Ergod. Th. & Dynam. Sys. (1990), **10**, 451–462
Printed in Great Britain

Connecting ergodicity and dimension in dynamical systems

C. D. CUTLER†

Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

(*Received* 30 *October* 1988 *and revised* 10 *January* 1989)

Abstract. In this paper we make precise the relationship between local or pointwise dimension and the dimension structure of Borel probability measures on metric spaces. Sufficient conditions for exact-dimensionality of the stationary ergodic distributions associated with a dynamical system are obtained. A counterexample is provided to show that ergodicity alone is not sufficient to guarantee exact-dimensionality even in the case of continuous maps or flows.

1. *Introduction and statement of results*

1.1. *Introduction*

In the study of chaos and dynamical systems there has been much interest in measure-dependent definitions of dimension. Of particular interest is the notion of local or pointwise dimension

$$\hat{\alpha}(x) = \lim_{\varepsilon \to 0^+} \log \mu(B(x, \varepsilon))/\log \varepsilon$$

where μ is an invariant occupation measure over the attractor and $B(x, \varepsilon)$ is the closed ball of radius ε centred at x. The question of existence of the pointwise limit and its possible relationships with other dimensions and quantities associated with dynamical systems (such as information dimension and Lyapunov exponents) have been discussed extensively in the literature. Young (1982) has shown that if the pointwise limit exists and is μ-a.s. equal to some constant α then α is in fact the information dimension of the attractor. (Here information dimension is taken to mean the infimum of the Hausdorff dimensions of all supporting sets for μ.) Various authors assume or conjecture the existence of a constant pointwise limit in the systems under study (see, for example, Farmer *et al.*, 1983 and Guckenheimer, 1984). Young (1982) has shown that in the case of C^2 diffeomorphisms of surfaces a constant pointwise limit does in fact exist and is related by formula to the Lyapunov exponents of the map.

In this paper we first construct a general framework describing the relationship between measures on metric spaces, Hausdorff dimension, and pointwise or local dimension. This framework is an extension to general metric spaces of results

† Research supported by the Natural Sciences and Engineering Research Council of Canada.

obtained by Cutler (1986) and Cutler and Dawson (1988) for measures on finite-dimensional Euclidean space. The key point here is that the pointwise or local dimension map completely describes the distribution of μ-mass with respect to Hausdorff dimension. These results are then used to examine the relationship between ergodicity and exact-dimensionality of measures. (A measure μ will be said to be *exact-dimensional* if its local dimension map is μ-a.s. constant.) Eckmann and Ruelle (1985) in their excellent review article on ergodic theory and chaos incorrectly suggest that ergodicity plus μ-a.s. existence of the pointwise limit is sufficient to ensure that μ is exact-dimensional. We provide a counterexample to this (in § 4) but establish sufficient conditions for ergodicity to imply exact-dimensionality. These conditions are met by any ergodic differentiable dynamical system.

1.2. *Statement of main results*

Let μ be a probability measure on the Borel sets of a metric space $\mathbf{X}$. Then there exists an associated distribution $\hat{\mu}$ on the extended real line $[0, \infty]$ which describes the manner in which μ-mass is distributed with respect to Hausdorff dimension. (If $\hat{\mu}$ consists of a unit mass at the value α then μ has no mass on sets of dimension less than α but can be supported on some set of dimension α.) Furthermore if $x \in \mathbf{X}$ is selected randomly according to μ then the nonnegative extended real-valued function defined on $\mathbf{X}$ by

$$\hat{\alpha}(x) = \liminf_{\varepsilon \to 0^+} \log \mu(B(x, \varepsilon))/\log \varepsilon \tag{1.2.1}$$

has distribution $\hat{\mu}$ on $[0, \infty]$. (From this point on we refer to $\hat{\alpha}$ as defined in (1.2.1) as the *local dimension map* of μ. It is easy to see that $\hat{\alpha}$ is a measurable function since $\{x \mid \mu(B(x, \varepsilon)) \geq c\}$ is a closed subset of $\mathbf{X}$ for each c.) Thus $\hat{\mu}$ consists of a unit mass at α if and only if $\hat{\alpha} = \alpha$ μ-a.s.

If μ is stationary and ergodic with respect to a mapping $T: \mathbf{X} \to \mathbf{X}$ then μ will be exact-dimensional provided T satisfies a certain weak local Lipschitz condition at μ-almost all x. This condition is stronger than continuity (but weaker than differentiability in spaces where such a notion is possible). An analogous result holds for continuous-time dynamical systems. Moreover, if a continuous-time system arises as the family of solutions to a differential equation $dx/dt = f(x)$ then the condition for exact-dimensionality is met whenever $f(x)$ satisfies the standard conditions required for existence and uniqueness of solutions.

2. *Dimension distributions and the local dimension map*

In Cutler (1986) the notion of the dimension distribution $\hat{\mu}$ associated with a measure μ on N-dimensional Euclidean space was introduced. Here we extend this idea, with the necessary minor modifications, to general metric spaces.

Let $\mathbf{X}$ be a metric space with metric d. If $E \subseteq \mathbf{X}$, $0 \leq \alpha < \infty$, and $\delta > 0$, then the *α, δ-outer Hausdorff measure* of E is defined by

$$H^{\alpha}_{\delta}(E) = \inf_{\substack{E \subseteq US_i \\ d(S_i) \leq \delta}} \sum_{i=1}^{\infty} d(S_i)^{\alpha}, \tag{2.1}$$

where $d(S) = \sup \{d(x, y) \mid x, y \in S\}$ is the diameter of S and the infimum in (2.1) is

https://doi.org/10.1017/S014338570000568X Published online by Cambridge University Press

taken over all countable coverings $\{S_i\}_i$ of E by subsets of $\mathbf{X}$ satisfying $d(S_i) \le \delta$ for all i. (Such a covering is called a *δ-covering* of E.) If no δ-covering of E exists (which may occur in a nonseparable space) then $H_\delta^\alpha(E) = \infty$.

The *Hausdorff α-outer measure* of E is then defined to be

$$H^\alpha(E) = \lim_{\delta \to 0^+} H_\delta^\alpha(E). \tag{2.2}$$

For fixed E, $H^\alpha(E)$ is a decreasing function of α and corresponding to each set E there exists a unique point α_0, $0 \le \alpha_0 \le \infty$, such that

$$\alpha_0 = \sup\{0 \le \alpha < \infty \mid H^\alpha(E) = \infty\} = \inf\{0 \le \alpha < \infty \mid H^\alpha(E) = 0\}. \tag{2.3}$$

α_0 is called the *Hausdorff dimension of E* and denoted dim (E). A detailed discussion of Hausdorff measures and dimension can be found in Rogers (1970).

By the Borel sets of $\mathbf{X}$ we will mean the σ-field generated by the open sets of $\mathbf{X}$.

THEOREM 2.1. *Let μ be a probability measure defined on the Borel sets of* $\mathbf{X}$. *Then there exists a unique probability distribution $\hat{\mu}$ on the Borel sets of the extended real line* $[0, \infty]$ *such that* $\hat{\mu}([0, \alpha]) = \sup\{\mu(E) \mid E \subseteq \mathbf{X},\ E \text{ Borel},\ \dim(E) \le \alpha\}$ *for each* $0 \le \alpha \le \infty$. *Thus $\hat{\mu}([0, \alpha])$ is the proportion of mass of μ concentrated on sets of dimension not exceeding α. We call $\hat{\mu}$ the* dimension distribution *of μ.*

Proof. The construction of μ follows the procedure in § 2 of Cutler (1986) (see pp. 1460–1462). The only change is that we must allow for the possibility $\hat{\mu}(\{\infty\}) > 0$ in metric spaces of infinite Hausdorff dimension. □

In Cutler and Dawson (1989) it was shown that when $\mathbf{X} = \mathbb{R}^N$ the local dimension map $\hat{\alpha}$ as defined in (1.2.1) satisfies

$$\dim(D_\alpha) \le \alpha \quad \text{where} \quad D_\alpha = \{x \mid \hat{\alpha}(x) \le \alpha\} \quad \text{and} \tag{2.4}$$

$$\mu(D_\alpha) \ge \hat{\mu}([0, \alpha]) \quad \text{for each } 0 \le \alpha \le \infty. \tag{2.5}$$

As a consequence of (2.4) and (2.5) we conclude $\mu(D_\alpha) = \hat{\mu}([0, \alpha])$ and hence $\hat{\alpha}(x)$ has distribution $\hat{\mu}$ when x has distribution μ. However the proofs of (2.4) and (2.5) provided for the Euclidean case do not extend to more general metric spaces. The key to the more general situation is provided by a Vitali covering theorem for Hausdorff measures. If $E \subseteq \mathbf{X}$, a Vitali covering of E will be a collection $\mathcal{V}$ of closed subsets of $\mathbf{X}$ such that for each $x \in E$ and each $\varepsilon > 0$ there exists $V \in \mathcal{V}$ with $x \in V$ and $0 < d(V) \le \varepsilon$.

THEOREM 2.2. (Vitali Covering Theorem): *Let $E \subseteq \mathbf{X}$ and suppose $\mathcal{V}$ is a Vitali covering of E. Then there exists a finite or countable disjoint collection* $\{V_i)_i$ *of members of $\mathcal{V}$ which, for given $\alpha > 0$, satisfies either*

(i) $\sum_i d(V_i)^\alpha = \infty$ or

(ii) $H^\alpha\left(E \Big\backslash \bigcup_i V_i\right) = 0.$

Proof. The proof is that of Theorem 1.10(a) of Falconer (1985). Note that Falconer states the result for $\mathbb{R}^N$ and under a measurability restriction on E but the proof does not require any further conditions than those we have stated here. □

LEMMA 2.1. *Let μ be a probability measure on the Borel sets of* $\mathbf{X}$. *Then* (2.4) *holds. That is,* $\dim(D_\alpha) \le \alpha$ *for each* $0 \le \alpha \le \infty$.

Proof. The case $\alpha = \infty$ is immediate. Suppose $0 \le \alpha < \infty$ and let $\beta > \alpha$ be arbitrary and finite. For each $x \in D_\alpha$ and each positive integer n there exists a family $\mathscr{B}_{x,n}$ of closed balls centred at x or arbitrarily small diameter (and maximum diameter not exceeding $1/n$) such that $\log \mu(B(x, \varepsilon))/\log \varepsilon \le \beta$ for each $B(x, \varepsilon) \in \mathscr{B}_{x,n}$. (This gives $\mu(B(x, \varepsilon)) \ge \varepsilon^\beta \ge 2^{-\beta} d(B(x, \varepsilon))^\beta$ since $d(B(x, \varepsilon)) \le 2\varepsilon$.) The collection $\mathscr{V}_n = \bigcup_{x \in D_\alpha} \mathscr{B}_{x,n}$ is then a Vitali covering of D_α and by Theorem 2.2 there exists a disjoint sequence $\{V_{n,i}\}_i$ from $\mathscr{V}_n$ such that either: (i) $\sum_i d(V_{n,i})^\beta = \infty$ or (ii) $H^\beta(D_\alpha \backslash \bigcup_i V_{n,i}) = 0$. But since the $V_{n,i}$'s are disjoint and satisfy $d(V_{n,i})^\beta \le 2^\beta \mu(V_{n,i})$ (by construction of $\mathscr{V}_n$) and μ is a probability measure, we obtain

$$\sum_i d(V_{n,i})^\beta \le 2^\beta \sum_i \mu(V_{n,i}) \le 2^\beta < \infty.$$

Hence (ii) must hold. Now setting $V = \bigcap_n \bigcup_i V_{n,i}$ we obtain

$$H^\beta(D_\alpha \backslash V) = H^\beta\left(\bigcup_n \left(D_\alpha \Big\backslash \bigcup_i V_{n,i}\right)\right) \le \sum_n H^\beta\left(D_\alpha \Big\backslash \bigcup_i V_{n,i}\right) = 0$$

by the countable subadditivity of H^β. Writing $D_\alpha = (D_\alpha \backslash V) \cup (D_\alpha \cap V)$ we conclude $H^\beta(D_\alpha) = H^\beta(D_\alpha \cap V)$. But for each n $\{V_{n,i}\}_i$ is a $1/n$-covering of $D_\alpha \cap V$ satisfying $\sum_i d(V_{n,i})^\beta \le 2^\beta$ and hence $H^\beta(D_\alpha \cap V) \le 2^\beta < \infty$. Thus $\dim(D_\alpha) \le \beta$. As $\beta > \alpha$ was aribtrary we get $\dim(D_\alpha) \le \alpha$ as claimed. □

To show that (2.5) holds we will need the following:

LEMMA 2.2. *Let μ be a probability measure on the Borel sets of* $\mathbf{X}$. *Let* $0 \le \alpha \le \infty$. *Suppose* $E \subseteq D_\alpha^*$ *where* $D_\alpha^* = \{x \mid \hat{\alpha}(x) \ge \alpha\}$. *If* $\mu(E) > 0$ *then* $\dim(E) \ge \alpha$.

Proof. The case $\alpha = 0$ is immediate so suppose $\alpha > 0$ and let $0 < \beta < \alpha$ be arbitrary. Then for each $x \in E$ there exists $\varepsilon(x) > 0$ such that $0 < \varepsilon \le \varepsilon(x)$ implies $\log \mu(B(x, \varepsilon))/\log \varepsilon \ge \beta$. Hence we can write $E = \bigcup_{n=1}^\infty E_n$ where $E_n = \{x \in E \mid \varepsilon(x) \ge 1/n\}$. Noting that $E_n \subseteq E_{n+1}$the assumption $\mu(E) > 0$ then implies there exists $\theta > 0$ such that $\mu(E_n) > \theta$ for all sufficiently large n. Now fix n large and let $\{S_i\}_i$ be any $1/n$-covering of E. (If no such covering exists then $\dim(E) = \infty$ and the theorem is trivially true.) If $S_i \cap E_n \ne \phi$ choose $x_i \in S_i \cap E_n$ and define the closed ball $B_i = B(x_i, d(S_i))$. Note that $S_i \subseteq B_i$ and hence the collection of balls $\{B_i\}$ must cover E_n. Now since $x_i \in E_n$ and $d(S_i) \le 1/n$ we obtain $\log \mu(B_i) \le \beta \log d(S_i)$ and therefore $\mu(B_i) \le d(S_i)^\beta$. This gives the chain of inequalities

$$0 < \theta < \mu(E_n) \le \sum_{x_i} \mu(B_i) \le \sum_{x_i} d(S_i)^\beta \le \sum_{i=1}^\infty d(S_i)^\beta,$$

which shows that $H^\beta_{1/n}(E) > \theta$. Thus $H^\beta(E) > \theta$ and so $\dim(E) \ge \alpha$. □

Remark. A proof of Lemma 2.2 in the case $\mathbf{X} = \mathbb{R}^N$ has been given by Young (1982). Related problems have also been studied by Billingsley (1960, 1961).

THEOREM 2.3. *Let μ be a probability measure on the Borel sets of* $\mathbf{X}$. *Then both* (2.4) *and* (2.5) *hold and hence the local dimension map* $\hat{\alpha}(x)$ *defined in* (1.2.1) *has distribution* $\hat{\mu}$ *when* x *has distribution* μ.

https://doi.org/10.1017/S014338570000568X Published online by Cambridge University Press

Proof. From Lemma 2.1 we know (2.4) holds. To show (2.5) holds, let E be any Borel set of $\mathbf{X}$ satisfying $\dim(E) \le \alpha$. As $\hat{\mu}([0, \alpha])$ is the supremum of the μ-measures of such sets E, (2.5) will be proved if we show $\mu(E) \le \mu(D_\alpha)$. If $\alpha = \infty$ then $D_\alpha = \mathbf{X}$ so there is nothing to prove in this case. Therefore we assume $0 \le \alpha < \infty$. Now we can write $E = (E \cap D_\alpha) \cup (E \backslash D_\alpha)$ and for each $\beta > \alpha$ we have $(E \backslash D_\beta) \subseteq D_\beta^*$. Since $\dim(E \backslash D_\beta) \le \dim(E) \le \alpha < \beta$ it follows from Lemma 2.2 that $\mu(E \backslash D_\beta) = 0$. But $(E \backslash D_\beta) \uparrow (E \backslash D_\alpha)$ as $\beta \downarrow \alpha$ and hence we must also have $\mu(E \backslash D_\alpha) = 0$. Thus $\mu(E) = \mu(E \cap D_\alpha) \le \mu(D_\alpha)$ as required. □

3. *Sufficient conditions for exact-dimensionality in the ergodic case*

The following lemma expresses the notion of continuity in a way which will allow for easy comparison with the sufficient conditions of Theorems 3.1 and 3.2. As before $\mathbf{X}$ is a metric space with metric d.

LEMMA 3.1. *A function $T: \mathbf{X} \to \mathbf{X}$ is continuous at a point x if and only if there exists $R(\varepsilon) \ge 0$ such that $\varepsilon R(\varepsilon) \downarrow 0$ as $\varepsilon \downarrow 0$ and $d(T(x), T(y)) \le R(d(x, y))\, d(x, y)$ for all y in some neighbourhood of x.*

Proof. Clearly the stated conditions imply continuity at x. To show the converse, if T is continuous at x set

$$R(\varepsilon) = (1/\varepsilon) \sup \{d(T(x), T(y)) \mid y \in \mathbf{X}, d(x, y) \le \varepsilon\}.$$ □

We now consider the case of a discrete dynamical system.

THEOREM 3.1. *Let the probability measure μ be stationary and ergodic with respect to a measurable mapping $T: \mathbf{X} \to \mathbf{X}$. Suppose T satisfies the following condition*:

(3.1) *There exists $R(x, \varepsilon) \ge 0$ such that, for μ-almost all x,*
- (a) $d(T(x), T(y)) \le R(x, d(x, y))\, d(x, y)$ *for all y in some neighbourhood of x,*
- (b) $\varepsilon R(x, \varepsilon) \downarrow 0$ *as* $\varepsilon \downarrow 0$, *and*
- (c) $\lim_{\varepsilon \to 0^+} \log \varepsilon R(x, \varepsilon) / \log \varepsilon = 1$.

Then μ is exact-dimensional.

Proof. First note that, in view of Lemma 3.1, condition (3.1) is stronger than μ-a.s. continuity of T. Also note that without loss of generality we can assume $R(x, \varepsilon)$ is uniformly bounded below by some constant $\gamma > 0$ (otherwise let $\tilde{R}(x, \varepsilon) = \max(\gamma, R(x, \varepsilon))$ and note that (3.1.(a), (b), (c)) continue to hold with $\tilde{R}(x, \varepsilon)$ in place of $R(x, \varepsilon)$). This shows that (3.1.(c)) is simply a bound on the allowable rate of increase of $R(x, \varepsilon)$ as ε shrinks. In particular, (3.1) is weaker than requiring T to be locally Lipschitz. (We say T is *locally Lipschitz at* x if there exists a finite constant $R(x)$ such that $d(T(x), T(y)) \le R(x)\, d(x, y)$ for all y in some neighbourhood of x.) If $\mathbf{X}$ is a normed linear space and T is differentiable at x it is well known that T is also locally Lipschitz at x.

To prove the theorem we will show that condition (3.1) implies that $\hat{\alpha}$ is an invariant function. If $E \subseteq \mathbf{X}$ then let $I(E, y)$ denote the indicator function of E. From ergodicity we can write, for μ-almost all y:

$$\mu(B(x, \varepsilon)) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} I(B(x, \varepsilon), T^k(y)) \tag{3.2}$$

https://doi.org/10.1017/S014338570000568X Published online by Cambridge University Press

and

$$\mu(B(T(x), \delta)) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} I(B(T(x), \delta), T^k(y)). \tag{3.3}$$

Let A denote the subset of $\mathbf{X}$ for which (3.1.(a), (b), (c)) hold. (By assumption, $\mu(A)=1$.) Then for $x \in A$ we see that if $T^k(y) \in B(x, \varepsilon)$ *then*

$$d(T(x), T^{k+1}(y)) \le R(x, d(x, T^k(y)))\, d(x, T^k(y)) \le \varepsilon R(x, \varepsilon).$$

Thus $T^k(y) \in B(x, \varepsilon)$ implies $T^{k+1}(y) \in B(T(x), \varepsilon R(x, \varepsilon))$. In view of this and (3.2) and (3.3) with $\delta = \varepsilon R(x, \varepsilon)$, we conclude $\mu(B(T(x), \varepsilon R(x, \varepsilon))) \ge \mu(B(x, \varepsilon))$. Consequently for $x \in A$ we have

$$\begin{aligned}
\hat{\alpha}(T(x)) &= \liminf_{\varepsilon\to 0^+} \log \mu(B(T(x), \varepsilon))/\log \varepsilon \\
&\le \liminf_{\varepsilon\to 0^+} \log \mu(B(T(x), \varepsilon R(x, \varepsilon)))/\log \varepsilon R(x, \varepsilon) \\
&\le \liminf_{\varepsilon\to 0^+} \log \mu(B(x, \varepsilon))/\log \varepsilon R(x, \varepsilon) \\
&= \liminf_{\varepsilon\to 0^+} (\log \mu(B(x, \varepsilon))/\log \varepsilon)(\log \varepsilon/\log \varepsilon R(x, \varepsilon)) \\
&= (\liminf_{\varepsilon\to 0^+} \log \mu(B(x, \varepsilon))/\log \varepsilon)(\lim_{\varepsilon\to 0^+} \log \varepsilon/\log \varepsilon R(x, \varepsilon)) \\
&= \hat{\alpha}(x).
\end{aligned}$$

Thus we have established

$$\hat{\alpha}(T(x)) \le \hat{\alpha}(x) \quad \mu\text{-a.s.} \tag{3.4}$$

Now since μ is stationary with respect to T the distribution of $\hat{\alpha} \circ T$ is the same as that of $\hat{\alpha}$ when initial conditions are selected according to μ. This plus the pointwise inequality (3.4) forces $\hat{\alpha} \circ T = \hat{\alpha}$ μ-a.s. Thus $\hat{\alpha}$ is invariant, and ergodicity now implies the existence of a constant α (possibly $\alpha = \infty$) such that $\hat{\alpha}(x) = \alpha$ μ-a.s. □

The case of a continuous dynamical system is similar. We will need to modify condition (3.1) slightly.

LEMMA 3.2. *Let* $\{T^t\}$, $t \ge 0$, *be a semigroup of measurable mappings* $T^t : \mathbf{X} \to \mathbf{X}$ *with stationary distribution* μ. *If the semigroup satisfies*

(3.2) *There exists* $\delta > 0$ *and* $R^t(x, \varepsilon) \ge 0$ *for each* $0 \le t < \delta$ *such that, for* μ*-almost all* x,

- (a) $d(T^t(x), T^t(y)) \le R^t(x, d(x, y))\, d(x, y)$ *for all* y *in some neighbourhood of* x,
- (b) $\varepsilon R^t(x, \varepsilon) \downarrow 0$ *as* $\varepsilon \downarrow 0$, *and*
- (c) $\lim_{\varepsilon\to 0+} \log \varepsilon R^t(x, \varepsilon)/\log \varepsilon = 1$

then in fact for each $t \ge 0$ *there exists* $R^t(x, \varepsilon)$ *such that* (a), (b), *and* (c) *of* (3.2) *hold.*

Proof. We first show that if $0 \le s, u < \delta$ then there exists $R^{u+s}(x, \varepsilon)$ satisfying (3.2.(a), (b), (c)) for μ-almost all x. Define $R^{u+s}(x, \varepsilon)$ by

$$R^{u+s}(x, \varepsilon) = R^u(T^s(x), \varepsilon R^s(x, \varepsilon)) R^s(x, \varepsilon).$$

Letting $A = \{x \mid R^s(x, \varepsilon)$ and $R^u(T^s(x), \varepsilon)$ satisfy (3.2.(a), (b), (c))$\}$ then stationarity and assumption (3.2) applied at the time points u and s imply $\mu(A)=1$. Now if $x\in A$ then for all y in a sufficiently small neighbourhood of x we obtain

$$\begin{aligned}
d(T^{u+s}(x), T^{u+s}(y)) &= d(T^u(T^s(x)), T^u(T^s(y)))\\
&\le R^u(T^s(x), d(T^s(x), T^s(y)))\, d(T^s(x), T^s(y)) \quad \text{using (3.2.(a))}\\
&\le R^u(T^s(x), R^s(x, d(x,y))\, d(x,y))R^s(x, d(x,y))\, d(x,y) \quad \text{using (3.2.(a) and (b))}\\
&= R^{u+s}(x, d(x,y))\, d(x,y) \quad \text{by definition of } R^{u+s}.
\end{aligned}$$

Thus (3.2(a)) holds for R^{u+s}. It is straightforward to verify that (3.2.(b) and (c)) also hold. To complete the proof of the lemma we note that if $t\ge\delta$ then there exists $0<s<\delta$ and a positive integer k such that $t=ks$ and so the result follows by induction on k. □

THEOREM 3.2. *Let* $\{T^t\}$, $t\ge 0$, *be a semigroup of measurable mappings* $T^t:\mathbf{X}\to\mathbf{X}$ *with stationary, ergodic distribution* μ. *If the semi-group satisfies condition* (3.2) *then* μ *is exact-dimensional.*

Proof. The proof is similar to that of discrete case with sums replaced by integrals. Let $t\ge 0$ be arbitrary. Then for μ-almost all x and all $s\ge 0$, $T^{t+s}(y)\in B(T^t(x), \varepsilon R^t(x,\varepsilon))$ whenever $T^s(y)\in B(x,\varepsilon)$. From this and ergodicity we obtain, for μ-almost all y,

$$\begin{aligned}
\mu(B(T^t(x), \varepsilon R^t(x,\varepsilon))) &= \lim_{u\to\infty}\frac{1}{u}\int_0^u I(B(T^t(x), \varepsilon R^t(x,\varepsilon)), T^s(y))\, ds\\
&= \lim_{u\to\infty}\frac{1}{u}\int_0^u I(B(T^t(x), \varepsilon R^t(x,\varepsilon)), T^{t+s}(y))\, ds\\
&\ge \lim_{u\to\infty}\frac{1}{u}\int_0^u I(B(x,\varepsilon), T^s(y))\, ds\\
&= \mu(B(x,\varepsilon)).
\end{aligned}$$

Now as in the proof of Theorem 3.1 it follows that $\hat\alpha\circ T^t=\hat\alpha$ μ-a.s. Thus $\hat\alpha$ is an invariant function and μ is exact-dimensional. □

Remark. Suppose that $dx/dt=f(x)$ is a differential equation such that for each x_0 there exists a ball $B(x_0)$ and a constant $K(x_0)$ such that

$$\|f(x)-f(y)\|\le K(x_0)\|x-y\| \quad \text{for all } x, y\in B(x_0).$$

Then it is well known that for each x there exists a unique maximal solution $u_x(t)=T^t(x)$ (with $u_x(0)=x$) and that the solution curves depend continuously on the initial conditions to the extent that

$$\|T^t(x)-T^t(y)\|\le \exp(K(x)t)\|x-y\|$$

for all y sufficiently near x. Thus the generated dynamical system satisfies condition (3.2).

https://doi.org/10.1017/S014338570000568X Published online by Cambridge University Press

4. *A counterexample*

We construct a continuous function $T:[0,1]\to[0,1]$ and stationary, ergodic distribution μ such that μ has positive mass on two sets of different dimensions and the pointwise limit

$$\lim_{\varepsilon\to 0^+} \log \mu(B(x,\varepsilon))/\log \varepsilon \quad \text{exists } \mu\text{-a.s.}$$

The function T can be regarded as a special example of a shift map with respect to some generalized expansion of points in $[0,1]$. Let $r\geq 2$ be a fixed integer, $S=\{0,1,\ldots,r-1\}$, and, for each positive integer n, let $\{a_{n,i}\}$, $i=0,1,\ldots,r$, be a partition of $[0,1]$ i.e. $0=a_{n,0}<a_{n,1}<\cdots<a_{n,r-1}<a_{n,r}=1$. This splits $[0,1]$ into r intervals of length $d_{n,i}=a_{n,i+1}-a_{n,i}$, $i=0,1,\ldots,r-1$. Each point $x\in[0,1]$ then has an associated generalized expansion of digits from S with respect to these partitions. We set the first digit $I_1(x)=i_1$ if $a_{1,i_1}\leq x<a_{1,i_1+1}$. At the next stage each of the preceding r intervals is split into r segments according to the second-stage partition $\{a_{2,i}\}_i$ to determine the second digit $I_2(x)$. We obtain the left-closed interval

$$\{x \mid I_1(x)=i_1,\ I_2(x)=i_2\}=\{x \mid a_{1,i_1}+a_{2,i_2}\,d_{1,i_1}\leq x<a_{1,i_1}+a_{2,i_2+1}\,d_{1,i_1}\}.$$

This procedure is repeated for each n to produce the expansion $x=0\,.\,i_1\,i_2\cdots i_n\cdots$. (The usual r-adic expansion of numbers in $[0,1]$ is a special case of this with $a_{n,i}=i/r$ for each n and i.) Cutler (1988) has discussed a broader class of generalized expansions where additionally the number of splits at the nth stage may vary with n.

We will let $c_n(x)=\{y \mid I_1(y)=I_1(x),\ldots,I_n(y)=I_n(x)\}$. We also make the following restriction:

$$\inf_{0\leq i\leq r-1}\ \inf_n d_{n,i}=d>0. \tag{4.1}$$

Note that as a consequence of (4.1) each sequence of digits from S defines a unique $x\in[0,1]$. (Equivalently, for each x, the diameter

$$d(c_n(x))=d_{1,I_1(x)}\,d_{2,I_2(x)}\cdots d_{n,I_n(x)}\downarrow 0 \quad \text{as } n\to\infty.)$$

The corresponding shift $T:[0,1]\to[0,1]$ is the function mapping $0\,.\,i_1\,i_2\,i_3\cdots$ to $0\,.\,i_2\,i_3\,i_4\cdots$. If the splits $a_{n,i}$ do not depend on n it is possible to obtain a closed form expression for T as a piecewise linear map. In the general case of n-dependent partitions this is not possible but we still obtain the following continuity result.

THEOREM 4.1. *The shift T is continuous at all points of* $[0,1]$ *except the initial interval endpoints* $a_{1,1},\ldots,a_{1,r-1},1$.

Proof. Suppose first that x belongs to the interior of $c_n(x)$ for each n. Then, given n, there exists $\varepsilon>0$ such that the interval $B(x,\varepsilon)\subseteq c_n(x)$. Thus $|y-x|<\varepsilon$ implies $I_j(y)=I_j(x)$ for $j=1,\ldots,n$. It follows then that $I_j(T(y))=I_j(T(x))$ for $j=1,\ldots,n-1$, and hence, from (4.1), $|T(y)-T(x)|\leq(1-d)^{n-1}$. Thus T is continuous at x. If x is the lefthand endpoint of $c_n(x)$ for some $n\geq 2$ (and for no smaller value of n) we can write $x=0\,.\,i_1\cdots i_n\,0\,\bar{0}$ where $i_n\neq 0$. Then $T(x)=0\,.\,i_2\cdots i_n\,0\,\bar{0}$ and it is clear that points immediately to the right of x map near $T(x)$. Points immediately to the left of x also map near $T(x)$ due to the identity $0\,.\cdots i_n-1\ r-1\ \overline{r-1}=0\,.\cdots i_n\,0\,\bar{0}$ which follows from (4.1). Thus T is continuous at x. Continuity fails at the first-stage endpoints $a_{1,j}$ because $T(a_{1,j})=0$ while points to the left of $a_{1,j}$ map near 1. □

The following theorem shows that local dimension can be calculated by taking limits over the intervals $c_n(x)$.

THEOREM 4.2. *Let μ be any probability measure on the Borel sets of* $[0,1]$ *and suppose a family* $\{a_{n,i}\}_{n,i}$ *of partitions satisfies* (4.1). *Define* $\hat{\alpha}_0(x)=\liminf_{n\to\infty}\log\mu(c_n(x))/\log d(c_n(x))$. *Then* $\hat{\alpha}_0=\hat{\alpha}$ μ-a.s. (*In fact* (2.4) *and* (2.5) *hold with* $\hat{\alpha}_0$ *in place of* $\hat{\alpha}$.)

Proof. Note that for each x we have $c_n(x)\subseteq B(x, d(c_n(x)))$ and so

$$\begin{aligned}\hat{\alpha}_0(x) &\geq \liminf_{n\to\infty}\log\mu(B(x, d(c_n(x))))/\log d(c_n(x))\\ &\geq \hat{\alpha}(x). \qquad (4.2)\end{aligned}$$

(Consequently $\hat{\alpha}_0$ must satisfy (2.4).) We need only check that $\hat{\alpha}_0$ also satisfies (2.5) to conclude $\hat{\alpha}_0=\hat{\alpha}$ μ-a.s. To show (2.5) holds we need the analogue of Lemma 2.2 for $\hat{\alpha}_0$. The proof of Lemma 2.2 with $\hat{\alpha}_0$ in place of $\hat{\alpha}$ begins as the proof for $\hat{\alpha}$. We have $E\subseteq\{x\,|\,\hat{\alpha}_0(x)\geq\alpha\}$, $\mu(E)>0$, and $0<\beta<\alpha$. Let $\varepsilon_n(x)=d(c_n(x))$. Then for each $x\in E$ there exists $\varepsilon(x)>0$ such that $\varepsilon_n(x)\leq\varepsilon(x)$ implies $\mu(c_n(x))\leq\varepsilon_n(x)^\beta$. Thus we have $E=\bigcup_{n=1}^\infty E_n$ where $E_n=\{x\in E\,|\,\varepsilon(x)\geq 1/n\}$ and $\theta>0$ such that $\mu(E_n)>\theta$ for all sufficiently large n. At this point we note that the collection of intervals

$$\mathscr{C}=\{c_n(x)\,|\,x\in[0,1], n=1,2,\ldots\}$$

is a complete bounded Vitali covering of [0, 1] (see Cutler (1988) for the definition and discussion of a complete bounded Vitali covering). Thus from Theorem 2.2 of Cutler (1988) it follows that the correct Hausdorff dimension is still obtained for sets when the members of coverings are restricted to being members of $\mathscr{C}$. So let $\{c_i\}_i$ be any $1/n$-covering of E by members of $\mathscr{C}$. We then have the inequalities

$$\theta<\mu(E_n)\leq\sum_{c_i\cap E_n\neq\phi}\mu(c_i)\leq\sum d(c_i)^\beta$$

and hence we conclude $\dim(E)\geq\beta$ and therefore $\dim(E)\geq\alpha$ as required. The proof of (2.5) for $\hat{\alpha}_0$ then follows as in Theorem 2.3. □

Remark. By (4.1) and the argument used in (4.2) we can write

$$\begin{aligned}&\limsup_{n\to\infty}\log\mu(c_n(x))/\log d(c_n(x))\\ &\quad\geq\limsup_{n\to\infty}\log\mu(B(x, d(c_n(x))))/\log d(c_n(x))\\ &\quad=\limsup_{\varepsilon\to 0^+}\log\mu(B(x,\varepsilon))/\log\varepsilon. \qquad (4.3)\end{aligned}$$

This plus Theorem 4.2 shows that if the pointwise limit

$$\lim_{n\to\infty}\log\mu(c_n(x))/\log d(c_n(x)) \quad \text{exists } \mu\text{-a.s.}$$

then so does the pointwise limit $\lim_{\varepsilon\to 0^+}\log\mu(B(x,\varepsilon))/\log\varepsilon$.

To construct distributions μ which are stationary and ergodic with respect to the shift T, let P be any probability measure on the space of one-sided sequences $\Omega = \prod_{i=1}^{\infty} S$ with usual σ-algebra $\mathscr{F}$ generated by the cylinder sets. Then there is a corresponding induced measure $\mu = P\psi^{-1}$ on the Borel sets of $[0, 1]$. If P is stationary and ergodic with respect to the usual shift on the sequence space Ω then μ will be stationary and ergodic with respect to the shift T on $[0,1]$ $(\psi(i_1\, i_2\, i_3 \cdots) = 0\,.\, i_1\, i_2\, i_3 \cdots)$.

To obtain a choice of μ which is not exact-dimensional, set $S = \{0, 1, 2\}$ and let P be the distribution of a periodic Markov chain with state space S, transition matrix

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix},$$

and time-0 distribution given by the unique steady state $\bar{\pi} = (\frac{1}{4}\,\frac{1}{2}\,\frac{1}{4})$. It follows that P is stationary and ergodic with respect to the usual shift on Ω. Noting that P has period 2 we choose an n-dependent family of partitions $\{a_{n,i}\}_{n,i}$ on $[0, 1]$ which takes advantage of this fact by possessing distinct behaviours for odd and even values of n. Let $0 < \beta < \alpha < \frac{1}{2}$ and set

$$a_{n,1} = \alpha, \qquad a_{n,2} = 1 - \alpha \quad \text{for odd } n$$
$$a_{n,1} = \beta, \qquad a_{n,2} = 1 - \beta \quad \text{for even } n.$$

Clearly this partition satisfies (4.1) and thus Theorems 4.2 and 4.3 hold for the corresponding shift T and measure $\mu = P\psi^{-1}$. Note that

$$\begin{aligned} \text{if } I_1(x) = 0 \text{ or } 2, \quad \text{then} \quad & d(c_{2n}(x)) = \alpha^n(1-2\beta)^n \quad \mu\text{-a.s.} \\ \text{and} \quad & d(c_{2n-1}(x)) = \alpha^n(1-2\beta)^{n-1} \quad \mu\text{-a.s.} \end{aligned} \tag{4.4}$$

$$\begin{aligned} \text{if } I_1(x) = 1, \quad \text{then} \quad & d(c_{2n}(x)) = (1-2\alpha)^n\beta^n \quad \mu\text{-a.s.} \\ \text{and} \quad & d(c_{2n-1}(x)) = (1-2\alpha)^n\beta^{n-1} \quad \mu\text{-a.s.} \end{aligned} \tag{4.5}$$

Now from the Shannon–McMillan–Breiman Theorem and the known entropy for Markov shifts (see Billingsley, 1978) we obtain

$$\lim_{n\to\infty} -\log \mu(c_n(x))/n = (\log 2)/2 \quad \mu\text{-a.s.} \tag{4.6}$$

In fact (4.6) is not difficult to compute directly without resorting to entropy since explicit expressions along the lines of (4.4) and (4.5) can be obtained for $\mu(c_n(x))$.

Consequently from (4.4), (4.5), and (4.6) we obtain

$$\begin{aligned} \hat{\alpha}_0(x) &= \lim_{n\to\infty} \log \mu(c_n(x))/\log d(c_n(x)) \\ &= \begin{cases} \log 2/2 \log \alpha^{-1}(1-2\beta)^{-1} & \mu\text{-a.s. when } I_1(x) = 0 \text{ or } 2 \\ \log 2/2 \log (1-2\alpha)^{-1}\beta^{-1} & \mu\text{-a.s. when } I_1(x) = 1. \end{cases} \end{aligned} \tag{4.7}$$

This shows that the dimension distribution $\hat{\mu} = (1/2)\delta_{\alpha_1} + (1/2)\delta_{\alpha_2}$ where $\alpha_1 = \log 2/2 \log \alpha^{-1}(1-2\beta)^{-1}$, $\alpha_2 = \log 2/2 \log (1-2\alpha)^{-1}\beta^{-1}$, and δ_{α_1}, δ_{α_2} denote unit masses at α_1 and α_2 respectively. Furthermore, appealing to the Remark following Theorem 4.2, we see that the pointwise limit

$$\hat{\alpha}(x) = \lim_{\varepsilon\to 0^+} \log \mu(B(x, \varepsilon))/\log \varepsilon \quad \text{exists } \mu\text{-a.s.}$$

https://doi.org/10.1017/S014338570000568X Published online by Cambridge University Press

Note. The map T discussed above has two points of discontinuity. This can be modified to a completely continuous map $\tilde{T}$ in the following manner: Let $x = 0.i_1 i_2 i_3 \cdots$ be a point in $[0, 1]$. If $i_1 = 0$ or 2 then define $\tilde{T}(x) = T(x)$. If $i_1 = 1$, then $\tilde{T}(x)$ is defined by a two-step procedure: first exchange all 0's and 2's in the expansion of x (a mapping $x \to \tilde{x}$) then apply T to the resulting point $\tilde{x}$, i.e. $\tilde{T}(x) = T(\tilde{x})$. It is not difficult to see that in sequence space this procedure also defines a Markov chain with the same transition matrix **P** as the first procedure (so the measure μ on $[0, 1]$ remains unchanged). It is also straightforward to verify that the map $\tilde{T}$ is continuous at all points of $[0, 1]$.

Concluding remarks

(1) This example shows that the dimensional behaviour of a measure μ depends not only on ergodicity but on the degree to which the associated mapping T separates nearby points in a supporting set for μ.

(2) This same technique can obviously be extended to construct ergodic measures with positive mass on k different dimensions by using a chain of period k and appropriate partitions.

(3) If the chain is chosen to be aperiodic then $\mu = P\psi^{-1}$ will be exact-dimensional for any choice of partitions satisfying (4.1). A finite aperiodic Markov chain is mixing and from (4.1) and the 0–1 law there exists a finite constant c such that

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \log d_{j, I_j(x)} = c \quad \mu\text{-a.s.}$$

The Shannon–McMillan–Breiman theorem then gives $\hat{\alpha}(x) = h/|c|$ μ-a.s., where h is the entropy of the chain. This shows that T has better behaviour at points in the supporting sets of mixing distributions.

(4) Conditions (3.1)(a) and (3.2)(a) can be required to hold only for μ-almost all y in a neighbourhood of x. In fact further weakening of these conditions can be done but may not be of any practical importance.

Acknowledgement. The suggestion that the map T might be patched to make it continuous at all points was due to a participant of the 1990 Oberwolfach Measure Theory Conference, whose name is regrettably not known to the author.

Note added in proof. Since the set of atoms of μ has dimension 0, it can be temporarily removed from D_α in the proof of Lemma 2.1. Hence the balls in $\mathscr{B}_{x,n}$ must have positive diameter (thus validating the use of Theorem 2.2).

REFERENCES

Billingsley, P. (1960), Hausdorff dimension in probability theory. *Illinois J. Math.* **4** 187–209.
Billingsley, P. (1961), Hausdorff dimension in probability theory II. *Illinois J. Math.* **5** 291–298.
Billingsley, P. (1978), *Ergodic Theory and Information.* Krieger.
Cutler, C. D. (1986), The Hausdorff dimension distribution of finite measures in Euclidean space. *Canad. J. Math.* **38** (6) 1459–1484.

Cutler, C. D. and Dawson, D. A. (1989), Estimation of dimension for spatially-distributed data and related limit theorems. *J. Multivariate Anal.* **28** 115–198

Cutler, C. D. (1988), A note on equivalent interval covering systems for Hausdorff dimension on R. *Int. J. Math. Math. Sci.* **11** (4) 643–650.

Eckmann, J.-P. and Ruelle, D. (1985), Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.* **57** (3) Part 1, 617–656.

Falconer, K. J. (1985), *The Geometry of Fractal Sets.* Cambridge University Press.

Farmer, J. D., Ott, E. and Yorke, J. A. (1983), The dimension of chaotic attractors. *Physica* **7D** 153–180.

Guckenheimer, J. (1984), Dimension estimates for attractors. *Contemp. Math.* **28** 357–367.

Rogers, C. A. (1970), *Hausdorff Measures.* Cambridge University Press.

Young, L.-S. (1982), Dimension, entropy, and Lyapunov exponents. *Ergod. Th. & Dyn. Sys.* **2** 109–124.