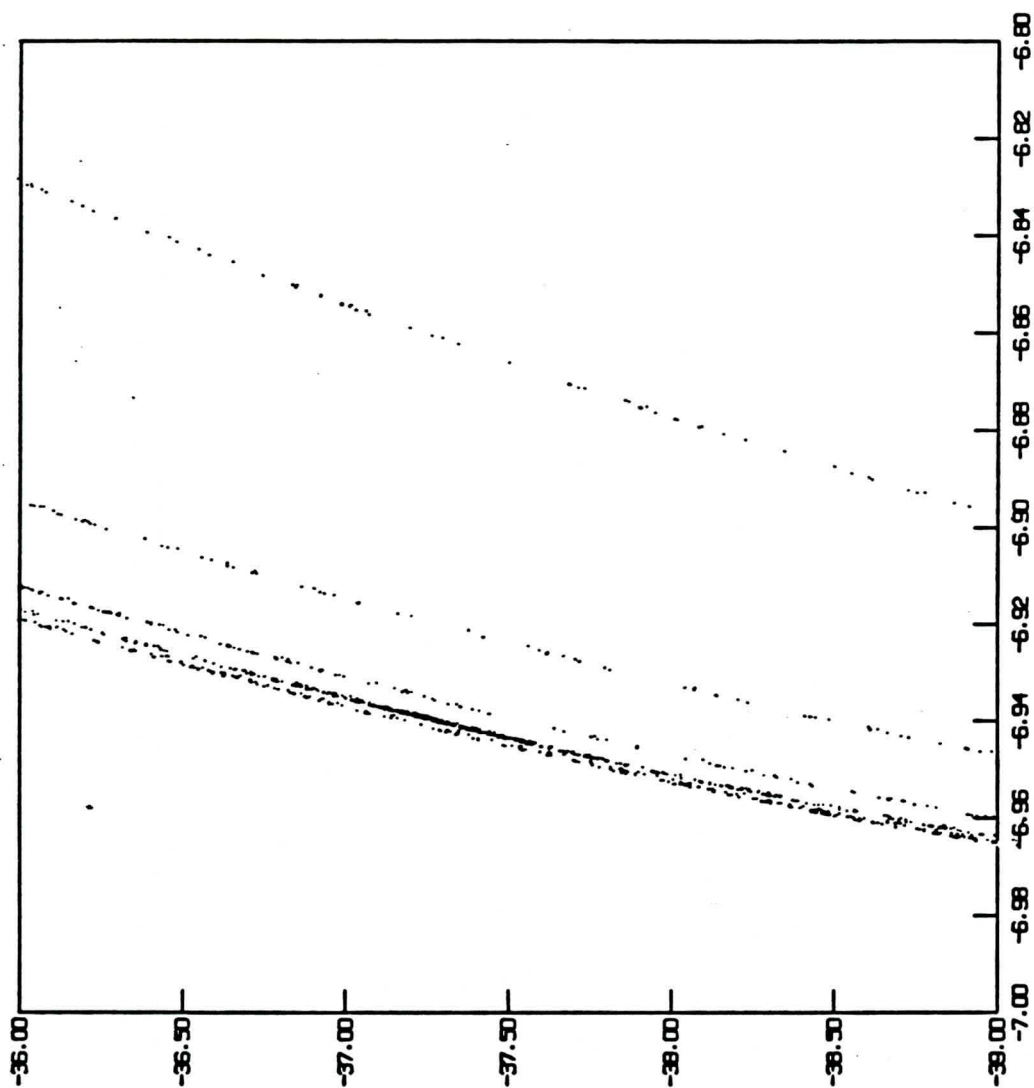


Figure 9

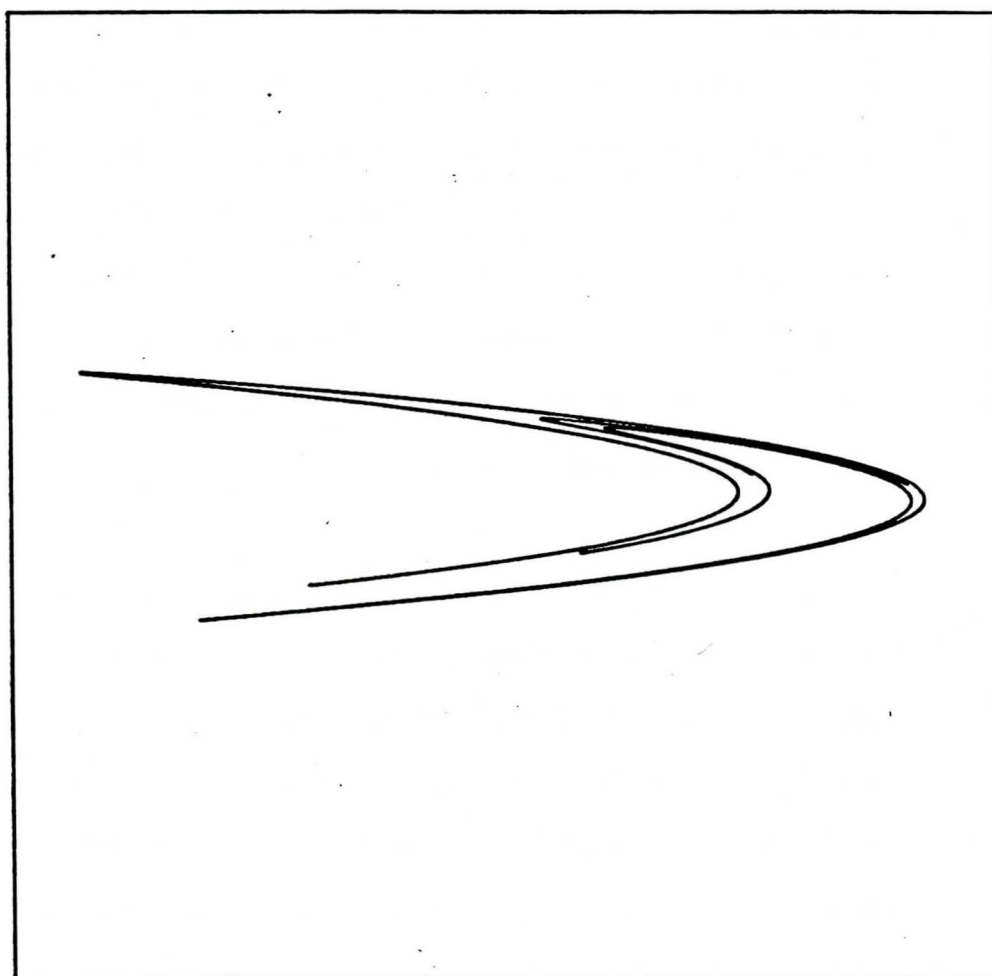


Figure 10

APPENDIX 1 Apparent Chaos in Digital Computers

In this appendix the effect of digitization on chaos is explored¹. It is often remarked that recent advances in the study of nonlinear systems in general and chaotic dynamics in particular are due in large part to the use of modern digital computers. Paradoxically, these machines cannot display truly chaotic trajectories; real numbers are represented by a finite number of bits which implies that the dynamics occur on a finite grid. All bounded motion is eventually periodic. In this sense, the best one can hope for is pandemonium; truly chaotic trajectories are not realizable on a digital computer just as identical initial conditions are not available on an analog computer.

The numerical grid imposes an inner cutoff to both geometrical scaling and sensitivity to initial conditions; by determining the number of distinct states, the grid imposes an upper bound on the cycle time of any deterministic dynamical system. This cycle time may, of course, be very large: the Goddard General Circulation Model (see Hansen, 1983) has something like $(2^{32})^{2^{15}}$ potential states (though the vast majority of these are of no physical relevance). In studies of strange (fractal) attractors, where the quantities of interest are defined in the limit of small scales, the effects of digitization are more pronounced. As an example, consider a simulation of the Hénon map (Equations 1.19) carried out in 32-bit precision (REAL*4 with an IBM 360 architecture; for details, see

¹This problem was suggested to me by Colin Sparrow.

Struble, 1975). A maximum of 2^{32} distinct values restricts the two-dimensional mapping to an orbit of at most 2^{64} iterations before the orbit repeats (or diverges). When the mapping is dissipative, the observed cycle time is significantly less than this.

There are several questions we can ask. How will this periodicity affect the representation of attractors? As the mesh becomes finer, will there be many periodic orbits or only a few? Perhaps even just one? If there are several, how do their structures compare? How are their basins of attraction divided? Does the relative probability of observing one attractor dominate? How machine-dependent are the answers to these questions? One approach to these questions is to examine the dynamics on grids that are sufficiently coarse that a complete catalog of behaviors can be computed, and then to observe how these behaviors change as the grid becomes finer. Consider a uniform grid with mesh ϵ where

$$\epsilon = \frac{1.0}{N}$$

and N is integer. To every real number X we associate an integer IX by

$$IX = [N \cdot X] = \left[\frac{X}{\epsilon} \right]$$

where $[y]$ is equal to the greatest integer less than or equal to y . Here the origin is a grid point. On this grid, the Hénon map² becomes

²a discussion of the Hénon map is given in §3.3. For more details, see Hénon (1976).

$$\begin{aligned}
 IX' &= [N + IY - ((IA \cdot IX)/N_A) \cdot IX]/N \\
 IY' &= [(IB \cdot IX)/N_B]
 \end{aligned}
 \tag{A1.1}$$

and the usual parameter choice ($a=1.4$, $b=0.3$) corresponds to

$$\begin{array}{ll}
 IA = 14 & N_A = 10 \\
 IB = 3 & N_B = 10
 \end{array}$$

Written in this way, the mapping may be iterated using only integer arithmetic.

Once N is specified, all periodic orbits may be identified, in principle, by tracing the path of every initial condition within a region which completely contains its image. While every orbit will eventually close upon itself, the many-to-one nature of the system will produce transient points to which a given trajectory will never return.

Consider the case $N = 100$. At this resolution it is practical to consider the entire self-mapping region and identify all the attracting orbits of the system. There are two: one of period 25, which attracts 99.56% of the initial conditions within the basin; the other a period 1 orbit corresponding to an unstable fixed point of the exact map (the orbits and their basins of attraction are shown in Figure A1.1). This result is not characteristic; for most values of N examined there exist about 6 orbits, two of which usually attract the majority of the initial conditions. An example of the manner in which the basins are interwoven is shown in Figure A1.2 using the

case $N = 110$. As N increases, the number of initial conditions in the bound region becomes too great for a complete search with our current resources. The difficulty is compounded by an increase in the length of transient with increasing N . The results given in Table A1.1 were computed by evaluating small regions of contiguous initial conditions. For the larger values of N , several such areas were evaluated. In general, different regions produces similar statistics for the high probability orbits, although some low-probability orbits were absent from some sample regions.

The use of the 32-bit integers in system [A1.1] places an upper limit on N near 10,000 (due to overflow in evaluating the quadratic term). It is interesting to note that, even at this value, the number of orbits is holding at around 6 or 7, with one or two of the orbits usually attracting 90% of the initial conditions. Throughout the range of grid spacings examined, the number of digitally periodic orbits remains small (say, of order 10), with the majority of points being drawn to one or two of these. The length of the most probable orbits generally increases with N ; however, the longest orbit is often not the most probable. Finally, we note that the self-similar structures of Figure 3.8 are already apparent with $N=10,000$.

The Hénon attractor shown in Figure 1.10 was calculated using 64-bit real numbers. The structure of the periodic orbits of these evolutions is more complicated, partly due to the nonuniformity of the REAL grid; since real numbers are represented in exponent-mantissa form, the mesh varies over the plane, being finer near the coordinates on axes. To get an idea of the effect of this variable

grid and test the machine dependence of the results, a search for periodic orbits using REAL*4 variables was performed. As the number of initial conditions precluded a complete search, p-random initial conditions in the neighborhood of the unstable fixed point were determined. The results for several machines are shown in Table A2.

It is notable that still only a few orbits are found, with one of intermediate length dominating. Note also that the comparison across machines yields similar results.

One source of these orbits is clear -- unstable periodic orbits (stabilized by the grid) coincide with some of the observed low period, low probability orbits. Examining the basins of attraction of the remaining orbits suggests that the orbits fall into two categories: those orbits which move along a stable manifold until the grid pinches it off, and those which fall onto the image of the attractor. It is the digitally periodic orbits of the last class onto which the majority of initial conditions evolve.

To test this conjecture we examined a region that was relatively free of portions of the stable manifold of the interior fixed point. From the figure on page 764 of Franceschini and Russi (1981) [this figure is mislabeled as Figure 3], it seems that the region $x = -0.4 \pm 0.001$, $y = -0.8 \pm 0.001$ would be a good choice. For $N = 10,000$, all but 14 of the initial conditions were drawn to orbit A or orbit B (see Table A1.1). The initial conditions that did evolve to orbit D were all clustered near the edge of the test region. The results are shown in Figure A1.4. (In the REAL*4 calculations on the Amdahl, the probability of observing the Amdahl

orbit A increased by 10% when the initial conditions were drawn from the region given above.) These observations are consistent with the idea that the long, low probability orbits capture initial conditions on a stable manifold; the trajectories from these initial conditions, moving along the stable manifold, are eventually mapped into a gridbox that has been visited previously. The majority of initial conditions appear to evolve onto one or two orbits which resemble the Hénon attractor.

Given the required periodicity of digital orbits, why do such simulations appear chaotic? There are several reasons, not the least of which is the length of the cycle time. For two just-distinguishable initial conditions, both of which approach the same numerical attractor, the rate of separation is still, on the average, exponential; two such orbits may reach the attractor with uncorrelated phases after transients of very different lengths.

In summary, the digitally periodic orbits of the Hénon mapping have been determined for several numerical grids. The length and number of these orbits is surprisingly small considering the relatively high resolution of the grids. The use of this approach to characterize strange attractors at the smallest available scales may provide the ultimate box counting method for probing their structure.

FIGURE CAPTIONS

Figure	Caption
A1.1	The basins of attraction for the two digitally periodic orbits of the Hénon mapping with $a = 1.4$ and $b = .30$ when evolved on a grid with $N = 100$. The statistics of each orbit are tabulated in Table 1. The basins shown correspond to (a) Orbit A and (b) Orbit B.
A1.2	Basins of attraction as described in Figure A1.1 for the case $N = 110$, (a) Orbit A (b) Orbit B.
A1.3	The digitally periodic orbits of the $N = 10,000$ grid. Statistics of the orbits are provided in Table 1 where (a) Orbit A, (b) Orbit B, (c) Orbit C, (d) Orbit D, (e) Orbit F. Orbit E, an unstable fixed point, is not shown.
A1.4	The basins corresponding to the orbits shown in Figure A1.3. In each case the initial conditions which evolve to an orbit in figure 3 are shown in the corresponding panel here. Initial conditions are from the region $-0.81 < x < -0.79$, $-0.81 < y < -.79$.
A1.5	The basins of (a) Orbit A, (b) Orbit B, and (c) Orbit D on the $N = 10,000$ grid. The region was chosen to avoid the stable manifold as described in the text. No initial conditions observed in this region evolved onto the other orbits.
Table A1.1	Statistics for the orbits observed on uniform numerical grids for a variety of grid resolutions.
Table A1.2	Statistics of the digitally periodic orbits observed on REAL*4 grids for a variety of computer architectures.

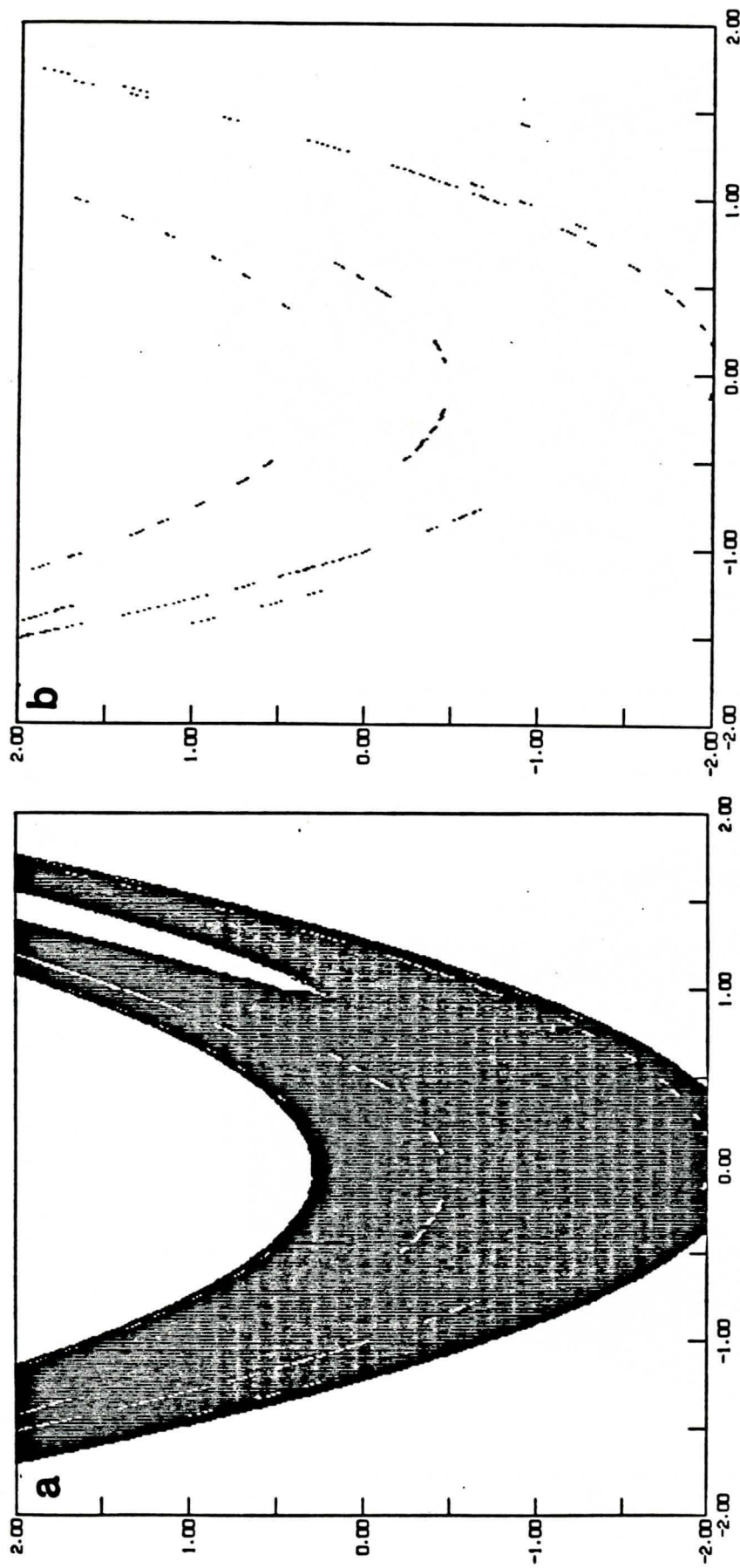


Figure 1

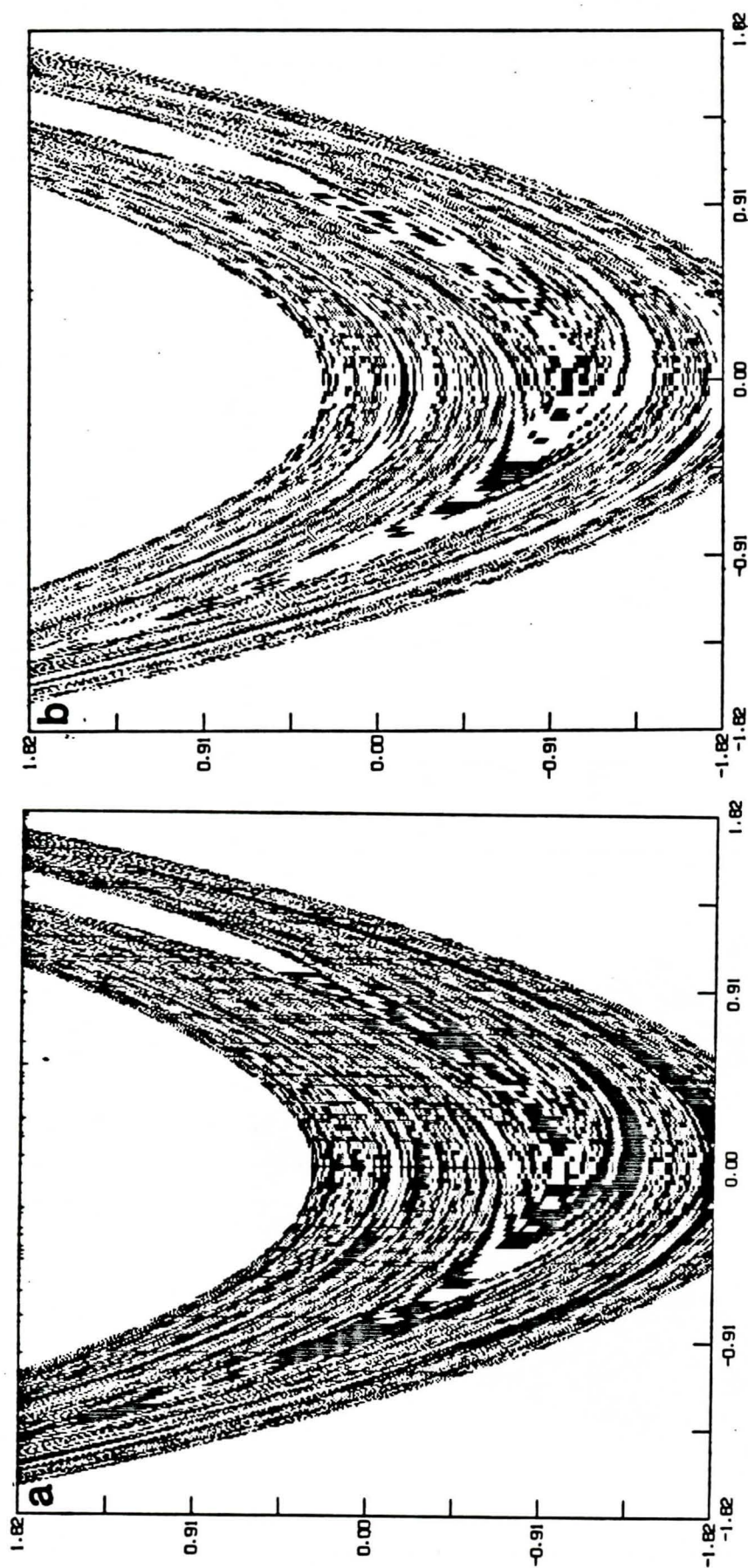


Figure 2

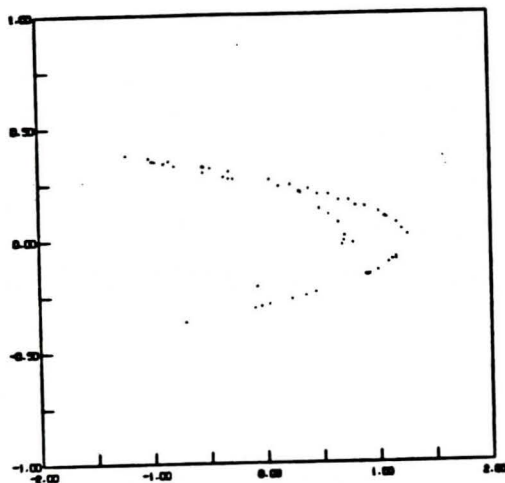
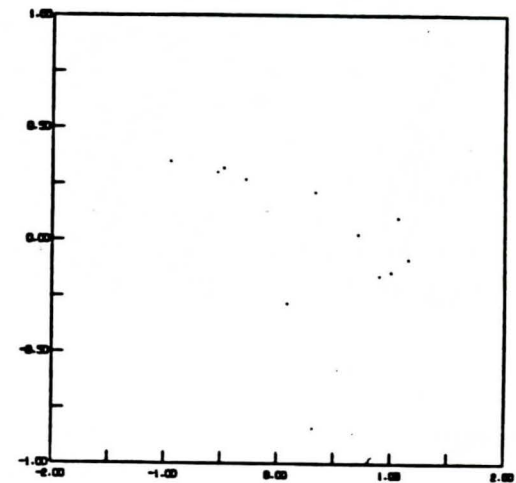
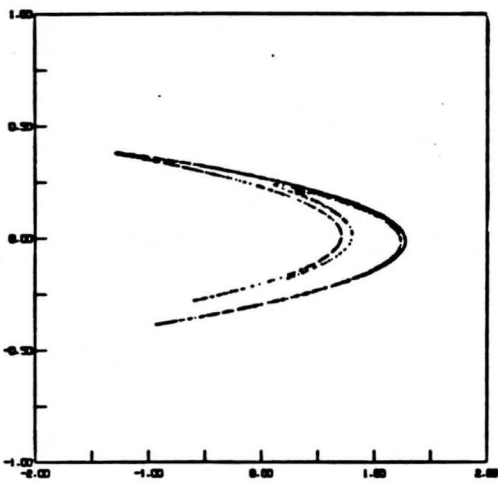
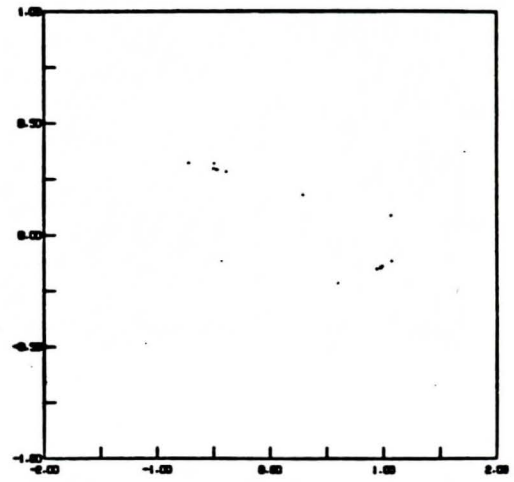
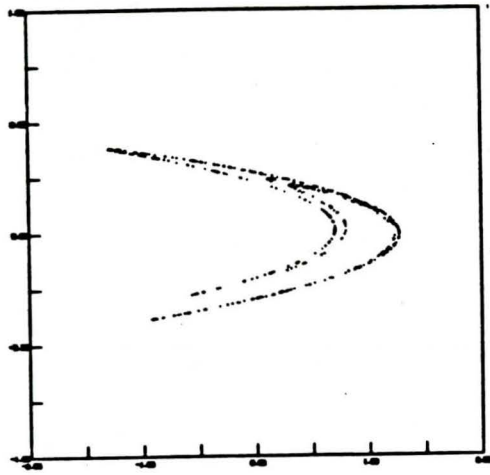


Figure 3

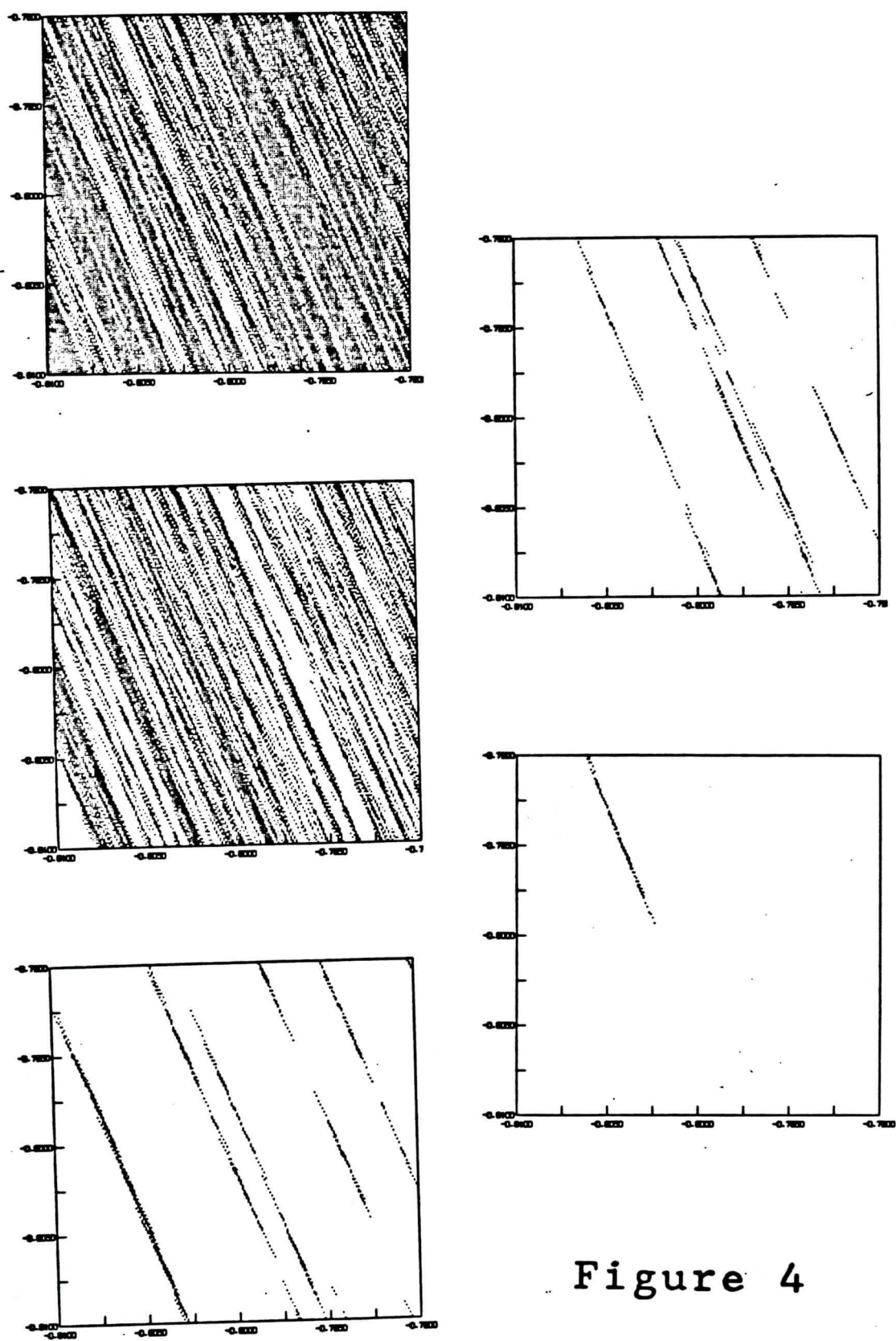


Figure 4

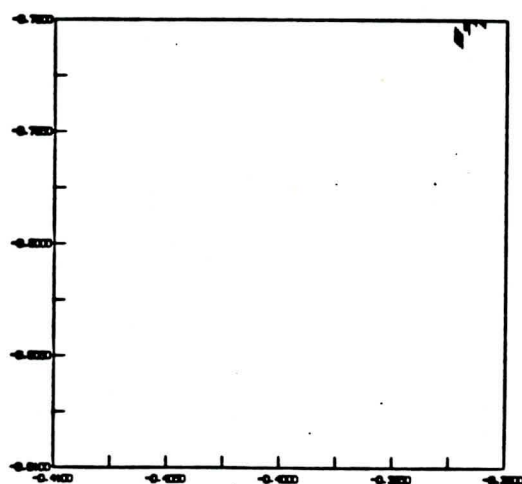
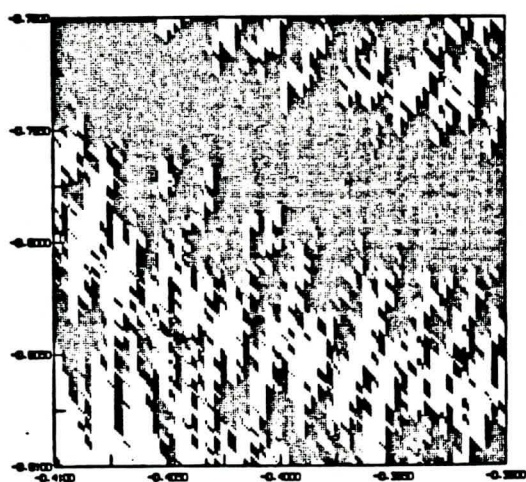
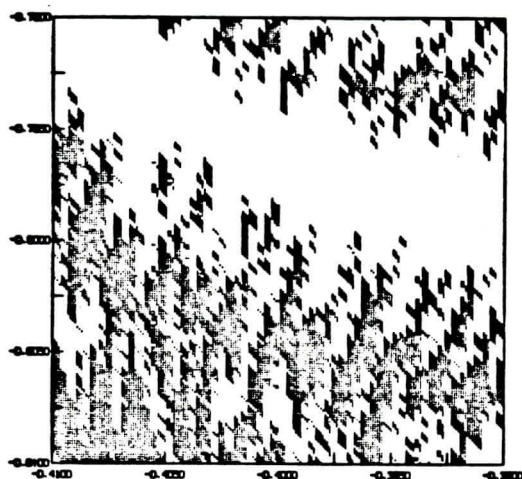


Figure 5

TABLE A1.1

※	Number of Initial Conditions	Number of Orbits	Orbit	Length	Relative Probability
10	508	3	A	7	.95
			B	2	.04
			C	1	.01
50	1352	2	A	15	>.999
			B	1*	<.001
100	55863	2	A	25	.996
			B	1†	.004
101	52831	3	A	9	.998
			B	2	.002
			C	1†	<.001
110	59176	4	A	7	.663
			B	4	.320
			C	13	.016
			D	1†	<.001
1,000	19468	7	A	190	.483
			B	202	.412
			C	102	.092
			D	7	.006
			E	2	.004
			F	10	.003
			G	7	<.001
5,000	85096	7	A	253	.717
			B	341	.129
			C	320	.125
			D	21	.011
			E	13	.006
			F	61	.006
			G	9	.005
10,000	10185	6	A	504	.584
			B	1376	.397
			C	14	.013
			D	58	.005
			E	1†	.001
			F	20	<.001

* fixed point with $x < 0$ † fixed point with $x > 0$

TABLE A1.2

Machine	Representations	Number of Initial Conditions	Number of Orbits	Orbit	Length	Relative Probability
AMDAHL V6 ^a	32 bit	20,000	7	A	4472	0.5490
				B	5690	0.2133
				C	224	0.1595
				D	1432	0.0401
				E	4735	0.0261
				F	816	0.0119
				G	25	0.0001
LSI 11/23 ^b	32 bit	40,000	10	A	8557	0.4421
				B	1073	0.4005
				C	5827	0.1215
				D	3987	0.0343
				E	21	0.0008
				F	770	0.0004
				G	72	0.0001
				H	34	0.0003
				I	1	1 case [†]
				J	156	1 case [†]
IBM PC/XT ^c	32 bit ^{††}	1680	4	A	19229	0.8861
				B	2469	0.0761
				C	2239	0.0338
				D	488	0.0042

[†] an integer in this column indicates the number of initial conditions tested which evolved onto the orbit.

^{††} the 8087 chip computes 60-bit results with rounding, whereas the Amdahl truncates; this finer resolution may account for the lengthier orbit.

^a See Struble (1975) for details.

^b See DEC (1982) for details.

^c See IBM (1984) for details.

Chapter II Analysis of Dimensionality

Introduction

Over the past 30 years, many methods have been developed to estimate the number of active degrees of freedom present in an empirical data set and, thereby, its topological dimension. Early algorithms were advanced by psychologists in order to quantify relationships between psychological observables (see Shepard 1962 a,b). These methods, similar to principal component analysis, suffer from the assumption that there is only linear structure in the data. Shepard and Carrol (1966) introduced a generalization assuming only local linearity (see also Fukunage and Olsen, 1971). From these beginnings various procedures were developed and applied to problems in signal processing and pattern recognition. Operationally, these methods may be separated into two groups. Members of the first class examine eigenvalues of some characteristic matrix such as the covariance matrix, constructed from the entire data set. An example of this type is singular systems analysis (see Bertero and Pike, 1982). Methods of the second class examine the scaling properties of a point set (vectors); an example of this type is the nearest neighbor method of Pettis *et al.* (1979), who also provide a comparison between the two methods. In general, those interested in dynamical systems theory have tended toward methods of the second type. Approaches include the direct application of box-counting ideas discussed in Chapter 1 (Mandelbrot, 1977; Russel *et al.* 1980), examination of the number of near neighbors as a function of distance

(Termonia and Alexandrowicz, 1983), and measurement of the correlation integral (Grassberger and Procaccia, 1983; see also §2.2). Multilinear regression techniques have been applied to determine the local structure of reconstructed attractors by Froehling *et al.* (1981). An additional approach for chaotic systems is to relate the dimension of the trajectory to the Lyapunov exponents of the flow via the Kaplan-Yorke conjecture (Kaplan and Yorke, 1979, see e.g. Lichtenberg and Lieberman, 1983, 394).

Recently, Broomhead and King (1986) proposed a procedure to determine the "qualitative information" contained in experimental data. Based on the theory of singular systems theory (see Pike, 1984), the Broomhead-King procedure estimates several parameters (the embedding dimension, noise level, eigenvectors) of a dynamical system from a time series. The Broomhead-King method is presented in §2.1 and applied to the Moore-Spiegel system. Anomalous results for the embedding dimension imply the method is not reliable in its present form. Through straightforward examples, these results are shown to be consistent with the conjecture that the estimate provided by the Broomhead-King method is related to the dimension of the space that the trajectory explores, not the dimension of the minimum embedding space.

Box-counting procedures that determine the fractal dimension of a data set are difficult to implement (Greenside *et al.*, 1982). An alternative to box-counting which provides a lower bound on the capacity (see §1.1) is the determination of the correlation exponent, ν . The favored method for this calculation is the Grassberger-

Procaccia algorithm (GPA, see Grassberger and Procaccia, 1983a,b). Our version of the GPA is introduced in §2.2 and used to demonstrate the lacunarity oscillation of Cantor sets in §2.3. We note that the presence of this oscillation may result in an incorrect interpretation of the correlation integral's properties.

The Grassberger-Procaccia algorithm has made it possible to calculate scaling exponents for many interesting systems. Implementation of the procedure is straightforward and there is strong enticement to treat it as a black box and process every available time series. In the final sections of this chapter, we demonstrate a lower bound on the amount of data required to unambiguously identify a correlation exponent of a given magnitude and we present supporting analytical arguments. It has been shown previously (Greenside *et al.*, 1982) that this amount increases exponentially with the dimension of the data set; here this growth is quantified and an additional, more restrictive bound is determined. In Figure 2.15 we present a disqualification table, which can assist in determining the significance of a particular result.

The phase space structure of an attractor is the true object of interest in many cases. The coordinates of the phase space trajectory of a numerical solution of a system of ODE's is known; experimentally neither these coordinates nor even the dimension of the phase space is known *a priori*. Packard *et al.* (1980) suggested that the phase space structure of an attractor could be determined from the time series of a single observable by calculating the derivatives of the measured variable and analyzing the trajectory in

this state space. A second approach, less sensitive to operational noise, was discussed by Takens (1981). In Takens' approach, the elements of the reconstruction vectors are consecutive values of the observable separated by a given delay time. Given a time series of measurements equally spaced by some sampling time τ_s

$$x(t_0), x(t_0+\tau_s), x(t_0+2\tau_s), x(t_0+3\tau_s), \dots,$$

the n -dimensional reconstruction consists of vectors

$$X(t) = \{x(t), x(t+\tau_d), \dots, x(t+n\tau_d)\}$$

where the delay time τ_d is an integral multiple of τ_s . Simply put, Taken's theorem states that a complete knowledge of the phase space variables is not required to reconstruct the attractor, that the reconstruction may be done in delay coordinates of a single observable, and that for ideal data, the result is not dependent on the delay time chosen.

For real data, the lack of complete precision places a lower bound on the delay time, measurements separated by less than this bound reflect only the operational noise in the analysis. In general, short delay times tend to project the data onto the diagonal which may then bias dimension calculating algorithms toward low values. When long delays are used, the decay of correlations may result in an overestimate of the scaling exponents. Determining the optimal delay for a given data set is a subject of current research; two methods to automate this choice involve either the first zero of the autocorrelation of the signal or the first minimum of its mutual

information (Fraser and Swinney, 1986). Signals of interest often have oscillating structure in their autocorrelation function and determination of the mutual information is a CPU intensive calculation. For low-dimensional reconstructions, when a signal is reconstructed in three-dimensional space (or four if a section is taken), it is useful to plot the data for several delay times and visually evaluate the constructions obtained.

Several delay reconstructions of the time series of z from the Moore-Spiegel system are shown in Figures 2.1 and 2.2. Each reconstruction is a collection of single points on the plane. The periodic nature of the $T=10.0$ trajectory is evident in Figure 2.1a. The trajectory is one-dimensional; it is self-intersecting in this two-dimensional reconstruction due to projection effects at this choice of delay time. The true phase space trajectory is a one-dimensional limit cycle in a three-dimensional space; a typical reconstruction in three-dimensional space would consist of a one-dimensional curve which did not self-intersect. Figure 2.1b is a similar reconstruction from a chaotic time series obtained with $T = 36.0$. Here, no simple structure is visible; as additional points are included, some sheetlike structures are implied as the figure generally blackens implying that the trajectory has a dimension of at least two. In Figure 2.2, several reconstructions of the $T = 26.0$ case are provided for different choices of delay time.

Once a delay time and an embedding dimension are chosen, the time series becomes a static object which may be treated by the methods of chapter 1, or the Grassberger-Procaccia algorithm. Before applying

these methods, we investigate a recent proposal to estimate the embedding dimension and optimal delay time.

SECTION 2.1 The Broomhead-King Procedure

In this section we review the Broomhead-King procedure and apply the method to the Moore-Spiegel system in several parameter regimes. These examples demonstrate that the method fails to accurately estimate the embedding dimension of the system. The similarity of the eigenvectors generated by a diverse set of examples is noted. These examples include both periodic and strange attractors, and the similarity of the eigenvectors indicates that they do not provide a customized set of coordinates for the reconstruction of attractors.

For band limited data, Broomhead and King examine the structure of the trajectory over time scales less than some window, τ_w , given by

$$\tau_w = \tau^* = 2\pi/\omega^* \quad (2.1)$$

where ω^* is the highest frequency which contains "significant power" in the Fourier spectrum of the time series. In many cases this definition of ω^* is sufficiently ambiguous to hamper the application of the method.

For a given τ_w , the structure of the trajectory within this window is reflected by the eigenvalue spectrum of the $n \times n$ covariance matrix of the trajectory, $E(n)$, given by

$$E(n) = \frac{1}{N} \begin{bmatrix} \sum_{i=1}^N x_i x_i & \sum_{i=1}^N x_i x_{i+1} & \cdots & \sum_{i=1}^N x_i x_{i+n-1} \\ \sum_{i=1}^N x_{i+1} x_i & \sum_{i=1}^N x_{i+1} x_{i+1} & \cdots & \sum_{i=1}^N x_{i+1} x_{i+n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^N x_{i+n-1} x_i & \sum_{i=1}^N x_{i+n-1} x_{i+1} & \cdots & \sum_{i=1}^N x_{i+n-1} x_{i+n-1} \end{bmatrix} \quad (2.2)$$

where

$$x_i = x(t_0 + (i-1) \tau_s)$$

and

$$\tau_s = \frac{1}{n} \tau_w \quad (2.3)$$

The idea here is that the structure of the trajectory at a given scale τ_w is examined in greater detail with increasing n . Once n exceeds the embedding dimension of the trajectory, the eigenvectors of E divide into two sets: those which describe the trajectory and are characterized by "large" eigenvalues, and those which result from uncertainty (either in the observations or the analysis) whose eigenvalues correspond to numerical zero's. If this were the case, the algorithm would provide three useful parameters: an upper bound on the dimension of the system, the noise level, and an optimized set of coordinates (the eigenvectors of E) with which to reconstruct the

attractor. Note that since τ_s varies with n , a different series is used for each value of n . Broomhead and King consider $n = 3, 4, 5, 7, 9, 13$, and 25. Construction of the covariance matrix is simplified somewhat by the choice $n = 2^j$ with $j = 1, 2, 3$, and so on. With these choices for n , each x series may be chosen from the same data set and the full resolution of the time series is used.

The spectrum of Moore-Spiegel Case III, $T=36.0$, is relatively flat for $\omega > 0.128$. Adopting this value as ω^* provides our first example. The normalized eigenvalue spectrum of a 40,000 point time series ($\tau_s = 0.002$) is shown in Figure 2.3. The n eigenvalues, σ_i , are first placed in descending order and then plotted against $\frac{i}{n}$ for $n = 4, 8, 16, 32$ and 64. The eigenvalues for each n are joined by a solid curve to ease identification. In Figure 2.3a, the first four eigenvalues appear to be above the "noise floor." Although the phase space of the Moore-Spiegel system is three-dimensional, this seems a reasonable estimate of the embedding dimension. Figure 2.3b shows the eigenvectors corresponding to the first 12 eigenvalues from a Broomhead-King analysis. When the analysis is repeated using only 30,000 points, the eigenvalue spectrum and first 6 eigenvalues change only slightly. As τ_w is decreased, fewer of the eigenvalues retain significant power; this is to be expected since for τ_w sufficiently small, the change in x will be small when compared to any noise in the system or the analysis; the power will be concentrated in a single eigenvalue corresponding to a constant x eigenvector. It is important to make the "correct" choice of ω^* . As noted by Broomhead and King for the Lorenz case, the eigenvectors resemble orthogonal

polynomials. We have observed eigenvectors with this same structure and ordering for many of the systems we have analyzed. This distinctive structure may assist in identifying the significant eigenvalues; qualitatively, the numerically zero eigenvalues correspond to less ordered looking eigenvectors.

Similar results for the $T=10.0$ series are shown in Figure 2.4. Here however, the trajectory lies on a one-dimensional limit cycle, the periodicity of which is evident from an examination of the Fourier series or the time series itself. The procedure fails to distinguish between this periodic orbit and the chaotic one. The $T=10.0$ time series is not band limited; its periodicity is evident at all scales in the Fourier spectrum. The analysis used a time window determined from the highest frequency containing power in excess of the equivalent white noise spectrum.

To simplify matters, consider the system

$$x(t) = A \sin(\omega_1 t) + B \sin(\omega_2 t) + C \sin(\omega_3 t) + \epsilon \text{Pr}(t)$$

where the function $\text{Pr}(t)$ represents a p-random "noise" term. $\text{Pr}(t)$ is evenly distributed between zero and one, the constant ϵ is used to adjust the strength of the noise. Consider the choice

$$A = 1.0 \quad B = 2.0 \quad C = 1.5$$

with

$$\omega_1 = \frac{4}{3} \omega_2 = \frac{10}{18} \omega_3$$

For $\epsilon = 0$, and τ_w less than 2π divided by the highest frequency above the white noise limit, six large eigenvalues are observed. For nonzero ϵ the signal becomes bandlimited. With $\epsilon = 10^{-2}$ and $\tau_s = 0.04$, the whitened spectrum at high frequencies indicates $\tau^* \approx 0.128$. The singular spectrum and eigenvectors of this case are given in Figure 2.5. We conclude that, as formulated, the method does not provide a reasonable estimate of the embedding dimension; while the rank of Ξ does provide an upper bound on M , it may vastly overestimate M .

Another algorithm advanced to study the intrinsic dimensionality of a data set examines how the number of nearest neighbors scales as a function of separation (Pettis *et al.*, 1979). This method has been applied to dynamical systems by Termonia (1983, see also Gluckenhammer, 1984). We shall not discuss this method other than to note that it is sensitive to noise in the distribution. For low-dimensional objects, an estimate of the noise level is available from a calculation of the correlation integral (Grassberger and Procaccia, 1983 b, also §2.2). As is shown in §2.4, this estimate is no longer available for high dimensional embeddings. Analysis of the covariance matrix, however, may be of use in estimating the noise level and hence the meaningful separation scales to be considered in a nearest neighbor computation.

SECTION 2.2 The Grassberger-Procaccia Algorithm

Box counting algorithms suffer from a host of maladies. One of the most serious is the prohibitively large quantities of computer memory required to partition embedding dimensions greater than three (Greenside *et al.*, 1982). While virtual computer storage has increased the range of these methods (Giorgilli, 1986), following chaotic trajectories on a grid stored in virtual memory has been implicated as a cause of thrashing¹ (N. Habra, personal communication). Secondly, since d_0 depends only on the geometry of the set, determining the fractal dimension requires a knowledge of the low probability features of the set requiring 'long' runs. At some level, the structure of low probability features is of less importance to the physics of the problem than the general structure of the probability distribution. While the information dimension is sensitive to the probability distribution, it also suffers the problems of box-counting; indeed the same difficulty is encountered in the calculation of $f(\alpha)$ as presented by Hasley *et al.* (1986). (An alternative method has recently been advanced by Pawelzik and Schuster, 1987).

To circumvent these difficulties, Grassberger and Procaccia (1983a) introduced the correlation integral

¹ Thrashing is said to occur in computers with virtual storage when the resources expended on overhead far exceed those used for calculations, caused, in this case, by a great deal of disk access.

$$C_2(\ell) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \theta(\ell - |\mathbf{x}_{ij}|) \quad (2.5)$$

where N is the number of points in the set, $\mathbf{x}_{ij}$ is the vector separating the i^{th} and j^{th} points, and θ is the Heavyside function. For fractal sets Grassberger and Procaccia (1983a,b) show

$$C_2(\ell) \sim \ell^\nu \quad (2.6)$$

where ν is called the correlation exponent. As it turns out, ν is equal to the generalized dimension d_2 (Hentschel and Procaccia, 1983, see also §1.3), and therefore provides a lower bound for the fractal dimension. Since the correlation exponent is sensitive to the probability distribution on the object; high probability regions are more heavily weighted than those of low probability and a good estimate of ν may be obtained without knowledge of the fine structure of the latter.

The direct evaluation of equation (1.8) for a many point approximation of the set in question provides an accurate method for determining ν . This is demonstrated in Figure 2.6 for the middle thirds Cantor set. An alternative method restricting the sum over j to a smaller number of "randomly" chosen base points, increases the accessible N for given computer resources thereby extending the range of length scales over which $C_2(\ell)$ may be calculated, but losing its fine structure.

A brief comment on the optimal implementation of the above methods is in order. As noted in Appendix A of Grassberger and Procaccia

(1983), execution time is significantly reduced when the logarithm is approximated by reading the exponent directly from memory. Computers generally represent real numbers in a hexadecimal format with exponent and mantissa juxtaposed within one or more words (for details of IBM machine architecture see Struble, 1975). The trick is to approximate the base-16 logarithm by the exponent alone, ignoring the contribution of the mantissa. While this does save time, its cost in resolution is great. We introduce an alternative method, binning the separation data in a (very large) integer array. The latter method yields very high resolution at the largest scales, revealing details of the fine structure of the oscillatory component (see §1.2) to be discussed in the following section. We have also investigated a method which reads the exponent and the first few bits of the mantissa and then estimates the logarithm using a look-up table for the contribution of the mantissa. The execution time required here was approximately a factor of three less than that required to evaluate the logarithm. The fastest method we have found is to use integer binning in conjunction with the Euclidian sup norm (the maximum component of the separation vector). The optimization problem is important in that the algorithm is presently limited by the amount of data that can usefully be considered.

Procedures that evaluate ν via Equations 2.5 and 2.6 are generically referred to as the Grassberger Procaccia Algorithm (GPA). The principle difficulty of the GPA is that it is an N^2 algorithm; the number of calculations required to determine $C_2(\ell)$ increases as the square of the number of data points. Coupled with the

exponential increase in the amount of data required to describe a surface for increasing dimension, reliable estimates of dimension via GPA are restricted to sets of small ν ($\nu \lesssim 5$). This limitation is quantified in §2.5. In Chapter 3 we suggest an algorithm whose computational time consumption scales as the number of points in the set and apply it to several examples.

Scaling representative of an attractor may also be biased in the reconstruction if the orbit is sampled too frequently; doing so results in a large contribution to $C_2(\ell)$ from dynamically correlated points (exaggerating the one-dimensional nature of small segments of the trajectory). The effect has been investigated by Theiler (1986).

Several things are done in the next 3 sections: we illustrate the GPA on Cantor sets where the embedding dimension is known *a priori*. The oscillatory component of scaling in these sets arises in a natural way. Next, additional effects which arise when analyzing real data are discussed for a simple model and the Moore-Spiegel system. Finally, arguments are presented which quantify the amount of data required to use the method.

SECTION 2.3: Evaluating Correlation Exponents of Simple Fractal Sets

In this section the GPA is applied to fractal distributions where the embedding dimension is known. This is the case for a physical distribution (a coastline or the large scale distribution of visible matter in the Universe) or the solution to a system of ODE's. Our first examples will be strictly self-similar Cantor sets such as the

middle thirds set discussed in Chapter 1. Cantor sets are the simplest self-similar sets with *noninteger* dimension. A strictly self-similar Cantor set may be approximated to arbitrarily high resolution in the following manner:

- 1) divide the unit interval in R^1 into s equal subintervals
- 2) delete all but r of the subintervals
- 3) record the endpoints of each subinterval as the members of this approximation to the set (without double counting)
- 4) iterate this procedure on each of the remaining subintervals until the desired accuracy is achieved.

After N iterations of this procedure, the approximation consists of $2r^N$ points². A set constructed in this way may be designated by an s -digit binary number with r ones and $s-r$ zeros, the zeros being located at the positions where segments were deleted. The middle thirds set is then designated as 101.

In these cases a great deal of information is available from the outset. The set may be embedded in 1 dimension; for the N th order approximation, the ratio of the outer cutoff to the inner cutoff is r^N and any operational uncertainty may be kept below this threshold. The specific cases described here are the sets 101, 101010001 and 101001001, which will be called sets I, II, and III respectively. All three sets are homogeneous fractals of dimension

$$d_q = d_{HB} = d_2 = \frac{\log 2}{\log 3}$$

²The total number of points will be less than $2r^N$ if any two segments are contiguous; we have also assumed that the first and final segments are retained.

for all q .

A graph of $\log C_2(\ell)$ versus $\log \ell$ for set I is shown in Figure 2.7. The oscillatory component of $C_2(\ell)$ for each of the three sets is evident. To examine it more closely we approximate $\psi(\ell)$ as

$$\psi(\ell) \approx \nu \log_2 \ell - \log_2 C_2(\ell). \quad (2.7)$$

The results for each of the three sets are shown in Figures 2.8 and 2.9. Computed estimates of the correlation exponent and period of $\psi(\ell)$ for each set are provided in the figure. They are determined by a least square fit to a range of data that is a multiple of $\log \sigma$. The oscillations reflect the manner in which the structure of the set changes over a range of $\log \sigma$. For set I, $r=3$ and the change is regular. Sets II and III are indistinguishable from set I by their d_q , but are easily distinguished from set I and each other by the structure of $\psi(\ell)$; note especially that the asymmetry of set II is reflected in Figure 2.9 b.

The origin of the oscillation in strictly self-similar sets may be easily understood. Using set I as an example, for every structure (e.g. point pair of separation) observed at scale ℓ , exactly two such structures will be observed at scale $\ell/3$, hence the details of the scaling structure which form the oscillations are strictly periodic. Appendix 2 demonstrates that these oscillations are strictly log-periodic and provides a handy construction mechanism for the study of the residual. The oscillation itself has structure on all scales and appears self-similar. More correctly, it is self-affine in that there is an arbitrary (unit dependent) factor that relates $\psi(\ell)$ to ℓ .

The fine structure of $\chi(\ell)$ depends on the statistical quantity

measured. In these cases, where it is determined from $C_2(\ell)$ for sets embedded in one dimension, χ reveals the lacunarity of the underlying set; a large range of χ with slope $-d_2$ indicates the absence of pairs of points which are separated by a distance in this range, or equivalently, the presence of lacuna of scale ℓ in the underlying set. For inhomogeneous fractals, the lacunarity interpretation of the $C_2(\ell)$ oscillation is less useful at large scales. An alternative is discussed in Chapter 5.

Guckenheimer (1984) noted the departure from linear scaling caused by the oscillations and questioned whether this "logarithmic correction" might limit the accuracy with which dimensions may be calculated. In Chapter 3, we take advantage of the regularity of the oscillation to increase the accuracy of such calculations.

Consider the Poincaré sections of the T=36.0 Moore-Spiegel system (Figures 1.8); each section contains approximately 10^4 points. The correlation integral of these data are shown in Figure 2.10 with $\psi(\ell)$ below. While there is an indication of an oscillation, the data are insufficient to identify it unambiguously. In these cases, as in the dissipative Hénon map (Equation 1.19) the oscillations are damped at large scales due to the interplay of many regions. The important point here is that deviation from linear scaling is expected; the residual of a linear least squares fit of $\log C_2(\ell)$ to $\log \ell$ is a poor measure of the quality of the scaling.

The best fit values in this case are

$$\begin{array}{lll} \nu_x & \approx & 1.08 \quad (10^{-1.5} < \ell < 1) \\ \nu_v & \approx & 1.10 \quad (10^{-2.5} < \ell < 1) \end{array}$$

where the subscript denotes the section taken: $x \equiv$ position and $v \equiv$ velocity . Since these sections come from the same attractor, their correlation exponents should be equal.

We have looked for the $\psi(\ell)$ function for several scaling distributions found in nature, generally without success. These are described in Chapter 4. An exception is found in the statistics of fluid turbulence. Here the oscillatory component of the scaling can be established in the cascade models (Novikov, 1964; Smith *et al.* 1986); it has been suggested (Smith *et al.*, 1986) that this oscillation is compatible with recent observations of the velocity structure functions of turbulence (Anselmet *et al.*, 1983). Similarly, velocity reconstructions of some turbulence data show the oscillations (Van de Water, 1987). The remainder of this chapter focuses on the difficulties of applying the GPA to more realistic data sets.

Section 2.4: Scaling limits in a Nonlacunar Set

In this section we address the question: given an ideal data set of unknown geometry, what is the minimum amount of data required to determine the correlation exponent of the set to within a desired degree of accuracy using the Grassberger Procaccia algorithm. The discussion centers upon the GPA determination of ν in sets for which no oscillation is visible. As mentioned above, this currently includes most attractors reconstructed from experimental time series. The "size" of a data set will be a central issue throughout this

section and is defined by the length of the time series and the optimal delay time, not the sampling rate. Thus the size will reflect the amount of information contained in the set.

Truly nonlacunar attractors do exist; a simple example is an attracting two torus. If the torus is covered densely and uniformly the $\psi(\ell)$ function of the correlation integral will tend to zero everywhere. It is this special case which is sometimes referred to as pure scaling; we will call this behavior linear scaling. In such a case, the prefactor A is a constant; and for $\ell \ll \ell_0$

$$\begin{aligned}\nu_s(\ell) &= \frac{d \log C_2(\ell)}{d \log \ell} \\ &= d(\nu \log \ell + \log A)/d \log \ell \\ &= \nu\end{aligned}\tag{2.8}$$

where the subscript s emphasizes that we are not dealing with the limiting behavior. When working with digitized data, the derivative must be evaluated carefully.

Consider the general case of a time series embedded in an N-dimensional space. The dimension of the record is approximated by calculating the correlation exponent of the embedding. If the orbit of this system in its (unknown) phase space lies on an M-dimensional manifold and M is greater than N, the projection into E^N will be locally space filling. When this procedure is repeated for increasing N, the conventional expectation is that for small N (that is $N < M$), $\nu_s \approx N$ while once N exceeds M, ν_s approaches a constant value assumed to be the correlation exponent of the underlying set. This approach works well in many instances and its use has been suggested in the case of small data sets (Abrahamson et al., 1985).

Our first example, perhaps the simplest dynamical system³, is

$$\dot{x} = c \quad (2.9)$$

where c is a real constant. The system corresponds to a particle in steady motion along a linear path. When the motion is restricted to a finite interval (by, say, periodic boundary conditions), the trajectory consists of a line segment. The width of the well is taken as the unit of length and the sampling rate is chosen to cover the interval uniformly. Noise is introduced to system 2.9 by adding p -random numbers evenly distributed between 0 and ϵ to each data point.

The resulting time series is embedded in E^2 and $C_2(\ell)$ is calculated via the GPA; a graph of $\log C_2(\ell)$ vs $\log \ell$ is shown in Figure 2.11. The expected $N < M$ behavior of $\nu_s(\ell)$ is shown in Figure 2.12 where four distinct regions are visible. For separations less than the nearest neighbor distance ($\log \ell < -11$), the set scales as a collection of individual points and $\nu_s(\ell)$ is equal to zero. Similarly for ℓ greater than the diameter of the set ($\log \ell > 1.0$), the entire distribution will have been included, $C_2(\ell) = 1$, and $\nu_s(\ell)$ is again zero. For sufficiently long, sufficiently clean data records reconstructed with good delay times, two additional regions are found between these limiting behaviors. The first ($\log \ell \approx -7$), whose existence is due to noise or uncertainty in the measurements, will range from the nearest neighbor distance up to the magnitude of

³The alternative choice of simple harmonic motion yields a similar result, but the scaling structure is complicated either by projection effects ($N=1$) or modifications due to large scale structure ($N=2$).

the noise. (For white noise the upper limit of the noise range is this magnitude multiplied by the square root of the embedding dimension, corresponding to the diameter of a hypercube.) By definition, random fluctuations must not lie on a finite dimensional surface (if this was the case, they would not be independent) and are therefore expected to appear space filling in any embedding dimension. As a result, one expects $\nu_s(\ell)$ to be equal to M in this region. The region of greatest interest lies above the noise region. Here ($\log \ell \approx -3$) the value of the correlation integral is governed by the object and, for nonlacunar objects, $\nu_s(\ell) \approx \nu$. We reiterate that for lacunar objects, $\psi(\ell)$ is not zero and $\nu_s(\ell)$ need not equal ν .

$\nu_s(\ell)$ is estimated by determining the slope of the least squares fit over a window of width τ centered on $\log \ell$, thus τ represents the logarithm of the ratio of the largest ℓ scale to the smallest. The choice of τ depends on the uniformity of the data set; for completely nonlacunar data such as the cat map (see below), we have found $\tau = 2$ to yield good results. Within each power of τ , $\nu_s(\ell)$ is evaluated a number of times. These running averages are displayed in Figure 2.13 as "+" marks.

The minimum window size, $\tau_{\min}$, corresponds to the finest scale at which the value of $C_2(\ell)$ is determined. The disconnected points in Figure 2.13 represent $\nu_s(\ell)$ as calculated by this separation. Generally, these points fall about the curve computed with $\tau=2$. The major deviations which occur at small scales result from the finite amount of data and the numerical resolution. For values of ℓ comparable with the nearest neighbor separation, the quantization

inherent in the correlation integral is observable. In this region, a small increment of ℓ introduces only a few additional separations into the sum; this results in discrete values in the change of $\log C_2(\ell)$ for each increment in $\log \ell$. Specifically

$$\delta_{ij} = \frac{\Delta \log C_2(\ell)}{\Delta \log \ell} = \frac{\log i - \log j}{N^2} \quad (2.10)$$

where

$$i = j, j+1, j+2, \dots$$

and

$$j = 1, 2, 3, \dots$$

For each value of j , lines of δ_{ij} are observable in Figure 2.13. As ℓ increases j increases, and thus δ_{ij} decreases for a given value of i . For ℓ values in the scaling region, δ_{ij} is small and the pointwise value of $\nu_s(\ell)$ is in close agreement with the $\tau=2$ value.

Rarely is the scaling region so distinct as it appears in Figure 2.12. Even when the system is ergodic and two-dimensional, the deviation of $\log C_2(\ell)$ from linearity may be large (an example is the standard map, see Appendix 5B). If $\nu_s(\ell)$ is to provide a good estimate of ν , τ must be enlarged or this method should be abandoned. On the other hand, though increasing the window reduces the fine structure in $\nu_s(\ell)$, it may also obscure the distinction between the noise region and the scaling region. Unless the noise level can be reduced significantly, the collection of additional data will not alleviate this problem, and this method will not yield a good estimate of ν .

When $\nu_s(\ell)$ has fine structure and the window size is large, the transition between the noise region and the scaling region becomes

indistinct. In the case of small data sets, it is crucial to determine which behavior is due to the data and which is due to the parameters of the data set (eg. the number of points). The large ℓ transition is investigated in the next section. Here we concentrate on distinguishing the scaling region from the noise. The point is that a truncated white noise signal will often appear to be scaling with an exponent less than E . Look again at the noise plateau of Figure 2.12. The mean value of $\nu_s(\ell)$ on the plateau is

$$\nu = 1.72 \quad \left\{ \begin{array}{l} \tau = 2.0 \\ \log \ell = -6.8 \end{array} \right.$$

The data is embedded in a two-dimensional space, yet the noise plateau is significantly less than two. Either the noise is not space filling or the $\nu_s(\ell)$ calculation is faulty. The dilemma is resolved by increasing the magnitude of the noise term until it dominates. The plateau value of $\nu_s(\ell)$ is then computed as a function of the size of the data set for several low dimensional embedding spaces. As a result we are able to quantify when a data set is "small" as a function of the embedding space dimension.

These results are summarized in Figure 2.14, which shows the $\tau=2$ plateau of $\nu_s(\ell)$ for several data set sizes. The data consists of p -random numbers generated by the additive congruent method⁴ (see Carnahan et al. 1969) evenly distributed on the interval $[0,1]$. The series is then embedded in three-dimensional space. In the figure,

⁴A delay of at least 2 was used for all embeddings with this generator as fine structure in $\nu_s(\ell)$ was found when consecutive values were taken.

$\nu_s(\ell)$ is plotted against $\log \ell$; the number of points in the set increases by a factor of two between each realization. For small N the plateau is poorly defined and estimated by the greatest $\tau = 2$ value of $\nu_s(\ell)$ - a procedure which overestimates the value of $\nu_s(\ell)$. For a given embedding dimension E , we define a small data set as one in which the value of $\nu_s(\lambda)$ for p -random data is less than $0.90 E$.⁵ Thus any data set with $N \leq 2^{12}$ is small when embedded in 3 dimensions. It is not our claim that a data set of this size is sufficient, only that with a smaller set one cannot distinguish the dynamical system in question from white noise. For example, the departure of the mean slope of the correlation integral computed from a 1000 point data set from $E = 3$ by 15% is not evidence of an underlying low-dimensional structure.

A more common approach with experimental results is to embed a series of a given length into successively higher dimensions and to interpret the deviation of ν_s from E as a sign of fractal structure. The observations above apply even more strongly for large E values and sets reconstructed from a time series; the relevant comparison is not to E but to ν_s for a random data set of the same length. To ease this comparison, we have amalgamated the results of a variety of p -random computations to produce Table 2.1. Here the value of maximum $\nu_s(\ell)$ is calculated as a function of N and E . For large enough E (for any N), the plateau narrows; the $\tau = 2$ and $\tau = 10$ maximum $\nu_s(\ell)$ values are given. Note this comparison with white

⁵Note that with the 10% limit it is not possible to distinguish the geometry of a smooth limit cycle from that of a fractal attractor for dimension exceeding ten.

noise is most reasonable when there is no correlation between successive point in the reconstruction; comparison with an appropriate red noise signal would yield a stricter, more useful bound.

Section 2.5: Edge effects as Strict Limits to Dimension Calculation

Given that there seem to be serious difficulties implementing the GPA with small data sets, we present an analytic argument to support this numerical result. First consider a slight modification of an argument due to Alan Wolf (personal communication, 1986; see also Wolf, 1984). In order to determine a scaling exponent one requires a certain range over which the data approximate some power law. Denoting the upper and lower limits of this region by ϵ_{hi} and ϵ_{low} , the requirement is then

$$\frac{\epsilon_{hi}}{\epsilon_{low}} \geq R \quad (2.11)$$

where R (the range of scales) remains to be chosen. The largest conceivable choice for ϵ_{hi} is the diameter of the set. Considering an M -dimensional unit cube,

$$\epsilon_{hi} = \sqrt{M} \quad (2.12)$$

With ϵ_{low} determined through Equation 2.11, a lower limit on the number of points may be determined by assuming a uniform density at this separation. Then

$$\epsilon_{low} = N^{-1/M} \quad (2.13).$$

Solving for N

$$N = \epsilon_{\text{low}}^M = (R/\epsilon_{hi})^M = (R/\sqrt{M})^M.$$

When present, the oscillations place a lower bound on R; R must be an integer multiple of the period of the oscillation. In practice Wolf found a Q of 30 to be required; it was with the hope of lowering this data limit that Wolf examined the Lyapunov exponents of dynamical systems. At present, it appears a similar limit may be required in the calculation of these exponents as well.

We now demonstrate a stronger bound on ϵ_{hi} , resulting in an increase of $N(M)$, due to the equal weighing of points near the surface of the object. Such an effect was predicted by Guckenheimer (1984). Consider a one-dimensional unit cube made up of points uniformly distributed on the unit interval⁶ as in the system given by Equation 2.9. The correlation integral is

$$C_2(\ell) = 2\ell - \ell^2 \quad 0 \leq \ell \leq 1 \quad (2.15)$$

Due to the complete uniformity of this example, the set is nonlacunar; thus no oscillatory corrections are present and one may approximate ν by $\nu_s(\ell)$. Substituting Equation 2.15 into Equation 2.8 we have

$$\nu_s(\ell) = (1 - \ell/2 + O(\ell^2)) \quad \text{for } \ell < 1. \quad (2.16)$$

As ℓ approaches one, ν_s becomes a poor estimator of ν . This is intuitive, as ℓ approaches one, more and more of the points lie

⁶ This argument followed from a discussion with J. Theiler during the Los Alamos Summer School in Nonlinear Science, 1986

within ℓ of both endpoints) and scale as isolated points; for ℓ greater than one, $C_2(\ell)$ is a constant. For ν_s to be accurate to within 10% requires $\ell \leq 0.1$.

For an M-dimensional hypercube

$$C_2(\ell) = \ell^M (2 - \ell) \quad (2.17)$$

and

$$\nu_s(\ell) = M (1 - \ell/2 + O(\ell^2)) \quad (2.18)$$

The argument may be made exact by using the Euclidean sup (ℓ_∞) norm. For a uniformly filled hypercube, $C_2(\ell)$ is equal to the probability that two randomly chosen elements will be separated by a distance less than ℓ . In the one dimensional case (the unit interval), this probability is equal to the fraction of the area of the unit square which is bounded by the lines $x - y = \ell$ and $y - x = \ell$. That is

$$P(|x-y| < \ell) = \ell (2 - \ell) \quad \text{for } 0 \leq \ell \leq 1.$$

This is equivalent to Equation 2.15. For an M dimensional cube, the probability that two points are separated by a distance less than ℓ is equal to the probability that each of the elements of the separation vector is less than ℓ . Since these are independent for points chosen at random

$$C_2(\ell) = (\ell (2 - \ell))^M \quad (2.19)$$

and

$$\nu_s(\ell) = M [1 - \ell/(2 - \ell)]. \quad (2.20)$$

Note that $\nu_s(\ell)$ always underestimates ν and define a quality factor, Q , such that $\nu_s(\ell) > Q \nu$. Using Equation 2.20 to define ϵ_{h_1} for a

given value of Q and the minimum number of uniformly distributed points requires as a function of R and Q is given by

$$N = \left[\frac{R(2-Q)}{2(1-Q)} \right]^M \quad (2.21)$$

In particular, assuming $R = 4$ is sufficient to determine a value of ν_s within 5% of ν yields

$$N = 42^D$$

Even with this conservative value of R , the number of points required increases so rapidly as to make application of the Grassberger-Procaccia algorithm impractical objects with dimensions greater than approximately 5.

The large scale transition is examined using Arnold's cat map [Arnold and Alvez, 1968]. The cat map is a two-dimensional, area preserving iterative map of the unit square onto itself. Given by

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix} \pmod{1}$$

it is both ergodic and mixing. Consider the embedding of the series $x_1, x_2, x_3, \dots$ with delay 1 in two-dimensions⁷. In the range

⁷embedding this series in higher dimensional spaces results in a sheet like structure which obscures the boundary effects to be demonstrated here; the question as to whether similar effects might occur in experimental data is open.

$(.01 < \ell < 1)$ $\nu_s(\ell)$ as calculated by Equation 2.18 agrees well with the $r=2$ value of ν determined from a 2^{14} point data set.

We emphasize again that we have used simple objects - cubes of uniform density. Indeed, hyperspheres are the least sensitive objects to this edge effect. With complicated objects the effect should be more pronounced. There exist certain density distributions which minimize this effect. Objects of this category have been constructed and studied by Theiler (1986); however, to assume their existence in any given experimental setting is unwarranted.

Finally note that the minimum number of points given by Equation 2.21 assumes a uniform covering of the entire set. For data from a time series, this quantity is related to the ratio of the length of the series to the reconstruction time, not the sampling rate. At some point, increasing the sampling rate does not effect the size of the data set.

Appendix 2: Explicit construction of the Cantor Oscillation

The simple structure of Cantor sets permits a straightforward recursive construction of $C_2(\ell)$, and thereby $\chi(\ell)$, to arbitrarily high accuracy. For concreteness we will examine the 101 set. The argument is easily extended. Consider the most developed oscillation at every finite approximation to the set, that is $C_2(\ell)$ over the interval $1/\sigma \leq \ell \leq 1$.

$C_2(\ell)$ is completely specified by the location of its discontinuities, x_i , and their strengths, w_i . The function is generated in a manner which parallels the method for constructing Cantor sets (see §1.2). The following notation allows us to deal with two sets of integers, simplifying and speeding numerical implementation of the method. The N^{th} approximation of the set will consist of $n = 2^{N+1}$ points and there will be a discontinuous increase in $C_2(\ell)$ at the 3^N values of ℓ given by

$$N_{x_i} = \frac{\alpha_i}{3^N}$$

and the magnitude of the i^{th} discontinuity will be

$$N_{w_i} = \frac{\beta_i}{2^{N+1}}$$

for $i = 0, 1, 2, \dots, 3^N$.

The problem is now reduced to finding the two sets of integers α_i and β_i . The correlation integral may then be evaluated by calculating

$$C_2(\ell) = \frac{1}{N^2} \sum' \beta_i$$

where Σ' is a sum over all i such that

$$N_{x_i} \leq \ell$$

or, equivalently

$$N_{\alpha_i} \leq 3^N \ell .$$

For $N = 1$, the four point approximation to the set yields

$$\begin{aligned} {}^1\alpha_0 &= 0 \\ {}^1\alpha_i &= i \\ {}^1b_0 &= 4 \\ {}^1b_i &= 2(4 - i). \end{aligned} \quad i = 1, 2, 3$$

At each generation, there exists at least one pair of points separated at every integer multiple of the smallest separation. Thus

$$N_{\alpha_i} = i. \quad i = 0, 1, \dots, 3^N$$

The N_{β_i} are slightly more complicated. The destruction operator may be thought of as introducing fine scale structures in the $N+1^{\text{st}}$ generation with twice the weight of the corresponding structure of the N^{th} generation on $(0 \leq \ell \leq 1)$. If the N^{th} generation is considered as a 3^N digit binary integer with 1's corresponding to points, the N^{th} generation structure in the range $1/3 \leq \ell \leq 1$, is

equal to the number of 1 bits remaining after a left shift of α_i bits and logical AND operation on the binary representation of the set; this results in a simple growth and decay process of the $N + 1^{\text{st}}$ generation β 's in this range. At the point $\ell = 1/3$, both processes contribute. In symbols

$${}^{N+1}\beta_i = \begin{cases} 2^N \beta_i & i = 0, 1, \dots, k-1 \\ 2^N \beta_i + 2 & i = k \\ {}^N\beta_{2k-i} & i = k+1, k+2, \dots, 2k-1 \\ 2^N \beta_0 & i = 2k \\ {}^N\beta_{i-2k} & i = 2k+1, 2k+2, \dots, 3^N \end{cases}$$

where $k = 3^{N-1}$.

FIGURE CAPTIONS

Figure	Caption
2.1	Delay reconstruction of $z(t)$ from the Moore-Spiegel system with the parameter values $R=100.0$ and (a) $T=10.0$, (b) $T=36.0$. In both cases $\tau_d = 0.20$.
2.2	The effect of varying the delay time is shown in reconstructions of a single time series of z from the Moore-Spiegel system in the case $R=100.0$ and $T=26.0$. The delay time increases with each panel, specifically (a) $\tau_d=0.20$, (b) $\tau_d=0.40$, (c) $\tau_d=0.80$, (d) $\tau_d=1.60$.
2.3	The (a) normalized eigenvalues and (b) corresponding eigenvectors generated by the Broomhead-King procedure applied to a time series from the Moore-Spiegel system in the case $R=100.0$ and $T=36.0$. In (b) the eigenvalue corresponding to each eigenvector is printed above each panel.
2.4	As in Figure 2.3 for the case $T=10.0$. Here the attractor is periodic.
2.5	As in Figure 2.3 for the system of three period harmonic oscillations with noise described in the text.
2.6	A linear plot of the correlation integral as a function of ℓ for the set 101.
2.7	The graph of $\log C_2(\ell)$ as a function of $\log(\lambda)$ for set 101.
2.8	$\psi(\ell)$ for the set 101.
2.9	$\psi(\ell)$ for the sets (a) 101010001 and (b) 10100101.
2.10	The (a) correlation integral and (b) $\psi(\ell)$ for a surface of section of the Moore-Spiegel system.
2.11	The correlation integral for a one-dimensional object.
2.12	$\nu_s(\ell)$ for the set of Figure 2.11.
2.13	$\nu_s(\ell)$ as in Figure 2.12. Here the δ_{ij} structure is shown.
2.14	$\nu_s(\ell)$ plateaus for an increasing numbers of points. The individual curves are labeled by the number of points in

the source data set. Note the slow increase in the maximum value.

Table 2.1 Disqualification table showing the results of embedding white p-random noise.

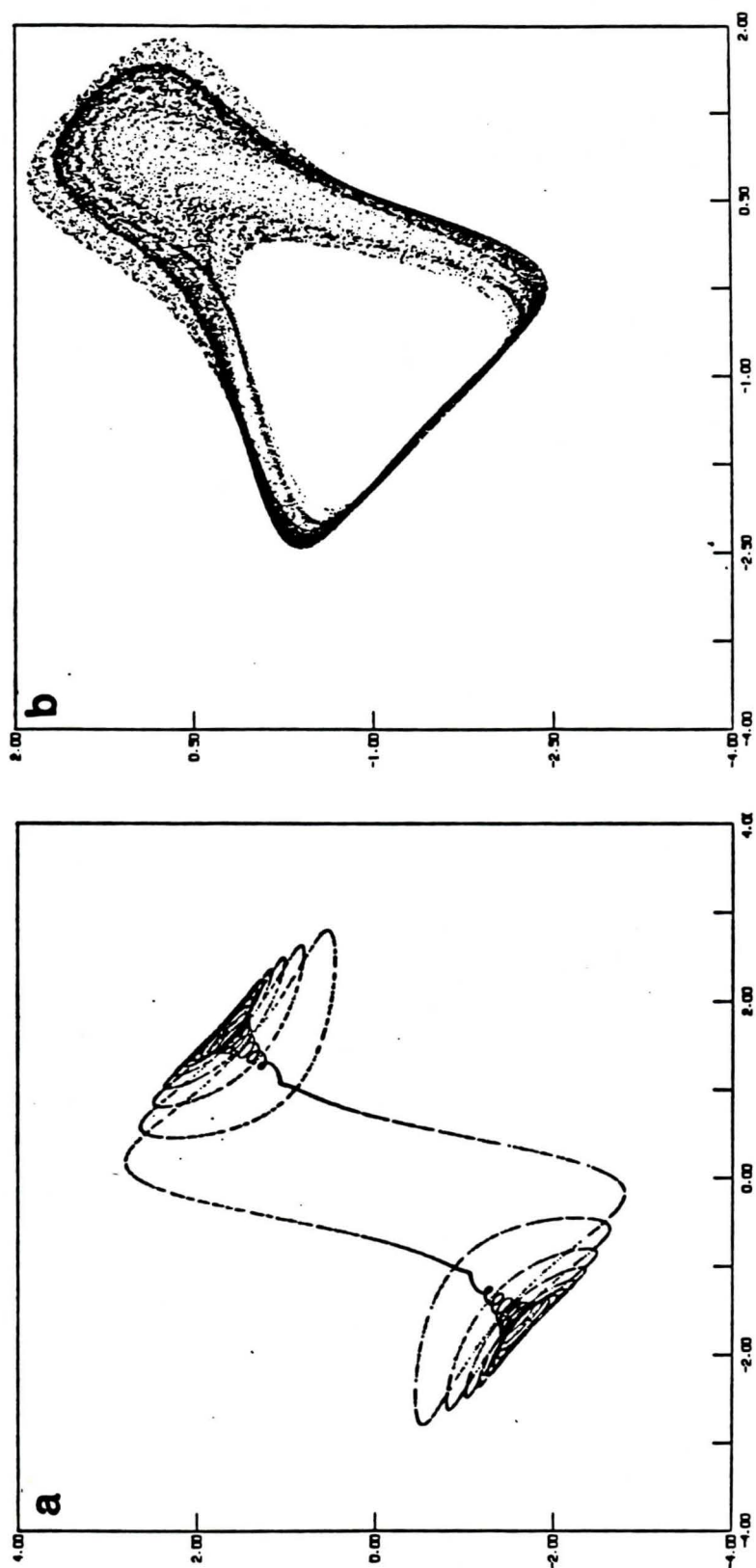


Figure 1

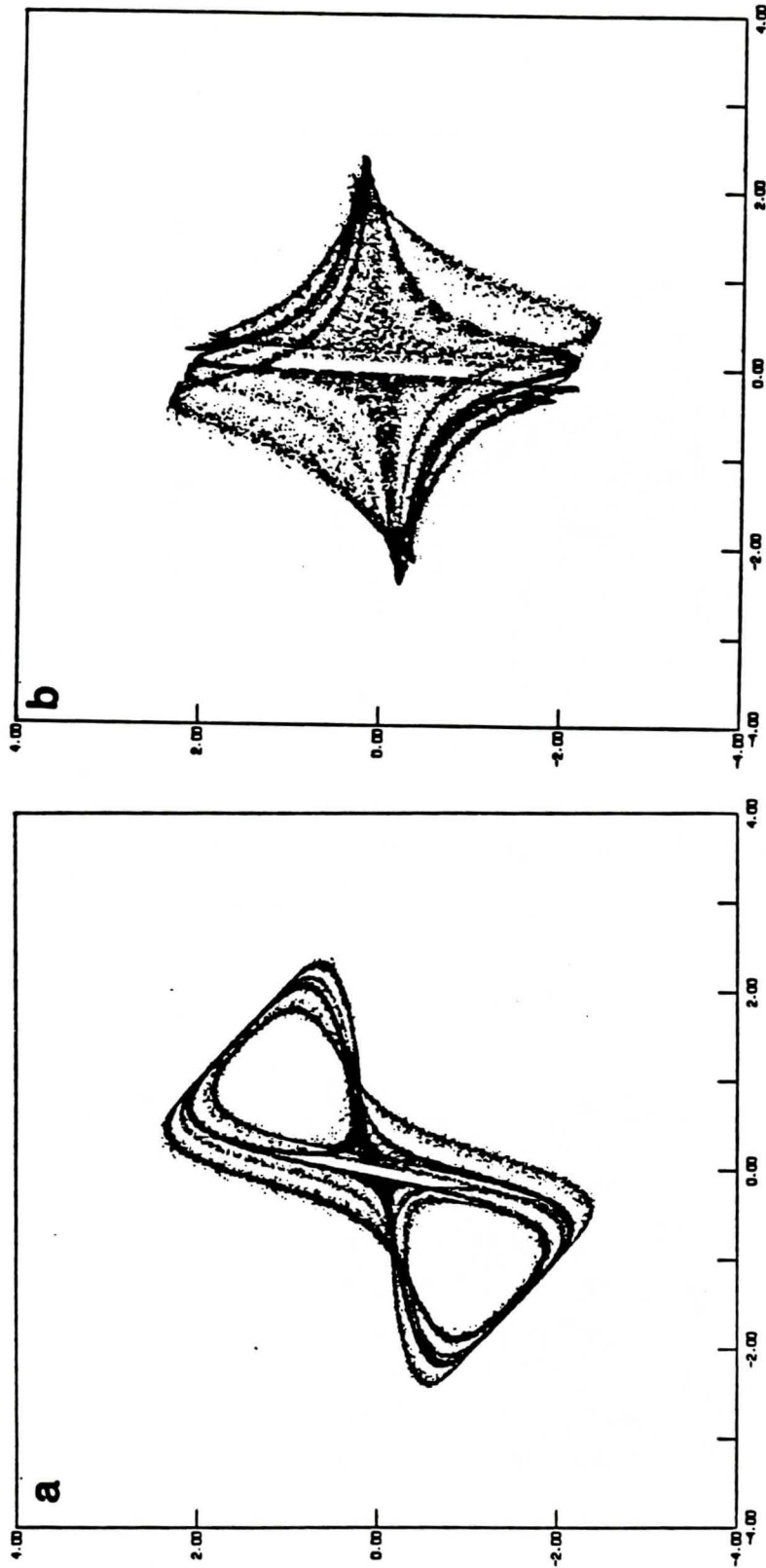


Figure 2

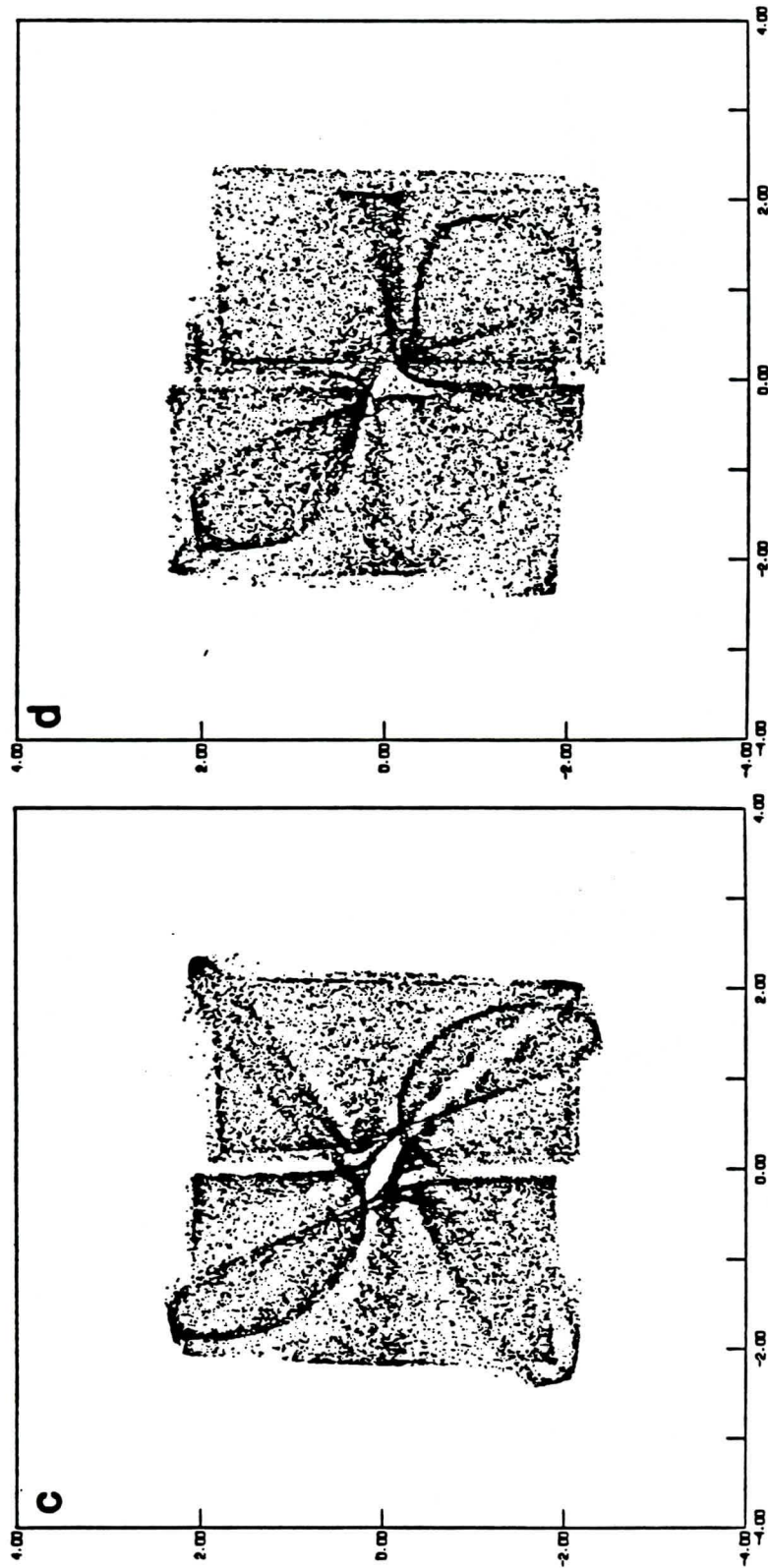


Figure 2

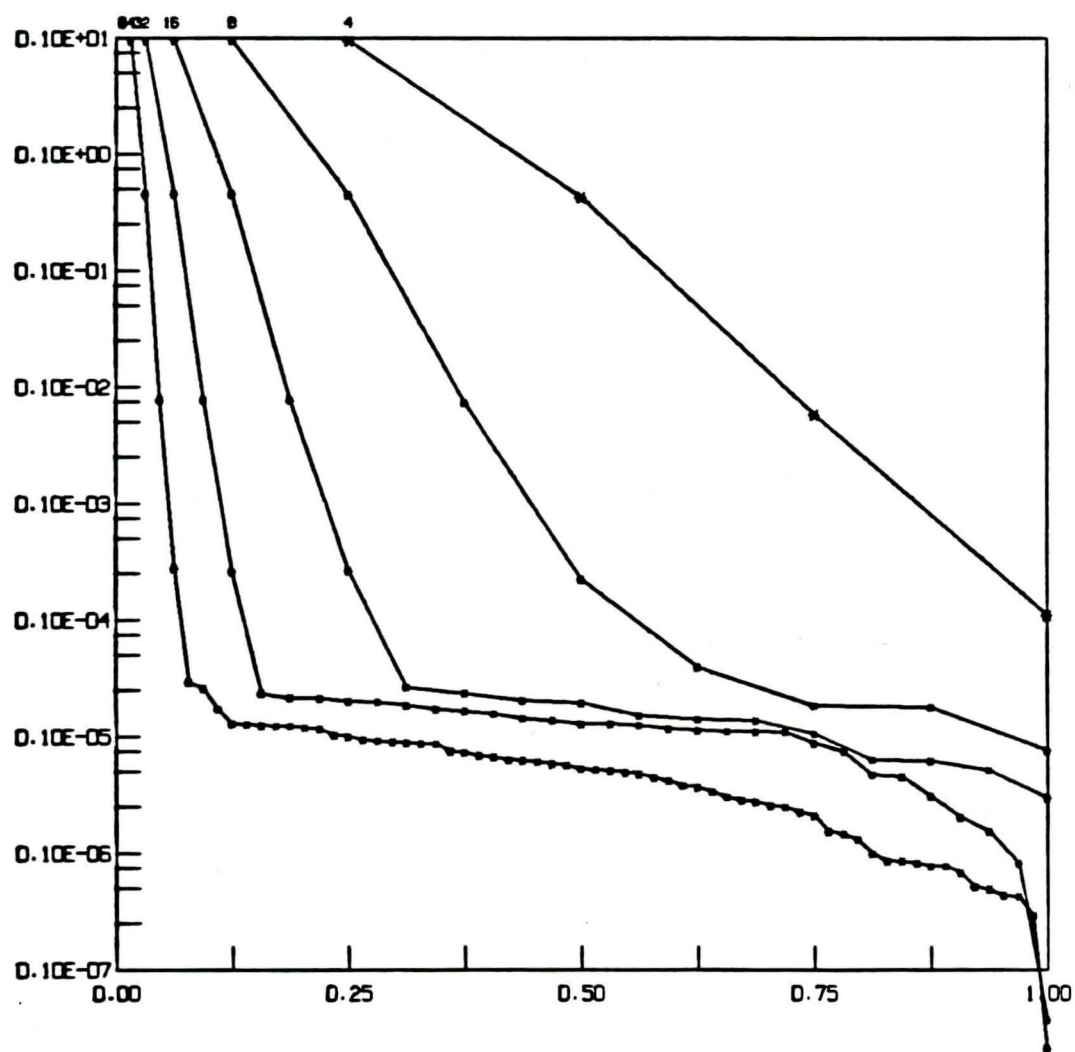


Figure 3a

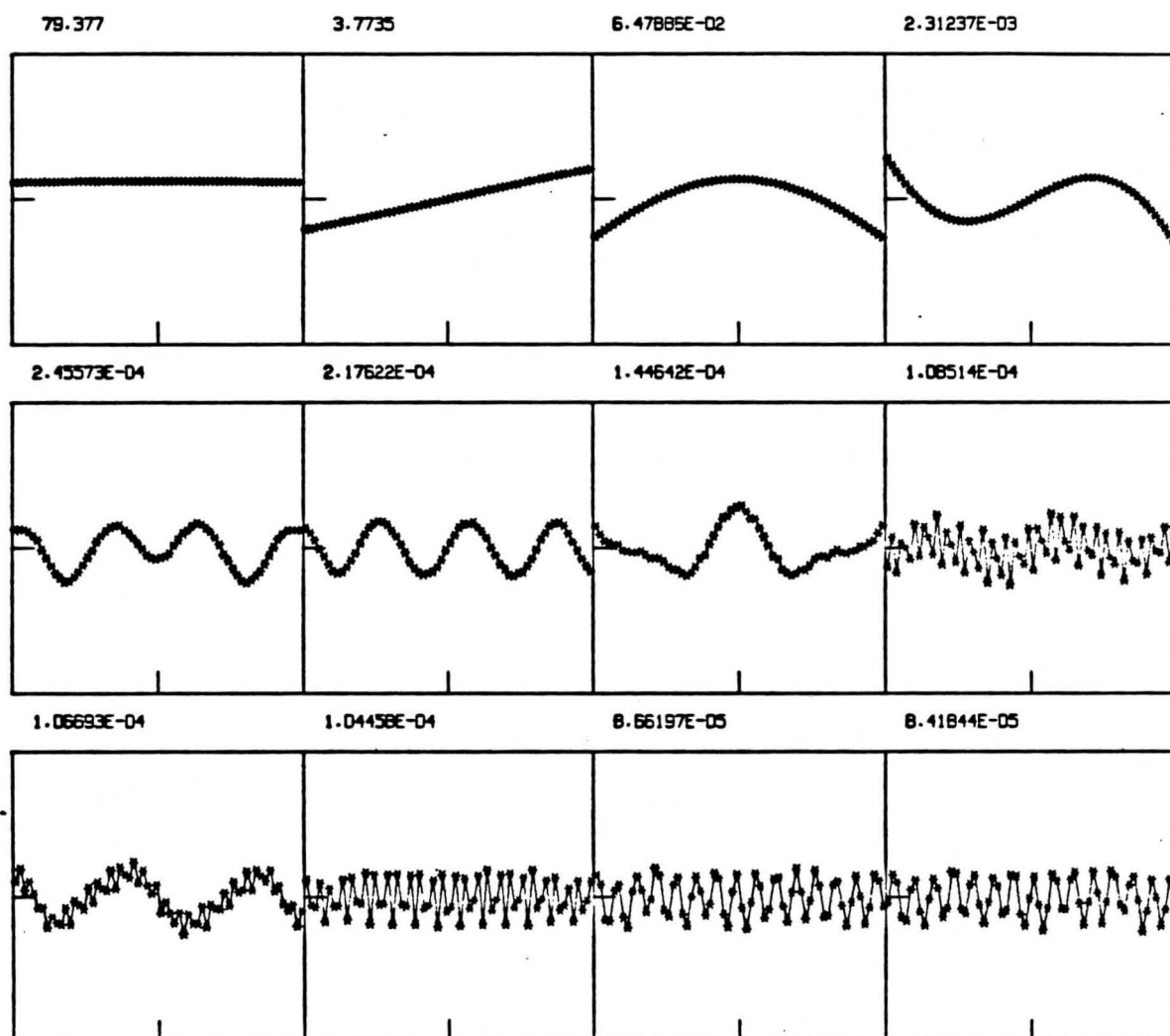


Figure 3b

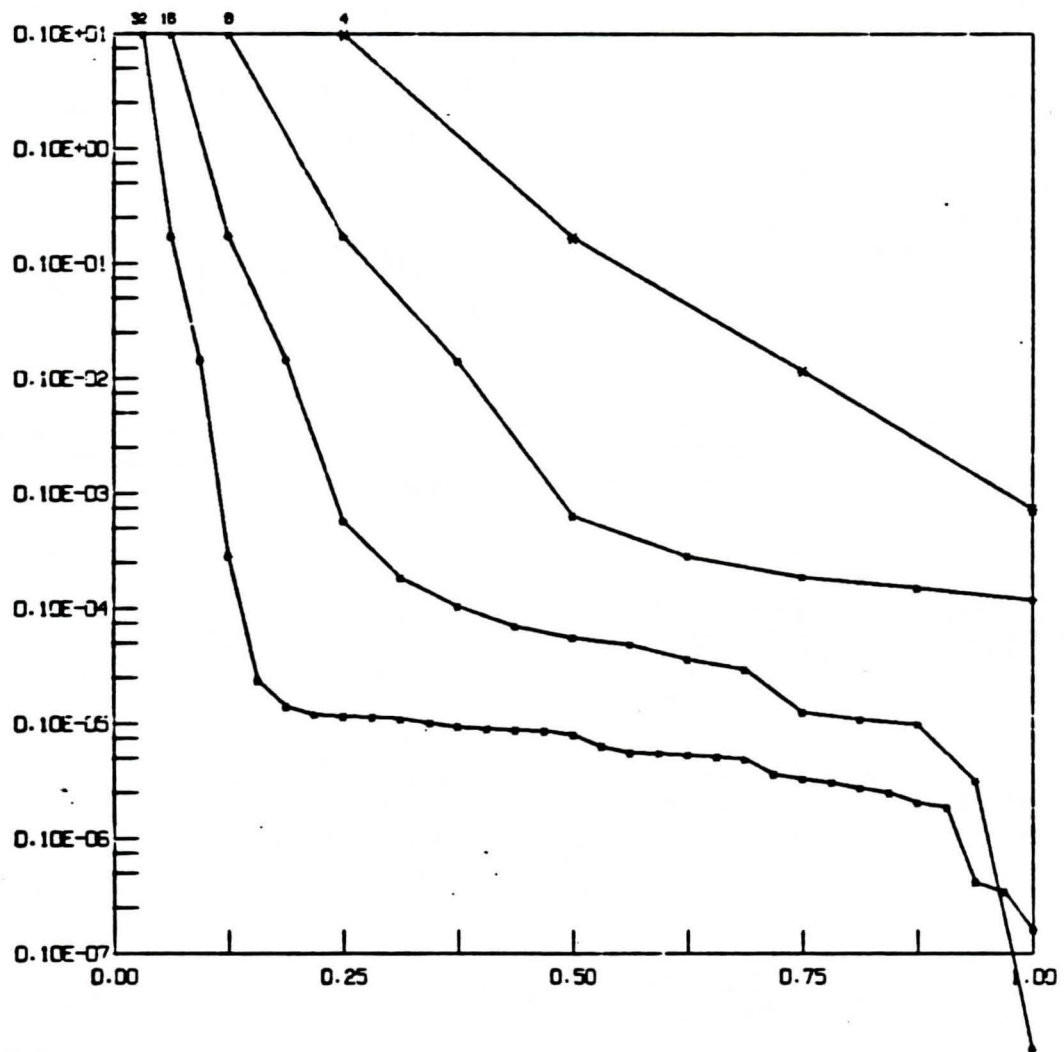


Figure 4

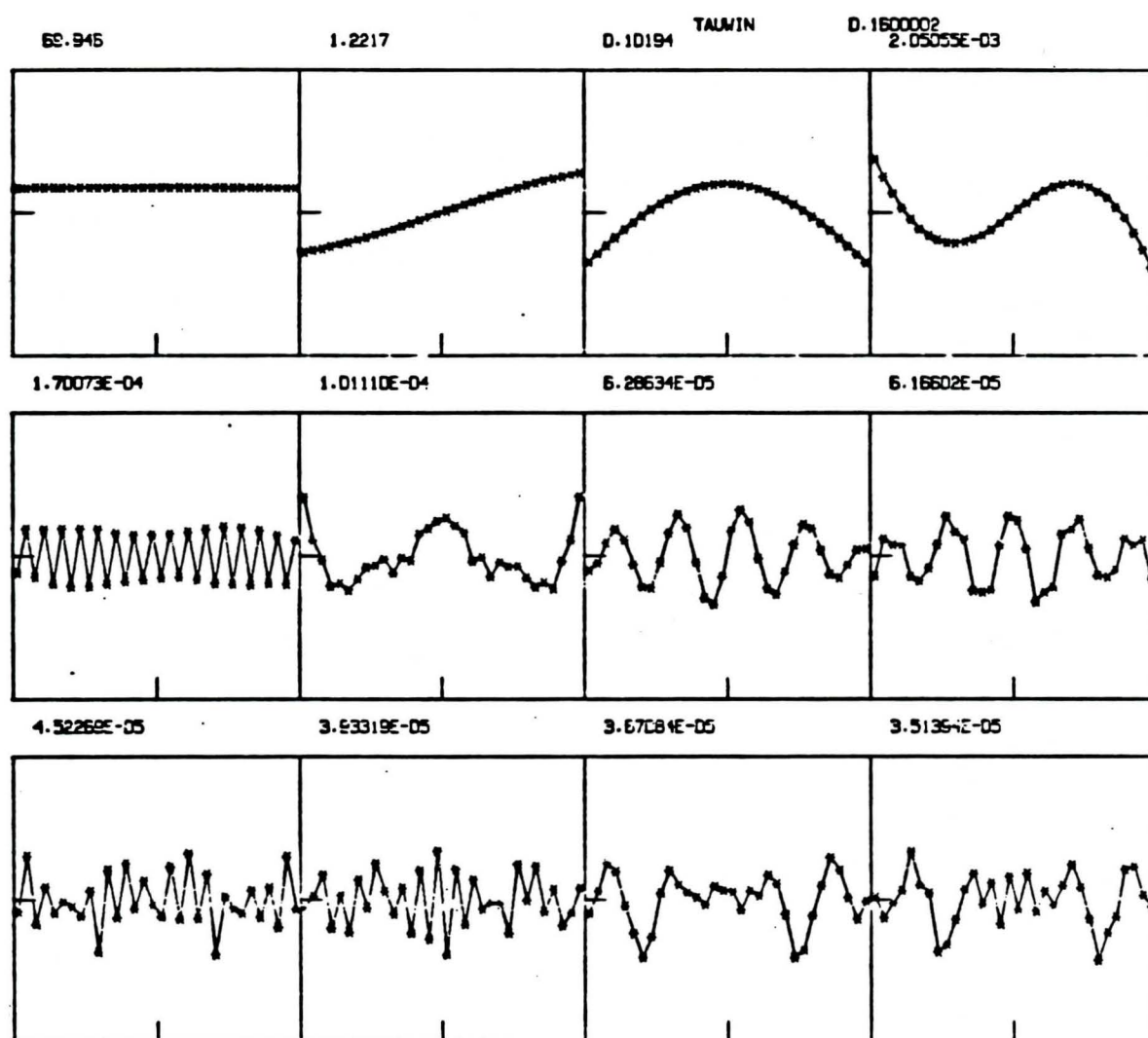


Figure 4

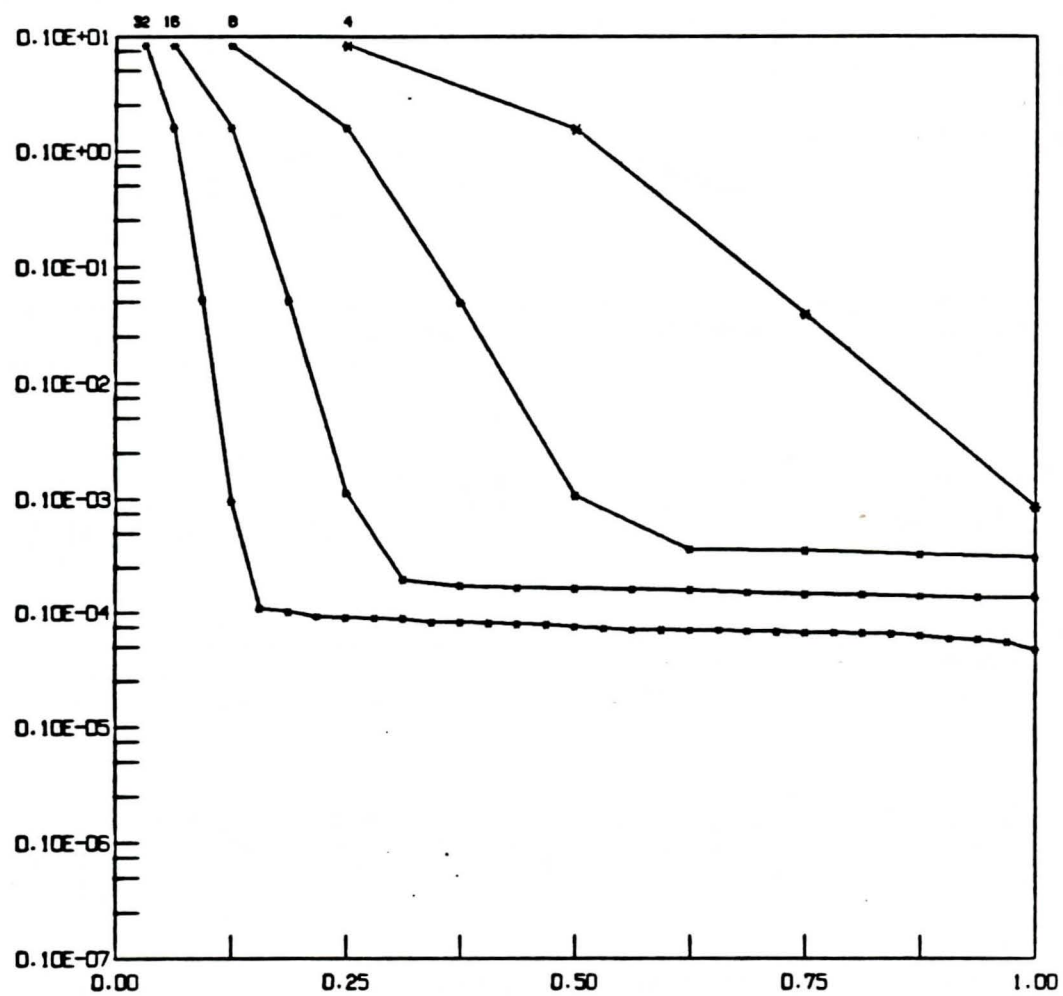


Figure 5

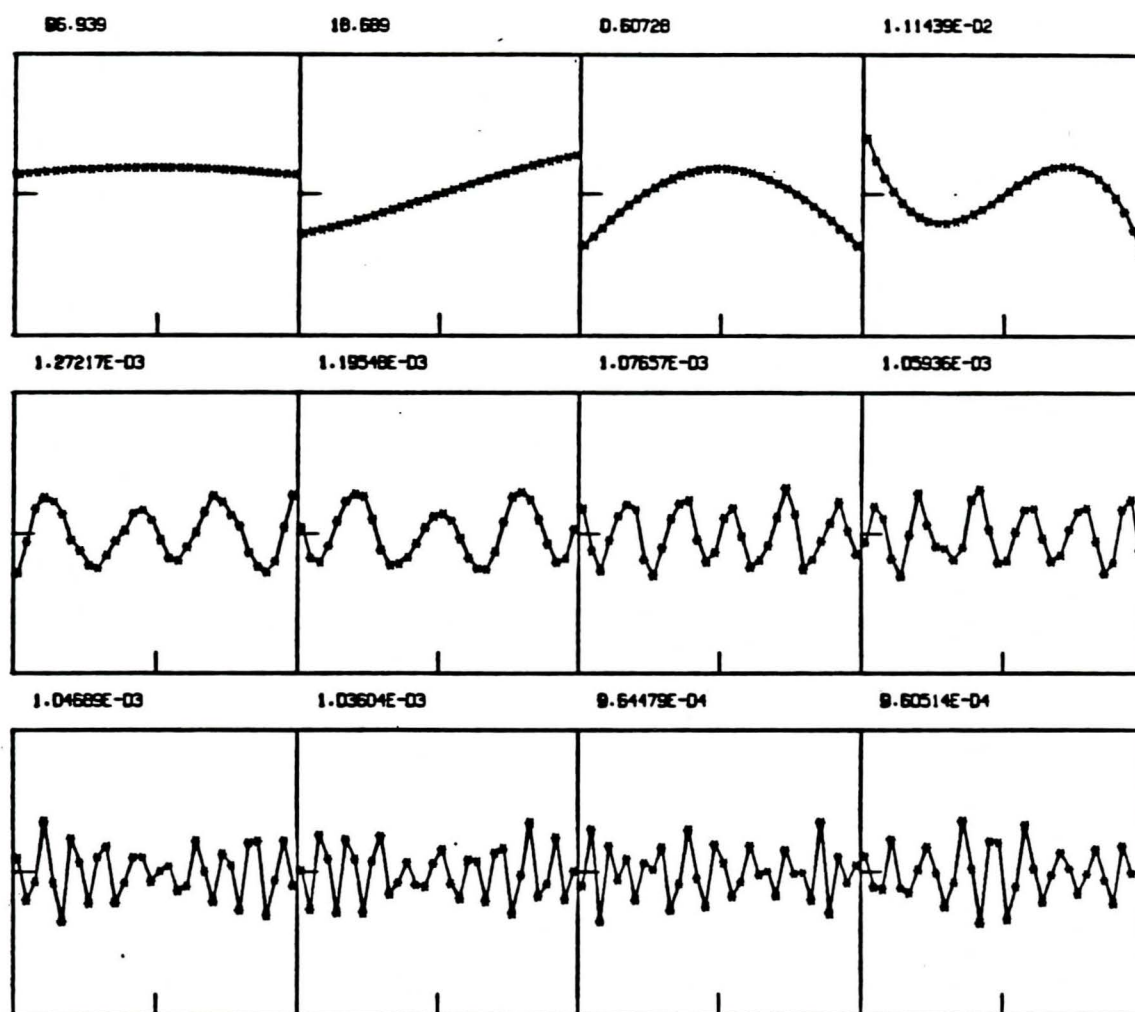


Figure 5

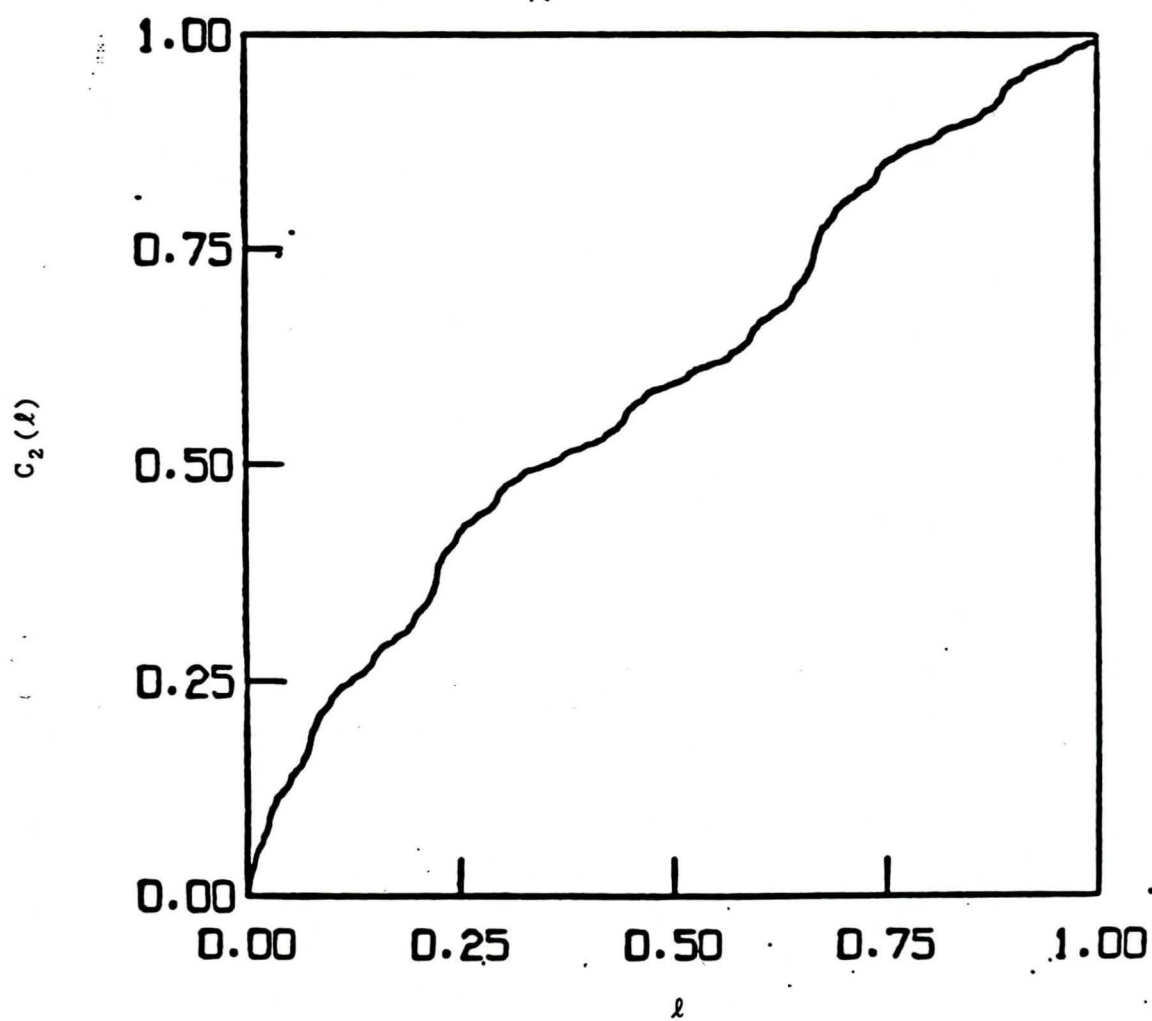


Figure 6

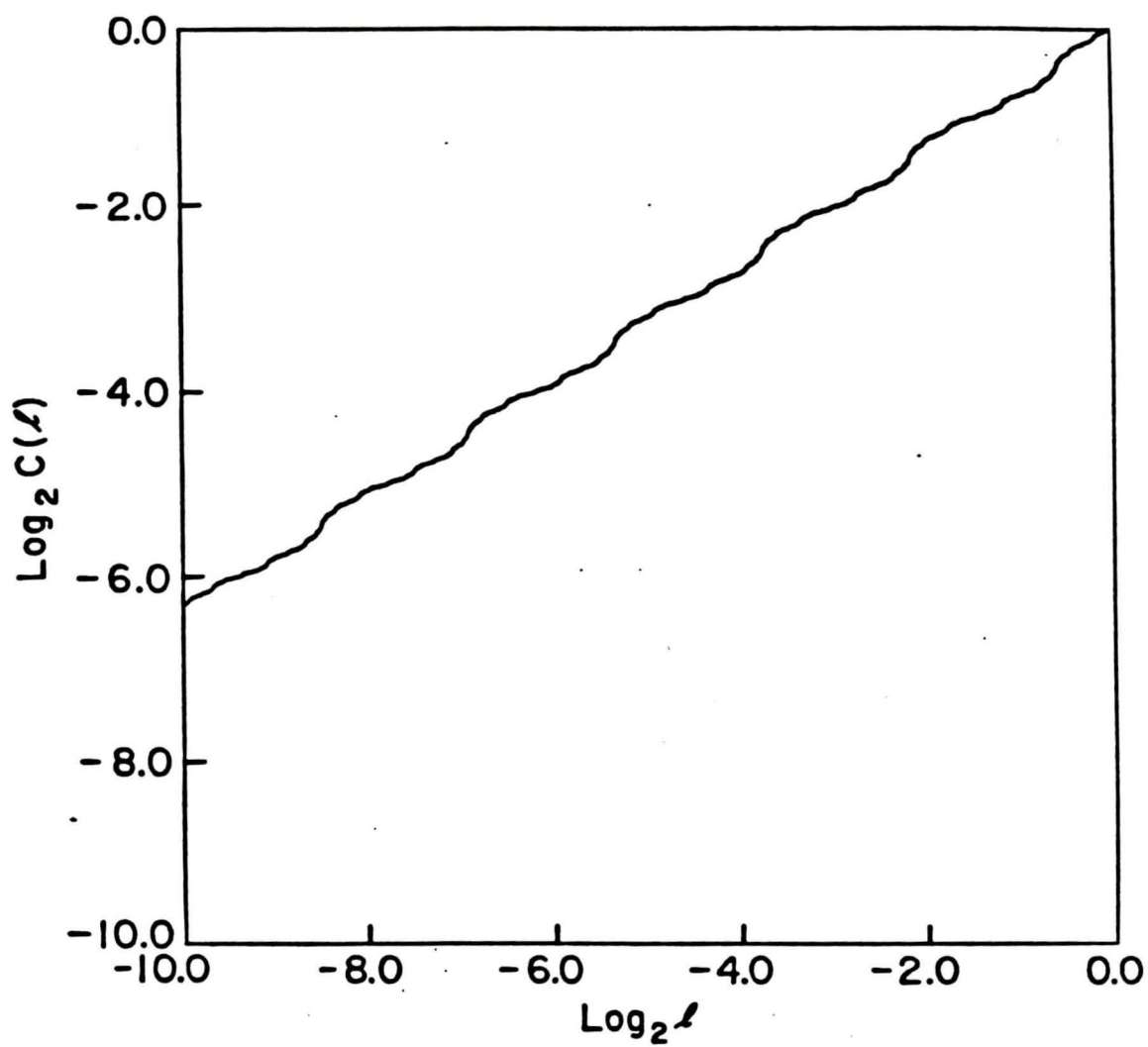


Figure 7

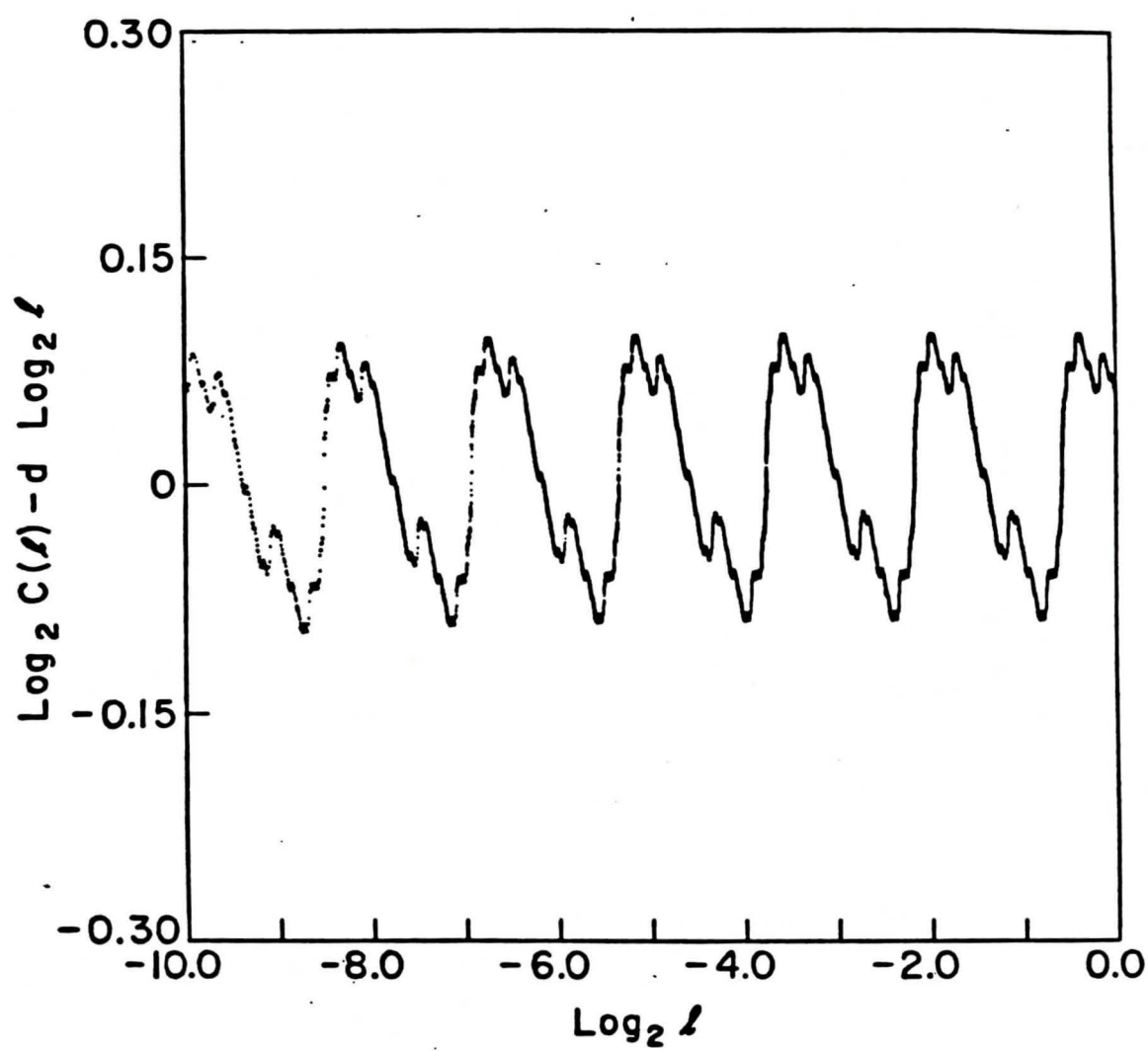


Figure 8

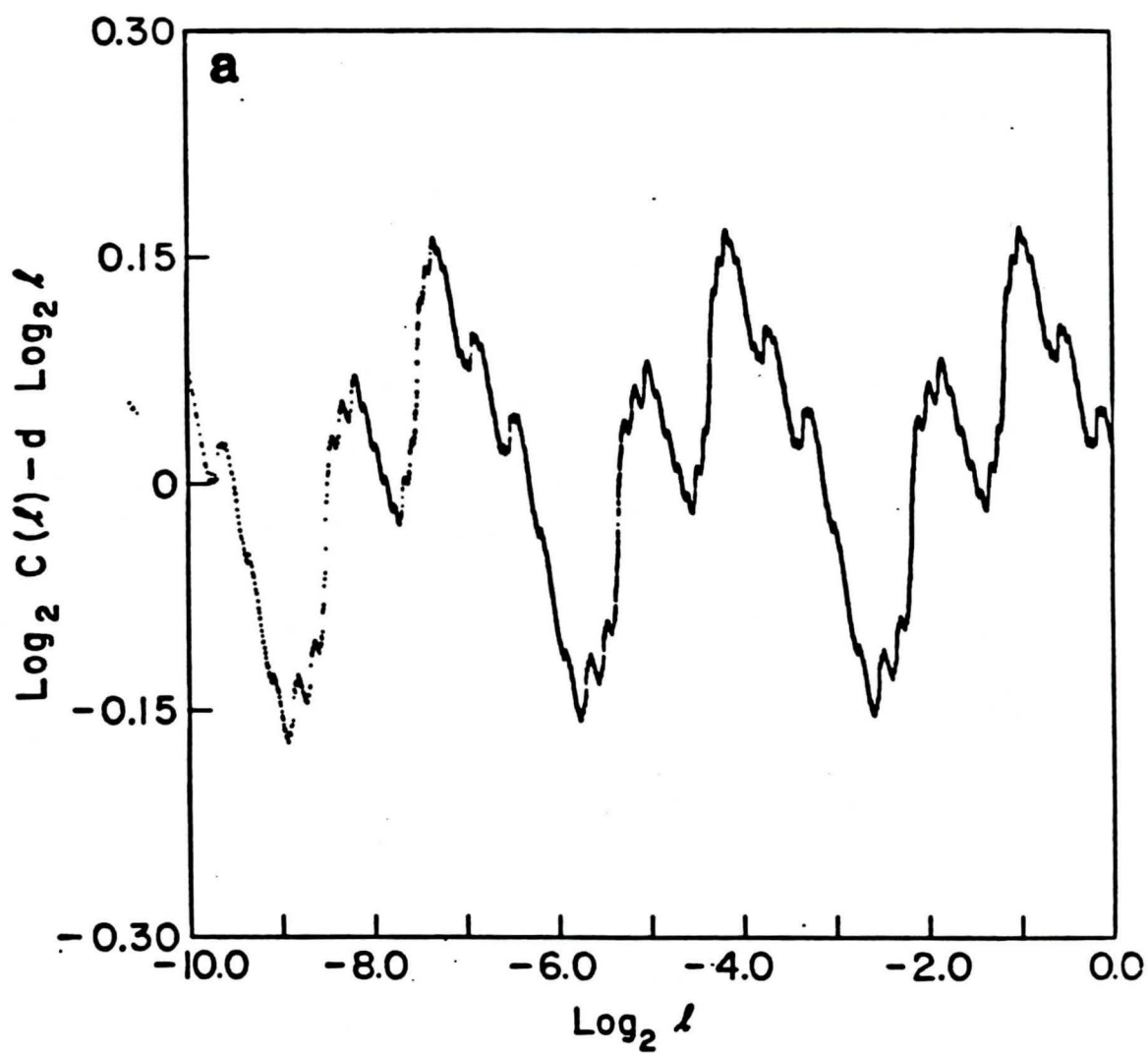


Figure 9

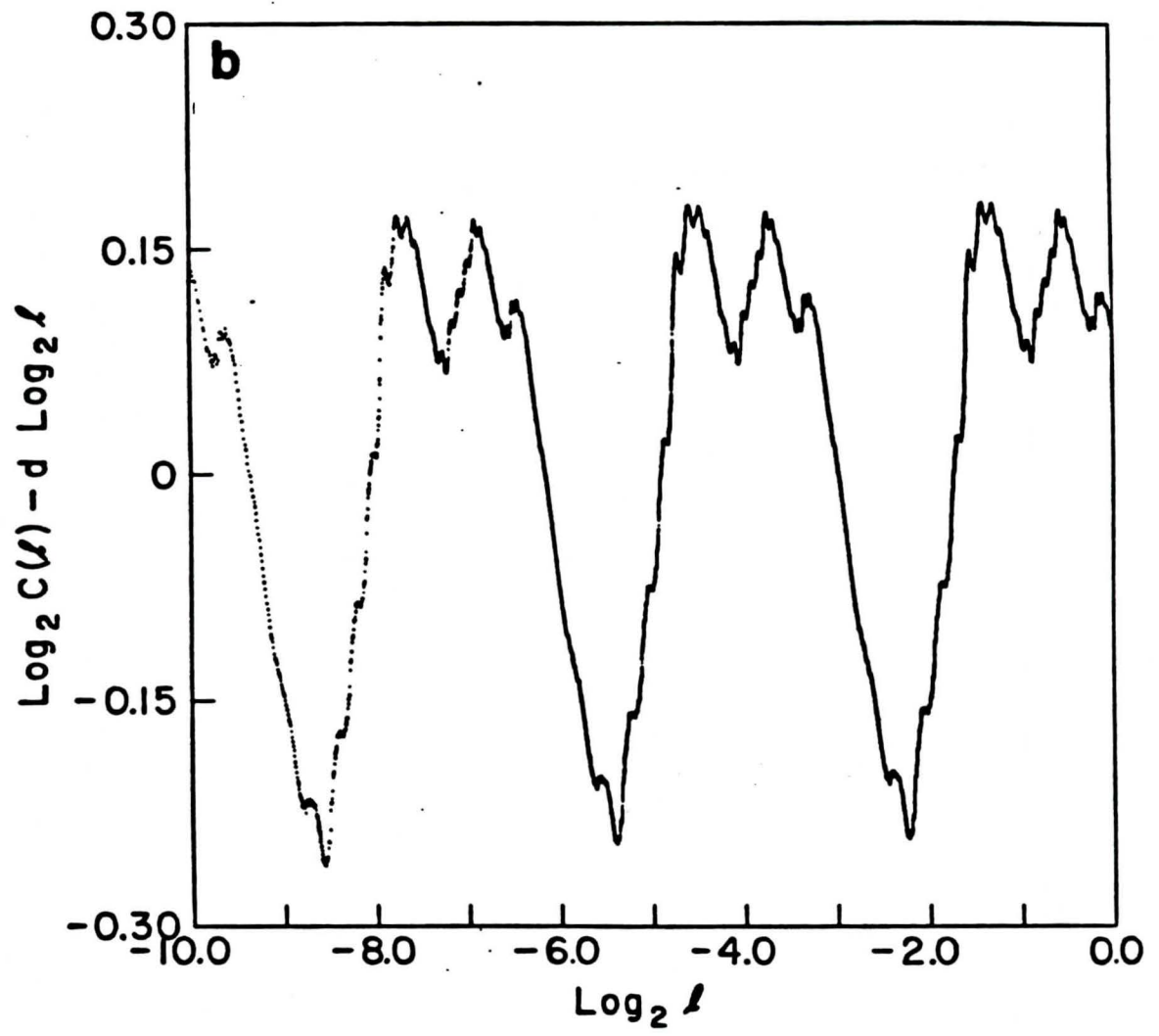
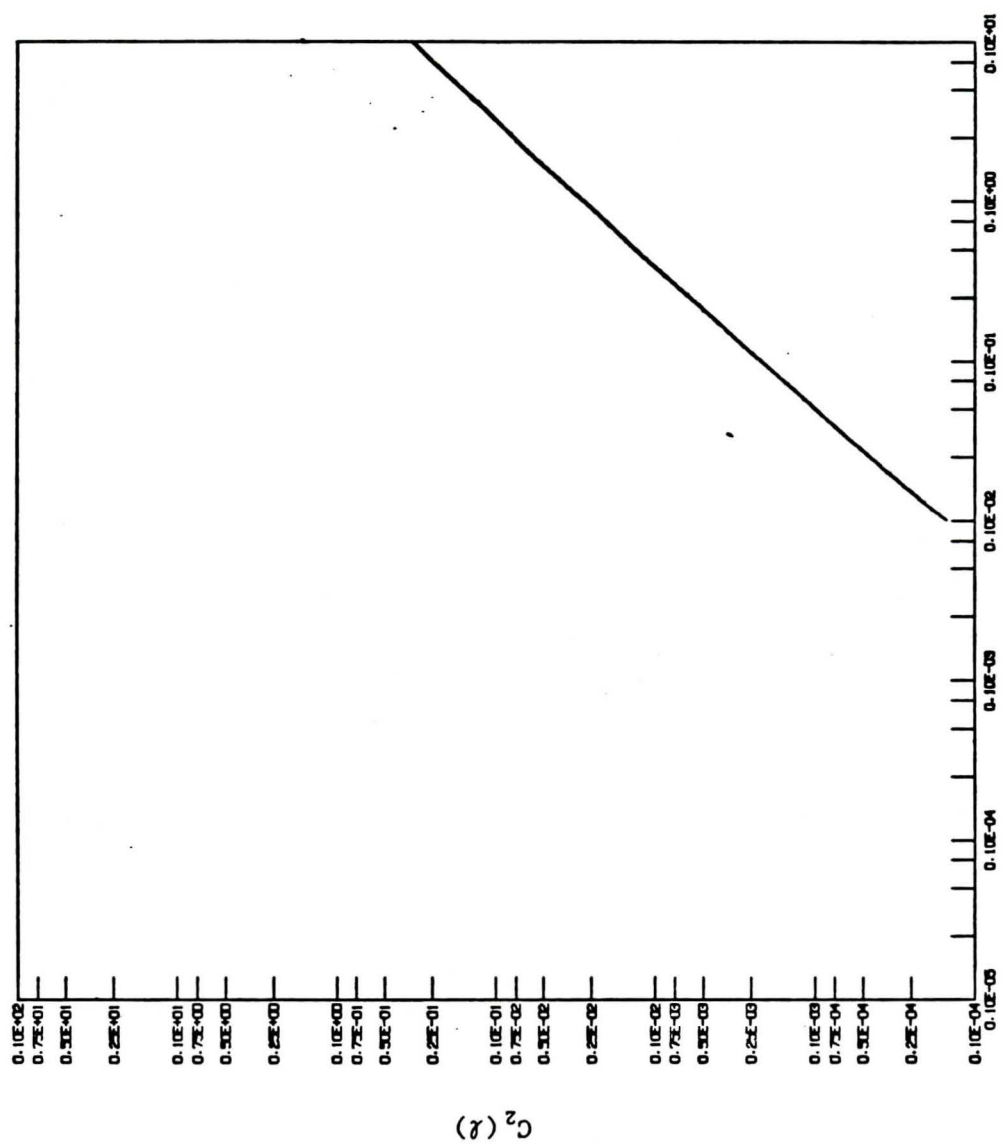


Figure 9



l
Figure 10

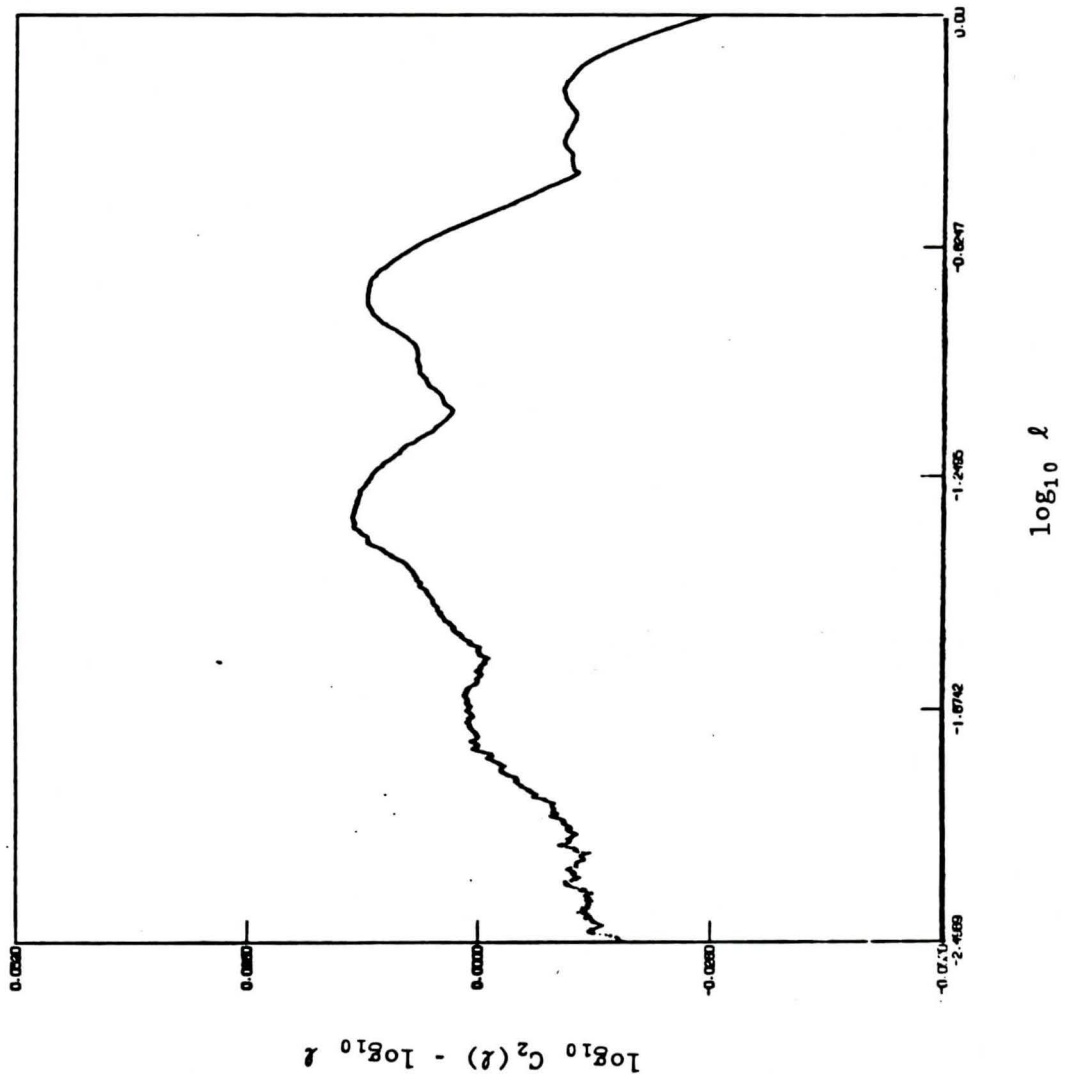


Figure 10

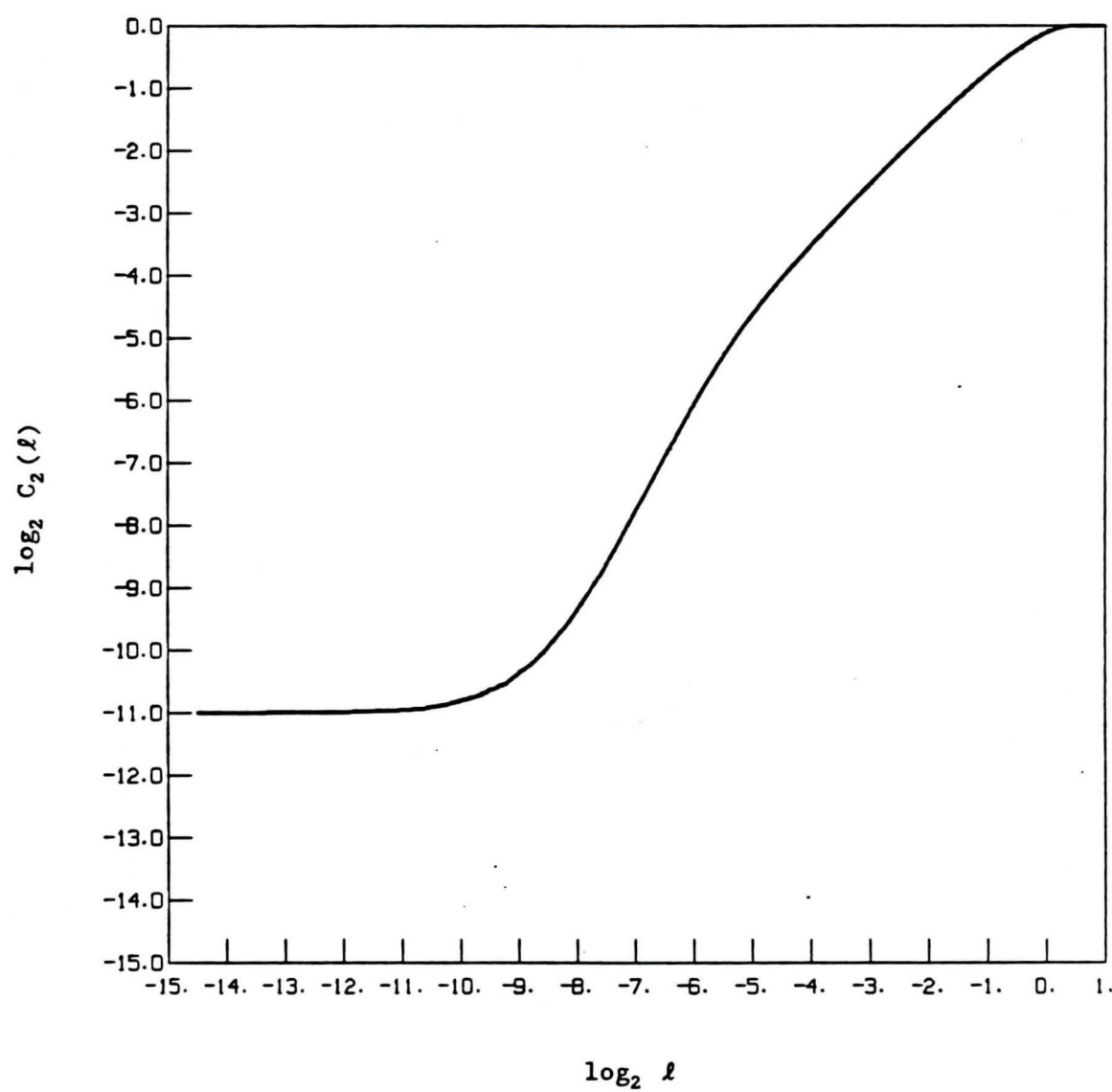


Figure 11

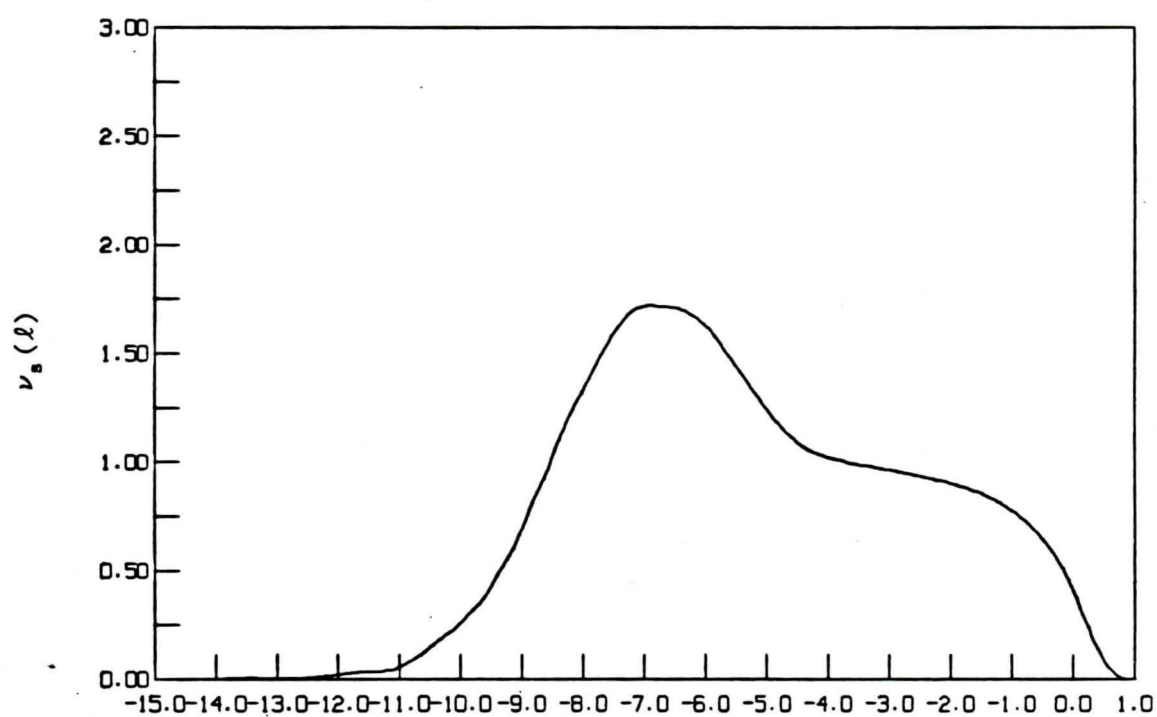


Figure 12

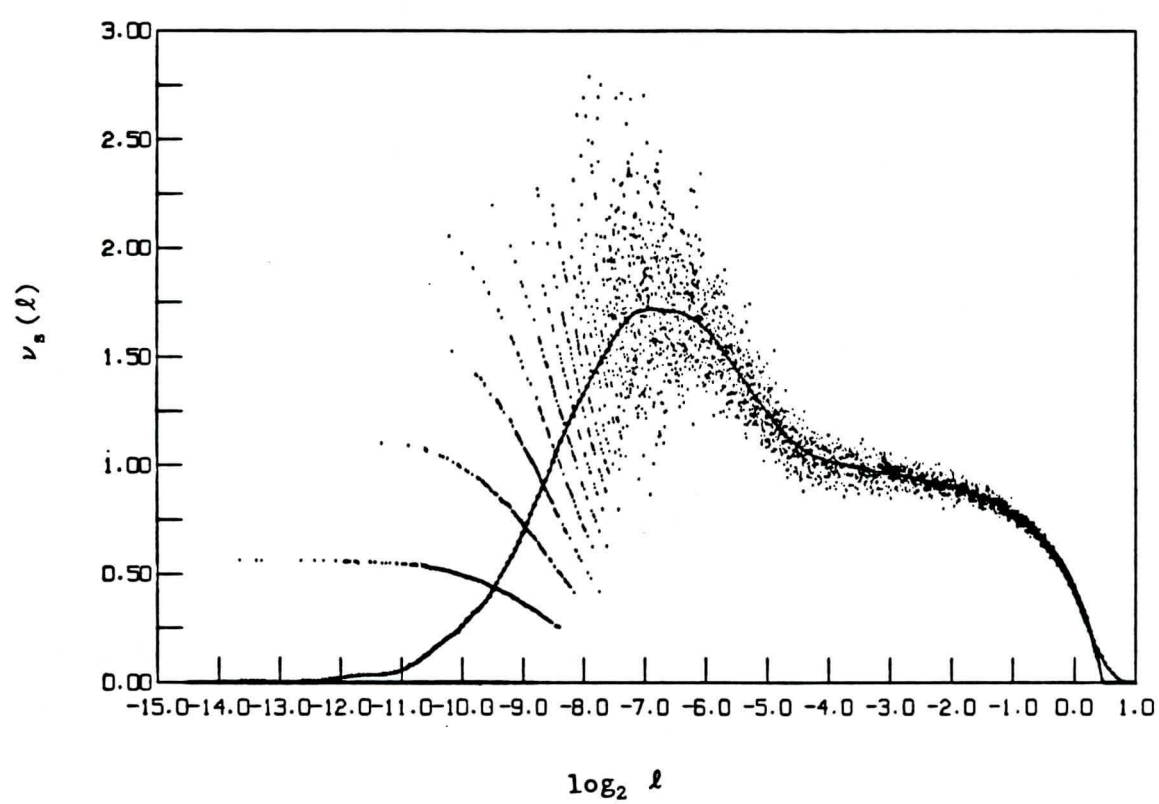


Figure 13

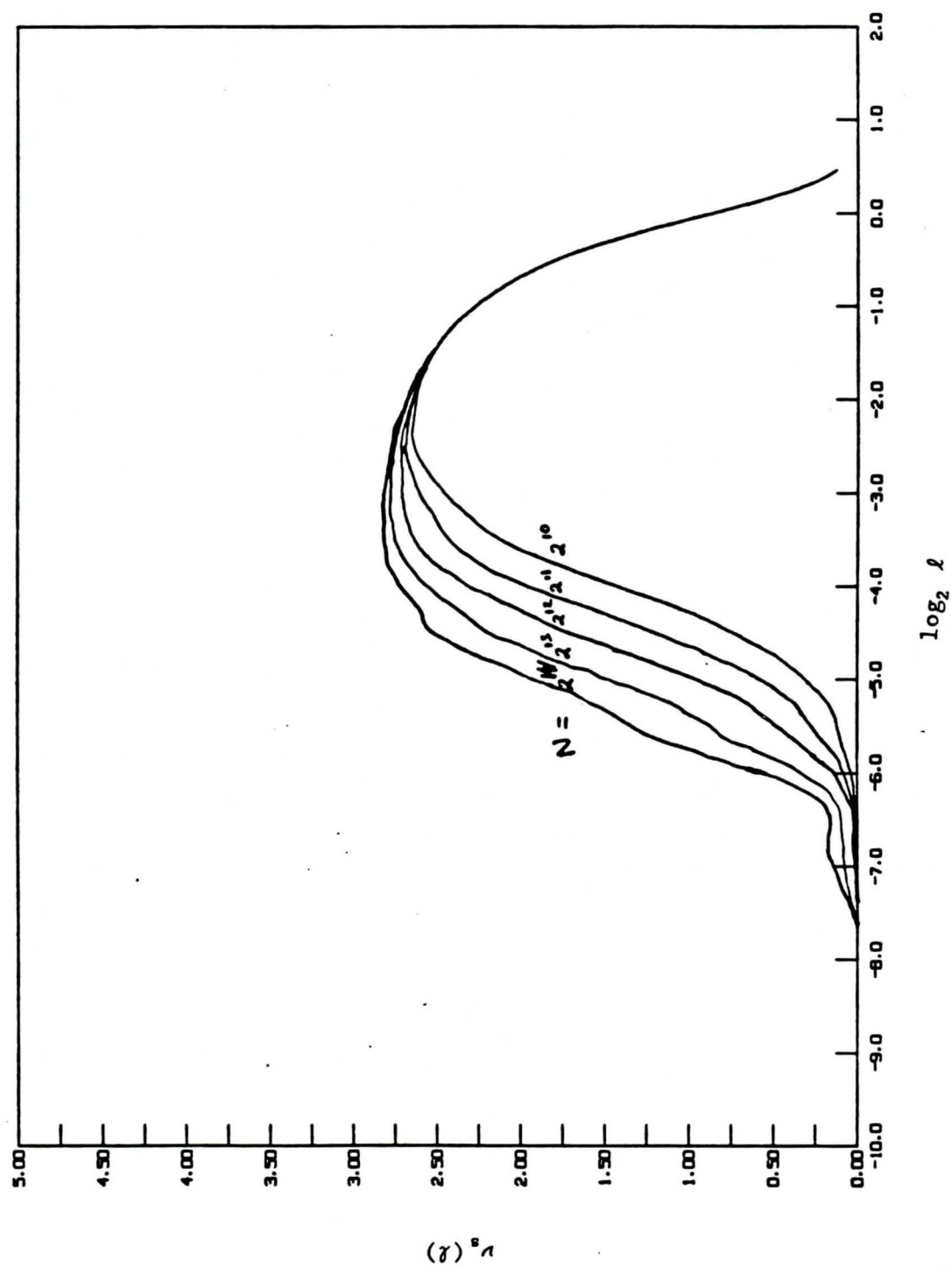


Figure 14

Table 2.1

E M B E D D I N G	6	4.27	4.58	4.85	4.95	5.06	5.16	
		3.03	3.36	3.68	3.89	4.11	4.28	
D D I N G	5	3.72	3.92	4.10	4.16	4.28	4.46	
		2.87	3.17	3.43	3.64	3.80	3.98	
I N G	4	3.12	3.26	3.40	3.45	3.55	3.61	
		2.63	2.84	3.05	3.18	3.30	3.41	
D I M E N S I O N	3	2.45	2.54	2.64	2.69	2.73	2.79	2.83
		2.24	2.36	2.50	2.58	2.65	2.72	2.75
E N S I O N	2	1.75	1.79	1.85	1.87	1.90	1.92	
		1.67	1.75	1.81	1.84	1.87	1.90	
I O N	1	0.957	0.973	0.971	0.987	0.987	0.988	
		0.937	0.955	0.965	0.972	0.980	0.985	
		2 ⁸	2 ⁹	2 ¹⁰	2 ¹¹	2 ¹²	2 ¹³	2 ¹⁴

Number of Data Points

CHAPTER III The Oscillatory Component of Scaling

Introduction

The calculation of scaling exponents is becoming a common tool with which to analyze systems which display complex behavior. Several difficulties remain in processes for the determination of the scaling exponents of inhomogeneous fractal sets.¹ Included are (1) the lack of reliable error estimates, (2) the intrinsic N^2 nature of the much-used Grassberger-Procaccia algorithm (GPA), and (3) the box-counting limitations of current $f(\alpha)$ schemes. In addition, there are no reliable schemes for calculating d_q for large negative q ; global methods have difficulty determining the properties of the least dense singularities. In this chapter we present a new method which alleviates the first two of these difficulties and may assist in solving the third. In Section 3.2 we demonstrate that the log-periodic ψ functions discussed in Chapters 1 and 2 originate from an oscillatory component in the general solution to the scaling equation. The main result of this chapter is a new characterization of the local scaling structure of fractal sets, focused, in the case of strange attractors, on the unstable periodic orbits. Using the periodicity (in $\log \ell$) of the ψ function, one may compute a reliable estimate of a local scaling exponent; this is illustrated in Section 3.3. This procedure is the first to supply meaningful error statistics on the exponent. Since only the local structure is considered, bounds on d_∞ are available. Computationally, the time

¹Definitions of the quantities discussed here are to be found in Chapter 1.

required by the algorithm goes as the number of points. The procedure may be interrupted to examine intermediate results and then restarted easily.

It is generally believed that the singular structure of strange attractors is closely related to the set comprised of all (unstable) periodic orbits. For the flows in a three-dimensional phase space such as the Lorenz and Moore-Spiegel systems, these (closed) trajectories have been investigated by Baker *et al.* (1971). In a time series generated by a weakly chaotic system, a trajectory is often observed to alternate (irregularly) among several cycles; these cycles resemble the periodic orbits. A comprehensive discussion of the importance of periodic orbits to the dynamics of the Lorenz system is given by Sparrow (1982). This behavior can be seen in the Moore-Spiegel attractor (Figure 1.3b), where the trajectory exhibits characteristics of one of several distinct motions between each crossing of the $z=0$ plane. In a two-dimensional mapping (or surface of section) the periodic orbits are isolated points. A method for determining both the periodic and near periodic points for 2-D mappings has been introduced by Cvitanović and Gunaratne (in preparation). If the singular structure is completely determined by the periodic points, it should be possible to determine the scaling structures of a set using the methods introduced in this chapter. A discussion of how this might be done is given in Section 3.5.

This chapter is structured as follows: in Section 3.1 we define the local correlation integral or Cantor function. Section 3.2 then demonstrates its application to familiar fractal sets; oscillatory

components of the Cantor function of those sets are observed and shown to arise from the general solution to the scaling equation. This departure from linear scaling is then utilized in Section 3.3 to determine the scaling exponents of the Feigenbaum set, where some of the scaling exponents may be computed analytically. The method is applied to a strange attractor from the Hénon system in Section 3.4. In § 3.5 we outline a procedure to relate these new scaling exponents to global quantities and discuss implications for $f(\alpha)$. The importance of the oscillations in physical systems is discussed in Chapter 4, where their implications for fractal models of turbulence and experimental observations are shown.

Section 3.1 : The Local Correlation Integral

In this section we introduce the local correlation integral and relate it to the full correlation integral. Consider a distribution of points, perhaps fractal, in E^n , the Euclidean space of dimension n . Define the local correlation integral or Cantor function, $C_1(\mathbf{x}, \ell)$, of the set of $\mathbf{X}$ as

$$C_1(\mathbf{x}, \ell) = \frac{1}{N} \sum_{i=1}^N \theta(\ell - |\mathbf{x}_i - \mathbf{x}|) \quad (3.1)$$

where $\mathbf{x}$ is a vector in E^n , N is the number of points in the set, $\mathbf{x}_i$ is the position vector of the i^{th} point, and θ is the Heavyside

function. In order to get a feel for the scaling properties of $C_1(x, \ell)$, we examine its behavior in several familiar geometries. In each case we assume that N is sufficiently large that there exists a range of scales between the inner and outer cutoff; it is in this range that the scaling behavior is observed. The discrete nature of the sets may be made explicit by including a density factor in the equations below; this is omitted in the interest of clarity. All lengths have been normalized by the outer cutoff. In the following cases the set is considered to lie in E^3 .

If X consists of an isolated point located at x ,

$$C_1(x, \ell) = 1 = \ell^0 . \quad x \in X$$

Alternatively in the case X is a line

$$C_1(x, \ell) = 2 \ell^1 \quad x \in X;$$

when X is a plane

$$C_1(x, \ell) = \pi \ell^2 \quad x \in X;$$

and when X is the full 3-dimensional space

$$C_1(x, \ell) = \frac{4}{3} \pi \ell^3 \quad x \in X.$$

In each case

$$C_1(x, \ell) = A \ell^\mu \quad (3.2)$$

where μ is equal to the topological dimension of the object. In these nonlacunar, Euclidean examples, the various definitions of dimension given in Chapter 1 correspond to the same numerical value and A is a constant. Note that, in general, μ is not equal to d_1 .

Now consider $C_1(x, \ell)$ for sets of fractional Hausdorff dimension.

First, the 101 Cantor set: Figure 3.1a shows $C_1(x, \ell)$ for a 2^{10} point approximation of the 101 set about $x = 0$. This is a classic devil's staircase (a curve of zero derivative everywhere except at points corresponding to members of the Cantor set itself). As ℓ approaches zero, $C_1(0, \ell) \sim \ell^d$ where $d = \log 2 / \log 3$ is equal to the Hausdorff dimension of the underlying set. Figure 3.2 shows log-log plots of the local correlation integral for (a) the 101 Cantor set about $x = 0$, (b) the $\log(4)/\log(3)$ Koch curve (see Mandelbrot, 1977, Chapter 2) about $x = (\sin(\pi/3)/3, 0)$, and (c) for the Hénon attractor about its unstable fixed point x_{us} . The topological dimension of the Cantor set is zero, while that of the Koch curve is one; the topological dimension of the Hénon attractor appears to be one, but no proof of this is available (Farmer, 1982). In each case μ exceeds the topological dimension of the set, implying that the sets are fractal over the range of scales considered.

In Figure 3.3 the local correlation integral of the 101 Cantor set is plotted for several values of x . The vertical offset of each base point is proportional to its distance from the origin. For these choices (see below), the scaling of $C_1(x, \ell)$ at small ℓ is independent of x . Observation at macroscopic scales and the presence of the finite inner cutoff will introduce a dependence on the particular choice of x . Both the strength and weakness of the local correlation integral lies in this x dependence; when evaluated at a "dynamically" interesting point, $C_1(x, \ell)$ yields important information unavailable in the correlation integral $C_2(\ell)$ (defined below). The lack of dependence of μ on x for approximations to the 101 set is due to the

uniformity of this set; for certain elements of a homogeneous fractal there exists an effective outer cutoff inside which $\mu = d_0$. For the 101 set this cutoff, when it exists, is the largest power of $1/3$ which evenly divides $|x|$. If the set is approximated using the decimation procedure of §2.3, a point which is first recorded in the N^{th} generation scales with $\mu = d_0$ for all $(\ell < (1/3)^N)$. For members of the set whose symbolic representation does not terminate in a repeated symbol, no such bound exists. Simply put, if a point is an end point of a reduced image of the full set, it will scale like $x=0$ for sufficiently small ℓ .

To relate the Cantor function, $C_1(x, \ell)$, to the correlation integral, $C_2(\ell)$, recall the definition of the correlation integral (Equation 2.5) for a set of N points embedded in E^N .

$$C_2(\ell) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \theta(\ell - x_{ij}) \quad (3.3)$$

where

$$x_{ij} = |x_i - x_j|$$

is the separation of the i^{th} and j^{th} points. $C_2(\ell)$ is simply the integral of $C_1(x, \ell)$ taken over all members of the set. That is,

$$C_2(\ell) = \frac{1}{N} \int C_1(x_i, \ell) \delta^n(x - x_i) dx \quad (3.4)$$

where the x_i are the elements of the set and δ^n is the n dimensional Dirac delta function. A summation of the repeated index is implied. For sets whose topological dimension is zero, the integral of equation 3.4 reduces to a sum over the elements of the set.

As noted in Chapter 2, the double sum in Equation 3.3 makes computation of the correlation integral difficult for large N . This difficulty is alleviated in the computation of the Cantor function. Evaluating $C_1(x, \ell)$ differs fundamentally from restricting the summation over i in $C_2(\ell)$ to a small number of randomly selected points in that $C_1(x, \ell)$ is evaluated at specific, dynamically significant points.

Section 3.2: Intrinsic Oscillations in Statistics on Fractal Sets

The self-similar nature of fractals implies the presence of log periodic oscillations about the mean scaling behavior of the local correlation integral. These oscillations may be observed in Figures 3.2. Computing the linear least squares fit of $\log C_1(x, \ell)$ as a function of $\log \ell$ for the 101 Cantor set data of Figure 3.2a yields a slope:

$$\mu = 0.630931 \quad \text{for} \quad \frac{1}{2} \left(\frac{1}{3} \right)^6 < \ell < \frac{1}{2} .$$

A regular oscillation, periodic in $\log \ell$, is apparent in the graph of the deviation from the linear trend in each of the examples in Figure 3.2. The oscillation in the scaling of $\log C_1(x_{us}, \ell)$ for the Hénon attractor is considered in Section 3.4.

Such oscillatory corrections have been anticipated and observed by other authors. The earliest reference of which we are aware was by De Bruijn (1948), who discusses log-periodic oscillations in the context of a number theoretic partition problem. An oscillatory

factor of this type was noted by E.A. Novikov (in a footnote of Novikov, 1966) in his model of turbulence. It was this model that laid the foundation for future self-similar mathematical models of turbulence, in which the implication of an oscillatory term was lost. In the late 1960's, de Vaucouleurs (1970, 1971) postulated an oscillatory correction to the "clumpyness" of the mass distribution in the universe. These two topics are treated more fully in Chapter 4. Badii and Politi (1985) found the oscillations in their analysis of the Zaslavski attractor using a method based on mean nearest neighbor distances. They surmised the oscillations were intrinsic to fractal sets.

An oscillatory term is to be expected from the general solution to the scaling equation (Smith, Fournier, and Spiegel, 1986). It may be used to advantage when determining scaling parameters and ignored only at the peril of the investigator.

Consider a statistical moment $C(\ell)$ on the set X where C depends on a separation scale ℓ . As our notation indicates, both C_1 and C_2 are examples. A self-similar fractal set may be expected to satisfy the scaling law

$$C(\ell) = \frac{1}{\rho} C(\sigma\ell) \quad (3.4)$$

for some ρ and σ . The usual power law

$$C(\ell) = A \ell^d, \quad (3.5)$$

where

$$d = \frac{\log \rho}{\log \sigma} \quad (3.6)$$

and A is a constant is only a particular solution of Equation 3.4.