Nonlinearity **10** (1997) 1031–1046. Printed in the UK PII: S0951-7715(97)80912-7

How projections affect the dimension spectrum of fractal measures

Brian R Hunt†§ and Vadim Yu Kaloshin‡‖

† Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742-2431, USA

‡ Department of Mathematics, Princeton University, Princeton, NJ 08544-1000, USA

Received 8 January 1996, in final form 19 June 1997
Recommended by P Grassberger

Abstract. We introduce a new potential-theoretic definition of the dimension spectrum D_q of a probability measure for $q > 1$ and explain its relation to prior definitions. We apply this definition to prove that if $1 < q \leqslant 2$ and μ is a Borel probability measure with compact support in $\mathbb{R}^n$, then under almost every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, the q-dimension of the image of μ is $\min(m, D_q(\mu))$; in particular, the q-dimension of μ is preserved provided $m \geqslant D_q(\mu)$. We also present results on the preservation of information dimension D_1 and pointwise dimension. Finally, for $0 \leqslant q < 1$ and $q > 2$ we give examples for which D_q is not preserved by any linear transformation into $\mathbb{R}^m$. All results for typical linear transformations are also proved for typical (in the sense of prevalence) continuously differentiable functions.

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

1. Introduction

In the study of chaotic dynamical systems there has been much interest in various notions of the dimension of an attractor. Experimentalists usually measure the dimension of an attractor by creating low-dimensional 'embeddings' of the attractor from data. Does such a measurement indicate the true dimension of the attractor in its natural (often high-dimensional, or even infinite-dimensional) state space? We prove a result giving an affirmative answer to this question for part of the dimension spectrum D_q that is commonly used by dynamicists, and give examples that indicate why the answer may be negative for the other parts of the spectrum. We also provide a new definition for part of the spectrum ($q > 1$). This definition is related to the known potential-theoretic definitions of Hausdorff and correlation (D_2) dimensions and may allow further analysis of the mathematical properties of the dimension spectrum.

One way to define the dimension of a chaotic attractor is to view it as a set of points. The fractal dimension of a set can be defined in a number of ways, including the box-counting, Hausdorff, and packing dimensions (see, for example, [Fal, Mat3]). For an arbitrary set these dimensions may differ, but they are generally expected to coincide for an attracting set of a dynamical system. However, it is often useful to consider other notions of dimension that take into account the distribution of points induced by the dynamics on the attractor.

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

Thus we associate to each attractor its 'natural measure'—that is, the probability measure induced by the statistics of a typical trajectory that approaches the attractor (see, for example, [FOY, ER]. The natural measure is not known to exist for arbitrary attractors, but for Axiom A attractors it exists [Sin, BoR] and is often called the Sinai–Ruelle–Bowen (SRB) measure). The dimension of a probability measure can also be defined in several ways. Measure-dependent dimensions of attractors can be more readily related to the dynamics on the attractor and more easily measured from numerical or experimental data than the dimensions of the attracting set.

One important notion of the dimension of a probability measure is the 'information dimension' [BaR, Far]. For natural measures associated with chaotic attractors, the information dimension is generally expected to coincide with several other notions of the dimension of a measure [Y, FOY] (the coincidence is now known for invariant measures of $C^{1+\beta}$ diffeomorphisms with nonzero Lyapunov exponents [BPS]). The information dimension is also conjectured for typical attractors to be equal to the 'Lyapunov dimension', which is defined in terms of the Lyapunov exponents [KY, FKYY, AY]. A more complete characterization of the fractal (or 'multifractal') structure of a chaotic attractor is obtained by considering the 'dimension spectrum' D_q, which includes both the box-counting dimension (D_0) and information dimension (D_1) [Re, Gr, HP, HJKPS]. In addition the correlation dimension D_2 has received much individual attention and is perhaps the easiest dimension in the spectrum to estimate from data [GP1, GP2, C1, Pe].

In [DGOSY, SY] there appears a definition of correlation dimension D_2,

$$D_2(\mu) = \sup\left\{s : \iint \frac{\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)}{|x-y|^s} < \infty\right\} \tag{1.1}$$

which is related to the well-known potential theoretic definition of Hausdorff dimension introduced by Frostman [Fr] (see also [Kah, Fal, Mat3]). We present a generalization of (1.1) to the dimension spectrum D_q. For $q > 1$ the q-dimension of a measure can be defined as

$$D_q(\mu) = \sup\left\{s : \int \left(\int \frac{\mathrm{d}\mu(y)}{|x-y|^s}\right)^{q-1} \mathrm{d}\mu(x) < \infty\right\}. \tag{1.2}$$

This definition allows the application of techniques developed by Kaufmann [Kau] and Mattila [Mat1, Mat2], who studied properties of intersections and projections of fractal sets in terms of the Hausdorff dimension, to part of the dimension spectrum. In this paper we concentrate on the property of dimension preservation under typical projections.

We say that a particular dimension $D(\mu)$ of a Borel probability measure μ on $\mathbb{R}^n$ is 'preserved under typical projections' (or simply 'preserved') if for all μ and all $m < n$, almost every orthogonal projection $P : \mathbb{R}^n \to \mathbb{R}^m$ (with respect to the Haar measure on the group of orthogonal projections) has the property that $D(P(\mu)) = \min(m, D(\mu))$. (The measure $P(\mu)(S)$ of a set $S \subset \mathbb{R}^m$ is defined to be $\mu(P^{-1}(S))$.) In particular, if $m \geqslant D(\mu)$ then the dimension of $P(\mu)$ is the same as the dimension of μ for almost every P. It has long been known that the Hausdorff dimension of a set is preserved in this sense [Mar, Kau, Mat1]. More recently it was shown that the correlation dimension D_2 and, under some restrictions on μ the information dimension D_1, are preserved [SY], but that there are sets in $\mathbb{R}^n$ for which the packing dimension and box-counting dimension D_0 are not preserved under any projection to a lower-dimensional space [J, FH, SY]. (However, the constructed sets are pathological from the point of view of dynamical systems. Recall that the box-counting, Hausdorff, and packing dimensions are expected to coincide for typical attractors of dynamical systems; if they do, then since the Hausdorff dimension is a lower bound on the other two dimensions, and it is preserved under typical projections, the packing and box-counting dimensions are preserved as well.) See also [HT, FM, Ol2] for recent results on the dimension of projections, products, and cross-sections of measures.

It is important to know whether the dimensions used in the analysis of experimental data are preserved under typical projections. Assume that m simultaneous measurements are made of a system with n-dimensional state space. Typically n is a lot larger than m. The measured quantities are a function $f : \mathbb{R}^n \to \mathbb{R}^m$ of the state of the system; thus if the measurements are recorded over time then the distribution of points observed in $\mathbb{R}^m$ is the image under f of the natural measure μ on the attractor of the system. We think of f as a 'generalized projection'—though f may not be linear, if f is C^1 then we can think of f as a diffeomorphism $x \mapsto (x, f(x))$ followed by a projection $(x, f(x)) \mapsto x$. All of the dimensions we consider in this paper are preserved under diffeomorphisms. Thus whether or not the dimension spectrum computed from the measurements should be expected to reflect the true dimension spectrum of μ depends on whether the dimension spectrum is preserved under typical projections.

Our main result is as follows.

Theorem 1.1. *Let μ be a Borel probability measure on $\mathbb{R}^n$ with compact support and let q satisfy $1 < q \leqslant 2$. Then for almost every linear transformation L from $\mathbb{R}^n$ to $\mathbb{R}^m$ (in the sense of the Lebesgue measure on the space of $m \times n$ matrices), $D_q(L(\mu)) = \min(m, D_q(\mu))$.*

It follows that the same is true for almost every orthogonal projection of $\mathbb{R}^n$ to $\mathbb{R}^m$ (in the sense of the Haar measure). Indeed, with an appropriate change of coordinates in both $\mathbb{R}^n$ and $\mathbb{R}^m$, every linear transformation of rank r can be put into the form of an $r \times r$ unit matrix with all other entries zero (see, for example, [AGV]). Since almost every linear $L : \mathbb{R}^n \to \mathbb{R}^m$ has rank m and since the dimension spectrum is preserved under linear changes of coordinates, preservation under almost every orthogonal projection is equivalent to preservation under almost every linear transformation. Furthermore, we extend theorem 1.1 to hold for 'almost every' C^1 function f from $\mathbb{R}^n$ to $\mathbb{R}^m$, in the sense of prevalence (see section 2.2 for a definition).

For $0 \leqslant q < 1$ and $q > 2$ we will give examples of μ for which D_q is not preserved by any linear transformation to a lower dimensional space. For $q > 2$ there is a particularly simple example, which we present here. Let μ be the uniformly distributed probability measure on the unit circle S^1 in the plane $\mathbb{R}^2$; then $D_q(\mu) = 1$ for all $q \geqslant 0$. Every orthogonal projection $P : \mathbb{R}^2 \to \mathbb{R}$ yields the same measure $P(\mu)$, which has a singularity at ± 1 (see figure 1) such that the measure of an interval of length ϵ near ± 1 is proportional to $\epsilon^{1/2}$ for small ϵ. We will show that as a result, $D_q(P(\mu)) < 1$ for all $q > 2$. (In fact, $P(\mu)$ can be thought of as the invariant measure for the quadratic map $x \mapsto 1 - 2x^2$, for which the dimension spectrum was examined in [OWY].) We emphasize that this example is not at all pathological from the point of view of dynamical systems; indeed the natural measure of an attracting periodic orbit in $\mathbb{R}^n$ has the same properties under projection into $\mathbb{R}$.

Many authors consider a different 'dimension spectrum' $f(\alpha)$, defined for a given μ to be the Hausdorff dimension of the set of points at which μ has pointwise dimension α (see section 4 for a definition of pointwise dimension). In some cases $f(\alpha)$ is the Legendre transform of $(q-1)D_q(\mu)$:

$$f(\alpha) = \inf_q(\alpha q - (q-1)D_q(\mu))$$

(see, for example, the introduction of [BMP] for a heuristic derivation of this formula). How generally this 'multifractal formalism' applies is a subject of ongoing investigation

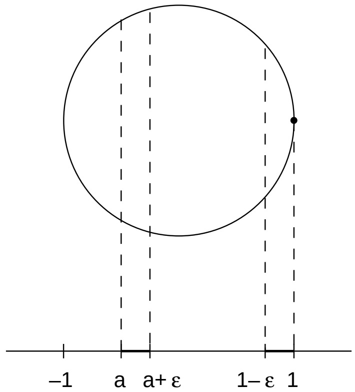

Figure 1. Projection of the unit circle onto a line.

[HJKPS, CLP, Ra, B, L, Po, BMP, CM, Sim, Ol1, Ri, PW]. Thus it is unclear what our results indicate for the preservation of the $f(\alpha)$ spectrum under typical projections.

In section 2.1 we give a formal definition of the dimension spectrum D_q, and prove that (1.2) is equivalent to previous definitions of the lower q-dimension. In section 2.2 we discuss the notion of prevalence [HSY], in terms of which we generalize our results for almost every linear transformation to hold for 'almost every' C^1 function. In section 3 we prove theorem 1.1, and in section 4 we consider the preservation of pointwise and information dimensions. Finally, in section 5 we give a counterexample to the preservation of D_q for $0 \leqslant q < 1$ and we indicate how higher-dimensional analogues of the counterexample for $q > 2$ given above can be constructed. In particular, we give examples of measures for which D_q decreases for large q under all projections into $\mathbb{R}^m$, even when m is many times larger than the dimension of the support of the measure.

2. Preliminaries

2.1. The dimension spectrum

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

$$C_q(\mu, \varepsilon) = \int [\mu(B(x, \varepsilon))]^{q-1} \, \mathrm{d}\mu(x)$$

where $B(x, \varepsilon)$ is the closed ball of radius ε centred at x.

Definition 2.1. For $q \geqslant 0, q \neq 1$, the lower and upper q-dimensions of μ are

$$\begin{aligned} D_q^-(\mu) &= \liminf_{\varepsilon \to 0} \frac{\log C_q(\mu, \varepsilon)}{(q-1)\log \varepsilon} \\ D_q^+(\mu) &= \limsup_{\varepsilon \to 0} \frac{\log C_q(\mu, \varepsilon)}{(q-1)\log \varepsilon}. \end{aligned} \tag{2.1}$$

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

It is expected that $D_q(\mu)$ exists (that is, $D_q^-(\mu) = D_q^+(\mu)$) for sufficiently regular fractal measures μ, in particular those that arise from smooth dynamical systems. For such a measure μ, the function $q \to D_q(\mu)$ is called the 'dimension spectrum' of μ. A fundamental property of the spectra $D_q^{\pm}$ is that they are nonincreasing as a function of q and continuous except possibly at $q = 1$ (see, for example, [B, St]). The case $q = 1$ is discussed in section 4.

Definition 2.1 is the same as Pesin's definition [Pe] of the 'generalized spectrum of dimensions' with the notational conversion $D_q^- = \underline{\gamma}_{q-1}$ and $D_q^+ = \overline{\gamma}_{q-1}$. Our notation for the dimension spectrum follows the physics literature originating with [Gr, HP], which use instead the following definition. For $\varepsilon > 0$, cover the support of μ with a grid of cubes with edge length ε. Let $N(\varepsilon)$ be the number of cubes that intersect the support of μ, and let the measures of these cubes be $p_1, p_2, \ldots, p_{N(\varepsilon)}$. Write

$$\begin{aligned} D_q^-(\mu) &= \liminf_{\varepsilon\to 0} \frac{\sum_{i=1}^{N(\varepsilon)} p_i^q}{(q-1)\log\varepsilon} \\ D_q^+(\mu) &= \limsup_{\varepsilon\to 0} \frac{\sum_{i=1}^{N(\varepsilon)} p_i^q}{(q-1)\log\varepsilon}. \end{aligned} \tag{2.2}$$

For $q \geqslant 0$, $q \neq 1$ these limits are independent of the choice of ε-grids, and give the same values as definition 2.1 (this is proved for $q > 1$ in [Ri] and can be proved for $0 \leqslant q < 1$ as well). On the other hand, Riedi [Ri] shows that the limits in (2.2) can be grid-dependent for $q < 0$, so a more careful definition is needed in that case. We consider only $q \geqslant 0$ in this paper.

For $q > 1$ we introduce a potential-theoretic definition of the lower q-dimension D_q^-. Of course it then also provides a definition for D_q when the latter exists. For $s \geqslant 0$ consider the function

$$\varphi_s(\mu, x) = \int |x-y|^{-s} \mathrm{d}\mu(y) \tag{2.3}$$

which is called the s-potential of the measure μ at the point x. Let us define the (s, q)-energy of μ to be

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

In proposition 2.1 below we show that $I_{s,q}(\mu)$ is finite when $s < D_q^-(\mu)$ and infinite when $s > D_q^-(\mu)$.

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

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

Frostman [Fr] (see also [Fal, Mat3]) showed that the Hausdorff dimension of a Borel set $S \subset \mathbb{R}^n$ is equal to the supremum of s for which there exists a Borel probability measure μ with $\mu(S) = 1$ and $I_s(\mu) < \infty$, and this characterization is used by Kaufmann [Kau] and Mattila [Mat1] to prove their results on the preservation of the Hausdorff dimension. Also, the lower correlation dimension $D_2^-(\mu)$ can be expressed [DGOSY, SY] as the supremum of s for which $I_s(\mu) < \infty$, and this in turn is used by Sauer and Yorke [SY] to establish the preservation of the correlation dimension. Proposition 2.1 generalizes this notion to the lower-dimension spectrum D_q^- for $q > 1$, and thus allows the methods of potential theory to be applied to this part of the spectrum.

Proposition 2.1. *If $q > 1$ and μ is a Borel probability measure, then*

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

Proof. We first show that if $s > D_q^-(\mu)$, then $I_{s,q}(\mu)$ is infinite. Choose $\delta > 0$ such that $s - \delta > D_q^-(\mu)$. It follows that there exists a sequence $\{\varepsilon_k\}_{k=1}^{\infty}$ such that $\varepsilon_k \to 0$ as $k \to \infty$ and $C_q(\mu, \varepsilon_k) > \varepsilon_k^{(s-\delta)(q-1)}$. Note that for all $\varepsilon > 0$,

$$I_{s,q}(\mu) = \int [\varphi_s(\mu, x)]^{q-1} \, \mathrm{d}\mu(x) \geqslant \int \left(\int_{|x-y|<\varepsilon} \frac{\mathrm{d}\mu(y)}{|x-y|^s} \right)^{q-1} \mathrm{d}\mu(x)$$
$$\geqslant \varepsilon^{-s(q-1)} C_q(\mu, \varepsilon).$$

Thus $I_{s,q}(\mu) \geqslant \varepsilon_k^{-s(q-1)} C_q(\mu, \varepsilon_k) \geqslant \varepsilon_k^{-\delta(q-1)}$ for all k, and letting $k \to \infty$ we have that $I_{s,q}(\mu) = \infty$.

Next we show that $I_{s,q}(\mu)$ is finite for $s < D_q^-(\mu)$. We assume that $\mu(\{x\}) = 0$ for all $x \in \mathbb{R}^n$, since otherwise $D_q^-(\mu) = 0$ and $I_{s,q}(\mu)$ is trivially seen to be finite for $s < 0$. It then follows that

$$\varphi_s(\mu, x) = \int_{|x-y|>1} \frac{\mathrm{d}\mu(y)}{|x-y|^s} + \sum_{n=0}^{\infty} \int_{2^{-n-1}<|x-y|\leqslant 2^{-n}} \frac{\mathrm{d}\mu(y)}{|x-y|^s}$$
$$\leqslant 1 + \sum_{n=0}^{\infty} 2^{s(n+1)} \mu(B(x, 2^{-n})).$$

Case 1, $1 < q \leqslant 2$. Since $0 < q - 1 \leqslant 1$, by Jensen's inequality

$$I_{s,q}(\mu) = \int [\varphi_s(\mu, x)]^{q-1} \, \mathrm{d}\mu(x) \leqslant \int \left(1 + \sum_{n=0}^{\infty} [2^{s(n+1)} \mu(B(x, 2^{-n}))]^{q-1} \right) \mathrm{d}\mu(x)$$
$$= 1 + \sum_{n=0}^{\infty} 2^{s(n+1)(q-1)} C_q(\mu, 2^{-n}).$$

Choose $\delta > 0$ such that $s + \delta < D_q^-(\mu)$. There exists ε_0 such that $C_q(\mu, \varepsilon) < \varepsilon^{(s+\delta)(q-1)}$ for any $\varepsilon < \varepsilon_0$. Then when $2^{-n} < \varepsilon_0$, we have

$$2^{s(n+1)(q-1)} C_q(\mu, \varepsilon) \leqslant 2^{(s-n\delta)(q-1)}$$

and the sum over n converges geometrically; thus $I_{s,q}(\mu) < \infty$.

Case 2, $2 < q < \infty$. By the triangle inequality for the L^{q-1} norm $\| \cdot \|_{q-1}$ with respect to the measure μ,

$$[I_{s,q}(\mu)]^{1/(q-1)} = \|\varphi_s(\mu)\|_{q-1} \leqslant 1 + \sum_{n=0}^{\infty} \|2^{s(n+1)} \mu(B(x, 2^{-n}))\|_{q-1}$$
$$\leqslant 1 + \sum_{n=0}^{\infty} 2^{s(n+1)} [C_q(\mu, 2^{-n})]^{1/(q-1)}.$$

We proceed as in case 1 to show that the sum converges geometrically and hence that $I_{s,q}(\mu)$ is finite; this completes the proof. □

Remark. Using the argument of [SY] (see the proof of proposition 2.3 therein) one can prove also that

$$D_q^-(\mu) = \sup \left\{ s : \int_0^{\infty} \varepsilon^{-s(q-1)-1} C_q(\mu, \varepsilon) \, \mathrm{d}\varepsilon < \infty \right\} \qquad \text{for } q > 1$$
$$D_q^+(\mu) = \inf \left\{ s : \int_0^{\infty} \varepsilon^{-s(q-1)-1} C_q(\mu, \varepsilon) \, \mathrm{d}\varepsilon < \infty \right\} \qquad \text{for } 0 \leqslant q < 1.$$

Again, if $D_q(\mu)$ exists then it too equals the right-hand side of the equalities given above.

2.2. Prevalence

It is desirable to extend the results of this paper that hold for almost every linear transformation (or equivalently, almost every orthogonal projection) from $\mathbb{R}^n$ to $\mathbb{R}^m$ to hold also for 'almost every' smooth function between these spaces. Indeed the measurements taken from an experiment may not be linear functions of the state variables of the system describing the experiment. An obstacle to such a result is the lack of an analogue of Lebesgue or Haar measure on the infinite-dimensional space $C^1(\mathbb{R}^n, \mathbb{R}^m)$.

One option for describing generic properties on infinite-dimensional spaces is to use the topological notion of a *generic* or *residual* set: a set that contains a countable intersection of open dense sets. The drawback of this approach is that topological genericity is not equivalent to the notion of 'Lebesgue almost every' on finite-dimensional spaces; indeed there are residual sets with Lebesgue measure zero (see, for example, [Ox]). We use instead the notion of prevalence [HSY], which is equivalent to Lebesgue almost every on finite-dimensional space.

To motivate the definition of prevalence we recall the following consequence of the Fubini theorem. *A subset P of $\mathbb{R}^k \times \mathbb{R}^\ell$ has full $(\ell + k)$-dimensional Lebesgue measure if and only if for Lebesgue almost every $x \in \mathbb{R}^\ell$ the intersection $P \cap (\mathbb{R}^k \times \{x\})$ has full k-dimensional Lebesgue measure*. We think of an infinite-dimensional space X as the Cartesian product of a finite-dimensional space Z and a complement space Y so that every point $x \in X$ as a vector has unique representation $x = z + y$, $z \in Z$, $y \in Y$. Roughly speaking, a subset P of X is called *prevalent* if for every $y \in Y$ the intersection $P \cap Z + y$ has full Lebesgue measure ($Z + y$ is finite-dimensional). We cannot say 'for Lebesgue almost every $y \in Y$' because Y is infinite-dimensional.

The formal definition of prevalence is as follows. A Borel measure μ on the complete metric linear space X, with the property that $\mu(U)$ is positive and finite for some compact set $U \subset X$, is said to be *transverse* to a Borel set $S \subset X$ if $\mu(S+x) = 0$ for every $x \in X$. A Borel set $P \subset X$ is called *prevalent* if there exists a measure transverse to the complement of P. (In the previous paragraph we took μ to be the Lebesgue measure on the finite-dimensional space L.) The definition of prevalence can be extended to infinite-dimensional manifolds that are not vector spaces; see [Kal].

3. Proof of the preservation theorem

In this section we prove theorem 1.1 and the following analogue for smooth functions. Recall that if μ is a Borel probability measure on $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^m$ is a measurable function then we define the measure $f(\mu)$ by the equation $f(\mu)(S) = \mu(f^{-1}(S))$ for all Borel sets $S \subset \mathbb{R}^m$.

Theorem 3.1. *Let μ be a Borel probability measure on $\mathbb{R}^n$ with compact support and let q satisfy $1 < q \leqslant 2$. Assume $D_q(\mu)$ exists; that is, $D_q^-(\mu) = D_q^+(\mu)$. Then for a prevalent set of functions $f \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, the q-dimension of $f(\mu)$ exists and is given by $D_q(f(\mu)) = \min(m, D_q(\mu))$.*

Remark. In the case that $D_q^-(\mu) \neq D_q^+(\mu)$, theorems 1.1 and 3.1 apply with D_q replaced by D_q^-.

Theorems 1.1 and 3.1 are both implied by the following result (theorem 1.1 follows from applying it to arbitrary linear transformations f_0.)

Proposition 3.2. *Let μ be a Borel probability measure on $\mathbb{R}^n$ with compact support, and assume that $D_q(\mu)$ exists for some $1 < q \leqslant 2$. Let Z denote the set of $m \times n$ matrices whose* *entries are bounded by* 1 *in absolute value. Then for all* $f_0 \in C^1(\mathbb{R}^n, \mathbb{R}^m)$*, the function* $f_L = f_0 + L$ *satisfies* $D_q(f_L(\mu)) = \min(m, D_q(\mu))$ *for Lebesgue almost every* $L \in Z$*.*

Proof. First we claim that $D_q^-(f(\mu)) \leqslant D_q^-(\mu)$ and $D_q^+(f(\mu)) \leqslant D_q^+(\mu)$ for all C^1 functions $f : \mathbb{R}^n \to \mathbb{R}^m$. Because μ has compact support, f has a uniform Lipschitz constant, say λ, on the support of μ. Then for all x in the support of μ we have $f(\mu)(B(f(x), \varepsilon)) = \mu(f^{-1}(B(f(x), \varepsilon))) \geqslant \mu(B(x, \varepsilon/\lambda))$, and hence $C_q(f(\mu), \varepsilon) \geqslant C_q(\mu, \varepsilon/\lambda)$ for all $q > 1$. Our claim then follows from (2.1).

It suffices to prove that $D_q^-(f_L(\mu)) \geqslant \min(m, D_q^-(\mu))$ for almost every $L \in Z$, because then if $D_q^-(\mu) = D_q^+(\mu)$, we know $D_q^+(f_L(\mu)) \leqslant \min(m, D_q^+(\mu)) = \min(m, D_q^-(\mu)) \leqslant D_q^-(f_L(\mu))$, and hence there is equality, for almost every $L \in Z$. In turn by proposition 2.1 it suffices to show for $0 \leqslant s < m$ that $I_{s,q}(\mu) < \infty$ implies that $I_{s,q}(f_L(\mu)) < \infty$ for almost every $L \in Z$.

Now

$$I_{s,q}(f_L(\mu)) = \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^m} \frac{\mathrm{d}f(\mu)(v)}{|u - v|^s} \right)^{q-1} \mathrm{d}f(\mu)(u)$$
$$= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \frac{\mathrm{d}\mu(y)}{|f_L(x) - f_L(y)|^s} \right)^{q-1} \mathrm{d}\mu(x).$$

Thus

$$\int_Z I_{s,q}(f_L(\mu))\,\mathrm{d}L = \int_Z \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \frac{\mathrm{d}\mu(y)}{|f_L(x) - f_L(y)|^s} \right)^{q-1} \mathrm{d}\mu(x)\,\mathrm{d}L$$
$$= \int_{\mathbb{R}^n} \int_Z \left(\int_{\mathbb{R}^n} \frac{\mathrm{d}\mu(y)}{|f_L(x) - f_L(y)|^s} \right)^{q-1} \mathrm{d}L\,\mathrm{d}\mu(x).$$

Since $1 \leqslant q < 2$, we have $0 < q - 1 \leqslant 1$, and thus by the Hölder inequality,

$$\int_Z I_{s,q}(f_L(\mu))\,\mathrm{d}L \leqslant \int_{\mathbb{R}^n} \left(\int_Z \int_{\mathbb{R}^n} \frac{\mathrm{d}\mu(y)}{|f_L(x) - f_L(y)|^s}\,\mathrm{d}L \right)^{q-1} \mathrm{d}\mu(x)$$
$$= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \int_Z \frac{\mathrm{d}L}{|f_L(x) - f_L(y)|^s}\,\mathrm{d}\mu(y) \right)^{q-1} \mathrm{d}\mu(x). \tag{3.1}$$

To estimate the inner integral we use a slight variation of lemma 2.5 from [SY], which is based on the fact that $\int |y|^{-s}\,\mathrm{d}y$ converges in $\mathbb{R}^m$ if and only if $s < m$.

Lemma 3.3. *Let A be a linear transformation from $\mathbb{R}^k$ to $\mathbb{R}^m$, let b be a vector in $\mathbb{R}^m$, and let Z be a cube in $\mathbb{R}^k$. Assume that the image of Z under A contains a cube of width δ in $\mathbb{R}^m$. Then for $0 \leqslant s < m$,*

$$\int_Z \frac{\mathrm{d}z}{|Az + b|^s} \leqslant \frac{C}{\delta^s}$$

where C is a constant depending only on Z, s, k and m.

We apply the lemma to the cube $Z \subset \mathbb{R}^{mn}$ described in the statement of the proposition, the linear transformation A that takes an $m \times n$ matrix L to $L(x - y)$, and the vector $b = f_0(x) - f_0(y)$ for a given x and y in $\mathbb{R}^n$. Note that some coordinate of $x - y$, say the jth coordinate, has a magnitude of at least $|x - y|/\sqrt{n}$. It follows that the set of matrices in Z that are zero except in the jth column takes $x - y$ to a cube of width at least $2|x - y|/\sqrt{n}$. Applying lemma 3.3 with $\delta = 2|x - y|/\sqrt{n}$, we find that

$$\int_Z \frac{\mathrm{d}L}{|f_L(x) - f_L(y)|^s} = \int_Z \frac{\mathrm{d}L}{|f_0(x) - f_0(y) + L(x - y)|^s} \leqslant \frac{K}{|x - y|^s}$$

where K is a constant depending only on s, n and m. It follows from (3.1) that

$$\int_Z I_{s,q}(f_L(\mu))\,\mathrm{d}L \leqslant \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}\frac{K\,\mathrm{d}\mu(y)}{|x-y|^s}\right)^{q-1}\mathrm{d}\mu(x) < \infty$$

provided $I_{s,q}(\mu) < \infty$. In particular, the integrand $I_{s,q}(f_L(\mu))$ is finite for almost every $L \in Z$, and the proof is complete. □

Remark. The results of this section can be extended to apply, for invariant measures of smooth dynamical systems, to almost every smooth delay-coordinate function with the same hypotheses as in [SY].

4. Preservation of pointwise and information dimensions

In this section we discuss the preservation of information and pointwise dimension under typical projections. Let μ be a Borel probability measure on $\mathbb{R}^n$ with compact support.

Definition 4.1. For each point x in the support of μ we define the *lower and upper pointwise dimensions* of μ at x to be

$$\alpha_\mu^-(x) = \liminf_{\varepsilon\to 0}\frac{\log\mu(B(x,\varepsilon))}{\log\varepsilon}$$
$$\alpha_\mu^+(x) = \limsup_{\varepsilon\to 0}\frac{\log\mu(B(x,\varepsilon))}{\log\varepsilon}$$

where $B(x,\varepsilon)$ denotes the ball of radius ε centred at x. If $\alpha_\mu^-(x) = \alpha_\mu^+(x)$, then their common value $\alpha_\mu(x)$ is called the *pointwise dimension* of μ at x.

If μ is an ergodic invariant measure of a $C^{1+\beta}$ diffeomorphism with nonzero Lyapunov exponents, then $\alpha_\mu(x)$ exists and is constant for almost every x with respect to μ [BPS]. In this case the information and Hausdorff dimensions of μ (defined below) coincide with the common value of the pointwise dimension [Y]. In [SY] it is shown that in this case, the information dimension and the 'almost everywhere' value of the pointwise dimension are preserved by typical projections provided the pointwise dimension of the projected measure exists almost everywhere. We will show how to eliminate the latter assumption, and discuss preservation of these dimensions in the more general case when the pointwise dimension is not constant almost everywhere.

In [SY] it is observed that the lower pointwise dimension is equal to the supremum of values s for which the s-potential of μ at x (given by (2.3)) is finite; that is,

$$\alpha_\mu^-(x) = \sup\{s : \varphi_s(\mu,x) < \infty\} = \inf\{s : \varphi_s(\mu,x) = \infty\}.$$

The proof of this formula is similar to the proof of proposition 2.1 and we omit it. We use the above characterization of the lower pointwise dimension to prove the following result. In the case of linear transformation, this theorem follows from the results of [HT].

Theorem 4.1. *Let μ be a Borel probability measure on $\mathbb{R}^n$ with compact support. For a prevalent set of C^1 functions (also, for almost every linear transformation) $f : \mathbb{R}^n \to \mathbb{R}^m$,*

$$\alpha_{f(\mu)}^-(f(x)) = \min(m, \alpha_\mu^-(x))$$

for almost every x with respect to μ. If in addition $\alpha_\mu(x)$ exists for almost every x, then for almost every f the pointwise dimension of $f(\mu)$ at $f(x)$ exists and is given by

$$\alpha_{f(\mu)}(f(x)) = \min(m, \alpha_\mu(x))$$

for almost every x.

Proof. We use the scheme of proposition 3.2, proving the results hold for $f_L = f_0 + L$ for an arbitrary C^1 function f_0 and for almost every $L \in Z$, where Z is the set of $m \times n$ matrices with entries of magnitude at most 1. As in the proof of proposition 3.2, we begin by observing that the upper and lower pointwise dimensions cannot increase under a smooth function f. Furthermore, $\alpha^+_{f(\mu)}(f(x)) \leqslant m$ for almost every x with respect to μ. (Though it is possible for a measure on $\mathbb{R}^m$ to have an upper pointwise dimension larger than m at some points in its support, one can show that the set of such points cannot have positive measure [C2].) It follows that we need only prove for almost every $L \in Z$ that $\alpha^-_{f_L(\mu)}(f_L(x)) \geqslant \min(m, \alpha^-_\mu(x))$ for almost every x.

For a given real s and positive integer N, let S_N be the set of x in the support of μ for which $\varphi_s(\mu, x) \leqslant N$. Notice that the set of all x for which $\alpha^-_\mu(x) > s$ is contained in the union of S_N for $N \geqslant 1$. Much as in the proof of proposition 3.2, we have that

$$\begin{aligned}\int_Z \int_{S_N} \varphi_s(f_L(\mu), f_L(x))\,\mathrm{d}\mu(x)\,\mathrm{d}L &= \int_Z \int_{S_N} \int_{\mathbb{R}^n} \frac{\mathrm{d}\mu(y)}{|f_L(x) - f_L(y)|^s}\,\mathrm{d}\mu(x)\,\mathrm{d}L \\ &= \int_{S_N} \int_{\mathbb{R}^n} \int_Z \frac{\mathrm{d}L}{|f_L(x) - f_L(y)|^s}\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(x) \\ &\leqslant K \int_{S_N} \int_{\mathbb{R}^n} \frac{\mathrm{d}\mu(y)}{|x-y|^s}\,\mathrm{d}\mu(x) \leqslant KN\end{aligned}$$

where K depends only on s, n, and m. We conclude that for almost every $L \in Z$, the integral $\int_{S_N} \varphi_s(f_L(\mu), f_L(x))\,\mathrm{d}\mu(x)$ is finite, or in other words $\varphi_s(f_L(\mu), f_L(x)) < \infty$ for almost every x in S_N with respect to μ. Taking the union over N, we have for almost every $L \in Z$ that $\varphi_s(f_L(\mu), f_L(x)) < \infty$, or equivalently $\alpha^-_{f_L(\mu)}(f_L(x)) \geqslant s$, for almost every x with $\alpha^-_\mu(x) > s$.

We have shown that for all real s,

$$\mu(\{x \in \mathbb{R}^n : \alpha^-_\mu(x) > s > \alpha^-_{f_L(\mu)}(f_L(x))\}) = 0$$

for almost every $L \in Z$. Taking the union over all rational s we conclude that for almost every $L \in Z$,

$$\mu(\{x \in \mathbb{R}^n : \alpha^-_\mu(x) > \alpha^-_{f_L(\mu)}(f_L(x))\}) = 0$$

for if $\alpha^-_\mu(x) > \alpha^-_{f_L(\mu)}(f_L(x))$ then there exists a rational s between the two values. The proof is now finished. □

Next we turn to the case of information (or Rényi) dimension. The following definition is equivalent to standard definition [BaR, Far].

Definition 4.2. The *lower and upper information dimensions* of the Borel probability measure μ are defined to be

$$D^-_1(\mu) = \liminf_{\varepsilon \to 0} \frac{\int \log \mu(B(x,\varepsilon))\,\mathrm{d}\mu(x)}{\log \varepsilon}$$

$$D^+_1(\mu) = \limsup_{\varepsilon \to 0} \frac{\int \log \mu(B(x,\varepsilon))\,\mathrm{d}\mu(x)}{\log \varepsilon}.$$

If $D^-_1(\mu) = D^+_1(\mu)$, then their common value $D_1(\mu)$ is called the *information dimension* of μ.

It can be shown [Ga, C1] that

$$\int \alpha^+_\mu(x)\,\mathrm{d}\mu(x) \geqslant D^+_1(\mu) \geqslant D^-_1(\mu) \geqslant \int \alpha^-_\mu(x)\,\mathrm{d}\mu(x)$$

(in fact from the definitions given here these inequalities are a simple consequence of the Fatou lemma). It follows that if the pointwise dimension of μ exists almost everywhere, then the information dimension of μ exists and is given by the average of the pointwise dimension; that is,

$$D_1(\mu) = \int \alpha_\mu(x)\,\mathrm{d}\mu(x). \tag{4.1}$$

The following result is an immediate consequence of theorem 4.1 and (4.1).

Corollary 4.2. *Let μ be a Borel probability measure on $\mathbb{R}^n$ with compact support. If the pointwise dimension $\alpha_\mu(x)$ exists and does not exceed m for almost every x with respect to μ, then for a prevalent set of C^1 functions (also, for almost every linear transformation) $f : \mathbb{R}^n \to \mathbb{R}^m$, the information dimension of $f(\mu)$ exists and is given by*

$$D_1(f(\mu)) = D_1(\mu).$$

Note that if $\alpha_\mu(x) > m$ for a positive measure set of x, then $D_1(f(\mu)) < D_1(\mu)$ for all C^1 functions $f : \mathbb{R}^n \to \mathbb{R}^m$, even though $D_1(\mu)$ may be less than m. An example is when μ consists of a mass $\beta \in (0, 1)$ uniformly distributed over an n-dimensional volume and a point mass of magnitude $1 - \beta$. Then $D_1(\mu) = \beta n$, and for $1 \leqslant m \leqslant n - 1$ we can choose β so that $D_1(\mu) < m$, while $D_1(f(\mu)) \leqslant \beta m < D_1(\mu)$ for all smooth $f : \mathbb{R}^n \to \mathbb{R}^m$.

Finally we observe that the Hausdorff dimension $D_H(\mu)$ of a measure μ is defined in two different ways in the literature, but either definition leads to preservation under typical projections. Some authors define $D_H(\mu)$ to be the infimum over all sets S with $\mu(S) = 1$ of the Hausdorff dimension of S, while others require only that $\mu(S) > 0$. By the results of Cutler [C1] on the distribution of pointwise dimensions, it follows that the first definition corresponds to the essential supremum of the lower pointwise dimension $\alpha_\mu^-(x)$ while the second definition corresponds to the essential infimum. In either case, theorem 4.1 implies that $D_H(\mu)$ is preserved under typical projections (or prevalent smooth functions).

5. Examples of nonpreservation

5.1. Nonpreservation for $0 \leqslant q < 1$

For $0 \leqslant q < 1$ we exhibit a probability measure μ supported on a compact set $Q \subset \mathbb{R}^n$ such that for every C^1 function $f : \mathbb{R}^n \to \mathbb{R}^m$ with $n > m$, we have $D_q(f(\mu)) < D_q(\mu)$. The set Q is the same as was described in [FH, SY] as an example for which box-counting (D_0) and packing dimensions are not preserved under any projection. Fix $\alpha > 0$ and let Q consist of the origin together with sets Q_k for $k = 1, 2, \ldots$, where each Q_k is a finite set of points chosen from the sphere centred at the origin with radius $k^{-\alpha}$. To form Q_k choose as many points as possible on the sphere subject to the restriction that every pair of points in Q_k be at least $\alpha k^{-1-\alpha}$ apart; the number of points chosen can be bounded between $c_1 k^{n-1}$ and $c_2 k^{n-1}$ where $c_1 < c_2$ depend only on n and α. Notice that $\alpha k^{-1-\alpha}$ is also a lower bound on the spacing between the spheres of radius $j^{-\alpha}$ for $j = 1, 2, \ldots, k$; thus all points of $Q_1 \cup Q_2 \cup \cdots \cup Q_k$ are at least $\alpha k^{-1-\alpha}$ apart.

Next fix $\beta > 0$ and define μ by assigning a mass $M k^{-1-\beta}$ to each Q_k, where

$$M = \left(\sum_{k=1}^{\infty} k^{-1-\beta} \right)^{-1}$$

and distributing the mass within Q_k equally to each point. Let $\varepsilon_k = \alpha k^{-1-\alpha}$. Since all points of $Q_1 \cup Q_2 \cup \cdots \cup Q_k$ are at least ε_k apart, we have for $0 \leqslant q < 1$ that

$$C_q(\mu, \varepsilon_k) \geqslant \sum_{j=1}^{k} M j^{-1-\beta} \left(\frac{M j^{-1-\beta}}{c_1 j^{n-1}} \right)^{q-1} = M^q c_1^{1-q} \sum_{j=1}^{k} j^{n(1-q)-\beta q-1}.$$

For a given $q \in [0, 1)$ we can choose β small enough that $n(1-q) - \beta q > 0$. Then it follows that

$$C_q(\mu, \varepsilon_k) \geqslant c_3 k^{n(1-q)-\beta q}$$

where c_3 depends only on n, q, α, and β. Consequently

$$D_q^-(\mu) = \liminf_{k\to\infty} \frac{\log C_q(\mu, \varepsilon_k)}{(q-1)\log \varepsilon_k} \geqslant \frac{n(1-q)-\beta q}{(1+\alpha)(1-q)}.$$

(In fact one can show that $D_q(\mu)$ exists and equals the right side of the above equation.)

Now let f be a C^1 function from $\mathbb{R}^n$ to $\mathbb{R}^m$, with $n > m$, and let λ be a Lipschitz constant for f on the closed unit ball $B(0, 1)$. Our goal is to obtain an upper bound on $D_q^+(f(\mu))$, for which we need an upper bound on $C_q(f(\mu), \delta)$ for small δ. Let $\delta_k = k^{-n/m-\alpha}$, and let μ_k be the restriction of μ to $Q_1 \cup Q_2 \cup \cdots \cup Q_k$. It is not hard to show that

$$C_q(f(\mu), \delta_k) \leqslant C_q(f(\mu_k), \delta_k) + C_q(f(\mu - \mu_k), \delta_k). \tag{5.1}$$

To estimate the first term on the right side of (5.1), notice that $f(\mu_k)$ is supported on the finite set $f(Q_1 \cup Q_2 \cup \cdots \cup Q_k)$. Though we have no guarantee that these points are at least δ_k of each other, since $0 \leqslant q < 1$ it follows from the definition of C_q that the widest possible distribution of the masses making up $f(\mu_k)$ yields the greatest possible value of $C_q(f(\mu_k), \delta_k)$. That is,

$$\begin{aligned} C_q(f(\mu_k), \delta_k) &\leqslant \sum_{j=1}^{k} \sum_{x\in Q_j} [\mu(x)]^{q-1}\mu(x) \\ &\leqslant \sum_{j=1}^{k} M j^{-1-\beta} \left(\frac{M j^{-1-\beta}}{c_2 j^{n-1}} \right)^{q-1} = M^q c_2^{1-q} \sum_{j=1}^{k} j^{n(1-q)-\beta q-1}. \end{aligned}$$

As before,

$$C_q(f(\mu_k), \delta_k) \leqslant c_4 k^{n(1-q)-\beta q}$$

where c_4 depends only on n, q, α, and β.

Next, to estimate the second term on the right side of (5.1), we observe that the support of $f(\mu - \mu_k)$ is contained in $B(0, \lambda k^{-\alpha})$, where we recall that λ is the Lipschitz constant of f. Again we reason that $C_q(f(\mu - \mu_k), \delta_k)$ is largest when the measure is spread as evenly as possible throughout $B(0, \lambda k^{-\alpha})$; that is, if we take

$$\begin{aligned} f(\mu - \mu_k)(B(x, \delta_k)) &= \left(\frac{\delta_k}{\lambda k^{-\alpha}} \right)^m f(\mu - \mu_k)(B(0, \lambda k^{-\alpha})) \\ &= \left(\frac{k^{-n/m}}{\lambda} \right)^m \sum_{j=k}^{\infty} M j^{-1-\beta} \leqslant \lambda^{-m} k^{-n} c_5 k^{-\beta} \end{aligned}$$

where c_5 depends only on β. Hence

$$C_q(f(\mu - \mu_k), \delta_k) \leqslant (\lambda^{-m} k^{-n} c_5 k^{-\beta})^{q-1} c_5 k^{-\beta} = c_6 k^{n(1-q)-\beta q}$$

where c_6 depends only on m, q, β, and λ.

Using the results of the previous two paragraphs we conclude that

$$D_q^+(f(\mu)) = \limsup_{k\to\infty} \frac{\log C_q(\mu, \delta_k)}{(q-1)\log \delta_k} \leqslant \frac{n(1-q) - \beta q}{(n/m+\alpha)(1-q)}.$$

In particular, $D_q^+(f(\mu)) < D_q^-(\mu)$ for all C^1 functions $f : \mathbb{R}^n \to \mathbb{R}^m$.

The above example seems pathological from the point of view of dynamical systems. Indeed since box-counting dimension and Hausdorff dimension are expected to coincide for attractors, and since Hausdorff dimension is preserved under typical projections, it seems reasonable to expect that when μ is the natural invariant measure of an attractor that $D_q(\mu)$ is preserved under typical projections for $0 \leqslant q < 1$ as well as for $1 \leqslant q \leqslant 2$.

5.2. *Nonpreservation for $q > 2$*

Next we consider the case $q > 2$. In the introduction we stated that if μ is the uniform probability measure on the unit circle in $\mathbb{R}^2$, then for every orthogonal projection $P : \mathbb{R}^2 \to \mathbb{R}$ we have for all $q > 2$ that $D_q(P(\mu)) < D_q(\mu) = 1$. To see this, choose $\varepsilon \in (0, 1)$ and consider the contribution to $C_q(P(\mu), \varepsilon)$ from a neighbourhood of an endpoint $x = 1$ of the support of $P(\mu)$. Notice that for $x \in [1-\varepsilon, 1]$, we have $P(\mu)(B(x, \varepsilon)) \geqslant \mu([1-\varepsilon, 1]) \geqslant \sqrt{2\varepsilon}/\pi$. It follows that $C_q(P(\mu), \varepsilon) \geqslant (\sqrt{2\varepsilon}/\pi)^q$, and thus

$$D_q(P(\mu)) = \lim_{\varepsilon\to 0} \frac{\log C_q(P(\mu), \varepsilon)}{(q-1)\log \varepsilon} \leqslant \frac{q}{2(q-1)}.$$

(Strictly speaking, we have not proved that the limit exists, but it is not hard to show in this case.) In particular, for $q > 2$ the right side of the inequality is less than 1. The fact that $D_q(\mu) = 1$ is easy to check.

The above argument can easily be extended to replace the circle by an arbitrary C^2 closed curve and P by an arbitrary C^2 function. The example can also be generalized to consider the projection of the uniform probability measure on a generic d-dimensional compact manifold in $\mathbb{R}^n$ into $\mathbb{R}^m$. In this case D_q is decreased for q sufficiently large when $d \leqslant m < 2d$; how large q must be depends on n, m, and d. On the other hand, if $m > 2d$ then almost every projection is an embedding (that is, a diffeomorphism from the manifold to its image) [W, SYC] and in particular D_q is preserved for all $q \geqslant 0$. (In fact, D_q is preserved also for $m = 2d$, though this is not a direct consequence of the embedding theorem.)

Based on embedding results, one is tempted to think that the situation is similar for arbitrary Borel probability measures μ on $\mathbb{R}^n$. For instance, one might conjecture that if $m \geqslant 2D_0(\mu)$, then $D_q(\mu)$ is preserved by typical projections for all $q > 1$. However, this is not the case, even if 2 is replaced by a larger constant. To illustrate we give an example of a probability measure μ supported on a C^1 curve in $\mathbb{R}^n$ such that every projection P into $\mathbb{R}^{n-1}$, we have $D_q(P(\mu)) < D_q(\mu) = 1$ for $q > n$. This example is a generalization of the previous example, which represents $n = 2$.

Let $g : [0, 1] \to \mathbb{R}^n$ be a continuous function that maps the unit interval onto the unit sphere in $\mathbb{R}^n$. Since the sphere is $(n-1)$-dimensional, it is possible to construct g to be Hölder continuous with exponent $1/(n-1)$; that is, for some constant K we have

$$|g(t) - g(t')| \leqslant K|t - t'|^{1/(n-1)}$$

for all $t, t' \in [0, 1]$. Let $h(t) = \int_0^t g(t')\,\mathrm{d}t'$, and let μ be the image under h of the uniform measure on $[0, 1]$. Since $h'(t) = g(t)$ has magnitude 1 for all t, it is not hard to show that $D_q(\mu) = 1$ for all $q \geqslant 0$. Now consider a linear transformation L into $\mathbb{R}^{n-1}$; some point on the unit sphere in $\mathbb{R}^n$ must be annihilated by L, and thus there exists $t_0 \in [0, 1]$ such that $L(g(t_0)) = 0$. It follows that

$$|L(h(t)) - L(h(t_0))| = \left| \int_{t_0}^{t} L(g(t'))\,\mathrm{d}t' \right| \leqslant \int_{t_0}^{t} \|L\| K |t' - t_0|^{1/(n-1)}\,\mathrm{d}t'$$

$$= \|L\| K \frac{n-1}{n} |t - t_0|^{n/(n-1)}.$$

Then for ε sufficiently small,

$$L(\mu)(B(L(h(t_0)), \varepsilon)) \geqslant C\varepsilon^{(n-1)/n}$$

where C depends only on L, K, and n. As in the example above, we can then argue that

$$D_q(L(\mu)) \leqslant \frac{(n-1)q}{n(q-1)}$$

and in particular $D_q(L(\mu)) < 1$ for $q > n$.

What sort of positive result can be given for $q > 2$? We can show for integer $k \geqslant 2$ that $D_q(\mu)$ is preserved by typical projections onto $\mathbb{R}^m$ for $1 < q \leqslant k$ provided $m \geqslant (k-1)D_q(\mu)$, and the previous example shows that this bound is sharp. We do not know precisely the best possible result for non-integer q.

Acknowledgments

We thank J A Yorke for fruitful discussions. VYuK is grateful to the Institute for Physical Science and Technology, University of Maryland for its hospitality. BRH was supported by the 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). VYuK was supported in part by the International Science Foundation and the Russian Government grant No M98300, a grant from INTAS and by the Russian Fund of Fundamental Investigation grant No 95-01-00229.

References

[AY] Alexander J C and Yorke J A 1984 Fat baker's transformations *Ergod. Theor. Dynam. Syst.* **4** 1–23

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

[BaR] Balatoni J and Rényi A 1956 Remarks on entropy *Publ. Math. Inst. Hung. Acad. Sci.* **1** 9–40 (in Hungarian with English summary)

Balatoni J and Rényi A 1976 *Selected Papers of Alfréd Rényi* vol 1 Budapest: Akadémiai Kiadó) pp 558–86 (English translation)

[BPS] Barreira L, Pesin Ya and Schmeling J 1996 On the pointwise dimension of hyperbolic measures: a proof of the Eckmann–Ruelle conjecture *Elec. Res. Ann. Am. Math. Soc.* **2** 69–72

Barreira L, Pesin Ya and Schmeling J Dimension of hyperbolic measures: a proof of the Eckmann–Ruelle conjecture *Preprint*

[B] Beck C 1990 Upper and lower bounds on the Renyi dimensions and the uniformity of multifractals *Physica* D **41** 67–78

[BoR] Bowen R and Ruelle D 1975 The ergodic theory of Axiom-A flows *Inv. Math.* **29** 181–202

[BMP] Brown G, Michon G and Peyrière J 1992 On the multifractal analysis of measures *J. Stat. Phys.* **66** 775–90

[CM] Cawley R and Mauldin D 1992 Multifractal decomposition of Moran fractals *Adv. Math.* **92** 196–236

[CLP] Collet P, Lebowitz J L and Porzio A The dimension spectrum of some dynamical systems *J. Stat. Phys.* **47** 609–44

[C1] Cutler C D 1991 Some results on the behavior and estimation of the fractal dimension of distributions on attractors *J. Stat. Phys.* **62** 651–708

[C2] Cutler C D 1995 Strong and weak duality principles for fractal dimension in Euclidean space *Math. Proc. Camb. Phil. Soc.* **118** 393–410

[DGOSY] Ding M, Grebogi C, Ott E, Sauer T and Yorke J A 1993 Estimating correlation dimension from a chaotic time series: when does plateau onset occur? *Physica* D **69** 404–24

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

[Fal] 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

[FM] Falconer K and Mattila P 1996 The packing dimension of projections and sections of measures *Math. Proc. Camb. Phil. Soc.* **119** 695–713

[Far] Farmer J D 1982 Information dimension and the probabilistic structure of chaos *Z. Naturforsch.* a **37** 1304–25

[FOY] Farmer J D, Ott E and Yorke J A 1983 The dimension of chaotic attractors *Physica* D **7** 153–80

[FKYY] Fredrickson P, Kaplan J, Yorke E D and Yorke J A 1983 The Lyapunov dimension of strange attractors *J. Diff. Eqns* **49** 185–207

[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

[Ga] Gács P 1973 Hausdorff dimension and probability distributions *Periodica Mathematica Hungarica* **3** 59–71

[Gr] Grassberger P 1983 Generalized dimension of strange attractors *Phys. Lett.* **97A** 227–31

[GP1] Grassberger P and Procaccia I 1983 Characterization of strange attractors *Phys. Rev. Lett.* **50** 346–9

[GP2] Grassberger P and Procaccia I Measuring the strangeness of strange attractors *Physica* D **7** 189–208

[HJKPS] Halsey T C, Jensen M H, Kadanoff L P, Procaccia I and Shraiman B I 1986 Fractal measures and their singularities: the characterization of strange sets *Phys. Rev.* A **33** 1141–51

[HP] Hentschel H G and Procaccia I 1983 The infinite number of generalized dimensions of fractals and strange attractors *Physica* D **8** 435–44

[HT] Hu X and Taylor S J 1994 Fractal properties of products and projections of measures in $\mathbb{R}^d$ *Math. Proc. Camb. Phil. Soc.* **115** 527–44

[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. Fen. A1 Math. Diss.* **99** 1–34

[Kah] Kahane J-P 1985 Ensembles aleatoires et dimensions *Recent Progress in Fourier Analysis (North-Holland Mathematics Studies 111)* ed I Peral and J-L Rubio de Francia (Amsterdam: North-Holland) pp 65–121

[KY] Kaplan J L and Yorke J A 1979 Chaotic behavior of multidimensional difference equations *Functional Differential Equations and Approximation of Fixed Points (Lecture Notes in Mathematics 730)* ed H-O Peitgen and H-O Walter (New York: Springer) pp 204–27

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

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

[L] Lopes A O 1990 Dimension spectra and a mathematical model for phase transition *Adv. Appl. Math.* **11** 475–502

[Mar] Marstrand J M 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. Fennicae* A **1** 227–44

[Mat2] Mattila P 1981 Integralgeometric properties of capacities *Trans. Am. Math. Soc.* **266** 539–54

[Mat3] Mattila P 1995 *Geometry of Sets and Measures in Euclidean Spaces* (Cambridge: Cambridge University Press)

[OWY] Ott E, Withers W D and Yorke J A 1984 Is the dimension of chaotic attractors invariant under coordinate changes? *J. Stat. Phys.* **36** 687–97

[Ol1] Olsen L 1995 A multifractal formalism *Adv. Math.* **116** 82–196

[Ol2] Olsen L 1996 Multifractal dimensions of product measures *Math. Proc. Camb. Phil. Soc.* **120** 709–34

[Ox] Oxtoby J C 1971 *Measure and Category* (New York: Springer)

[Pe] Pesin Ya 1993 On rigorous mathematical definitions of correlation dimension and generalized spectrum for dimensions *J. Stat. Phys.* **71** 529–47

[PW] Pesin Ya and Weiss H 1997 The multifractal analysis of Gibbs measures: motivation, mathematical foundation, and examples *Chaos* **7** 89–106

[Po] Porzio A 1990 The dimension spectrum of Axiom A attractors *J. Stat. Phys.* **58** 923–37

[Ra] Rand D 1989 The singularity spectrum $f(\alpha)$ for cookie-cutters *Ergod. Theor. Dynam. Syst.* **9** 527–41

[Re] Rényi A 1970 *Probability Theory* (Amsterdam: North-Holland)

[Ri] Riedi R H 1995 An improved multifractal formalism and self-affine measures *J. Math. Anal. Appl.* **189** 462–90

[SY] Sauer T and Yorke J A Are the dimensions of a set and its image equal under typical smooth functions? *Ergod. Theor. Dynam. Syst.* at press

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

[Sim] Simpelaere D 1994 Dimension spectrum of Axiom A diffeomorphisms. II. Gibbs measures *J. Stat. Phys.* **76** 1359–75

[Sin] Sinai Ya G 1972 Gibbs measures in ergodic theory *Russ. Math. Surv.* **27** 21–70

[St] Strichartz R S 1993 Self-similar measures and their Fourier transforms III *Indiana Univ. Math. J.* **42** 367–411

[W] Whitney H 1936 Differentiable manifolds *Ann. Math.* **37** 645–80

[Y] Young L-S 1982 Dimension, entropy and Lyapunov exponents *Ergod. Theor. Dynam. Syst.* **2** 109–24