ABSTRACT

Title of Dissertation: INFINITE-DIMENSIONAL DYNAMICAL SYSTEMS AND PROJECTIONS

William Ott, Doctor of Philosophy, 2004

Dissertation directed by: Professor James A. Yorke
Department of Mathematics

We address three problems arising in the theory of infinite-dimensional dynamical systems. First, we study the extent to which the Hausdorff dimension and the dimension spectrum of a fractal measure supported on a compact subset of a Banach space are affected by a typical mapping into a finite-dimensional Euclidean space. We prove that a typical mapping preserves these quantities up to a factor involving the thickness of the support of the measure. Second, we prove a weighted Sobolev-Lieb-Thirring inequality and we use this inequality to derive a physically relevant upper bound on the dimension of the global attractor associated with the viscous lake equations. Finally, we show that in a general setting one may deduce the accuracy of the projection of a dynamical system solely from observation of the projected system.

Infinite-Dimensional Dynamical Systems and Projections

by

William Ott

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2004

Advisory Committee:

Professor James A. Yorke, Chairman/Advisor
Professor Brian Hunt, Advisor
Professor David Levermore, Advisor
Professor Edward Ott
Professor Daniel Rudolph

DEDICATION

I dedicate this dissertation to my parents. The gift of unwavering love and support has no equal.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Introduction

The connection between the theory of dynamical systems and the long-time behavior of solutions of *a priori* infinite-dimensional continuum systems described by partial differential equations is of great importance. Indeed, the application of dynamical ideas to areas of mathematical physics such as turbulence theory and fluid dynamics depends on this relationship. One views the equation of interest as the generator of a semiflow or a flow on a suitable function space. Ergodic theory and dimension theory may then be brought to bear on the analysis of asymptotic behavior in both the deterministic and stochastic contexts.

My research is based on two general lines of inquiry. Intrinsic questions concern the nature of the flow and its asymptotic properties. Examples of such problems include global attractor existence, attractor dimension estimates, inertial manifold existence, and the ergodic properties of invariant measures. The second line of inquiry deals with measurement and reconstruction from experimental data or a finite-dimensional truncation of the flow. One effectively projects the phase space onto a finite-dimensional space in order to reconstruct dynamical objects of interest and compute dynamical invariants. How accurately does the projection of the dynamical system reflect the dynamical system itself? Can the accuracy of the projection be deduced solely from observation of the projected system? We address problems with origins in both lines of inquiry.

In Chapter 2, we study the extent to which the Hausdorff dimension and the dimension spectrum of a fractal measure supported on a compact subset of a Banach space are affected by a typical mapping into a finite-dimensional Euclidean space. Let X be a compact subset of a Banach space B with thickness exponent $\tau(X)$ and Hausdorff dimension $\dim_H(X)$. Let M be any subspace of the Borel measurable functions from B to $\mathbb{R}^m$ that contains the space of linear functions and is contained in the space of locally Lipschitz functions. We prove that for almost every (in the sense of prevalence) function $f \in M$, one has $\dim_H(f(X)) \geqslant \min\{m, \dim_H(X)/(1+\tau(X))\}$. We also prove an analogous result for a certain part of the dimension spectra of Borel probability measures supported on X. The factor $1/(1+\tau(X))$ can be improved to $1/(1+\tau(X)/2)$ if B is a Hilbert space. Since dimension cannot increase under a locally Lipschitz function, these theorems become dimension preservation results when $\tau(X) = 0$. We conjecture that many of the attractors associated with the evolution equations of mathematical physics have zero thickness. The sharpness of our results in the case $\tau(X) \neq 0$ is discussed.

In Chapter 3, we consider the motion of an incompressible fluid confined to a shallow basin with varying bottom topography. A two-dimensional shallow water model has been derived from a three-dimensional anisotropic eddy viscosity model and has been shown to be globally well posed in [40]. The dynamical system associated with the shallow water model is studied. We show that this system possesses a global attractor and that the Hausdorff and box-counting dimensions of this attractor are bounded above by a value proportional to the weighted L^2-

norm of the wind forcing function. A weighted Sobolev-Lieb-Thirring inequality plays the key role in obtaining the dimension estimate.

In Chapter 4, we study the extent to which the accuracy of a projection may be deduced solely from observation of the projected system. Let A be a compact invariant set for a map f on $\mathbb{R}^n$ and let $\phi : \mathbb{R}^n \to \mathbb{R}^m$ where $n > m$ be a "typical" smooth map. When can we say that A and $\phi(A)$ are similar, based only on knowledge of the images in $\mathbb{R}^m$ of trajectories in A? For example, under what conditions on $\phi(A)$ (and the induced dynamics thereon) are A and $\phi(A)$ homeomorphic? Are their Lyapunov exponents the same? Or, more precisely, which of their Lyapunov exponents are the same? We address these questions with respect to both the general class of smooth mappings ϕ and the subclass of delay coordinate mappings.

In answering these questions, a fundamental problem arises about an arbitrary compact set A in $\mathbb{R}^n$. For $x \in A$, what is the smallest integer d such that there is a C^1 manifold of dimension d that contains all points of A that lie in some neighborhood of x? We define a tangent space T_xA in a natural way and show that the answer is $d = \dim(T_xA)$. As a consequence we obtain a Platonic version of the Whitney embedding theorem.

Chapter 2

The Effect of Projections on Fractal Sets and Measures in Banach Spaces

Written in collaboration with Brian Hunt and Vadim Kaloshin, the material in this chapter has been submitted for publication in *Ergodic Theory & Dynamical Systems*.

2.1 Introduction

Many infinite-dimensional dynamical systems have been shown to have compact finite-dimensional attractors [9, 20, 56, 61, 64]. Such attractors exist for a variety of the evolution equations of mathematical physics, including the Navier-Stokes system, various classes of reaction-diffusion systems, nonlinear dissipative wave equations, and complex Ginzburg-Landau equations. When an attractor is measured experimentally, one observes a 'projection' of the attractor into finite-dimensional Euclidean space. This technique of observation via projection leads to a natural and fundamental question. How accurately does the image of the attractor reflect the attractor itself? We address this question from a dimension-theoretic perspective and we consider the following problem. For an attractor of an infinite-dimensional dynamical system, how is its dimension affected by a typical projection into a finite-dimensional Euclidean space?

One may define the dimension of an attractor in many different ways. Setting aside dynamics, the attractor may be viewed as a compact set of points in a metric space. Viewing the attractor in this light, the dimension of the attractor may be defined as the box-counting dimension or the Hausdorff dimension of the attracting set. Measure-dependent notions of attractor dimension take into account the distribution of points induced by the dynamics and are thought to be more accurately measured from numerical or experimental data. One often analyzes the 'natural measure,' the probability measure induced by the statistics of a typical trajectory that approaches the attractor. A natural measure is not known to exist for arbitrary systems, but it does exist for Axiom A attractors and for certain classes of systems satisfying conditions weaker than uniform hyperbolicity. See [32, 67] for expository discussions of systems that are known to have natural measures.

The dimension spectrum (D_q spectrum) characterizes the multifractal structure of an attractor. Given a Borel measure μ with compact support X in some metric space, for $q \geqslant 0$ and $q \neq 1$ let

$$D_q(\mu) = \lim_{\epsilon \to 0} \frac{\log \int_X \left[\mu(B(x,\epsilon))\right]^{q-1} d\mu(x)}{(q-1)\log \epsilon}, \tag{2.1}$$

provided the limit exists, where $B(x,\epsilon)$ is the ball of radius ϵ centered at x. (If the limit does not exist, define $D_q^+(\mu)$ and $D_q^-(\mu)$ to be the lim sup and lim inf, respectively.) Let

$$D_1(\mu) = \lim_{q \to 1} D_q(\mu),$$

again provided the limit exists. This spectrum includes the box-counting dimension (D_0), the information dimension (D_1), and the correlation dimension (D_2).

In particular, when $q = 0$ the dimension depends only on the support X of μ and we write $D_0(X) = D_0(\mu)$. See Section 2.2 for a discussion of this definition and its relationship to other definitions of D_q in the literature.

The goal of this paper is to extend the following theorems, as much as possible, to infinite-dimensional Banach spaces. In all of the results in this paper, 'almost every' is in the sense of *prevalence*, a generalization of 'Lebesgue almost every' to infinite-dimensional spaces. See Section 2.2 and [33, 34] for details.

Theorem 2.1 ([58]). *Let $X \subset \mathbb{R}^n$ be a compact set. For almost every function $f \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, one has*

$$\dim_H(f(X)) = \min\{m, \dim_H(X)\}$$

where $\dim_H(\cdot)$ is the Hausdorff dimension.

Theorem 2.2 ([30]). *Let μ be a Borel probability measure on $\mathbb{R}^n$ with compact support and let q satisfy $1 < q \leqslant 2$. Assume that $D_q(\mu)$ exists. Then for almost every function $f \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, $D_q(f(\mu))$ exists and is given by*

$$D_q(f(\mu)) = \min\{m, D_q(\mu)\}.$$

For each result, the space C^1 can be replaced by any space that contains the linear functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ and is contained in the locally Lipschitz functions. Theorem 2.1 extends to smooth functions a result of Mattila [45] (generalizing earlier results of Marstrand [44] and Kaufmann [38]) that makes the same conclusion for almost every linear function from $\mathbb{R}^n$ into $\mathbb{R}^m$, in the sense of Lebesgue measure on the space of m-by-n matrices. Strictly speaking, Marstrand, Kaufmann, and Mattila considered orthogonal projections, but the analogous results for general linear projections follow immediately. Sauer and

Yorke [58] prove Theorem 2.2 for the correlation dimension (D_2) and recover (2.1) by invoking a variational principle for Hausdorff dimension [22]. Theorem 2.1 and its predecessors follow from a potential-theoretic characterization of the dimensions involved. Roughly speaking, the dimension is the largest exponent for which a certain singular integral converges. Theorem 2.2 follows from a similar characterization of D_q for $q > 1$ [30].

The potential-theoretic approach only leads to a dimension preservation result for D_q if $1 < q \leqslant 2$. For $0 \leqslant q < 1$ and $q > 2$, [30] gives examples for which D_q is not preserved by any linear transformation into $\mathbb{R}^m$. For $0 \leqslant q < 1$, the construction is based on the discovery by Kan [59, 58] of a class of examples for which the box dimension is not preserved by any C^1 function.

When the ambient space is not finite-dimensional, one does not expect a dimension preservation result analogous to Theorem 2.1 or Theorem 2.2 to hold. We use the thickness exponent to study the extent to which the dimension spectrum is affected by projection from a Banach space to $\mathbb{R}^m$. This exponent, defined precisely in Section 2.2 and denoted $\tau(X)$, measures how well a compact subset X of a Banach space B can be approximated by finite-dimensional subspaces of B, with smaller values of the thickness exponent indicating better approximability. In general one has $\tau(X) \leqslant D_0^+(X)$, the upper box-counting dimension of X, and equality is possible. We expect that the thickness exponent can be shown to be significantly smaller than the box-counting dimension for many attractors of infinite-dimensional systems. Studying the Hölder regularity of embeddings of infinite-dimensional fractal sets into finite-dimensional spaces, [31] establishes a bound on the amount the dimension may drop for a typical projection.

Theorem 2.3 ([31]). *Let B denote a Banach space. Let $X \subset B$ be a compact* *set with box-counting dimension d and thickness exponent $\tau(X)$. Let $m > 2d$ be an integer, and let α be a real number with*

$$0 < \alpha < \frac{m - 2d}{m(1 + \tau(X))}.$$

Then for almost every (in the sense of prevalence) bounded linear function (or C^1 function, or Lipschitz function) $f : B \to \mathbb{R}^m$ there exists $C > 0$ such that for all x, $y \in X$,

$$C|f(x) - f(y)|^\alpha \geqslant |x - y|. \tag{2.2}$$

For such a function f, one has

$$\frac{m - 2d}{m(1 + \tau(X))} \dim(X) \leqslant \dim(f(X)) \leqslant \dim(X)$$

where $\dim(X)$ *represents either the box-counting dimension or Hausdorff dimension.*

This theorem generalizes earlier results in [5] and [23].

For a function f satisfying (2.2), the factor by which the dimension may drop is the product of two terms, $(m-2d)/m$ and $1/(1+\tau(X))$. The first term depends on the embedding dimension m and converges to one as $m \to \infty$ while the second term depends intrinsically on X via its thickness. We prove that the Hausdorff dimension is preserved by a typical projection up to a factor of $1/(1 + \tau(X))$. In particular, the factor $(m - 2d)/m$ has been removed. We now state the main theorem for compact subsets of Banach spaces. Because of the possibility of dimension drop, the existence of $D_q(\mu)$ does not imply the existence of $D_q(f(\mu))$ for functions f satisfying the conclusion of the theorem. We therefore formulate the result in terms of the lower dimension D_q^-.

Banach Space Theorem. *Let B be a Banach space, and let M be any subspace of the Borel measurable functions from B to $\mathbb{R}^m$ that contains the space of linear* *functions and is contained in the space of locally Lipschitz functions. Let $X \subset B$ be a compact set with thickness exponent $\tau(X)$. Let μ be a Borel probability measure supported on X. For almost every $f \in M$, one has*

$$\dim_H(f(X)) \geqslant \min\left\{m, \frac{\dim_H(X)}{1+\tau(X)}\right\},$$

and, for $1 < q \leqslant 2$,

$$D_q^-(f(\mu)) \geqslant \min\left\{m, \frac{D_q^-(\mu)}{1+\tau(X)}\right\}. \tag{2.3}$$

Notice that for sets with thickness zero, the Banach space theorem is a dimension preservation result. Every compact set $X \subset \mathbb{R}^n$ has thickness zero. Thus, the Banach space theorem generalizes Theorems 2.1 and 2.2. Furthermore, it strengthens Theorem 2.2, because for a prevalent set of functions, (2.3) holds simultaneously for all $1 < q \leqslant 2$. On the other hand, suppose $\tau(X) > 0$. The Hausdorff dimension of X may be *noncomputable* in the sense that for any positive integer m and any subspace M of the Borel measurable functions from B to $\mathbb{R}^m$, $\dim_H(f(X)) < \dim_H(X)$ for all $f \in M$. In other words, the Hausdorff dimension of X cannot be ascertained from any finite-dimensional representation of X.

The proof of the Banach space theorem uses only the most general information about the structure of the dual space B'. In specific situations, additional knowledge about the structure of the dual space may yield improved theorems. We show that this does indeed happen in the Hilbert space setting.

Hilbert Space Theorem. *Let H be a Hilbert space, and let M be any subspace of the Borel measurable functions from H to $\mathbb{R}^m$ that contains the space of linear functions and is contained in the space of locally Lipschitz functions. Let $X \subset H$*

be a compact set with thickness exponent $\tau(X)$. Let μ be a Borel probability measure supported on X. For almost every $f \in M$, one has

$$\dim_H(f(X)) \geqslant \min\left\{m, \frac{\dim_H(X)}{1+\tau(X)/2}\right\},$$

and, for $1 < q \leqslant 2$,

$$D_q^-(f(\mu)) \geqslant \min\left\{m, \frac{D_q^-(\mu)}{1+\tau(X)/2}\right\}. \tag{2.4}$$

As we have mentioned, examples in [30] preclude similar results for $0 \leqslant q < 1$ and $q > 2$. The case $q = 1$ is of interest because it corresponds to the commonly used notion of information dimension, in the following sense. In general, the limit (2.1) need not exist. However, $D_q^-(\mu)$ is a nonincreasing function of q and is continuous for $q \neq 1$ [4]. From this it follows that (2.3) and (2.4) hold for $q = 1$ if we define

$$D_1^-(\mu) = \lim_{q \to 1^+} D_q^-(\mu).$$

Next, we consider the sharpness of the Banach and Hilbert space theorems. In [31], the authors give an example of a compact subset X of Hausdorff dimension d in ℓ^p for $1 \leqslant p < \infty$ such that for all bounded linear functions $\pi : \ell^p \to \mathbb{R}^m$,

$$\dim_H(\pi(X)) \leqslant \frac{d}{1+d/q}$$

where $q = p/(p-1)$. In these cases, $\tau(X) = d$. Thus, the Hausdorff dimension parts of the Banach and Hilbert space theorems are sharp, in the sense that there is no better bound in terms of $\tau(X)$ that holds for all such spaces (notice that $q = 2$ for the separable Hilbert space ℓ^2 and $q \to 1$ as $p \to \infty$).

On the other hand, when $p = 1$, q is infinite, and the example in [31] does not rule out the possibility of a dimension preservation result for subsets of ℓ^1 of arbitrary thickness. We demonstrate that such a result is not possible by constructing a compact subset X of Hausdorff dimension d in ℓ^1 such that for all bounded linear functions $\pi : \ell^1 \to \mathbb{R}$,

$$\dim_H(\pi(X)) \leqslant \frac{d}{1 + d/2}.$$

In light of this example, we are somewhat pessimistic regarding the existence of infinite-dimensional spaces for which a general dimension preservation theorem holds. It is thus natural to consider the following fundamental question. Suppose X represents the global attractor of a flow on a function space generated by an evolution equation. Under what hypotheses on the flow does one have $\tau(X) = 0$? If one assumes that the flow is sufficiently dissipative and smoothing, then X will have finite box dimension. We conjecture that similar *dynamical* hypotheses imply that $\tau(X) = 0$. Friz and Robinson [24] obtain a result of this type. They prove that if an attractor is uniformly bounded in the Sobolev space H^s on an appropriate bounded domain in $\mathbb{R}^m$, then its thickness is at most m/s. This result implies that certain attractors of the Navier-Stokes equations have thickness exponent zero. Roughly speaking, thickness is inversely proportional to smoothness.

Section 2.2 reviews prevalence, the dimension spectrum, and the thickness exponent. The main two theorems are presented and proved in Section 2.3. In Section 2.4 we describe the counterexample to the dimension preservation conjecture for subsets of ℓ^1 of arbitrary thickness.

2.2 Preliminaries

We discuss prevalence, the dimension spectrum, and the thickness exponent.

2.2.1 Prevalence

Mathematicians often use topological notions of genericity when formulating theorems in dynamical systems and topology. In topological terms, 'generic' refers to an open and dense subset of mappings, or to a countable intersection of such sets (a 'residual' subset). In finite-dimensional spaces, there exists considerable discord between the topological notion of genericity and the measure-theoretic notion of the size of a set (see [33, 51] for examples). Prevalence is intended to be a better analogue to "probability one" on function spaces where no Lebesgue or Haar measure exists.

To motivate the definition of prevalence on a Banach space B, consider how the notion of 'Lebesgue almost every' on $\mathbb{R}^n$ can be formulated in terms of the same notion on lower-dimensional spaces. Foliate $\mathbb{R}^n$ by k-dimensional planes, which by an appropriate choice of coordinates we think of as translations of $\mathbb{R}^k \subset \mathbb{R}^n$ by elements of $\mathbb{R}^{n-k}$. If 'Lebesgue almost every' translation of $\mathbb{R}^k$ intersects a Borel set $S \subset \mathbb{R}^n$ in full k-dimensional Lebesgue measure, then S has full n-dimensional Lebesgue measure by the Fubini theorem. If $\mathbb{R}^n$ is replaced by an infinite-dimensional space B, we cannot formulate the same condition because the space of translations of a k-dimensional subspace is infinite-dimensional. However, we can impose the stronger condition that *every* translation of the subspace intersects S in a set of full Lebesgue measure. A preliminary notion of prevalence is obtained by declaring that a Borel set $S \subset B$ is prevalent if there exists some finite k and some k-dimensional subspace V such that every translation of V intersects S in a set of full k-dimensional Lebesgue measure. In order to ensure that a countable intersection of prevalent sets is prevalent, we must enlarge the space of measures under consideration beyond Lebesgue measure supported on finite-dimensional subspaces.

Definition 2.4. A Borel set $S \subset B$ is said to be prevalent if there exists a measure μ on B such that

1. $0 < \mu(C) < \infty$ for some compact subset C of B, and
2. the set $S - x$ has full μ-measure (that is, the complement of $S - x$ has measure 0) for all $x \in B$.

A non-Borel set that contains a prevalent Borel set is also prevalent.

The measure μ may be a Lebesgue measure on a finite-dimensional subspace of B. More generally, one may think of μ as describing a family of perturbations in B. In this sense, S is prevalent if for all $x \in B$, choosing a perturbation at random with respect to μ and adding it to x yields a point in S with probability one. Prevalent sets share several of the desirable properties of residual sets. A prevalent subset of B is dense and the countable intersection of prevalent sets is prevalent. See [33] for details. One may formulate a notion of prevalence appropriate for spaces without a linear structure [36]. This notion applies to the space of diffeomorphisms of a compact smooth manifold.

2.2.2 The Dimension Spectrum

Let μ be a Borel probability measure on a metric space X. For $q \geqslant 0$ and $\epsilon > 0$ define

$$C_q(\mu, \epsilon) = \int_X [\mu(B(x, \epsilon))]^{q-1} \, d\mu(x)$$

where $B(x, \epsilon)$ is the open ball of radius ϵ centered at x.

Definition 2.5. For $q \geqslant 0$, $q \neq 1$, the lower and upper q-dimensions of μ are

$$D_q^-(\mu) = \liminf_{\epsilon \to 0} \frac{\log C_q(\mu, \epsilon)}{(q-1)\log(\epsilon)},$$
$$D_q^+(\mu) = \limsup_{\epsilon \to 0} \frac{\log C_q(\mu, \epsilon)}{(q-1)\log(\epsilon)}.$$

If $D_q^-(\mu) = D_q^+(\mu)$, their common value is denoted $D_q(\mu)$ and is called the q-dimension of μ.

For a measure μ such that $D_q(\mu)$ exists, the function $q \to D_q(\mu)$ is called the dimension spectrum of μ. For $q = 0$, D_0^- and D_0^+ depend only on X. We write $D_0^-(X)$ and $D_0^+(X)$ for the lower and upper 0-dimensions of X. For $\epsilon > 0$, let $n(X, \epsilon)$ be the minimum number of ϵ-balls required to cover X. Written in terms of $n(X, \epsilon)$, $D_0^-(X)$ and $D_0^+(X)$ are given by

$$D_0^-(X) = \liminf_{\epsilon \to 0} \frac{\log n(X, \epsilon)}{\log(1/\epsilon)},$$
$$D_0^+(X) = \limsup_{\epsilon \to 0} \frac{\log n(X, \epsilon)}{\log(1/\epsilon)}.$$

The values $D_0^-(X)$ and $D_0^+(X)$ are therefore equal to the lower and upper box-counting dimensions of X, respectively.

For measures on $\mathbb{R}^n$, one encounters the following alternative definition of the dimension spectrum [27, 28, 54]. For $\epsilon > 0$, cover the support of μ with a grid of cubes with edge length ϵ. Let $N(\epsilon)$ be the number of cubes that intersect the support of μ, and let the measure of these cubes be $p_1, p_2, \ldots, p_{N(\epsilon)}$. Write

$$D_q^-(\mu) = \liminf_{\epsilon \to 0} \frac{\sum_{i=1}^{N(\epsilon)} p_i^q}{(q-1)\log(\epsilon)},$$
$$D_q^+(\mu) = \limsup_{\epsilon \to 0} \frac{\sum_{i=1}^{N(\epsilon)} p_i^q}{(q-1)\log(\epsilon)}.$$

For $q \geqslant 0$, $q \neq 1$, these limits are independent of the choice of ϵ-grids, and give the same values as Definition 2.5. See [55] for a proof of this equivalence for $q > 1$.

The grid definition of the dimension spectrum is not appropriate for measures on general metric spaces. We therefore adopt Definition 2.5 as the natural notion in the general case.

A potential-theoretic definition of the lower q-dimension $D_q^-(\mu)$ for $q > 1$ is introduced in [30]. For $s \geqslant 0$ the s-potential of the measure μ at the point x is given by

$$\varphi_s(\mu, x) = \int_X |x-y|^{-s}\, d\mu(y).$$

Definition 2.6. The (s,q)-energy of μ, denoted $I_{s,q}(\mu)$, is given by

$$I_{s,q}(\mu) = \int_X [\varphi_s(\mu, x)]^{q-1}\, d\mu(x) = \int_X \left(\int_X \frac{d\mu(y)}{|x-y|^s} \right)^{q-1} d\mu(x).$$

For $q = 2$, the (s,q)-energy of μ reduces to the more standard notion of the s-energy of μ, written

$$I_s(\mu) = \int_X \varphi_s(\mu, x)\, d\mu(x) = \int_X \int_X \frac{d\mu(x)d\mu(y)}{|x-y|^s}.$$

Sauer and Yorke [58] show that the lower correlation dimension $D_2^-(\mu)$ can be expressed as

$$D_2^-(\mu) = \sup\{s : I_s(\mu) < \infty\}. \tag{2.5}$$

This characterization of $D_2^-(\mu)$ is used to establish the preservation of correlation dimension. The following proposition generalizes (2.5) to the lower-dimension spectrum for $q > 1$.

Proposition 2.7 ([30]). *If $q > 1$ and μ is a Borel probability measure, then*

$$D_q^-(\mu) = \sup\{s \geqslant 0 : I_{s,q}(\mu) < \infty\}.$$

2.2.3 The Thickness Exponent

Let B denote a Banach space.

Definition 2.8. The thickness exponent $\tau(X)$ of a compact set $X \subset B$ is defined as follows. Let $d(X, \epsilon)$ be the minimum dimension of all finite-dimensional subspaces $V \subset B$ such that every point of X lies within ϵ of V; if no such V exists, then $d(X, \epsilon) = \infty$. Let

$$\tau(X) = \limsup_{\epsilon \to 0} \frac{\log d(X, \epsilon)}{\log(1/\epsilon)}.$$

There is no general relationship between the thickness exponent and the Hausdorff dimension. A finite-dimensional disk has thickness exponent zero but can have arbitrarily high dimension. A countable set, which necessarily has Hausdorff dimension zero, can have positive thickness. For example, one can show that the compact subset $\{0, e_2/\log 2, e_3/\log 3, \ldots\}$ of the real Hilbert space with basis $\{e_1, e_2, \ldots\}$ has an infinite thickness exponent. A definitive statement may be made concerning the box-counting dimension D_0.

Lemma 2.9 ([31]). *Let $X \subset B$ be a compact set. Then $\tau(X) \leqslant D_0^+(X)$.*

Proof. Recall that the box-counting dimension $D_0^+(X)$ may be expressed similarly to $\tau(X)$, but in terms of the minimum number of $n(X, \epsilon)$ of ϵ-balls required to cover X. For any such cover, X lies within ϵ of the space spanned by the centers of the balls. Thus $d(X, \epsilon) \leqslant n(X, \epsilon)$ for each $\epsilon > 0$, and the desired inequality follows. $\square$

2.3 Main Results

We begin with the main results for general Banach spaces.

Theorem 2.10. *Let B be a Banach space, and let M be any subspace of the Borel measurable functions from B to $\mathbb{R}^m$ that contains the bounded linear functions. Let $X \subset B$ be a compact set with thickness exponent $\tau(X)$, and let μ be a Borel probability measure supported on X. For almost every function $f \in M$,*

$$D_q^-(f(\mu)) \geqslant \min\left\{m, \frac{D_q^-(\mu)}{1+\tau(X)}\right\}$$

for all $q \in (1, 2]$.

Corollary 2.11. *Assume in addition that M is contained in the space of locally Lipschitz functions, that $\tau(X) = 0$, and that $D_q(\mu)$ exists ($D_q^-(\mu) = D_q^+(\mu)$) for all $q \in (1, 2]$. Then for almost every function $f \in M$, $D_q(f(\mu))$ exists and equals $\min\{m, D_q(\mu)\}$ for all $q \in (1, 2]$.*

Remark 2.12. For $r \geqslant 1$, the space $M = C^r(B, \mathbb{R}^m)$ satisfies the hypotheses of Theorem 2.10 and Corollary 2.11.

The corollary follows immediately from Theorem 2.10 and the fact that for all μ and all locally Lipschitz f, $D_q^+(f(\mu)) \leqslant \min\{m, D_q^+(\mu)\}$.

Corollary 2.13. *Let B be a Banach space. Let $X \subset B$ be a compact set with thickness exponent $\tau(X)$. For almost every function $f \in M$,*

$$\dim_H(f(X)) \geqslant \min\left\{m, \frac{\dim_H(X)}{1+\tau(X)}\right\}. \tag{2.6}$$

Proof. Let $\mathcal{M}(X)$ denote the set of Borel probability measures on X. The Hausdorff dimension of X may be expressed in terms of the lower correlation dimension of measures supported on X via the variational principle [22]

$$\dim_H(X) = \sup_{\mu \in \mathcal{M}(X)} D_2^-(\mu).$$

For each $i \in \mathbb{N}$, there exists $\mu_i \in \mathcal{M}(X)$ such that $D_2^-(\mu_i) > \dim_H(X) - 1/i$. Applying Theorem 2.10, there exists a prevalent set $P_i \subset M$ of functions such that for $f \in P_i$,

$$D_2^-(f(\mu_i)) \geqslant \min\left\{m, \frac{D_2^-(\mu_i)}{1+\tau(X)}\right\}.$$

The set $\bigcap_{i=1}^{\infty} P_i$ is prevalent. For $f \in \bigcap_{i=1}^{\infty} P_i$, the bound (2.6) follows from the variational principle. $\square$

Remark 2.14. No analogue of Corollary 2.13 holds for the box-counting dimension. Let $n > m$ be integers and let $d \leqslant m$. Sauer and Yorke [58] construct a compact set $A \subset \mathbb{R}^n$ such that $D_0^+(A) = d$ and $D_0^+(f(A)) < d$ for every $f \in C^1(\mathbb{R}^n, \mathbb{R}^m)$.

Proof of Theorem 2.10. Fix $1 < q \leqslant 2$. Let $L \subset M$ denote the space of bounded linear functions from B into $\mathbb{R}^m$. We construct a 'Banach brick' $Q \subset L$ of perturbations and a probability measure λ on Q. For $f \in M$ and $\pi \in Q$, write $f_\pi = f + \pi$. Utilizing the potential-theoretic description of $D_q^-(\mu)$ for $1 < q \leqslant 2$, we must show that for any $f \in M$, $t > 0$, and $0 \leqslant s < \min\{m, t/(1+\tau(X))\}$,

$$I_{t,q}(\mu) < \infty \Rightarrow I_{s,q}(f_\pi(\mu)) < \infty \tag{2.7}$$

for λ-almost every $\pi \in Q$. The result follows because we can choose t arbitrarily close to $D_q^-(\mu)$.

We define the Banach brick Q as follows. For $j \in \mathbb{N}$, let $d_j = d(X, 2^{-j})$ and let $V_j \subset B$ be a subspace of dimension d_j such that every point of X lies within 2^{-j} of V_j. Fix $\sigma > \tau(X)$. By Definition 2.8 of $\tau(X)$, there exists $C_1 > 0$, depending only on X and σ, such that $d_j \leqslant C_1 2^{j\sigma}$. Let S_j be the closed unit ball in the dual space V_j' of V_j. There is no natural embedding of V_j' into B', but it follows from the Hahn-Banach theorem that there exists an isometric embedding of V_j'

into B'. As such, we can think of S_j as a subset of B'. On the other hand, V_j' is linearly isomorphic to $\mathbb{R}^{d_j}$, and S_j corresponds to a convex set $U_j \subset \mathbb{R}^{d_j}$. The uniform (Lebesgue) probability measure on U_j induces a measure λ_j on S_j. Define the Banach brick Q by

$$Q = \left\{ \pi = (\pi_1, \ldots, \pi_m) : \pi_i = \sum_{j=1}^{\infty} j^{-2} \phi_{ij} \text{ with } \phi_{ij} \in S_j \ \forall j \right\}.$$

Since each $S_j \subset B'$ is compact, $Q \subset L$ is compact. Let λ be the probability measure on Q that results from choosing the elements ϕ_{ij} randomly and independently with respect to the measures λ_j on the sets S_j. (While the term "brick" suggests that Q is the product of compact sets $j^{-2}S_j$ that are all transverse to each other, these sets may have nontrivial intersection, in which case Q and λ are still well-defined.)

Choose $\rho > \sigma > \tau(X)$. We will show that for $0 \leqslant s < m$,

$$I_{s(1+\rho),q}(\mu) < \infty \Rightarrow I_{s,q}(f_\pi(\mu)) < \infty$$

for λ-almost every $\pi \in Q$. Since ρ and σ can be arbitrarily close to $\tau(X)$, this implies (2.7). Computing the (s, q)-energy of $f_\pi(\mu)$, we have

$$\begin{aligned} I_{s,q}(f_\pi(\mu)) &= \int_{\mathbb{R}^m} \left[\int_{\mathbb{R}^m} \frac{df_\pi(\mu)(v)}{|u-v|^s} \right]^{q-1} df_\pi(\mu)(u) \\ &= \int_B \left[\int_B \frac{d\mu(y)}{|f_\pi(x) - f_\pi(y)|^s} \right]^{q-1} d\mu(x). \end{aligned}$$

Integrating the energy over Q and using the Fubini/Tonelli theorem and the fact that $0 < q-1 \leqslant 1$, we have

$$\begin{aligned}
\int_Q I_{s,q}(f_\pi(\mu))\,d\lambda(\pi) &= \int_Q \int_B \left[\int_B \frac{d\mu(y)}{|f_\pi(x)-f_\pi(y)|^s}\right]^{q-1} d\mu(x)d\lambda(\pi) \\
&= \int_B \int_Q \left[\int_B \frac{d\mu(y)}{|f_\pi(x)-f_\pi(y)|^s}\right]^{q-1} d\lambda(\pi)d\mu(x) \\
&\leqslant \int_B \left[\int_Q \int_B \frac{d\mu(y)}{|f_\pi(x)-f_\pi(y)|^s}\,d\lambda(\pi)\right]^{q-1} d\mu(x) \\
&= \int_B \left[\int_B \left(\int_Q \frac{d\lambda(\pi)}{|f_\pi(x)-f_\pi(y)|^s}\right) d\mu(y)\right]^{q-1} d\mu(x).
\end{aligned}$$

We now estimate the interior integral.

Lemma 2.15 (Banach Perturbation Lemma). *If $s < m$, there exists a constant C_2 depending only on s, σ, and ρ, such that for all x, $y \in X$,*

$$\int_Q \frac{d\lambda(\pi)}{|f_\pi(x)-f_\pi(y)|^s} \leqslant \frac{C_2}{\min\{|x-y|,1\}^{s(1+\rho)}}.$$

Proof. Set $\zeta = \min\{|x-y|,1\}$. Choose $j \in \mathbb{N}$ such that $2-\log_2\zeta \leqslant j \leqslant 3-\log_2\zeta$. There exist points $\gamma_j(x)$ and $\gamma_j(y)$ in V_j satisfying $|x-\gamma_j(x)| \leqslant 2^{-j}$ and $|y-\gamma_j(y)| \leqslant 2^{-j}$. Estimating the distance between $\gamma_j(x)$ and $\gamma_j(y)$, we have

$$|\gamma_j(x)-\gamma_j(y)| \geqslant |x-y| - 2^{-j+1} \geqslant |x-y| - \frac{\zeta}{2} \geqslant \frac{|x-y|}{2}.$$

For $\pi \in Q$, write $\pi = \xi_j + j^{-2}\phi_j$ where $\phi_j = (\phi_{1j},\ldots,\phi_{mj}) \in S_j^m$ and $\xi_j = (\xi_{1j},\ldots,\xi_{mj})$ with

$$\xi_{ij} = \sum_{\substack{k\in\mathbb{N}\\ k\neq j}} k^{-2}\phi_{ik}$$

for each i. We fix ξ_j and integrate over $\phi_j \in S_j^m$. We have

$$\begin{aligned}
&\int_{S_j^m} \frac{d\lambda_j^m(\phi_j)}{|f_{\xi_j+j^{-2}\phi_j}(x) - f_{\xi_j+j^{-2}\phi_j}(y)|^s} \\
&= \int_{S_j^m} \frac{d\lambda_j^m(\phi_j)}{|f_{\xi_j}(x) - f_{\xi_j}(y) + j^{-2}\phi_j(x-y)|^s} \\
&\leqslant \int_{S_j^m} \frac{d\lambda_j^m(\phi_j)}{|j^{-2}\phi_j(x-y)|^s} \\
&= j^{2s} \int_{S_j^m} \frac{d\lambda_j^m(\phi_j)}{|\phi_j(x-y)|^s}.
\end{aligned}$$

Let $P \subset B'$ be the annihilator of $x - y$. By the Hahn-Banach theorem, there exists $\psi \in B'$ such that $\psi(x-y) = |x-y|$ and $\|\psi\|_{B'} = 1$. By restricting P and ψ to V_j, we may think of them as belonging to V_j', and hence also to $\mathbb{R}^{d_j}$. Notice that

$$\frac{|\psi(\gamma_j(x) - \gamma_j(y))|}{|\gamma_j(x) - \gamma_j(y)|} \geqslant \frac{|x-y| - \zeta/2}{|x-y| + \zeta/2} \geqslant \frac{|x-y|/2}{3|x-y|/2} = \frac{1}{3},$$

so $\|\psi\|_{V_j'} \geqslant \frac{1}{3}$. Let b be such that $\|b\psi\|_{V_j'} = 1$ and set $\widetilde{\psi} = b\psi$. By convexity, S_j contains the cones with base $P \cap S_j$ and vertices $\widetilde{\psi}$ and $-\widetilde{\psi}$. Let C_j be the union of this pair of cones and let $\widetilde{\lambda}_j$ denote the restriction of λ_j to C_j. We show that

$$j^{2s} \int_{S_j^m} \frac{d\lambda_j^m(\phi_j)}{|\phi_j(x-y)|^s} \leqslant j^{2s} \left(\frac{\int_{C_j^m} \frac{d\widetilde{\lambda}_j^m(\phi_j)}{|\phi_j(x-y)|^s}}{\int_{C_j^m} d\widetilde{\lambda}_j^m(\phi_j)} \right). \tag{2.8}$$

Let $S_j^+ = \{\gamma \in S_j : \gamma = p + t\widetilde{\psi}$ for some $p \in P$ and some $t \geqslant 0\}$. Let C_j^+ be the cone with base $P \cap S_j$ and vertex $\widetilde{\psi}$. We define functions $g : [0,1] \to \mathbb{R}$ and $h : [0,1] \to \mathbb{R}$ giving the normalized volumes of slices of S_j^+ and C_j^+. For $t \in [0,1]$, let $g(t)$ be the $(d_j - 1)$-dimensional volume of $S_j^+ \cap (P + t\widetilde{\psi})$ normalized by $\int_{S_j^+} d\lambda_j(\gamma)$, and let $h(t)$ be the (d_j-1)-dimensional volume of $C_j^+ \cap (P + t\widetilde{\psi})$ normalized by $\int_{C_j^+} d\widetilde{\lambda}_j(\gamma)$. Since $\int_0^1 g(t)\,dt = \int_0^1 h(t)\,dt = 1$ and $g(1) \geqslant h(1)$, there exists $c \in (0,1)$ such that $g(c) = h(c)$. It follows from the Brunn-Minkowski inequality [25] that the function $g^{1/(d_j-1)}$ is concave. The function $h^{1/(d_j-1)}$ is linear, so we must have $g(t) \leqslant h(t)$ for $t \leqslant c$ and $g(t) \geqslant h(t)$ for $t \geqslant c$. Observe that if $k : [0,1] \to \mathbb{R}$ is decreasing, then

$$\int_0^1 h(t)k(t)\,dt \geqslant \int_0^1 g(t)k(t)\,dt.$$

Verifying this, we have

$$\begin{aligned}\int_0^c (h(t)-g(t))k(t)\,dt &\geqslant k(c)\int_0^c (h(t)-g(t))\,dt \\ &= k(c)\int_c^1 (g(t)-h(t))\,dt \\ &\geqslant \int_c^1 (g(t)-h(t))k(t)\,dt.\end{aligned}$$

The inequality (2.8) follows from the bound

$$\begin{aligned}&\int_0^1 \cdots \int_0^1 |\alpha|^{-s} g(\alpha_1)\cdots g(\alpha_m)\,d\alpha_1\cdots d\alpha_m \\ &\leqslant \int_0^1 \cdots \int_0^1 |\alpha|^{-s} h(\alpha_1)\cdots h(\alpha_m)\,d\alpha_1\cdots d\alpha_m,\end{aligned} \tag{2.9}$$

where $\alpha = (\alpha_1, \ldots, \alpha_m)$. The bound (2.9) follows from the observation because $|\alpha|^{-s}$ is decreasing in each of its m arguments. Let W_j be the right side of (2.8). In order to estimate W_j, we use the $(P, \widetilde{\psi})$ foliation given by

$$C_{j,i} = \{C_{j,i} \cap (P + \alpha_i \widetilde{\psi}) : \alpha_i \in [-1,1]\}$$

for each $i = 1, \ldots, m$.

Lemma 2.16 (Banach Integral Asymptotics). *Let $m \in \mathbb{N}$ and $s < m$. There exists a constant K, independent of $n \in \mathbb{N}$, such that*

$$\frac{\int_0^1 \cdots \int_0^1 \frac{(1-\alpha_1)^{n-1}\cdots(1-\alpha_m)^{n-1}}{|\alpha|^s}\,d\alpha_1\cdots d\alpha_m}{\int_0^1 \cdots \int_0^1 (1-\alpha_1)^{n-1}\cdots(1-\alpha_m)^{n-1}\,d\alpha_1\cdots d\alpha_m} \leqslant Kn^s, \tag{2.10}$$

where $\alpha = (\alpha_1, \ldots, \alpha_m)$.

Proof. Since $e^{-z} \geqslant 1 - z$ for all real z, and the denominator of (2.10) is n^{-m}, the ratio of integrals in (2.10) is bounded above by

$$n^m \int_0^\infty \cdots \int_0^\infty \frac{\exp\left(-\sum_{i=1}^m \alpha_i(n-1)\right)}{|\alpha|^s}\, d\alpha_1 \cdots d\alpha_m.$$

Setting $u_i = \alpha_i(n-1)$, this becomes

$$n^m(n-1)^{s-m} \int_0^\infty \cdots \int_0^\infty \frac{\exp\left(-\sum_{i=1}^m u_i\right)}{|u|^s}\, du_1 \cdots du_m.$$

Since $|u|^{-s}$ is integrable in a neighborhood of 0 for $s < m$, the lemma is established. □

We are now in position to complete the proof of Lemma 2.15. Estimating the ratio of integrals in W_j using the $(P, \widetilde{\psi})$ foliation, it follows from Lemma 2.16 with $n = d_j \leqslant C_1 2^{j\sigma}$ that there exists K, independent of j, such that

$$\begin{aligned}
W_j &\leqslant j^{2s} K |x-y|^{-s} \left(C_1 2^{j\sigma}\right)^s \\
&\leqslant K C_1^s j^{2s} |x-y|^{-s} \left(2^j\right)^{\sigma s} \\
&\leqslant K C_1^s j^{2s} |x-y|^{-s} \left(8\zeta^{-1}\right)^{\sigma s} \\
&\leqslant 8^{\sigma s} K C_1^s j^{2s} \zeta^{-s(1+\sigma)} \\
&\leqslant 8^{\sigma s} K C_1^s \left(3 - \log_2 \zeta\right)^{2s} \zeta^{-s(1+\sigma)}.
\end{aligned}$$

Thus, since $\rho > 0$, there exists C_2 such that

$$W_j \leqslant \frac{C_2}{\zeta^{s(1+\rho)}}.$$

We have established that

$$\int_{S_j^m} \frac{d\lambda_j^m(\phi_j)}{|f_{\xi_j + j^{-2}\phi_j}(x) - f_{\xi_j + j^{-2}\phi_j}(y)|^s} \leqslant \frac{C_2}{\zeta^{s(1+\rho)}}$$

for all ξ_j, and hence by integrating over ξ_j that

$$\int_Q \frac{d\lambda(\pi)}{|f_\pi(x) - f_\pi(y)|^s} \leqslant \frac{C_2}{\zeta^{s(1+\rho)}}.$$

The proof of the perturbation lemma is complete. □

Returning to the proof of Theorem 2.10, recall that $0 \leqslant s < \min\{m, t/(1+\tau(X))\}$ and $\rho > \sigma > \tau(X)$ have been fixed. Applying the perturbation lemma, we have

$$\begin{aligned}\int_Q I_{s,q}(f_\pi(\mu))\, d\lambda(\pi) &\leqslant \int_B \left[\int_B \left(\int_Q \frac{d\lambda(\pi)}{|f_\pi(x) - f_\pi(y)|^s}\right) d\mu(y)\right]^{q-1} d\mu(x) \\ &\leqslant \int_B \left[\int_B \frac{C_2}{\min\{|x-y|, 1\}^{s(1+\rho)}}\, d\mu(y)\right]^{q-1} d\mu(x).\end{aligned}$$

Therefore,

$$I_{s(1+\rho),q}(\mu) < \infty \Rightarrow I_{s,q}(f_\pi(\mu)) < \infty$$

for λ-almost every $\pi \in Q$. Since ρ and σ can be arbitrarily close to $\tau(X)$, this implies (2.7) for fixed t. Because we can choose t arbitrarily close to $D_q^-(\mu)$, there exists a prevalent set $P_q \subset M$ such that for $f \in P_q$,

$$D_q^-(f(\mu)) \geqslant \min\left\{m, \frac{D_q^-(\mu)}{1+\tau(X)}\right\}.$$

Let $\{q_i\}$ be a countable dense subset of $(1, 2]$. The set $\bigcap_{i=1}^\infty P_{q_i}$ is prevalent. For $f \in \bigcap_{i=1}^\infty P_{q_i}$, the continuity of D_q^- on $(1, 2]$ implies that

$$D_q^-(f(\mu)) \geqslant \min\left\{m, \frac{D_q^-(\mu)}{1+\tau(X)}\right\}$$

for all $1 < q \leqslant 2$. □

The proof of the perturbation lemma uses only the convexity of S_j. In specific cases, additional information about the structure of the dual space may lead to an improved perturbation lemma and hence to an improvement of the factor $1/(1+\tau(X))$. We establish such an improvement for Hilbert spaces.

Theorem 2.17. *Let H be a Hilbert space, and let M be any subspace of the Borel measurable functions from H to $\mathbb{R}^m$ that contains the bounded linear functions.*

Let $X \subset H$ be a compact set with thickness exponent $\tau(X)$, and let μ be a Borel probability measure supported on X. For almost every function $f \in M$,

$$D_q^-(f(\mu)) \geqslant \min\left\{m, \frac{D_q^-(\mu)}{1+\tau(X)/2}\right\}$$

for all $q \in (1, 2]$.

Corollary 2.18. *Let H be a Hilbert space. Let $X \subset H$ be a compact set with thickness exponent $\tau(X)$. For almost every function $f \in M$,*

$$\dim_H(f(X)) \geqslant \min\left\{m, \frac{\dim_H(X)}{1+\tau(X)/2}\right\}.$$

Remark 2.19. For the example from [31] discussed in the introduction, this Hausdorff dimension estimate is sharp.

Proof of Theorem 2.17. Let $L \subset M$ denote the space of bounded linear functions from H into $\mathbb{R}^m$. We must show that for any $f \in M$ and $0 \leqslant s < \min\{m, t/(1+\tau(X)/2)\}$,

$$I_{t,q}(\mu) < \infty \Rightarrow I_{s,q}(f_\pi(\mu)) < \infty$$

for λ-almost every $\pi \in Q$. The construction of the Hilbert brick Q follows that of the Banach brick. Notice that each S_j is isometric to a Euclidean ball. The dual space V_j' embeds canonically into $H' = H$: an element of V_j' acts on an element of H by composition with the orthogonal projection onto V_j. Let $\rho > \sigma > \tau(X)$. We will show that for $0 \leqslant s < m$,

$$I_{s(1+\rho),q}(\mu) < \infty \Rightarrow I_{s,q}(f_\pi(\mu)) < \infty$$

for λ-almost every $\pi \in Q$. The proof of this implication follows the argument given in the proof of Theorem 2.10. We only need to apply the following improved perturbation lemma. □

Lemma 2.20 (Hilbert Perturbation Lemma). *If $s < m$, there exists a constant C_3, depending only on s, σ, and ρ, such that for all x, $y \in X$,*

$$\int_Q \frac{d\lambda(\pi)}{|f_\pi(x) - f_\pi(y)|^s} \leqslant \frac{C_3}{\min\{|x-y|, 1\}^{s(1+\rho/2)}}.$$

Proof. Set $\zeta = \min\{|x-y|, 1\}$. Select j as before and note that

$$\int_{S_j^m} \frac{d\lambda_j^m(\phi_j)}{|f_{\xi_j + j^{-2}\phi_j}(x) - f_{\xi_j + j^{-2}\phi_j}(y)|^s} \leqslant j^{2s} \int_{S_j^m} \frac{d\lambda_j^m(\phi_j)}{|\phi_j(\gamma_j(x) - \gamma_j(y))|^s}.$$

Lemma 2.21 (Hilbert Integral Asymptotics). *There exists $K > 0$, independent of $n \in \mathbb{N}$, such that for $s < m$,*

$$\frac{\int_0^1 \cdots \int_0^1 \frac{(1-\alpha_1^2)^{\frac{n-1}{2}} \cdots (1-\alpha_m^2)^{\frac{n-1}{2}} \, d\alpha_1 \cdots d\alpha_m}{|\alpha|^s}}{\int_0^1 \cdots \int_0^1 (1-\alpha_1^2)^{\frac{n-1}{2}} \cdots (1-\alpha_m^2)^{\frac{n-1}{2}} \, d\alpha_1 \cdots d\alpha_m} \leqslant K n^{\frac{s}{2}}.$$

Proof. The proof is similar to that of Lemma 2.16 and is left to the reader. □

Let P be the annihilator of $\gamma_j(x) - \gamma_j(y)$ in V_j'. Foliating S_j into leaves parallel to P and using Lemma 2.21 with $n = d_j \leqslant C_1 2^{j\sigma}$, we have

$$\begin{aligned}
&j^{2s} \int_{S_j^m} \frac{d\lambda_j^m(\phi_j)}{|\phi_j(\gamma_j(x) - \gamma_j(y))|^s} \\
&\leqslant K j^{2s} |\gamma_j(x) - \gamma_j(y)|^{-s} \left(C_1 2^{j\sigma}\right)^{s/2} \\
&\leqslant 2^s K j^{2s} |x-y|^{-s} \left(C_1 2^{j\sigma}\right)^{s/2} \\
&\leqslant 2^s K C_1^{s/2} j^{2s} |x-y|^{-s} \left(2^j\right)^{\sigma s/2} \\
&\leqslant 2^s K C_1^{s/2} j^{2s} |x-y|^{-s} \left(8\zeta^{-1}\right)^{\sigma s/2} \\
&\leqslant 2^s 8^{\sigma s/2} K C_1^{s/2} j^{2s} \zeta^{-s(1+\sigma/2)} \\
&\leqslant C_3 \zeta^{-s(1+\rho/2)}
\end{aligned}$$

for some $C_3 > 0$. We have established that

$$\int_{S_j^m} \frac{d\lambda_j^m(\phi_j)}{|f_{\xi_j + j^{-2}\phi_j}(x) - f_{\xi_j + j^{-2}\phi_j}(y)|^s} \leqslant \frac{C_3}{\zeta^{s(1+\rho/2)}}$$

for all ξ_j, and hence by integrating over ξ_j that

$$\int_Q \frac{d\lambda(\pi)}{|f_\pi(x) - f_\pi(y)|^s} \leqslant \frac{C_3}{\zeta^{s(1+\rho/2)}}.$$

The proof of the perturbation lemma is complete. □

2.4 Nonpreservation of Hausdorff Dimension

Theorems 2.10 and 2.17 are sharp in the following sense. Given $d > 0$, $1 \leqslant p \leqslant \infty$, and a positive integer m, there is a compact subset X of Hausdorff dimension d in ℓ^p such that for all bounded linear functions $\pi : \ell^p \to \mathbb{R}^m$,

$$\dim_H(\pi(X)) \leqslant \frac{d}{1 + d/q},$$

where $q = p/(p-1)$ [31]. The cases $p = \infty$ and $p = 2$ show respectively that Theorems 2.10 and 2.17 are sharp for bounded linear functions on these particular Banach spaces. On the other hand, this class of examples does not rule out a dimension preservation result in ℓ^1.

Here we construct a compact subset X of Hausdorff dimension d in ℓ^1 such that for all bounded linear functions $\pi : \ell^1 \to \mathbb{R}$,

$$\dim_H(\pi(X)) \leqslant \frac{d}{1 + d/2}.$$

Let $\{e_i\}$ be the standard basis of ℓ^1, and let $\lambda = 2^{-1/d}$. Consider the inductively constructed sets X_k, defined as follows. Let $X_0 = \{0\}$ and $X_1 = \{\pm p\}$, where

$$p = \frac{1}{2}(e_1 - e_2).$$

For the next step, construct the two points

$$p_0 = \frac{\lambda}{4}(e_3 - e_4 + e_5 - e_6), \text{ and}$$
$$p_1 = \frac{\lambda}{4}(e_3 + e_4 - e_5 - e_6).$$

Attach these points to the nodes of X_1, forming the set

$$X_2 = \{p \pm p_0,\ -p \pm p_1\}.$$

We now describe the construction of X_{k+1} given X_k. Let

$$\alpha_k = 1 + \sum_{i=0}^{k-1} 2^{2^i}.$$

Define the collection of 2^k points

$$\left\{p_{\beta_1\beta_2\cdots\beta_k} : \beta_1, \beta_2, \ldots, \beta_k \in \{0,1\}\right\}$$

by setting

$$p_{\beta_1\beta_2\cdots\beta_k} = \frac{\lambda^k}{2^{2^k}} \sum_{i=0}^{2^{2^k}-1} (-1)^{\left[2^{-\gamma_{\beta_1\cdots\beta_k}}\cdot i\right]} e_{\alpha_k+i},$$

where $\gamma_{\beta_1\cdots\beta_k}$ is the integer in $[0, 2^k)$ whose binary representation is $\beta_1\cdots\beta_k$; that is,

$$\gamma_{\beta_1\cdots\beta_k} = \beta_1 2^{k-1} + \beta_2 2^{k-2} + \cdots + \beta_k.$$

Notice that $\|p_{\beta_1\cdots\beta_k}\|_{\ell^1} = \lambda^k$. Attach these points to the nodes of X_k, forming

$$X_{k+1} = \left\{(-1)^{\beta_1} p + (-1)^{\beta_2} p_{\beta_1} + \cdots + (-1)^{\beta_{k+1}} p_{\beta_1\cdots\beta_k} : \beta_1, \ldots, \beta_{k+1} \in \{0,1\}\right\}.$$

Figure 2.1 illustrates the third step in the construction. Let X be the set of all limit points of

$$\bigcup_{k=0}^{\infty} X_k.$$

Equivalently,

$$X = \left\{(-1)^{\beta_1} p + (-1)^{\beta_2} p_{\beta_1} + (-1)^{\beta_3} p_{\beta_1\beta_2} + \cdots : \beta_1, \beta_2, \beta_3, \ldots \in \{0,1\}\right\}.$$

Proposition 2.22. *For the set $X \subset \ell^1$ constructed above,*

$$\dim_H(X) = D_0^+(X) = \frac{\log 2}{\log(1/\lambda)} = d.$$

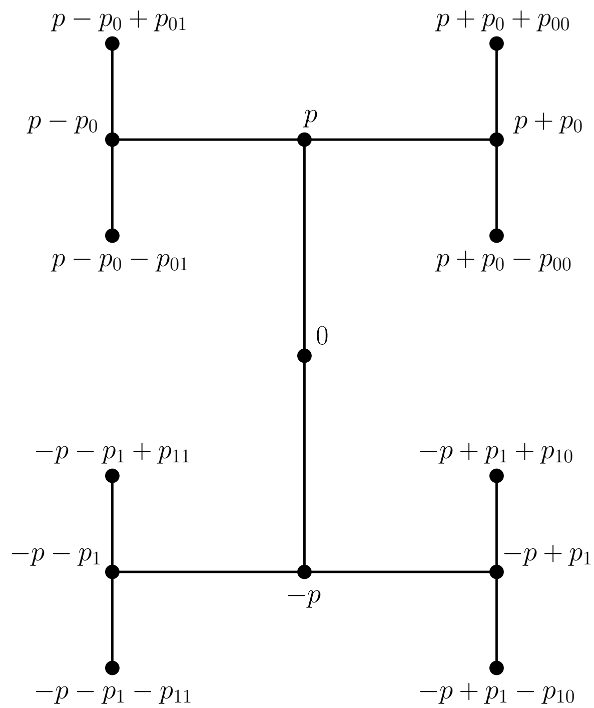

Figure 2.1: The sets X_0, X_1, X_2, and X_3 consist of the nodes of the binary tree above.

Proof. The set X can be covered by 2^k balls of radius $\lambda^k/(1-\lambda)$ centered at the points of X_k, so $\dim_H(X) \leqslant D_0^+(X) \leqslant d$. To show that $\dim_H(X) \geqslant d$, we apply Frostman's lemma [22, 47]. The binary tree X may be identified with the set of binary strings $S = \big\{\beta = \beta_1\beta_2\beta_3\cdots : \beta_1, \beta_2, \beta_3, \ldots \in \{0,1\}\big\}$. Consider the measure μ on X induced by the uniform probability measure on S. Since every two points in X corresponding to different initial strings $\beta_1\cdots\beta_k\beta_{k+1}$ and $\beta_1\cdots\beta_k\beta'_{k+1}$ must lie at least $2\lambda^k$ apart, the measure of a ball of radius less than λ^k is at most the measure of all strings in S starting with a given $\beta_1\cdots\beta_{k+1}$, which is $2^{-(k+1)} = (\lambda^k)^d/2$. By Frostman's lemma, $\dim_H(X) \geqslant d$. $\square$

Proposition 2.23. *For every bounded linear map $\pi : \ell^1 \to \mathbb{R}$,*

$$\dim_H(\pi(X)) \leqslant \frac{d}{1+d/2}.$$

Proof. Let $s = d/(1+d/2) = (1/d+1/2)^{-1}$. Let $\pi \in \ell^\infty$ and assume $\|\pi\|_{\ell^\infty} = 1$. We will show for each $k \geqslant 0$ that $\pi(X)$ can be covered by a collection of 2^k intervals $\mathcal{C}_k = \{I_0, I_1, \ldots, I_{2^k-1}\}$ such that

$$\lim_{k\to\infty} \max_{I\in\mathcal{C}_k} \operatorname{diam}(I) = 0$$

and

$$\sum_{j=0}^{2^k-1} \operatorname{diam}(I_j)^s$$

remains bounded as $k \to \infty$. It then follows that the s-dimensional Hausdorff measure of $\pi(X)$ is finite, and therefore that the Hausdorff dimension of $\pi(X)$ is at most s, as desired. The proposition is trivially true if $s \geqslant 1$, so assume henceforth that $s < 1$. Then by convexity,

$$2^{-k} \sum_{j=0}^{2^k-1} \operatorname{diam}(I_j)^s \leqslant \left(2^{-k} \sum_{j=0}^{2^k-1} \operatorname{diam}(I_j)\right)^s,$$

so it suffices to show that

$$\begin{aligned} 2^{-k} \sum_{j=0}^{2^k-1} \operatorname{diam}(I_j) &\leqslant C_4 2^{-k/s} \\ &= C_4 2^{-k/2-k/d} \\ &= C_4 2^{-k/2} \lambda^k \end{aligned}$$

for some constant C_4 independent of k.

Each interval I_j will be the convex hull of the image under π of the part P_j of X corresponding to point j in X_k. As in the proof of Proposition 2.22, P_j is contained in a ball of radius $\lambda^k/(1-\lambda)$. Thus in effect, we want to show that on average (over j), π contracts P_j by a factor proportional to $2^{-k/2}$. Recall that

$$N_k = \left\{(-1)^{\beta_{k+1}} p_{\beta_1 \cdots \beta_k} : \beta_1, \ldots, \beta_{k+1} \in \{0,1\}\right\}$$

is the set of points used to perturb the 2^k points of X_k to form the 2^{k+1} points of X_{k+1}. We seek an asymptotic bound on the quantity

$$Z_k = \sup_{\|\pi\|_{\ell^\infty}=1} \frac{1}{2^{k+1}} \sum_{s \in N_k} \frac{|\pi(s)|}{\|s\|_{\ell^1}} = \sup_{\|\pi\|_{\ell^\infty}=1} \frac{1}{2^{k+1}} \sum_{s \in N_k} \frac{|\pi(s)|}{\lambda^k}.$$

Lemma 2.24. *There exists $C_5 > 0$ such that $Z_k \leqslant C_5 2^{-k/2}$.*

Proof. For each $\beta_1 \cdots \beta_k \in \{0,1\}^k$, N_k contains $p_{\beta_1 \cdots \beta_k}$ and $-p_{\beta_1 \cdots \beta_k}$. Define

$$N_k^+ = \left\{p_{\beta_1 \cdots \beta_k} : \beta_1, \ldots, \beta_k \in \{0,1\}\right\}.$$

We reindex the elements of N_k^+ by $\gamma_{\beta_1 \cdots \beta_k}$, obtaining $N_k^+ = \{p_i : i = 0, \ldots, 2^k - 1\}$. For each $\pi = (\pi_i) \in \ell^\infty$, there exists a permutation σ such that $\pi_\sigma = (\pi_{\sigma(i)})$ satisfies the positivity condition

$$\pi_\sigma(p_i) \geqslant 0$$

for all $i = 0, \ldots, 2^k - 1$. Therefore, we express Z_k in terms of N_k^+, yielding

$$Z_k = \sup_{\|\pi\|_{\ell^\infty}=1} \frac{1}{2^k} \sum_{i=0}^{2^k-1} \frac{\pi(p_i)}{\lambda^k}.$$

Think of the points of N_k^+ as the rows of a $2^k \times 2^{2^k}$ matrix. The entry in row i, column j of this matrix (starting the numbering at $i = 0$ and $j = 0$) is

$$p_{ij} = \frac{\lambda^k}{2^{2^k}} (-1)^{\left[\frac{j}{2^i}\right]} e_{\alpha_k + j}.$$

Let (s_{ij}) be the associated matrix of signs, defined by

$$s_{ij} = (-1)^{\left[\frac{j}{2^i}\right]}.$$

The set of columns of (s_{ij}) maps bijectively onto the set of vectors

$$\big\{\big((-1)^{\rho_1}, \ldots, (-1)^{\rho_k}\big) : \rho_1, \ldots, \rho_k \in \{0,1\}\big\}. \tag{2.11}$$

We construct an element $\pi^* \in \ell^\infty$ as follows. For $0 \leqslant j < 2^{2^k}$, set

$$\pi^*_{\alpha_k+j} = \begin{cases} 1, & \text{if } \sum_{i=0}^{2^k-1} p_{ij} \geqslant 0; \\ -1, & \text{if } \sum_{i=0}^{2^k-1} p_{ij} < 0, \end{cases}$$

and set $\pi^*_l = 0$ for $l < \alpha_k$ and $l \geqslant \alpha_k + 2^{2^k}$. Writing

$$r_{ij} = s_{ij} e_{\alpha_k+j} \text{ and } r_i = \sum_{j=0}^{2^{2^k}-1} r_{ij},$$

we have

$$Z_k = \frac{1}{2^k} \sum_{i=0}^{2^k-1} \frac{\pi^*(p_i)}{\lambda^k} = \frac{1}{2^k 2^{2^k}} \sum_{i=0}^{2^k-1} \pi^*(r_i).$$

Since the columns of (s_{ij}) correspond bijectively to (2.11), Z_k may be related to the expected value of a binomially distributed random variable. Let Y be a binomial random variable such that for $0 \leqslant m \leqslant 2^k$, the probability that $Y = m$ is given by

$$\binom{2^k}{m} \left(\frac{1}{2}\right)^{2^k}.$$

Summing over m, we have

$$\begin{aligned} Z_k &= \frac{1}{2^k 2^{2^k}} \sum_{i=0}^{2^k-1} \pi^*(r_i) \\ &= \frac{1}{2^k 2^{2^k}} \sum_{m=0}^{2^k} \binom{2^k}{m} |2^k - 2m| \\ &= \frac{1}{2^{2^k}} \sum_{m=0}^{2^k} \binom{2^k}{m} |1 - 2m/2^k| \\ &= E[|1 - 2Y/2^k|], \end{aligned}$$

where $E[\cdot]$ denotes the expectation. By the central limit theorem, there exists $C_5 > 0$ such that

$$E[|1 - 2Y/2^k|] \leqslant C_5 2^{-k/2}.$$

The proof of Lemma 2.24 is complete. $\square$

Returning to the proof of Proposition 2.23, we show that for each $k \geqslant 0$, $\pi(X)$ can be covered by 2^k intervals $I_0, \ldots, I_{2^k-1}$ such that

$$2^{-k} \sum_{j=0}^{2^k-1} \operatorname{diam}(I_j) \leqslant C_4 2^{-k/2} \lambda^k$$

for some constant C_4 independent of k. Fix $k \geqslant 0$. For each string $\beta_1 \cdots \beta_k$, the subtree

$$\begin{aligned} X^{\beta_1 \cdots \beta_k} = \Big\{ &(-1)^{\beta_1} p + (-1)^{\beta_2} p_{\beta_1} + \cdots + (-1)^{\beta_k} p_{\beta_1 \cdots \beta_{k-1}} + (-1)^{\beta_{k+1}} p_{\beta_1 \cdots \beta_k} \\ &+ (-1)^{\beta_{k+2}} p_{\beta_1 \cdots \beta_{k+1}} + \cdots : \beta_{k+1}, \beta_{k+2}, \ldots \in \{0, 1\} \Big\} \end{aligned}$$

can be covered by an interval $I_j = I_{\gamma_{\beta_1 \cdots \beta_k}}$ containing

$$\pi \left((-1)^{\beta_1} p + (-1)^{\beta_2} p_{\beta_1} + \cdots + (-1)^{\beta_k} p_{\beta_1 \cdots \beta_{k-1}} \right)$$

of length

$$\sum_{i=1}^{\infty} \sum_{\beta_{k+1} \cdots \beta_{k+i}} \left| \pi \left((-1)^{\beta_{k+i}} p_{\beta_1 \cdots \beta_{k+i-1}} \right) \right|.$$

Applying Lemma 2.24, we have

$$\begin{aligned} 2^{-k} \sum_{j=0}^{2^k-1} \operatorname{diam}(I_j) &= 2^{-k} \sum_{j=0}^{2^k-1} \sum_{i=1}^{\infty} \sum_{\beta_{k+1} \cdots \beta_{k+i}} \left| \pi \left((-1)^{\beta_{k+i}} p_{\beta_1 \cdots \beta_{k+i-1}} \right) \right| \\ &= \sum_{i=1}^{\infty} 2^{-k} \sum_{j=0}^{2^k-1} \sum_{\beta_{k+1} \cdots \beta_{k+i}} \left| \pi \left((-1)^{\beta_{k+i}} p_{\beta_1 \cdots \beta_{k+i-1}} \right) \right| \\ &\leqslant \sum_{n=0}^{\infty} 2^{n+1} \lambda^{k+n} \cdot C_5 \cdot 2^{-(k+n)/2} \\ &= 2 C_5 \lambda^k 2^{-k/2} \sum_{n=0}^{\infty} \left(\sqrt{2} \lambda \right)^n. \end{aligned}$$

The assumption that $s < 1$ implies that $\lambda < 1/\sqrt{2}$. Setting

$$C_4 = 2C_5 \sum_{n=0}^{\infty} \left(\sqrt{2}\lambda\right)^n = \frac{2C_5}{1 - \sqrt{2}\lambda},$$

we have

$$2^{-k} \sum_{j=0}^{2^k - 1} \operatorname{diam}(I_j) \leqslant C_4 2^{-k/2} \lambda^k. \tag{2.12}$$

Finally, (2.12) implies that $\operatorname{diam}(I_j) \leqslant C_4 2^{k(1-1/s)}$ for each $j = 0, \ldots, 2^k - 1$. $\square$

Acknowledgements. The first author thanks Jacob Sterbenz for numerous insightful discussions.

Chapter 3

The Global Attractor Associated with the Viscous Lake Equations

The material in this chapter has been accepted for publication in *Nonlinearity.*

3.1 Introduction

We study the asymptotic behavior of the solutions of a two-dimensional shallow water model with eddy viscosity for basins with varying bottom topography. The shallow water model has been derived from a three-dimensional anisotropic eddy viscosity model and has been shown to be globally well-posed in [40]. The derivation exploits two main scaling assumptions. First, one assumes that the ratio of the horizontal fluid velocity to the gravity wave speed is small, while the ratio of the length scale of the top surface height variation to the basin depth is much smaller still. Second, one assumes that the basin is shallow compared with the horizontal length scales of interest. The viscous shallow water model refines the lake system [6] and the great lake system [7]. These systems are derived from three-dimensional Euler flow under the same scaling assumptions. As Levermore and Sammartino [40] point out, the lake and great lake systems neglect several physical phenomena of crucial dynamical importance. The effects of viscous stresses are restored in the viscous lake system.

The viscous shallow water model bears considerable structural resemblance to the two-dimensional incompressible Navier-Stokes system. The study of the attractor associated with the Navier-Stokes equations has motivated a considerable amount of the theory of infinite-dimensional dynamical systems. Consider first the two-dimensional incompressible Navier-Stokes system on a bounded domain with Dirichlet boundary conditions. Invoking a Sobolev-Lieb-Thirring inequality, one may show [9, 12, 64] that the dimension of the global attractor is bounded above by a constant multiple of the Grashof number G, a nondimensional quantity proportional to the L^2-norm of the forcing function. The Sobolev-Lieb-Thirring inequalities play an important role in the estimation of the trace of certain linear operators arising in the study of infinite-dimensional dynamical systems and have led to sharp bounds on attractor dimension in terms of the physical data. Lieb and Thirring [41] prove the first such inequality, a powerful generalization of the Sobolev-Gagliardo-Nirenberg inequalities for a finite family of functions which are orthonormal in $L^2(\mathbb{R}^n)$. Systems amenable to dynamical systems methods include reaction-diffusion equations, nonlinear dissipative wave equations, complex Ginzburg-Landau equations, and various fluid models.

Now consider the Navier-Stokes system on the torus $\mathbb{T}^2$. Using an L^∞ estimate of Constantin on collections of functions whose gradients are orthonormal [10], one may improve the previous bound and show that the dimension of the global attractor is bounded above by a value proportional to $G^{2/3}(1 + \log G)^{1/3}$ in the space-periodic case [13, 14, 15]. This estimate is consistent up to a logarithmic correction with the predictions of the conventional theory of turbulence due to

Constantin, Foias, and Temam [13].

One strives to establish sharp bounds on the attractor dimension, for physical interpretation becomes especially significant once such bounds have been established. Research in this direction has followed two streams of thought. Liu [42] derives a lower bound in terms of the Grashof number when the domain is the torus $\mathbb{T}^2$. A family of external forces is constructed such that

$$\dim(\mathcal{A}) \geqslant \gamma G^{2/3}.$$

Therefore, in the space-periodic case, the best available lower and upper bounds agree up to a logarithmic correction. Alternatively, one may study a flow on the elongated domain $\Omega_\alpha = [0, 2\pi/\alpha] \times [0, 2\pi]$ and investigate the aspect-ratio limit $\alpha \to 0$. In the space-periodic case, a sharp estimate exists. Babin and Vishik [2] choose a specific volume force for which a simple stationary solution can be found. An estimate on the number of unstable modes around the stationary solution yields the lower bound

$$\dim(\mathcal{A}) \geqslant \frac{\gamma_1}{\alpha}.$$

Ziane [68] establishes the sharpness of this lower bound by employing a version of the Sobolev-Lieb-Thirring inequalities for elongated domains to derive the upper bound

$$\dim(\mathcal{A}) \leqslant \frac{\gamma_2}{\alpha}.$$

Doering and Wang [16] show that an application of a Lieb-Thirring inequality with the domain-dependence of the prefactors carefully controlled produces a sharp dependence of the attractor dimension on the length of the channel for certain channel flows. The derivation of a sharp estimate in the case of a general bounded domain with Dirichlet boundary conditions remains an open problem.

Given the structural similarity between the Navier-Stokes equations and the shallow water model, one suspects that a physically significant upper bound may be established for the dimension of the attractor $\mathcal{A}$ of the shallow water system. We initiate the study of this question in the present work. The Hausdorff and box-counting dimensions of $\mathcal{A}$ are shown to be bounded above by a value proportional to the weighted L^2-norm of the wind forcing function. The key technical innovation is the use of a new weighted Sobolev-Lieb-Thirring inequality. This weighted inequality is crucial because the natural function spaces for the shallow water system are the energy spaces with Lebesgue measure weighted by the basin depth function.

Many interesting questions remain open. Is the linear-in-norm bound derived in the present work sharp? Does this bound agree with any qualitative theoretical picture? In particular, how does the attractor dimension scale with the aspect ratio? Illumination of the physical significance of the scaling of an attractor dimension estimate becomes especially meaningful when the estimate is sharp. The use of inequalities akin to the L^∞ estimates of Constantin [10] may lead to an improved dimension estimate. Finally, for simplified geometries one might obtain a sharp result via an argument similar in spirit to the work of Doering and Wang [16] on channel flows.

The paper is organized as follows. In Section 3.2 we introduce the shallow water model and discuss its mathematical structure. The existence of a global attractor for the shallow water system is established in Section 3.3. Section 3.4 contains the derivation of the main attractor dimension estimate. We present the weighted Sobolev-Lieb-Thirring inequality in Section 3.5.

3.2 The Shallow Water Model

We consider an incompressible fluid that is confined to a three-dimensional basin by a uniform gravitational field of magnitude g. In terms of the standard Cartesian coordinates with the positive z-axis oriented upward, the basin is defined by its orthogonal projection onto the xy-plane, Ω, and by its bottom. The bottom is defined by $z = -b(\boldsymbol{x})$ for $\boldsymbol{x} = (x, y) \in \Omega$. The domain $\Omega \subset \mathbb{R}^2$ is assumed to be bounded with a smooth boundary $\partial\Omega$. We assume that b is a positive, smooth function over $\overline{\Omega}$. Let the free top surface of the fluid at time t be given by $z = h(\boldsymbol{x}, t)$. We assume that the free top surface never meets the bottom and that the average level of the top surface is $z = 0$. The domain occupied by the fluid at time t, denoted $\Sigma(t)$, is given by

$$\Sigma(t) = \{(\boldsymbol{x}, z) \in \mathbb{R}^3 : \boldsymbol{x} \in \Omega \; - b(\boldsymbol{x}) < z < h(\boldsymbol{x}, t)\}.$$

The shallow water model governs the evolution of $\boldsymbol{u}(\boldsymbol{x}, t)$, the horizontal fluid velocity averaged vertically over $\boldsymbol{x} \in \Omega$ at time t, and the top surface height $h(\boldsymbol{x}, t)$. The system of equations is as follows.

$$\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \nabla_{\boldsymbol{x}} \boldsymbol{u} + g \nabla_{\boldsymbol{x}} h = b^{-1} \nabla_{\boldsymbol{x}} \cdot [b\nu(\nabla_{\boldsymbol{x}} \boldsymbol{u} + (\nabla_{\boldsymbol{x}} \boldsymbol{u})^T - \nabla_{\boldsymbol{x}} \cdot \boldsymbol{u} \boldsymbol{I})] - \eta \boldsymbol{u} + \boldsymbol{f},$$

$$\nabla_{\boldsymbol{x}} \cdot (b\boldsymbol{u}) = 0,$$

$$\boldsymbol{u}(\boldsymbol{x}, 0) = \boldsymbol{u}_0(\boldsymbol{x}),$$

$$\boldsymbol{u} \cdot \boldsymbol{n} = 0 \quad (\text{for } \boldsymbol{x} \in \partial\Omega),$$

$$\nu \boldsymbol{t} \cdot (\nabla_{\boldsymbol{x}} \boldsymbol{u} + (\nabla_{\boldsymbol{x}} \boldsymbol{u})^T) \cdot \boldsymbol{n} = -\beta \boldsymbol{t} \cdot \boldsymbol{u} \quad (\text{for } \boldsymbol{x} \in \partial\Omega).$$

Here $\nu(\boldsymbol{x})$ and $\eta(\boldsymbol{x})$ are a positive eddy viscosity coefficient and a non-negative turbulent drag coefficient defined over Ω, $\boldsymbol{I}$ is the 2×2 identity, $\boldsymbol{f}(\boldsymbol{x}, t)$ is the wind forcing defined over $\Omega \times [0, \infty)$, $\boldsymbol{n}(\boldsymbol{x})$ and $\boldsymbol{t}(\boldsymbol{x})$ are the outward unit normal and a unit tangent to $\partial\Omega$ at $\boldsymbol{x}$ and $\beta(\boldsymbol{x})$ is a non-negative turbulent boundary drag coefficient defined on $\partial\Omega$.

We reformulate the shallow water equations as an abstract evolution equation governing the velocity field $\boldsymbol{u}$. It is natural to work with Sobolev spaces weighted by the function b. The scalar-valued spaces are denoted L_b^p, $W_b^{s,p}$, and H_b^s with norms $\|\cdot\|_{L_b^p}$, $\|\cdot\|_{W_b^{s,p}}$, and $\|\cdot\|_{H_b^s}$, respectively. The vector-valued counterparts are given by $\mathbf{L}_b^p$, $\mathbf{W}_b^{s,p}$, and $\mathbf{H}_b^s$. The inner product between $\boldsymbol{u}, \boldsymbol{v} \in \mathbf{L}_b^2$ is denoted $(\boldsymbol{u}, \boldsymbol{v})$ and is defined by

$$(\boldsymbol{u}, \boldsymbol{v}) = \int_\Omega b\boldsymbol{u} \cdot \boldsymbol{v}\, d\boldsymbol{x} = \int_\Omega \boldsymbol{u}(\boldsymbol{x}) \cdot \boldsymbol{v}(\boldsymbol{x})\, d\lambda(\boldsymbol{x}),$$

where λ denotes the two-dimensional Lebesgue measure weighted by b. We define the spaces

$$H = \{\boldsymbol{u} : \boldsymbol{u} \in \mathbf{L}_b^2,\ \nabla_{\boldsymbol{x}} \cdot (b\boldsymbol{u}) = 0,\ \boldsymbol{n} \cdot \boldsymbol{u} = 0 \text{ for } \boldsymbol{x} \in \partial\Omega\},$$
$$V = \{\boldsymbol{u} : \boldsymbol{u} \in \mathbf{H}_b^1,\ \nabla_{\boldsymbol{x}} \cdot (b\boldsymbol{u}) = 0,\ \boldsymbol{n} \cdot \boldsymbol{u} = 0 \text{ for } \boldsymbol{x} \in \partial\Omega\}.$$

When there is no possibility of confusion we write $|\cdot| = \|\cdot\|_{\mathbf{L}_b^2}$ and $\|\cdot\| = \|\cdot\|_{\mathbf{H}_b^1}$.

Assume $\beta(\boldsymbol{x}) \geqslant \kappa(\boldsymbol{x})$ for all $\boldsymbol{x} \in \partial\Omega$, where κ is the curvature of $\partial\Omega$. Suppose that b and ν are smooth, positive functions such that $b\nu \geqslant C > 0$ for some constant C. Under these assumptions, the bilinear form $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ defined by

$$a(\boldsymbol{u}, \boldsymbol{v}) = \frac{1}{2} \int_\Omega b\nu(\nabla_{\boldsymbol{x}}\boldsymbol{u} + (\nabla_{\boldsymbol{x}}\boldsymbol{u})^T - \nabla_{\boldsymbol{x}} \cdot \boldsymbol{u}\boldsymbol{I}) : (\nabla_{\boldsymbol{x}}\boldsymbol{v} + (\nabla_{\boldsymbol{x}}\boldsymbol{v})^T - \nabla_{\boldsymbol{x}} \cdot \boldsymbol{v}\boldsymbol{I})\, d\boldsymbol{x}$$
$$+ \int_\Omega b\nu\eta\boldsymbol{u} \cdot \boldsymbol{v}\, d\boldsymbol{x} + \int_{\partial\Omega} b\nu\beta\boldsymbol{u} \cdot \boldsymbol{v}\, ds$$

is coercive; that is, there exists $\alpha > 0$ such that $a(\boldsymbol{u}, \boldsymbol{u}) \geqslant \alpha\|\boldsymbol{u}\|^2$ for all $\boldsymbol{u} \in V$. By the Lax-Milgram theorem, the operator $A : V \to V'$ defined by

$$\langle A\boldsymbol{u}, \boldsymbol{v}\rangle = a(\boldsymbol{u}, \boldsymbol{v}) \quad (\boldsymbol{u}, \boldsymbol{v} \in V)$$

maps V isomorphically onto V'. This operator is a linear unbounded operator on H with dense domain $D(A) = \mathbf{H}_b^2 \cap V$. The inverse operator A^{-1} is self-adjoint and compact by virtue of Rellich's theorem. Thus there exists an orthonormal basis of H and a sequence (λ_j) such that

$$\begin{cases} 0 < \lambda_1 \leqslant \lambda_2 \leqslant \cdots, \ \lambda_j \to \infty, \\ A\boldsymbol{w}_j = \lambda_j \boldsymbol{w}_j \ \forall j. \end{cases}$$

We define the trilinear form $(\cdot, \cdot, \cdot)$ on V by

$$(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = \int_\Omega b\boldsymbol{u} \cdot \nabla_{\boldsymbol{x}} \boldsymbol{v} \cdot \boldsymbol{w} \, d\boldsymbol{x},$$

and the corresponding bilinear operator $B(\cdot, \cdot) : V \times V \to V'$ by

$$\langle B(\boldsymbol{u}, \boldsymbol{v}), \boldsymbol{w} \rangle = (\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}).$$

The shallow water system is equivalent to the evolution equation

$$\partial_t \boldsymbol{u} + A\boldsymbol{u} + B(\boldsymbol{u}, \boldsymbol{u}) = \boldsymbol{f}, \tag{3.1}$$

coupled with initial data

$$\boldsymbol{u}(\boldsymbol{x}, 0) = \boldsymbol{u}_0(\boldsymbol{x}). \tag{3.2}$$

The shallow water system is shown to be globally well-posed in [40]. The following is established therein.

Theorem 3.1 ([40]). *Let Ω be smooth. Suppose that $b(\boldsymbol{x})$, $\nu(\boldsymbol{x})$, and $\eta(\boldsymbol{x})$ are non-negative functions over $\overline{\Omega}$. Suppose that b and ν are smooth, that $b\nu \geqslant C > 0$ for some constant C, and that $\beta(\boldsymbol{x}) \geqslant \kappa(\boldsymbol{x})$ on $\partial\Omega$, where $\kappa(\boldsymbol{x})$ is the curvature of $\partial\Omega$ at $\boldsymbol{x}$. Let $\boldsymbol{f} \in \mathbf{L}_b^2$ and let $T > 0$. If $\boldsymbol{u}_0 \in H$, then there exists a unique*

$$\boldsymbol{u} \in C([0,T], H) \cap L^2([0,T], V)$$

that satisfies (3.1) and (3.2). If $\boldsymbol{u}_0 \in \mathbf{H}_b^2 \cap V$, then one has moreover that

$$\boldsymbol{u} \in L^\infty([0,T], \mathbf{H}_b^2) \cap C([0,T], V),$$

and

$$\partial_t \boldsymbol{u} \in L^\infty([0,T], H) \cap L^2([0,T], V).$$

We define the semigroup $S(\cdot)$ of continuous operators on H as follows. For fixed $t \geqslant 0$, $S(t) : H \to H$ is given by $S(t)\boldsymbol{u}_0 = \boldsymbol{u}(t)$.

3.3 The Attractor

To demonstrate the existence of the global attractor $\mathcal{A}$ associated with $\{S(t) : t \geqslant 0\}$, we show that the semigroup is dissipative and uniformly asymptotically compact. Dissipativity in this context is characterized by the existence of a bounded absorbing set in H. The existence proof relies on standard techniques. We include the argument to fix notation and to establish estimates that are needed for the dimension calculation.

Definition 3.2. Let $\mathcal{C} \subset H$. We say that $\mathcal{C}$ is absorbing in H if for each bounded set $\mathcal{B} \subset H$ there exists $t_1(\mathcal{B})$ such that $S(t)\mathcal{B} \subset \mathcal{C}$ for all $t \geqslant t_1(\mathcal{B})$.

Definition 3.3. The semigroup $S(\cdot)$ is said to be uniformly asymptotically compact if for each bounded set $\mathcal{B} \subset H$ there exists $t_0(\mathcal{B})$ such that

$$\bigcup_{t \geqslant t_0} S(t)\mathcal{B}$$

is relatively compact in H.

We establish the uniform asymptotic compactness of the semigroup by establishing the existence of a bounded absorbing set in V and noting that V embeds compactly into H. One uses energy methods to produce absorbing sets in H and V.

3.3.1 Absorbing Set in H

We will need the following orthogonality relation.

Lemma 3.4. *For $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w} \in V$ one has*

$$(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = -(\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}),$$

and thus one has the orthogonality relation

$$(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{v}) = 0.$$

By Sobolev embeddings there exists a constant c_1 such that

$$|\boldsymbol{u}| \leqslant c_1 \|\boldsymbol{u}\|.$$

Now $\|A^{-1}\|_{L(V',V)} \leqslant \frac{1}{\alpha}$ so one has

$$\|\boldsymbol{u}\| \leqslant \frac{1}{\alpha}\|A\boldsymbol{u}\|_{V'} \leqslant \frac{c_1}{\alpha}|A\boldsymbol{u}|.$$

Set $c_2 = \frac{c_1}{\alpha}$ and $c_3 = c_1^2$. Taking the scalar product of (3.1) with $\boldsymbol{u}$ in H, we obtain

$$\frac{1}{2}\frac{d}{dt}|\boldsymbol{u}|^2 + a(\boldsymbol{u}, \boldsymbol{u}) = (\boldsymbol{f}, \boldsymbol{u}).$$

Bounding the right-hand side, we have

$$\begin{aligned}
(\boldsymbol{f}, \boldsymbol{u}) &\leqslant |\boldsymbol{f}|\,|\boldsymbol{u}| \\
&\leqslant c_1 |\boldsymbol{f}|\,\|\boldsymbol{u}\| \\
&\leqslant \frac{\alpha}{2}\|\boldsymbol{u}\|^2 + \frac{c_1^2}{2\alpha}|\boldsymbol{f}|^2 \quad \text{(Young's inequality).}
\end{aligned}$$

We obtain

$$\frac{d}{dt}|\boldsymbol{u}|^2 + \alpha\|\boldsymbol{u}\|^2 \leqslant \frac{c_1^2}{\alpha}|\boldsymbol{f}|^2,$$
$$\frac{d}{dt}|\boldsymbol{u}|^2 + \frac{\alpha}{c_1^2}|\boldsymbol{u}|^2 \leqslant \frac{c_1^2}{\alpha}|\boldsymbol{f}|^2,$$
$$\frac{d}{dt}|\boldsymbol{u}|^2 \leqslant \left(-\frac{\alpha}{c_3}\right)|\boldsymbol{u}|^2 + \left(\frac{c_3}{\alpha}|\boldsymbol{f}|^2\right).$$

An application of the classical Gronwall inequality yields the estimate

$$|\boldsymbol{u}(t)|^2 \leqslant |\boldsymbol{u}_0|^2 \exp\left(-\frac{\alpha}{c_3}t\right) + \frac{c_3^2}{\alpha^2}|\boldsymbol{f}|^2\left[1 - \exp\left(-\frac{\alpha}{c_3}t\right)\right].$$

Taking the upper limit, one obtains

$$\overline{\lim_{t\to\infty}}\,|\boldsymbol{u}(t)| \leqslant \frac{c_3}{\alpha}|\boldsymbol{f}| := \rho_0.$$

We conclude that $\mathcal{B}_H(0,\rho)$, the metric ball in H of radius ρ, is absorbing for $\rho > \rho_0$. For fixed $\rho > \rho_0$ and a bounded set $\mathcal{B} \subset H$, there exists $t_1(\mathcal{B},\rho)$ such that $S(t)\mathcal{B} \subset \mathcal{B}_H(0,\rho)$ for all $t \geqslant t_1(\mathcal{B},\rho)$.

3.3.2 Absorbing Set in V

We need the following continuity property of the trilinear form $(\cdot,\cdot,\cdot)$.

Lemma 3.5. *There exists a constant k such that for $\boldsymbol{u} \in V$, $\boldsymbol{v} \in D(A)$, and $\boldsymbol{w} \in H$ one has*

$$|(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})| \leqslant k|\boldsymbol{u}|^{\frac{1}{2}}\|\boldsymbol{u}\|^{\frac{1}{2}}\|\boldsymbol{v}\|^{\frac{1}{2}}|A\boldsymbol{v}|^{\frac{1}{2}}|\boldsymbol{w}|. \tag{3.3}$$

Proof. The proof is based on two key facts. The first is an interpolation inequality known as Ladyzhenskaya's inequality.

Lemma 3.6. *For $\boldsymbol{u} \in \mathbf{H}_b^1(\Omega)$ one has*

$$\|\boldsymbol{u}\|_{\mathbf{L}_b^4} \leqslant c_4|\boldsymbol{u}|^{\frac{1}{2}}\|\boldsymbol{u}\|^{\frac{1}{2}}.$$

We also need an elliptic regularity estimate for the strong Stokes problem associated with the shallow water system. It is shown in [40] that for $\boldsymbol{f} \in \mathbf{L}_b^p$, $p \in (1, \infty)$, the strong Stokes problem admits a unique solution $\boldsymbol{u} \in \mathbf{W}_b^{2,p}$ satisfying

$$\|\boldsymbol{u}\|_{\mathbf{W}_b^{2,p}} \leqslant c(\|\boldsymbol{f}\|_{\mathbf{L}_b^p} + \|\boldsymbol{u}\|_{\mathbf{L}_b^p}). \tag{3.4}$$

Notice that A^{-1}, the operator mapping $\mathbf{L}_b^2$ data to the solution of the strong Stokes problem, may be extended as a linear continuous operator from $\mathbf{L}_b^p(\Omega)$ into $\mathbf{W}_b^{2,p}(\Omega)$ for all $p \in (1, \infty)$. For $\boldsymbol{u} \in V$, $\boldsymbol{v} \in D(A)$, and $\boldsymbol{w} \in H$, one has

$$\begin{aligned}
\left| \int_\Omega \boldsymbol{u} \cdot \nabla_{\boldsymbol{x}} \boldsymbol{v} \cdot \boldsymbol{w} \, d\lambda \right| &\leqslant \sum_{i,j=1}^{2} \int_\Omega |u_i (D_i v_j) w_j| \, d\lambda \\
&\leqslant \sum_{i,j=1}^{2} \|u_i\|_{L_b^4} \|D_i v_j\|_{L_b^4} |w_j| \\
&\leqslant \sum_{i,j=1}^{2} c_4^2 |u_i|^{\frac{1}{2}} \|u_i\|^{\frac{1}{2}} |D_i v_j|^{\frac{1}{2}} \|D_i v_j\|^{\frac{1}{2}} |w_j| \\
&\leqslant c_4^2 \left(\sum_{i=1}^{2} |u_i| \, \|u_i\| \right)^{\frac{1}{2}} \left(\sum_{i,j=1}^{2} |D_i v_j| \, \|D_i v_j\| \right)^{\frac{1}{2}} \left(\sum_{j=1}^{2} |w_j|^2 \right)^{\frac{1}{2}} \\
&\leqslant c_4^2 |\boldsymbol{u}|^{\frac{1}{2}} \|\boldsymbol{u}\|^{\frac{1}{2}} \|\boldsymbol{v}\|^{\frac{1}{2}} \|\boldsymbol{v}\|_{\mathbf{H}_b^2(\Omega)}^{\frac{1}{2}} |\boldsymbol{w}| \\
&\leqslant c_4^2 c^{\frac{1}{2}} (1 + c_2)^{\frac{1}{2}} |\boldsymbol{u}|^{\frac{1}{2}} \|\boldsymbol{u}\|^{\frac{1}{2}} \|\boldsymbol{v}\|^{\frac{1}{2}} |A\boldsymbol{v}|^{\frac{1}{2}} |\boldsymbol{w}|.
\end{aligned}$$

Setting $k = c_4^2 c^{\frac{1}{2}} (1 + c_2)^{\frac{1}{2}}$, the lemma is established. □

We are now in position to establish the existence of an absorbing set in V. Taking the scalar product of (3.1) with $A\boldsymbol{u}$ gives

$$\frac{1}{2} \frac{d}{dt} a(\boldsymbol{u}, \boldsymbol{u}) + |A\boldsymbol{u}|^2 = (\boldsymbol{f}, A\boldsymbol{u}) - (\boldsymbol{u}, \boldsymbol{u}, A\boldsymbol{u}).$$

Applying the continuity estimate (3.3), we obtain

$$\begin{aligned}
|(\boldsymbol{u}, \boldsymbol{u}, A\boldsymbol{u})| &\leqslant (|A\boldsymbol{u}|^{\frac{3}{2}})(k |\boldsymbol{u}|^{\frac{1}{2}} \|\boldsymbol{u}\|) \\
&\leqslant \frac{3}{8} |A\boldsymbol{u}|^2 + 2k^4 |\boldsymbol{u}|^2 \|\boldsymbol{u}\|^4 \quad \text{(Young's inequality)}.
\end{aligned}$$

Bounding the scalar product $(\boldsymbol{f}, A\boldsymbol{u})$, one has

$$(\boldsymbol{f}, A\boldsymbol{u}) \leqslant |\boldsymbol{f}|\,|A\boldsymbol{u}| \leqslant \frac{|A\boldsymbol{u}|^2}{4} + |\boldsymbol{f}|^2.$$

Collecting these estimates, we obtain

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}a(\boldsymbol{u},\boldsymbol{u}) + \frac{3}{8}|A\boldsymbol{u}|^2 &\leqslant |\boldsymbol{f}|^2 + 2k^4|\boldsymbol{u}|^2\|\boldsymbol{u}\|^4 \\ &\leqslant |\boldsymbol{f}|^2 + 2k^4|\boldsymbol{u}|^2\|\boldsymbol{u}\|^2\left(\frac{a(\boldsymbol{u},\boldsymbol{u})}{\alpha}\right),\end{aligned}$$

and we conclude that

$$\frac{d}{dt}a(\boldsymbol{u},\boldsymbol{u}) \leqslant 2|\boldsymbol{f}|^2 + c_5|\boldsymbol{u}|^2\|\boldsymbol{u}\|^2 a(\boldsymbol{u},\boldsymbol{u}),$$

where $c_5 = \frac{4k^4}{\alpha}$. In order to control $\|\boldsymbol{u}(t)\|$ as $t \to \infty$ we invoke the uniform Gronwall lemma.

Lemma 3.7 (Uniform Gronwall). *Let g, h, and y be three positive locally integrable functions on $[t_1,\infty)$ such that y is absolutely continuous on $[t_1,\infty)$ and which satisfy*

$$\frac{dy}{dt} \leqslant gy + h,$$
$$\int_t^{t+r} g(s)\,ds \leqslant a_1, \quad \int_t^{t+r} h(s)\,ds \leqslant a_2, \quad \int_t^{t+r} y(s)\,ds \leqslant a_3$$

for $t \geqslant t_1$, where r, a_1, a_2, and a_3 are positive constants. Then

$$y(t+r) \leqslant \left(\frac{a_3}{r} + a_2\right)\exp(a_1) \quad (t \geqslant t_1).$$

Fix $\rho > \rho_0$ and $r > 0$. Let $\mathcal{B} \subset H$ be a bounded subset of H. As we have seen, there exists $t_1(\mathcal{B},\rho)$ such that $S(t)\mathcal{B} \subset \mathcal{B}_H(0,\rho)$ for all $t \geqslant t_1(\mathcal{B},\rho)$. We apply the uniform Gronwall lemma with

$$\begin{cases} y = a(\boldsymbol{u},\boldsymbol{u}) \\ g = c_5|\boldsymbol{u}|^2\|\boldsymbol{u}\|^2 \\ h = 2|\boldsymbol{f}|^2 \end{cases}$$

by producing constants a_1, a_2, and a_3 valid for $t \geqslant t_1(\mathcal{B}, \rho)$. One must first bound the integral of $\|\boldsymbol{u}\|^2$ over time intervals $[t, t+r]$ with $t \geqslant t_1(\mathcal{B}, \rho)$. Recall the inequality

$$\frac{d}{dt}|\boldsymbol{u}|^2 + \alpha\|\boldsymbol{u}\|^2 \leqslant \frac{c_1^2}{\alpha}|\boldsymbol{f}|^2.$$

Integrating in time, we obtain

$$\begin{aligned}
\int_t^{t+r} \frac{d}{ds}|\boldsymbol{u}|^2\, ds + \int_t^{t+r} \alpha\|\boldsymbol{u}\|^2\, ds &\leqslant \frac{c_1^2}{\alpha}|\boldsymbol{f}|^2 r,\\
\int_t^{t+r} \|\boldsymbol{u}\|^2\, ds &\leqslant \frac{c_1^2}{\alpha^2}|\boldsymbol{f}|^2 r + \frac{|\boldsymbol{u}(t)|^2}{\alpha}\\
&\leqslant \frac{c_1^2}{\alpha^2}|\boldsymbol{f}|^2 r + \frac{\rho^2}{\alpha}.
\end{aligned}$$

The constants a_1, a_2, and a_3 are defined as follows:

$$\begin{aligned}
\int_t^{t+r} g(s)\, ds &\leqslant c_5\rho^2 \int_t^{t+r} \|\boldsymbol{u}\|^2\, ds\\
&\leqslant c_5\rho^2 \left(\frac{c_1^2}{\alpha^2}|\boldsymbol{f}|^2 r + \frac{\rho^2}{\alpha}\right) := a_1,\\
\int_t^{t+r} h(s)\, ds &= \int_t^{t+r} 2|\boldsymbol{f}|^2\, ds = 2|\boldsymbol{f}|^2 r := a_2,\\
\int_t^{t+r} a(\boldsymbol{u}, \boldsymbol{u})\, ds &\leqslant \int_t^{t+r} M\|\boldsymbol{u}(s)\|^2\, ds \quad (a(\boldsymbol{u}, \boldsymbol{u}) \leqslant M\|\boldsymbol{u}\|^2)\\
&\leqslant M\left(\frac{c_1^2}{\alpha^2}|\boldsymbol{f}|^2 r + \frac{\rho^2}{\alpha}\right) := a_3.
\end{aligned}$$

The uniform Gronwall lemma yields the bound

$$\alpha\|\boldsymbol{u}\|^2 \leqslant a(\boldsymbol{u}(t), \boldsymbol{u}(t)) \leqslant \left(\frac{a_3}{r} + a_2\right)\exp(a_1),$$

valid for every $t \geqslant t_1(\mathcal{B}, \rho) + r$. We conclude that the ball in V of radius

$$\left[\left(\frac{a_3}{r} + a_2\right)\frac{\exp(a_1)}{\alpha}\right]^{\frac{1}{2}}$$

is absorbing.

3.4 Upper Bound on the Attractor Dimension

3.4.1 Uniform Lyapunov Exponents

Fix $T = 1$. According to the ergodic theory of dynamical systems, the attractor $\mathcal{A}$ is the support of a measure μ that is invariant under the action of $S(T)$. The multiplicative ergodic theorem of Oseledec implies the existence of classical Lyapunov exponents for μ-almost every $\boldsymbol{u} \in \mathcal{A}$. Because the classical Lyapunov exponents may fail to exist, we employ the concept of uniform Lyapunov exponents (see [9, 64]).

Definition 3.8. The semigroup $\{S(t)\}$ is said to be uniformly quasidifferentiable on $\mathcal{A}$ if for $t \geqslant 0$ and $\boldsymbol{u} \in \mathcal{A}$ there exists a bounded linear operator $L(t, \boldsymbol{u}) : H \to H$, the quasidifferential, such that

$$\frac{|S(t)\boldsymbol{v} - S(t)\boldsymbol{u} - L(t, \boldsymbol{u})(\boldsymbol{v} - \boldsymbol{u})|}{|\boldsymbol{v} - \boldsymbol{u}|} \leqslant \gamma(t, |\boldsymbol{v} - \boldsymbol{u}|) \text{ for } \boldsymbol{v} \in \mathcal{A}$$

where $\gamma(t, s) \to 0$ as $s \to 0$.

Proposition 3.9. *The semigroup $\{S(t)\}$ associated with the shallow water system is uniformly quasidifferentiable on $\mathcal{A}$. The quasidifferential $L(t, \boldsymbol{u}(t))$ solves the linear variational equation*

$$\begin{cases} \partial_t \boldsymbol{\xi} = F'(\boldsymbol{u}(t))\boldsymbol{\xi} \\ \boldsymbol{\xi}(\boldsymbol{x}, 0) = \boldsymbol{v}(\boldsymbol{x}) \end{cases} \tag{3.5}$$

uniquely, where F' denotes the Fréchet derivative of F. One has the uniform bound

$$\sup_{\boldsymbol{u} \in \mathcal{A}} \|L(T, \boldsymbol{u})\|_{\mathcal{L}(H,H)} < \infty.$$

Proof. The result follows from the implicit function theorem and is analogous to the corresponding result for the semigroup associated with the two-dimensional Navier-Stokes system. See Theorem 7.1.1 of [3] or Chapter 13 of [56]. $\square$

This proposition implies that the uniform Lyapunov exponents, denoted μ_j, are well-defined. We relate these exponents to the evolution of the volume element. Fix $\boldsymbol{u}_0 \in \mathcal{A}$. Let $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ be m elements of H and let $\boldsymbol{\xi}_i$ denote the solution of the variational equation with initial data $\boldsymbol{v}_i$. The volume element satisfies the evolution equation

$$\|\boldsymbol{\xi}_1(t) \wedge \cdots \wedge \boldsymbol{\xi}_m(t)\|_{\bigwedge_H^m} = \|\boldsymbol{v}_1 \wedge \cdots \wedge \boldsymbol{v}_m\|_{\bigwedge_H^m} \exp\left(\int_0^t \operatorname{Tr}\left(F'(\boldsymbol{u}(\tau)) \circ Q_m(\tau)\right) d\tau\right)$$

where $Q_m(t) = Q_m(t, \boldsymbol{u}_0; \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$ is the orthogonal projector onto the space spanned by $\boldsymbol{\xi}_1(t), \ldots, \boldsymbol{\xi}_m(t)$. We introduce the quantities

$$q_m(t) = \sup_{\boldsymbol{u}_0 \in \mathcal{A}} \sup_{\substack{\boldsymbol{v}_i \in H \\ |\boldsymbol{v}_i| \leqslant 1 \\ i=1,\ldots,m}} \left(\frac{1}{t}\int_0^t \operatorname{Tr}\left(F'(S(\tau)\boldsymbol{u}_0) \circ Q_m(\tau)\right) d\tau\right),$$

$$q_m = \varlimsup_{t \to \infty} q_m(t).$$

The uniform Lyapunov exponents satisfy

$$\mu_1 + \cdots + \mu_m \leqslant q_m.$$

For the shallow water model we will establish the bound

$$q_m \leqslant \psi(m) := -\gamma_1 m^2 + \gamma_2$$

for some $\gamma_1 > 0$, $\gamma_2 > 0$. Applying Theorem III.3.2 of [9], one concludes that the Hausdorff and box dimensions of $\mathcal{A}$ are bounded above by

$$N + \frac{\psi(N)}{\psi(N) - \psi(N+1)}$$

where N is the smallest integer such that $\psi(N+1) < 0$ and $\psi(N) \geqslant 0$.

3.4.2 The Estimate

The variational equation (3.5) is equivalent to

$$\frac{d\boldsymbol{\xi}}{dt} + A\boldsymbol{\xi} + B(\boldsymbol{u}, \boldsymbol{\xi}) + B(\boldsymbol{\xi}, \boldsymbol{u}) = 0.$$

Fix $\tau > 0$. Let $\{\boldsymbol{\varphi}_j(\tau) : j = 1, \ldots, m\}$ be an orthonormal basis of $Q_m(\tau)H$. One has

$$\begin{aligned}\operatorname{Tr}\big(F'(S(\tau)\boldsymbol{u}_0) \circ Q_m(\tau)\big) &= \sum_{j=1}^{m} (F'(\boldsymbol{u}(\tau))\boldsymbol{\varphi}_j(\tau), \boldsymbol{\varphi}_j(\tau)) \\ &= -\sum_{j=1}^{m} (A\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_j) - \sum_{j=1}^{m} (\boldsymbol{\varphi}_j, \boldsymbol{u}, \boldsymbol{\varphi}_j).\end{aligned}$$

Notice that the first term has the good sign. Gaining control of the second term is the key to the estimate. Now

$$\sum_{j=1}^{m} (\boldsymbol{\varphi}_j, \boldsymbol{u}, \boldsymbol{\varphi}_j) = \int_\Omega \sum_{j=1}^{m} \sum_{i,k=1}^{2} \varphi_{ji}(\boldsymbol{x}) D_i u_k(\boldsymbol{x}) \varphi_{jk}(\boldsymbol{x})\, d\lambda(\boldsymbol{x}),$$

whence for almost every $\boldsymbol{x} \in \Omega$ we have

$$\left| \sum_{j=1}^{m} \sum_{i,k=1}^{2} \varphi_{ji}(\boldsymbol{x}) D_i u_k(\boldsymbol{x}) \varphi_{jk}(\boldsymbol{x}) \right| \leqslant |\nabla \boldsymbol{u}(\boldsymbol{x})| \rho(\boldsymbol{x}),$$

where

$$|\nabla \boldsymbol{u}(\boldsymbol{x})| = \left(\sum_{i,k=1}^{2} |D_i u_k(\boldsymbol{x})|^2 \right)^{\frac{1}{2}},$$

$$\rho(\boldsymbol{x}) = \sum_{i=1}^{2} \sum_{j=1}^{m} (\varphi_{ji}(\boldsymbol{x}))^2.$$

Thus

$$\begin{aligned}\left| \sum_{j=1}^{m} (\boldsymbol{\varphi}_j, \boldsymbol{u}, \boldsymbol{\varphi}_j) \right| &\leqslant \int_\Omega |\nabla \boldsymbol{u}(\boldsymbol{x})| \rho(\boldsymbol{x})\, d\lambda(\boldsymbol{x}) \\ &\leqslant |\rho|\, |\nabla \boldsymbol{u}| \quad \text{(Hölder)} \\ &\leqslant |\rho|\, \|\boldsymbol{u}\|.\end{aligned}$$

At this point we have established the estimate

$$\operatorname{Tr}\big(F'(\boldsymbol{u}(\tau)) \circ Q_m(\tau)\big) \leqslant -\sum_{j=1}^{m}(A\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_j) + |\rho|\,\|\boldsymbol{u}\|.$$

Applying the weighted Sobolev-Lieb-Thirring inequality (3.11), there exists d_1 independent of the family $\{\boldsymbol{\varphi}_j\}$ and of m such that

$$\int_\Omega \rho(\boldsymbol{x})^2\, d\lambda(\boldsymbol{x}) \leqslant d_1 \left(\sum_{j=1}^{m} a(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_j)\right).$$

Set $\omega = 1/c_1^2$. By the variational principle and the spectral estimate (3.10), there exists d_2 such that

$$\sum_{j=1}^{m} a(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_j) \geqslant d_2 \omega m^2.$$

Substituting, we have

$$\begin{aligned}
\operatorname{Tr}\big(F'(\boldsymbol{u}(\tau)) \circ Q_m(\tau)\big) &\leqslant -\sum_{j=1}^{m}(A\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_j) + \|\boldsymbol{u}\| \left(d_1 \sum_{j=1}^{m} a(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_j)\right)^{\frac{1}{2}} \\
&\leqslant -\sum_{j=1}^{m}(A\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_j) + \frac{\|\boldsymbol{u}\|^2 d_1}{2} + \frac{1}{2}\sum_{j=1}^{m} a(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_j) \\
&= -\frac{1}{2}\sum_{j=1}^{m} a(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_j) + \frac{\|\boldsymbol{u}\|^2 d_1}{2} \\
&\leqslant -\frac{d_2 \omega m^2}{2} + \frac{\|\boldsymbol{u}\|^2 d_1}{2}
\end{aligned}$$

and therefore

$$\frac{1}{t}\int_0^t \operatorname{Tr}\big(F'(\boldsymbol{u}(\tau)) \circ Q_m(\tau)\big)\, d\tau \leqslant -\frac{d_2 \omega m^2}{2} + \frac{d_1}{2}\left(\frac{1}{t}\int_0^t \|\boldsymbol{u}(\tau)\|^2\, d\tau\right).$$

Define

$$\epsilon := \alpha\omega \varlimsup_{t\to\infty} \sup_{\boldsymbol{u}_0 \in \mathcal{A}} \frac{1}{t}\int_0^t \|\boldsymbol{u}(s)\|^2\, ds.$$

Integrating the estimate

$$\frac{d}{dt}|\boldsymbol{u}|^2 + \alpha\|\boldsymbol{u}\|^2 \leqslant \frac{1}{\alpha\omega}|\boldsymbol{f}|^2$$

in time, one has

$$\frac{1}{t}|\boldsymbol{u}(t)|^2 + \frac{\alpha}{t}\int_0^t \|\boldsymbol{u}(s)\|^2\,ds \leqslant \frac{1}{t}|\boldsymbol{u}_0|^2 + \frac{1}{\alpha\omega}|\boldsymbol{f}|^2.$$

It follows that

$$\epsilon \leqslant G^2\alpha\omega^2,$$

where

$$G := \frac{|\boldsymbol{f}|}{\alpha\omega}.$$

We conclude that

$$q_m(t) \leqslant -\frac{d_2\omega m^2}{2} + \frac{d_1}{2}\sup_{\boldsymbol{u}_0\in\mathcal{A}}\frac{1}{t}\int_0^t \|\boldsymbol{u}\|^2\,d\tau,$$

and thus

$$q_m = \overline{\lim_{t\to\infty}}\, q_m(t) \leqslant -\gamma_1 m^2 + \gamma_2,$$

where

$$\begin{cases} \gamma_1 = \dfrac{d_2\omega}{2}, \\ \gamma_2 = \dfrac{d_1\epsilon}{2\alpha\omega}. \end{cases}$$

Applying the aforementioned Theorem III.3.2 of [9], one sees that the Hausdorff and box dimensions of $\mathcal{A}$ are bounded above by

$$\left(\frac{\gamma_2}{\gamma_1}\right)^{\frac{1}{2}}.$$

Notice that

$$\left(\frac{\gamma_2}{\gamma_1}\right)^{\frac{1}{2}} \leqslant \left(\frac{d_1}{d_2}\right)^{\frac{1}{2}} G.$$

3.5 The Weighted Lieb-Thirring Inequality and the Spectral Estimate

We prove the spectral estimate for the operator A and outline the proof of the weighted Sobolev-Lieb-Thirring inequality.

Proposition 3.10. *There exists a constant κ_1 such that the eigenvalues λ_j of the operator A satisfy*

$$\lambda_j \geqslant \kappa_1 j.$$

Proof. The argument follows the proof of Theorem 4.11 of [11]. Recall that $(\boldsymbol{w}_j)$ denotes the sequence of eigenfunctions of A corresponding to the sequence (λ_j) of eigenvalues of A. Let $\alpha_1, \dots, \alpha_j \in \mathbb{R}$ and let

$$\boldsymbol{w} = \sum_{k=1}^{j} \alpha_k \boldsymbol{w}_k.$$

Interpolating between $\mathbf{L}^2_b(\Omega)$ and $\mathbf{H}^2_b(\Omega)$, one has

$$\|\boldsymbol{w}\|_{\mathbf{L}^\infty_b(\Omega)} \leqslant k_1 |\boldsymbol{w}|^{1/2}_{\mathbf{L}^2_b(\Omega)} \|\boldsymbol{w}\|^{1/2}_{\mathbf{H}^2_b(\Omega)}.$$

The Agmon-Douglis-Nirenberg elliptic regularity estimate (3.4) gives

$$\|\boldsymbol{w}\|_{\mathbf{H}^2_b(\Omega)} \leqslant c(1 + c_2)|A\boldsymbol{w}|$$

and hence

$$\|\boldsymbol{w}\|_{\mathbf{L}^\infty_b(\Omega)} \leqslant k_2 |\boldsymbol{w}|^{\frac{1}{2}} |A\boldsymbol{w}|^{\frac{1}{2}}. \tag{3.6}$$

Bounding $|A\boldsymbol{w}|^2$, we have

$$\begin{aligned} |A\boldsymbol{w}|^2 &= \sum_{k=1}^{j} \lambda_k^2 \alpha_k^2 \\ &\leqslant \lambda_j^2 \sum_{k=1}^{j} \alpha_k^2. \end{aligned}$$

Applying this bound to (3.6), one has

$$\begin{aligned} \|\boldsymbol{w}\|_{\mathbf{L}^\infty_b(\Omega)} &\leqslant k_2 \left(\sum_{k=1}^{j} \alpha_k^2 \right)^{\frac{1}{4}} \lambda_j^{\frac{1}{2}} \left(\sum_{k=1}^{j} \alpha_k^2 \right)^{\frac{1}{4}} \\ &= k_2 \lambda_j^{\frac{1}{2}} \left(\sum_{k=1}^{j} \alpha_k^2 \right)^{\frac{1}{2}}. \end{aligned}$$

We have established that $|\boldsymbol{w}(\boldsymbol{x})|^2 \leqslant k_3\lambda_j \sum_{k=1}^{j} \alpha_k^2$ for almost every $\boldsymbol{x} \in \Omega$. In fact, this holds for all $\boldsymbol{x} \in \Omega$ by Sobolev embeddings. Let $1 \leqslant i \leqslant 2$. One has

$$\left|\sum_{k=1}^{j} \alpha_k w_k^{(i)}(\boldsymbol{x})\right|^2 \leqslant |\boldsymbol{w}(\boldsymbol{x})|^2 \leqslant k_3\lambda_j \sum_{k=1}^{j} \alpha_k^2.$$

Setting $\alpha_k = w_k^{(i)}(\boldsymbol{x})$, we obtain

$$\sum_{k=1}^{j} |w_k^{(i)}(\boldsymbol{x})|^2 \leqslant k_3\lambda_j.$$

Summing over i,

$$\sum_{k=1}^{j} |\boldsymbol{w}_k(\boldsymbol{x})|^2 \leqslant 2k_3\lambda_j$$

for each $\boldsymbol{x} \in \Omega$. Integration over Ω yields the spectral estimate. $\square$

Proposition 3.11 (Weighted Lieb-Thirring Inequality). *Let* $\{\boldsymbol{\varphi}_j \in V,\ j = 1, \ldots, m\}$ *be an orthonormal set in* H*. For almost every* $\boldsymbol{x} \in \Omega$ *set*

$$\rho(\boldsymbol{x}) = \sum_{j=1}^{m} |\boldsymbol{\varphi}_j(\boldsymbol{x})|^2.$$

For p *satisfying* $1 < p \leqslant 2$ *one has*

$$\left(\int_\Omega \rho(\boldsymbol{x})^{\frac{p}{p-1}}\, d\lambda(\boldsymbol{x})\right)^{p-1} \leqslant \kappa_2 \left(\sum_{j=1}^{m} a(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_j)\right)$$

where κ_2 *is independent of the family* $\{\boldsymbol{\varphi}_j\}$ *and of* m*.*

Proof. One checks that the arguments given in [41] and the appendix of [64] may be adapted to the case in which the weighted measure λ replaces the Lebesgue measure. We proceed initially by assuming that the operator A satisfies the following hypotheses.

- **(H1)** There exists a constant κ_1 such that the eigenvalues λ_j of the operator A satisfy $\lambda_j \geqslant \kappa_1 j$.

- **(H2)** For each $r > 0$, the operator $(A+r)^{-1} \in \mathcal{L}(V', V)$ extends as a linear continuous operator from $\mathbf{L}_b^s(\Omega)$ into $V \cap \mathbf{W}_b^{2,s}(\Omega)$ for $1 < s < \infty$. This operator considered as an operator on $\mathbf{L}_b^2(\Omega)$ is positive.

- **(H3)** The eigenfunctions $\boldsymbol{w}_j$ of A are uniformly bounded in $\mathbf{L}_b^\infty$.

Hypothesis $H3$ is very strong as it is not true in general and often very difficult to verify when true. Donnelly [17] shows that if an n-dimensional compact Riemannian manifold M admits an isometric circle action, and if the metric is generic, then one has eigenfunctions of the Laplacian corresponding to the eigenvalue γ_k satisfying

$$\|\phi_k\|_\infty \geqslant C\gamma_k^{\frac{n-1}{8}}\|\phi\|_2.$$

Let $p > 2$ and let $f \in L_b^p(\Omega)$. The form

$$a(\boldsymbol{u}, \boldsymbol{v}) + \int_\Omega (f+\alpha)\boldsymbol{u}\cdot\boldsymbol{v}\, d\lambda$$

is bilinear, continuous, and coercive on V for an appropriate choice of the translate α. Therefore, H has an orthonormal basis consisting of eigenfunctions of the Schrödinger-type operator $A_f = A+f$. Let $(\mu_j(f))$ denote the increasing sequence of eigenvalues of $A+f$. Using the Birman-Schwinger inequality [60], one obtains an estimate on the negative part of the spectrum of $A+f$ in terms of a phase space integral involving f. For $0 < \beta \leqslant 1$, there exists $\gamma_1 = \gamma_1(\beta)$ such that

$$\sum_{\mu_j<0} |\mu_j| \leqslant \gamma_1 \left[\int_\Omega (f_-(\boldsymbol{x}))^{\beta+1} d\lambda\right]^{\frac{1}{\beta}}.$$

This spectral estimate makes crucial use of (H3). The weighted Sobolev-Lieb-Thirring inequality now follows by setting $f = -\alpha\rho^{1/(p-1)}$ for an appropriate value of α and studying the unbounded operator A_f^m on $\bigwedge^m H$ defined by

$$A_f^m(\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_m) = (A_f\boldsymbol{u}_1 \wedge \boldsymbol{u}_2 \wedge \cdots \wedge \boldsymbol{u}_m) + \cdots + (\boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_{m-1} \wedge A_f\boldsymbol{u}_m).$$

The general weighted Sobolev-Lieb-Thirring inequality reduces to the case of the negative Laplacian with periodic boundary conditions, an operator for which (H1)-(H3) hold. □

Remark 3.12. See [26] for other interesting generalizations of the Sobolev-Lieb-Thirring inequalities.

Acknowledgements. The author thanks C.D. Levermore for posing the problem and for numerous insightful discussions, R. Pego for clarifying the spectral theory of elliptic operators, and J. Sterbenz for pointing out the estimates in [17].

Chapter 4

Learning About Reality From Observation

Written in collaboration with James A. Yorke, the material in this chapter has been published by SIAM. Publication data: W. Ott and James A. Yorke, *Learning About Reality From Observation*, SIAM J. Applied Dynamical Systems **2** (2003), no. 3, 297-322.

4.1 Introduction

In *The Republic*, Plato writes of people who are chained in a cave for all of their lives, unable to observe life directly. Behind these people a fire burns and real objects cast shadows on the cave wall for them to see. Forced to base their knowledge of reality on inferences made from the shadows, they equate the shadows with reality. While philosophers may vigorously debate epistemological theory, it is certainly true that experimentalists are limited to observations that may not encode the full complexity of their systems.

As Ruelle and Takens have observed, it is very difficult to directly observe all aspects of the evolution of a high dimensional dynamical system such as a turbulent flow. Out of necessity, it is frequently the case that experimentalists study such systems by measuring a relatively low number of different quantities.

We assume that all measurements have infinite precision in what follows. A central experimental question is the following.

Question 4.1. Is the measured data sufficient for us to understand the evolution of the dynamical system? In particular, does the measured data contain enough information to reconstruct dynamical objects of interest and recover coordinate independent dynamical properties such as attractor dimension and Lyapunov exponents? How many exponents can be determined?

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a map and suppose $A \subset \mathbb{R}^n$ is a compact invariant set. Let $\phi : \mathbb{R}^n \to \mathbb{R}^m$ be a smooth map. We always assume $m > 0$. We think of ϕ as a measurement function measuring m physical quantities, and for each point x in the state space $\mathbb{R}^n$ we say that $\phi(x)$ is the **measurement** associated with x. Motivated by an experimental point of view, we say that **observations are deterministic** if there exists an induced map $\bar{f}$ on $\phi(A)$ such that the following diagram commutes:

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "\phi"'] & A \arrow[d, "\phi"] \\ \phi(A) \arrow[r, "\bar{f}"] & \phi(A) \end{tikzcd}$$

The dynamics generated by $\bar{f}$ may be thought of as the shadows that traverse Plato's hypothetical cave wall. The global goal is to infer as much as possible about the dynamical system f from knowledge of the induced dynamics. In the absence of induced dynamics, experimenters increase m by either making more measurements or using delay coordinate maps. Assuming $\bar{f}$ exists, there is a considerable literature on how to compute the Lyapunov exponents associated with the induced system. Do these values correspond to those of the full system? What do we need to check to see this? We would like to state theorems of the following type.

Prototypical Theorem. *For a typical measurement map ϕ, if the induced map $\bar{f}$ exists and has certain properties, then the measurement map ϕ preserves dynamical objects of interest and dynamical invariants of the full system may be computed from the induced dynamics.*

Under what conditions do our observations allow us to make predictions? James Clerk Maxwell wrote of the fundamental importance of continuous dependence on initial data [8, 35]:

> "It is a metaphysical doctrine that from the same antecedents follow the same consequents. No one can gainsay this. But it is not of much use in a world like this, in which the same antecedents never again concur, and nothing ever happens twice.... The physical axiom which has a somewhat similar aspect is 'That from like antecedents follow like consequents'."

We ask what we can conclude if observations are deterministic and if the induced map $\bar{f}$ is continuous. Using a translation invariant concept of "almost every" on infinite dimensional vector spaces described in Section 4.2, we obtain the main C^0 conclusion.

Notation 4.2. For a map ψ we denote the restriction of ψ to a subset S of the domain of ψ by $\psi[S]$. Notice that this notation is not standard.

Let $\mathrm{Fix}(\bar{f})$ and $\mathrm{Per}_2(\bar{f})$ denote the collection of fixed points and period two points, respectively, of $\bar{f}$.

C^0 **Theorem.** *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a map and let A be a compact invariant set. For almost every map $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, there is an induced map $\bar{f}$ satisfying*

1. $\bar{f}$ *is continuous and invertible, and*

2. $\mathrm{Fix}(\bar{f})$ *and* $\mathrm{Per}_2(\bar{f})$ *are countable*

if and only if the following hold.

1. *The measurement map* ϕ *is one to one on* A*.*

2. *The sets* $\mathrm{Fix}(f[A])$ *and* $\mathrm{Per}_2(f[A])$ *are countable.*

3. *The map* $f[A]$ *is continuous and invertible.*

Remark 4.3. If one can infer a property of A from a corresponding property of $\phi(A)$, we say that the property is **observable**. The boundedness of A is observable in the sense that if A is unbounded, then $\phi(A)$ is unbounded for almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$. This applies to each of the embedding theorems in this paper.

Remark 4.4. Our goal is to obtain results with few or preferably no assumptions on f and A. Hypotheses should instead be placed on the observed objects, $\phi(A)$ and $\bar{f}$. This point of view motivates the definition of a Platonic result.

Definition 4.5. A result is said to be **Platonic** if it contains no hypotheses on the dynamical system f aside from the assumption of a finite-dimensional Euclidean phase space.

Does a typical measurement function preserve differential structure? If f is a diffeomorphism, A is a smooth submanifold of $\mathbb{R}^n$ and $\dim(A)$ is known a priori, one may appeal to the Whitney embedding theorem [29]. This theorem states that if A is a compact C^r k-dimensional manifold, where $r \geq 1$, then there is a C^r embedding of A into $\mathbb{R}^m$ where $m \geq 2k+1$. This situation is generic in the sense that the set of embeddings of A is open and dense in $C^r(A, \mathbb{R}^m)$. However, the experimentalist lacking a priori knowledge of the structure of A cannot rely on embedding theorems of Whitney type.

In Section 4.3 we define a notion of tangent space, denoted T_xA, suitable for a general compact subset A of $\mathbb{R}^n$ and we prove a manifold extension theorem. This result allows us to prove a Platonic version of the Whitney embedding theorem and to formulate a notion of diffeomorphism on A equivalent to the notion of injective immersion on A. We formulate our C^1 embedding theorems using this notion of diffeomorphism. Our Platonic C^1 theorem states that for almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, the existence of an invertible quasidifferentiable (see Section 4.6) induced map $\bar{f}$ on $\phi(A)$ satisfying mild assumptions implies that ϕ is a diffeomorphism on A.

Platonic C^1 Theorem. *Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is a map. For almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, if there exists an invertible quasidifferentiable (see Section 4.6) induced map $\bar{f}$ on $\phi(A)$ satisfying*

1. $\mathrm{Fix}(\bar{f})$ *and* $\mathrm{Per}_2(\bar{f})$ *are countable,*
2. $\dim T_y(\phi(A)) < m \; \forall y \in \phi(A)$, *and*
3. $D\bar{f}(y)[T_y\phi(A)]$ *is invertible* $\forall y \in \phi(A)$,

then the measurement mapping ϕ is a diffeomorphism on A.

It is difficult for a scientist to measure a large number of independent quantities simultaneously. For this reason one introduces the class of delay coordinate mappings. This mapping class was introduced into the literature by Takens [63].

Definition 4.6. Let $g \in C^1(\mathbb{R}^n, \mathbb{R})$. The **delay coordinate map** $\phi(f,g) : \mathbb{R}^n \to \mathbb{R}^m$ is given by

$$\phi(f,g)(x) = (g(x), g(f(x)), \ldots, g(f^{m-1}(x)))^{\mathrm{T}}$$

Analogs of several of our embedding results hold for the class of delay coordinate mappings. Since the delay coordinate mappings form a subspace of $C^1(\mathbb{R}^n, \mathbb{R}^m)$, it should be stressed that the delay coordinate results do not follow from the corresponding results about almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$. The following result addresses the observation of differentiable dynamics.

Delay Coordinate Map Theorem. *Let f be a diffeomorphism on $\mathbb{R}^n$ and let A be a compact invariant set. For almost every $g \in C^1(\mathbb{R}^n, \mathbb{R})$, if there is a quasidifferentiable induced map $\bar{f}$ satisfying*

1. *$\bigcup_{i=1}^{2m} \mathrm{Per}_i(\bar{f})$ is countable and*
2. *for each $p \in \{1, \ldots, m\}$ and $y \in \mathrm{Per}_p(\bar{f})$ we have*

$$D\bar{f}^p(y)[T_y\phi(f,g)(A)] \neq \gamma \cdot I \text{ for every } \gamma \in \mathbb{R},$$

then the delay map $\phi(f,g)$ is a diffeomorphism on A.

Assume that f and $\bar{f}$ are quasidifferentiable and invertible on A and $\phi(A)$, respectively, with invertible quasiderivatives at each point $x \in A$ and $y \in \phi(A)$. Suppose that ϕ is a diffeomorphism on A. We say that a Lyapunov exponent $\lambda(y,v)$ of $\bar{f}$ at $y \in \phi(A)$ is **true** if it does not depend on the choice of quasiderivative $D\bar{f}$ and if it is also a Lyapunov exponent of f at $\phi^{-1}(y) \in A$. The works of Eckmann, Ruelle, Sano and Sawada provide heuristic computational procedures for obtaining m Lyapunov exponents for a trajectory (y_k) of $\bar{f}$. They use the subset of measurement mappings generated by so-called delay coordinate mappings, the mapping class considered in the famous, fundamental paper of Takens [63]. In particular, the Eckmann and Ruelle algorithm (ERA) [18] uses a linear fitting of the tangent map and has proven to be computationally efficient in giving the complete Lyapunov spectrum of many dynamical systems. Mera and Morán [48] find conditions ensuring the convergence of this algorithm for a smooth dynamical system on a $C^{1+\alpha}$ submanifold supporting an ergodic invariant Borel probability measure. Our exponent characterization theorem establishes a rigorous connection between the observed Lyapunov exponents and the Lyapunov exponents of $f[A]$. Under our assumptions, an observed Lyapunov exponent $\lambda(y, v)$ is a true Lyapunov exponent if and only if $v \in T_y\phi(A)$.

Suppose A is a manifold of dimension d. Implementation of the full Eckmann and Ruelle algorithm yields m observed Lyapunov exponents, d of which are true. The remaining $m - d$ exponents are spurious, artifacts of the embedding process. In order to identify the d true exponents, one must either devise a method to identify the spurious exponents a fortiori or modify ERA to completely avoid the computation of spurious exponents. Several authors propose a modified ERA in which the tangent maps are computed only on the tangent spaces and not on the ambient space $\mathbb{R}^m$. Mera and Morán [49] discuss the convergence of the modified ERA. This technique eliminates the computation of spurious exponents but requires that tangent spaces be computed along orbits. We propose a new technique based on the exponent characterization theorem that allows for the a fortiori determination of the spurious exponents without requiring the computation of tangent spaces along orbits. We describe this algorithm in Section 4.7 following the statement of the exponent characterization theorem.

4.1.1 The case of linear f and ϕ

We illustrate our ansatz with the case where f and ϕ are linear.

Proposition 4.7. *Let f be linear on $\mathbb{R}^n$ and let A be an invariant subspace on which f is an isomorphism. If the restriction of f to A is not a scalar multiple of the identity, then for almost every $\phi \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)$ in the sense of Lebesgue measure, there is an induced map on $\phi(A)$ if and only if ϕ is an isomorphism on A.*

Key issues are raised by this proposition. Notice that if there exists $c \in \mathbb{R}$ for which $f(x) = cx$ for all $x \in A$, then $y \mapsto cy$ is the induced map on $\phi(A)$ even if ϕ is not one to one on A. Since this is a theory of observation, when possible the assumptions should be verifiable from observation. The following alternative version of the proposition transfers the assumption onto the induced dynamics in a manner that will be followed throughout this paper.

Proposition 4.8. *Let f be linear on $\mathbb{R}^n$ and let A be an invariant subspace on which f is an isomorphism. For almost every $\phi \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)$, there is an induced map on $\phi(A)$ and this induced map is not identically a scalar multiple of the identity if and only if ϕ is an isomorphism on A and the restriction of f to A is not a scalar multiple of the identity.*

Remark 4.9. The hypothesis that f is an isomorphism on A is observable in the sense mentioned earlier. The key point is that if $f[A]$ is not one-to-one, then for almost every $\phi \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)$ there does not exist an injective induced map $\bar{f}$ on $\phi(A)$.

4.1.2 What does "typical" mean?

The conclusions of the linear propositions hold for almost every linear ϕ with respect to Lebesgue measure. In the general situation we will consider the space of C^1 measurement mappings. In order to prove versions of our prototypical theorem, we must first clarify what we mean by a "typical" measurement mapping ϕ. The notion of typicality may be cast in topological terms. In this setting, "typical" would be used to refer to an open and dense subset or a residual subset of mappings. For example, consider the topological Kupka-Smale theorem.

Definition 4.10. Let M be a smooth, compact manifold. A diffeomorphism $f \in \mathrm{Diff}^r(M)$ is said to be **Kupka-Smale** if

1. The periodic points of f are hyperbolic.

2. If p and q are periodic points of f, then $W^s(p)$ is transverse to $W^u(q)$.

Theorem 4.11 (Kupka-Smale [52]). *The set of Kupka-Smale diffeomorphisms is residual in* $\mathrm{Diff}^r(M)$.

The topological notion of typicality is not the appropriate conceptualization for the experimentalist interested in a probabilistic result on the likelihood of a given property in a function space. Any Cantor set of positive measure illustrates the difference between the topological and measure theoretic notions of a small set. The discord between topological typicality and probabilistic typicality is also evident in the following dynamical examples.

Example 4.12. Arnold [1] studied the family of circle diffeomorphisms

$$f_{\omega,\epsilon}(x) = x + \omega + \epsilon \sin(x) \pmod{2\pi},$$

where $0 \leq \omega \leq 2\pi$ and $0 \leq \epsilon < 1$ are parameters. For each ϵ we define the set

$$S_\epsilon = \{\omega \in [0, 2\pi] : f_{\omega,\epsilon} \text{ has a stable periodic orbit}\}.$$

For $0 < \epsilon < 1$, the set S_ϵ is a countable union of disjoint open intervals (one for each rational rotation number) and is an open dense subset of $[0, 2\pi]$. However, the Lebesgue measure of S_ϵ converges to 0 as $\epsilon \to 0$.

There are even more striking examples where the Baire categorical and measure theoretic notions of typicality yield diametrically opposite conclusions about the size of a set.

Example 4.13. Misiurewicz [50] proved that the mapping $z \mapsto e^z$ on the complex plane is topologically transitive, implying that a residual set of initial points yield dense trajectories. On the other hand, Lyubich [43] and Rees [53] proved that Lebesgue almost every initial point has a trajectory whose limit set is a subset of the real axis.

Finally, we consider Lyapunov exponents. This example is particularly relevant because the work of Eckmann, Ruelle, Sano, and Sawada on the computation of these exponents motivated this paper.

Example 4.14 (Lyapunov Exponents). Let $f : M \to M$ be a C^1 diffeomorphism on a compact finite-dimensional Riemannian manifold M. For $(x, v) \in TM$, $\|v\| \neq 0$, the number

$$\lim_{n \to \pm\infty} \frac{1}{n} \log \|Df^n(x)v\|$$

should the limit exist is called the Lyapunov exponent of f at (x, v), denoted $\lambda(x, v)$. We say that $x \in M$ is a **regular point** for f if there are Lyapunov exponents

$$\lambda_1(x) > \cdots > \lambda_l(x)$$

and a splitting

$$T_xM = \bigoplus_{i=1}^{l} E_i(x)$$

of the tangent space to M at x such that

$$\lim_{n\to\pm\infty} \frac{1}{n} \log \|Df^n(x)u\| = \lambda_j(x) \quad (u \in E_j(x) \setminus \{0\} \text{ and } 1 \leq j \leq l).$$

While the periodic points of f are always regular points, frequently the set of regular points is a topologically small subset of M. Quite often this set is Baire first category and it may even be finite [65]. From a measure theoretic point of view the situation is completely different.

Theorem 4.15 (Oseledec Multiplicative Ergodic Theorem [65, 37]). *The set of regular points for f has full measure with respect to any f-invariant Borel probability measure on M.*

The Oseledec theorem holds in the more general context of measurable cocycles over invertible measure-preserving transformations of a Lebesgue space (X, μ) [37]. Let $f : X \to X$ be an invertible measure preserving transformation and let $L : X \to GL(n, \mathbb{R})$ be a measurable cocycle over X. If

$$\log^+ \|L^{\pm 1}(x)\| \in L^1(X, \mu)$$

then almost every $x \in X$ is a regular point for (f, L).

The following example illustrates that Lyapunov exponents may not exist for a residual set of points. Let $p > 1$ and $q > 1$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$ and $p \neq q$.

Consider the Markov map $f : [0,1] \to [0,1]$ defined by

$$f(x) = \begin{cases} px, & \text{if } 0 \leqslant x < \frac{1}{p}; \\ qx - \frac{q}{p}, & \text{if } \frac{1}{p} \leqslant x \leqslant 1. \end{cases}$$

This transformation represents the full shift on two symbols with probabilities $1/p$ and $1/q$. Lebesgue measure is invariant under f and ergodic, thus the Lyapunov exponent at Lebesgue almost every $x \in [0,1]$ exists and is equal to

$$\frac{\log(p)}{p} + \frac{\log(q)}{q}$$

by virtue of the Birkhoff ergodic theorem. On the other hand, we claim that no Lyapunov exponent exists for a residual set of points. For $n \in \mathbb{N}$, set

$$V_{p,n}(x) = \frac{1}{n}(|\{0 \leqslant i \leqslant n-1 : f^i(x) \in [0, 1/p)\}|).$$

Fix $\alpha > 1/p$ and $\beta < 1/p$. Define for each $N \in \mathbb{N}$ the sets $C_N = \{x : \exists n \geqslant N$ for which $V_{p,n}(x) \geqslant \alpha\}$ and $D_N = \{x : \exists n \geqslant N$ for which $V_{p,n}(x) \leqslant \beta\}$. The set C_N contains an open interval to the right of each preimage of $1/p$, and thus C_N contains an open and dense subset of $[0,1]$. Similarly, D_N contains an open interval to the left of each preimage of $1/p$, and thus D_N also contains an open and dense subset of $[0,1]$. No Lyapunov exponent exists for points in the residual set

$$\bigcap_{N=1}^{\infty} C_N \cap D_N$$

because $V_{p,n}(x)$ does not converge for such points.

Motivated by the probabilistic interpretation of typicality, we will use the notion of prevalence developed in [33, 34]. See the references given in [34] for closely related concepts. The notion of prevalence generalizes the translation invariant concept of Lebesgue full measure to infinite-dimensional Banach spaces.

4.1.3 Overview of this paper

Section 4.2 develops the relevant prevalence theory and demonstrates that cardinality and boundedness are observable properties. In §4.3 we define a notion of tangent space suitable for general compact subsets of $\mathbb{R}^n$ and we prove the manifold extension theorem. The manifold extension theorem is used in §4.4 to derive a Platonic version of the Whitney embedding theorem. We present our embedding theorems in §4.5 and §4.6 and our results on delay coordinate mappings and Lyapunov exponents in §4.7.

4.1.4 The Transference Method

Schematically our embedding theorems are developed in the following way. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a dynamical system and let A be a compact invariant set. We want to require no regularity assumptions about f nor do we wish to assume that f is invertible. For a map g, a subset D of the domain of g and any property L, write $(g, L; D)$ to indicate that the restriction of g to D has property L. Let $\mathcal{S}$ denote a collection of properties of a dynamical system. Let $\mathcal{Q}$ denote a collection of properties of maps in the measurement function space $C^1(\mathbb{R}^n, \mathbb{R}^m)$. For example, $\mathcal{Q}$ might consist of the assertion that $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ is a homeomorphism on A. We are interested in the ability of the observer to make inferences; that is, in results of the form

$$(\bar{f}, \mathcal{L}; \phi(A)) \Rightarrow (\phi, \mathcal{Q}) \text{ for almost every } \phi, \tag{4.1}$$

where $\mathcal{L}$ is a collection of properties of $\bar{f}$. In other words, the existence of an induced map $\bar{f}$ satisfying properties $\mathcal{L}$ implies that ϕ satisfies properties $\mathcal{Q}$. We first prove

$$(f, \mathcal{S}; A) \Rightarrow ((\bar{f}, \mathcal{L}_1; \phi(A)) \Leftrightarrow (\phi, \mathcal{Q})) \text{ for a.e. } \phi.$$

The Platonic version of the theorem is obtained by replacing each assumption on f with one on $\bar{f}$. For $P \in \mathcal{S}$, we replace the assumption

$$(f, P; A)$$

with one on $\bar{f}$, giving

$$(\bar{f}, \mathcal{L}_1 \cup \mathcal{S}; \phi(A)) \Leftrightarrow ((\phi, \mathcal{Q}) \text{ and } (f, \mathcal{S}; A)) \text{ for a.e. } \phi.$$

In particular, (4.1) holds with $\mathcal{L} = \mathcal{L}_1 \cup \mathcal{S}$. In essence the Platonic version has been obtained by transferring the hypotheses $(f, P; A)$ for $P \in \mathcal{S}$ onto the induced dynamics. Prevalence statements allow for these transfers. Properties for which this program may be implemented are said to be **observable**.

4.2 Prevalence

Let V be a complete metric linear space.

Definition 4.16. A Borel measure μ on V is said to be **transverse** to a Borel set $S \subset V$ if the following holds:

1. There exists a compact set $U \subset V$ for which $0 < \mu(U) < \infty$, and
2. for every $v \in V$ we have $\mu(S + v) = 0$.

For example, μ might be Lebesgue measure supported on a finite-dimensional subspace of V.

Definition 4.17. A Borel set $S \subset V$ is called **shy** if there exists a measure transverse to S. More generally, a subset of V is called shy if it is contained in a shy Borel set. The complement of a shy set is called a **prevalent** set.

A subset of $\mathbb{R}^n$ is shy if and only if it has Lebesgue measure zero. For a map ϕ contained in a prevalent subset S of a linear function space V, we say that ϕ is typical. Employing the language of the finite dimensional case, we say that **almost every** element of V lies in S (in the sense of prevalence).

Using the notion of prevalence, researchers have reformulated several topological and dynamical theorems. Sauer, Yorke, and Casdagli prove in [59] a prevalence version of the Whitney embedding theorem.

Theorem 4.18 (Prevalence Whitney Embedding Theorem [59]). *Let A be a compact subset of $\mathbb{R}^n$ of box dimension d and let m be an integer greater than $2d$. For almost every smooth map $\phi : \mathbb{R}^n \to \mathbb{R}^m$,*

1. *ϕ is one to one on A and*
2. *ϕ is an immersion on each compact subset C of a smooth manifold contained in A.*

This theorem is not Platonic because the dimension assumption is on A. In Section 4.4 we prove a Platonic Whitney embedding theorem as a corollary of the manifold extension theorem.

The reformulation of a genericity theorem of Kupka-Smale type requires a notion of prevalence for nonlinear function spaces such as the space of diffeomorphisms of a compact smooth manifold. Kaloshin in [36] develops such a notion and proves a prevalence version of the Kupka-Smale theorem for diffeomorphisms.

4.2.1 Cardinality Preservation

In Sections 4.5, 4.6 and 4.7 we will need to know how a typical smooth projection affects the cardinality of a set. We show that for a set $A \subset \mathbb{R}^n$, A and $\phi(A)$ have the same cardinality for a.e. $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$. We begin by assuming that A is a countable set.

Proposition 4.19. *Let $A \subset \mathbb{R}^n$ be countable. Almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ is one to one on A. In particular, if A is countably infinite, then $\phi(A)$ is also countably infinite for almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$.*

Proof. We write $A = \{x_i : i \in \mathbb{N}\}$. For $i \neq j$ let $C_{ij} = \{\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m) : \phi(x_i) = \phi(x_j)\}$. We first show that C_{ij} is shy. Let $B(x_i, r_i)$ be a metric ball such that $x_j \notin B(x_i, r_i)$. Let $\beta : \mathbb{R}^n \to \mathbb{R}$ be a C^∞ map such that

1. $\beta > 0$ on $B(x_i, r_i)$ and
2. $\mathrm{supp}(\beta) = \overline{B(x_i, r_i)}$.

Let $v \in \mathbb{R}^m$ be a nonzero vector and let μ be the Lebesgue measure supported on the one dimensional subspace

$$\{tv\beta : t \in \mathbb{R}\}.$$

For any $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, it is evident that $\phi + tv\beta \in C_{ij}$ for at most one $t \in \mathbb{R}$. Thus C_{ij} is a shy subset of $C^1(\mathbb{R}^n, \mathbb{R}^m)$ because μ is transverse to it. The set

$$\bigcap_{\substack{i,j \in \mathbb{N} \\ i \neq j}} C^1(\mathbb{R}^n, \mathbb{R}^m) \setminus C_{ij}$$

consists of functions that map A injectively into $\mathbb{R}^m$. This set is prevalent because the countable intersection of prevalent sets is prevalent (see [33]). $\square$

Plato would have us consider the prisoner's question where the cardinality of A is not known a priori. For a typical ϕ, does the countability of $\phi(A)$ imply the countability of A? The next proposition answers this question affirmatively with the help of the following lemma.

Lemma 4.20. *Let $A_0 \subset \mathbb{R}^n$ be an uncountable set. Lebesgue almost every function $\phi \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)$ maps A_0 to an uncountable set.*

Proof. It suffices to consider the scalar case $m = 1$. For each $\phi \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$ there exists a unique vector $v \in \mathbb{R}^n$ such that $\phi(x) = (x, v)$ for all $x \in \mathbb{R}^n$. Suppose by way of contradiction that the set

$$\{\phi \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}) : \phi(A_0) \text{ is countable}\}$$

has positive measure. This implies that there exist n linearly independent vectors $\{v_i : i = 1, \ldots, n\}$ such that the functions ϕ_{v_i} given by $x \mapsto (x, v_i)$ map A_0 to a countable set. Let A_1 be an uncountable subset of A_0 such that $\phi_{v_1}(A_1) = \{y_1\}$. Inductively construct a collection of sets $\{A_i : i = 1, \ldots, n\}$ satisfying

1. A_i is uncountable for each i,
2. $A_i \subset A_{i-1}$ for each i, and
3. $\phi_{v_i}(A_i) = \{y_i\}$.

We have $\phi_{v_i}(A_n) = \{y_i\}$ for each i, so A_n consists of one point. This contradiction establishes the lemma. $\square$

Proposition 4.21. *Let A_0 be an uncountable set. For almost every*

$$\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m),$$

$\phi(A_0)$ is uncountable.

Proof. Once again it suffices to consider the scalar case $m = 1$. We show that the set

$$S = \{\phi \in C^1(\mathbb{R}^n, \mathbb{R}) : \phi(A_0) \text{ is countable}\}$$

is shy. Let $\{\phi_{e_i}\}$ be a basis for $\mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$ and let μ be the Lebesgue measure on $\mathbb{R}^n$. Write $\alpha = (\alpha_i)$ for a vector in $\mathbb{R}^n$ and for $\phi \in C^1(\mathbb{R}^n, \mathbb{R})$ set

$$\phi_\alpha := \phi + \sum_{i=1}^{n} \alpha_i \phi_{e_i}.$$

If S is not shy, there exists some $g \in S$ such that

$$\mu\{\alpha : g_\alpha(A_0) \text{ is countable}\} > 0$$

where μ denotes n dimensional Lebesgue measure. Without loss of generality assume that $g(A_0)$ is countable. There is at least one point y such that $g^{-1}(y) \cap A_0$ is uncountable. Shrinking A_0 if necessary, without loss of generality we may assume that g maps A_0 to a single point; that is, g is constant on A_0. There exist n linearly independent vectors $\{v_i\}$ such that the functions $\phi_{v_i} + g$ map A_0 to a countable set. As in the proof of (4.20) we inductively construct a collection of sets $\{A_i : i = 1, \ldots, n\}$ satisfying

1. A_i is uncountable for each i,

2. $A_i \subset A_{i-1}$ for each i, and

3. $(\phi_{v_i} + g)(A_i) = \{y_i\}$.

We have $(\phi_{v_i} + g)(A_n) = \{y_i\}$ for each i, so A_n consists of one point. This contradiction establishes the proposition. $\square$

4.2.2 Preservation of Unboundedness

We now consider the question of how a typical smooth projection affects the boundedness of a set. For a typical ϕ, does the boundedness of $\phi(A)$ imply that A is bounded?

Proposition 4.22 (Unboundedness Preservation). *Assume $A \subset \mathbb{R}^n$ is unbounded. Then $\phi(A)$ is unbounded for almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$.*

Proof. It suffices to assume $m = 1$. We show that the set

$$V = \{\phi \in C^1(\mathbb{R}^n, \mathbb{R}) : \phi(A) \text{ is bounded}\}$$

is shy. As above, let μ be the Lebesgue measure on $\mathbb{R}^n$ and for $\phi \in C^1(\mathbb{R}^n, \mathbb{R})$ and $(\alpha_i) \in \mathbb{R}^n$ write

$$\phi_\alpha := \phi + \sum_{i=1}^{n} \alpha_i \phi_{e_i}.$$

If V is not shy, there exists some $g \in V$ such that

$$\mu\{\alpha : g_\alpha(A) \text{ is bounded}\} > 0.$$

Without loss of generality assume that $g(A) \subset [-d, d]$ for some $d > 0$. There exist n linearly independent vectors $\{v_i\}$ and scalars $c_i > 0$ such that the functions $g + \phi_{v_i}$ map A into $[-c_i, c_i]$. Thus A is contained in the set

$$\bigcap_{i=1}^{n} \phi_{v_i}^{-1}([-c_i - d, c_i + d]),$$

a bounded solid polygon. This contradiction establishes the proposition. $\square$

Remark 4.23. We conclude that for a typical $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, the boundedness of $\phi(A)$ implies that A is bounded. That is, the boundedness of A is an observable property.

4.3 Enveloping Manifolds

Let A be a compact subset of $\mathbb{R}^n$ and let $x \in A$. We say that a C^1 manifold M is an enveloping manifold for A at x if there exists a neighborhood $N(x)$ of x such that $M \supset N(x) \cap A$ and if the dimension of M is minimal with respect to this property. We demonstrate the existence of a C^1 enveloping manifold M for each $x \in A$.

Definition 4.24. Let D_xA be the set of all directions v for which there exist sequences (y_i) and (z_i) in A such that $y_i \to x$, $z_i \to x$, and $\frac{z_i - y_i}{\|z_i - y_i\|} \to v$. The tangent space at x relative to A, denoted T_xA, is the smallest linear space containing D_xA.

We note that this is one of the two obvious ways to define the tangent space at a point in an arbitrary compact subset of $\mathbb{R}^n$. The other would be to fix $y_i = x$ in the above definition, but the resulting tangent space would be too small for our purposes. In general neither the tangent space itself nor its dimension will vary continuously with $x \in A$. Nevertheless, the tangent space varies upper semicontinuously with $x \in A$. More precisely, we have

Lemma 4.25. *The function $x \mapsto \dim(T_xA)$ is upper semicontinuous on A. In fact, T_xA depends upper semicontinuously on $x \in A$ in the sense that if $x_i \to x$ where $x_i \in A$ and $v_i \to v$ where $v_i \in T_{x_i}A$ then $v \in T_xA$. In other words, $\{(x, v) : x \in A,\ v \in T_xA\}$ is a closed subset of $\mathbb{R}^n \times \mathbb{R}^n$. If T_xA has constant dimension on a set $A_0 \subset A$, then T_xA is continuous on A_0 in the same sense.*

Definition 4.26. The **tangent dimension** of A, denoted $\dim_T(A)$, is given by

$$\dim_T(A) = \max_{x \in A}(\dim T_xA).$$

Example 4.27. In Figure 4.1 the tangent space T_pA is two-dimensional while T_xA is one-dimensional for all other points $x \in A$. Choosing $(y_i) \subset A$ and $(z_i) \subset A$ such that $y_i \to p$, $z_i \to p$, and y_i and z_i lie on a vertical line for each i, we obtain the tangent vector $v \in T_pA$. Thus $\dim_T(A) = 2$.

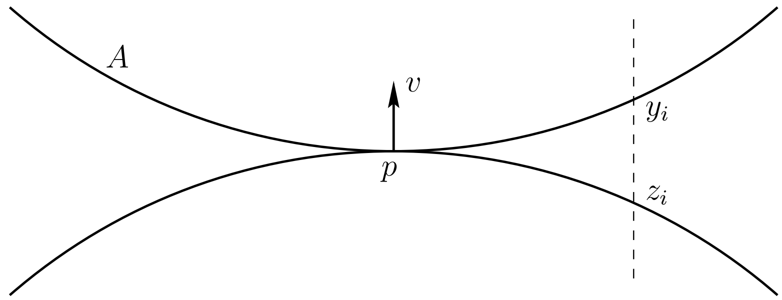

Figure 4.1: A Cusp

We are now in position to state a surprising theorem.

Theorem 4.28 (Manifold Extension Theorem). *For each $x \in A$ there exists an enveloping manifold M for A at x with $T_xM = T_xA$.*

Conjecture 4.29. We believe that integrability is an intrinsic feature of the definition of the tangent space. We therefore conjecture that a global version of the manifold extension theorem holds. Namely, there exists a manifold M such that $\dim(M) = \dim_T(A)$ and $A \subset M$.

Proof. Recall that for a map ψ we denote the restriction of ψ to a subset S of the domain of ψ by $\psi[S]$. Let $m = \dim(T_xA)$. There exists a compact neighborhood N of x such that $\dim(T_yA) \leq m$ for all $y \in N \cap A$. Let π denote the orthogonal projection of $\mathbb{R}^n$ onto T_xA. The projection map π induces the splitting $\mathbb{R}^n = T_xA \oplus E_x$. Using this splitting write (p, q) for points in $\mathbb{R}^n$. If $((p_i, q_i))$ is a sequence such that $(p_i, q_i) \in N \cap A$ for each i and $(p_i, q_i) \to x$ then $\frac{\|q_{i+1}-q_i\|}{\|p_{i+1}-p_i\|} \to 0$. We may assume N has been chosen sufficiently small so that π maps T_yA injectively into T_xA for each $y \in N \cap A$ and that $\pi[N \cap A]$ is one to one. Hence we may define ψ on $\pi(N \cap A)$ by $\psi(p) := q$ for $(p, q) \in N \cap A$. Repeated use of our main technical tool, the Whitney extension theorem, will allow us to extend ψ to a C^1 function defined on a neighborhood in T_xA of $\pi(A \cap N)$. We first state a C^1 version of the Whitney extension theorem for compact domains.

Definition 4.30. Let $Q \subset \mathbb{R}^m$ be a compact set and assume $f : Q \to \mathbb{R}^k$ and $L : Q \to \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^k)$ are given functions.

Notation 4.31.

1. $R(y, z) := \frac{f(z)-f(y)-L(y)\cdot(z-y)}{\|z-y\|}$ (for all $y, z \in Q,\ y \neq z$).

2. For $\delta > 0$, set

$$\rho(\delta) := \sup_{\substack{y,z \in Q \\ 0<\|z-y\|\leq\delta}} \|R(y, z)\|.$$

The pair (f, L) is said to be a **Whitney $\mathbf{C^1}$ function pair** on Q if f and L are continuous and if ρ satisfies

$$\rho(\delta) \to 0 \text{ as } \delta \to 0. \tag{4.2}$$

Notice that (4.2) is equivalent to the following uniformity condition stated by Whitney in [66]: Given any $w \in Q$ and $\epsilon > 0$, there exists $\delta > 0$ such that if $y \in Q$ and $z \in Q$ satisfy $\|y - w\| < \delta$ and $\|z - w\| < \delta$, then $\|R(y, z)\| \leq \epsilon$.

Theorem 4.32 (Whitney Extension Theorem [21, 39, 66]). *Given a Whitney C^1 function pair (f, L) defined on a compact subset Q of $\mathbb{R}^m$, there exists a C^1 function $\tilde{f} : \mathbb{R}^m \to \mathbb{R}^k$ such that $\tilde{f} = f$ and $D\tilde{f} = L$ on Q.*

We now continue the proof of our manifold extension theorem. Let

$$d(y) = \dim(T_y A)$$

for $y \in A \cap N$. For $k \leq m$ let $X_k = \{y \in N \cap A : d(y) = k\}$. We first find a function whose graph is a C^1 manifold which envelops X_m. For each $y \in N \cap A$, the tangent space $T_y A$ may be viewed as a subspace of $T_x A \oplus E_x = \mathbb{R}^n$. For $y \in X_m$ define the linear operator $L_m(y) : T_x A \to E_x$ as follows. For $(v, w) \in D_y A$ let $L_m(y)v = w$. By linearity $L_m(y)$ is determined on $T_y A$. The linear operator $L_m(y)$ depends continuously on $y \in X_m$ since $T_y A$ depends continuously on $y \in X_m$ by (4.25). The function pair (ψ, L_m) is Whitney C^1 on $\pi(X_m)$ because the uniformity condition of Whitney is implied by (4.24). Notice that the Whitney extension theorem can now only be used to extend $\psi[\pi(X_m)]$ because no obvious candidate exists for $L(y)$ for $y \notin X_m$. By applying the Whitney extension theorem, extend ψ to a function $\tilde{\psi}_1$ defined on $\pi(N)$. Notice that if $X_m = N \cap A$, the result is proved since the graph of $\tilde{\psi}_1$ constitutes an enveloping manifold for A at x.

The general case is handled inductively. Construct $\tilde{\psi}_1$ as above and make the nonlinear change of variable $(p, q) \to (p, q - \tilde{\psi}_1(p)) := (p, \psi_2(p))$. Consider the map $\psi_2[\pi(X_m) \cup \pi(X_{m-1})]$ and let $y \in \mathrm{graph}(\psi_2[\pi(X_m) \cup \pi(X_{m-1})])$. The tangent space $T_y(\mathrm{graph}(\psi_2[\pi(A)]))$ may be viewed as a subspace of $T_x A \oplus E_x = \mathbb{R}^n$. Define the linear map $L_{m-1}(y) : T_x A \to E_x$ as follows. If $y \in \mathrm{graph}(\psi_2[\pi(X_m)])$, set $L_{m-1}(y) \equiv 0$. If $y \in \mathrm{graph}(\psi_2[\pi(X_{m-1})])$, enlarge $T_y(\mathrm{graph}(\psi_2[\pi(A)]))$ to a linear space $\tilde{T}_y$ of dimension m by adjoining one vector in $T_x A$ orthogonal to $T_y(\mathrm{graph}(\psi_2[\pi(A)]))$. For $(v, w) \in \tilde{T}_y$ let $L_{m-1}(y)v = w$. The linear operator $L_{m-1}(y)$ depends continuously on $y \in \mathrm{graph}(\psi_2[\pi(X_m) \cup \pi(X_{m-1})])$ by (4.25). The function pair (ψ_2, L_{m-1}) is Whitney C^1 on $\pi(X_m) \cup \pi(X_{m-1})$ because the uniformity condition of Whitney is implied by (4.24). By applying the Whitney extension theorem, extend $\psi_2[\pi(X_m)\cup\pi(X_{m-1})]$ to a function $\tilde{\psi}_2$ defined on $\pi(N)$. Make the nonlinear change of variables $(p,q)\to(p,q-\tilde{\psi}_2(p))=(p,\psi_3(p))$.

Assume now that the functions $\tilde{\psi}_1,\tilde{\psi}_2,\dots,\tilde{\psi}_{k-1}$ and ψ_k have been constructed. Consider the map

$$\psi_k\left[\bigcup_{i=m-k+1}^{m}\pi(X_i)\right].$$

For each point y in the set

$$\operatorname{graph}\left(\psi_k\left[\bigcup_{i=m-k+1}^{m}\pi(X_i)\right]\right)$$

the tangent space $T_y(\operatorname{graph}(\psi_k[\pi(A)]))$ may be viewed as a subspace of $T_xA\oplus E_x=\mathbb{R}^n$. Define the linear map $L_{m-k+1}(y):T_xA\to E_x$ as follows. If $y\in\operatorname{graph}(\psi_k[\pi(X_m)\cup\cdots\cup\pi(X_{m-k+2})])$, set $L_{m-k+1}(y)\equiv 0$. On the other hand, if $y\in\operatorname{graph}(\psi_k[\pi(X_{m-k+1})])$, enlarge $T_y(\operatorname{graph}(\psi_k[\pi(A)]))$ to a linear space $\tilde{T}_y$ of dimension m by adjoining $k-1$ vectors in T_xA orthogonal to

$$T_y(\operatorname{graph}(\psi_k[\pi(A)])).$$

For $(v,w)\in\tilde{T}_y$ let $L_{m-k+1}(y)v=w$. By (4.24) and (4.25) the function pair

$$(\psi_k,L_{m-k+1})$$

is Whitney C^1 on the set

$$\bigcup_{i=m-k+1}^{m}\pi(X_i).$$

By applying the Whitney extension theorem, extend the function

$$\psi_k\left[\bigcup_{i=m-k+1}^{m}\pi(X_i)\right]$$

to a function $\tilde{\psi}_k$ defined on $\pi(N)$. Make the change of variables $(p,q)\to(p,q-\tilde{\psi}_k(p)):=(p,\psi_{k+1}(p))$. After $m+1$ steps we obtain a map

$$\Psi:=\sum_{i=1}^{m+1}\tilde{\psi}_i$$

defined on $\pi(N)$. The graph of Ψ constitutes an enveloping manifold M for A at x. $\square$

Remark 4.33. Although our inductive procedure is canonical, observe that the Whitney extension theorem makes no claim of uniqueness. Assume that (f, L_1) and (f, L_2) are Whitney C^1 function pairs defined on a compact subset Q of $\mathbb{R}^m$ as in (4.32). Let $y \in \text{graph}(f)$ and let π denote the orthogonal projection of $\mathbb{R}^m \times \mathbb{R}^k$ onto $\mathbb{R}^m$. The tangent space $T_y(\text{graph}(f))$ may be viewed as a subspace of $\mathbb{R}^m \times \mathbb{R}^k$. The linear operators $L_1(y)$ and $L_2(y)$ must satisfy $L_1(y)v = L_2(y)v = w$ for all $(v, w) \in T_y(\text{graph}(f))$. However, $L_1(y)$ and $L_2(y)$ are determined only for $(v, w) \in T_y(\text{graph}(f))$. If $v \notin \pi(T_y(\text{graph}(f)))$, then $L_1(y)$ and $L_2(y)$ may be such that $L_1(y)v \neq L_2(y)v$.

4.4 Platonic Embedology

Recall the prevalence version of the Whitney embedding theorem.

Theorem 4.34 (Prevalence Whitney Embedding Theorem [59]). *Let A be a compact subset of $\mathbb{R}^n$ of box dimension d and let m be an integer greater than $2d$. For almost every smooth map $\phi : \mathbb{R}^n \to \mathbb{R}^m$,*

1. *ϕ is one to one on A and*
2. *ϕ is an immersion on each compact subset C of a smooth manifold contained in A.*

The manifold extension theorem implies a Platonic version of this result. We need a notion of diffeomorphism appropriate for a general compact subset A of $\mathbb{R}^n$.

Definition 4.35. We say that a measurement map $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ is a **diffeomorphism** on A if ϕ is injective on A and if for each $x \in A$ there exists an enveloping manifold M for A at x that is mapped diffeomorphically onto an enveloping manifold for $\phi(A)$ at $\phi(x)$.

We are now in position to formulate the Platonic Whitney embedding theorem.

Theorem 4.36 (Platonic Whitney Embedding Theorem). *Let $A \subset \mathbb{R}^n$ be compact. For almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, if $\phi(A)$ satisfies $\dim_T \phi(A) < \frac{m}{2}$, then ϕ is a diffeomorphism on A.*

Conjecture 4.37. The Platonic Whitney embedding theorem remains valid under the weaker assumption that $\dim_T \phi(A) < m$.

The proof of this result requires an understanding of the relationship between the box dimension of A and the dimension of the tangent spaces $T_x A$ for $x \in A$. Working only with the definitions, the relationship is unclear. Illumination is provided by the manifold extension theorem.

Lemma 4.38. *Let $A \subset \mathbb{R}^n$ be compact. For each $x \in A$, there exists a neighborhood N of x such that $\dim(T_x A) \geqslant \dim_B(A \cap N)$.*

Proof. Fix $x \in A$. By the manifold extension theorem, there exists an enveloping manifold M for A at x and a neighborhood N of x such that $M \supset N \cap A$. The set $N \cap A$ is contained in a C^1 manifold of dimension $\dim(T_x A)$ and therefore $\dim(T_x A) \geqslant \dim_B(A \cap N)$. □

We now commence with the proof of the Platonic Whitney embedding theorem. Suppose there exists $x \in A$ such that $\dim(T_x A) \geqslant \frac{m}{2}$. In this case we would have that $\dim(T_{\phi(x)}\phi(A)) \geqslant \frac{m}{2}$ for almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ as a consequence of the fact that almost every linear transformation has full rank. Therefore we may assume that $\dim(T_x A) < \frac{m}{2}$ $\forall x \in A$. By the manifold extension theorem and the compactness of A, A is contained in a finite union $\bigcup_{i=1}^{k} M_i$ of enveloping manifolds such that $\dim(M_i) < \frac{m}{2}$ for each i. Box dimension is finitely stable, so one has

$$\dim_B(A) \leqslant \dim_B \left(\bigcup_{i=1}^{k} M_i \right) = \max_i \dim_B(M_i) < \frac{m}{2}.$$

The prevalence version of the Whitney embedding theorem (4.18) implies that almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ is a diffeomorphism on A.

Remark 4.39. Suppose one only knows that $\dim_B(\phi(A)) < \frac{m}{2}$ for a typical ϕ. It is difficult to draw any conclusions in this case. Sauer and Yorke [58] exhibit a compact subset A of $\mathbb{R}^{10}$ with $\dim_B(A) = 3.5$ such that $\dim_B(\phi(A)) < 3$ for every $\phi \in C^1(\mathbb{R}^{10}, \mathbb{R}^6)$.

4.5 Observing A Continuous Dynamical System

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a dynamical system and let A be a compact invariant set. We make no a priori regularity assumptions about f. Let $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ and let $B \subset \mathbb{R}^n$ be an open metric ball. Recall that if there exists a map $\bar{f} : \phi(A) \to \phi(A)$ such that for $x \in A$ the diagram

$$\begin{array}{ccc} A & \xrightarrow{f} & A \\ \phi \big\downarrow & & \big\downarrow \phi \\ \phi(A) & \xrightarrow{\bar{f}} & \phi(A) \end{array}$$

commutes, then we say that $\bar{f}$ is the **induced map** associated with f.

Remark 4.40. If f is continuous, then the existence of $\bar{f}$ implies the continuity of $\bar{f}$.

Definition 4.41. The pair $(x_1, x_2) \in A \times A$ is **coincident** if $\phi(x_1) = \phi(x_2)$. The pair $(x_1, x_2) \in A \times A$ is said to be **dynamically separated** by B if

1. (x_1, x_2) is coincident and

2. $x_1 \notin B$, $x_2 \notin B$, $f(x_1) \in B$ and $f(x_2) \notin B$.

Definition 4.42. Let S_B be the set of maps ϕ in $C^1(\mathbb{R}^n, \mathbb{R}^m)$ for which the following hold:

1. There exists some pair (x_1, x_2) dynamically separated by B, and

2. for all such pairs we have $\phi(f(x_1)) = \phi(f(x_2))$.

Lemma 4.43. *The set S_B is a shy subset of $C^1(\mathbb{R}^n, \mathbb{R}^m)$.*

Proof. We construct a measure transverse to S_B. Let $\beta : \mathbb{R}^n \to \mathbb{R}$ be a C^∞ map such that $\beta > 0$ on B and $\text{supp}(\beta) = \bar{B}$. Let $v \in \mathbb{R}^m$ be a nonzero vector. Let μ be the Lebesgue measure supported on the one dimensional subspace

$$\{tv\beta : t \in \mathbb{R}\} .$$

For any $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ it is evident that $\phi + tv\beta \in S_B$ for at most one $t \in \mathbb{R}$. Thus S_B is shy because μ is transverse to it. $\square$

Definition 4.44. Let $\text{Fix}(f)$ denote the set of fixed points of f. Let $\text{Per}_2(f)$ denote the set of periodic points of f of period 2.

Proposition 4.45. *Suppose $f[A]$ is continuous and invertible. Assume that the sets $\text{Fix}(f[A])$ and $\text{Per}_2(f[A])$ are countable. For almost every map $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ the following are equivalent:*

(1) *The map ϕ is one to one on A.*

(2) *The induced map $\bar{f}$ exists (and is therefore continuous).*

Proof.

((1) $\Rightarrow$ (2)) Define $\bar{f} := \phi \circ f \circ \phi^{-1}$.

((2) $\Rightarrow$ (1)) Let $\{B_i\}$ be a countable collection of open metric balls such that if $x,\, y \in A$ satisfy $x \neq y$ then there exists some B_i such that $x \in B_i$ and $y \notin B_i$. Consider the following three sets:

$$\begin{cases} G_1 = \{\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m) : \phi \text{ is one to one on } \operatorname{Fix}(f[A])\} \\ G_2 = \{\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m) : \phi \text{ is one to one on } \operatorname{Per}_2(f[A])\} \\ G_3 = \bigcap_{i=1}^{\infty} (S_{B_i})^C \end{cases}$$

The set G_1 is a prevalent subset of $C^1(\mathbb{R}^n, \mathbb{R}^m)$ by Proposition 4.19 because the fixed points of $f[A]$ are countable. Similarly, G_2 is prevalent. The set G_3 is a prevalent subset of $C^1(\mathbb{R}^n, \mathbb{R}^m)$ because $(S_{B_i})^C$ is prevalent for each i by (4.43) and because the countable intersection of prevalent sets is prevalent (see [33]). Thus $G_1 \cap G_2 \cap G_3$ is a prevalent subset of $C^1(\mathbb{R}^n, \mathbb{R}^m)$. Let $\phi \in G_1 \cap G_2 \cap G_3$ and assume that ϕ is not one to one on A. It follows that no induced map $\bar{f}$ exists. Since $\phi \notin S_{B_i}$ for all i, there exists a metric ball B_i and a coincident pair (x_1, x_2) dynamically separated by B_i such that $\phi(f(x_1)) \neq \phi(f(x_2))$. $\square$

Proposition 4.21 allows us to improve this result by transferring the dynamical hypotheses onto the induced dynamics. We need a lemma indicating that the existence of a point of discontinuity of $f[A]$ precludes the existence of a continuous induced map for a typical measurement function.

Lemma 4.46. *Suppose $f[A]$ is discontinuous at some point $x \in A$. Then for a.e. $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, no continuous induced map exists.*

Theorem 4.47. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a map. For almost every map $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, there is an induced map $\bar{f}$ satisfying*

1. $\bar{f}$ is continuous and invertible, and

2. $\mathrm{Fix}(\bar{f})$ *and* $\mathrm{Per}_2(\bar{f})$ *are countable*

if and only if the following hold.

1. The measurement map ϕ is one to one on A.

2. The sets $\mathrm{Fix}(f[A])$ *and* $\mathrm{Per}_2(f[A])$ *are countable.*

3. The map $f[A]$ is continuous and invertible.

Proof. We employ the transference method. If $f[A]$ is continuous and invertible and $\mathrm{Fix}(f[A])$ and $\mathrm{Per}_2(f[A])$ are countable sets, then (4.45) implies the result. If $\mathrm{Fix}(f[A])$ or $\mathrm{Per}_2(f[A])$ is uncountable then Proposition 4.21 implies that the statement of the theorem holds for almost every ϕ. Lemma 4.46 implies the result if $f[A]$ is discontinuous at some point. If $f[A]$ is not invertible, then for almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ no invertible induced map exists. $\square$

We now consider the possibility of recovering differential information.

4.6 Observing Differentiable Dynamics

Assume that f is a diffeomorphism on $\mathbb{R}^n$. The concept of a measurement function ϕ being an immersion on A usually requires A to be a manifold, but there is now an obvious extension.

Definition 4.48. We say the map $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ is an **immersion** on A if $D\phi(x)[T_xA] : T_xA \to T_{\phi(x)}\phi(A)$ is one to one for each $x \in A$.

Motivated by the theory of infinite-dimensional dynamical systems, we formulate our C^1 results using the notion of quasidifferentiability.

Definition 4.49. The function f is said to be quasidifferentiable on the set A if $f[A]$ is continuous and if for each $x \in A$ there exists a linear map $Df(x) : \mathbb{R}^n \to \mathbb{R}^n$, the quasiderivative of f at x, such that

$$\frac{f(x_i) - f(y_i) - Df(x)(x_i - y_i)}{\|x_i - y_i\|} \to 0$$

for all sequences $(x_i) \subset A$ and $(y_i) \subset A$ such that $x_i \to x$ and $y_i \to x$.

Remark 4.50. The function f is Whitney C^1 if and only if f is quasidifferentiable and the quasiderivative varies continuously. Since continuity is observable, the C^1 embedding results to follow may be formulated with 'Whitney C^1' in place of 'quasidifferentiable.'

We would like to prove under the assumptions of (4.45) that for almost every ϕ, the existence of a quasidifferentiable induced map $\bar{f}$ implies that ϕ is an injective immersion on A. However, one extra hypothesis on f is needed; namely, that for each $x \in \mathrm{Fix}(f[A])$ we have

$$Df(x)[T_x A] \neq \gamma \cdot I \text{ for every } \gamma \in \mathbb{R}.$$

To see the need for this hypothesis, suppose that f is the identity map, A is countable, and there exists $x \in A$ such that $\dim(T_x A) = n > m$. In this case, the identity map on $\phi(A)$ is the induced map for every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, yet every ϕ fails to be immersive at x.

Consider a countable set $\{B_i = B(y_i, r_i)$ of open metric balls in $\mathbb{R}^n$ that separates points. Let $T(A) = \{(x, v) : x \in A,\, v \in T_x A\}$.

Definition 4.51. Let W_{B_i} be the set of measurement maps in $C^1(\mathbb{R}^n, \mathbb{R}^m)$ with the following properties:

1. There exists some point $(x, v) \in T(A)$ such that $v \neq 0$, $x \notin B(y_i, 2r_i)$, $f(x) \in B(y_i, r_i)$, $D\phi(x)v = 0$, and

2. for all such points we have $D\phi(f(x)) \circ Df(x)v = 0$.

Lemma 4.52. *The set W_{B_i} is shy.*

Proof. Let $F_1, \ldots, F_t$ be a basis for the nm dimensional space of linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$. Let $\beta : \mathbb{R}^n \to \mathbb{R}$ be a C^∞ map with the following properties:

$$\begin{cases} (1) \quad \beta(x) = 1 \text{ for } x \in B(y_i, \frac{5}{4}r_i) \\ (2) \quad \operatorname{supp}(\beta) = \overline{B(y_i, \frac{3}{2}r_i)} \\ (3) \quad 0 < \beta \leq 1 \text{ on } B(y_i, \frac{3}{2}r_i) \end{cases}$$

Let P be the subspace of $C^1(\mathbb{R}^n, \mathbb{R}^m)$ spanned by the collection $\{\beta F_i : i = 1, \ldots, t\}$ and endow P with Lebesgue measure. For any ϕ, the set of vectors (α_i) for which

$$\phi + \beta \sum_{i=1}^{t} \alpha_i F_i \in W_{B_i}$$

is a subset of P of measure zero. □

Lemma 4.53. *Let $x \in \operatorname{Fix}(f[A])$ and assume that $Df(x)[T_xA] \neq \gamma \cdot I$ for all $\gamma \in \mathbb{R}$. The set Z_x of measurement mappings satisfying*

1. $\ker(D\phi(x)) \cap T_xA \neq \{0\}$ *and*

2. $Df(x)(\ker(D\phi(x)) \cap T_xA) \subset \ker(D\phi(x))$

is a shy subset of $C^1(\mathbb{R}^n, \mathbb{R}^m)$.

Proof. Consider the orthogonal decomposition $\mathbb{R}^n = T_xA \oplus E_x$. Let L be the subset of $\operatorname{Lin}(\mathbb{R}^n, \mathbb{R}^m)$ consisting of maps that vanish on E_x and have norm at most one. Endow L with the normalized Lebesgue probability measure μ. For any $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, we claim that

$$\mu(\{F \in L : \phi + F \in Z_x\}) = 0. \tag{4.3}$$

If $\dim T_x A \leq m$ then (4.3) follows from the fact that almost every linear transformation has full rank. If $\dim T_x A > m$, then it suffices to consider the scalar case $m = 1$. Let $d = \dim(T_x A)$ and let $\{\phi_{e_i}\}$ be an orthonormal basis for $\mathrm{Lin}(T_x A, \mathbb{R})$, the unit ball of which we identify with L. Let ϕ_w represent $D\phi(x)[T_x A]$ with respect to the basis $\{\phi_{e_i}\}$. For a map $\phi_v \in \mathrm{Lin}(T_x A, \mathbb{R})$ such that $v + w \neq 0$, it is necessary that $v + w$ be an eigenvector of $Df(x)[T_x A]^T$ in order to have

$$Df(x)(\ker(\phi_{v+w}) \cap T_x A) \subset \ker(\phi_{v+w}).$$

If $Df(x)[T_x A]^T$ does not have an eigenvalue of multiplicity d, then (4.3) holds. Finally, notice that $Df(x)[T_x A]^T$ has an eigenvalue of multiplicity d if and only if $Df(x)[T_x A]$ is a scalar multiple of the identity. □

Proposition 4.54. *Suppose f is a diffeomorphism on $\mathbb{R}^n$. Assume that*

$$\mathrm{Fix}(f[A]) \textit{ and } \mathrm{Per}_2(f[A])$$

are countable sets. Assume that for each $x \in \mathrm{Fix}(f[A])$ we have

$$Df(x)[T_x A] \neq \gamma \cdot I \textit{ for every } \gamma \in \mathbb{R}.$$

Then for almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, if there is a quasidifferentiable induced map $\bar{f}$ then the measurement map ϕ is an injective immersion on A.

Proof. Consider the following sets:

$$\begin{cases} G_4 = \displaystyle\bigcap_{i=1}^{\infty} (W_{B_i})^C \\ G_5 = \displaystyle\bigcap_{x \in \mathrm{Fix}(f[A])} (Z_x)^C \end{cases}$$

The sets G_4 and G_5 are prevalent by (4.52) and (4.53) respectively. For ϕ in the prevalent set

$$\bigcap_{j=1}^{5} G_j,$$

the existence of a quasidifferentiable induced map $\bar{f}$ implies that ϕ is an injective immersion on A. $\square$

Once again Proposition 4.21 allows us to transfer some of the hypotheses of this theorem onto the induced dynamics.

Theorem 4.55. *Suppose f is a diffeomorphism on $\mathbb{R}^n$. For almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, if there is a quasidifferentiable induced map satisfying*

1. $\mathrm{Fix}(\bar{f})$ *and* $\mathrm{Per}_2(\bar{f})$ *are countable and*

2. *For each $y \in \mathrm{Fix}(\bar{f})$, $D\bar{f}(y)[T_y\phi(A)] \neq \gamma \cdot I$ for every $\gamma \in \mathbb{R}$*

then the following hold.

1. *The measurement map ϕ is an injective immersion on A.*

2. $\mathrm{Fix}(f[A])$ *and* $\mathrm{Per}_2(f[A])$ *are countable.*

3. *For each $x \in \mathrm{Fix}(f[A])$, $Df(x)[T_xA] \neq \gamma \cdot I$ for every $\gamma \in \mathbb{R}$.*

Proof. It suffices to consider the cases in which the hypotheses of Proposition 4.54 fail to hold. If $\mathrm{Fix}(f[A]) \cup \mathrm{Per}_2(f[A])$ is uncountable, then for almost every ϕ there cannot exist an induced map satisfying $\mathrm{Fix}(\bar{f})$ and $\mathrm{Per}_2(\bar{f})$ are countable by Proposition 4.21. Suppose there exist $x \in \mathrm{Fix}(f[A])$ and $\gamma \in \mathbb{R}$ such that

$$Df(x)[T_xA] = \gamma \cdot I.$$

For almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, $D\phi(x)[T_xA]$ has full rank. If $\dim(T_xA) \geqslant m$ then the full rank of $D\phi(x)[T_xA]$ implies that $D\phi(x)$ maps T_xA onto $T_{\phi(x)}\phi(A)$ and therefore the existence of a quasidifferentiable induced map would imply

$$D\bar{f}(\phi(x))[T_{\phi(x)}\phi(A)] = \gamma \cdot I.$$

If $\dim(T_xA) < m$ then the full rank of $D\phi(x)[T_xA]$ implies that $D\phi(x)$ maps T_xA injectively into $T_{\phi(x)}\phi(A)$ and therefore surjectively onto $T_{\phi(x)}\phi(A)$. In this case, the existence of a quasidifferentiable induced map would imply

$$D\bar{f}(\phi(x))[T_{\phi(x)}\phi(A)] = \gamma \cdot I.$$

□

Using the manifold extension theorem we strengthen this theorem by utilizing the previously introduced notion of a diffeomorphism on A. We recall that definition here.

Definition 4.56. We say that a measurement map $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ is a **diffeomorphism** on A if ϕ is injective on A and if for each $x \in A$ there exists an enveloping manifold M for A at x that is mapped diffeomorphically onto an enveloping manifold for $\phi(A)$ at $\phi(x)$.

Theorem 4.57. *Suppose f is a diffeomorphism on $\mathbb{R}^n$. For almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, if there is a quasidifferentiable induced map $\bar{f}$ satisfying*

1. $\mathrm{Fix}(\bar{f})$ *and* $\mathrm{Per}_2(\bar{f})$ *are countable and*

2. *For each* $y \in \mathrm{Fix}(\bar{f})$, $D\bar{f}(y)[T_y\phi(A)] \neq \gamma \cdot I$ *for every* $\gamma \in \mathbb{R}$

then the following hold.

1. *The measurement map ϕ is a diffeomorphism on A.*

2. $\mathrm{Fix}(f[A])$ *and* $\mathrm{Per}_2(f[A])$ *are countable.*

3. *For each $x \in \mathrm{Fix}(f[A])$, $Df(x)[T_xA] \neq \gamma \cdot I$ for every $\gamma \in \mathbb{R}$.*

Remark 4.58. Mera and Morán [48] provide a test for determining whether or not observed trajectories of $\bar{f}$ are consistent with the assumption that $\bar{f}$ belongs to a certain regularity class.

The C^1 Theorem (4.57) is not Platonic because we assume that f is a diffeomorphism on $\mathbb{R}^n$. We formulate a Platonic version of the C^1 Theorem by selecting new hypotheses on the induced map $\bar{f}$. The key modification is the replacement of the dynamical assumption on the nature of $D\bar{f}(y)[T_y\phi(A)]$ for $y \in \mathrm{Fix}(\bar{f})$ with the structural assumption that $\dim T_y(\phi(A)) < m\ \forall y \in \phi(A)$. The smoothness of f becomes an observable in this new setting. After presenting several technical preliminaries, we state and prove the main result. We assume only that f is a map throughout this section.

Lemma 4.59. *If $\dim T_x(A) \geqslant m$ for some $x \in A$, then for almost every $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ one has $\dim T_{\phi(x)}\phi(A) \geqslant m$.*

Proof. The result follows from the fact that almost every linear transformation from one finite-dimensional vector space to another has full rank. □

Lemma 4.60. *Suppose there exist sequences $(x_i) \subset A$, $(y_i) \subset A$, and $x \in A$ such that $x_i \to x$, $y_i \to x$ and $\frac{x_i - y_i}{\|x_i - y_i\|} \to v \in T_xA$, but*

$$\left(\frac{f(x_i) - f(y_i)}{\|x_i - y_i\|}\right)$$

does not converge to a vector in $\mathbb{R}^n$*. For almost every* $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$*, there does not exist a quasidifferentiable induced map* $\bar{f}$ *on* $\phi(A)$ *with* $\dim T_y\phi(A) < m \; \forall y \in \phi(A)$*.*

Proof. We need to consider two cases. Assume that the sequence

$$\left(\frac{f(x_i) - f(y_i)}{\|x_i - y_i\|}\right) \tag{4.4}$$

has two limit points, v_1 and v_2. There cannot exist a quasidifferentiable induced map $\bar{f}$ on $\phi(A)$ if $v \notin \ker(D\phi(x)[T_xA])$ and $v_1 - v_2 \notin \ker(D\phi(f(x))[T_{f(x)}A])$. This condition is prevalent and therefore the lemma holds in the first case. Now suppose that the sequence (4.4) tends to infinity. If either $\dim(T_xA) \geqslant m$ or $\dim(T_{f(x)}A) \geqslant m$, then Lemma 4.59 implies that for almost every ϕ one does not have $\dim T_y\phi(A) < m \; \forall y \in \phi(A)$. If both $\dim(T_xA) < m$ and $\dim(T_{f(x)}A) < m$, then for almost every ϕ it follows that $D\phi(x)[T_xA]$ and $D\phi(f(x))[T_{f(x)}A]$ are injective. For such a ϕ, the existence of a quasidifferentiable induced map $\bar{f}$ on $\phi(A)$ would imply

$$\frac{\bar{f} \circ \phi(x_i) - \bar{f} \circ \phi(y_i)}{\|\phi(x_i) - \phi(y_i)\|} = \frac{\phi \circ f(x_i) - \phi \circ f(y_i)}{\|\phi(x_i) - \phi(y_i)\|} \to \infty,$$

a contradiction. ∎

Theorem 4.61 (Platonic C^1 Theorem). *Suppose* $f : \mathbb{R}^n \to \mathbb{R}^n$ *is a map. For almost every* $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m)$*, if there exists an invertible quasidifferentiable induced map* $\bar{f}$ *on* $\phi(A)$ *satisfying*

1. $\mathrm{Fix}(\bar{f})$ *and* $\mathrm{Per}_2(\bar{f})$ *are countable,*
2. $\dim T_y(\phi(A)) < m \; \forall y \in \phi(A)$*, and*
3. $D\bar{f}(y)[T_y\phi(A)]$ *is invertible* $\forall y \in \phi(A)$*,*

then the following hold.

1. *The measurement mapping ϕ is a diffeomorphism on A.*

2. *The mapping $f[A]$ is invertible.*

3. *The sets $\mathrm{Fix}(f[A])$ and $\mathrm{Per}_2(f[A])$ are countable.*

4. *The dynamical system f is quasidifferentiable on A and $Df(x)[T_xA]$ is invertible for all $x \in A$.*

5. *For each $x \in A$, $\dim(T_xA) < m$.*

Proof. See Sections 4.5 and 4.6 for the definitions of the sets G_1, G_2, G_3, and G_4. Let

$$G_6 = \{\phi \in C^1(\mathbb{R}^n, \mathbb{R}^m) : D\phi(x)[T_xA] \text{ is injective for each } x \in \mathrm{Fix}(f[A])\}.$$

If $\mathrm{Fix}(f[A])$ is countable and $\dim(T_xA) < m$ for each $x \in A$, then G_6 is prevalent. We employ the transference method to prove the Platonic C^1 Theorem.

If f satisfies conclusions (2), (3), (4), and (5), then for ϕ in the prevalent set

$$\left(\bigcap_{j=1}^{4} G_j\right) \bigcap G_6,$$

the existence of a quasidifferentiable induced map $\bar{f}$ on $\phi(A)$ implies that ϕ is an injective immersion on A. If $f[A]$ is not invertible, then for almost every ϕ, no invertible induced map exists. If $\mathrm{Fix}(f[A]) \cup \mathrm{Per}_2(f[A])$ is uncountable, then Proposition 4.21 implies that no induced map satisfying hypothesis (1) exists for almost every ϕ. If there exists $x \in A$ for which $\dim(T_xA) \geqslant m$, then Lemma 4.59 implies that $\dim T_{\phi(x)}\phi(A) \geqslant m$ for almost every ϕ and for such ϕ hypothesis (2) is not satisfied.

Suppose f is not quasidifferentiable on A. If $f[A]$ is not continuous, then Lemma 4.46 implies that for almost every ϕ there does not exist a quasidifferentiable induced map $\bar{f}$ on $\phi(A)$. If f fails to be quasidifferentiable on A because the hypotheses of Lemma 4.60 are satisfied, then this lemma implies that for a.e. ϕ there does not exist a quasidifferentiable induced map $\bar{f}$ on $\phi(A)$ with $\dim T_y\phi(A) < m \; \forall y \in \phi(A)$. The remaining possibility is that for some $x \in A$ there exists a nonlinear map taking T_xA into $T_{f(x)}A$. For a.e. ϕ, this precludes the existence of a quasidifferentiable induced map $\bar{f}$. Finally, suppose f is quasidifferentiable on A but $Df(x)[T_xA]$ is not invertible for some $x \in A$. In this case for a.e. ϕ there does not exist a quasidifferentiable induced map $\bar{f}$ on $\phi(A)$ satisfying hypothesis (3). □

We finish with theorems concerning delay coordinate mappings and Lyapunov exponents.

4.7 Delay Coordinate Mappings and Lyapunov Exponents

We state delay coordinate embedding versions of our results and prove the exponent characterization theorem.

4.7.1 Delay Coordinate Maps

The following theorems do not follow from the previously established corresponding theorems for the general class of smooth measurement mappings because the delay coordinate mappings form a subspace of $C^1(\mathbb{R}^n, \mathbb{R}^m)$. Nevertheless, their veracity is established using essentially the same reasoning.

Theorem 4.62. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a map. For almost every $g \in C^1(\mathbb{R}^n, \mathbb{R})$, there is an induced map $\bar{f}$ satisfying*

1. *$\bar{f}$ is continuous and invertible, and*

2. *$\bigcup_{i=1}^{2m} \mathrm{Per}_i(\bar{f})$ is countable*

if and only if the following hold.

1. *The delay coordinate map $\phi(f, g)$ is one to one on A.*

2. *The set $\bigcup_{i=1}^{2m} \mathrm{Per}_i(f[A])$ is countable.*

3. *The map $f[A]$ is continuous and invertible.*

Theorem 4.63. *Let f be a diffeomorphism on $\mathbb{R}^n$. For a.e. $g \in C^1(\mathbb{R}^n, \mathbb{R})$, if there is a quasidifferentiable induced map $\bar{f}$ satisfying*

1. *$\bigcup_{i=1}^{2m} \mathrm{Per}_i(\bar{f})$ is countable and*

2. *for each $p \in \{1, \dots, m\}$ and $y \in \mathrm{Per}_p(\bar{f})$ we have*

$$D\bar{f}^p(y)[T_y\phi(f, g)(A)] \neq \gamma \cdot I \text{ for every } \gamma \in \mathbb{R}$$

then the following hold.

1. *The delay coordinate map $\phi(f, g)$ is a diffeomorphism on A.*

2. *The set $\bigcup_{i=1}^{2m} \mathrm{Per}_i(f[A])$ is countable.*

3. *For each $p \in \{1, \dots, m\}$ and each $x \in \mathrm{Per}_i(f[A])$, we have*

$$Df^p(x)[T_xA] \neq \gamma \cdot I \text{ for every } \gamma \in \mathbb{R}.$$

4.7.2 Lyapunov Exponents

We conclude Section 4.7 with a discussion of Lyapunov exponents. Assume f and $\bar{f}$ are quasidifferentiable and invertible on A and $\phi(A)$, respectively, with invertible quasiderivatives at each point $x \in A$ and $y \in \phi(A)$. Suppose ϕ is a diffeomorphism on A. Assume $y \in \phi(A)$ is a regular point for $\bar{f}$ and recall that this implies the existence of a decomposition

$$\mathbb{R}^m = \bigoplus_{i=1}^{l} E_i(y)$$

such that

$$\lim_{k \to \pm\infty} \frac{1}{k} \log \|D\bar{f}^k(y)v\| = \lambda_j(y) \quad (v \in E_j(y) \setminus \{0\} \text{ and } 1 \leq j \leq l).$$

Since the set of regular points $R(\bar{f})$ is invariant in the sense that

1. $y \in R(\bar{f}) \Rightarrow \bar{f}^k(y) \in R(\bar{f})$ for all $k \in \mathbb{Z}$ and

2. $D\bar{f}^{\pm 1}(E_i(y)) = E_i(\bar{f}^{\pm 1}(y))$ for $i = 1, \ldots, l$,

we associate the Lyapunov exponents $\lambda_1 > \cdots > \lambda_l$ with the trajectory (y_k). Counting multiplicities, there are m Lyapunov exponents associated with (y_k) and we label them $\chi_1, \ldots, \chi_m$ such that

$$\chi_1 \geqslant \chi_2 \geqslant \cdots \geqslant \chi_m.$$

In light of Remark 4.33 following the manifold extension theorem, we make the following definitions.

Definition 4.64. We say that a Lyapunov exponent $\lambda(y, v)$ of $\bar{f}$ is a **tangent** Lyapunov exponent if $v \in T_y\phi(A)$. A Lyapunov exponent $\lambda(y, v)$ of $\bar{f}$ is said to be a **transverse** Lyapunov exponent if it is not a tangent exponent.

Definition 4.65. A Lyapunov exponent $\lambda(y, v)$ of $\bar{f}$ is said to be a **true** Lyapunov exponent if it does not depend on the choice of quasiderivative $D\bar{f}$ and if it is also a Lyapunov exponent of f at $\phi^{-1}(y)$. We say that a Lyapunov exponent $\lambda(y, v)$ of $\bar{f}$ is **spurious** if there exists a quasiderivative $D\bar{f}$ for which

$$\lim_{k\to\pm\infty} \frac{1}{k} \log \|D\bar{f}^k(y)v\|$$

either does not exist or is not a Lyapunov exponent of f at $\phi^{-1}(y)$.

Theorem 4.66 (Exponent Characterization Theorem). *Assume f and $\bar{f}$ are quasidifferentiable and invertible on A and $\phi(A)$, respectively, with invertible quasiderivatives at each point $x \in A$ and $y \in \phi(A)$. Suppose ϕ is a diffeomorphism on A. Assume that $y \in \phi(A)$ is a regular point for $\bar{f}$ such that* $\dim T_z\phi(A) = \dim T_y\phi(A)$ *for all $z \in \overline{(y_k)}$. The following characterizations hold for a Lyapunov exponent $\lambda(y, v)$ of $\bar{f}$.*

1. *If the exponent $\lambda(y, v)$ is tangent then it is a true exponent.*

2. *If the exponent $\lambda(y, v)$ is transverse then it is a spurious exponent.*

The tangent exponents of $\bar{f}$ correspond to the tangent exponents of f.

Remark 4.67. The tangent space $T_y\phi(A)$ admits the decomposition

$$T_y\phi(A) = \bigoplus_{i=1}^{l} V_i(y)$$

where $V_i(y)$ is a subspace of $E_i(y)$ for $i = 1, \ldots, l$.

Remark 4.68. From a computational point of view, one is interested in constructing algorithms to efficiently and accurately compute the Lyapunov spectrum and identify the true exponents. The existing technique ([18, 57, 49]) requires that one modify the Eckmann and Ruelle algorithm by computing the tangent maps only on the tangent spaces and not on the ambient space $\mathbb{R}^m$. Assuming A is a smooth submanifold, Mera and Morán [49] state conditions under which this modified ERA converges. Clearly this technique eliminates the computation of spurious exponents. However, one has to compute the tangent spaces along the entire orbit. In light of the exponent characterization theorem, we propose a new algorithm that eliminates the need to compute these tangent spaces.

Definition 4.69. A **forward filtration** of $\mathbb{R}^m$ is a nested collection of subspaces

$$\emptyset = F_0(y) \subset F_1(y) \subset F_2(y) \subset \cdots \subset F_m(y) = \mathbb{R}^m$$

such that

$$\lim_{k \to +\infty} \frac{1}{k} \log \|D\bar{f}^k(y)v\| = \chi_{m-j+1}$$

for $v \in F_j(y) \setminus F_{j-1}(y)$.

Definition 4.70. A **backward filtration** of $\mathbb{R}^m$ is a nested collection of subspaces

$$\emptyset = B_0(y) \subset B_1(y) \subset B_2(y) \subset \cdots \subset B_m(y) = \mathbb{R}^m$$

such that

$$\lim_{k \to -\infty} \frac{1}{k} \log \|D\bar{f}^k(y)v\| = \chi_j$$

for $v \in B_j(y) \setminus B_{j-1}(y)$.

Suppose that forward and backward filtrations have been computed. Assume that one may determine computationally if a given $(m-1)$-dimensional subspace of $\mathbb{R}^m$ contains $T_y\phi(A)$. For $j = 1, \ldots, m$, compute the Lyapunov vector

$$v_j \in B_j \cap F_{m-j+1}.$$

We now fix j and determine if $v_j \in T_y\phi(A)$. If $\text{Span}\{v_i : i \neq j\} \supset T_y\phi(A)$ then $v_j \notin T_y\phi(A)$. If $\text{Span}\{v_i : i \neq j\} \not\supset T_y\phi(A)$ then $v_j \in T_y\phi(A)$ and χ_j is a true Lyapunov exponent. The true Lyapunov exponents and $T_y\phi(A)$ have been determined. It would be interesting to compare the performance of this algorithm to that of existing ERA techniques.

Proof. Statement (1) follows from the fact that ϕ is a diffeomorphism on A. We establish (2) with a perturbation argument. Let $\alpha > 1$ and let $d = \dim T_y\phi(A)$. For each $z \in \overline{(y_k)}$ there exists an enveloping manifold M_z for $\phi(A)$ at z with $T_zM_z = T_z\phi(A)$ and $\dim(M_z) = d$. Let

$$\{B(z, r_z) : z \in \overline{(y_k)}\}$$

be a collection of metric balls such that

$$B(z, r_z) \cap \phi(A) \subset \text{Int}(M_z).$$

By compactness there exists a finite subcover

$$\{B(z_i, \frac{r_{z_i}}{2}) : i = 1, \ldots, N\}$$

of $\overline{(y_k)}$. We inductively construct a sequence $\{D\bar{f}_k : k = 1, \ldots, N\}$ of perturbations of $D\bar{f}$. Let $\beta : \mathbb{R}^m \to \mathbb{R}$ be a C^∞ map such that

$$\begin{cases} (1) \quad 1 \leq \beta \leq \alpha, \\ (2) \quad \beta(z) = \alpha \text{ for } z \in B(z_1, \frac{r_{z_1}}{2}), \text{ and} \\ (3) \quad \beta(z) = 1 \text{ on } \mathbb{R}^m \setminus B(z_1, r_{z_1}). \end{cases}$$

For each $z \in B(z_1, r_{z_1}) \cap M_{z_1}$, $\mathbb{R}^m$ admits the orthogonal decomposition

$$\mathbb{R}^m = T_z(M_{z_1}) \oplus E_z.$$

Using this decomposition we define $D\bar{f}_1$ as follows.

1. $D\bar{f}_1[\phi(A) \cap \mathbb{R}^m \setminus B(z_1, r_{z_1})] = D\bar{f}[\phi(A) \cap \mathbb{R}^m \setminus B(z_1, r_{z_1})]$

2. For $z \in \phi(A) \cap B(z_1, r_{z_1})$, define $D\bar{f}_1(z)$ by

$$D\bar{f}_1(z)v = \begin{cases} D\bar{f}(z)v, & \text{if } v \in T_z(M_{z_1}); \\ \beta(z)D\bar{f}(z)v, & \text{if } v \in E_z. \end{cases}$$

In this fashion we inductively construct the family of perturbations $\{D\bar{f}_k : k = 1, \ldots, N\}$. For $v \in (T_y\phi(A))^\perp$ we have

$$\varliminf_{k\to\infty} \frac{1}{k} \log \|D\bar{f}_N^k(y)v\| \geq \lambda(y, v) + \log(\alpha).$$

Since $\alpha > 1$ was arbitrary, it follows that if $\lambda(y, v)$ is transverse then it is spurious.

$\square$

Acknowledgements. The authors would like to thank Dr. Michael Brin for his numerous insightful comments on a draft. We thank Jaroslav Stark for directing us to the Whitney extension theorem and David Broomhead for a discussion concerning the nature of tangent spaces associated with arbitrary sets.

BIBLIOGRAPHY

[1] V.I. Arnold, *Geometrical Methods in the Theory of Ordinary Differential Equations*, Springer-Verlag, 1983.

[2] A.V. Babin and M.I. Vishik, *Attractors of Partial Differential Equations and Estimates of their Dimensions*, Russian Math. Survey **38** (1983), 151-213.

[3] A.V. Babin and M.I. Vishik, *Attractors of Evolution Equations*, North-Holland, 1992.

[4] C. Beck, *Upper and Lower Bounds on the Renyi Dimensions and the Uniformity of Multifractals*, Physica D **41** (1990), 67-78.

[5] A. Ben-Artzi, A. Eden, C. Foias and B. Nicolaenko, *Hölder Continuity for the Inverse of Mañé's Projection*, J. Math. Anal. Appl. **178** (1993), no. 1, 22-29.

[6] R. Camassa, D. Holm and C.D. Levermore, *Long-Time Effects of Bottom Topography in Shallow Water*, Physica D **98** (1996), 258-286.

[7] R. Camassa, D. Holm and C.D. Levermore, *Long-Time Shallow-Water Equations with a Varying Bottom*, J. Fluid Mech. **349** (1997), 173-189.

[8] L. Campbell and W. Garnett, *The Life of James Clerk Maxwell*, Macmillan, London, 1882; reprinted by Johnson Reprint, New York, 1969.

[9] V. Chepyzhov and M. Vishik, *Attractors for Equations of Mathematical Physics*, American Mathematical Society, 2002.

[10] P. Constantin, *Collective L^∞ Estimates for Families of Functions with Orthonormal Derivatives*, Indiana University Math. J. **36** (1987), 603-615.

[11] P. Constantin and C. Foias, *Navier-Stokes Equations*, The University of Chicago Press, 1988.

[12] P. Constantin, C. Foias and R. Temam, *Attractors Representing Turbulent Flows*, Mem. Amer. Math. Soc. **53** (1985).

[13] P. Constantin, C. Foias and R. Temam, *On the Dimension of the Attractors in Two-Dimensional Turbulence*, Physica D **30** (1988), 284-296.

[14] C. Doering and J. D. Gibbon, *Note on the Constantin-Foias-Temam Attractor Dimension Estimate for Two-Dimensional Turbulence*, Physica D **48** (1991), 471-480.

[15] C. Doering and J. D. Gibbon, *Applied Analysis of the Navier-Stokes Equations*, Cambridge University Press, 1995.

[16] C. Doering and X. Wang, *Attractor Dimension Estimates for Two-Dimensional Shear Flows*, Physica D **123** (1998), 206-222.

[17] H. Donnelly, *Bounds for Eigenfunctions of the Laplacian on Compact Riemannian Manifolds*, Journal of Functional Analysis **187** (2001), 247-261.

[18] J.-P. Eckmann, S.O. Kamphorst, D. Ruelle, and S. Ciliberto, *Lyapunov Exponents from Time Series*, Phys. Rev. A **34** (1986), 4971-4979.

[19] J.-P. Eckmann and D. Ruelle, *Ergodic Theory of Chaos and Strange Attractors*, Rev. Mod. Phys. **57** (1985), 617-656.

[20] A. Eden, C. Foias, B. Nicolaenko and R. Temam, *Exponential Attractors for Dissipative Evolution Equations*, John Wiley & Sons, 1994.

[21] L. Evans and R. Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press, 1992.

[22] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, Wiley, New York, 1990.

[23] C. Foias and E. Olson, *Finite Fractal Dimension and Hölder-Lipschitz Parametrization*, Indiana Univ. Math. J. **45** (1996), no. 3, 603-616.

[24] P. Friz and J. Robinson, *Smooth Attractors have Zero "Thickness,"* J. Math. Anal. Appl. **240** (1999), no. 1, 37-46.

[25] R. J. Gardner, *The Brunn-Minkowski Inequality*, Bull. Amer. Math. Soc. **39** (2002), 355-405.

[26] J.-M. Ghidaglia, M. Marion and R. Temam, *Generalization of the Sobolev-Lieb-Thirring Inequalities and Applications to the Dimension of Attractors*, Differential and Integral Equations **1** (1988), no. 1, 1-21.

[27] P. Grassberger, *Generalized Dimension of Strange Attractors*, Phys. Lett. **97A** (1983), 227-231.

[28] H.G. Hentschel and I. Procaccia, *The Infinite Number of Generalized Dimensions of Fractals and Strange Attractors*, Physica D **8** (1983), 435-444.

[29] M. Hirsch, *Differential Topology*, Springer-Verlag, 1976.

[30] B. Hunt and V. Kaloshin, *How Projections Affect the Dimension Spectrum of Fractal Measures*, Nonlinearity **10** (1997), 1031-1046.

[31] B. Hunt and V. Kaloshin, *Regularity of Embeddings of Infinite-Dimensional Fractal Sets into Finite-Dimensional Spaces*, Nonlinearity **12** (1999), 1263-1275.

[32] B. Hunt, J. Kennedy, T.Y. Li, and H. Nusse, *SLYRB Measures: Natural Invariant Measures for Chaotic Systems*, Phys. D **170** (2002), no. 1, 50-71.

[33] B. Hunt, T. Sauer and J. Yorke, *Prevalence: A Translation-Invariant "Almost Every" on Infinite-Dimensional Spaces*, Bulletin of the AMS **27** (1992), 217-238.

[34] B. Hunt, T. Sauer and J. Yorke, *Prevalence: An Addendum*, Bulletin of the AMS **28** (1993), no. 2, 306-307.

[35] B. R. Hunt and J. A. Yorke, *Maxwell on Chaos*, Nonlinear Science Today **3** (1993), 2-4.

[36] V. Yu. Kaloshin, *Some Prevalent Properties of Smooth Dynamical Systems*, Proc. Steklov Inst. Math. **213** (1996), no. 2, 115-140.

[37] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, 1995.

[38] R. Kaufmann, *On Hausdorff Dimension of Projections*, Mathematatika **15** (1968), 153-155.

[39] S.G. Krantz and H.R. Parks, *The Geometry of Domains in Space*, Birkhäuser, 1999.

[40] C.D. Levermore and M. Sammartino, *A Shallow Water Model with Eddy Viscosity for Basins with Varying Bottom Topography*, Nonlinearity **14** (2001), 1493-1515.

[41] E. Lieb and W. Thirring, *Inequalities for the Moments of the Eigenvalues of the Schrödinger Hamiltonian and their Relation to Sobolev Inequalities*, in Studies in Mathematical Physics, E. Lieb, B. Simon and A. Wightman Eds., Princeton University Press, Princeton, 1976.

[42] V. Liu, *A Sharp Lower Bound on the Dimension of the Global Attractor of the Two-Dimensional Navier-Stokes Equations*, Comm. Math. Phys. **158** (1993), 327-339.

[43] M. Lyubich, *On the Typical Behavior of the Trajectories of the Exponential*, Russian Math. Surveys **41** (1986), no. 2, 207-208.

[44] J.M. Marstrand, *Some Fundamental Geometrical Properties of Plane Sets of Fractional Dimension*, Proc. London Math. Soc. **4** (1954), 257-302.

[45] P. Mattila, *Hausdorff Dimension, Orthogonal Projections and Intersections with Planes*, Ann. Acad. Sci. Fennicae A **1** (1975), 227-244.

[46] P. Mattila, *Integralgeometric Properties of Capacities*, Trans. Am. Math. Soc. **266** (1981), 539-554.

[47] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces*, Cambridge University Press, Cambridge, 1995.

[48] M. Mera and Manuel Morán, *Convergence of the Eckmann and Ruelle Algorithm for the Estimation of Lyapunov Exponents*, Ergodic Theory and Dynamical Systems **20** (2000), 531-546.

[49] M. Mera and Manuel Morán, *Degrees of Freedom of a Time Series*, Journal of Statistical Physics **106** (2002), no. 1-2, 125-145.

[50] M. Misiurewicz, *On Iterates of e^z*, Ergodic Theory Dynamical Systems **1** (1981), 103-106.

[51] J.C. Oxtoby, *Measure and Category*, Springer, New York, 1971.

[52] J. Palis, Jr. and W. de Melo, *Geometric Theory of Dynamical Systems*, Springer-Verlag, 1982.

[53] M. Rees, *The Exponential Map is Not Recurrent*, Math. Z. **191** (1986), 593-598.

[54] A. Rényi, *Probability Theory*, North-Holland, Amsterdam, 1970.

[55] R.H. Riedi, *An Improved Multifractal Formalism and Self-Affine Measures*, J. Math. Anal. Appl. **189** (1995), 462-490.

[56] J. Robinson, *Infinite-Dimensional Dynamical Systems*, Cambridge University Press, 2001.

[57] M. Sano and Y. Sawada, *Measurement of the Lyapunov Spectrum from a Chaotic Time Series*, Phys. Rev. Lett. **55** (1985), 1082-1085.

[58] T. Sauer and J. Yorke, *Are the Dimensions of a Set and its Image Equal Under Typical Smooth Functions?*, Ergodic Theory and Dynamical Systems **17** (1997), no. 4, 941-956.

[59] T. Sauer, J.A. Yorke, and M. Casdagli, *Embedology*, Journal of Statistical Physics **65** (1991), no. 3-4, 579-616.

[60] J. Schwinger, *On the Bound States of a Given Potential*, Proc. Nat. Acad. Sci. **47** (1961), 122-129.

[61] G. Sell and Y. You, *Dynamics of Evolutionary Equations*, Springer, 2002.

[62] E. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, 1983.

[63] F. Takens, *Detecting Strange Attractors in Turbulence*, In: Dynamical Systems and Turbulence, Lecture Notes in Mathematics **898**, Springer-Verlag, Berlin, 1981.

[64] R. Temam, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Springer, 1997.

[65] M. Viana, *A Proof of Oseledets' Theorem*, `http://w3.impa.br/~viana/`.

[66] H. Whitney, *Analytic Extensions of Differentiable Functions Defined in Closed Sets*, Trans. Amer. Math. Soc. **36** (1934), 63-89.

[67] L.-S. Young, *What are SRB measures, and which dynamical systems have them? Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays.*, Journal of Statistical Physics **108** (2002), no. 5-6, 733-754.

[68] M. Ziane, *Optimal Bounds on the Dimension of the Attractor of the Navier-Stokes Equations*, Physica D **105** (1997), 1-19.