Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

Website: http://AIMsciences.org

pp. **X–XX**

ERGODIC OPTIMIZATION

Oliver Jenkinson

School of Mathematical Sciences
Queen Mary, University of London
Mile End Road
London E1 4NS, UK.

(Communicated by Aim Sciences)

Abstract. Let f be a real-valued function defined on the phase space of a dynamical system. Ergodic optimization is the study of those orbits, or invariant probability measures, whose ergodic f-average is as large as possible.

In these notes we establish some basic aspects of the theory: equivalent definitions of the maximum ergodic average, existence and generic uniqueness of maximizing measures, and the fact that every ergodic measure is the unique maximizing measure for some continuous function. Generic properties of the support of maximizing measures are described in the case where the dynamics is hyperbolic. A number of problems are formulated.

Contents

1. **Introduction.** Ergodic theory is concerned with the iteration of measure preserving transformations T of probability spaces $(X, \mathfrak{B}, \mu)$, and in particular with

2000 *Mathematics Subject Classification.* Primary: 37D20; Secondary: 37A05, 37D35, 37E10.
Key words and phrases. maximizing measure.

time averages

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) \tag{1}$$

of μ-integrable functions $f : X \to \mathbb{R}$.

Topological dynamics is concerned with the iteration of continuous self-maps $T : X \to X$, where X is a compact metric space. The ergodic theory of topological dynamics is the study of $\mathcal{M}_T$, the collection of those Borel probability measures on X which are preserved by such a self-map T. There is always at least one T-invariant measure, though the dynamics of T is potentially more interesting if $\mathcal{M}_T$ is a large set (e.g. this is the case if T is hyperbolic). If $\mathcal{M}_T$ is a singleton $\{\mu\}$, and $f : X \to \mathbb{R}$ is continuous, then $\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \int f\, d\mu$ for all $x \in X$. On the other hand if $\mathcal{M}_T$ is not a singleton then there are many continuous functions f for which the time average (1) is not a constant function[1] of x (e.g. if T is hyperbolic then the range of possible values of (1) is a closed interval, typically of positive length). In this case it is natural to ask what the largest and smallest values of (1) are, and to understand those points x (or rather their orbits) which attain these extreme values.

The ergodic theorem allows a reformulation of this problem in terms of invariant measures: the objects of interest are then those T-invariant probability measures which maximize, or minimize, the space average $\int f\, d\mu$ over all $\mu \in \mathcal{M}_T$. These are the *maximizing measures* and *minimizing measures* for the function f (with respect to the dynamical system $T : X \to X$). Of course it is sufficient to restrict attention to *maximizing* measures, because any minimizing measure for f is a maximizing measure for $-f$.

We shall refer to the above circle of problems as *ergodic optimization*. More broadly it might refer to the optimization of ergodic (time or space) averages of a real-valued function f in any situation where the problem has a sense. In general neither f nor T need be continuous, and X need not be a compact, or even a metric, space; however these assumptions are convenient ones which simplify the development of the theory. Equally the dynamical system need not be a single map; it might be a flow, or more generally the topological action of a suitable (e.g. amenable) group or semi-group. Many of the results in these notes are valid in this more general context.

The fundamental question of ergodic optimization is: what do maximizing measures look like? Roughly speaking, this question can be interpreted in two ways. The first interpretation concerns *specific* problems: for a given map $T : X \to X$, and a specific function $f : X \to \mathbb{R}$, we would like to determine precisely its maximizing measure(s) and its maximum ergodic average. This is most interesting when $\mathcal{M}_T$ is large, for example when T is hyperbolic. Perhaps surprisingly, this sort of problem is often tractable (see §§2.2, 2.4), in spite of the large number of candidate maximizing measures. The second interpretation is more general: under appropriate hypotheses on T we would like to determine common properties of f-maximizing measures, for f varying within some large function space. For example, is the maximizing measure unique? Is it supported on a periodic orbit? Is it fully supported?

[1] Strictly speaking (1) need not define a function on X, since there may be points x for which the limit (1) does not exist. Nevertheless there do exist $x, x' \in X$ such that the time averages along both orbits exist but are not equal.

Does it have positive entropy? As we shall see, a more coherent picture emerges by restricting attention to *generic* functions.

These notes reflect the two interpretations of the fundamental question, though more emphasis is placed on the general theory, in particular generic properties of maximizing measures. In §§2 and 3 we give complete proofs of most of the important basic theory: the equivalent definitions of maximum ergodic average, and the existence and generic uniqueness of maximizing measures. We sketch the proof of the fact that every ergodic measure is the unique maximizing measure for some continuous function.

The explicit problems considered concern maximum hitting frequencies for intervals in §2.2, and maximizing measures for degree-one trigonometric polynomials in §2.4. An application to the characterisation of expanding maps is described in §2.5.

When T is hyperbolic, more can be said about the *support* of the maximizing measure for a generic function, although the nature of this knowledge depends sharply on the ambient function space. In C^0 the support of the maximizing measure is generically large (§4.1), while in spaces of more regular functions it is generically thin (§4.2), and indeed is conjecturally a periodic orbit (§4.3).

The definition of a maximizing measure is reminiscent of the thermodynamic notion of an *equilibrium measure*. An equilibrium measure for the function f is by definition an invariant probability measure which maximizes $\int f\, d\mu + h(\mu)$ over all $\mu \in \mathcal{M}_T$, where $h(\mu)$ denotes the entropy of μ. There is a sense in which ergodic optimization can be regarded as a limiting case of thermodynamic formalism, and there are some interesting open questions concerning the relation between the two theories. However there are many respects in which they differ. One striking difference arises when T is hyperbolic and f is a sufficiently regular (e.g. Lipschitz) continuous function. The equilibrium measure for such an f is fully supported, with a Bernoulli natural extension and in particular positive entropy, so to a large extent it reflects the chaotic nature of T. In general, however, the equilibrium measure cannot be described in closed form, and it is impossible to determine explicitly any of its generic points. By contrast the maximizing measure for f is typically strictly ergodic, and therefore not fully supported. It often admits a very explicit description, in which case its generic points can be precisely identified. As noted above, it is conjectured that the maximizing measure for a Lipschitz function is typically a periodic orbit measure.

This conjecture is one of several open questions described in these notes, and the plentiful supply of unsolved problems reflects the youthful nature of the subject. Despite the naturalness of the optimization problem, and the potential for applications, ergodic optimization only began to develop during the 1990s. Undoubtedly one factor hampering its development any earlier than the 1980s was the unavailability of sufficiently powerful computers. Computer experiment is a key tool for gaining insight into the detailed structure of maximizing measures, and is helpful in suggesting which of their properties are typical. Some of the most striking results have been discovered with the aid of a computer, and without one it is hard to imagine that these would even have been conjectured.

The material described in these notes is the work of a number of authors, and a detailed guide to the literature is included as §5.

Terminology. Uniquely ergodic invariant sets, and in particular periodic orbits, play an important role in ergodic optimization (see §2.2, §2.4, §4). For such sets it will often be convenient to blur the distinction between the set itself and the unique invariant measure supported on it. So an invariant measure supported on a periodic orbit will simply be referred to as a *periodic orbit* whenever no confusion is likely to arise.

2. **Maximizing measures and maximum ergodic averages.**

2.1. **The basic problem.** Suppose we are given a map $T : X \to X$ on some set X, and a real-valued function $f : X \to \mathbb{R}$. Let $S_n f = \sum_{i=0}^{n-1} f \circ T^i$. Our basic problem will be to determine the maximum value of the time average

$$\lim_{n\to\infty} \frac{1}{n} S_n f(x) \,. \tag{2}$$

In other words, we are interested in the supremum of (2) over all choices of initial position x. Moreover we would like to know which points x, if any, attain the supremum.

Whilst conveying the essence of our problem, the above formulation is ill-posed, since in general the limit (2) need not exist. There are two obvious, and equally valid, ways of rectifying this. The first is to define $\mathrm{Reg}(f,T)$ as the set of $x \in X$ for which the limit (2) exists, and then consider the quantity

$$\beta(f) = \sup_{x\in \mathrm{Reg}(f,T)} \lim_{n\to\infty} \frac{1}{n} S_n f(x) \,. \tag{3}$$

If $\mathrm{Reg}(f,T)$ is empty then we define $\beta(f) = -\infty$.

The second is to simply set

$$\gamma(f) = \sup_{x\in X} \limsup_{n\to\infty} \frac{1}{n} S_n f(x) \,. \tag{4}$$

Clearly $\beta(f) \le \gamma(f)$. The points x which attain the supremum in either (3) or (4), as well as their orbits $\{x, T(x), T^2(x), \ldots\}$, might be called *maximizing.* Similarly, both $\beta(f)$ and $\gamma(f)$ might be called *maximum ergodic averages*, though we shall reserve this terminology for cases where they are known to coincide (see Definition 2.3).

A third plausible definition of the maximum ergodic average of f is

$$\delta(f) = \limsup_{n\to\infty} \frac{1}{n} \sup_{x\in X} S_n f(x) \,,$$

where we only maximize over orbit *segments*, and then let the segment length increase. Clearly $\delta(f) \ge \gamma(f)$. If $\delta(f) \ne +\infty$ (which is the case if f is bounded above, say) then $\sup_{x\in X} S_n f(x)$ is finite for all sufficiently large n, and is a subadditive sequence of reals, so the limit $\lim_n \frac{1}{n} \sup_{x\in X} S_n f(x) \in [-\infty, \infty)$ exists and equals $\inf_n \frac{1}{n} \sup_{x\in X} S_n f(x)$.

2.2. **An example: maximum hitting frequencies.** Pick a subset $A \subset X$, and let $f = \chi_A$ be the associated characteristic function. In this case ergodic averages are *hitting frequencies*: the limit

$$\lim_{n\to\infty} \frac{1}{n} S_n f(x) = \lim_{n\to\infty} \frac{1}{n} \# \left\{ 1 \le i \le n : T^i x \in A \right\} \,, \tag{5}$$

if it exists, is the frequency with which the orbit of x hits (i.e. lands in) the set A. So

$$\gamma(f) = \sup_{x\in X} \limsup_{n\to\infty} \frac{1}{n} \# \left\{ 1 \le i \le n : T^i x \in A \right\} \tag{6}$$

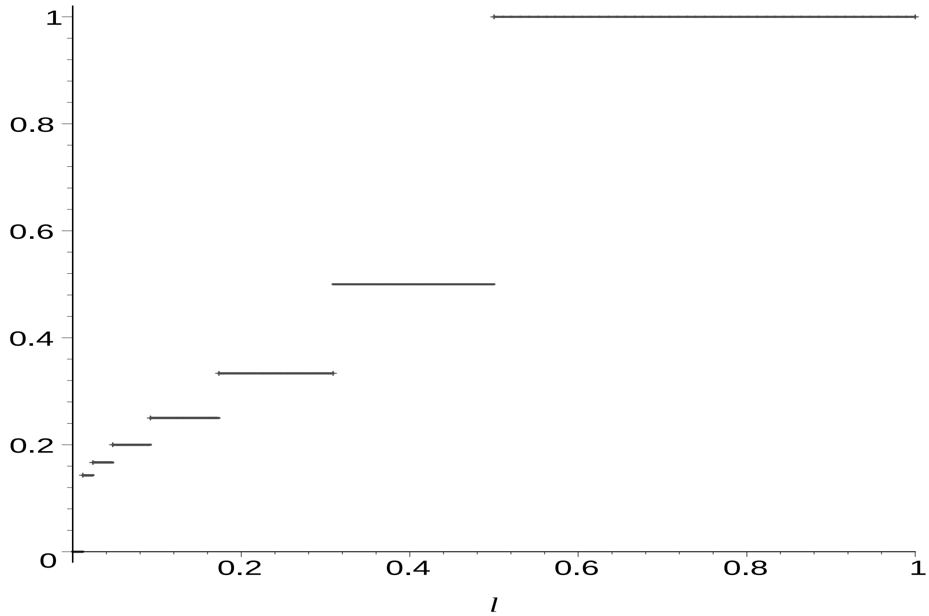

FIGURE 1. $\alpha(l)$ for the Ulam-von Neumann map

and

$$\beta(f) = \sup_{x\in \mathrm{Reg}(f,T)} \lim_{n\to\infty} \frac{1}{n} \# \left\{ 1 \le i \le n : T^i x \in A \right\} \tag{7}$$

can be interpreted as *maximum hitting frequencies* for the set A. Alternatively, the reciprocals $\beta(f)^{-1}$, $\gamma(f)^{-1}$ may be interpreted as *fastest mean return times* to A (see [J5]).

For example let $T : [0,1] \to [0,1]$ be the Ulam-von Neumann map $T(x) = 4x(1-x)$. Let $A = A_l = [(1-l)/2, (1+l)/2]$ be the closed interval of length l centred at the point $1/2$. It can be shown that $\delta(\chi_{A_l}) = \gamma(\chi_{A_l}) = \beta(\chi_{A_l})$ for all l, and we shall use $\alpha(l)$ to denote this common value. We wish to study the way in which $\alpha(l)$ varies with l. If l is sufficiently large then there are orbits which remain in A_l for all time, so that $\alpha(l) = 1$. Indeed if $l \ge 1/2$ then the fixed point at $3/4$ lies in A_l, so $\alpha(l) = 1$ for $1/2 \le l \le 1$. On the other hand $\alpha(0) = 0$: no orbit visits $A_0 = \{1/2\}$ with positive frequency, since the point $1/2$ is not periodic. So $\alpha : [0,1] \to [0,1]$ is a non-decreasing function, increasing from the value $\alpha(0) = 0$ to the value $\alpha(1) = 1$. In [J5] it is shown that $\alpha(l)$ only takes values $1/n$, for $n \ge 1$ an integer (see Figure 1), and is discontinuous at the points $l_n = \sin\left(\frac{\pi}{2(2^n+1)}\right)$.

2.3. **Basic theory.** So far our ergodic averages have been *time averages*. It will be convenient to relate these to space averages, via Birkhoff's ergodic theorem. For this we shall henceforth assume that X is a *topological space*, and that both $T : X \to X$ and $f : X \to \mathbb{R}$ are Borel measurable. Let $\mathcal{M}_T$ denote the collection of all T-invariant Borel probability measures on X. Clearly $\mathcal{M}_T$ is a convex set. Provided f is integrable with respect to every measure in $\mathcal{M}_T$ (in particular this will be the case if f is bounded) we can define

$$\alpha(f) = \sup_{\mu \in \mathcal{M}_T} \int f \, d\mu \,, \tag{8}$$

with $\alpha(f) = -\infty$ if $\mathcal{M}_T = \emptyset$.

The quantity $\alpha(f)$ is now a fourth candidate definition of *maximum ergodic average*, alongside $\beta(f)$, $\gamma(f)$, and $\delta(f)$. In general $\alpha(f)$, $\beta(f)$, $\gamma(f)$, $\delta(f)$ do not coincide (see [JMU2]), though it will be useful to impose conditions on X and T which ensure that they do coincide.

If X is a non-empty compact metric space, and $T : X \to X$ is continuous, then the set $\mathcal{M}_T$ is non-empty, by the Krylov-Bogolioubov Theorem [Wa2, Cor. 6.9.1], and when equipped with the weak* topology it is compact [Wa2, Thm. 6.10]. We shall be interested in the case where f is either the characteristic function of a closed subset (as in §2.2), or is continuous.

Proposition 2.1. *Let $T : X \to X$ be a continuous map on a compact metric space. If $f : X \to \mathbb{R}$ is either continuous, or a characteristic function of a closed subset, then*

$$\alpha(f) = \beta(f) = \gamma(f) = \delta(f) \neq \pm\infty \,.$$

Proof. To see that $\alpha(f) \le \beta(f)$, suppose on the contrary that there exists an invariant measure $\mu \in \mathcal{M}_T$ for which $\int f \, d\mu > \beta(f)$. The ergodic decomposition theorem (see [Ph1, Ch. 10], [Wa2, p. 34, and p. 153, Remark (2)]) means we may assume μ to be ergodic, and Birkhoff's ergodic theorem then guarantees the existence of an $x \in X$ for which

$$\int f \, d\mu = \lim_{n\to\infty} \frac{1}{n} S_n f(x) \le \beta(f) \,,$$

a contradiction.

The inequality $\beta(f) \le \gamma(f) \le \delta(f)$ is immediate from the definitions, so it remains to show that $\delta(f) \le \alpha(f)$. The compactness of X means that the set $\mathcal{M}$ of Borel probability measures on X is compact with respect to the weak* topology (cf. [Wa2, Thm. 6.5]). If $\mu_n := \frac{1}{n}\sum_{i=0}^{n-1} \delta_{T^i x_n}$, where x_n is such that

$$\max_{x\in X} \frac{1}{n} S_n f(x) = \frac{1}{n} S_n f(x_n) = \int f \, d\mu_n \,,$$

then the sequence (μ_n) has a weak* accumulation point μ. It is easy to see that in fact $\mu \in \mathcal{M}_T$.

Without loss of generality we shall suppose that $\mu_n \to \mu$ in the weak* topology. If f is continuous, this means that

$$\delta(f) = \lim_{n\to\infty} \int f \, d\mu_n = \int f \, d\mu \le \alpha(f) \,,$$

while if $f = \chi_A$ where A is closed then [Bil, Thm. 2.1] implies that

$$\delta(f) = \lim_{n\to\infty} \int f \, d\mu_n = \lim_{n\to\infty} \mu_n(A) \le \mu(A) = \int f \, d\mu \le \alpha(f) \,.$$

Since f is bounded, $\alpha(f) \neq \pm\infty$. □

The key ingredients in the proof of Proposition 2.1 were the application of the ergodic theorem, and the upper semi-continuity of the functional $\mu \mapsto \int f \, d\mu$. This suggests the following generalisation:

Proposition 2.2. *Let $T : X \to X$ be a continuous map on a compact metric space. If $f : X \to \mathbb{R}$ is upper semi-continuous then*

$$\alpha(f) = \beta(f) = \gamma(f) = \delta(f) \in [-\infty, \infty)\,.$$

Proof. Since f is upper semi-continuous, and X is compact, then f is bounded above. Consequently $\int f\, d\mu$ is well-defined for all (T-invariant) probability measures μ, and does not equal $+\infty$ (though might equal $-\infty$). In particular $\alpha(f) = \sup_{\mu\in\mathcal{M}_T} \int f\, d\mu \in [-\infty, \infty)$.

We claim that the map

$$\begin{aligned} \mathcal{M}_T &\longrightarrow [-\infty, \infty) \\ \mu &\longmapsto \int f\, d\mu \end{aligned}$$

is upper semi-continuous: if $\mu_n \to \mu$ in $\mathcal{M}_T$ then $\int f\, d\mu \geq \limsup_{n\to\infty} \int f\, d\mu_n$.

Now there is a sequence of continuous functions $f_i : X \to \mathbb{R}$ with $f_i \geq f_{i+1}$ for all i, and such that $\lim_{i\to\infty} f_i(x) = f(x)$ (this monotone approximation of an upper semi-continuous function by continuous ones is possible in any perfectly normal topological space [Ton], and in particular in any metric space [Tie]; see [Eng, 1.7.15 (c)]).

If μ is such that $\int f\, d\mu > -\infty$ then the monotone convergence theorem implies that $\lim_{i\to\infty} \int (f - f_i)\, d\mu = 0$. So if $\varepsilon > 0$ then $\int (f - f_i)\, d\mu > -\varepsilon$ for i sufficiently large, and for any such i we have

$$\begin{aligned} \int f\, d\mu &= \int (f - f_i)\, d\mu + \int f_i\, d\mu \\ &> -\varepsilon + \int f_i\, d\mu \\ &= -\varepsilon + \limsup_{n\to\infty} \int f_i\, d\mu_n \\ &\geq -\varepsilon + \limsup_{n\to\infty} \int f\, d\mu_n\,. \end{aligned}$$

But $\varepsilon > 0$ was arbitrary, so

$$\int f\, d\mu \geq \limsup_{n\to\infty} \int f\, d\mu_n\,, \tag{9}$$

as required.

If $\int f\, d\mu = -\infty$ then we must show that $\limsup_{n\to\infty} \int f\, d\mu_n = -\infty$ as well. Now $\int \max(f, -j)\, d\mu \geq -j > -\infty$, so for any $j \in \mathbb{N}$ we can replace f by $\max(f, -j)$ in (9) to deduce that

$$\limsup_{n\to\infty} \int f\, d\mu_n \leq \limsup_{n\to\infty} \int \max(f, -j)\, d\mu_n \leq \int \max(f, -j)\, d\mu\,. \tag{10}$$

But $\int f\, d\mu = -\infty$, so $\int \max(f, -j)\, d\mu \to -\infty$ as $j \to \infty$. Letting $j \to \infty$ in (10) gives $\limsup_{n\to\infty} \int f\, d\mu_n = -\infty$, as required.

Having proved the upper semi-continuity of $\mu \mapsto \int f\, d\mu$, the proof of Proposition 2.1 can be followed almost verbatim to show that $\delta(f) \leq \alpha(f)$. Since $\beta(f) \leq \gamma(f) \leq \delta(f)$ is trivially true, it remains to prove that $\alpha(f) \leq \beta(f)$. If not then there exists $\mu \in \mathcal{M}_T$ for which $\int f\, d\mu > \beta(f)$, so in particular $\int f\, d\mu > -\infty$. Moreover $\int f\, d\mu < +\infty$, since f is bounded above. So $f \in L^1(\mu)$. As in Proposition 2.1 we may assume that μ is ergodic, and then apply the ergodic theorem to find an $x \in X$ for which

$$\int f\,d\mu = \lim_{n\to\infty} \frac{1}{n} S_n f(x) \le \beta(f)\,,$$

a contradiction. $\square$

Henceforth all our triples (X, T, f) will be as in Proposition 2.2, so that $\alpha(f) = \beta(f) = \gamma(f) = \delta(f)$.

Definition 2.3. Let X be a compact metric space, and suppose that $T : X \to X$ is continuous. The quantity

$$\alpha(f) = \sup_{\mu \in \mathcal{M}_T} \int f\,d\mu$$

is called the *maximum ergodic average* for f. A measure $\mu \in \mathcal{M}_T$ is called f-*maximizing* (or simply *maximizing*) if $\int f\,d\mu = \alpha(f)$, and the collection of all f-maximizing measures is denoted by $\mathcal{M}_{\max}(f)$.

Proposition 2.4. *Let $T : X \to X$ be a continuous map on a compact metric space, and suppose that $f : X \to \mathbb{R}$ is upper semi-continuous.*
(i) *There exists at least one f-maximizing measure.*
(ii) *$\mathcal{M}_{\max}(f)$ is a compact metrizable simplex.*
(iii) *The extreme points of $\mathcal{M}_{\max}(f)$ are precisely those f-maximizing measures which are ergodic. In particular, there is at least one ergodic f-maximizing measure.*

Proof. The set $\mathcal{M}_T$ is compact in the weak* topology, and $\mu \mapsto \int f\,d\mu$ is upper semi-continuous with respect to this topology, as shown in the proof of Proposition 2.2. Consequently there is at least one element $m \in \mathcal{M}_T$ for which $\int f\,dm = \sup_{\mu\in\mathcal{M}_T} \int f\,d\mu = \alpha(f)$, so $\mathcal{M}_{\max}(f) \neq \emptyset$.

The remaining properties are simple consequences of the fact that $\mu \mapsto \int f\,d\mu$ is affine (with the obvious convention that $-\infty + r = -\infty$ for all $r \in [-\infty, \infty)$) and upper semi-continuous, together with the fact that $\mathcal{M}_T$ is a compact metrizable simplex whose extreme points are the ergodic T-invariant probability measures (see e.g. [Ph1, Ch. 10]). $\square$

2.4. **An example: Sturmian measures.** Consider the map $T(x) = 2x \pmod 1$ on the circle $\mathbb{T}$. Every closed semi-circle contains the support of one, and only one, T-invariant probability measure (see [BS]). Any such measure is called *Sturmian*. One interpretation of this result is that the maximum hitting frequency (see §2.2) for every closed semi-circle is equal to one. However, the role of Sturmian measures in ergodic optimization is a deeper one, as we shall soon see.

Sturmian measures form a one-parameter family, and can be characterised in various ways. Most fundamental is the relation with circle rotations $R_\varrho : x \mapsto x + \varrho \pmod 1$. It turns out that for any angle ϱ there is one and only one Sturmian measure s_ϱ such that $T|_{\mathrm{supp}(s_\varrho)}$ is combinatorially equivalent to R_ϱ. In particular, if $\varrho = \frac{p}{q}$ is rational then s_ϱ is a periodic orbit of period q; for example $s_{2/5}$ is the orbit $\{\frac{5}{31}, \frac{10}{31}, \frac{20}{31}, \frac{9}{31}, \frac{18}{31}\}$. All orbits of period 1, 2 or 3 are Sturmian, but the period-4 orbit $\{\frac{1}{5}, \frac{2}{5}, \frac{4}{5}, \frac{3}{5}\}$ is not. Among periodic orbits, the Sturmian ones become increasingly rare as the period increases. If ϱ is irrational then the support of s_ϱ is a Cantor set.

Sturmian measures arise naturally in many branches of mathematics, and appear to play an important role in ergodic optimization: for many “naturally occurring” functions $f : \mathbb{T} \to \mathbb{R}$ the maximizing measure turns out to be Sturmian. This was first discovered [B1, J1, J2] in the context of the family $f_\theta(x) = \cos 2\pi(x - \theta)$ of degree-one trigonometric polynomials. Here the f_θ-maximizing measure is always some Sturmian measure s_ϱ, and conversely every Sturmian measure is f_θ-maximizing for some θ. This was proved by Bousch [B1] after conjectures in [J1, J2].

In particular there is a well-defined function $\theta \mapsto \varrho(\theta)$, where $s_{\varrho(\theta)}$ denotes the f_θ-maximizing measure. This function is weakly increasing, though not a bijection; it is locally constant on a countable infinity of intervals, each corresponding to a rational value of ϱ, i.e. to a *periodic* Sturmian measure. Thus periodic orbits are *stably maximizing* within the family f_θ: for any p/q, the set $D_{p/q} = \{\theta \in \mathbb{T} : s_{p/q}$ is f_θ-maximizing$\}$ has non-empty interior. Moreover, the union $\cup_{p/q\in\mathbb{Q}} D_{p/q}$ is dense in parameter space $\mathbb{T}$. So within the family f_θ, the property of having a *periodic* maximizing measure is (topologically) *generic*. This result motivates a related conjecture (see §4.3): that in various infinite-dimensional function spaces the maximizing measure is generically periodic. The property of having a periodic maximizing measure is also generic in a measure-theoretic sense: the parameters θ for which the f_θ-maximizing measure is non-periodic form a set of zero Lebesgue measure (indeed zero Hausdorff dimension).

A geometric picture of the phenomenon of periodic orbits being stably maximizing is obtained if we realise T as the squaring map $T : z \mapsto z^2$ on the unit circle $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ in the complex plane. To any T-invariant measure μ on S^1 we assign its *barycentre* (or first Fourier coefficient) $b(\mu) = \int z\, d\mu(z)$, some point in the closed unit disc. The barycentre set $\Omega = \{b(\mu) : \mu \in \mathcal{M}_T\}$ is easily seen to be compact, convex, and symmetric about the real axis. It is completely determined by its boundary points, and a short calculation shows that $b(\mu)$ lies on $\partial\Omega$ if and only if μ is f_θ-maximizing for some θ; in this case $b(\mu)$ has maximal component in the $2\pi\theta$ direction (i.e. its projection to the line through the origin making angle $2\pi\theta$ with the positive real axis is larger than for any other barycentre).

Thus the boundary $\partial\Omega$ is precisely the barycentre locus $\{b(s_\varrho) : \varrho \in \mathbb{T}\}$ of Sturmian measures. This boundary turns out to be non-differentiable at a countable dense subset, the points of non-differentiability being precisely the barycentres of *periodic* Sturmian measures (see Figure 2, where $\partial\Omega$ is approximated by periodic Sturmian barycentres of low period). In other words, for a whole interval's worth of angles θ, such a barycentre $b(s_{p/q})$ has maximal component in the $2\pi\theta$ direction, reflecting the fact that $D_{p/q}$ is an interval with non-empty interior.

2.5. **An application: when is a map expanding?** Let M be a compact Riemannian manifold. A C^1 map $T : M \to M$ is called *expanding* if there exists $\lambda > 1$ and $N \in \mathbb{N}$ such that

$$\|D_x T^n(v)\| \geq \lambda^n \|v\| \quad \text{for all } n \geq N,\ x \in M,\ v \in T_x M\,. \tag{11}$$

Clearly if T has a critical point (i.e. a point $c \in M$ such that $D_c T(v) = 0$ for some non-zero $v \in T_c M$) then it is not expanding. On the other hand if T has no critical points then there is a simple[2] necessary and sufficient condition for it to be expanding:

[2] To simplify further we prove Proposition 2.5 in the case where M is the circle, but see the remarks after its proof.

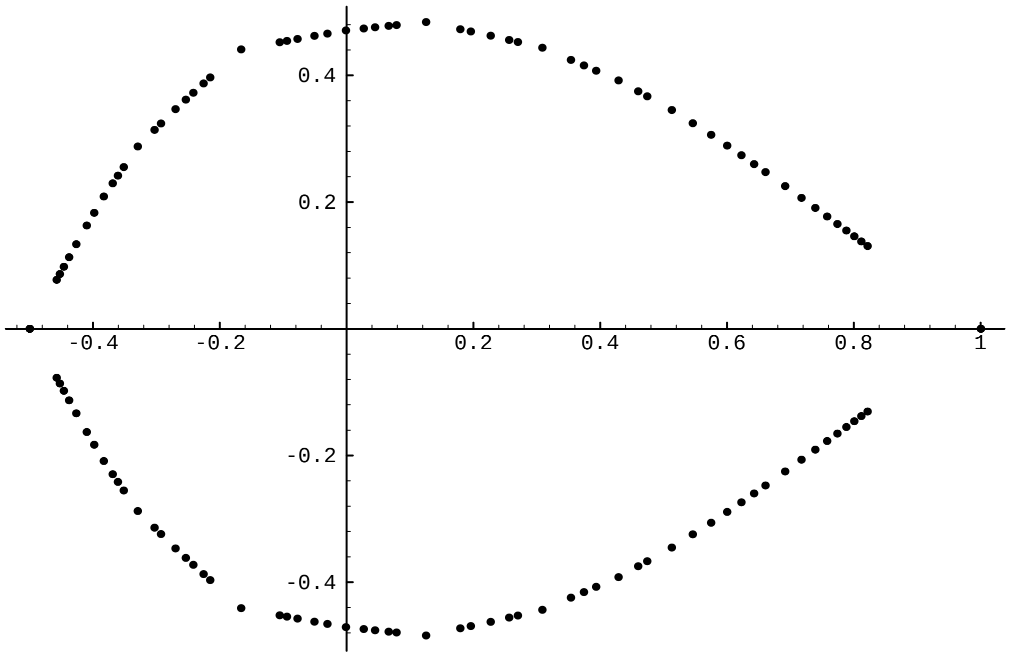

FIGURE 2. Approximation to the boundary of Ω

Proposition 2.5. *Let $T : \mathbb{T} \to \mathbb{T}$ be a C^1 map on the circle $\mathbb{T}$, and suppose that T has no critical points. Then T is an expanding map if and only if $\int \log |T'|\, d\mu > 0$ for all $\mu \in \mathcal{M}_T$.*

Proof. Suppose T is expanding. If n is sufficiently large, $|(T^n)'(x)| \geq \lambda^n$ for all $x \in \mathbb{T}$, and hence $\frac{1}{n}\sum_{i=0}^{n-1} \log |T'(T^i x)| \geq \log \lambda$ by the chain rule. Defining $f = -\log |T'|$ we see that

$$\gamma(f) = \sup_{x\in\mathbb{T}} \limsup_{n\to\infty} \frac{1}{n} S_n f(x) \leq -\log \lambda\,.$$

But T is continuous, as is f (since T is C^1 and has no critical point), so Proposition 2.1 implies that $\gamma(f) = \alpha(f)$. Therefore

$$\sup_{\mu\in\mathcal{M}_T} \int -\log |T'|\, d\mu = \alpha(f) = \gamma(f) \leq -\log \lambda < 0\,,$$

or in other words $\inf_{\mu\in\mathcal{M}_T} \int \log |T'|\, d\mu > 0$, as required.

Conversely, suppose $\int \log |T'|\, d\mu > 0$ for all $\mu \in \mathcal{M}_T$. Since $\log |T'|$ is continuous then so is the functional $\mu \mapsto \int \log |T'|\, d\mu$ defined on the compact space $\mathcal{M}_T$. It follows that $\eta := \inf_{\mu\in\mathcal{M}_T} \int \log |T'|\, d\mu > 0$. Therefore

$$\lim_{n\to\infty} \max_{x\in\mathbb{T}} \frac{1}{n} S_n f(x) = \delta(f) = \alpha(f) = \sup_{\mu\in\mathcal{M}_T} \int -\log |T'|\, d\mu = -\eta < 0\,,$$

by Proposition 2.1. In particular there exists $N \in \mathbb{N}$ such that if $n \geq N$ then

$$-\frac{1}{n} \sum_{i=0}^{n-1} \log |T'(T^i x)| = \frac{1}{n} S_n f(x) \leq -\frac{\eta}{2}$$

for all $x \in \mathbb{T}$. By the chain rule this means that

$$|(T^n)'(x)| \geq (e^{\eta/2})^n \quad \text{for all } n \geq N,\, x \in \mathbb{T}\,,$$

so T is expanding with $\lambda = e^{\eta/2} > 1$. ☐

Proposition 2.5 is due to [Cao] (see also [AAS]), who in fact proved the analogous result for C^1 maps T on any d-dimensional compact Riemannian manifold M: provided T has no critical points, it is expanding if and only if all Lyapunov exponents of all invariant measures are strictly positive. This result can be proved as above by replacing T by the projective cocycle $X \times \mathbb{RP}^{d-1} \to X \times \mathbb{RP}^{d-1}$, $(x,l) \mapsto (Tx, A(x)l)$, where $A(x)$ is the map on $\mathbb{RP}^{d-1}$ induced by $D_xT : \mathbb{R}^d \to \mathbb{R}^d$, and defining[3] $f : X \times \mathbb{RP}^{d-1} \to \mathbb{R}$ by $f(x,l) = -\log \|D_xTv\|$ where $v \in l$ is such that $\|v\| = 1$ (this is essentially the approach of [CLR]).

In spite of Proposition 2.5, for a particular map T it may be hard to determine whether or not T is expanding, since the smallest N satisfying (11) might be extremely large (see [Liv] for a related discussion in the context of piecewise smooth maps). More generally (i.e. irrespective of whether or not T is expanding), a non-trivial problem is to find the T-invariant measure(s) with minimum Lyapunov exponent. Such information is of interest to physicists, for example in connection with scarring of quantum eigenstates (see e.g. [Kap]), and the thermodynamic formalism for Lorentz gases [BD]. In general not much is known about invariant measures with minimum Lyapunov exponent (though see [CLT]), or more generally about maximizing (T-invariant) measures for functions f_T depending on the map T.

3. **Uniqueness of maximizing measures.**

3.1. **Generic uniqueness.** We have noted (Proposition 2.4 (i)) that the existence of an f-maximizing measure is always guaranteed, provided $T : X \to X$ is a continuous map on a compact metric space and $f : X \to \mathbb{R}$ is continuous[4]. Uniqueness, on the other hand, is certainly not guaranteed, unless $\mathcal{M}_T$ is a singleton. For example if f is a constant function then every invariant measure is f-maximizing.

We would like to say, however, that a "typical" function f *does* have a unique maximizing measure. More generally we may wish to assert that a typical function f has a certain property $\mathcal{P}$, where for our purposes $\mathcal{P}$ will always relate to the set of f-maximizing measures. That is, in a given function space E we would like to find a "large" subset E' such that every $f \in E'$ has the property $\mathcal{P}$. In any topological space E, a subset E' which is both *open* and *dense* would certainly be regarded as a large subset.

The following result asserts that if we make a rather strong assumption on the map T, then for a very natural class of function spaces E, the set

$$\mathcal{U}(E) = \{f \in E : \text{ there is a unique } f\text{-maximizing measure}\}$$

is open and dense in E. We use $C^0 = C^0(X)$ to denote the space of continuous real-valued functions on X, a real Banach space when equipped with the supremum norm $\|f\|_\infty = \sup_{x\in X} |f(x)|$.

---

[3]Similarly, for an arbitrary linear cocycle the T-invariant measure(s) with largest leading Lyapunov exponent can be related to measures invariant under the projective cocycle which are f-maximizing for a function f defined on $X \times \mathbb{RP}^{d-1}$. There are a number of interesting questions related to such "Lyapunov maximizing" measures (see also [SS]).

[4]Henceforth f will always be continuous rather than merely upper semi-continuous.

Proposition 3.1. *Let X be a compact metric space, and $T : X \to X$ a continuous map which has only finitely many ergodic invariant measures. Let E be a topological vector space which is densely and continuously embedded*[5] *in C^0. Then $\mathcal{U}(E)$ is open and dense in E.*

Proof. Let $\{\mu_1, \dots, \mu_N\}$ be the ergodic measures for T, and define

$$F_i = \{f \in E : \mu_i \text{ is } f\text{-maximizing}\}$$

for each $1 \le i \le N$. The complement $\mathcal{U}(E)^c$ can be expressed as the finite union

$$\mathcal{U}(E)^c = \bigcup_{i<j} F_i \cap F_j \,, \tag{12}$$

so to prove that $\mathcal{U}(E)$ is open it suffices to show that each F_i is closed. To this end suppose that $\{f_\alpha\}$ is a net in F_i, with $f_\alpha \to f$ in E. The continuous embedding of E in C^0 means that $f_\alpha \to f$ in C^0, and hence $\int f_\alpha \, d\mu \to \int f \, d\mu$ for any $\mu \in \mathcal{M}$. Now $\int f_\alpha \, d\mu_i \ge \int f_\alpha \, d\mu$ for every $\mu \in \mathcal{M}$, so $\int f \, d\mu_i \ge \int f \, d\mu$ for every $\mu \in \mathcal{M}$. That is, $f \in E$ is such that μ_i is f-maximizing. Therefore $f \in F_i$, so F_i is indeed closed.

Since E is densely embedded in C^0, for any $i < j$ there exists $g = g_{ij} \in E$ such that $\int g \, d\mu_i \ne \int g \, d\mu_j$. If $f \in F_i \cap F_j$ then for every $\varepsilon > 0$ the function $f + \varepsilon g$ does not lie in $F_i \cap F_j$. Therefore $F_i \cap F_j$ has empty interior whenever $1 \le i, j \le N$ are such that $i < j$, and from (12) it follows that $\mathcal{U}(E)$ is dense in E. □

The hypothesis that T has only finitely many ergodic measures is rather restrictive, and we would like to establish some analogue of Proposition 3.1 for more general maps T. In this case, as well as for certain other properties $\mathcal{P}$, it is too ambitious to hope to find an open and dense subset E' such that every $f \in E'$ has the property $\mathcal{P}$. More realistically one might search for a *residual* subset E' such that every $f \in E'$ has the property $\mathcal{P}$. A residual set is by definition one which contains a countable intersection of open dense subsets. We say that $\mathcal{P}$ is a *generic* property if there is some residual subset E' such that every element of E' has the property $\mathcal{P}$. We say that E is a *Baire space* if every residual subset of E is dense in E; in particular every complete metric space, hence every Banach space, has this property, by the Baire category theorem [Roy, p. 158].

The following Theorem 3.2 gives very general conditions under which the property of having a unique maximizing measure is a generic one; there are no extra hypotheses on the continuous map T, and the assumptions on E are as in Proposition 3.1.

Theorem 3.2. *Let $T : X \to X$ be a continuous map on a compact metric space. Let E be a topological vector space which is densely and continuously embedded in C^0. Then $\mathcal{U}(E)$ is a countable intersection of open and dense subsets of E.*

If moreover E is a Baire space, then $\mathcal{U}(E)$ is dense in E.

To prove Theorem 3.2 we first require some more notation and a preliminary lemma.

Definition 3.3. For $f, g \in C^0$, define

$$\alpha(g \,|\, f) = \max_{\mu \in \mathcal{M}_{\max}(f)} \int g \, d\mu \,, \tag{13}$$

[5] Less formally: E is dense as a vector subspace of C^0, and the topology on E is stronger than that on C^0.

the *relative maximum ergodic average of* g *given* f. Define

$$\mathcal{M}_{\max}(g\,|\,f) = \left\{\mu \in \mathcal{M}_{\max}(f) : \int g\,d\mu = \alpha(g\,|\,f)\right\}. \tag{14}$$

Lemma 3.4. *For any* $f, g \in C^0$,

$$\left\{\int g\,d\mu : \mu \in \mathcal{M}_{\max}(f+\varepsilon g)\right\} \longrightarrow \{\alpha(g\,|\,f)\} \quad \textit{as } \varepsilon \searrow 0\,, \tag{15}$$

the convergence being in the Hausdorff metric[6].

Proof. For all $\varepsilon > 0$, the set $\{\int g\,d\mu : \mu \in \mathcal{M}_{\max}(f+\varepsilon g)\}$ is some compact interval $[a_\varepsilon^-, a_\varepsilon^+]$. To prove the lemma it is enough to show that $\lim_{\varepsilon\searrow 0} a_\varepsilon^- = \alpha(g\,|\,f) = \lim_{\varepsilon\searrow 0} a_\varepsilon^+$, and for this it suffices to prove that if $a_\varepsilon \in \{\int g\,d\mu : \mu \in \mathcal{M}_{\max}(f+\varepsilon g)\}$ then $\lim_{\varepsilon\searrow 0} a_\varepsilon = \alpha(g\,|\,f)$. Writing $a_\varepsilon = \int g\,dm_\varepsilon$ for some $m_\varepsilon \in \mathcal{M}_{\max}(f+\varepsilon g)$, it is in turn enough to show that any weak* accumulation point of m_ε, as $\varepsilon \searrow 0$, belongs to $\mathcal{M}_{\max}(g\,|\,f)$. If m is such an accumulation point, with $m_{\varepsilon_i} \to m$ for some sequence $\varepsilon_i \searrow 0$, then

$$\int f + \varepsilon_i g\,dm_{\varepsilon_i} \geq \int f + \varepsilon_i g\,d\mu \quad \text{for all } \mu \in \mathcal{M}_T\,. \tag{16}$$

Letting $i \to \infty$ gives $\int f\,dm \geq \int f\,d\mu$ for all $\mu \in \mathcal{M}_T$, i.e. $m \in \mathcal{M}_{\max}(f)$. If $\mu \in \mathcal{M}_{\max}(f)$ then

$$\int f\,d\mu \geq \int f\,dm_{\varepsilon_i} \quad \text{for all } i \geq 1\,, \tag{17}$$

so combining (16), (17) gives

$$\varepsilon_i \int g\,dm_{\varepsilon_i} \geq \varepsilon_i \int g\,d\mu \quad \text{for all } \mu \in \mathcal{M}_{\max}(f)\,.$$

Dividing by the positive constant ε_i and letting $i \to \infty$ gives $\int g\,dm \geq \int g\,d\mu$ for all $\mu \in \mathcal{M}_{\max}(f)$, and therefore $m \in \mathcal{M}_{\max}(g\,|\,f)$, as required. $\square$

Proof of Theorem 3.2. Since X is a compact metric space, C^0 is separable [Wa2, Thm. 0.19]. But E is densely embedded in the separable space C^0, so there is a countable subset of E which is dense[7] in C^0 (if $\{f_n : n \in \mathbb{N}\}$ is a dense subset of C^0, choose $e_{n,i} \in E$ such that $e_{n,i} \to f_n$ (in C^0) as $i \to \infty$, and then $\{e_{n,i} : (n,i) \in \mathbb{N}^2\}$ is dense in C^0). If $\{g_i\}_{i=1}^\infty$ denotes this countable subset of E then a measure μ is uniquely determined by how it integrates the family $\{g_i\}_{i=1}^\infty$, by the Riesz representation theorem [Roy, p. 357]. Consequently $\mathcal{M}_{\max}(f)$ is a singleton if and only if the closed interval

$$M_i(f) := \left\{\int g_i\,d\mu : \mu \in \mathcal{M}_{\max}(f)\right\}$$

is a singleton for every $i \geq 1$.

If we define

$$E_{i,j} := \left\{f \in E : |M_i(f)| \geq j^{-1}\right\},$$

[6]The Hausdorff metric d_H (on non-empty compact subsets of $\mathbb{R}$) is defined by $d_H(A,B) = \max_{a\in A}\min_{b\in B}|a-b| + \max_{b\in B}\min_{a\in A}|a-b|$.

[7]In fact it suffices that the linear span of this countable subset be dense in C^0.

where $|\cdot|$ denotes length, then the complement $\mathcal{U}(E)^c$ can be written as

$$\begin{aligned}\mathcal{U}(E)^c &= \{f \in E : \mathcal{M}_{\max}(f) \text{ is not a singleton}\}\\ &= \{f \in E : |M_i(f)| > 0 \text{ for some } i \in \mathbb{N}\}\\ &= \bigcup_{j=1}^{\infty} \{f \in E : |M_i(f)| \ge j^{-1}, \text{ for some } i \in \mathbb{N}\}\\ &= \bigcup_{j=1}^{\infty}\bigcup_{i=1}^{\infty} \{f \in E : |M_i(f)| \ge j^{-1}\}\\ &= \bigcup_{j=1}^{\infty}\bigcup_{i=1}^{\infty} E_{i,j}\,.\end{aligned}$$

We claim that each $E_{i,j}$ is closed and has empty interior, from which it will follow that

$$\mathcal{U}(E) = \{f \in E : \mathcal{M}_{\max}(f) \text{ is a singleton}\} = \bigcap_{j=1}^{\infty}\bigcap_{i=1}^{\infty} E_{i,j}^c$$

is a countable intersection of open and dense subsets of E.

To show that $E_{i,j}$ is closed in E, let $\{f_\alpha\}$ be a net in $E_{i,j}$ with $f_\alpha \to f$ in E. We can write $M_i(f_\alpha) = [\int g_i\, d\mu_\alpha^-, \int g_i\, d\mu_\alpha^+]$ for measures $\mu_\alpha^\pm \in \mathcal{M}_{\max}(f_\alpha) \subset \mathcal{M}_T$. The weak* compactness of $\mathcal{M}_T$ means there exist $\mu^-, \mu^+ \in \mathcal{M}_T$ such that

$$\mu_\alpha^- \to \mu^- \qquad \text{and} \qquad \mu_\alpha^+ \to \mu^+ \tag{18}$$

along convergent sub-nets. This weak* convergence implies, in particular, that

$$\int g_i\, d\mu_\alpha^- \to \int g_i\, d\mu^- \qquad \text{and} \qquad \int g_i\, d\mu_\alpha^+ \to \int g_i\, d\mu^+\,.$$

Now $\int g_i\, d\mu_\alpha^+ - \int g_i\, d\mu_\alpha^- = |M_i(f_\alpha)| \ge 1/j$ for all α, so $\int g_i\, d\mu^+ - \int g_i\, d\mu^- \ge 1/j$. Therefore if we can show that $\mu^\pm \in \mathcal{M}_{\max}(f)$ it will follow that $|M_i(f)| \ge 1/j$, i.e. that $f \in E_{i,j}$, and hence that $E_{i,j}$ is closed. Now E is continuously embedded in C^0, so $f_\alpha \to f$ in C^0. Therefore $\int f_\alpha\, d\mu_\alpha^- = \int (f_\alpha - f)\, d\mu_\alpha^- + \int f\, d\mu_\alpha^- \to \int f\, d\mu^-$, and $\int f_\alpha\, dm \to \int f\, dm$ for every $m \in \mathcal{M}_T$. Since $\int f_\alpha\, d\mu_\alpha^- \ge \int f_\alpha\, dm$ for all $m \in \mathcal{M}_T$, we deduce that $\int f\, d\mu^- \ge \int f\, dm$ for all $m \in \mathcal{M}_T$, i.e. that μ^- is f-maximizing. The same argument shows that μ^+ is f-maximizing, so we are done.

To see that each $E_{i,j}$ has empty interior, let $f \in E_{i,j}$ be arbitrary. Now Lemma 3.4 tells us that

$$M_i(f + \varepsilon g_i) = \left\{ \int g_i\, d\mu : \mu \in \mathcal{M}_{\max}(f + \varepsilon g_i) \right\} \longrightarrow \{\alpha(g_i \,|\, f)\}$$

as $\varepsilon \searrow 0$, so in particular $|M_i(f + \varepsilon g_i)| < 1/j$ for $\varepsilon > 0$ sufficiently small. Therefore $f + \varepsilon g_i \notin E_{i,j}$ for $\varepsilon > 0$ sufficiently small, so f is not an interior point of $E_{i,j}$. $\square$

Many of the standard Banach spaces satisfy the hypotheses of Theorem 3.2. Since several of these will be important later on, it is convenient to define them here.

Definition 3.5. For any $0 < \alpha \le 1$, a function $g : X \to Y$ between metric spaces is called *α-Hölder* if there exists $K > 0$ such that $d_Y(g(x), g(x')) \le K d_X(x, x')^\alpha$ for all $x, x' \in X$. A 1-Hölder function is called *Lipschitz*, and it will be notationally convenient to say that any continuous function is 0-*Hölder*. For an α-Hölder function g, let $|g|_\alpha$ denote the infimum of those K which satisfy the above inequality. In particular $|g|_0 = \|g\|_\infty$.

For a compact metric space X, and for any $0 \le \alpha \le 1$, let $C^{0,\alpha} = C^{0,\alpha}(X)$ denote the set of α-Hölder functions $g : X \to \mathbb{R}$. Each $C^{0,\alpha}$ is a Banach space when equipped with the norm $\|g\|_\alpha := \max(\|g\|_\infty, |g|_\alpha)$.

If X is also a C^r smooth manifold, for $r \in \mathbb{N}$, then let $C^{r,\alpha}$ denote the set of functions which are r times continuously differentiable and whose r-th order derivative is α-Hölder. This is a Banach space when equipped with the norm $\|g\|_{r,\alpha} := \max(\|g\|_\infty, \|Dg\|_\infty, \ldots, \|D^r g\|_\infty, |D^r g|_\alpha)$.

We shall often make some statement about the space $C^{r,\alpha} = C^{r,\alpha}(X)$ for some, or all, $0 \le \alpha \le 1$ and $r \in \mathbb{Z}_{\ge 0}$, where X denotes a compact metric space. If $r = 0$ then such a statement will always have a sense, while if $r \ge 1$ then of course the statement is only being asserted for those X which are also C^r manifolds. For example the following is a consequence of Theorem 3.2.

Corollary 3.6. *Let $T : X \to X$ be a continuous map on a compact metric space. For all $0 \le \alpha \le 1$ and $r \in \mathbb{Z}_{\ge 0}$, a generic function in $C^{r,\alpha}$ has a unique maximizing measure.*

3.2. **Ergodic measures are uniquely maximizing.** Clearly every invariant measure $\mu \in \mathcal{M}_T$ is f-maximizing for *some* continuous function f; indeed if f is a constant then *every* $\mu \in \mathcal{M}_T$ is f-maximizing. It is more difficult for μ to be the *unique* maximizing measure for some continuous[8] f; by Proposition 2.4 (iii) a necessary condition for this is that μ be ergodic. In fact the ergodicity of μ is also a sufficient condition:

Theorem 3.7. *Let $T : X \to X$ be a continuous map on a compact metric space. For any ergodic measure $\mu \in \mathcal{M}_T$ there exists a continuous function $f : X \to \mathbb{R}$ such that μ is the unique f-maximizing measure.*

Proof. The full details are a little technical, so we just provide a sketch proof. As mentioned previously, the extreme points of the convex set $\mathcal{M}_T$ are precisely the ergodic invariant measures. Since $\mathcal{M}_T$ is a (compact metrizable) *simplex*, it can be shown that any extreme point μ is actually an *exposed* point. This means there is a (weak*) continuous affine functional $l : \mathcal{M}_T \to \mathbb{R}$ such that $l(\mu) = 0$ and $l(\nu) < 0$ for $\nu \in \mathcal{M}_T \setminus \{\mu\}$. It can be shown that this functional extends to a (weak*) continuous linear functional on the space of signed measures on X, and therefore is necessarily of the form $l : m \mapsto \int f\, dm$ for some $f \in C^0$. □

Theorem 3.7 is in fact a special case of the following result.

Theorem 3.8. *Let $T : X \to X$ be a continuous map on a compact metric space. Let $\mathcal{E}$ be a non-empty collection of ergodic T-invariant probability measures which is closed as a subset of $\mathcal{M}_T$. There exists a continuous function $f : X \to \mathbb{R}$ such that $\mathcal{M}_{\max}(f)$ is equal to the closed convex hull of $\mathcal{E}$.*

Despite Theorem 3.7, for a particular ergodic measure μ it might be difficult to explicitly exhibit a continuous function whose unique maximizing measure is μ. For example the following problem is open.

Problem 3.9. *Let $T(x) = 2x \pmod 1$. Explicitly exhibit a continuous function $f : \mathbb{T} \to \mathbb{R}$ whose unique maximizing measure is Lebesgue measure.*

[8] It is easy to see that every ergodic measure is the unique maximizing measure for some *bounded measurable* (rather than continuous) function f: we could take $f = \chi_{G(\mu)}$, for example, where $G(\mu)$ is the set of μ-generic points (i.e. those x such that $\frac{1}{n} S_n g(x) \to \int g\, d\mu$ for all $g \in C^0$).

Definition 3.10. For a measure μ on X, its *support*, denoted $\mathrm{supp}(\mu)$, is the smallest closed subset $S \subset X$ with $\mu(S) = 1$. If $\mu \in \mathcal{M}_T$ then it is easily shown that $\mathrm{supp}(\mu)$ is a T-invariant set. If $\mathrm{supp}(\mu) = X$ then we say that μ has *full support*.

A measure $\mu \in \mathcal{M}_T$ is called *strictly ergodic* if the restricted dynamical system $T|_{\mathrm{supp}(\mu)} : \mathrm{supp}(\mu) \to \mathrm{supp}(\mu)$ is uniquely ergodic (i.e. μ is its only invariant measure)[9].

Every strictly ergodic measure is ergodic, and for a strictly ergodic measure μ it is easy to explicitly exhibit a continuous function whose unique maximizing measure is μ. For example we may define $f(x) = -d(x, \mathrm{supp}(\mu))$. If the space X is a smooth manifold then we can find *smooth* functions f whose unique maximizing measure is μ, simply by choosing f to attain its maximum on, and only on, the closed set $\mathrm{supp}(\mu)$. Of course elementary constructions of this kind fail in the analytic category. For example if X is the circle $\mathbb{T}$ or the interval $[0, 1]$ then a non-constant real analytic function $f : X \to \mathbb{R}$ cannot attain its maximum on a set with accumulation points. So unless the invariant measure μ is periodic, f cannot attain its maximum only on $\mathrm{supp}(\mu)$. This suggests the following:

Problem 3.11. *Let T be a continuous map on either the circle or the interval. Suppose that T is not uniquely ergodic. For any given non-periodic strictly ergodic measure μ, can we always find a real analytic function f whose unique maximizing measure is μ?*

For the circle map $T(x) = 2x \pmod 1$ it is known that there exist real analytic functions whose unique maximizing measure is strictly ergodic but non-periodic. As noted in §2.4, examples of such functions can be found within the one-parameter family $f_\theta(x) = \cos 2\pi(x-\theta)$ (see [B1, J1, J2]); for certain values of θ the maximizing measure is a Sturmian measure supported on a Cantor set. Non-periodic Sturmian measures are, in a sense, the closest to periodic among all non-periodic measures; for example the symbolic complexity[10] of a non-periodic Sturmian orbit (which is a generic orbit for the corresponding Sturmian measure) is as small as it can be among non-periodic orbits. Non-periodic measures with higher complexity can also arise as maximizing measures for (higher degree) trigonometric polynomials; for example measures which are combinatorially equivalent to an interval exchange can occur (cf. [Br2, HJ]). All these measures seem to have rather low symbolic complexity, however, so that the following question is open.

Problem 3.12. *If $T(x) = 2x \pmod 1$, can a positive entropy T-invariant measure uniquely maximize a real analytic function f?*

4. **The support of a maximizing measure.** Our overall aim is to understand the nature of maximizing measures. By Theorem 3.2, in various large function spaces a

---

[9]The use of the term *strictly ergodic* to describe a measure is a little non-standard; more usually a closed invariant set is called strictly ergodic if it is both uniquely ergodic and minimal. However there is an obvious one-to-one correspondence between strictly ergodic measures and strictly ergodic sets.

[10]The complexity function $p(n)$ is defined in terms of the natural symbolic coding of T. If $\omega(x) \in \{0,1\}^{\mathbb{N}}$ is the dyadic expansion for $x \in \mathbb{T}$ then $p(n) = p_x(n)$ denotes the number of distinct length-n subwords of $\omega(x)$. A classical result of Morse & Hedlund [MH] (see also [PF, Ch. 6] for example) asserts that among sequences which are not eventually periodic, the lowest possible complexity is $p_x(n) = n + 1$, and this is attained if and only if x is Sturmian.

"typical" function has a unique maximizing measure. This is in harmony with the results for the one-parameter family of functions $f_\theta(x) = \cos 2\pi(x-\theta)$ described in §2.4, where in fact *every* function has a unique maximizing measure. As noted in §2.4, within this family a typical maximizing measure is *periodic*. The attempt to generalise this result to large function spaces will be described in this section. The main conjecture is that for any suitably hyperbolic map $T : X \to X$, a generic Lipschitz function has a periodic maximizing measure (see Conjecture 4.11). This conjecture is still open, though partial results towards it have been obtained (see §§4.2, 4.3). There is an analogous conjecture (that periodic maximizing measures are generic) in the space $C^{r,\alpha} = C^{r,\alpha}(X)$ for any $(r,\alpha) > (0,0)$. In $C^0 = C^{0,0}$, however, the typical behaviour is very different: in this case the maximizing measure is known to be generically of *full support* (see §4.1).

Throughout this section we shall assume that X is a compact metric space, and that the continuous map $T : X \to X$ is *(uniformly) hyperbolic* with *local product structure*. Our definitions are similar to those of Mañé [M1, §IV.9] and Ruelle [Rue1, §7.1]. First let us suppose that $T : X \to X$ is a homeomorphism, and for $\varepsilon > 0$ define the *ε-stable set* and the *ε-unstable set* of a point $x \in X$ by

$$W^s_\varepsilon(x) = \{y \in X : d(T^n y, T^n x) \le \varepsilon \text{ for all } n \ge 0\}$$

and

$$W^u_\varepsilon(x) = \big\{y \in X : d(T^{-n} y, T^{-n} x) \le \varepsilon \text{ for all } n \ge 0\big\}$$

respectively. We say that T has *local product structure* if there exist $\delta, \varepsilon > 0$ such that if $x, y \in X$ satisfy $d(x,y) < \delta$ then $W^s_\varepsilon(x) \cap W^u_\varepsilon(y)$ is a singleton.

The homeomorphism $T : X \to X$ is called *hyperbolic* if there exist $\varepsilon > 0$, $C > 0$, and $0 < \lambda < 1$ such that for all $x \in X$,

$$d(T^n y, T^n z) \le C\lambda^n \quad \text{for} \quad n \ge 0\,,\ y, z \in W^s_\varepsilon(x)\,,$$

and

$$d(T^{-n} y, T^{-n} z) \le C\lambda^n \quad \text{for} \quad n \ge 0\,,\ y, z \in W^u_\varepsilon(x)\,.$$

To handle the case where $T : X \to X$ is non-invertible it is convenient to work with its *natural extension*. The set

$$\widetilde{X} = \left\{ (x_i) \in \prod_{i \le 0} X \ :\ T(x_{i-1}) = x_i \text{ for all } i \le 0 \right\}$$

is equipped with the metric $\widetilde{d}((x_i),(y_i)) = \sup_{i\le 0} d(x_i,y_i)/2^{|i|}$, which makes it a compact space. The *natural extension* of T is defined to be the homeomorphism $\widetilde{T} : \widetilde{X} \to \widetilde{X}$ given by $\widetilde{T}((x_i)) = (T(x_i))$. We say that $T : X \to X$ is *hyperbolic*[11] *with local product structure* if its natural extension has these properties.

The above definitions are general enough to include all the familiar examples of smooth hyperbolic maps. For example if $X \subset M$ is a hyperbolic set (in the sense of [KH, Defn. 6.4.1]) with local product structure[12] for a diffeomorphism $T : M \to M$, then $T|_X$ is a hyperbolic homeomorphism with local product structure. In particular this is the case if T is an Anosov diffeomorphism (in which case $X = M$), or if X is the non-wandering set of an Axiom A diffeomorphism [Bow, Ch. 3]. Anosov endomorphisms [Pr], expanding maps [M1, §III.1], and locally maximal hyperbolic repellers [KH, §6.4] are examples of hyperbolic maps with local product structure.

[11] Ruelle [Rue2] prefers the terminology *pre-hyperbolic* in the non-invertible case.

[12] Or equivalently (see [KH, Prop. 6.4.21, Thm. 18.4.1]) X is a *locally maximal* hyperbolic set.

The symbolic codings of all these smooth maps, namely one-sided and two-sided subshifts of finite type (see [PP]), are also hyperbolic with local product structure.

If T is also *(topologically) transitive*, i.e. there exists an $x \in X$ whose orbit is dense in X, then we have the following important result of Sigmund [Sig].

Lemma 4.1. *Let $T : X \to X$ be transitive, and hyperbolic with local product structure. Let K be a proper closed T-invariant subset of X, and suppose $\mu \in \mathcal{M}_T$ is such that supp$(\mu) \subset K$. Then there is a sequence of periodic orbits μ_n, disjoint from K, such that $\mu_n \to \mu$ in the weak* topology.*

4.1. **Generic properties in C^0.** The main result here is that a generic C^0 function f is such that every f-maximizing measure has full support (i.e. $\mathrm{supp}(\mu) = X$ for all $\mu \in \mathcal{M}_{\max}(f)$).

Theorem 4.2. *Let $T : X \to X$ be transitive, and hyperbolic with local product structure. Then*

$$FS(C^0) := \{f \in C^0 : \textit{every } f\textit{-maximizing measure has full support}\}$$

is a residual subset of C^0.

If X is infinite then $FS(C^0)$ has empty interior.

Proof. If X is finite then by transitivity it is a single periodic orbit, so $\mathcal{M}_T$ is a singleton. The unique invariant measure is equi-distributed on X, and in particular has full support, so $FS(C^0) = C^0$.

If X is infinite then let $\{K_i\}$ be a sequence of *proper* compact subsets of X such that $K_1 \subset K_2 \subset \ldots$ and $\cup_i K_i = X$. Such a sequence exists because X is a compact metric space, so has a countable base consisting of proper open subsets. Let

$$C_i = \{f \in C^0 : \text{ some } \mu \in \mathcal{M}_{\max}(f) \text{ has } \mathrm{supp}(\mu) \subset K_i\}\,.$$

Then the complement $FS(C^0)^c$ can be written as the union

$$FS(C^0)^c = \bigcup_{i=1}^{\infty} C_i\,.$$

Our aim is to show that each C_i is closed and has empty interior (i.e. that each C_i^c is open and dense).

First we check that each C_i is closed. Suppose that $f_n \in C_i$, and $f_n \to f$ in C^0. Let $\mu_n \in \mathcal{M}_{\max}(f_n)$ be such that $\mathrm{supp}(\mu_n) \subset K_i$. If $\mu \in \mathcal{M}_T$ is any weak* accumulation point of the sequence μ_n then it is easily seen that μ is f-maximizing. If $\mu_{n_j} \to \mu$ as $j \to \infty$ then since K_i is closed, [Bil, Thm. 2.1] implies that $\mu(K_i) \geq \lim_{j\to\infty} \mu_{n_j}(K_i) = 1$. Therefore $\mathrm{supp}(\mu) \subset K_i$, and hence $f \in C_i$, so C_i is indeed closed.

Now we show that C_i has empty interior. First note that in fact $C_i = \{f \in C^0 : \text{ some } \mu \in \mathcal{M}_{\max}(f) \text{ has } \mathrm{supp}(\mu) \subset \bigcap_{n=0}^{\infty} T^{-n}K_i\}$, since $\mathrm{supp}(\mu)$ is a closed T-invariant set for every $\mu \in \mathcal{M}_T$, and $\bigcap_{n=0}^{\infty} T^{-n}K_i$ is the largest closed T-invariant subset of K_i. Therefore we may assume, without loss of generality, that K_i itself is invariant, i.e. that $TK_i \subset K_i$.

If $f \in C_i$ then there exists $\mu \in \mathcal{M}_{\max}(f)$ with $\mathrm{supp}(\mu) \subset K_i$. By Lemma 4.1, μ can be weak* approximated by periodic orbits which do not intersect the proper closed T-invariant set K_i. So for any $\varepsilon > 0$ we can find a periodic orbit measure μ_ε, with $\mathrm{supp}(\mu_\varepsilon)$ disjoint from K_i, such that $\int f\, d\mu_\varepsilon \geq \alpha(f) - \varepsilon$.

We want to find a continuous function f_ε, close to f, such that no measure ν with support in K_i is f_ε-maximizing. For this it will suffice to show that $\int f_\varepsilon\, d\mu_\varepsilon >$

$\int f_\varepsilon \, d\nu$. Such an f_ε can be constructed as a perturbation of f, by increasing its value on $\mathrm{supp}(\mu_\varepsilon)$ and leaving it unchanged on K_i. More precisely, the disjointness of the closed sets $\mathrm{supp}(\mu_\varepsilon)$ and K_i ensures there exists a Urysohn function $g_\varepsilon \in C^0$ such that $g_\varepsilon \equiv 0$ on K_i, $g_\varepsilon \equiv 1$ on $\mathrm{supp}(\mu_\varepsilon)$, and $0 \le g_\varepsilon \le 1$ everywhere. Define $f_\varepsilon = f + 2\varepsilon g_\varepsilon$. Then

$$\begin{aligned}\int f_\varepsilon \, d\mu_\varepsilon &= \int f \, d\mu_\varepsilon + 2\varepsilon \int g_\varepsilon \, d\mu_\varepsilon \\ &= \int f \, d\mu_\varepsilon + 2\varepsilon \\ &\ge \alpha(f) + \varepsilon \,,\end{aligned}$$

while if $\nu \in \mathcal{M}_T$ is such that $\mathrm{supp}(\nu) \subset K_i$ then

$$\begin{aligned}\int f_\varepsilon \, d\nu &= \int f \, d\nu + 2\varepsilon \int g_\varepsilon \, d\nu \\ &= \int f \, d\nu \\ &\le \alpha(f) \,.\end{aligned}$$

So a measure ν with $\mathrm{supp}(\nu) \subset K_i$ cannot be f_ε-maximizing, because $\int f_\varepsilon \, d\nu < \int f_\varepsilon \, d\mu_\varepsilon$, and therefore $f_\varepsilon \notin C_i$. But $\|f_\varepsilon - f\|_\infty = 2\varepsilon$, so $f_\varepsilon \to f$ as $\varepsilon \to 0$. Therefore f is not an interior point of C_i, so C_i has empty interior. Thus $FS(C^0)$ is a residual subset of C^0.

The fact that $FS(C^0)$ has empty interior when X is infinite will follow from Theorem 4.5, which asserts in particular that those Lipschitz functions which do *not* have a fully supported maximizing measure form a dense subset of $C^{0,1}$. This subset is densely embedded in C^0, because $C^{0,1}$ is, so $FS(C^0)$ does not have interior. □

Since the intersection of two residual sets is itself residual, we may combine Theorems 3.2 and 4.2 to deduce:

Corollary 4.3. *Let $T : X \to X$ be transitive, and hyperbolic with local product structure. A generic C^0 function has a unique maximizing measure, and this measure has full support.*

In spite of Corollary 4.3, the following generalisation of Problem 3.9 is open:

Problem 4.4. *Let $T : X \to X$ be any transitive hyperbolic map with local product structure. Find an explicit example of a continuous function with a* unique *maximizing measure of full support.*

The transitivity assumption in Corollary 4.3 and Problem 4.4 is clearly a necessary one: without it there are no fully supported uniquely maximizing measures, since any such measure μ is necessarily ergodic, by Proposition 2.4 (iii), and so $T|_{\mathrm{supp}(\mu)}$ is transitive.

One of the reasons why Problems 3.9 and 4.4 are open is that any continuous function with a unique fully supported maximizing measure must necessarily be rather irregular; in particular it cannot be Hölder, by results to be described in §4.2.

4.2. **Generic properties in $C^{r,\alpha}$.** Now we turn to the generic properties of maximizing measures in the spaces $C^{r,\alpha}$, for $(r,\alpha) > (0,0)$ (i.e. any space in the scale $\{C^{r,\alpha} : (r,\alpha) \in \mathbb{Z}_{\ge 0} \times [0,1]\}$ *except* for the space $C^0 = C^{0,0}$ of continuous functions). It turns out that these properties are very different from those in C^0.

Theorem 4.5. *Let $T : X \to X$ be transitive, and hyperbolic with local product structure. Suppose that X does not consist of a single periodic orbit. If $(r, \alpha) > (0, 0)$ then the set*

$$NFS(C^{r,\alpha}) := \{f \in C^{r,\alpha} : f \text{ has no fully supported maximizing measure}\}$$

is open and dense in $C^{r,\alpha}$.

Indeed the only functions in $C^{r,\alpha}$ with a fully supported maximizing measure are those for which every invariant measure is maximizing, namely those in the proper closed subspace

$$EC(C^{r,\alpha}) := \{f \in C^{r,\alpha} : f = c + \varphi - \varphi \circ T \text{ for some } c \in \mathbb{R},\, \varphi \in C^0\}$$

of essential coboundaries.

Definition 4.6. Let $f : X \to \mathbb{R}$ be continuous. We say a continuous function $\widetilde{f}$ is a *normal form*[13] for f if $\int \widetilde{f}\, d\mu = \int f\, d\mu$ for all $\mu \in \mathcal{M}_T$, and $\widetilde{f} \le \alpha(f)$.

The usefulness of a normal form is evident: the condition that $\int \widetilde{f}\, d\mu = \int f\, d\mu$ for all $\mu \in \mathcal{M}_T$ means that the maximizing measures for $\widetilde{f}$ are the same as the maximizing measures for f, and in particular that $\alpha(\widetilde{f}) = \alpha(f)$, so in fact $\widetilde{f} \le \alpha(\widetilde{f})$. So if a normal form $\widetilde{f}$ exists then the f-maximizing measures are identified as precisely those invariant measures whose support is contained in the set of maxima of $\widetilde{f}$. This reduction is useful in *specific* problems (for example the one described in §2.4), where explicit information about the set of maxima of $\widetilde{f}$ may be available. It is also useful in the more general context of this section: for example Theorem 4.5 will follow readily from the following important *normal form theorem*.

Theorem 4.7. *Let $T : X \to X$ be transitive, and hyperbolic with local product structure. Every Hölder function $f : X \to \mathbb{R}$ has a normal form: there exists $\varphi \in C^0$ such that $f + \varphi - \varphi \circ T \le \alpha(f)$.*

Proof. We shall give a proof in the special case where T is an expanding[14] map. The proof of the general case can be found in Bousch [B2]. We shall also assume that the function f is Lipschitz; this is simply to ease the exposition, the proof for more general Hölder functions being almost identical.

Suppose we can find a continuous function φ satisfying the equation

$$c + \varphi(x) = \max_{y \in T^{-1}(x)} (f + \varphi)(y) \tag{19}$$

for all $x \in X$ and for some constant $c \in \mathbb{R}$. Replacing x by Tx we see that

$$\begin{aligned} c + \varphi(Tx) &= \max_{y \in T^{-1}(Tx)} (f + \varphi)(y) \\ &\ge (f + \varphi)(x)\,. \end{aligned}$$

[13]This terminology arises because the condition that $\int f\, d\mu = \int g\, d\mu$ for all $\mu \in \mathcal{M}_T$ defines an equivalence relation on C^0, and a normal form is a privileged member of its equivalence class because its maximizing measures are readily apparent. Note, however, that in general a normal form is not unique.

[14]Here our definition of *expanding* is that there exist $\gamma > 1$, $\delta > 0$, such that if $d(x, y) < \delta$ then $d(Tx, Ty) \ge \gamma d(x, y)$. This differs from the definition of §2.5 in two respects: there is no differentiability assumption on T, and the expansion is witnessed *on the first iterate*. Although the apparent loss of generality can be easily remedied, we shall not concern ourselves with this since in any case we are not striving for the full generality of [B2].

That is, $(f+\varphi-\varphi\circ T)(x)\le c$, with equality if and only if the point x is such that $\max_{y\in T^{-1}(Tx)}(f+\varphi)(y)=(f+\varphi)(x)$. We claim that the set

$$Z=\left\{x\in X:\max_{y\in T^{-1}(Tx)}(f+\varphi)(y)=(f+\varphi)(x)\right\}$$

of such points contains a non-empty compact T-invariant set, from which it follows that $c=\alpha(f)$, and hence that

$$f+\varphi-\varphi\circ T$$

is a normal form for f. To prove the claim note that each $x\in Z$ has at least one pre-image in Z, so every finite intersection $\cap_{n=0}^{N}T^{-n}Z$ is non-empty. Therefore the compact T-invariant set $\cap_{n=0}^{\infty}T^{-n}Z$ is the intersection of a decreasing sequence of non-empty compacta, hence is itself non-empty.

It remains to show that there does exist a continuous φ satisfying (19). By introducing the (nonlinear) operator M_f, defined by

$$M_f\varphi(x)=\max_{y\in T^{-1}(x)}(f+\varphi)(y)\,, \tag{20}$$

we can re-cast (19) as a "fixed point equation"

$$M_f\varphi=\varphi+c\,. \tag{21}$$

The existence of a continuous solution to (21) can be proved in a number of ways. The approach here is based on [B2] and consists of two steps: (i) show that M_f has "approximate" fixed points $\varphi_\lambda\in C^0$ satisfying $\varphi_\lambda=M_f(\lambda\varphi_\lambda)$, for any $0\le\lambda<1$; (ii) after quotienting by constants, the family $(\varphi_\lambda)_{0\le\lambda<1}$ has an accumulation point φ which moreover satisfies (21).

First note that the operator M_f is non-increasing on the space C^0:

$$\|M_f\varphi-M_f\psi\|_\infty\le\|\varphi-\psi\|_\infty \tag{22}$$

for all $\varphi,\psi\in C^0$. Indeed if $M_f\varphi(x)=(f+\varphi)(y)$ for $y\in T^{-1}(x)$, and $M_f\psi(x)=(f+\psi)(z)$ for $z\in T^{-1}(x)$, then $(f+\psi)(z)\ge(f+\psi)(y)$, so

$$\begin{aligned}M_f\varphi(x)-M_f\psi(x)&=(f+\varphi)(y)-(f+\psi)(z)\\&\le(f+\varphi)(y)-(f+\psi)(y)\\&=\varphi(y)-\psi(y)\\&\le\|\varphi-\psi\|_\infty\,,\end{aligned}$$

and the reverse inequality is obtained by exchanging the roles of φ and ψ, so (22) follows. The operator $\varphi\mapsto M_f(\lambda\varphi)$ is therefore a strict contraction for $0\le\lambda<1$, and by the contraction mapping principle it has a unique fixed point $\varphi_\lambda\in C^0$.

Since T is expanding, and f is Lipschitz, the functional equation (21) can be used to deduce that each φ_λ is also Lipschitz. More precisely, if f has Lipschitz constant $K>0$ (i.e. $|f(x)-f(y)|\le Kd(x,y)$), and $\theta^{-1}=\gamma>1$ is an expanding constant for T (i.e. $d(Tx,Ty)\ge\theta^{-1}d(x,y)$ whenever x,y are sufficiently close), then it is not hard to show that φ_λ has Lipschitz constant $K\theta/(1-\lambda\theta)$. Importantly this means that the family $(\varphi_\lambda)_{0\le\lambda<1}$ is *uniformly* Lipschitz (with common Lipschitz constant $K\theta/(1-\theta)$).

Uniform Lipschitzness gives both local and global control over this family. The global control is that the oscillation

$$\mathrm{Osc}(\varphi_\lambda)=\max_{x,x'\in X}|\varphi_\lambda(x)-\varphi_\lambda(x')|\le\mathrm{diam}(X)\times K\theta/(1-\theta)$$

is bounded independently of λ. Consequently the functions $\varphi_\lambda^* := \varphi_\lambda - \min_x \varphi_\lambda(x)$, translated so that their minimum value is always zero, form a uniformly bounded family. The local control is the fact that $(\varphi_\lambda)_{0\le\lambda<1}$ is an equicontinuous family, hence so is $(\varphi_\lambda^*)_{0\le\lambda<1}$. By the Ascoli-Arzelá theorem [Roy, p. 169], the uniformly bounded equicontinuous family $(\varphi_\lambda^*)_{0\le\lambda<1}$ has an accumulation point φ in C^0.

If $c_\lambda := \min_x \varphi_\lambda(x)$ then

$$\begin{aligned} M_f(\lambda\varphi_\lambda^*) &= M_f(\lambda(\varphi_\lambda - c_\lambda)) \\ &= M_f(\lambda\varphi_\lambda) - \lambda c_\lambda \\ &= \varphi_\lambda - \lambda c_\lambda \\ &= \varphi_\lambda^* + c_\lambda(1-\lambda)\,, \end{aligned}$$

and since M_f is continuous we may let $\lambda \nearrow 1$ along an appropriate subsequence to see that $M_f\varphi = \varphi + c$, where $c = \lim c_\lambda(1-\lambda)$. This proves (21), and completes the proof of the theorem. □

Before proving Theorem 4.5 we require one extra ingredient, the following well known result of Livšic [Livš], which can in fact be deduced from Theorem 4.7.

Lemma 4.8. *Let $T : X \to X$ be transitive, and hyperbolic with local product structure. Suppose $f : X \to \mathbb{R}$ is Hölder. If $\int f\,d\mu = 0$ for every $\mu \in \mathcal{M}_T$ then there exists $\varphi \in C^0$ such that $f = \varphi - \varphi\circ T$.*

Proof. Since f is Hölder, and $\alpha(f) = 0$, Theorem 4.7 implies there exists $\psi \in C^0$ such that $f + \psi - \psi\circ T \le 0$. Similarly $\alpha(-f) = 0$ so there exists $\varphi \in C^0$ such that $-f + \varphi - \varphi\circ T \le 0$. Adding the two inequalities gives $(\psi+\varphi) \le (\psi+\varphi)\circ T$. But T is transitive, so $\psi+\varphi$ is constant. Therefore $f = \varphi - \varphi\circ T$, as required. □

Of course the converse of Livšic's result is trivially true: if $f = \varphi - \varphi\circ T$ then $\int f\,d\mu = 0$ for all $\mu \in \mathcal{M}_T$.

We can now prove Theorem 4.5, that for $(r,\alpha) > (0,0)$, an open dense subset of functions in $C^{r,\alpha}$ have no fully supported maximizing measure.

Proof of Theorem 4.5. By Theorem 4.7, every $f \in C^{r,\alpha}$ has a normal form $f + \varphi - \varphi\circ T$. Therefore the f-maximizing measures are precisely those invariant measures whose support is contained in the set of maxima of $f+\varphi-\varphi\circ T$. So an f-maximizing measure is fully supported if and only if every $x \in X$ is a maximum for $f+\varphi-\varphi\circ T$, i.e. if and only if $f + \varphi - \varphi\circ T$ is a constant, i.e. if and only if f is an essential coboundary.

The set $EC(C^{r,\alpha})$ of essential coboundaries is a *proper* subset of $C^{r,\alpha}$: since X is not a single periodic orbit it is easy to find an $f \in C^{r,\alpha}$, and two periodic orbit measures μ_1, μ_2, such that $\int f\,d\mu_1 \ne \int f\,d\mu_2$. But $EC(C^{r,\alpha})$ is also a vector subspace, so must have empty interior. To see that $EC(C^{r,\alpha})$ is closed in $C^{r,\alpha}$, let $f_n = c_n + \varphi_n - \varphi_n\circ T \in EC(C^{r,\alpha})$ and suppose that $f_n \to f$ in the topology of $C^{r,\alpha}$. Now $\int f_n\,d\mu = c_n$ for each $\mu \in \mathcal{M}_T$, and the sequence c_n converges to some real number c. In fact $\int f\,d\mu = c$ for all $\mu \in \mathcal{M}_T$, since $f_n \to f$ in C^0 and μ is a continuous functional on C^0. Since f is Hölder, Lemma 4.8 means there exists $\varphi \in C^0$ such that $f - c = \varphi - \varphi\circ T$, so $f \in EC(C^{r,\alpha})$. □

Corollary 4.9. *Suppose that $T : X \to X$ is hyperbolic with local product structure, and X does not consist of a single periodic orbit. Suppose that $(r,\alpha) > (0,0)$.*
(i) *A generic function in $C^{r,\alpha}$ has a unique maximizing measure, and this measure is not fully supported.*

(ii) *If $T : X \to X$ is transitive then a generic function in $C^{r,\alpha}$ has a unique maximizing measure, and this measure is strictly ergodic.*

Proof. By Theorem 3.2 the set $\mathcal{U}(C^{r,\alpha})$, consisting of $C^{r,\alpha}$ functions with a unique maximizing measure, is residual in $C^{r,\alpha}$. If $T : X \to X$ is not transitive then no ergodic measure is fully supported, hence neither is the unique maximizing measure for each function in $\mathcal{U}(C^{r,\alpha})$. Now suppose that $T : X \to X$ is transitive, and that $f \in \mathcal{U}(C^{r,\alpha})$. By Theorem 4.7 an invariant measure is f-maximizing if and only if its support lies in the set of maxima of a normal form $f+\varphi-\varphi\circ T$. But there is only one f-maximizing measure, μ say, so in particular μ is the only invariant measure whose support is contained in $\mathrm{supp}(\mu)$. So μ is strictly ergodic, and therefore not fully supported. $\square$

Remark 4.10. The transitivity assumption in Corollary 4.9 (ii) can be weakened. In fact the same conclusion holds if the non-wandering set $\Omega(T)$ is hyperbolic and such that every transitive closed invariant subset of $\Omega(T)$ is contained in a transitive closed invariant subset of $\Omega(T)$ which has local product structure. In particular this is the case if $\Omega(T)$ is the union $\Omega_1\cup\ldots\cup\Omega_N$ of finitely many closed, pairwise disjoint, T-invariant sets Ω_i, each of which is transitive and has local product structure. An Axiom A map (i.e. a C^1 map on an open subset of a compact Riemannian manifold whose non-wandering set is hyperbolic and equals the closure of the set of periodic points) has this property (see [Sma], [Bow], [Rue2]).

4.3. **Generic periodic maximization?** By Corollary 4.9, if $T : X \to X$ is hyperbolic with local product structure, and $(r, \alpha) > (0, 0)$, then for a generic function in $C^{r,\alpha}$, the maximizing measure is unique, and supported on a proper closed invariant subset of X. A major conjecture is that this result can be strengthened to assert that the support of the maximizing measure is actually a periodic orbit. For simplicity we shall focus attention on this conjecture in the particular case of Lipschitz functions (i.e. $C^{r,\alpha} = C^{0,1}$), and always assume T to be transitive.

Conjecture 4.11. *Let $T : X \to X$ be transitive, and hyperbolic with local product structure. Let $\mathrm{Per}(C^{0,1})$ denote the set of Lipschitz functions with a periodic maximizing measure. Then $\mathrm{Per}(C^{0,1})$ contains an open dense subset of $C^{0,1}$.*

In fact it is already known that $\mathrm{Per}(C^{0,1})$ contains an open set. Indeed if μ is a periodic orbit for *any* continuous map T then the set

$$C^{0,1}(\mu) = \{f \in C^{0,1} : \mu \in \mathcal{M}_{\max}(f)\}$$

has interior:

Proposition 4.12. *Let $T : X \to X$ be a continuous map on a compact metric space. If μ is a periodic orbit for T then $C^{0,1}(\mu)$ is a closed set with non-empty interior.*

In particular, $\mathrm{Per}(C^{0,1})$ has non-empty interior.

Proof. If μ is *any* T-invariant probability measure then $C^{0,1}(\mu)$ is easily seen to be closed as a subset of C^0, and hence as a subset of $C^{0,1}$.

We shall prove that $C^{0,1}(\mu)$ has interior in the special case where μ is a *fixed point* p, leaving the reader to extend the result to more general periodic orbits. We claim that the Lipschitz function $f(x) = -d(x, p)$ is in the interior of $C^{0,1}(\mu)$. Note that f attains its unique maximum value 0 at the fixed point p, so its unique maximizing measure is μ.

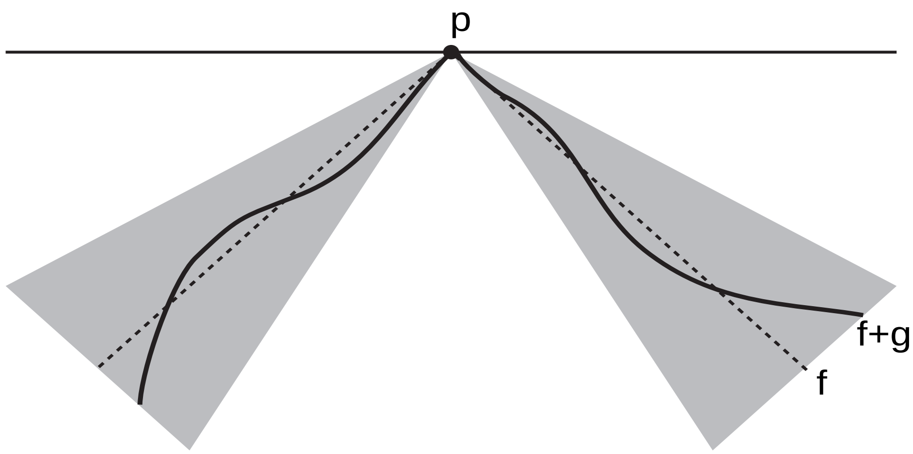

FIGURE 3. A fixed point p is stably maximizing

Now suppose that $f+g$ is a sufficiently small Lipschitz perturbation of f (i.e. g has small Lipschitz norm). Since adding a constant to a function does not change its maximizing measure, we may suppose that g, and hence $f+g$, also vanishes at p. But $f+g$ is Lipschitz-close to f, so its graph must lie within the shaded cones in Figure 3. Therefore $f+g$ also attains its unique maximum at the fixed point p, and so μ is also the unique maximizing measure for $f+g$. More formally, if $\|g\|_{C^{0,1}} \le 1/2$, say, then $g(x) = g(x) - g(p) \le \frac{1}{2}d(x,p)$, so

$$(f+g)(x) = -d(x,p) + g(x) \le -\frac{1}{2}d(x,p)\,.$$

So $f+g$ is a non-positive function attaining its unique maximum value 0 at p, and hence $f+g \in C^{0,1}(\mu)$. □

In view of Proposition 4.12, the outstanding part of Conjecture 4.11 is to show that the interior of $\mathrm{Per}(C^{0,1})$ is dense in $C^{0,1}$. In fact since $\mathrm{Per}(C^{0,1})$ is the union of the $C^{0,1}(\mu)$, for μ periodic, and each of these sets is the closure of its interior, it is enough to show that $\mathrm{Per}(C^{0,1})$ is itself dense in $C^{0,1}$. Progress towards this goal has been made by Yuan & Hunt [YH], who proved:

Theorem 4.13. *Let T be an Axiom A diffeomorphism. If $\mu \in \mathcal{M}_T$ is non-periodic then $C^{0,1}(\mu)$ has empty interior.*

In the case of smooth expanding maps of the circle, Contreras, Lopes & Thieullen [CLT] identified, for each $0 < \alpha < 1$, a certain subspace $C^{\alpha+}$ of $C^{0,\alpha}$ on which the analogue of Conjecture 4.11 can be established. This space consists of those functions f such that for all $\varepsilon > 0$ there exists $\delta > 0$ such that if $d(x,y) < \delta$ then $|f(x) - f(y)| < \varepsilon d(x,y)^\alpha$. Equipped with the $C^{0,\alpha}$ topology, $C^{\alpha+}$ is a closed subspace of $C^{0,\alpha}$, with infinite dimension and infinite codimension.

Theorem 4.14. *Let $T : \mathbb{T} \to \mathbb{T}$ be an expanding map of the circle, and let $\mathrm{Per}(C^{\alpha+})$ denote the set of functions in $C^{\alpha+}$ with a periodic maximizing measure. Then $\mathrm{Per}(C^{\alpha+})$ contains an open dense subset of $C^{\alpha+}$.*

In the case where $T : X \to X$ is a Bernoulli shift, Bousch [B2] has established the analogue of Conjecture 4.11 in a certain Banach space containing all Hölder functions. This space, denoted $\mathrm{Wal}(X,T)$, consists of all functions satisfying *Walters' condition*, a certain "dynamical Hölder" condition first studied by Walters [Wa1]. More precisely, a function satisfies Walters' condition if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for all $n \in \mathbb{N}$, $x, y \in X$,

$$d_n(x,y) \leq \delta \Rightarrow |S_n f(x) - S_n f(y)| \leq \varepsilon\,,$$

where $d_n(x,y) = \max_{0\leq i\leq n-1} d(T^i x, T^i y)$. The space $\mathrm{Wal}(X,T)$ is a Banach space (see [B2]) when equipped with the norm $\|f\| = \|f\|_\infty + |f|_W$, where

$$|f|_W = \sup\left\{|S_n f(x) - S_n f(y)| : n \in \mathbb{N}, d_n(x,y) \leq \mathrm{diam}(X)/2\right\}\,.$$

Bousch [B2] proves:

Theorem 4.15. *Let $T : X \to X$ be a Bernoulli shift, and let* $\mathrm{Per}(\mathrm{Wal})$ *denote the set of functions in* $\mathrm{Wal}(X,T)$ *with a periodic maximizing measure. Then* $\mathrm{Per}(\mathrm{Wal})$ *contains an open dense subset of* $\mathrm{Wal}(X,T)$.

The proof of Theorem 4.15 uses the fact that, since X is a Bernoulli shift equipped with the compact-open topology, the (countable) set of characteristic functions of cylinder sets is dense in $\mathrm{Wal}(X,T)$. This fact has no analogue if $T : X \to X$ is a *smooth* hyperbolic map, so that Theorem 4.15 does not immediately generalise to the smooth setting.

5. **Bibliographical notes. §1:** *Maximizing measures* have also been called *maximal* [J1], or *optimal* [HO, J4, YH], and some authors (e.g. [CLT]) prefer minimization rather than maximization, especially if the problem is derived from classical mechanics (e.g. as in [Fa, M2, M3, Mat]).

§2: The notions of *maximum hitting frequency* and *fastest mean return time* in §2.2 were introduced in [J5]. Propositions 2.1, 2.2 and 2.4 in §2.3 are fairly routine, though the extension to upper semi-continuous f does not seem to be in the literature. For continuous f, Proposition 2.1 appears in [YH, Lemmas 2.3, 2.4] (a similar result is [CG, Thm. 2.1]), while an analogue for non-compact X appears in [JMU2]. The results in §2.4 are due to Bousch [B1], and had been conjectured in [J1, J2]. The earliest experimental work on this family of functions can be found in [CG], where the maximizing measure for $\theta = 1/2$ is determined, and the maximizing measure for $\theta = 1/4$ is conjectured. More systematic experiments were reported in [HO], where much of the structure of maximizing measures was uncovered.

§3: Theorem 3.2 is something of a folklore result. Several authors (e.g. [CG, J1]) had noted that maximizing measures are tangent functionals to the convex functional $\alpha : B \to \mathbb{R}$, and that if B is a separable Banach space then a theorem of Mazur [Maz] implies that a residual subset of functions $f \in B$ have a unique tangent functional, hence a unique maximizing measure. Theorem 3.2 is more general; in particular it applies to non-separable spaces such as $C^{0,\alpha}$. Our method of proof is an elaboration of the one used in [B2, Prop. 9] to prove the case $E = C^0$. A version of Theorem 3.2 valid for Banach spaces appears in [CLT]. Theorem 3.7 was first proved, in a more general setting, in [IP] (see also [Ph2]). In the context of maximizing measures, Theorems 3.7 and 3.8 are proved in [J6].

§4: Theorem 4.2 is from [BJ, §3]. The normal form theorem, Theorem 4.7, is due to Bousch [B2] in the generality stated here. In fact he established the result more generally for maps T with *weak local product structure* and functions f satisfying *Walters' condition*. Our statement of Theorem 4.7 follows because if T is hyperbolic then it has weak local product structure and every Hölder function is Walters[15]. Our proof of Theorem 4.7 in the special case of expanding maps is based on [B2] (see also [JMU3]). An alternative proof in this context is to show that the operator M_f preserves some ball in the space of Lipschitz functions modulo constants, then use the Schauder-Tychonov fixed point theorem (see [B1, J4]).

The earliest version of Theorem 4.7 seems to be in the unpublished preprint of Conze & Guivarc'h [CG], in which it is proved for T a subshift of finite type and f Hölder. Savchenko [Sav] rediscovered the result in the same context, and like Conze & Guivarc'h his proof used the observation that certain maximizing measures can be seen as "zero temperature limits" of equilibrium measures (see [Br1, Coe, CLT, J3, JMU1, PS1] for further investigations of such zero temperature limits[16]). Bousch [B1] gave a more direct proof of Theorem 4.7 in the context of circle expanding maps, as did Contreras, Lopes & Thieullen [CLT], who were inspired by an analogous result of Mañé [M2, M3] in the context of certain Lagrangian systems first considered by Mather [Mat] (see Fathi [Fa] for a strengthening of Mañé's result). A version of Theorem 4.7 for functions of summable variation on finite alphabet subshifts of finite type appears in [J4], and extra hypotheses on f allow a generalisation to the case of infinite alphabets [JMU3]. Lopes & Thieullen [LT1] established a version of Theorem 4.7 for T an Anosov diffeomorphism, as well as an analogue for Anosov flows [LT2]. A similar result for Anosov flows has also been obtained by Pollicott & Sharp [PS2]. Souza [Sou] has proved a version of Theorem 4.7 in the case where T is an interval map with an indifferent fixed point.

The proof of Lemma 4.8 given here appears in [B2], while the standard proof can be found in [Livš], [PP, p. 45]. Perhaps the most natural context for Livšic's theorem is as a special case of a result of Bousch [B3] which asserts that if T is hyperbolic and f is Hölder then there exists $\varphi \in C^0$ such that $(f + \varphi - \varphi \circ T)(X) = [-\alpha(-f), \alpha(f)]$.

Acknowledgments. These notes grew out of a course given at the Ecole Pluri-thématique de Théorie Ergodique at CIRM in May 2004. I would like to thank the organisers for the invitation to speak there, and for their hospitality during my stay.

REFERENCES

[AAS] J. F. Alves, V. Araújo, & B. Saussol, On the uniform hyperbolicity of some non-uniformly hyperbolic systems, *Proc. Amer. Math. Soc.*, **131** (2003), 1303–1309.

[BD] H. van Beijeren & J. R. Dorfman, A note on the Ruelle pressure for a dilute disordered Sinai billiard, *J. Stat. Phys.*, **108** (2002), 767–785.

[Bil] P. Billingsley, *Convergence of probability measures (second edition)*, Wiley, 1999.

[B1] T. Bousch, Le poisson n'a pas d'arêtes, *Ann. Inst. Henri Poincaré (Proba. et Stat.)* **36**, (2000), 489–508.

[B2] T. Bousch, La condition de Walters, *Ann. Sci. ENS*, **34**, (2001), 287–311.

[B3] T. Bousch, Un lemme de Mañé bilatéral, *Comptes Rendus de l'Académie des Sciences de Paris, série I*, **335** (2002), 533–536.

[BJ] T. Bousch & O. Jenkinson, Cohomology classes of dynamically non-negative C^k functions, *Invent. Math.*, **148** (2002), 207–217.

[15] The class of functions satisfying Walters' condition depends on the map T. For more general maps with weak local product structure a Hölder function need not satisfy Walters' condition.

[16] The main open question here is whether or not the equilibrium measures for the function tf always *converge* as $t \to \infty$. Convergence is guaranteed whenever $\mathcal{M}_{\max}(f)$ is a singleton, or more generally if there is a unique maximizing measure which maximizes entropy within $\mathcal{M}_{\max}(f)$.

[Bow] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Springer LNM, **470**, Berlin-Heidelberg-New York, 1975.

[Br1] J. Brémont, *On the behaviour of Gibbs measures at temperature zero*, *Nonlinearity* **16** (2003), 419–426.

[Br2] J. Brémont, Dynamics of injective quasi-contractions, *Ergod. Th. & Dyn. Sys.*, to appear.

[BS] S. Bullett and P. Sentenac, Ordered orbits of the shift, square roots, and the devil's staircase, *Math. Proc. Camb. Phil. Soc.*, **115** (1994), 451–481.

[Cao] Y. Cao, Non-zero Lyapunov exponents and uniform hyperbolicity, *Nonlinearity*, **16** (2003), 1473–1479.

[CLR] Y. Cao, S. Luzzatto & I. Rios, A minimum principle for Lyapunov exponents and a higher-dimensional version of a theorem of Mañé, *Qual. Theory Dyn. Syst.*, to appear.

[Coe] Z. N. Coelho, Entropy and ergodicity of skew-products over subshifts of finite type and central limit asymptotics, *Ph.D. Thesis*, Warwick University, (1990).

[CLT] G. Contreras, A. Lopes, & Ph. Thieullen, Lyapunov minimizing measures for expanding maps of the circle, *Ergod. Th. & Dyn. Sys.*, **21** (2001), 1379–1409.

[CG] J.-P. Conze & Y. Guivarc'h, Croissance des sommes ergodiques, *manuscript*, circa 1993.

[Eng] R. Engelking, *General Topology*, Heldermann Verlag, Berlin, 1989.

[Fa] A. Fathi, Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens, *C. R. Acad. Sci., Paris, Sér. I, Math.* **324**, No.9, (1997) 1043–1046.

[HJ] E. Harriss & O. Jenkinson, in preparation.

[HO] B. R. Hunt and E. Ott, Optimal periodic orbits of chaotic systems occur at low period, *Phys. Rev. E*, **54** (1996), 328–337.

[IP] R. B. Israel & R. R. Phelps, Some convexity questions arising in statistical mechanics, *Math. Scand.*, **54** (1984), 133–156.

[J1] O. Jenkinson, Conjugacy rigidity, cohomological triviality, and barycentres of invariant measures, *Ph. D. thesis*, Warwick University, 1996.

[J2] O. Jenkinson, Frequency locking on the boundary of the barycentre set, *Experimental Mathematics*, **9** (2000), 309–317.

[J3] O. Jenkinson, Geometric barycentres of invariant measures for circle maps, *Ergod. Th. & Dyn. Sys.*, **21** (2001), 511–532.

[J4] O. Jenkinson, Rotation, entropy, and equilibrium states, *Trans. Amer. Math. Soc.*, **353** (2001), 3713–3739.

[J5] O. Jenkinson, Maximum hitting frequency and fastest mean return time, *Nonlinearity*, **18** (2005), 2305–2321.

[J6] O. Jenkinson, Every ergodic measure is uniquely maximizing, *preprint*.

[JMU1] O. Jenkinson, R. D. Mauldin & M. Urbański, Zero temperature limits of Gibbs-equilibrium states for countable alphabet subshifts of finite type, *J. Stat. Phys.*, **119** (2005), 765–776.

[JMU2] O. Jenkinson, R. D. Mauldin & M. Urbański, Ergodic optimization for non-compact dynamical systems, *preprint*.

[JMU3] O. Jenkinson, R. D. Mauldin & M. Urbański, Ergodic optimization for countable alphabet subshifts of finite type, *preprint*.

[Kap] L. Kaplan, Wavefunction intensity statistics from unstable periodic orbits, *Phys. Rev. Lett.*, **80** (1998), 2582–2585.

[KH] A. Katok & B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, Cambridge University Press, 1995.

[Liv] C. Liverani, Rigorous numerical investigation of the statistical properties of piecewise expanding maps. A feasibility study, *Nonlinearity*, **14** (2001), 463–490.

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

[LT1] A. Lopes & Ph. Thieullen, Sub-actions for Anosov diffeomorphisms, *Geometric methods in dynamics II. Astérisque vol. 287*, 2003

[LT2] A. Lopes & Ph. Thieullen, Sub-actions for Anosov flows, *Ergod. Th. & Dyn. Sys.*, **25** (2005), 605–628.

[M1] R. Mañé, *Ergodic theory and differentiable dynamics*, Springer-Verlag, 1987.

[M2] R. Mañé, On the minimizing measures of Lagrangian dynamical systems, *Nonlinearity*, **5** (1992), 623–638.

[M3] R. Mañé, Generic properties and problems of minimizing measures of Lagrangian systems, *Nonlinearity*, **9** (1996), 273–310.

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

[Maz] S. Mazur, Über konvexe Mengen in linearen normierten Räumen, *Studia Math.*, **4** (1933), 70–84.

[MH] M. Morse and G. A. Hedlund, Symbolic Dynamics II. Sturmian Trajectories, *Amer. J. Math.*, **62** (1940), 1–42.

[PF] N. Pytheas Fogg, *Substitutions in dynamics, arithmetics and combinatorics*, Springer Lecture Notes in Mathematics vol. 1794, 2002.

[Ph1] R. R. Phelps, *Lectures on Choquet's theorem*, Math. Studies, no. 7, Van Nostrand, 1966.

[Ph2] R. R. Phelps, Unique equilibrium states, in *Dynamics and randomness (Santiago, 2000)*, 219–225, Nonlinear Phenom. Complex Systems, 7, Kluwer Acad. Publ., Dordrecht, 2002.

[PP] W. Parry & M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics,* Astérisque, 187–188, 1990.

[Pr] F. Przytycki, Anosov endomorphisms, *Studia Math.*, **58** (1976), 249–285.

[PS1] M. Pollicott & R. Sharp, Rates of recurrence for $\mathbb{Z}^q$ and $\mathbb{R}^q$ extensions of subshifts of finite type, *Jour. London Math. Soc.*, **49** (1994), 401–416.

[PS2] M. Pollicott & R. Sharp, Livsic theorems, maximizing measures and the stable norm, *Dyn. Syst.*, **19** (2004), 75–88.

[Roy] H. L. Royden, *Real Analysis*, 3rd edition, Maxwell Macmillan, 1988.

[Rue1] D. Ruelle, *Thermodynamic Formalism*, Reading, Mass., Addison-Wesley, 1978.

[Rue2] D. Ruelle, *Elements of differentiable dynamics and bifurcation theory*, Academic Press, 1989.

[Sav] S. V. Savchenko, Homological inequalities for finite topological Markov chains, *Funct. Anal. Appl.*, **33** (1999), 236–238.

[Sig] K. Sigmund, Generic properties of invariant measures for Axiom A diffeomorphisms, *Invent. Math.*, **11** (1970), 99–109.

[Sma] S. Smale, Differentiable dynamical systems, *Bull. Amer. Math. Soc.*, **73** (1967), 747–817.

[Sou] R. Souza, Sub-actions for weakly hyperbolic one-dimensional systems, *Dyn. Syst.*, **18** (2003), 165–179.

[SS] R. Sturman & J. Stark, Semi-uniform ergodic theorems and applications to forced systems, *Nonlinearity*, **13** (2000), 113–143.

[Tie] H. Tietze, Über Funktionen, die auf einer abgeschlossenen Menge stetig sind, *J. Reine Angew. Math.*, **145** (1914), 9–14.

[Ton] H. Tong, Some characterizations of normal and perfectly normal spaces, *Duke Math. J.*, **19** (1952), 289–292.

[Wa1] P. Walters, Invariant measures and equilibrium states for some mappings which expand distances, *Trans. Amer. Math. Soc.*, **236** (1978), 127–153.

[Wa2] P. Walters, *An introduction to ergodic theory*, Springer, 1981.

[YH] G. Yuan & B. R. Hunt, Optimal orbits of hyperbolic systems, *Nonlinearity*, **12** (1999), 1207–1224.

E-mail address: `omj@maths.qmul.ac.uk`