Nonlinearity **12** (1999) 1263–1275. Printed in the UK PII: S0951-7715(99)97363-2

Regularity of embeddings of infinite-dimensional fractal sets into finite-dimensional spaces

Brian R Hunt† and Vadim Yu Kaloshin‡

† Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA
‡ Department of Mathematics, Princeton University, Princeton, NJ 08544-1000, USA

E-mail: bhunt@ipst.umd.edu and kaloshin@math.princeton.edu

Received 9 September 1998, in final form 4 June 1999
Recommended by Professor P Cvitanovic

Abstract. We consider the image of a fractal set X in a Banach space under typical linear and nonlinear projections π into $\mathbb{R}^N$. We prove that when N exceeds twice the box-counting dimension of X, then almost every (in the sense of prevalence) such π is one-to-one on X, and we give an explicit bound on the Hölder exponent of the inverse of the restriction of π to X. The same quantity also bounds the factor by which the Hausdorff dimension of X can decrease under these projections. Such a bound is motivated by our discovery that the Hausdorff dimension of X need not be preserved by typical projections, in contrast to classical results on the preservation of a Hausdorff dimension by projections between finite-dimensional spaces. We give an example for any positive number d of a set X with box-counting and Hausdorff dimension d in the real Hilbert space ℓ^2 such that for *all* projections π into $\mathbb{R}^N$, no matter how large N is, the Hausdorff dimension of $\pi(X)$ is less than d (and in fact, is less than two, no matter how large d is).

AMS classification scheme numbers: 28A20, 58F11, 31A15, 94A17, 49Q15, 60B05

1. Introduction

Many infinite-dimensional dynamical systems have been shown to have compact finite-dimensional attractors [Mal, Man1, FT, BV, L, CF, CFT1, NST1, Te1, CFT2, NST2, H, Te2, EFNT]. When such an attractor is measured experimentally, what is actually observed is, colloquially speaking, a 'projection' or 'embedding' of the attractor into a Euclidean space. (However, there is no reason to expect that the map from the state space of the system to Euclidean space is linear, nor *a priori* that it gives a one-to-one representation of the attractor.) In this scenario, a fundamental question is how accurately the image of the attractor reflects the attractor itself.

Past results on the projection of a fractal set X in (finite-dimensional) Euclidean space ensure that if the dimension of the image space is sufficiently large, almost every projection: (a) is one-to-one on X [W, Man2, SYC]; and (b) preserves the Hausdorff dimension of X [Mar, Kau, Mat1]. Further, Mañé showed that (a) holds for a topologically generic set of projections to a space of sufficiently high dimension when X is a finite-dimensional compact subset of a Banach space [Man2]. In this paper we show that Mañé's result holds for almost every projection in the measure-theoretic sense of prevalence [HSY], but that for some X, (b) does not hold. We also estimate the Hölder exponent of the inverse of a typical projection, and

0951-7715/99/051263+13$30.00

as a result obtain a lower bound on the Hausdorff dimension of the image of X under typical projections.

Let B be a Banach space, and let X be a compact subset of B with finite box-counting dimension $\dim_B(X)$. The box-counting dimension (also known as the capacity, entropic, fractal, or Minkowski dimension) is defined as follows. Let $n(X, \varepsilon)$ be the minimum number of ε-balls needed to cover X, and let

$$\dim_B(X) = \limsup_{\varepsilon\to 0} \frac{\log n(X, \varepsilon)}{\log(1/\varepsilon)}.$$

The box-counting dimension of X is an upper bound on its Hausdorff dimension $\dim_H(X)$, and in many cases the two dimensions coincide (see for example [Fa, Mat2]).

In section 2 we present an example for all $d > 0$ of a compact subset X of a real Hilbert space H with $\dim_B(X) = \dim_H(X) = d$ such that for all positive integers N and all bounded linear functions $\pi : H \to \mathbb{R}^N$, the Hausdorff dimension of $\pi(X)$ is at most $d/(1 + d/2)$. In particular, $\dim_H(\pi(X))$ is less than d, and further it is less than two no matter how large d is. As our main result below implies, the drop in dimension is not generally due to π failing to be one-to-one on X, but rather is due to $\pi^{-1}|_{\pi(X)}$, the inverse of π on $\pi(X)$, failing to be Lipschitz. In other words, distances between nearby points on X are shrunk by π, and the closer the points, the greater the shrinking factor.

To better understand how projections may distort a compact set X, despite being one-to-one on X, we consider how much regularity can be proved for the inverse of such projections. The existence of a dense set of projections with Hölder continuous inverse was proved in the case that X belongs to $\mathbb{R}^M$ in [BEFN] and for $X \subset H$ in [FO]; in the finite-dimensional case an explicit bound on the Hölder exponent is given. Though these results are stated for orthogonal projections, they also hold for the general class of linear projections considered in this paper, since every linear projection π into $\mathbb{R}^N$ with full rank can be expressed (in a unique way) as the composition of an orthogonal projection of rank N (onto the subspace orthogonal to the null space of π) and an invertible linear map (from the aforementioned subspace to $\mathbb{R}^N$) [AGV]. In section 3 we extend these results to almost every linear projection, where in the infinite-dimensional case we mean 'almost every' in the sense of prevalence [HSY]; see section 1.1 for a definition of prevalence. We also give a bound on the Hölder exponent in the infinite-dimensional case.

Our main result is as follows.

Theorem 1.1. *Let $X \subset B$ be a compact set with box-counting dimension d. Let $N > 2d$ be an integer, and let*

$$\alpha_0 = \frac{N - 2d}{N(1 + d)}.$$

Then, for all $\alpha \in (0, \alpha_0)$ and for almost every (in the sense of prevalence) bounded linear function $\pi : B \to \mathbb{R}^N$ there exists $C > 0$ such that for all $x, y \in X$,

$$C|\pi(x) - \pi(y)|^\alpha \geqslant |x - y|.$$

In particular, almost every π is one-to-one on X and $\pi^{-1}|_{\pi(X)}$ is Hölder continuous with exponent α.

This result is a special case of theorem 3.9, which is stated and proved in section 3. In theorem 3.6 we prove the same result in the case that B is a real Hilbert space, but with the improved exponent

$$\alpha_0 = \frac{N - 2d}{N(1 + d/2)}.$$

In [BEFN], the bound α_0 on the Hölder exponent of the inverse (for the case of $X \subset \mathbb{R}^M$) is simply $(N-2d)/N$, which approaches 1 as $N \to \infty$. In the case of a real Hilbert space, the exponent α_0 that we obtain approaches $1/(1+d/2)$ as $N \to \infty$, and this is precisely the factor by which the dimension drops in the example of section 2. In fact, we have the following corollary.

Corollary 1.2. *Under the hypotheses of theorem 1.1, for almost every* π

$$\frac{N-2d}{N(1+d)} \dim_B(X) \leqslant \dim_B(\pi(X)) \leqslant \dim_B(X)$$

and

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

In a real Hilbert space, the factor $(1+d)$ *can be replaced by* $(1+d/2)$ *in both cases.*

Corollary 1.2 applies to many other notions of dimension as well, including measure-dependent notions such as information and correlation dimension. See [HK] for a discussion of how these various dimensions transform under projections between finite-dimensional spaces. The corollary is simply a consequence of the fact that Hölder continuity of $\pi^{-1}|_{\pi(X)}$ with exponent α implies that the inverse image of the intersection of a ball of radius ε with $\pi(X)$ is contained in a ball of radius $C\varepsilon^\alpha$ for some constant $C > 0$. Thus, for example, in the case of box-counting dimension, $n(X, C\varepsilon^\alpha) \leqslant n(\pi(X), \varepsilon)$.

Let us discuss why we consider both box-counting and the Hausdorff dimension in this paper. Neither theorem 1.1 nor its finite-dimensional analogues [Man2, SYC] are valid if d represents the Hausdorff dimension of X. The appendix by Kan in [SYC] describes an example of a set $X \subset \mathbb{R}^M$ with Hausdorff dimension zero such that every linear function $\pi : \mathbb{R}^M \to \mathbb{R}^N$ with $N < M$ fails to be one-to-one on X. On the other hand, in contrast with the Hausdorff dimension there are finite-dimensional examples in which the box-counting dimension of X is not preserved under any such π [FH, J, SY].

The lower bound on the box-counting dimension for typical finite-dimensional projections given in [FH] suggests that corollary 1.2 is not sharp, and undoubtedly the lower bound can be further improved for the Hausdorff dimension and other notions of dimension that are preserved under typical projections in finite-dimensional spaces (as described in [HK]). Indeed, if the example in section 2 represents the 'worst case', then the factor $(N-2d)/N$ could be eliminated from the Hausdorff dimension bound. Nonetheless, the example in section 2 shows that the bound in corollary 1.2 is sharp in the limit $N \to \infty$, and thus so is the bound in theorem 1.1.

Despite this asymptotic sharpness of the bounds in theorem 1.1 and corollary 1.2, we can improve them substantially for some X by replacing the box-counting dimension d in the denominator by another non-negative quantity that depends on X. In section 3 we introduce the 'thickness exponent' τ of X, and show (in theorem 3.9) that the factor $(1+d)$ in both the results above can be replaced by $(1+\tau)$. The thickness exponent measures how well X can be approximated by finite-dimensional subspaces, with smaller values of τ indicating better approximability. In general, τ is bounded above by d, and the example in section 2 shows that equality is possible. However, we expect that τ can be shown to be significantly smaller than d for many attractors of infinite-dimensional systems. Indeed, Friz and Robinson [FR] 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 exponent is at most m/s, and point out that this implies that certain attractors of the Navier–Stokes equations have thickness exponent zero. See also, chapter 4 of [EFNT].

Finally, in section 4 we discuss results for almost every delay-coordinate function.

1.1. Prevalence

We prove theorem 1.1 not only for almost every bounded linear function $\pi : H \to \mathbb{R}^N$, but also for almost every C^1 function and almost every Lipschitz function. In all of these cases, we mean 'almost every' in the sense of prevalence [HSY]. Prevalence extends the notion of 'Lebesgue almost every' from Euclidean spaces to infinite-dimensional spaces. In this respect it differs from the commonly used topological notion of a *generic* or *residual* set, one that contains a countable intersection of open dense sets—in Euclidean spaces, a residual set may have Lebesgue measure zero.

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, using the Fubini theorem. Let us foliate $\mathbb{R}^n$ by k-dimensional planes, which we think of as translations of $\mathbb{R}^k \subset \mathbb{R}^n$ by elements of $\mathbb{R}^{n-k}$. If the 'Lebesgue almost every' translation of $\mathbb{R}^k$ intersects $S \subset \mathbb{R}^n$ in full (k-dimensional) Lebesgue measure, then S has full (n-dimensional) Lebesgue measure. If we replace $\mathbb{R}^n$ by an infinite-dimensional space B, we cannot formulate the same condition (because the space of translations is infinite dimensional), but we can impose the stronger condition that every translation of $\mathbb{R}^k$ intersects S in full Lebesgue measure.

Here is the precise definition of prevalence in a Banach space B.

Definition 1.3. *We say that a Borel subset $S \subset B$ is prevalent if there exists a compactly supported probability measure μ such that $\mu(S + x) = 1$ for all $x \in B$. A non-Borel set that contains a prevalent Borel set is also prevalent.*

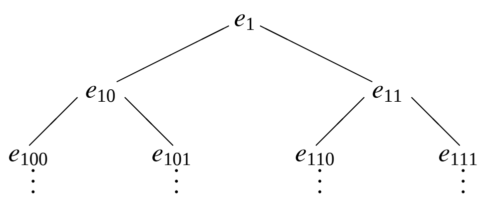
Equivalently, we can allow μ to be any measure that is positive and finite on some compact subset of B; in that case we require that every translate $S + x$ of S have full measure. In particular, μ can be the Lebesgue measure on any finite-dimensional subspace of B, as in the previous paragraph. Allowing more general measures in our definitions ensures that a countable intersection of prevalent sets is prevalent; see [HSY] for details. To prove the 'prevalent' results of this paper, we construct a compactly supported probability measure μ such that for all π_0 in the given function space, $\pi_0 + \pi$ has the desired properties for almost every π with respect to μ.

2. Nonpreservation of Hausdorff dimension

In this section we present an example of a compact set X of Hausdorff and box-counting dimension d in the real separable Hilbert space $H = \ell^2$ such that for all N and all bounded linear functions $\pi : H \to \mathbb{R}^N$,

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

Let $\{e_i\}_{i=1}^{\infty}$ be an orthonormal basis of H. Consider the following tree based on writing the positive integers i in base 2:



Let $\lambda = 2^{-1/d}$, and consider the countable collection of sets $\{X_k\}_{k=1}^{\infty}$ defined as follows. First, X_1 contains one point e_1. Next, X_2 contains two points: $e_1 + \lambda e_{10}$ and $e_1 + \lambda e_{11}$; these two points correspond to two different branches of length one in the tree. Next, X_3 contains four points: $e_1 + \lambda e_{10} + \lambda^2 e_{100}$, $e_1 + \lambda e_{10} + \lambda^2 e_{101}$, $e_1 + \lambda e_{11} + \lambda^2 e_{110}$, $e_1 + \lambda e_{11} + \lambda^2 e_{111}$; these four points correspond to branches of length two.

More formally, let S consist of all infinite strings $\theta = i_1 i_2 i_3 \ldots$ of binary digits with $i_1 = 1$, and let θ_n denote the truncated string $i_1 i_2 \ldots i_n$. Then

$$X_n = \{e_{\theta_1} + \lambda e_{\theta_2} + \cdots + \lambda^{n-1} e_{\theta_n} : \theta \in S\}.$$

Notice that X_n contains 2^{n-1} points, corresponding to a branch of length $n-1$ in the tree.

Let X be the set of all limit points of $\cup_i X_i$, or in other words

$$X = \{e_{\theta_1} + \lambda e_{\theta_2} + \lambda^2 e_{\theta_3} + \cdots : \theta \in S\}.$$

Proposition 2.1. *Both* $\dim_B(X)$ *and* $\dim_H(X)$ *are equal to* $d = \log 2/\log(1/\lambda)$.

Proof. First, $\dim_H(X) \leqslant \dim_B(X) \leqslant d$ because X can be covered by 2^{n-1} balls centred at the points of X_n with radius $\lambda^n/(1-\lambda)$. To show that $\dim_H(X) \geqslant d$, 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 θ_n and θ_n' of length n must lie at least λ^{n-1} apart, the measure of a ball of diameter λ^{n-1}, is at most the measure of all strings in S starting with a given θ_n, that is, $2^{1-n} = (\lambda^{n-1})^d$. By Frostman's lemma [Fr, Fa, Mat2], $\dim_H(X) \geqslant d$. □

Proposition 2.2. *For all bounded linear functions* $\pi : H \to \mathbb{R}^N$*, the Hausdorff dimension of* $\pi(X)$ *is at most* $d/(1 + d/2)$.

Proof. Each coordinate of π is a bounded linear functional from H to $\mathbb{R}$, and hence can be expressed as the inner product with a vector in H. In other words, if $x = x_1 e_1 + x_{10} e_{10} + x_{11} e_{11} + \cdots \in H$, then $\pi(x) = a_1 x_1 + a_{10} x_{10} + \cdots$ where $a_1, a_{10}, \ldots \in \mathbb{R}^N$ and $M = |a_1|^2 + |a_{10}|^2 + \cdots < \infty$.

Intuitively speaking, there are 2^{n-1} different strings θ_n of length n, so most of the $|a_{\theta_n}|$ should be bounded above by a constant times $2^{-n/2}$. Thus, under π, distances of order λ^n contract to distances of order $2^{-n/2}\lambda^n = \lambda^{n(d/2+1)}$ at most, and the dimension of X drops by at least a factor of $d/2 + 1$.

To be more precise, let $x = e_{\theta_1} + \lambda e_{\theta_2} + \cdots$ be a point in X, and let $y = e_{\theta_1} + \lambda e_{\theta_2} + \cdots + \lambda^{n-1} e_{\theta_n} \in X_n$. Then

$$|\pi(x) - \pi(y)| = |\lambda^n a_{\theta_{n+1}} + \lambda^{n+1} a_{\theta_{n+2}} + \cdots| \leqslant \frac{\lambda^n}{1-\lambda} \max_{j>n} |a_{\theta_j}|.$$

Thus, for all $x \in X$ corresponding to the initial string θ_n of length n, $\pi(x)$ is contained in a ball centred at $\pi(y)$ with radius $(\lambda^n/(1-\lambda))|a_{\hat{\theta}_n}|$ where $|a_{\hat{\theta}_n}|$ is the maximum value of $|a_{\theta_j}|$ over all finite strings θ_j of length more than n starting with θ_n.

It follows that $\pi(X)$ is contained in the union of 2^{n-1} balls, the sum of whose radii to the power $d/(1 + d/2)$ is

$$\frac{1}{(1-\lambda)^{d/(1+d/2)}} \sum_{\theta_n} \lambda^{nd/(1+d/2)} |a_{\hat{\theta}_n}|^{d/(1+d/2)}.$$

The sum can be bounded by the Hölder inequality with exponents $1 + d/2$ and $1 + 2/d$:

$$\begin{aligned}\sum_{\theta_n} \lambda^{nd/(1+d/2)} |a_{\hat{\theta}_n}|^{d/(1+d/2)} &\leqslant \left(\sum_{\theta_n} \lambda^{nd}\right)^{1/(1+d/2)} \left(\sum_{\theta_n} |a_{\hat{\theta}_n}|^2\right)^{1/(1+2/d)} \\ &\leqslant (2^{n-1}2^{-n})^{1/(1+d/2)} M^{1/(1+2/d)}.\end{aligned}$$

The key above is that the index of $a_{\hat{\theta}_n}$ is different for each θ_n.

We have constructed a cover of $\pi(X)$ by balls of radius at most $\sqrt{M}\lambda^n/(1-\lambda)$ such that the sum of the radii to the $d/(1+d/2)$ power is bounded above independently of n. Therefore, the Hausdorff dimension of $\pi(X)$ is at most $d/(1+d/2)$. □

Remark 2.3. *The set X can also be constructed in ℓ^p for $1 \leqslant p < \infty$. In that case, the sum of the qth powers of the coefficients of π converges, where $q = p/(p-1)$, and the Hausdorff dimension of $\pi(X)$ is shown in the same way as above to be at most $d/(1+d/q)$.*

3. Proof of the main results

3.1. The finite-dimensional case

We begin by proving a finite-dimensional version of theorem 1.1. This result extends a similar result for a dense set of projections in [BEFN]. It also serves to illustrate some of the key ideas in the proof of the infinite-dimensional version.

Theorem 3.1. *Let $X \subset \mathbb{R}^M$ be a compact set with box-counting dimension d. Let $N > 2d$ be an integer, and let α be a real number with*

$$0 < \alpha < \frac{N-2d}{N}.$$

Let $f : \mathbb{R}^M \to \mathbb{R}^N$ be a Lipschitz function. Then for the 'Lebesgue almost every' linear function $\pi : \mathbb{R}^M \to \mathbb{R}^N$, there exists $C > 0$ such that for all $x, y \in X$,

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

Remark 3.2. *With $f = 0$, theorem 3.1 is a result about almost every linear function. With f allowed to vary over a space of nonlinear functions, such as the C^1 functions or Lipschitz functions, theorem 3.1 is a result about almost every function from that space, in the sense of prevalence.*

Proof. We consider each linear function $\pi : \mathbb{R}^M \to \mathbb{R}^N$ as an $N \times M$ matrix $\{a_1, a_2, \ldots, a_M\}$ of (column) vectors $a_m \in \mathbb{R}^N$, where a_m represents the image of the mth basis vector of $\mathbb{R}^M$ under π. Let Q be the 'unit cube' consisting of all $\{a_m\}$ with $|a_m| \leqslant 1$ for each m.

Choose $\varepsilon \in (0, 1)$. Let Z_ε be the set of points in $(x, y) \in X \times X$ with $|x - y| \geqslant \varepsilon^{\alpha}$. Let Q_ε be the set of $\pi \in Q$ for which $|f(x) + \pi(x) - (f(y) + \pi(y))| \leqslant \varepsilon$ for some $(x, y) \in Z_\varepsilon$. Let L be a Lipschitz constant that is valid for all $f + \pi$ with $\pi \in Q$.

Lemma 3.3. *For all $\delta > d$, the Lebesgue measure μ of Q_ε is bounded as*

$$\mu(Q_\varepsilon) \leqslant C_0 \varepsilon^{N(1-\alpha)-2\delta}$$

where C_0 depends only on L, M, N, X, δ.

Proof. Since $\dim_B(X) = d$, there exists $C_1 > 0$ depending only on X and δ such that X can be covered by at most $C_1\varepsilon^{-\delta}$ balls of radius $\varepsilon/2$ for all $\varepsilon \in (0, 1)$. It follows that $Z_\varepsilon \subset X \times X$ can be covered by at most $C_1^2\varepsilon^{-2\delta}$ balls of radius ε. Let Y be the intersection of Z_ε with one of these balls.

Choose points $(x, y), (x_0, y_0) \in Y$. Since $|(x, y) - (x_0, y_0)| \leqslant 2\varepsilon$, if $|f(x_0) + \pi(x_0) - (f(y_0) + \pi(y_0))| \geqslant (4L + 1)\varepsilon$ then $|f(x) + \pi(x) - (f(y) + \pi(y))| \geqslant \varepsilon$ for all $(x, y) \in Y$.

Since $|x_0 - y_0| \geqslant \varepsilon^\alpha$, for some $1 \leqslant m \leqslant M$ the mth coordinate of $x_0 - y_0$ has an absolute value of at least $M^{-1/2}\varepsilon^\alpha$. Then for fixed $a_1, \ldots, a_{m-1}, a_{m+1}, \ldots, a_M$, the Lebesgue measure of values of a_m for which $|f(x_0) + \pi(x_0) - (f(y_0) + \pi(y_0))| \leqslant (4L+1)\varepsilon$ is bounded above by

$$\rho = C_2 \left(\frac{(4L+1)\varepsilon}{M^{-1/2}\varepsilon^\alpha} \right)^N = C_2(4L+1)^N M^{N/2} \varepsilon^{N(1-\alpha)},$$

where C_2 is the volume of the unit ball in $\mathbb{R}^N$. Thus, by the Fubini theorem, the Lebesgue measure of $\{a_1, \ldots, a_M\} \in Q$ for which $|f(x_0) + \pi(x_0) - (f(y_0) + \pi(y_0))| \leqslant (4L+1)\varepsilon$ is bounded above by $C_2^{M-1}\rho$.

We have shown that the Lebesgue measure of $\pi \in Q$ for which $|f(x) + \pi(x) - (f(y) + \pi(y))| \leqslant \varepsilon$ for some $(x, y) \in Y$ is at most $C_2^{M-1}\rho$. Since Z_ε is covered by at most $C_1^2 \varepsilon^{-2\delta}$ such sets Y,

$$\mu(Q_\varepsilon) \leqslant C_1^2 \varepsilon^{-2\delta} C_2^{M-1} \rho = C_1^2 C_2^M (4L+1)^N M^{N/2} \varepsilon^{N(1-\alpha)-2\delta},$$

which completes the proof of the lemma. □

Since $\alpha < (N - 2d)/N$, we can choose $\delta > d$ such that the exponent of ε in the bound on $\mu(Q_\varepsilon)$ is positive. It follows that the sum of $\mu(Q_{2^{-j}})$ for $j = 1, 2, 3, \ldots$ is finite. Then, by the Borel–Cantelli lemma, 'Lebesgue almost every' $\pi \in Q$ belongs to only finitely many $Q_{2^{-j}}$. For such π, the conclusion of the theorem follows. Finally, choose $\pi_1, \pi_2, \pi_3, \ldots$ such that every linear function from $\mathbb{R}^M$ to $\mathbb{R}^N$ can be written as $\pi_i + \pi$ for some $i \geqslant 1$ and $\pi \in Q$. Applying what we have just proved to each $f + \pi_i$ completes the proof of the theorem. □

3.2. The thickness exponent

Now we consider, instead of $\mathbb{R}^M$, an infinite-dimensional Banach space B.

Definition 3.4. *The thickness exponent $\tau(X)$ of a compact set $X \subset B$ is defined as follows. Let $d(X, \varepsilon)$ 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 $d(X, \varepsilon) = \infty$. Let*

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

The lemma below shows that the thickness exponent of X is bounded above by its box-counting dimension; however, there is no corresponding lower bound, nor any general relationship between the thickness exponent and the Hausdorff dimension. For example, a finite-dimensional disk has thickness exponent zero but can have arbitrarily high dimension. And 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.

Lemma 3.5. *Let $X \subset B$ be a compact set with box-counting dimension d. Then $\tau(X) \leqslant d$.*

Proof. Recall that the box-counting dimension is defined similarly to $\tau(X)$ but in terms of the minimum number $n(X, \varepsilon)$ of ε-balls required to cover X. For any such cover, X lies within ε of the space spanned by the centres of the balls. Thus $d(X, \varepsilon) \leqslant n(X, \varepsilon)$, and the desired inequality immediately follows. □

In fact, one can show that if X belongs to a finite-dimensional C^r manifold, then $\tau(X) \leqslant d/r$. Roughly speaking, one covers X with $n(X, \varepsilon)$ balls of radius ε, then within each ball one approximates the manifold containing X to within ε^r by a Taylor polynomial.

One then bounds $d(X, \varepsilon^r)$ by a constant times $n(X, \varepsilon)$, where the constant is the number of terms in the Taylor polynomial and depends only on r and the dimension of the manifold. Also, at the end of the introduction we discussed a result from [FR] that shows, in a different sense, that 'thickness is inversely proportional to smoothness'.

3.3. The infinite-dimensional case

First we present our main result in the case of a real Hilbert space H.

Theorem 3.6. *Let $X \subset H$ be a compact set with box-counting dimension d and thickness exponent τ. Let $N > 2d$ be an integer, and let α be a real number with*

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

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

$$C|f(x) - f(y)|^\alpha \geqslant |x - y|.$$

Proof. In the proof of theorem 3.1, we showed that the desired inequality holds for almost every perturbation of a Lipschitz function f by a linear function π in the 'unit cube' Q-represented by matrices with columns belonging to the unit ball of $\mathbb{R}^N$. There is no such cube in the space of the bounded linear functions $\pi : H \to \mathbb{R}^N$, so instead we use a 'Hilbert brick' Q-defined as follows.

For $j = 1, 2, 3, \ldots$, let $d_j = d(X, 2^{-j\alpha}/3)$ and let $V_j \subset H$ be a subspace of dimension d_j such that every point of X lies within $2^{-j\alpha}/3$ of V_j. From the definition of τ it follows that for all $\sigma > \tau$, there exists $C_3 > 0$, depending only on X and σ, such that $d_j \leqslant C_3 2^{j\alpha\sigma}$. Let S_j be the closed unit ball in the dual space of V_j, which we embedded into the dual space of H in the natural way—an element of the dual space of V_j acts on an element of H by composition with the orthogonal projection from H onto V_j.

Every bounded linear function $\pi : H \to \mathbb{R}^N$ can be written $\pi = (\pi_1, \pi_2, \ldots, \pi_N)$ where each π_n belongs to the dual space of H, and $\pi(x) = (\pi_1(x), \ldots, \pi_N(x))$ for all $x \in H$. Let

$$Q = \left\{ (\pi_1, \ldots, \pi_N) : \text{ for all } n, \pi_n = \sum_{j=1}^{\infty} j^{-2} \phi_{nj} \text{ with } \phi_{nj} \in S_j \text{ for all } j \right\}.$$

Notice that Q is compact. Let μ be the probability measure on Q that results from choosing each ϕ_{nj} randomly (and independently) with respect to the uniform probability measure on S_j.

Now choose $\varepsilon = 2^{-j}$ for some $j \geqslant 1$. Refer to the second paragraph of the proof of theorem 3.1 for the definitions of Z_ε, Q_ε, and L. In terms of μ we formulate the analogue of lemma 3.3.

Lemma 3.7. *For all $\delta > d$,*

$$\mu(Q_{2^{-j}}) \leqslant C_0 j^{2N} 2^{-j[N(1-\alpha(1+\sigma/2))-2\delta]}$$

where C_0 depends only on L, N, X, δ, σ.

Proof. Refer to the first two paragraphs of the proof of lemma 3.3 for the definitions of C_1, Y, x_0, and y_0. We wish to bound the probability ρ that $\pi \in Q$ chosen randomly with respect to μ satisfies

$$|f(x_0) + \pi(x_0) - (f(y_0) + \pi(y_0))| = |f(x_0) - f(y_0) + \pi(x_0 - y_0)| \leqslant (4L + 1)\varepsilon.$$

Now ρ is bounded above by the maximum probability over all $(4L+1)\varepsilon$-balls that $\pi(x_0 - y_0)$ lies in the given ball, which is in turn bounded above by the the product over $n = 1, 2, \ldots, N$ of the maximum probability ρ_n over all intervals of length $(8L+2)\varepsilon$ that the nth coordinate of $\pi(x_0 - y_0)$ lies in the given interval. By symmetry, ρ_n is independent of n.

Recall that $x_0, y_0 \in X$, and that every point of X lies within $2^{-j\alpha}/3 = \varepsilon^\alpha/3$ of V_j. Let z be the orthogonal projection of $x_0 - y_0$ onto V_j; then since $|x_0 - y_0| \geqslant \varepsilon^\alpha$, it follows that $|z| \geqslant \varepsilon^\alpha/3$. Notice that $\phi_{nj}(x_0 - y_0) = \phi_{nj}(z)$. Then by the Fubini theorem, ρ_n is bounded above by the maximum probability that $j^{-2}\phi_{nj}(z)$ lies in a given $(8L+2)\varepsilon$-interval. Recall that ϕ_{nj} is distributed uniformly in the d_j-dimensional unit ball S_j.

Because we are working in a Hilbert space, S_j (together with its action on z) is equivalent to a Euclidean ball. Hence, ρ_n is bounded above by the maximum probability that ϕ_{nj} lies between two given parallel planes of dimension $d_j - 1$, a distance $(8L+2)\varepsilon j^2 |z|^{-1}$ apart. Let v_k denote the volume of the k-dimensional unit ball; then

$$\rho_n \leqslant \frac{v_{d_j-1}(8L+2)\varepsilon j^2 |z|^{-1}}{v_{d_j}} \leqslant C_4 d_j^{1/2}(8L+2)3j^2\varepsilon^{1-\alpha},$$

where C_4 is an absolute constant (representing the maximum of $v_{k-1}/(k^{1/2}v_k)$ over all k).

Next, since $\rho \leqslant \rho_n^N$, $\varepsilon = 2^{-j}$, and $d_j \leqslant C_3 2^{j\alpha\sigma}$,

$$\rho \leqslant C_4^N C_3^{N/2} 2^{j\alpha\sigma N/2}(8L+2)^N 3^N j^{2N} 2^{-j(1-\alpha)N} = C_5 j^{2N} 2^{-jN(1-\alpha(1+\sigma/2))}$$

where C_5 depends only on L, N, X, σ.

Then by the same reasoning as in the proof of lemma 3.3,

$$\mu(Q_{2^{-j}}) \leqslant C_1^2 \varepsilon^{-2\delta} \rho \leqslant C_1^2 C_5 j^{2N} 2^{-j[N(1-\alpha(1+\sigma/2))-2\delta]},$$

which completes the proof of the lemma. □

Next, choose δ sufficiently close to d and σ sufficiently close to τ that

$$\alpha < \frac{N - 2\delta}{N(1+\sigma/2)}.$$

Then the exponent of 2^{-j} in the bound on $\mu(Q_{2^{-j}})$ is positive, and as in the proof of theorem 3.1 it follows from the Borel–Cantelli lemma that the conclusion of the theorem holds for $f + \pi$ for almost every $\pi \in Q$ with respect to μ. Therefore the conclusion of the theorem holds for a prevalent set of f in any space of functions from H to $\mathbb{R}^N$ that is contained in the Lipschitz functions and contains the bounded linear functions. □

As with theorem 1.1, we have the following corollary.

Corollary 3.8. *Under the hypotheses of theorem 3.6, for almost every f*

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

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

Thus, the maximum factor by which the dimension of X drops under typical projections into $\mathbb{R}^N$ is close to one if N is large and τ is small.

Finally, we consider the general case of a Banach space B. Lemma 3.5 and the following theorem imply theorem 1.1.

Theorem 3.9. *Let $X \subset B$ be a compact set with box-counting dimension d and thickness exponent τ. Let $N > 2d$ be an integer, and let α be a real number with*

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

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

$$C|f(x) - f(y)|^\alpha \geqslant |x - y|.$$

Proof. We proceed exactly as in the proof of theorem 3.6, except that we use lemma 3.10 below in place of lemma 3.7, replacing $1+\sigma/2$ with $1+\sigma$, and that we replace the denominator three in the definition of d_j and V_j with five (this is necessary for the proof of the lemma). There is no natural embedding of the dual space of V_j into the dual space of B, but it follows from the Hahn–Banach theorem that there exists such an embedding that is isometric. The uniform probability measure on S_j is defined as follows. Choose a basis for S_j and, by coordinate representation, identify S_j with a convex set $U_j \subset \mathbb{R}^{d_j}$. The uniform probability measure on U_j induces a probability measure on S_j (which is independent of the choice of basis, though we do not use this fact). □

Lemma 3.10. *For all $\delta > d$,*

$$\mu(Q_{2^{-j}}) \leqslant C_0 j^{2N} 2^{-j[N(1-\alpha(1+\sigma))-2\delta]}$$

where C_0 depends only on L, N, X, δ, σ.

Proof. We proceed as in the proof of lemma 3.7, except that we must define z differently and make a different estimate of ρ_n. First of all, since three was replaced by five above, we now know that x_0 and y_0 each lie within $\varepsilon^\alpha/5$ of some points $x, y \in V_j$. Let $z = x - y$; then $|z| \geqslant 3\varepsilon^\alpha/5$. By the Hahn–Banach theorem, there exists $\psi \in S_j$ such that $\psi(z) = |z|$ and $|\psi| = 1$. It follows that $\psi(x_0 - y_0) \geqslant |z| - |x_0 - x| - |y - y_0| \geqslant \varepsilon^\alpha/5$.

We bound ρ_n by the maximum probability that $j^{-2}\phi_{nj}(x_0 - y_0)$ lies in a given interval of length $(8L+2)\varepsilon$. Within the dual space of B, let P be the subspace that annihilates $x_0 - y_0$. Then ρ_n is bounded above by the measure of S_j that lies between $\lambda_1\psi + P$ and $\lambda_2\psi + P$ for some real numbers λ_1 and λ_2 with

$$|\lambda_1 - \lambda_2| = \frac{(8L+2)\varepsilon j^2}{\psi(x_0 - y_0)} \leqslant 5(8L+2) j^2 \varepsilon^{1-\alpha}.$$

Recall that the uniform probability measure on S_j was defined by identifying S_j with a convex set $U_j \subset \mathbb{R}^{d_j}$; let us now think of P and ψ as lying in $\mathbb{R}^{d_j}$. Let h be the Euclidean norm of ψ and let θ be the angle that ψ makes with P. Let A be the maximum $(d_j - 1)$-dimensional volume of the intersection R of a translate of P with U_j. Then ρ_n is bounded above by $5(8L+2)\varepsilon^{1-\alpha} j^2 h A \sin\theta$ divided by the (d_j-dimensional) volume of U_j.

Now ψ and $-\psi$ both belong to U_j, and by convexity U_j contains the cones with vertices ψ and $-\psi$ and base R. The sum of the heights of these cones is $2h\sin\theta$, and thus the volume of U_j is at least $d_j^{-1} A 2h \sin\theta$. Therefore

$$\rho_n \leqslant 5(4L+1)\varepsilon^{1-\alpha} j^2 d_j,$$

and as in the proof of lemma 3.7,

$$\mu(Q_{2^{-j}}) \leqslant C_1^2 \varepsilon^{-2\delta} \rho_n^N \leqslant C_1^2 2^{2j\delta} 5^N (4L+1)^N j^{2N} 2^{-j(1-\alpha)N} C_3^N 2^{j\alpha\sigma N}.$$

Combining the terms that are exponential in j yields the desired result. □

The analogue of corollary 3.8 is the following.

Corollary 3.11. *Under the hypotheses of theorem 3.9, for almost every f*

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

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

4. Delay-coordinate version of results

Recall that one motivation for results like those in the previous section is to describe the relation between an attractor and its representation in an N-dimensional 'embedding space'. In an experimental situation, this representation could be made by recording N independent measurements of the state of the system simultaneously over a period of time. Since the theorems above apply to 'almost every' representation, in applying them the assumption would be that the choice of measurements is somehow 'typical'.

However, in practice it can be difficult to make many independent measurements of the system state. It is more common to make a single measurement at each time and use the *delay-coordinate method*. To describe the method, assume one makes a series of measurements in increments of time T (called the *delay*). The result is a time series $f(x_t)$, where $t = t_0, t_0 + T, t_0 + 2T, \ldots,$ f is a measurement function defined on the state space, and x_t is the state of the system at time t. The delay-coordinate 'embedding' is obtained by constructing the N-dimensional vectors whose coordinates consist of N consecutive numbers in the time series.

More formally, let ϕ be a flow on a Banach space B generated by a Lipschitz vector field, and let T be positive number. Let $f : B \to \mathbb{R}$ be a Lipschitz function. Then the *delay-coordinate function* $F : B \to \mathbb{R}^N$ is defined by

$$F(x) = (f(x), f(\phi^{-T}x), \ldots, f(\phi^{-(N-1)T}x)).$$

For a more detailed discussion of the delay-coordinate method see [Ta, ER, SYC].

Notice that the delay-coordinate functions F form a subspace of the Lipschitz functions from B to $\mathbb{R}^N$, and therefore our results about almost every Lipschitz function do not necessarily apply to delay-coordinate functions. See [SYC] for a description of difficulties that can arise, for instance when the flow has an orbit of period T. Nonetheless, one can obtain a positive result with some assumptions about ϕ and T. We state an analogue of theorem 3.1, based on theorem 2.5 of [SYC]; the proof simply combines the arguments of these two theorems.

Theorem 4.1. *Let ϕ be a flow on an open set $U \subset \mathbb{R}^M$, and let $X \subset U$ be a compact set with box-counting dimension d. Let $N > 2d$ be an integer, and let α be a real number with*

$$0 < \alpha < \frac{N - 2d}{N}.$$

Let $T > 0$ be a real number, and assume that X contains only finitely many fixed points of ϕ and no periodic orbits of period $T, 2T, 3T, \ldots, NT$. Then for a 'Lebesgue almost every' bounded linear function (or almost every, in the sense of prevalence, C^1 or Lipschitz function) $f : \mathbb{R}^M \to \mathbb{R}$, and corresponding delay-coordinate function $F : \mathbb{R}^M \to \mathbb{R}^N$, there exists $C > 0$ such that for all $x, y \in X$,

$$C|F(x) - F(y)|^\alpha \geqslant |x - y|.$$

As remarked in [SYC], the hypotheses of this theorem can be satisfied by making T sufficiently small. Further, it follows from the Kupka–Smale theorem [Ku, S, PM] (see also [Kal] for a prevalent version) that for generic flows, the hypotheses are satisfied for all but finitely many T.

We also have the following corollary about the effect of typical delay-coordinate functions on the dimension of X, though it seems likely in this finite-dimensional case that a better result is possible.

Corollary 4.2. *Under the same hypotheses, for almost every f*

$$\frac{N - 2d}{N} \dim(X) \leqslant \dim(F(X)) \leqslant \dim(X),$$

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

We feel that it would be of considerable interest to extend the results of this section to infinite-dimensional case; that is, to prove a delay-coordinate analogue of theorem 3.6 and/or theorem 3.9.

Acknowledgments

We thank J C Robinson and J A Yorke for their valuable input, and the referees for their careful readings and helpful comments. The second author is grateful to the Institute for Physical Science and Technology, University of Maryland for its hospitality. US Department of Energy (Mathematical, Information, and Computational Sciences Division, High Performance Computing and Communications Program) and by the National Science Foundation (Divisions of Mathematical Sciences and Physics).

References

[AGV] Arnold V I, Gusein-Zade S M and Varchenko A N 1985 *Singularities of Differentiable Maps* vol 1 (Boston, MA: Birkhauser)

[BV] Babin A V and Vishik M I 1983 Regular attractors of semigroups and evolution equations *J. Math. Pures Appl.* **62** 441–91

[BEFN] Ben-Artzi A, Eden A, Foias C and Nicolaenko B 1993 Hölder continuity for the inverse of Mañé's projection *J. Math. Anal. Appl.* **178** 22–9

[CF] Constantin P and Foias C 1985 Global Lyapunov exponents, Kaplan–Yorke formulae and the dimension of the attractors for 2D Navier–Stokes equations *Commun. Pure Appl. Math.* **38** 1–27

[CFT1] Constantin P, Foias C and Temam R 1985 Attractors representing turbulent flows *Mem. Am. Math. Soc.* **53** 314

[CFT2] Constantin P, Foias C and Temam R 1988 On the dimension of the attractors in two-dimensional turbulence *Physica* D **30** 284–96

[ER] Eckmann J-P and Ruelle D 1985 Ergodic theory of chaos and strange attractors *Rev. Mod. Phys.* **57** 617–56

[EFNT] Eden A, Foias C, Nicolaenko B and Temam R 1994 *Exponential Attractors for Dissipative Evolution Equations (Res. Appl. Math.)* (New York: Wiley)

[Fa] Falconer K 1990 *Fractal Geometry: Mathematical Foundations and Applications* (New York: Wiley)

[FH] Falconer K and Howroyd J D 1996 Projection theorems for box and packing dimensions *Math. Proc. Camb. Phil. Soc.* **119** 287–95

[FO] Foias C and Olson E 1996 Finite fractal dimension and Hölder–Lipschitz parametrization *Indiana Univ. Math. J.* **45** 603

[FT] Foias C and Temam R 1979 Some analytic and geometric properties of the solutions of the Navier–Stokes equations *J. Math. Pures Appl.* **58** 339–68

[FR] Friz P and Robinson J Smooth attractors have zero 'thickness' *Preprint*

[Fr] Frostman O 1935 Potential d'équilibre et capacité des ensembles avec quelques applications à la théorie des fonctions *Meddel. Lunds Univ. Math. Sem.* **3** 1–118

[H] Hale J K 1988 Asymptotic behaviour of dissipative systems *(Math. Surveys and Monographs vol 25)* (Providence, RI: American Mathematical Society)

[HK] Hunt B R and Yu Kaloshin V 1997 How projections affect the dimension spectrum of fractal measures *Nonlinearity* **10** 1031–46

[HSY] Hunt B R, Sauer T and Yorke J A 1992 Prevalence: a translation-invariant almost every for infinite dimensional spaces *Bull. Am. Math. Soc.* **27** 217–38

Hunt B R, Sauer T and Yorke J A 1993 Prevalence: an addendum *Bull. Am. Math. Soc.* **28** 306–7

[J] Järvenpää M 1994 On the upper Minkowski dimension, the packing dimension, and orthogonal projections *Ann. Acad. Sci. Fenn. Math. Diss.* **99** 1–34

[Kal] Yu Kaloshin V 1996 Some prevalent properties of smooth dynamical systems *Differential Equations with Complex and Real Time, (Proc. Steklov Inst. Math. vol 213)* ed Yu S Il'yashenko (Moscow: Maik Nauka)

[Kau] Kaufmann R 1968 On Hausdorff dimension of projections *Mathematika* **15** 153–5

[Ku] Kupka I 1963 Contribution à la théorie des champs génériques *Contrib. Diff. Eqn* **2** 457–84

Kupka I 1964 *Contrib. Diff. Eqn* **3** 411–20

[L] Ladyzhenskaya O A 1985 On the finiteness of the dimension of bounded invariant sets for the Navier–Stokes equations and other related dissipative systems *J. Sov. Math.* **28** 714–25

[Mal] Mallet-Paret J 1976 Negatively invariant sets of compact maps and an extension of a theorem of Cartwright *J. Diff. Eqns* **22** 331

[Man1] Mañé R 1976 Reduction of semilinear parabolic equations to finite dimensional C^1 flows *Geometry and Topology (Lecture Notes in Mathematics vol 597)* (Berlin: Springer) pp 361–78

[Man2] Mañé R 1981 On the dimension of the compact invariant sets of certain non-linear maps *Dynamical Systems and Turbulence (Lecture Notes in Mathematics vol 898)* (Berlin: Springer) pp 230–42

[Mar] Marstrand J M 1954 Some fundamental geometrical properties of plane sets of fractional dimension *Proc. London Math. Soc.* **4** 257–302

[Mat1] Mattila P 1975 Hausdorff dimension, orthogonal projections and intersections with planes *Ann. Acad. Sci. Fenn.* A **1** 227–44

[Mat2] Mattila P 1995 *Geometry of Sets and Measures in Euclidean Spaces* (Cambridge: Cambridge University Press)

[NST1] Nicolaenko B, Scheurer B and Temam R 1985 Some global dynamical properties of the Kuramoto–Sivashinsky equations *Physica* D **16** 155–83

[NST2] Nicolaenko B, Scheurer B and Temam R 1988 Some global dynamical properties of a class of pattern formation equations *Commun. Part. Diff. Eqns* **14** 245–97

[PM] Palis J and de Melo W 1982 *Geometric Theory of Dynamical Systems: An Introduction* (New York: Springer) translated by A K Manning

[SY] Sauer T and Yorke J A 1997 Are the dimensions of a set and its image equal under typical smooth functions? *Ergod. Theor. Dynam. Sys.* **17** 941–56

[SYC] Sauer T, Yorke J A and Casdagli M 1993 Embedology *J. Stat. Phys.* **71** 529–47

[S] Smale S 1963 Stable manifolds for differential equations and diffeomorphisms *Ann. Sci. Norm. Sup. Pisa* **18** 97–116

[Ta] Takens F 1981 Detecting strange attractors in turbulence *Dynamical Systems and Turbulence (Lecture Notes in Mathematics vol 898)* (New York: Springer) pp 366–81

[Te1] Temam R 1985 Attractors for Navier–Stokes equations *Nonlinear Partial Differential Equations and Their Applications vol VII (Research Notes in Mathematics vol 122)* (London: Pitman)

[Te2] Temam R 1988 *Infinite Dimensional Dynamical Systems in Mechanics and Physics (Appl. Math. Sci. vol 68)* (New York: Springer)

[W] Whitney H 1963 Differentiable manifolds *Ann. Math.* **37** 645–80