ROTATION, ENTROPY, AND EQUILIBRIUM STATES

OLIVER JENKINSON

Abstract. For a dynamical system (X, T) and function $f : X \to \mathbb{R}^d$ we consider the corresponding generalised rotation set. This is the convex subset of $\mathbb{R}^d$ consisting of all integrals of f with respect to T-invariant probability measures. We study the entropy $H(\varrho)$ of rotation vectors ϱ, and relate this to the directional entropy $\mathcal{H}(\varrho)$ of Geller & Misiurewicz [15]. For (X, T) a mixing subshift of finite type, and f of summable variation, we prove that if the rotation set is strictly convex then the functions $\mathcal{H}$ and H are in fact one and the same. For those rotation sets which are not strictly convex we prove that $\mathcal{H}(\varrho)$ and $H(\varrho)$ can differ only at non-exposed boundary points ϱ.

Contents

1. Introduction

Consider a continuous map $T : X \to X$ of a compact metrizable space X, and a continuous function $f : X \to \mathbb{R}^d$ with coordinate functions $f_1, \dots , f_d$. If $\mathcal{M}$ denotes the set of T-invariant Borel probability measures, then f induces the map $f_* : \mathcal{M} \to \mathbb{R}^d$ given by

$$f_*(\mu) = \left(\int f_1 \, d\mu, \dots , \int f_d \, d\mu \right) .$$

We call $f_*(\mu)$ the *rotation vector* of the measure $\mu \in \mathcal{M}$. The image $f_*(\mathcal{M})$ is clearly compact and convex, since $\mathcal{M}$ is convex and weak* compact, and f_* is affine and continuous. We call $f_*(\mathcal{M})$ the *rotation set*. For $\varrho \in f_*(\mathcal{M})$, the fibre $f_*^{-1}(\varrho)$ is

1991 Mathematics Subject Classification: Primary 54H20, 58F11, 28D20; Secondary: 58F15.

called the *rotation class* of ϱ. This general definition of a rotation set has recently been studied by Blokh [8], Geller & Misiurewicz [15], and Ziemian [59].

Rotation sets were originally defined in the context of degree-one circle maps [40], or more generally d-dimensional torus maps homotopic to the identity [24], both of which are generalisations of Poincaré's [46] notion of rotation number for circle homeomorphisms. Such sets describe the asymptotic "drift" of orbits in the universal cover $\mathbb{R}^d$. Here $f = F \circ \pi^{-1} - \pi^{-1}$, where F is some fixed lift of the torus map to $\mathbb{R}^d$, and π^{-1} some local inverse of the natural covering map $\pi : \mathbb{R}^d \to \mathbb{T}^d$. An open question is to determine which compact convex sets can arise as such rotation sets. Kwapisz [25] has shown that any polygon whose vertices are at rational points in the plane is the rotation set of some homeomorphism of the two-torus. However such rotation sets need not be polygons [26].

A different rotation set arises in the following way, motivated by Aubry-Mather theory [1], [37] and the work of Schwartzman [50]. Let M be a compact orientable Riemannian surface of genus g, with corresponding real homology space $H_1(M, \mathbb{R}) = \mathbb{R}^{2g}$. Let $\mathcal{M}$ denote the set of probability measures on the unit tangent bundle T^1M which are invariant under the geodesic flow. For $\mu \in \mathcal{M}$ we can define a linear functional Φ_μ on 1-forms by $\Phi_\mu(\omega) = \int_{T^1M} \omega \, d\mu$. Flow invariance of μ implies the functional is well-defined on the real (de Rham) cohomology space $H^1(M, \mathbb{R})$, and hence by duality can be considered as an element of $H_1(M, \mathbb{R})$. By the Hodge theorem (see [57]) we can choose a basis $\omega_1, \ldots, \omega_b$ of harmonic forms for $H^1(M, \mathbb{R})$, where b denotes the first Betti number of M. The map $\mu \mapsto \Phi_\mu$ is continuous and affine, so its image in $\mathbb{R}^{2g}$ is compact and convex. By choosing the coordinate functions of $f : T^1M \to \mathbb{R}^{2g}$ to be the harmonic forms ω_i, we see this image is the rotation set $f_*(\mathcal{M})$. In fact this rotation set is precisely the unit ball for the stable norm (or Federer norm) on $H_1(M, \mathbb{R})$ (see [30], [31], [35], [36] for more details). For surfaces of genus $g \geq 2$, Massart [35] [36] has shown this rotation set is never strictly convex and never smooth. In particular, if a point of strict convexity on the boundary corresponds to a measure supported on a closed geodesic then the boundary is non-differentiable at that point. See [30], [31] for the case where the compact surface M is replaced by a punctured hyperbolic torus of finite area. There are various generalisations in which the geodesic flow is replaced by a Lagrangian flow [3], [4], [32], [33], [38].

Another type of rotation set is the *barycentre set* (see [9], [20], [21], [22]) of a circle map $T : S^1 \to S^1$. If $\mathcal{M}$ denotes the set of T-invariant probability measures then the *barycentre* $f_*(\mu) = \int_{S^1} z \, d\mu(z)$ of a measure $\mu \in \mathcal{M}$ reflects its average weight around the circle. The properties of the corresponding set $f_*(\mathcal{M})$ depend on T. For example if $T(z) = z^k$, $k \geq 2$, then $f_*(\mathcal{M})$ is strictly convex but its boundary is non-differentiable at a countable dense subset, and corresponds to a certain one-parameter family of zero entropy measures (see [9], [22]). The points of non-differentiability correspond to measures supported on certain periodic orbits.

A familiar rotation set is given by the spectrum of Lyapunov exponents of a smooth map $T : X \to X$ of a Riemannian manifold. To every T-invariant measure μ we associate its Lyapunov exponent $\int \log ||DT_x||_x \, d\mu(x)$, and the set of all such exponents is a closed interval. This one-dimensional rotation set is less interesting geometrically, though challenging problems are to determine which measures correspond to the endpoints of the interval, and the way in which (maximal) entropy varies over the interval. There are various other *multifractal spectra*, many of which can be cast in terms of rotation sets (see [44], and the additional comments at the end of this Introduction).

Geller & Misiurewicz [15] recently considered rotation sets in a general setting, and introduced the notions of *directional* and *lost* ergodic measures. An ergodic measure μ is *directional* if all measures whose support is contained in the support of μ have the same rotation vector as μ. Otherwise an ergodic μ is called *lost*. The *directional entropy* $\mathcal{H}(\varrho)$ of a rotation vector ϱ is then defined as the supremum of the entropies of those directional measures with rotation vector ϱ. We clearly have $\mathcal{H}(\varrho) \leq H(\varrho)$, where $H(\varrho)$ is defined as the supremum of the entropies of *all* measures in the rotation class $f_*^{-1}(\varrho)$.

Geller & Misiurewicz [15] consider in detail the case of a mixing subshift of finite type, where the function f is constant on length-one cylinder sets. They show that the entropy of any ergodic measure μ can be approximated arbitrarily well by a directional measure with the same rotation vector. Furthermore, provided μ has its rotation vector in the (relative) interior of the rotation set, its entropy can be approximated arbitrarily well by a lost measure with the same rotation vector.

One of the goals of this paper is to sharpen and extend these results in various directions. We first consider the very general case where the function f is merely continuous, and the dynamical system (X, T) has upper semi-continuous entropy map h. This guarantees that every continuous function $g : X \to \mathbb{R}$ has at least one equilibrium state (see [55], p. 224).

The upper semi-continuity of h is guaranteed, for example, if T is (topologically) expansive (see [55]), or more generally if T admits a finite generating partition α (i.e. for all $\mu \in \mathcal{M}$, the limiting refinement $\vee_{i=0}^{\infty} T^{-i}\alpha$ agrees with the Borel σ-algebra up to sets of μ-measure zero) such that boundaries of sets in α have zero measure for every invariant measure (see [23], Cor. 4.2.5). In particular this includes all symbolic systems, since here the natural generating partition (by length-one cylinders) consists of sets without boundary, and hence as a special case all subshifts of finite type. Upper semi-continuity of h is also guaranteed (see [39]) if T is a C^∞ map of a compact manifold.

In this context we prove (Theorem 1) that the rotation set is the closure of the set of rotation vectors of a distinguished d-parameter family of *equilibrium states*.

Next we specialize to the case of (X, T) a mixing subshift of finite type. Here our assumption on f is that it is of summable variation (in particular this includes all locally constant functions) and cohomologically full (a non-triviality hypothesis which ensures the rotation set has interior as a subset of $\mathbb{R}^d$). We prove (Theorem 3) that for ϱ in the interior of the rotation set, the entropy $H(\varrho)$ is achieved (uniquely) by a lost measure in this rotation class. Moreover, this lost measure is always an equilibrium state from our distinguished d-parameter family.

Next we consider the boundary of the rotation set. If ϱ is an *exposed* point of the boundary, we show (Proposition 5) that its rotation class $f_*^{-1}(\varrho)$ cannot contain *any* lost measures. Moreover (Theorem 4), the entropy $H(\varrho)$ is attained by at least one directional measure in $f_*^{-1}(\varrho)$, so that for exposed ϱ we have the equality $\mathcal{H}(\varrho) = H(\varrho)$.

We then show (Theorem 5) that if ϱ is an *interior* point of the rotation set, then again we have the equality $\mathcal{H}(\varrho) = H(\varrho)$. That is, there exist directional measures in $f_*^{-1}(\varrho)$ with entropy arbitrarily close to (but not equal to, by Theorem 3) $H(\varrho)$.

Therefore the two entropy functions $\mathcal{H}$ and H (which for general dynamical systems (X, T) are *not* the same) are essentially the same in the case where (X, T)

is a subshift of finite type. The only rotation vectors ϱ at which $\mathcal{H}(\varrho)$ and $H(\varrho)$ might differ are non-exposed points of the boundary. So in particular if the rotation set is *strictly* convex then H and $\mathcal{H}$ coincide completely (Corollary 7).

In sections 10 and 11 we consider the possible behaviour of $\mathcal{H}$ and H at non-exposed boundary points ϱ. Here $H(\varrho)$ may be attained uniquely by a lost measure, uniquely by a directional measure, by both a lost and a directional measure, or by neither. In particular, we may or may not have $\mathcal{H}(\varrho) = H(\varrho)$.

Our general approach is motivated by thermodynamic formalism, in particular the theory of pressure and equilibrium states developed by Ruelle [48], [49] and Walters [54], [55]. After briefly recalling some notions from convex geometry in section 2, we develop some thermodynamic ideas in section 3. We relate (Theorem 1) the rotation set to the image of the subdifferential of a certain pressure function, and prove it can be parametrised by the rotation vectors of a d-dimensional family of equilibrium states. In section 4 we consider subshifts of finite type, and review facts about equilibrium states for sufficiently regular (potential) functions. In section 5 we prove a new version of a powerful lemma due independently to Bousch and Mañé, which is later used to study the boundary of the rotation set. In section 6 we consider the entropy function H, and show (Theorem 2, Theorem 3) it is completely determined by the d-parameter family of equilibrium states. In section 7 we show how lost and directional measures are related to various regions of the rotation set. Section 8 is the key section. Here we prove (Theorem 4, Theorem 5, Theorem 6) that in fact the entropy functions H and $\mathcal{H}$ coincide, except possibly at non-exposed boundary points. In section 9 we specialize to locally constant functions, in which case the rotation set is a polyhedron [59]. We give a new proof (Theorem 7) of a result of Marcus & Tuncel [34] that each face of the polyhedron itself corresponds to a subshift of finite type. Sections 10 and 11 consist of examples of the range of pathological behaviour possible at non-exposed boundary points.

Our approach builds on some ideas used in [20] and [21]. While preparing this paper we learned that other authors have used a similar approach in studying related problems. In particular we mention the recent article of Babillot & Ledrappier [2] on the asymptotic distribution of periodic orbits for hyperbolic flows, and the work of Tuncel and co-workers (see for example [34], [43] and the references therein) on the beta function of Markov chains. There are also connections with multifractal analysis (see Pesin [44], Fan & Feng [14]), in particular the following very general multifractal spectrum proposed by Barreira, Pesin & Schmeling [5] (see also Appendix IV of [44]). Let $g : Y \to [-\infty, \infty]$ be defined on some subset $Y \subset X$, and G a real-valued function defined on the collection of subsets of Y. Then the function $\mathcal{F} : g(X) \to \mathbb{R}$ defined by $\mathcal{F}(\varrho) = G(g^{-1}\varrho)$ is called the (g, G)-*multifractal spectrum*. To relate this to our context we let $g(x) = \lim_{n\to\infty} \frac{1}{n}\sum_{r=0}^{n-1} f(T^r x)$ on the subset $Y \subset X$ where these ergodic averages exist, so that $g(Y)$ is then the rotation set of f. Setting $G = h_{top}$ recovers our entropy function H as the (g, G)-multifractal spectrum. To obtain the directional entropy $\mathcal{H}$ as a multifractal spectrum we define $G(Z)$ to be the supremum of the entropies $h(\mu)$ of those directional measures μ with $\mu(Z) = 1$.

Acknowledgments. I thank Thierry Bousch for a copy of [9], for helpful discussions, and for a careful reading of a preliminary version of this paper. I thank Ai Hua Fan for a copy of [14], and for useful comments on a preliminary version of

this paper. I thank Evgeny Verbitski for pointing out an error in a previous version of this paper, and for a copy of [52].

2. Convex geometry

Let us recall a few notions from convex geometry (see [18], Ch. 2).

Our ambient space will be $\mathbb{R}^d$, where for $u, v \in \mathbb{R}^d$ we denote the Euclidean inner product by $u.v = u_1 v_1 + \ldots + u_d v_d$. A subset $K \subset \mathbb{R}^d$ is *convex* if for every $u, v \in K$ we have that $\lambda u + (1-\lambda)v \in K$ for all $\lambda \in (0,1)$. A point $u \in K$ is an *extremal* point of K if $v, w \in K$, $\lambda \in (0,1)$, and $u = \lambda v + (1-\lambda)w$ imply that $u = v = w$.

A hyperplane $P = \{u \in \mathbb{R}^d : u.v = a\}$ is said to *cut* K provided both open half-spaces determined by P contain points of K. We say P *supports* K if its distance from K is zero yet it does not cut K. A set $F \subset K$ is called a *face* of K if $F = K \cap P$ for some supporting hyperplane P. A point $u \in K$ is called *exposed* if the singleton $\{u\}$ is a face of K. A compact convex set K is strictly convex if and only if every point on the boundary ∂K is an exposed point.

All exposed points are extremal points, but not conversely. Consider, for example, $K = \{(v_1, v_2) \in \mathbb{R}^2 : v_1^2 + v_2^2 \le 1\} \cup \{(v_1, v_2) \in \mathbb{R}^2 : |v_1| \le 1 \text{ and } -1 \le v_2 \le 0\}$. Here the points $(1,0)$ and $(-1,0)$ are extremal but not exposed.

If K has a finite number of extremal points it is called a *polyhedron*. In this case every extremal point is also exposed.

The *interior*, denoted $int(A)$, of a set $A \subset \mathbb{R}^d$ is the union of all subsets of A which are open as subsets of $\mathbb{R}^d$. The *relative interior*, denoted $ri(A)$, is the interior of A with respect to the topology of the smallest affine subspace of $\mathbb{R}^d$ containing A.

3. A family of equilibrium states

By a *dynamical system* (X, T) we will mean a continuous surjection $T : X \to X$ of a compact metrizable space X. Almost all of our definitions will make sense in this context. For proving results, however, our minimum assumption will be that the corresponding entropy map $h : \mathcal{M} \to \mathbb{R}$ is upper semi-continuous. As previously noted, this is true for a large class of systems. Our minimum assumption on the function $f : X \to \mathbb{R}^d$ is that it is continuous. At various points we will make extra hypotheses about T and f so as to prove stronger results.

Definition 1. Given a continuous function $g : X \to \mathbb{R}$ we define its *pressure* $P(g)$ with respect to T to be

$$P(g) = \sup_{\mu \in \mathcal{M}} \left(h(\mu) + \int g \, d\mu \right),$$

where $h(\mu)$ denotes the entropy of μ. If $m \in \mathcal{M}$ satisfies $P(g) = h(m) + \int g \, dm$ then it is called an *equilibrium state* for g. Let ES_g denote the set of all equilibrium states for g.

The set ES_g is convex, and the extremal points are precisely the ergodic equilibrium states. Upper semi-continuity of entropy means ES_g is compact and non-empty ([55] p. 224).

Definition 2. Given $v \in \mathbb{R}^d$ we let $v.f$ denote the function $v_1 f_1 + \ldots + v_d f_d$. Define $p : \mathbb{R}^d \to \mathbb{R}$ by $p(v) = P(v.f)$. Of course p depends on both T and f, though for ease of notation we suppress this dependence.

Let $\mathcal{M}_f(v) = ES_{v.f}$ denote the set of equilibrium states of $v.f$, so that $\mathcal{M}_f$ is a map from $\mathbb{R}^d$ to the power set of $\mathcal{M}$. We will be interested in the d-parameter family $\mathcal{M}_f(\mathbb{R}^d)$. We call this the *family of f-equilibrium states.*

Since pressure P is convex ([55] p. 214) then so is $p : \mathbb{R}^d \to \mathbb{R}$. Although p need not be differentiable, its convexity is enough to ensure the existence of directional derivatives at each point.

Definition 3. Let

$$p'(v; h) = \lim_{t \downarrow 0} \frac{p(v + th) - p(v)}{t}$$

denote the directional derivative of p at the point v in the direction $h \in \mathbb{R}^d$.

A vector $u \in \mathbb{R}^d$ is a *subgradient* of p at the point $v \in \mathbb{R}^d$ if

$$p(v + h) \geq p(v) + u.h \quad \text{for all } h \in \mathbb{R}^d.$$

The set of all subgradients of p at v is called the *subdifferential* of p at v, and is denoted by $\partial p(v)$. Let $\partial p(\mathbb{R}^d)$ denote $\cup_{v \in \mathbb{R}^d} \partial p(v)$ (i.e. a subset of $\mathbb{R}^d$).

Clearly $\partial p(v)$ is a closed convex subset of $\mathbb{R}^d$. It reduces to a point if p is differentiable at v, in which case $\partial p(v)$ is the gradient $\nabla p(v)$.

We now prove an explicit formula for the set $\partial p(v)$. Note that if T is a mixing subshift of finite type and f is Hölder (in particular these conditions guarantee that p is differentiable everywhere) then this formula can be derived by applying perturbation theory to the Ruelle transfer operator ([49] p. 133, [42] p. 60).

Proposition 1. *Let (X, T) be a dynamical system for which the entropy map is upper semi-continuous, and suppose $f : X \to \mathbb{R}^d$ is continuous. Then $\partial p(v) = f_*(\mathcal{M}_f(v))$ for all $v \in \mathbb{R}^d$.*

Proof. We claim that for all $h \in \mathbb{R}^d$,

$$p'(v; h) = \sup \{ f_*(m).h : m \in \mathcal{M}_f(v) \}. \tag{1}$$

Indeed this follows from Lemma 1 of [56], where the functions f, g in [56] are replaced by $v.f$ and $h.f$ respectively (since upper semi-continuity of entropy means $\mathcal{M}_f(v)$ coincides with the set of tangent functionals to pressure at the function $v.f$).

By [47], Thm. 23.2, equation (1) gives that for each $m \in \mathcal{M}_f(v)$, the rotation vector $f_*(m)$ is a subgradient of p at the point v. For the converse, note that (1), together with compactness of $\mathcal{M}_f(v)$, implies each point u on the *boundary* of the convex set $\partial p(v)$ can be written $u = f_*(m)$ for some $m \in \mathcal{M}_f(v)$. But then the convexity of the two sets $\partial p(v)$ and $\mathcal{M}_f(v)$ means that in fact *every* point in $\partial p(v)$ is of the form $f_*(m)$ for some $m \in \mathcal{M}_f(v)$, thus proving the result.

□

Corollary 1. *Let (X, T) be a dynamical system for which the entropy map is upper semi-continuous, and suppose $f : X \to \mathbb{R}^d$ is continuous. Then p is differentiable at the point $v \in \mathbb{R}^d$ if and only if $f_*(\mathcal{M}_f(v))$ is a singleton.*

Proposition 2. *Let (X, T) be a dynamical system for which the entropy map is upper semi-continuous, and suppose $f : X \to \mathbb{R}^d$ is continuous. Then $f_*(\mathcal{M}) \subset \overline{\partial p(\mathbb{R}^d)}$.*

Proof. If $\mu \in \mathcal{M}$ then the definition of pressure gives

$$p(v) = P(v_1 f_1 + \ldots + v_d f_d) \geq h(\mu) + v_1 \int f_1 \, d\mu + \ldots + v_d \int f_d \, d\mu. \tag{2}$$

Now let $A_\mu : \mathbb{R}^d \to \mathbb{R}$ be the affine map

$$A_\mu(v_1, \ldots, v_d) = h(\mu) + v_1 \int f_1 \, d\mu + \ldots + v_d \int f_d \, d\mu.$$

Then $Graph(A_\mu)$ is a hyperplane in $\mathbb{R}^{d+1}$, and the inequality (2) means $Graph(p)$ lies above this hyperplane (possibly touching it tangentially).

This implies that the gradient of the hyperplane is contained in the closure of all subdifferentials of p, since if not then the graphs would intersect transversally. That is, the vector $f_*(\mu)$ belongs to the closure of $\partial p(\mathbb{R}^d)$. But $\mu \in \mathcal{M}$ was arbitrary, so $f_*(\mathcal{M}) \subset \overline{\partial p(\mathbb{R}^d)}$. □

Theorem 1. *Let (X,T) be a dynamical system for which the entropy map is upper semi-continuous, and suppose $f : X \to \mathbb{R}^d$ is continuous. Then*

$$f_*(\mathcal{M}) = \overline{f_*(\mathcal{M}_f(\mathbb{R}^d))} = \overline{\partial p(\mathbb{R}^d)}.$$

Proof. Clearly $f_*(\mathcal{M}_f(\mathbb{R}^d)) \subset f_*(\mathcal{M})$ since $\mathcal{M}_f(\mathbb{R}^d) \subset \mathcal{M}$. Compactness of $f_*(\mathcal{M})$ means in fact $\overline{f_*(\mathcal{M}_f(\mathbb{R}^d))} \subset f_*(\mathcal{M})$. Now Proposition 1 gives $\partial p(\mathbb{R}^d) = f_*(\mathcal{M}_f(\mathbb{R}^d))$, so Proposition 2 gives $f_*(\mathcal{M}) \subset \overline{f_*(\mathcal{M}_f(\mathbb{R}^d))}$. □

We would like to discuss the interior of $f_*(\mathcal{M})$, and show it is a subset of (and in good cases is equal to) $f_*(\mathcal{M}_f(\mathbb{R}^d))$. To avoid making vacuous statements we consider the *relative interior* of $f_*(\mathcal{M})$ (see §2).

Corollary 2. *Let (X,T) be a dynamical system for which the entropy map is upper semi-continuous, and suppose $f : X \to \mathbb{R}^d$ is continuous. Then $ri(f_*(\mathcal{M})) \subset f_*(\mathcal{M}_f(\mathbb{R}^d)) = \partial p(\mathbb{R}^d)$.*

Proof. By Proposition 1 we have the equality $f_*(\mathcal{M}_f(\mathbb{R}^d)) = \partial p(\mathbb{R}^d)$. Now $\partial p(\mathbb{R}^d)$ need not be a convex set, but it is almost convex in the sense that

$$ri(dom(\hat{p})) \subset \partial p(\mathbb{R}^d) \subset dom(\hat{p}). \tag{3}$$

Here $\hat{p} : \mathbb{R}^d \to [-\infty, \infty]$ denotes the convex conjugate of p, defined by $\hat{p}(v) = \sup\{u.v - p(u) : u \in \mathbb{R}^d\}$, and $dom(\hat{p}) = \{v \in \mathbb{R}^d : \hat{p}(v) < \infty\}$ is its effective domain (see [47] §12).

Taking the closure followed by the relative interior of all sets in (3) gives the equality $ri(\overline{\partial p(\mathbb{R}^d)}) = ri(dom(\hat{p}))$. Combining this equality with (3) gives

$$ri(\overline{\partial p(\mathbb{R}^d)}) \subset \partial p(\mathbb{R}^d). \tag{4}$$

Now we just take relative interiors of the equality established in Theorem 1 to obtain $ri(f_*(\mathcal{M})) = ri(\overline{\partial p(\mathbb{R}^d)})$, then apply (4) to complete the proof. □

Corollary 3. *Let (X,T) be a dynamical system for which the entropy map is upper semi-continuous, and suppose $f : X \to \mathbb{R}^d$ is continuous. If p is strictly convex then $int(f_*(\mathcal{M})) = f_*(\mathcal{M}_f(\mathbb{R}^d))$.*

Proof. Strict convexity of p implies $\partial p(\mathbb{R}^d)$ is both open in $\mathbb{R}^d$ and convex ([47] p. 227), so by Proposition 1, $f_*(\mathcal{M}_f(\mathbb{R}^d))$ is open and convex. Thus we have $int\left(\overline{f_*(\mathcal{M}_f(\mathbb{R}^d))}\right) = f_*(\mathcal{M}_f(\mathbb{R}^d))$, and the result follows from Theorem 1. □

4. Subshifts of finite type

Consider the finite set $\{1, \dots, k\}$ with the discrete topology. Let A be a $k \times k$ matrix of zeros and ones. Define $X = X_A$ to be the set of all two-sided sequences $x = (x_n)_{n=-\infty}^{\infty}$ for which $x_n \in \{1, \dots, k\}$ and $A(x_n, x_{n+1}) = 1$ for all $n \in \mathbb{Z}$. We give X_A the Tychonov product topology, and call it the *two-sided subshift of finite type* defined by A.

We have a corresponding *one-sided subshift of finite type* X_A^+ defined as the set of all one-sided sequences $x = (x_n)_{n=1}^{\infty}$ for which $x_n \in \{1, \dots, k\}$ and $A(x_n, x_{n+1}) = 1$ for all $n \geq 1$. Clearly X_A is the natural extension of X_A^+. If $x = (x_n) \in X_A^+$ and $1 \leq i \leq k$ is such that $A(i, x_1) = 1$ we let ix denote the element $y = (y_n) \in X_A^+$ defined by $y_1 = i$, $y_{n+1} = x_n$ for $n \geq 1$.

The only map we ever consider on either a one-sided or a two-sided subshift of finite type X is the *shift map* defined by $(Tx)_n = x_{n+1}$. For this reason we also refer to the pair (X, T) as a *subshift of finite type.*

A subshift of finite type is *(topologically) mixing* if the matrix A is aperiodic (i.e. there exists $n \in \mathbb{N}$ with $A^n(i,j) > 0$ for all $1 \leq i, j \leq k$). A mixing subshift of finite type has positive topological entropy and a unique measure of maximal entropy, [41]. The set $\mathcal{M}$ of invariant measures is an infinite dimensional simplex (the so-called Poulsen simplex, see [27], [16]).

For a subshift of finite type X, let $C(X)$ denote the space of continuous real-valued functions on X, equipped with the uniform norm $|\cdot|_\infty$.

Given $x = (x_i) \in X$ we let $\pi_n(x) = (x_1, \dots, x_n)$ if X is one-sided, and $\pi_n(x) = (x_{-n}, \dots, x_n)$ if X is two-sided. A *cylinder* of length n is any set of the form $\{y \in X : \pi_n(y) = \pi_n(x)\}$ for some $x \in X$. For $f : X \to \mathbb{R}^d$ we define

$$\mathrm{var}_n(f) = \max_{\pi_n(x) = \pi_n(y)} |f(x) - f(y)|,$$

where $|\cdot|$ is the Euclidean norm on $\mathbb{R}^d$. We say f has *summable variation* if $V(f) := \sum_{n=1}^{\infty} \mathrm{var}_n(f) < \infty$. A key property of functions of summable variation is that

$$\text{if } \pi_n(x) = \pi_n(y) \text{ then } |f^n(x) - f^n(y)| \leq V(f), \tag{5}$$

where we use the notation $f^n(x) = \sum_{r=0}^{n-1} f(T^r x)$.

We say a continuous function $f : X \to \mathbb{R}^d$ is *Hölder* of exponent $\theta \in (0,1)$ if there exists $K > 0$ such that $\mathrm{var}_n(f) < K\theta^n$. Note that any Hölder continuous function has summable variation.

If a real-valued function $\varphi : X \to \mathbb{R}$ is Hölder of exponent θ we define $|\varphi|_\theta = \sup\{\theta^{-n}\mathrm{var}_n(\varphi) : n \geq 0\}$. Let $\mathcal{F}_\theta(X)$ denote the space of θ-Hölder functions equipped with the Banach norm $||\cdot||_\theta$ defined by $||\varphi||_\theta = |\varphi|_\theta + |\varphi|_\infty$.

Let $C'(X)$ (resp. $\mathcal{F}'_\theta(X)$) denote the quotient space defined by identifying those elements of $C(X)$ (resp. $\mathcal{F}_\theta(X)$) which differ by a constant. Note that $var_n(\cdot)$ is well-defined on both $C'(X)$ and $\mathcal{F}'_\theta(X)$. Moreover $|\cdot|_\theta$ (which is not a norm on $\mathcal{F}_\theta(X)$) lifts to a Banach norm (which we will also denote by $|\cdot|_\theta$) on $\mathcal{F}'_\theta(X)$. We equip $C'(X)$ with the quotient norm (which we also denote by $|\cdot|_\infty$) defined by

$$|\Phi|_\infty = \min\{|\varphi|_\infty : \varphi \in \Phi\} = \frac{1}{2}\mathrm{var}_0(\Phi). \tag{6}$$

We say two real-valued functions are *essentially cohomologous* if their difference is a function of the form $\psi \circ T - \psi + c$ for some bounded Borel measurable function ψ and some $c \in \mathbb{R}$. If $c = 0$ we say the functions are *cohomologous.* The function

ψ is called the *cobounding function*. We say a real-valued function is an *essential coboundary* if it is cohomologous to a constant. We say a vector-valued function $f : X \to \mathbb{R}^d$ is *cohomologically full* if its coordinate functions $f_1, \dots, f_d$ are cohomologically independent (i.e. if all non-trivial linear combinations of the f_i are not essential coboundaries). We will often assume that our function f is cohomologically full, to guarantee that as a subset of $\mathbb{R}^d$ it has interior (see Corollary 4). Of course this is no essential restriction, as we can always choose some maximal cohomologically independent subset $f_1, \dots, f_{d'}$, $0 \le d' < d$, in which case $f_*(\mathcal{M})$ will lie in some d'-dimensional hyperplane, which we can identify with $\mathbb{R}^{d'}$.

All our results will hold for both one-sided and two-sided mixing subshifts of finite type, though for convenience we will always work with one-sided systems. The justification for this is the following construction, due in the Hölder case to Sinai [51], and in the summable variation case to Coelho & Quas [11]. Given a two-sided subshift of finite type X_A, and a function $f : X_A \to \mathbb{R}^d$ of summable variation (resp. an element of $\mathcal{F}_\theta(X_A)$), there exists a function $f^+ : X_A \to \mathbb{R}^d$ of summable variation (resp. an element of $\mathcal{F}_{\sqrt{\theta}}(X)$) which is cohomologous to f, and such that $f^+(x) = f^+(y)$ whenever $x_n = y_n$ for all $n \ge 1$. Thus f^+ only depends on 'future' coordinates, and so can be considered as a function on the one-sided space X_A^+. Since f, f^+ are cohomologous, their integrals with respect to any invariant measure are identical.

Lemma 1. *If (X,T) is a mixing subshift of finite type, and $f : X \to \mathbb{R}^d$ is a cohomologically full function with summable variation, then*
(a) Each $\mathcal{M}_f(v)$ contains a single measure, $m_{v.f}$ say.
(b) Each $m_{v.f}$ is fully supported.
(c) T has a Bernoulli natural extension with respect to $m_{v.f}$, and in particular the entropy $h(m_{v.f})$ is positive.
(d) If $v, v' \in \mathbb{R}^d$ with $v \neq v'$ then the equilibrium states $m_{v.f}, m_{v'.f}$ are distinct.
(e) The function p defined by $p(v) = P(v.f)$ is strictly convex.
(f) If f is a Hölder continuous function then p is a real-analytic function of v.

Proof. For (a), (b), (c) see [53], Cor. 3.3(i) (and [11], Theorem 2, for the proof of (c) in the two-sided case).

For (d), since f is cohomologically full then $v.f$ and $v'.f$ are not essentially cohomologous. By Corollary 3.3(ii) of [53] this ensures their equilibrium states are distinct.

(e) follows easily from (d).

The proof of (f) uses the fact that if f is Hölder then the Ruelle transfer operator associated to $v.f$ has $e^{p(v)}$ as a simple isolated maximal eigenvalue (see [49]). The result then follows from standard analytic perturbation theory (see Appendix V of [42]).

□

Remark 1. Mere continuity of f does not ensure $\mathcal{M}_f(v)$ is a singleton (see [49], Cor. 3.17(b), and also [10], [19]).

Combining Corollary 3 with Lemma 1 gives

Corollary 4. *Let (X,T) be a mixing subshift of finite type. Suppose $f : X \to \mathbb{R}^d$ is a cohomologically full function with summable variation. Then $f_*(\mathcal{M}) \subset \mathbb{R}^d$ has interior, and $int(f_*(\mathcal{M})) = f_*(\mathcal{M}_f(\mathbb{R}^d))$.*

5. The Bousch-Mañé cohomology lemma

Definition 4. Given a dynamical system (X,T) and $g : X \to \mathbb{R}$ a continuous function, we say a measure $m \in \mathcal{M}$ is *g-optimal* if $\int g\,dm = \sup_{\mu\in\mathcal{M}} \int g\,d\mu$. We call $Q(g) = \sup_{\mu\in\mathcal{M}} \int g\,d\mu$ the *optimal ergodic average* for g.

The weak* compactness of $\mathcal{M}$ ensures a g-optimal measure exists, but in general it need not be unique. We now present a lemma which characterises optimal measures. The result is due to Bousch [9], who proved it for T the doubling map of the circle, and g Lipschitz continuous. Our proof is adapted from his. Bousch attributed the lemma to Mañé, who had obtained a similar result, by different methods, in the context of Lagrangian flows [33]. The support of a measure $m \in \mathcal{M}$, denoted by *supp*(m), is compact and T-invariant.

Lemma 2. *Let (X,T) be a subshift of finite type. Let $g : X \to \mathbb{R}$ be of summable variation. A measure $m \in \mathcal{M}$ is g-optimal if and only if there exists $\varphi \in C(X)$ such that*

$$\varphi(Tz) + Q(g) = \varphi(z) + g(z) \quad \textit{for all } z \in \textit{supp}(m). \tag{7}$$

Furthermore, if X is one-sided and $g \in \mathcal{F}_\theta(X)$ then any such φ also belongs to $\mathcal{F}_\theta(X)$. If X is two-sided and $g \in \mathcal{F}_\theta(X)$ then any such φ belongs to $\mathcal{F}_{\sqrt{\theta}}(X)$.

Proof. We will prove the result in the one-sided case, and then use the construction described in section 4 to deduce it for the two-sided case. Let $g : X \to \mathbb{R}$ be of summable variation. Define the (non-linear) operator $M_g : C(X) \to C(X)$ by

$$M_g(\varphi)(x) = \max_{y\in T^{-1}x} (\varphi(y) + g(y)).$$

Note $M_g(\varphi+c) = M_g(\varphi)+c$ for $c \in \mathbb{R}$, so that M_g is well-defined on the quotient space $C'(X)$. We will show that M_g has a fixed point when acting on $C'(X)$.

Let Λ denote the set of $\Phi \in C'(X)$ for which $\mathrm{var}_n(\Phi) \le \sum_{j=n+1}^\infty \mathrm{var}_j(g)$ for all $n \ge 0$. Note immediately that Λ is convex and uniformly closed. Λ is equicontinuous since if $\pi_n(x) = \pi_n(y)$ then for all $\Phi \in \Lambda$,

$$|\Phi(x) - \Phi(y)| \le \mathrm{var}_n(\Phi) \le \sum_{j=n+1}^\infty \mathrm{var}_j(g) \to 0 \text{ as } n \to \infty.$$

Using (6) we see Λ is uniformly bounded since for all $\Phi \in \Lambda$,

$$|\Phi|_\infty = \frac{1}{2}\mathrm{var}_0(\Phi) \le \frac{1}{2}\sum_{j=1}^\infty \mathrm{var}_j(g).$$

By the Ascoli-Arzela theorem we deduce Λ is uniformly compact.

Next we will show that $M_g(\Lambda) \subset \Lambda$. Suppose $\pi_n(x) = \pi_n(y)$. Given $\Phi \in \Lambda$, let the symbols $1 \le i_x, i_y \le k$ be such that $M_g\Phi(x) = (g+\Phi)(i_x x)$ and $M_g\Phi(y) = (g+\Phi)(i_y y)$ (note i_x, i_y need not be unique). Here for notational convenience we write g to denote its equivalence class in $C'(X)$. Since X is of finite type, and $\pi_n(x) = \pi_n(y)$, then we know $i_x y \in X$ and $i_y x \in X$, so we have

$$\begin{aligned} M_g(\Phi)(x) - M_g(\Phi)(y) &= (g+\Phi)(i_x x) - (g+\Phi)(i_y y) \\ &= (g+\Phi)(i_x x) - (g+\Phi)(i_y x) + (g+\Phi)(i_y x) - (g+\Phi)(i_y y) \\ &\ge -\mathrm{var}_{n+1}(g+\Phi). \end{aligned}$$

Similarly we can show $M_g(\Phi)(x) - M_g(\Phi)(y) \le \mathrm{var}_{n+1}(g+\Phi)$, and hence that

$$\begin{aligned} |M_g\Phi(x) - M_g\Phi(y)| \quad &\le \mathrm{var}_{n+1}(g+\Phi) \\ &\le \mathrm{var}_{n+1}(g) + \mathrm{var}_{n+1}(\Phi) \\ &\le \sum_{j=n+1}^{\infty} \mathrm{var}_j(g). \end{aligned}$$

Therefore $\mathrm{var}_n(M_g(\Phi)) \le \sum_{j=n+1}^{\infty} \mathrm{var}_j(g)$ so that $M_g(\Phi) \in \Lambda$, as required.

Thus M_g is a continuous map from the compact convex set Λ to itself, and hence by the Schauder-Tychonov theorem it admits a fixed point $\Phi \in \Lambda \subset C'(X)$.

That is, there exists $\varphi \in C(X)$ and $C_g \in \mathbb{R}$ such that $M_g(\varphi) = \varphi + C_g$. That is,

$$\varphi(x) + C_g = \max_{y \in T^{-1}x} \varphi(y) + g(y). \tag{8}$$

Now let $z \in X$ be arbitrary. Substituting Tz for x in (8) we obtain

$$\varphi(Tz) + C_g = \max_{y \in T^{-1}(Tz)} \varphi(y) + g(y) = \varphi(z) + g(z) + r(z) \tag{9}$$

where

$$r(z) := \max_{y \in T^{-1}(Tz)} (\varphi(y) + g(y)) - (\varphi(z) + g(z))\,.$$

Note that $r \ge 0$. Consider the set $r^{-1}(0)$ of zeros of r. It is certainly non-empty, since any $z \in X$ has at least one pre-image in $r^{-1}(0)$. Indeed this observation ensures any finite intersection $\cap_{n=0}^{N} T^{-n}(r^{-1}(0))$ is non-empty, and hence that $\cap_{n=0}^{\infty} T^{-n}(r^{-1}(0))$ is also non-empty. Moreover, this subset of $r^{-1}(0)$ is T-invariant, and therefore supports at least one T-invariant probability measure. That is, there exists at least one $m \in \mathcal{M}$ such that $supp(m) \subset r^{-1}(0)$.

Now let $\mu \in \mathcal{M}$ be arbitrary. Integrating (9) with respect to μ gives

$$-\int r\,d\mu = \int g\,d\mu - C_g. \tag{10}$$

Now the left-hand side of (10) is non-positive, since $r \ge 0$. Therefore we deduce that $\int g\,d\mu \le C_g$ for all $\mu \in \mathcal{M}$. However if now $m \in \mathcal{M}$ satisfies $supp(m) \subset r^{-1}(0)$ then the left-hand side of (10) is zero, so that $\int g\,dm = C_g$. Therefore we have that $C_g = Q(g)$, the optimal ergodic average for g.

We have proved that $\int g\,dm = Q(g)$ if and only if $supp(m) \subset r^{-1}(0)$, which holds if and only if

$$\varphi(Tz) + Q(g) = \varphi(z) + g(z) \quad \text{for all } z \in supp(m).$$

That is, m is g-optimal if and only if (7) holds.

Note that (8) gives $\mathrm{var}_n\varphi \le \mathrm{var}_{n+1}\varphi + \mathrm{var}_{n+1}g$. If $g \in \mathcal{F}_\theta(X)$ then iterating this inequality we see that $\mathrm{var}_n\varphi \le \sum_{m=1}^{\infty} \mathrm{var}_{n+m} g \le \sum_{m=1}^{\infty} |g|_\theta \theta^{n+m} = \frac{\theta|g|_\theta}{1-\theta}\theta^n$, so that in fact $\varphi \in \mathcal{F}_\theta(X)$.

If X is two-sided we first find a function $g^+ \in \mathcal{F}_{\sqrt{\theta}}$ which is cohomologous to g and depends only on future coordinates, and then apply the above argument. □

The following are immediate consequences of the Bousch-Mañé lemma.

Corollary 5. *Let (X,T) be a subshift of finite type, and $g : X \to \mathbb{R}$ be of summable variation. Suppose $\mu \in \mathcal{M}$ is g-optimal, and $m \in \mathcal{M}$ satisfies $supp(m) \subset supp(\mu)$. Then m is also g-optimal.*

Corollary 6. *Let (X,T) be a subshift of finite type, and $g : X \to \mathbb{R}$ be of summable variation. If there is a unique g-optimal measure $\mu \in \mathcal{M}$, then the restriction of T to $supp(\mu)$ is uniquely ergodic.*

6. An entropy function

Definition 5. Given a dynamical system (X,T) we define the *entropy function H* on $f_*(\mathcal{M})$ by

$$H(\varrho) = \sup\{h(\mu) : f_*(\mu) = \varrho\}.$$

Note immediately that H is concave, and therefore continuous, since h is upper semi-continuous. In section 8 we will define a slightly different entropy function $\mathcal{H}$, the *directional entropy*, which in general will be discontinuous at the boundary of the rotation set.

Theorem 2. *Let (X,T) be a dynamical system for which the entropy map is upper semi-continuous, and suppose $f : X \to \mathbb{R}^d$ is continuous. Let $\varrho \in ri(f_*(\mathcal{M}))$.*
(a) There exists $v \in \mathbb{R}^d$ for which $f_*^{-1}(\varrho) \cap \mathcal{M}_f(\mathbb{R}^d) \subset \mathcal{M}_f(v)$
(b) A measure $m \in f_*^{-1}(\varrho)$ satisfies $h(m) = H(\varrho)$ if and only if $m \in f_*^{-1}(\varrho) \cap \mathcal{M}_f(\mathbb{R}^d)$.

Proof. Note that $f_*^{-1}(\varrho) \cap \mathcal{M}_f(\mathbb{R}^d)$ is non-empty by Corollary 2.

(a) Suppose (a) is false. Then there exists $v \neq v'$, and a measure m which belongs to $\mathcal{M}_f(v)$ but not to $\mathcal{M}_f(v')$, and a measure m' which belongs to $\mathcal{M}_f(v')$ but not to $\mathcal{M}_f(v)$, but such that $m, m' \in f_*^{-1}(\varrho)$. This implies that $\int v.f \, dm = \int v.f \, dm'$, but the fact that m is an equilibrium state for $v.f$, yet m' is not, gives $h(m) > h(m')$. Similarly we have $\int v'.f \, dm = \int v'.f \, dm'$, and now the fact that m' is an equilibrium state for $v'.f$, yet m is not, gives $h(m') > h(m)$. This is a contradiction.

(b) Let $m \in f_*^{-1}(\varrho) \cap \mathcal{M}_f(\mathbb{R}^d)$, and let $v \in \mathbb{R}^d$ be such that $m \in \mathcal{M}_f(v)$. Now let μ be some measure in $f_*^{-1}(\varrho)$ which does not belong to the family $\mathcal{M}_f(\mathbb{R}^d)$. Then $\int v.f \, dm = \int v.f \, d\mu$, but the fact that m is an equilibrium state for $v.f$, whereas μ is not, gives $h(m) > h(\mu)$. □

With a tighter restriction on the dynamics we have the following sharper result.

Theorem 3. *Let (X,T) be a mixing subshift of finite type, and suppose $f : X \to \mathbb{R}^d$ is cohomologically full and has summable variation. Let $\varrho \in int(f_*(\mathcal{M}))$. Then*
(a) $f_*^{-1}(\varrho)$ intersects $\mathcal{M}_f(\mathbb{R}^d)$ at a single measure m,
(b) m is the unique measure in $f_*^{-1}(\varrho)$ satisfying $h(m) = H(\varrho)$.

Proof.
(a) From Theorem 2 we know there exists $v \in \mathbb{R}^d$ for which $f_*^{-1}(\varrho) \cap \mathcal{M}_f(\mathbb{R}^d) \subset \mathcal{M}_f(v)$. Now suppose $v \neq v'$ satisfies $f_*^{-1}(\varrho) \cap \mathcal{M}_f(\mathbb{R}^d) \subset \mathcal{M}_f(v')$. But by Lemma 1 we know $v.f$, $v'.f$ have distinct equilibrium states. So the argument used to prove part (a) of Theorem 2 again provides a contradiction.

Thus there exists a unique $v \in \mathbb{R}^d$ for which $f_*^{-1}(\varrho) \cap \mathcal{M}_f(\mathbb{R}^d) \subset \mathcal{M}_f(v)$. But since f has summable variation then $\mathcal{M}_f(v)$ contains a single measure, by Lemma 1.
(b) This follows immediately from part (a) above, and part (b) of Theorem 2. □

Remark 2. If f is locally constant (the case treated by [14], [59]), then (after a standard re-coding to ensure f only depends on two coordinates) each equilibrium state $m_{v.f}$ is a Markov measure, whose explicit form is given in [42] p. 27. Consequently there are simple formulae for both $f_*(m_{v.f})$ and $h(m_{v.f})$ (see [55] p. 103), and we can obtain an explicit formula for the entropy function H. For example suppose X is the full shift on the alphabet $\{0,1\}$, and $f = (\chi_{[1]}, \chi_{[11]})$ (i.e. the coordinate functions are characteristic functions of the cylinders [1] and [11]). A computation gives that the rotation set $f_*(\mathcal{M})$ is the triangle with vertices at $(0,0)$, $(0,\frac{1}{2})$, and $(1,1)$. The entropy function is

$$\begin{aligned} H(\varrho_1,\varrho_2) &= \varrho_1\log\varrho_1 + (1-\varrho_1)\log(1-\varrho_1) - \varrho_2\log\varrho_2 - 2(\varrho_1-\varrho_2)\log(\varrho_1-\varrho_2) \\ &\quad -(1-2\varrho_1+\varrho_2)\log(1-2\varrho_1+\varrho_2). \end{aligned}$$

7. Directional and lost measures

Geller & Misiurewicz [15] introduced the following concepts, the terminology once more inspired by the notion of rotation sets for torus maps.

Definition 6. An ergodic measure $\mu \in \mathcal{M}$ is called *directional* if $f_*(m) = f_*(\mu)$ for all $m \in \mathcal{M}$ satisfying $supp(m) \subset supp(\mu)$. An ergodic measure $\mu \in \mathcal{M}$ is called *lost* if it is not directional.

Proposition 3. *Let (X,T) be a dynamical system with upper semi-continuous entropy map, and suppose $f : X \to \mathbb{R}^d$ is continuous. Suppose $\mu \in \mathcal{M}$. Then μ is directional if either of the two following conditions hold*

(a) for all i, the restricted coordinate function $f_i|_{supp(\mu)}$ is an essential coboundary for the dynamical system $(supp(\mu), T|_{supp(\mu)})$,

(b) the restriction of T to $supp(\mu)$ is uniquely ergodic.

Proof. Clear from the definitions. □

Proposition 4. *Let (X,T) be a mixing subshift of finite type, and $f : X \to \mathbb{R}^d$ be of summable variation, with at least one of its coordinate functions not an essential coboundary. Then*

(a) Any fully supported ergodic measure $\mu \in \mathcal{M}$ is lost,

(b) Any f-equilibrium state $m_{v.f} \in \mathcal{M}_f(\mathbb{R}^d)$ is lost.

Proof. (a) We must show there exists some ergodic measure $m \in \mathcal{M}$ with $f_*(m) \neq f_*(\mu)$. If this were *not* true then in particular every atomic measure $m \in \mathcal{M}$ would satisfy $f_*(m) = f_*(\mu)$. Then a theorem of Livsic [29] (the proof in [42], Prop. 3.7 is easily adapted to the case of functions of summable variation), would imply every coordinate function f_i is an essential coboundary, a contradiction.

(b) This is because every equilibrium state corresponding to a function of summable variation is fully supported, by Lemma 1. □

Proposition 5. *Let (X,T) be a subshift of finite type, $f : X \to \mathbb{R}^d$ be of summable variation, and ϱ an exposed point of the rotation set $f_*(\mathcal{M})$. Then every ergodic measure in the rotation class $f_*^{-1}(\varrho)$ is directional, and $f_*^{-1}(\varrho)$ contains at least one ergodic (and hence directional) measure.*

Proof. (a) Since ϱ is an exposed point of $f_*(\mathcal{M})$ then there exists some normal $v \in \mathbb{R}^d$ such that the supporting hyperplane $\{w \in \mathbb{R}^d : w.v = w.\varrho\}$ intersects $f_*(\mathcal{M})$ at the single point ϱ. Thus if $\mu \in f_*^{-1}(\varrho)$ then μ is $v.f$-optimal, and for any other $\nu \in \mathcal{M}$ which is $v.f$-optimal we have $f_*(\nu) = \varrho = f_*(\mu)$.

Now by Corollary 5 any $m \in \mathcal{M}$ with $supp(m) \subset supp(\mu)$ is also $v.f$-optimal, and hence $f_*(m) = f_*(\mu)$.

(b) Since ϱ is exposed in $f_*(\mathcal{M})$ it is certainly extremal in $f_*(\mathcal{M})$. Let $\mu \in \mathcal{M}$ be an extremal point of the convex set $f_*^{-1}(\varrho)$. Suppose μ is not ergodic. Then there exist distinct measures $\mu_1, \mu_2 \in \mathcal{M}$ and $\alpha \in (0,1)$ such that

$$\mu = \alpha\mu_1 + (1-\alpha)\mu_2. \tag{11}$$

Thus $f_*(\mu) = \alpha f_*(\mu_1) + (1-\alpha) f_*(\mu_2)$, and extremality of $\varrho = f_*(\mu)$ implies $f_*(\mu_1) = \varrho = f_*(\mu_2)$. But now since $\mu_1, \mu_2 \in f_*^{-1}(\varrho)$, (11) implies μ is not extremal in $f_*^{-1}(\varrho)$, a contradiction. Thus μ must be ergodic. Hence μ is directional, by (a). □

8. Directional entropy

The following entropy function was defined by Geller & Misiurewicz [15].

Definition 7. For $\varrho \in f_*(\mathcal{M})$ the *directional entropy* $\mathcal{H}(\varrho)$ at the point ϱ is given by

$$\mathcal{H}(\varrho) = \sup\left\{h(\mu) : \mu \in f_*^{-1}(\varrho) \text{ is directional}\right\},$$

where we define the supremum of the empty set to be zero.

Note straightaway that $\mathcal{H}(\varrho) \le H(\varrho)$. In this section, for (X,T) a mixing subshift of finite type, $f : X \to \mathbb{R}^d$ cohomologically full and of summable variation, we will prove that in fact $\mathcal{H}$ and H coincide on large portions of $f_*(\mathcal{M})$. Indeed we will see that the only rotation vectors ϱ where the equality $\mathcal{H}(\varrho) = H(\varrho)$ does *not* always hold are those points on the boundary of the rotation set which are non-exposed. If the rotation set $f_*(\mathcal{M})$ is *strictly* convex then the set of non-exposed boundary points is empty, so that $\mathcal{H}$ and H will completely coincide (Corollary 7).

Theorem 4. *Let (X,T) be a subshift of finite type, and suppose $f : X \to \mathbb{R}^d$ has summable variation. If ϱ is an exposed point of the rotation set then there exists at least one directional measure $m \in f_*^{-1}(\varrho)$ with $h(m) = H(\varrho)$. Consequently the directional entropy $\mathcal{H}(\varrho)$ is equal to the entropy $H(\varrho)$ at all exposed points ϱ of $f_*(\mathcal{M})$.*

Proof. Let ϱ be an exposed point of $f_*(\mathcal{M})$. Clearly $H(\varrho) \ge \mathcal{H}(\varrho)$, so it remains to show $H(\varrho) \le \mathcal{H}(\varrho)$. Since by Proposition 5 all ergodic measures in $f_*^{-1}(\varrho)$ are directional, it will suffice to show there exists some ergodic measure $\mu \in f_*^{-1}(\varrho)$ with $h(\mu) = H(\varrho)$.

Now $f_*^{-1}(\varrho)$ is closed, and hence compact, so that there exists $m \in f_*^{-1}(\varrho)$ with $h(m) = H(\varrho)$. But extremality of ϱ means that almost every measure in the ergodic decomposition of m must belong to $f_*^{-1}(\varrho)$, so since entropy h is affine then there is some ergodic (and hence directional) $\mu \in f_*^{-1}(\varrho)$ with $h(\mu) \ge h(m) = H(\varrho)$. Thus $\mathcal{H}(\varrho) \ge H(\varrho)$. □

Now we turn to the interior of the rotation set. For notational convenience let us assume in the rest of this section that our subshift of finite type $X = X_A^+$ is one-sided. The necessary amendments for the two-sided case are routine. First we introduce some notation.

For $x \in X$ we call $\lim_{n\to\infty} \frac{1}{n} f^n(x)$ the *rotation vector* of x, provided this limit exists.

For $r > 0$ and $v \in \mathbb{R}^d$, let $B_r(v) = \{u \in \mathbb{R}^d : |u - v| < r\}$ denote the open ball of radius r.

By a length-n *symbolic block* we mean an element $b = (b_1, \ldots, b_n) \in \pi_n(X)$. We say this block is *periodic* if $A(b_n, b_1) = 1$. The *concatenation* of two symbolic blocks $b \in \pi_n(X), b' \in \pi_{n'}(X)$, where $A(b_n, b'_1) = 1$, is defined to be the element $bb' = (b_1, \ldots, b_n, b'_1, \ldots, b'_n) \in \pi_{n+n'}(X)$. The concatenation of more than two blocks is defined inductively in the obvious way. For $b \in \pi_n(X)$ we define its *follower set* $F(b)$ as the set of $x \in X$ satisfying $A(b_n, x_1) = 1$. If the block $b \in \pi_n(X)$ is periodic then it can be concatenated with itself. We let $\overline{b} \in X$ denote the infinite concatenation $bbb\ldots$, and note that this is a periodic point of X. There is a unique measure $\mu \in \mathcal{M}$ supported on the periodic orbit generated by $\overline{b}$, so we define the *rotation vector* $f_*(b)$ *of the block* b to be the rotation vector $f_*(\mu)$ of the corresponding measure. We call such measures *periodic orbit measures.* If the support of a periodic orbit measure has cardinality n, then each point in its support is periodic of least period n, and is the infinite concatenation of some length-n periodic block. Thus to each periodic orbit measure we can associate n distinct periodic blocks.

For a mixing subshift of finite type, the collection of periodic orbit measures is weak* dense in $\mathcal{M}$ (see [12], p. 196). Let $a \in \{1, \ldots, k\}$, and consider the collection of those periodic orbit measures with at least one associated block beginning with the symbol a. A simple modification of the proof in [12], p. 196, gives that this collection is also weak* dense in $\mathcal{M}$. The continuity of f then implies that the set of rotation vectors of periodic blocks beginning with symbol a is dense in $f_*(\mathcal{M})$. Together with the compactness of $f_*(\mathcal{M})$, this means that for any $\varepsilon > 0$ there is a finite collection C of periodic blocks, each block beginning with symbol a, such that

$$f_*(\mathcal{M}) \subset \bigcup_{c \in C} B_\varepsilon(f_*(c)).$$

That is, the finite set $\{f_*(c) : c \in C\}$ is ε-dense in $f_*(\mathcal{M})$. By taking m-fold self-concatenations of appropriate elements of C, for varying m, we may assume all blocks in C have a common length M. The above discussion is summarised in the following Lemma.

Lemma 3. *Let X be a mixing subshift of finite type, with alphabet $\{1, \ldots, k\}$. Let $f : X \to \mathbb{R}^d$ be continuous. For $\varepsilon > 0$, $a \in \{1, \ldots, k\}$, and for arbitrarily large $M \in \mathbb{N}$, there exists a finite collection C of length-M periodic blocks, each beginning with the symbol a, whose rotation vectors are ε-dense in $f_*(\mathcal{M})$.*

Our definition of $H(\varrho)$ is variational. However, Takens & Verbitski [52] prove that there is an equivalent definition of $H(\varrho)$ in terms of the growth rate of certain (n, ε)-separated sets. A minor modification of their proof allows us to define $H(\varrho)$ as the growth rate of the number of periodic blocks beginning with symbol a, and with rotation vector close to ϱ. More precisely, we have the following version of Theorem 5.1 in [52]:

Lemma 4. *Let X be a mixing subshift of finite type, with alphabet $\{1, \ldots, k\}$. Let $f : X \to \mathbb{R}^d$ be continuous. For any $a \in \{1, \ldots, k\}$, $\varrho \in f_*(\mathcal{M}), r > 0, n \in \mathbb{N}$, let $M(\varrho, r, n)$ be the number of length-n periodic blocks beginning with symbol a whose rotation vectors lie in $B_r(\varrho)$. Then*

$$H(\varrho) = \lim_{r \to 0} \limsup_{n \to \infty} \frac{1}{n} \log M(\varrho, r, n).$$

Theorem 5. *Let X be a mixing subshift of finite type. Let $f : X \to \mathbb{R}^d$ be of summable variation, and cohomologically full. For $\varrho \in int(f_*(\mathcal{M}))$ we have that*

$\mathcal{H}(\varrho) = H(\varrho)$. That is, the directional entropy $\mathcal{H}$ coincides with the entropy function H on the interior of the rotation set.

Proof. Let $\varrho \in int(f_*(\mathcal{M}))$. Given any $\alpha > 0$ we will construct a subshift $Y = Y_\alpha$ with topological entropy $h_{top}(Y_\alpha) > H(\varrho) - \alpha$, and such that *every* $y \in Y_\alpha$ has rotation vector ϱ. Consequently any ergodic measure supported on Y_α will be directional, and in the fibre $f_*^{-1}(\varrho)$. Moreover, the variational principle (see Walters [54], [55]) means we can find an ergodic measure μ_α, with $supp(\mu_\alpha) \subset Y_\alpha$, such that $h(\mu_\alpha) > H(\varrho) - \alpha$.

Since $\alpha > 0$ was arbitrary we deduce that

$$\mathcal{H}(\varrho) \geq \sup_{\alpha>0} h(\mu_\alpha) \geq \sup_{\alpha>0} (H(\varrho) - \alpha) = H(\varrho),$$

and hence that $\mathcal{H}(\varrho) = H(\varrho)$.

The *idea* of the proof is as follows. We will construct the subshift $Y = Y_\alpha$ by an inductive procedure, ensuring the strong condition that

$$\sup_{y\in Y} |f^n(y) - n\varrho| = O(1) \text{ as } n \to \infty. \tag{12}$$

In fact Y will be the orbit closure of a certain set $Z \subset X$. To construct Z we will introduce two finite collections B, C of symbolic blocks, where all blocks in both collections will begin with some fixed symbol, a say. All blocks in C will be of the same length M. All blocks in B will be of the same length N, where N is a large integer multiple of M. We will think of the collection B as an *alphabet*, where each block $b \in B$ has rotation vector close to ϱ, and the cardinality $|B|$ is large enough to ensure Y has high topological entropy.

To each $\underline{b} = (b_i) \in B^{\mathbb{N}}$ we will identify a collection $Z(\underline{b})$ of points $x \in X$ of the form $x = b_1c_1b_2c_2\ldots$, where $c_i \in C$. The fact that all blocks in B and C begin with the same symbol a ensures that any such x does belong to our subshift of finite type X. Given $\underline{b}$, the j^{th} block c_j will depend on the $b_i, i \leq j$, and $c_i, i < j$, and will be chosen to maintain $|f^n(x) - n\varrho| = O(1)$ (rather than the $|f^n(x) - n\varrho| = O(n)$, which would occur if x were constructed solely from blocks in B). We will think of the blocks $c \in C$ as being *corrective blocks*, since the c_i are chosen so as to reduce the distances $|f^n(x) - n\varrho|$. For any $\underline{b} \in B^{\mathbb{N}}$ we will show there exists at least one x of the above form, so that $Z(\underline{b})$ is non-empty. The set Z will be the union of the $Z(\underline{b})$.

Now we come to the details of the proof.

Step 1. Fix various constants in terms of ϱ and α.

Fix some $q \in \mathbb{N}$ such that

$$q > \frac{2(H(\varrho) - \alpha)}{\alpha}. \tag{13}$$

Note that q will be the *ratio* of the lengths of the (large) *alphabet* blocks in B and the (smaller) *corrective* blocks in C.

Since ϱ is an interior point of $f_*(\mathcal{M})$, we can choose $r > 0$ such that

$$B_{3rq}(\varrho) \subset f_*(\mathcal{M}). \tag{14}$$

Fix $\varepsilon > 0$ such that

$$\varepsilon < rq. \tag{15}$$

Now choose a natural number N (which will be the common length of our alphabet blocks) which simultaneously satisfies the following four conditions:

(a) N is an integer multiple of q,

(b) $V(f)/N < r - \varepsilon/q$,

(c) There is a collection C of periodic blocks, each of length $M := N/q$, each beginning with the same symbol a, whose rotation vectors are ε-dense in $f_*(\mathcal{M})$,

(d) There is a collection B of length-N periodic blocks, of cardinality $|B| \geq e^{(H(\varrho)-\alpha/2)N}$, such that each $b \in B$ begins with the symbol a and satisfies $f_*(b) \in B_r(\varrho)$.

Note that to satisfy (a)-(d) above, N must merely be a sufficiently large multiple of q. Lemma 3 and Lemma 4 guarantee that the conditions (c), (d) can be satisfied.

Step 2. Construction of points x^j.

Let $n_0 = 0$ and $n_j = j(N+M) = jN(1+1/q)$ for $j \geq 1$.

Given $\underline{b} = (b_i) \in B^{\mathbb{N}}$, we will show it is always possible to choose $(c_i) \in C^{\mathbb{N}}$, such that for the sequence of points $x^j = b_1 c_1 \dots b_j c_j \overline{b_{j+1}} \in X$, the length-$n_i$ Birkhoff sums $f^{n_i}(x^j)$, for $i = 0, \dots, j$, are some bounded (independent of both j and $\underline{b}$) distance from $n_i \varrho$. The points x^j will converge to $x = b_1 c_1 b_2 c_2 \dots \in X$, and we will similarly control the Birkhoff sums $f^n(x)$, for all $n \in \mathbb{N}$.

More precisely, we claim that for all $\underline{b} = (b_i) \in B^{\mathbb{N}}$ there exists a (not necessarily unique) $(c_i) \in C^{\mathbb{N}}$ such that for all $j \geq 0$, the point $x^j = b_1 c_1 \dots b_j c_j \overline{b_{j+1}}$ satisfies

$$|f^{n_i}(x^j) - n_i \varrho| < rN \text{ for all } i = 0, \dots, j. \tag{16}$$

Now (16) is trivially true for $j = 0$, so to prove (16) for general j it will suffice to show:

Claim. If $(b_i) \in B^{\mathbb{N}}$, and if $(c_i)_{i=1}^{j-1} \in C^{j-1}$ satisfies

$$|f^{n_i}(b_1 c_1 \dots b_j c_j u) - n_i \varrho| < rN \text{ for } u \in F(c_j), i = 0, \dots, j, \tag{17}$$

then there exists $c_{j+1} \in C$ such that

$$|f^{n_i}(b_1 c_1 \dots b_j c_j b_{j+1} c_{j+1} v) - n_i \varrho| < rN \text{ for } v \in F(c_{j+1}), i = 0, \dots, j+1. \tag{18}$$

Proof of Claim. Taking $u = b_{j+1} c v$ for any choice of $c \in C$, $v \in F(c)$ immediately gives us (18) for $i = 0, \dots, j$, by virtue of (17). So it remains to prove (18) holds for $i = j+1$.

Setting $i = j$ in (17) gives us

$$|f^{n_j}(b_1 c_1 \dots b_j c_j b_{j+1} c v) - n_j \varrho| < rN \tag{19}$$

Now (d) gives $|f^N(\overline{b_{j+1}}) - N\varrho| < rN$, and (5) gives

$$|f^N(\overline{b_{j+1}}) - f^N(b_{j+1} c v)| \leq V(f),$$

so combining these gives

$$|f^N(b_{j+1} c v) - N\varrho| < rN + V(f). \tag{20}$$

Combining (19), (20), and observing that

$$f^{n_j+N}(b_1 c_1 \dots b_j c_j b_{j+1} c v) = f^{n_j}(b_1 c_1 \dots b_j c_j b_{j+1} c v) + f^N(b_{j+1} c v),$$

we obtain

$$\left| f^{n_j+N}(b_1 c_1 \dots b_j c_j b_{j+1} c v) - (n_j + N)\varrho \right| < 2rN + V(f). \tag{21}$$

Now we want to compare the Birkhoff sum $f^{n_j+N}(b_1 c_1 \dots b_j c_j b_{j+1} c v)$ at time $(n_j + N)$ to the intended approximate location of the Birkhoff sum at time $n_{j+1} =$

n_j+N+M, namely the vector $n_{j+1}\varrho$. We must check whether it is possible, by a judicious choice of c, to (approximately) "join up" the vector $f^{n_j+N}(b_1c_1\dots b_jc_jb_{j+1}cv)$ to the vector $n_{j+1}\varrho$ in time $|c|=M=n_{j+1}-N$.

Now (21) implies that

$$\begin{aligned} n_{j+1}\varrho - &f^{n_j+N}(b_1c_1\dots b_jc_jb_{j+1}cv) \\ &= M\varrho + \left[(n_j+N)\varrho - f^{n_j+N}(b_1c_1\dots b_jc_jb_{j+1}cv)\right] \\ &\in B_{2rN+V(f)}(M\varrho), \end{aligned}$$

and "normalising by the required time M" we have that

$$\frac{n_{j+1}\varrho - f^{n_j+N}(b_1c_1\dots b_jc_jb_{j+1}cv)}{M} \in B_{(2rN+V(f))/M}(\varrho).$$

From (b) we know that $(2rN+V(f))/M = 2rq + V(f)q/N < 3rq$, so by (14) we deduce that, for all $c\in C$, $v\in F(c)$,

$$\frac{n_{j+1}\varrho - f^{n_j+N}(b_1c_1\dots b_jc_jb_{j+1}cv)}{M} \in f_*(\mathcal{M}). \tag{22}$$

By (c) we know the rotation vectors of the blocks in C are ε-dense in $f_*(\mathcal{M})$, so (22) means we can choose $c_{j+1}\in C$ such that

$$\left|\frac{(n_{j+1}\varrho - f^{n_j+N}(b_1c_1\dots b_jc_jb_{j+1}c_{j+1}v))}{M} - \frac{1}{M}f^M(\overline{c_{j+1}})\right| < \varepsilon. \tag{23}$$

Using (5) we estimate

$$\left|\frac{1}{M}f^M(\overline{c_{j+1}}) - \frac{1}{M}f^M(c_{j+1}v)\right| \le V(f)/M. \tag{24}$$

Combining (23), (24), multiplying by M, and observing that

$$f^{n_{j+1}}(b_1c_1\dots b_jc_jb_{j+1}c_{j+1}v) = f^{n_j+N}(b_1c_1\dots b_jc_jb_{j+1}c_{j+1}v)) + f^M(c_{j+1}v),$$

we derive

$$|f^{n_{j+1}}(b_1c_1\dots b_jc_jb_{j+1}c_{j+1}v) - n_{j+1}\varrho| < \varepsilon M + V(f). \tag{25}$$

Finally, by (b) we know $\varepsilon M + V(f) < rN$, so (25) becomes

$$|f^{n_{j+1}}(b_1c_1\dots b_jc_jb_{j+1}c_{j+1}v) - n_{j+1}\varrho| < rN.$$

This completes the proof of the Claim. To deduce (16) from the Claim we use induction. Since $x^j = b_1c_1\dots b_jc_ju$ with $u=\overline{b_{j+1}}$, and $x^{j+1} = b_1c_1\dots b_jc_jb_{j+1}c_{j+1}v$ with $v=\overline{b_{j+2}}$, the Claim tells us that *if* (16) is true for some j, then it is also true upon replacing j by $j+1$. But (16) is certainly true for $j=0$, thus by induction it is true for all $j\ge 0$.

Step 3. Construction of the subshift Y_α.

For $\underline{b}=(b_i)\in B^{\mathbb{N}}$, let x^j be a sequence constructed as in Step 2. Now (16) controls the Birkhoff sums at the particular instances of time n_j. However, since the gaps between the n_j are bounded (in fact constant) then we can easily control the Birkhoff sums at *all* time instances. Observing that $|f(w)-\varrho|\le 2|f|_\infty$ for all $w\in X$, and recalling that $n_j = j(N+M)$, we see that for all $j\ge 0$, and for all $n\le n_j$,

$$|f^n(x^j) - n\varrho| < rN + 2(N+M)|f|_\infty \tag{26}$$

The points x^j converge to $x = b_1c_1b_2c_2 \ldots \in X$. Indeed we have $\pi_n(x^j) = \pi_n(x)$ for all $n \le n_j$. By (5) this means that $|f^n(x) - f^n(x^j)| \le V(f)$ for all $n \le n_j$. Combining this with (26) gives that

$$\begin{aligned} |f^n(x) - n\varrho| &\le |f^n(x) - f^n(x^j)| + |f^n(x^j) - n\varrho| \\ &< V(f) + rN + 2(N+M)|f|_\infty. \end{aligned} \tag{27}$$

for all $j \ge 0$, $n \le n_j$, in other words for all $n \ge 0$.

For $\underline{b} = (b_i) \in B^{\mathbb{N}}$ we let $Z(\underline{b})$ denote the non-empty set of points x which are limits of sequences $x^j = b_1c_1 \ldots b_jc_j\overline{b_{j+1}}$ satisfying (16).

The right-hand side of (27) is independent of $\underline{b}$, so if we define

$$Z = \bigcup_{\underline{b} \in B^{\mathbb{N}}} Z(\underline{b}),$$

and set $K = V(f) + rN + 2(N+M)|f|_\infty$, then we obtain

$$|f^n(x) - n\varrho| < K \text{ for all } x \in Z, n \ge 0. \tag{28}$$

Moreover, if $x \in Z$, and $m, n \ge 0$, then (28) gives

$$\left| f^n(T^m x) - n\varrho \right| = \left| \left(f^{n+m}(x) - (n+m)\varrho\right) - \left(f^m(x) - m\varrho\right) \right| < 2K. \tag{29}$$

Now define

$$Y_\alpha = Y = \overline{\bigcup_{i \ge 0} T^i(Z)}.$$

Any $y \in Y$ is the limit of a sequence $T^{m_i}(z^i)$, for some $z^i \in Z$ and $m_i \ge 0$. Thus for $n \ge 0$ we can find $i \in \mathbb{N}$ such that $\pi_n(T^{m_i}z^i) = \pi_n(y)$, which by (5) then implies that $|f^n(y) - f^n(T^{m_i}z^i)| \le V(f)$. Combining this with (29) gives that for all $y \in Y$, and all $n \ge 0$,

$$|f^n(y) - n\varrho| \le |f^n(y) - f^n(T^{m_i}z^i)| + |f^n(T^{m_i}z^i) - n\varrho| < V(f) + 2K,$$

thus proving (12). In particular we see that each $y \in Y$ has rotation number ϱ, so that any ergodic measure supported on Y is directional.

Step 4. Estimating the topological entropy of Y_α.

We want to show that $h_{top}(Y_\alpha) > H(\varrho) - \alpha$. Indeed, since $|B| \ge e^{(H(\varrho)-\alpha/2)N}$, we can estimate the number $|\pi_{n_j}(Z)|$ of length-n_j blocks in Z, and then note that clearly $|\pi_{n_j}(Y_\alpha)| \ge |\pi_{n_j}(Z)|$. Therefore we have

$$\begin{aligned} |\pi_{n_j}(Y_\alpha)| \ge |\pi_{n_j}(Z)| &\ge |B|^j \\ &\ge e^{(H(\varrho)-\alpha/2)jN} \\ &= \exp\left[\frac{(H(\varrho)-\alpha/2)qn_j}{q+1}\right] \\ &> e^{(H(\varrho)-\alpha)n_j}, \end{aligned}$$

by our choice of q in (13).

It follows that

$$h_{top}(Y_\alpha) = \lim_{j\to\infty} \frac{1}{n_j} \log |\pi_{n_j}(Y_\alpha)| \ge H(\varrho) - \alpha.$$

By the variational principle [54], [55] we can find an ergodic $\mu \in \mathcal{M}$ supported on Y_α (hence directional and in the fibre $f_*^{-1}(\varrho)$) with entropy arbitrarily close to $H(\varrho) - \alpha$. Therefore $\mathcal{H}(\varrho) = H(\varrho)$. □

Proposition 6. *Let (X,T) be a mixing subshift of finite type, and $f : X \to \mathbb{R}^d$ be of summable variation and cohomologically full. If ϱ is in the interior of the rotation set $f_*(\mathcal{M})$ then there exists a unique measure $m \in f_*^{-1}(\varrho)$ with $h(m) = H(\varrho) = \mathcal{H}(\varrho)$, and this measure is lost.*

Proof. By Theorem 5 we know $\mathcal{H}(\varrho) = H(\varrho)$. By Theorem 3 there exists an f-equilibrium state $m \in \mathcal{M}_f(\mathbb{R}^d)$ such that $h(m) = H(\varrho)$, and m is the unique measure in this rotation class with this property. Moreover m is fully supported, by Lemma 1, and hence m is lost, by Proposition 4. □

Summarising the results of this section we have

Theorem 6. *Let (X,T) be a mixing subshift of finite type, and suppose $f : X \to \mathbb{R}^d$ is cohomologically full and has summable variation. Then $\mathcal{H}(\varrho) = H(\varrho)$ at all interior and exposed points ϱ of the rotation set $f_*(\mathcal{M})$. For ϱ in the interior of $f_*(\mathcal{M})$, $\mathcal{H}(\varrho) = H(\varrho)$ is attained by a unique measure in the rotation class $f_*^{-1}(\varrho)$, and this measure is lost. If ϱ is an exposed point of $f_*(\mathcal{M})$ then $\mathcal{H}(\varrho) = H(\varrho)$ is attained by at least one directional measure in $f_*^{-1}(\varrho)$, and is not attained by any lost measures in $f_*^{-1}(\varrho)$.*

Proof. This follows from Theorem 4, Theorem 5 and Proposition 6. □

Remark 3. The above theorem strengthens and generalises Corollary 5.2 and Proposition 5.4 of [15].

Corollary 7. *Let (X,T) be a mixing subshift of finite type, and suppose $f : X \to \mathbb{R}^d$ is cohomologically full with summable variation. If the rotation set $f_*(\mathcal{M})$ is strictly convex, then the functions $\mathcal{H}$ and H coincide.*

9. Rotation sets for locally constant functions

Let $X = X_A$ be a mixing subshift of finite type. We say $f : X \to \mathbb{R}^d$ is *locally constant* (or alternatively that f *depends on finitely many coordinates*) if $\mathrm{var}_n(\varphi) = 0$ for some n. Clearly such a function is of summable variation.

By *passing to a higher block presentation* of X (see [28] p. 12) we can consider any locally constant function as being constant on cylinders of length two (i.e. $\mathrm{var}_3(f) = 0$). Let us recall the details of this construction. Suppose f is constant on cylinders of length n, so that $f(x) = f(x_1, \dots, x_n)$ for all $x \in X$. We now define a new subshift of finite type Y whose alphabet consists of all allowed words of length $n-1$ in X, and where the only allowed transitions in Y are from a symbol $(x_1, \dots, x_{n-1})$ to a symbol $(x_2, \dots, x_n)$, where $(x_1, \dots, x_n)$ is an allowed word of length n for X. Then f can be considered as a function of the two variables $(x_1, \dots, x_{n-1})$, $(x_2, \dots, x_n)$ instead of the n variables $x_1, \dots, x_n$. The subshift of finite type Y is conjugate to X (where the conjugacy is given by the sliding block code defining the passage to a higher block presentation).

Let G be the *Markov graph* of (X,T). That is, G is the directed graph with vertices labelled by the alphabet B, and an arrow from i to j if and only if $A(i,j) = 1$. The transitivity of (X,T) means that for any two vertices i, j there is a path in G from i to j. A path $b_1, \dots, b_{n+1}$ in G is called a *loop* if $b_1 = b_{n+1}$, and can be identified with a periodic point $x \in X$ defined by $x_i = b_{i \pmod n}$. The corresponding periodic orbit supports a unique invariant measure $\mu \in \mathcal{M}$. We say a loop, and its corresponding periodic orbit, and the periodic points in the periodic orbit, are *elementary* if the loop is not the concatenation of two strictly smaller loops. There are finitely many elementary loops.

Any loop in G is formed by a finite number of concatenations of elementary loops. Any subset $L_1, \dots, L_s$ of elementary loops generates a directed graph G', which is a subgraph of G. This is the Markov graph for precisely one subshift of finite type, $X' \subset X$ say. We call X' the subshift of finite type *generated* by the loops $L_1, \dots, L_s$. Note that in general such subshifts of finite type need not be transitive, but the construction by closed loops ensures they are non-wandering.

The following result was proved by Ziemian [59].

Lemma 5. *Suppose (X,T) is a transitive subshift of finite type, and that $f : X \to \mathbb{R}^d$ is constant on cylinders of length two. Let $\mu_1, \dots, \mu_r$ be the invariant measures supported on the elementary periodic orbits of X. Then the rotation set $f_*(\mathcal{M})$ is a polyhedron whose extremal points are a subset of $f_*(\mu_1), \dots, f_*(\mu_r)$.*

Corollary 8. *Suppose (X,T) is a transitive subshift of finite type, and that $f : X \to \mathbb{R}^d$ is a locally constant function. Then the rotation set $f_*(\mathcal{M})$ is a polyhedron.*

Proof. Passing to a higher block presentation of X we can consider f as a function of two coordinates, and then use Lemma 5. □

We now consider the faces of rotation sets $f_*(\mathcal{M})$. For a given face F, the union of the supports of those measures whose rotation vectors lie in F is clearly a T-invariant subset of X. We now show that if f is locally constant then this invariant set is itself a subshift of finite type. This result is essentially proved (by a different method) in the article of Marcus & Tuncel [34], who defined the *weight-per-symbol polytope* of a Markov chain. This polytope is the rotation set of a locally constant function defined from the transition probabilities of the Markov chain.

Theorem 7. *Suppose (X,T) is a transitive subshift of finite type, and $f : X \to \mathbb{R}^d$ is constant on cylinders of length two. Let F be a face of the corresponding polyhedral rotation set $f_*(\mathcal{M})$, and $L_1, \dots, L_s$ those elementary loops of X whose rotation vectors $\varrho_1, \dots, \varrho_s$ lie in F. If $X_F \subset X$ denotes the non-wandering subshift of finite type generated by the loops $L_1, \dots, L_s$, then for any measure $\mu \in \mathcal{M}$,*

$$f_*(\mu) \in F \quad \textit{if and only if} \quad supp(\mu) \subset X_F.$$

Proof. Let $\mu_1, \dots, \mu_s$ be the invariant measures corresponding to the elementary loops $L_1, \dots, L_s$.

Let $v \in \mathbb{R}^d$ be orthogonal to the face F, and pointing outwards from $f_*(\mathcal{M})$. Then $f_*(\nu) \in F$ if and only if ν is an optimal measure for the function $g = v.f$. But by Lemma 2 this is true if and only if g is cohomologous to $Q(g)$ on $supp(\nu)$. Now f depends on two coordinates thus so does g, hence any cobounding function, $\varphi = \varphi_\nu$ say, must depend on one coordinate. So we have that $f_*(\nu) \in F$ if and only if

$$g(x_1, x_2) = \varphi(x_2) - \varphi(x_1) + Q(g) \text{ for all } x = (x_j) \in supp(\nu) \tag{30}$$

Suppose $y = (y_j) \in supp(\mu) \subset X_F$. Then (y_1, y_2) must appear as a substring in some elementary loop L_i, say, $1 \le i \le s$. In particular there is a periodic point $x = (x_j)$ in the periodic orbit corresponding to L_i such that $x_j = y_j$ for $j = 1, 2$.

Since $x \in supp(\mu_i)$, and $f_*(\mu_i) \in F$, then (30) holds for this x. Hence (30) holds with x replaced by y. Since $y \in supp(\mu)$ was arbitrary then $f_*(\mu) \in F$.

Conversely, suppose $y = (y_j) \in supp(\mu) \setminus X_F$. Shifting y if necessary we may suppose (y_1, y_2) does not appear as a substring of any of $L_1, \ldots, L_s$. However there must exist some elementary loop of X, call it L, in which (y_1, y_2) does appear. Then let $x = (x_j)$ be some periodic point in the periodic orbit corresponding to L such that $x_j = y_j$ for $j = 1, 2$. Note $x \in X \setminus X_F$. If m is the invariant measure corresponding to L then g is cohomologous to $f_*(m).v$ on $supp(m)$. Since $f_*(m) \notin F$ then $f_*(m).v < Q(g)$, so that (30) cannot hold for this x. Thus (30) cannot hold with x replaced by y. Thus μ is not g-optimal, and hence $f_*(\mu) \notin F$. $\square$

10. Non-exposed extremal points : examples

Let (X, T) be a mixing subshift of finite type, and $f : X \to \mathbb{R}^d$ be of summable variation. The results of the previous section classify the behaviour of $\mathcal{H}$ with respect to directional and lost measures at all interior and exposed points of the rotation set. In this section and the next we show that, by contrast, almost anything can happen at non-exposed points on the boundary of $f_*(\mathcal{M})$. All our examples are of functions $f : X \to \mathbb{R}^2$ of summable variation (indeed often locally constant). X is always a full shift, and apart from Example 2 and Example 7 can always be taken as the full shift on two symbols.

To obtain our examples we will repeatedly use certain constructions. Firstly, if δ is a metric on X which induces the product topology, and $Y \subset X$ is a subshift, then the function $g(x) = -\delta(x, Y)$ is Lipschitz continuous with respect to δ, and any g-optimal measures must have support contained in Y.

If our dynamical system (X, T) were smooth, we could also choose a smooth non-positive function g which vanishes on Y (indeed to obtain Example 2 we do precisely this, and then project down to a full shift). Of course this *method* would fail in the analytic category, though it is known that analytic functions can have their unique optimal measure supported on a Cantor set (see [9], [22]).

Secondly, if μ_1, μ_2 are distinct invariant measures we can certainly choose a continuous (indeed locally constant) function $\varphi : X \to \mathbb{R}$ with $\int \varphi \, d\mu_1 \neq \int \varphi \, d\mu_2$.

Thirdly, we will choose various subshifts $Y \subset X$ with a prescribed number of invariant measures. A convenient class of subshifts for realising this are Toeplitz subshifts (see for example [58]). These subshifts are always minimal, but can be very far from uniquely ergodic. Indeed any abstract simplex can arise as the set $\mathcal{M}_Y$ of invariant measures for a Toeplitz subshift Y [13]. In particular we will use that for any $n \in \mathbb{N}$, there exists a Toeplitz subshift (with positive entropy if necessary) with exactly n ergodic measures. Clearly if the rotation set of $g : Y \to \mathbb{R}$ is not a single point then all these ergodic measures are lost with respect to g.

Finally, we will also use a result of Grillenberger [17] that given any ergodic measure μ for some full shift, there is some *uniquely ergodic* measure μ' on the same full shift which has the same entropy as μ. By a uniquely ergodic measure μ' we mean that the dynamical system $(supp(\mu'), T)$ is uniquely ergodic in the usual sense (see [55], p. 158). That is, μ' is the only measure in $\mathcal{M}$ whose support is contained in $supp(\mu')$.

In this section we consider the possible behaviour at non-exposed extremal points. Such points ϱ must always contain an ergodic measure in their rotation class (indeed any extremal point of the rotation class is necessarily ergodic, by the argument used in part of the proof of Proposition 5), so the entropy function $H(\varrho)$ will always be realised by either a directional or a lost measure, or possibly both. We now show that indeed these three cases can all arise.

Example 1. *ϱ non-exposed extremal, $\mathcal{H}(\varrho) < H(\varrho)$, $H(\varrho)$ attained uniquely by a lost measure.*

Let $Y \subset X$ be some minimal subshift with precisely two ergodic measures μ_1, μ_2, both of positive entropy. Define $f_2(x) = -\delta(x, Y)$, and choose $f_1 : X \to \mathbb{R}$ such that $\int f_1\, d\mu_1 < \int f_1\, d\mu_2$. Let $f = (f_1, f_2)$. Then $\varrho_i = (\int f_1\, d\mu_i, 0)$, $i = 1, 2$, are extremal points of the rotation set $f_*(\mathcal{M})$, but the only ergodic measures in $f_*^{-1}(\varrho_i)$ are lost. Thus by Proposition 5 we immediately see that ϱ_i are non-exposed. Moreover $H(\varrho_i) > \mathcal{H}(\varrho_i) = 0$, and $H(\varrho_i)$ is attained by the entropy of the lost measure μ_i.

Example 2. *ϱ non-exposed extremal, $\mathcal{H}(\varrho) = H(\varrho)$, attained uniquely by a directional measure.*

Let (X, T) be the full shift on the four symbols $\{0, 1, 2, 3\}$, which we will identify, via 4-adic expansions, with the map $x \mapsto 4x \pmod 1$ on the unit interval. We will describe a C^1 function f_2 on the interval (indeed we can choose it C^∞) which will project to a Hölder function on X.

Let f_2 be a non-positive C^1 function which vanishes on the interval $[\frac{1}{3}, \frac{2}{3}]$ and on all points $\frac{1}{3} - \frac{1}{4^n}$, $n \geq 1$, but is strictly negative elsewhere. So projecting onto X, f_2 vanishes on the full shift on symbols $\{1, 2\}$, and all points $(x_j^{(n)})$ of the form $x_n^{(n)} = 0$, $x_j^{(n)} = 1$ for $j \neq n$.

Let $f_1(x) = x$ on the unit interval. Let $f = (f_1, f_2)$. Then the line segment from $(\frac{1}{3}, 0)$ to $(\frac{2}{3}, 0)$ lies on the boundary of the rotation set of f. In particular, the extremal point $\varrho = (\frac{1}{3}, 0)$ only contains one measure in its rotation class, namely the Dirac measure on the fixed point $\frac{1}{3}$ (or symbolically, the sequence consisting of all 1's). This measure is directional.

It remains to prove that ϱ is non-exposed. Consider the atomic measure μ_n supported on the period-n orbit $\frac{1}{3} - \frac{4^i}{4^n - 1}$, $i = 0, \ldots, n-1$. Clearly $f_*(\mu_n) \to (\frac{1}{3}, 0)$ as $n \to \infty$. The fact that f_2 is C^1 at its zeros means that for all $\varepsilon > 0$,

$$\int f_2\, d\mu_n > -\frac{\varepsilon}{4^n} \quad \text{for } n \text{ sufficiently large.}$$

By contrast we have

$$\int f_1 d\mu_n = \frac{1}{3} - \frac{1}{3n} \quad \text{for all } n.$$

Thus $\int f_2\, d\mu_n \to 0$ much faster than $\int f_1\, d\mu_n \to \frac{1}{3}$, so that the $f_*(\mu_n)$ approach ϱ at limiting angle 0 with the horizontal. Thus the only tangent to the rotation set at the point $\varrho = (\frac{1}{3}, 0)$ is the horizontal, but since this intersects $f_*(\mathcal{M})$ on a whole interval then ϱ is non-exposed.

Example 3. *ϱ non-exposed extremal, $\mathcal{H}(\varrho) = H(\varrho)$, attained by both a lost and a directional measure.*

Let $Y \subset X$ be a minimal subshift with precisely two ergodic measures μ_1, μ_2, both of the same entropy. Let $Y' \subset X$ be a uniquely ergodic subshift with invariant measure μ of the same entropy (see [17]) as μ_1, μ_2. Let $f_2(x) = -\delta(x, Y \cup Y')$, and choose f_1 such that $a := \int f_1\, d\mu = \int f_1\, d\mu_1 < \int f_1\, d\mu_2$. Let $f = (f_1, f_2)$. Then $\varrho = (a, 0)$ contains both the lost measure μ_1 and the directional measure μ in its rotation class $f_*^{-1}(\varrho)$. Thus $\mathcal{H}(\varrho) = H(\varrho) = h(\mu_1) = h(\mu)$. The point ϱ is extremal, and clearly non-exposed by Proposition 5.

11. Non-extremal boundary points : examples

In this section we will consider non-extremal points on the boundary of the rotation set,

Example 4. *ϱ non-extremal, $\mathcal{H}(\varrho) = H(\varrho)$ attained uniquely by a lost measure.*

Consider the function $f : X \to \mathbb{R}^2$ of Remark 2, with X the full shift on the alphabet $\{0, 1\}$. Here $f_*(\mathcal{M})$ is the triangle with vertices at $(0, 0)$, $(\frac{1}{2}, 0)$ and $(1, 1)$. The horizontal face of the rotation set corresponds to the subshift of finite type obtained by disallowing the block 11. The face from $(\frac{1}{2}, 0)$ to $(1, 1)$ corresponds to the subshift of finite type obtained by disallowing the block 00. Both these are mixing subshifts of finite type, so applying Proposition 6 to any ϱ in the (relative) interior of one of these faces we deduce that there is a lost measure in $f_*^{-1}(\varrho)$ which (uniquely) attains $\mathcal{H}(\varrho)$.

If ϱ is a non-extremal boundary point then there need not be any ergodic measures in its rotation class, as the next two examples show.

Example 5. *ϱ non-extremal, $\mathcal{H}(\varrho) = H(\varrho)$, but not attained by any ergodic measure.*

Take X and f as in Example 4. The face from $(0, 0)$ to $(1, 1)$ of the rotation set corresponds to the non-transitive subshift of finite type consisting of the two fixed points. In particular, if we choose a rotation vector ϱ lying in the relative interior of this face then the corresponding rotation class consists of a single measure, which is some non-trivial convex combination of the Dirac measures δ_0 , δ_1 supported on the fixed points of X. In particular there are no ergodic measures (and hence no lost or directional measures) with rotation number ϱ.

Example 6. *ϱ non-extremal, $\mathcal{H}(\varrho) = 0$, $H(\varrho) > 0$, but H not attained by any ergodic measure.*

Let X be the full shift on the alphabet $\{0, 1\}$. Let $Y_1, Y_2 \subset X$ be two distinct positive entropy uniquely ergodic subshifts, with corresponding invariant measures μ_1, μ_2. Let $f_2(x) = -\delta(x, Y_1 \cup Y_2)$, and let f_1 be such that $a = \int f_1 \, d\mu_1 < \int f_1 \, d\mu_2 = b$. Let $f = (f_1, f_2)$. Then the line segment from $(a, 0)$ to $(b, 0)$ lies on the boundary of the rotation set, but any point in its relative interior has no ergodic measures in its rotation class.

Remark 4. One might object that the absence of an ergodic measure from a rotation class, as in Example 6, is a rather trivial way for H to not be attained by a lost measure. In the next two examples, however, we show that this can occur even if the rotation class contains ergodic measures. Example 7 is in a sense optimal, in that one coordinate function is constant on cylinders of length two, and the other is constant on cylinders of length one. By Theorem 5.7 of [15] we cannot concoct such examples if *both* coordinate functions are constant on cylinders of length one. In such a case, if ϱ is non-extremal, and $f_*^{-1}(\varrho)$ contains an ergodic measure, then $f_*^{-1}(\varrho)$ must also contain a lost measure.

Example 7. *ϱ non-extremal, $\mathcal{H}(\varrho) = H(\varrho)$, attained by many directional measures but by no lost measures.*

Let X be the full shift on the alphabet $\{0,1,2,3\}$. Let $f_1[0] = 0 = f_1[1]$, $f_1[2] = -1$, $f_1[3] = 1$. Let f_2 be constant on length-two cylinders, with $f_2[i,j] = M(i,j)$ where

$$M = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Then $(0,1)$ is a non-extremal boundary point of the rotation set for $f = (f_1, f_2)$. Any ergodic measure supported on $[0] \cup [1]$ has rotation vector $(0,1)$, and all such measures are directional. These are the only ergodic measures with rotation vector $(0,1)$.

Example 8. *ϱ non-extremal, $\mathcal{H}(\varrho) = H(\varrho)$, attained uniquely by a directional measure.*

Let X be the full shift on the alphabet $\{0,1\}$. Let Y_1, Y_2, Y_3 be distinct positive entropy uniquely ergodic subshifts, with corresponding invariant measures μ_1, μ_2, μ_3. Suppose $h(\mu_2) \geq h(\mu_1)$ and $h(\mu_2) \geq h(\mu_3)$. Let $f_2(x) = -\delta(x, Y_1 \cup Y_2 \cup Y_3)$, and let f_1 be such that $a_1 < a_2 < a_3$, where a_i denotes $\int f_1 \, d\mu_i$. Let $f = (f_1, f_2)$. Then the line segment from $(a_1, 0)$ to $(a_3, 0)$ lies on the boundary of the rotation set, and the point $\varrho = (a_2, 0)$ lies in its relative interior. The only *ergodic* measure in the rotation class of ϱ is the directional measure μ_2, and since $h(\mu_2) \geq h(\mu_1)$, $h(\mu_2) \geq h(\mu_3)$, then $h(\mu_2) = \mathcal{H}(\varrho) = H(\varrho)$.

Example 9. *ϱ non-extremal, $\mathcal{H}(\varrho) = H(\varrho)$, attained by both a directional and a lost measure.*

Let X be the full shift on the alphabet $\{0,1\}$. Let Y be a minimal subshift with precisely three ergodic invariant measures, μ_1, μ_2, μ_3. Suppose $h(\mu_2) \geq h(\mu_1)$ and $h(\mu_2) \geq h(\mu_3)$. Let Y' be a uniquely ergodic subshift, with invariant measure μ'. Let $f_2(x) = -\delta(x, Y \cup Y')$, and let f_1 be such that $a_1 < a_2 < a_3$, where a_i denotes $\int f_1 \, d\mu_i$, and such that $\int f_1 d\mu' = a_2$. Let $f = (f_1, f_2)$. Then the line segment from $(a_1, 0)$ to $(a_3, 0)$ lies on the boundary of the rotation set, and the point $\varrho = (a_2, 0)$ lies in its relative interior. Both the directional measure μ' and the lost measure μ_2 lie in the rotation class of ϱ.

If both Y_1, Y_2 have zero topological entropy then we already have $H(\varrho) = \mathcal{H}(\varrho) = h(\mu') = h(\mu)$ as required.

To obtain a positive entropy example we choose Y such that μ_2 has positive entropy, then by [17] we can choose the (uniquely ergodic) measure μ' with $h(\mu') = h(\mu_2)$.

References

[1] S. Aubry and P. Y. Le Daeron, The discrete Frenkel-Kontorova model and its extensions 1. Exact results for the ground states, *Physica* **8D** (1983), 381–422.

[2] M. Babillot and F. Ledrappier, Lalley's theorem on periodic orbits of hyperbolic flows, *Ergod. Th. Dyn. Sys.*, **18** (1998), 17–39.

[3] V. Bangert, Mather sets for twist maps and geodesics on tori, *Dynamics Reported* **1** (1988), 1–45.

[4] V. Bangert, Minimal geodesics, *Ergod. Th. Dyn. Sys.* **10** (1990), 263–286.

[5] L. Barreira, Ya. Pesin, and J. Schmeling, On a general concept of multifractality: multifractal spectra for dimensions, entropies, and Lyapunov exponents. Multifractal rigidity, *Chaos*, **7:1** (1997), 27–38.

[6] L. Barreira and J. Schmeling, Invariant sets with zero measure and full Hausdorff dimension, *Electron. Res. Announc. Amer. Math. Soc.* **3** (1997), 114–118 (electronic).

[7] L. Barreira and J. Schmeling, Sets of "non-typical" points have full topological entropy and full Hausdorff dimension, *IST Preprint 14/97*, 1997.

[8] A. Blokh, Functional rotation numbers for one dimensional maps, *Trans. Amer. Math. Soc.*, **347** (1995), 499–513.

[9] T. Bousch, Le poisson n'a pas d'arêtes, *Ann. I. H. P. (Prob.)*, to appear.

[10] M. Bramson and S. Kalikow, Nonuniqueness in g-functions, *Israel J. Math.*, **84** (1993) 153–160.

[11] Z. Coelho and A. Quas, Criteria for $\overline{d}$-continuity, *Trans. Amer. Math. Soc.*, **350** (1998), 3257–3268.

[12] M. Denker, C. Grillenberger, and K. Sigmund, *Ergodic Theory on Compact Spaces*, Springer Lecture Notes in Mathematics, **527**, 1976.

[13] T. Downarowicz, The Choquet simplex of invariant measures for minimal flows, *Israel J. Math.*, **74** (1991), 241–256.

[14] A. H. Fan and D. J. Feng, Analyse multifractale de la récurrence sur l'espace symbolique, *C. R. Acad. Sci. Paris Sr. I Math.*, **327** (1998), 629–632.

[15] W. Geller and M. Misiurewicz, Rotation and entropy, *Trans. Amer. Math. Soc.*, **351** (1999), 2927–2948.

[16] E. Glasner and B. Weiss, Kazhdan's property T and the geometry of the collection of invariant measures, *Geom. Funct. Anal.*, **7** (1997), 917–935.

[17] C. Grillenberger, Construction of strictly ergodic systems I. Given entropy, *Z. Wahr. verw. Geb.*, **25** (1973), 323–334.

[18] B. Grünbaum, *Convex Polytopes*, Pure and Applied Mathematics vol. XVI, Interscience, 1967.

[19] F. Hofbauer, Examples for the nonuniqueness of the equilibrium state, *Trans. Amer. Math. Soc.*, **228** (1977) 223–241.

[20] O. Jenkinson, Conjugacy rigidity, cohomological triviality, and barycentres of invariant measures, *Ph.D. thesis, Warwick University*, 1996, `http://www.maths.qmw.ac.uk/~omj`

[21] O. Jenkinson, Geometric barycentres of invariant measures for circle maps, *Ergod. Th. Dyn. Sys.*, to appear.

[22] O. Jenkinson, Frequency-locking on the boundary of the barycentre set, *Exp. Math.*, to appear.

[23] G. Keller, *Equilibrium states in ergodic theory*, Cambridge University Press, 1998.

[24] S. Kim, R. S. MacKay, J. Guckenheimer, Resonance regions for families of torus maps, *Nonlinearity*, **2** (1989), 391–404.

[25] J. Kwapisz, Every convex polygon with rational vertices is a rotation set, *Ergod. Th. Dyn. Sys.* **12** (1992), 333–339.

[26] J. Kwapisz, A toral diffeomorphism with a nonpolygonal rotation set, *Nonlinearity* **8** (1995), 461–476.

[27] J. Lindenstrauss, G. H. Olsen, Y. Sternfeld, The Poulsen simplex, *Ann. Inst. Fourier (Grenoble)*, **28** (1978), 91–114.

[28] D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding*, Cambridge University Press, Cambridge, 1995.

[29] A. Livsic, Homology properties of Y-systems, *Math. Zametki*, **10** (1971), 758–763.

[30] G. McShane and I. Rivin, Simple curves on hyperbolic tori, *C. R. Acad. Sci. Paris Sr. I Math.*, **320** (1995), 1523–1528.

[31] G. McShane and I. Rivin, A norm on homology of surfaces and counting simple geodesics, *Internat. Math. Res. Notices*, (1995), 61–69 (electronic).

[32] R. Mañé, On the minimizing measures of Lagrangian dynamical systems, *Nonlinearity*, **5** (1992), 623–638.

[33] R. Mañé, Generic properties and problems of minimizing measures of Lagrangian systems, *Nonlinearity*, **9** (1996), 273–310.

[34] B. Marcus and S. Tuncel, The weight-per-symbol polytope and scaffolds of invariants associated with Markov chains, *Ergod. Th. Dyn. Sys.*, **11** (1991), 129–180.

[35] D. Massart, Normes stables des surfaces, *Ph.D. Thesis*, ENS Lyon, 1996.

[36] D. Massart, Stable norms of surfaces: local structure of the unit ball of rational directions, *Geom. Funct. Anal.* **7** (1997), 996–1010.

[37] J. Mather, Existence of quasiperiodic orbits for twist homeomorphisms of the annulus, *Topology* **21** (1982), 457–467.

[38] J. Mather, Action minimizing invariant measures for positive definite Lagrangian systems, *Math. Z.*, **207** (1991), 169–207.

[39] S. Newhouse, Continuity properties of entropy, Ann. of Math. (2), **129** (1989), 215–235.

[40] S. Newhouse, J. Palis, F. Takens, Bifurcations and stability of families of diffeomorphisms, *Inst. Haut. Études Sci. Publ. Math.* **57** (1983), 5–71.

[41] W. Parry, Intrinsic Markov Chains, *Trans. Amer. Math. Soc.,* **112** (1964), 55–65.

[42] W. Parry and M. Pollicott, *Zeta functions and the Periodic Orbit Structure of Hyperbolic Dynamics,* Astérisque, 187–188, 1990.

[43] W. Parry and S. Tuncel, On the classification of Markov chains by finite equivalence, *Ergod. Th. Dyn. Sys.*, **1** (1981), 303–335.

[44] Ya. Pesin, *Dimension Theory in Dynamical Systems*, University of Chicago Press, 1997.

[45] Ya. Pesin and B. Pitskel, Topological pressure and the variational principle for noncompact sets, *Functional Anal. Appl.* **18** (1984), 307-318.

[46] H. Poincaré, Sur les courbes définies par les équations différentielles, *Œuvres Complètes, tome 1*, Gauthier-Villars, Paris, 1952, 137–158.

[47] R. T. Rockafeller, *Convex Analysis*, Princeton University Press, 1970.

[48] D. Ruelle, Statistical mechanics on a compact set with $\mathbb{Z}^\nu$ action satisfying expansiveness and specification, *Trans. Amer. Math. Soc.*, **185** (1973), 237–253.

[49] D. Ruelle, *Thermodynamic Formalism*, Encyclopaedia of Mathematics and its Applications, vol. 5, Addison-Wesley, 1978.

[50] S. Schwartzman, Asymptotic cycles, *Ann. of Math. (2)* **66** (1957), 270–284.

[51] Y. G. Sinai, Gibbs measures in ergodic theory, *Russ. Math. Surv.*, **27** (1972), 21–70.

[52] F. Takens and E. Verbitski, On the variational principle for the topological entropy of certain non-compact sets, *preprint*, 1999, University of Groningen.

[53] P. Walters, Ruelle's operator theorem and g-measures, *Trans. Amer. Math. Soc.*, **214** (1975), 375–387.

[54] P. Walters, A variational principle for the pressure of continuous transformations, *Amer. J. Math.*, **97** (1976), 937–971.

[55] P. Walters, *An Introduction to Ergodic Theory*, Springer-Verlag 1982.

[56] P. Walters, Differentiability properties of the pressure of a continuous transformation on a compact metric space, J. London Math. Soc. (2), **46** (1992), 471–481.

[57] F. Warner, *Foundations of differentiable manifolds and Lie groups*, Graduate Texts in Mathematics, 94. Springer-Verlag, New York-Berlin, 1983.

[58] S. Williams, Toeplitz minimal flows which are not uniquely ergodic, *Z. Wahr. verw. Geb.*, **67** (1984), 95–107.

[59] K. Ziemian, Rotation sets for subshifts of finite type, *Fund. Math.*, **146** (1995), 189–201.

UPR 9016 CNRS, Institut de Mathématiques de Luminy, 163 avenue de Luminy, case 907, 13288 Marseille, cedex 9, France

E-mail address: omj@iml.univ-mrs.fr , omj@maths.warwick.ac.uk