Ergod. Th. & Dynam. Sys. (2019), **39**, 2593–2618 © *Cambridge University Press, 2018*

doi:10.1017/etds.2017.142

SURVEY

Ergodic optimization in dynamical systems

OLIVER JENKINSON

School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK
(e-mail: o.jenkinson@qmul.ac.uk)

(*Received* 13 *September* 2016 *and accepted in revised form* 27 *November* 2017)

Abstract. Ergodic optimization is the study of problems relating to maximizing orbits and invariant measures, and maximum ergodic averages. An orbit of a dynamical system is called f-maximizing if the time average of the real-valued function f along the orbit is larger than along all other orbits, and an invariant probability measure is called f-maximizing if it gives f a larger space average than any other invariant probability measure. In this paper, we consider the main strands of ergodic optimization, beginning with an influential model problem, and the interpretation of ergodic optimization as the zero temperature limit of thermodynamic formalism. We describe typical properties of maximizing measures for various spaces of functions, the key tool of adding a coboundary so as to reveal properties of these measures, as well as certain classes of functions where the maximizing measure is known to be Sturmian.

1. *Introduction*

For a real-valued function defined on the state space of a dynamical system, the topic of *ergodic optimization* revolves around understanding its largest possible ergodic average. Taking the dynamical system to be a map $T : X \to X$, and denoting the function by $f : X \to \mathbb{R}$, attention is focused on the supremum of time averages $\lim_{n\to\infty} (1/n) \sum_{i=0}^{n-1} f(T^i x)$ over those $x \in X$ for which the limit exists, or alternatively (and in nice cases equivalently) on the supremum of space averages $\int f\, d\mu$ over probability measures μ which are invariant under T.

In the most classical setting of a topological dynamical system, with X a compact metric space and $T : X \to X$ continuous, and if f is continuous, then the above suprema coincide. Indeed the common value is a maximum, as the weak* compactness of the set $\mathcal{M}_T$ of

T-invariant Borel probability measures guarantees some $m \in \mathcal{M}_T$ satisfying

$$\int f \, dm = \max_{\mu \in \mathcal{M}_T} \int f \, d\mu =: \beta(f), \tag{1}$$

and there exists $x \in X$ with $\lim_{n\to\infty} (1/n) \sum_{i=0}^{n-1} f(T^i x) = \beta(f)$, since m may be taken to be ergodic and x an m-generic point. Any such $m \in \mathcal{M}_T$ is called a *maximizing measure* for f, and $\beta(f)$ is the *maximum ergodic average*.

Ergodic optimization originated in the 1990s, with much early work focused on fixing a specific map T and studying the dependence of the maximizing measure on a function f which varied in some finite dimensional space V. Indeed a certain model problem (see §3) consisting of T the doubling map on the circle, and V the two-dimensional vector space of degree-one trigonometric polynomials, turned out to be influential: various subsequent results were suggested either by the behaviour of this model, or by the techniques used to understand it. In this model, any non-zero function in V has a unique maximizing measure, this measure is usually periodic (i.e. supported on a single periodic orbit), though not always periodic. The natural occurrence of non-periodic maximizing measures was itself somewhat surprising (and had ramifications in related areas [**32**, **33**]), while the apparent rarity of non-periodic maximizing measures anticipated the programme (described here in §7) of establishing analogous results for V an infinite-dimensional function space (e.g. the space of Lipschitz functions) and investigating further generic properties of maximizing measures (see §8).

The specific maximizing measures arising in the model problem of §3, so-called *Sturmian* measures, turned out to be unexpectedly ubiquitous in a variety of ergodic optimization problems (which we describe in §9), encompassing similar low-dimensional function spaces, certain infinite dimensional cones of functions, and problems concerning the joint spectral radius of matrix pairs. Various ideas used to resolve the model problem have been the subject of subsequent research; most notably, the prospect of adding a coboundary to f so as to reveal properties of its maximizing measure has been the cornerstone of much recent work (described in §§5 and 6), with many authors equally inspired by parallels with Lagrangian dynamical systems.

Another significant strand of research in ergodic optimization, again already present in early works, was its interpretation (see §4) as a limiting *zero temperature* version of the more classical thermodynamic formalism, with maximizing measures (referred to as *ground states* by physicists) arising as zero temperature accumulation points of equilibrium measures; work in this area has primarily focused on understanding convergence and non-convergence in the zero temperature limit.

2. *Fundamentals*

Let $\mathfrak{D}$ denote the set of pairs (X, T) where $X = (X, d)$ is a compact metric space and $T : X \to X$ is continuous. For $(X, T) \in \mathfrak{D}$, the set $\mathcal{M}_T$ of T-invariant Borel probability measures is compact when equipped with the weak* topology.

Let $\mathfrak{C}$ denote the set of triples (X, T, f), where $(X, T) \in \mathfrak{D}$ and $f : X \to \mathbb{R}$ is continuous. For X a compact metric space, let $C(X)$ denote the set of continuous real-valued functions on X, equipped with the supremum norm $\|f\|_\infty = \max_{x\in X} |f(x)|$. Let

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

Lip denote the set of Lipschitz real-valued functions on X, with $\mathrm{Lip}(f) := \sup_{x\neq y} |f(x) - f(y)|/d(x, y)$, and Banach norm $\|f\|_{\mathrm{Lip}} = \|f\|_\infty + \mathrm{Lip}(f)$.

Definition 2.1. For $(X, T, f) \in \mathfrak{C}$, the quantity $\beta(f) = \beta(T, f) = \beta(X, T, f)$ defined by

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

is the *maximum ergodic average*. Any $m \in \mathcal{M}_T$ satisfying $\int f\, dm = \beta(f)$ is an f*-maximizing measure*, and $\mathcal{M}_{\max}(f) = \mathcal{M}_{\max}(T, f) = \mathcal{M}_{\max}(X, T, f)$ denotes the collection of such measures.

While we adopt the convention that optimization means maximization, occasional mention will be made of the *minimum ergodic average*

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

and the set $\mathcal{M}_{\min}(f) = \{m \in \mathcal{M}_T : \int f\, dm = \alpha(f)\}$ of *minimizing measures* for f. The closed interval $[\alpha(f), \beta(f)] = \{\int f\, d\mu : \mu \in \mathcal{M}_T\}$ is the *set of ergodic averages*†.

The maximum ergodic average admits a number of alternative characterizations involving time averages (see e.g. [**87**, Proposition 2.2]).

PROPOSITION 2.2. *For $(X, T, f) \in \mathfrak{C}$, the maximum ergodic average $\beta(f)$ satisfies*

$$\beta(f) = \sup_{x\in X_{T,f}} \lim_{n\to\infty} \frac{1}{n} S_n f(x) = \sup_{x\in X} \limsup_{n\to\infty} \frac{1}{n} S_n f(x) = \limsup_{n\to\infty} \frac{1}{n} \sup_{x\in X} S_n f(x), \qquad (2)$$

where $S_n f = \sum_{i=0}^{n-1} f \circ T^i$, and $X_{T,f} = \{x \in X : \lim_{n\to\infty} (1/n) S_n f(x)$ exists$\}$.

The following is well known (see e.g. [**87**, Proposition 2.4]).

PROPOSITION 2.3. *If $(X, T, f) \in \mathfrak{C}$ then:*
(i) *there exists at least one f-maximizing measure;*
(ii) *$\mathcal{M}_{\max}(f)$ is compact;*
(iii) *$\mathcal{M}_{\max}(f)$ is a simplex, and in particular convex;*
(iv) *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.*

In §§7 and 8 we shall consider *typical* properties of maximizing measures in various spaces V of real-valued functions on X. The following result (see e.g. [**87**, Theorem 3.2], and in other forms see [**27**, **51**, **52**]) guarantees that, for all of the function spaces V considered, *uniqueness* of the maximizing measure is typical in V (though clearly there exist $f \in V$ such that $\mathcal{M}_{\max}(f)$ is not a singleton, provided $\mathcal{M}_T$ is not a singleton, most obviously $f \equiv 0$).

THEOREM 2.4. (Typical uniqueness of maximizing measures) *If $(X, T) \in \mathfrak{D}$, and V is a topological vector space which is densely and continuously embedded in $C(X)$, then $\{f \in V : \mathcal{M}_{\max}(f)$ is a singleton$\}$ is a residual subset of V.*

† This set, and its generalization for f taking values in higher-dimensional spaces, is often referred to as the *rotation set* (see e.g. [**20**, **65**, **69**, **84**, **85**, **104**, **105**, **160**]), while in the context of multifractal analysis it is sometimes referred to as the *spectrum* of (Birkhoff) ergodic averages.

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

If T and f are continuous, but X is non-compact, a number of difficulties potentially arise. Assuming $\mathcal{M}_T$ is non-empty, we may define $\beta(f) = \sup_{\mu\in\mathcal{M}_T} \int f\, d\mu$, though in general there need not exist any maximizing measures, and any one of the equalities in (2) may fail to hold (see e.g. [**95**]). The most commonly studied example of a non-compact X is a countable alphabet subshift of finite type, where a number of sufficient conditions have been given for the existence of maximizing measures (see e.g. [**17**, **18**, **78**, **94**, **95**]), while [**56**] includes applications to (non-compact) Julia sets $X \subset \mathbb{C}$ for maps T in the exponential family.

Note that versions of ergodic optimization have also been investigated in discrete time settings slightly different from the one described here, notably the case where $\mathcal{M}_T$ is a singleton (see [**41**]), in the context of non-conventional ergodic averages (see [**8**]), or when the optimization is over a restricted subset of $\mathcal{M}_T$ (see [**162**]). Generalizations of ergodic optimization include optimal tracking for dynamical systems (see [**122**]), and ergodic dominance (see §9).

3. *A model problem*

The map $T(x) = 2x \pmod 1$ on the circle $X = \mathbb{R}/\mathbb{Z}$ is a standard example of a hyperbolic dynamical system, and the functions $f(x) = \cos 2\pi x$ and $g(x) = \sin 2\pi x$ are arguably the most natural non-constant functions on X. While the f-maximizing measure is easily seen to be the Dirac measure at the fixed point 0, the g-maximizing measure is rather less obvious (it turns out to be the periodic measure on the orbit coded by 0001). This standard choice of T, and the naturalness of f and g, prompted several early authors to investigate those T-invariant measures which are maximizing for functions in the two-dimensional vector space V spanned by f and g. A rather complete understanding of this model problem has been provided by Bousch [**26**], following earlier partial progress [**52**, **75**, **76**, **81**–**83**], and the results in this case already point to some more universal features of ergodic optimization.

While the space of degree-one trigonometric polynomials V is two-dimensional, the fact that a measure is maximizing for $v \in V$ if and only if it is maximizing for cv, where $c > 0$, renders the problem a one-dimensional one; specifically, to identify the maximizing measures for functions in V it suffices to determine the maximizing measures for functions on the unit sphere in V, i.e. those of the form $v_\theta(x) = (\cos 2\pi\theta) f(x) + (\sin 2\pi\theta) g(x) = \cos 2\pi(x-\theta)$, for $\theta \in \mathbb{R}/\mathbb{Z}$.

It turns out that every v_θ has a unique maximizing measure, and that this measure is typically periodic: for Lebesgue almost every $\theta \in \mathbb{R}/\mathbb{Z}$, the v_θ-maximizing measure is supported on a single periodic orbit. Periodic maximizing measures are also typical in the topological sense: the set $\{\theta \in \mathbb{R}/\mathbb{Z} : \mathcal{M}_{\max}(v_\theta) \text{ is a periodic singleton}\}$ contains an open dense subset of $\mathbb{R}/\mathbb{Z}$, and consequently $\{v \in V : \mathcal{M}_{\max}(v) \text{ is a periodic singleton}\}$ contains an open dense subset of $V \equiv \mathbb{R}^2$. In summary, this model problem exhibits *typically periodic optimization*, a phenomenon which has subsequently been established for various natural (infinite-dimensional) function spaces V (see §7 for further details).

More can be said about the maximizing measures arising in this specific model problem. The only *periodic* measures which are maximizing for some degree-one trigonometric polynomials are those on which the action of T is combinatorially equivalent to a

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

rational rotation, while the non-periodic measures which are maximizing for some $v \in V$ correspond to irrational rotations (their support is a T-invariant Cantor set reminiscent of those arising for so-called Denjoy counterexamples in the theory of degree-one circle maps, cf. e.g. [**155**]). More precisely, the maximizing measures for (non-zero) functions in V are *Sturmian* measures: the Sturmian measure of rotation number $\varrho \in \mathbb{R}/Z$ is the push forward of the Lebesgue measure on X under the map $x \mapsto \sum_{n\geq 0} \chi_{[1-\varrho,1)}(\{x + n\varrho\})/2^{n+1}$, where $\{\cdot\}$ denotes reduction modulo 1. For example, all T-invariant measures supported on a periodic orbit of period < 4 are Sturmian, though the measure supported on $\{1/5, 2/5, 3/5, 4/5\} \equiv 0011$ is not, and periodic orbits supporting Sturmian measures become increasingly rare as the period grows (see e.g. [**4**, **26**, **43**, **82**, **89**, **90**, **134**] for further details on Sturmian measures and orbits). Bousch [**26**] showed that every Sturmian measure arises as the maximizing measure for some v_θ, and that if ϱ is irrational then $\theta = \theta(\varrho)$ is unique.

The fact that Sturmian measures are precisely the maximizing measures for this model problem does rely, to an extent, on the particular choice of f and g, though the presence of Sturmian measures is not altogether surprising: it has subsequently been shown that Sturmian measures arise naturally as maximizing measures in a variety of similar settings, as will be described in §9.

4. *Ergodic optimization as zero temperature thermodynamic formalism*

Given $(X, T, f) \in \mathfrak{C}$, the *pressure* $P(f) = P(T, f)$ is defined as

$$P(f) = \sup_{m\in\mathcal{M}_T} \left(\int f\, dm + h(m)\right), \tag{3}$$

where $h(m)$ denotes the entropy of m. Any $m \in \mathcal{M}_T$ attaining the supremum in (3) is called an *equilibrium measure* (denoted by m_f if it is unique) for the function f (which in this context is referred to as a *potential*). If f is replaced by tf for $t \in \mathbb{R}$, then the entropy term in the supremum (3) loses relative importance as $t \to \infty$ (the thermodynamic interpretation of the parameter t is as an *inverse temperature*, so that letting $t \to \infty$ is referred to as a *zero temperature limit*). For large values of t, an equilibrium measure for tf is almost maximizing for f, in that its integral is close to the maximum ergodic average $\beta(f)$. More precisely, a number of early authors [**48**, **51**, **52**, **81**, **83**, **138**] observed, in various broadly similar settings (with T hyperbolic and f Hölder continuous, so that m_{tf} exists and is unique) that the family (m_{tf}) has at least one accumulation point m as $t \to \infty$, that m is an f-maximizing measure, and that $\lim_{t\to\infty} h(m_{tf}) = h(m) = \max\{h(\mu) : \mu \in \mathcal{M}_{\max}(f)\}$ (i.e. any zero temperature accumulation point is of maximal entropy among the set of f-maximizing measures). Indeed these conclusions are true in wider generality: if X is compact, and the entropy map $\mu \mapsto h(\mu)$ is upper semi-continuous†, then every continuous function has at least one equilibrium measure (see [**157**, Theorem 9.13(iv)]), and it is not hard to establish the following result.

† Upper semi-continuity of entropy holds if T is expansive (see [**157**]), or more generally if T admits a finite generating partition (see [**101**, Corollary 4.2.5]); in particular, this includes all symbolic systems. Upper semi-continuity is also guaranteed (see [**135**]) whenever T is a C^∞ map of a compact manifold.

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

THEOREM 4.1. (Zero temperature limits as maximal entropy maximizing measures) *Let $(X, T, f) \in \mathfrak{C}$ be such that the entropy map on $\mathcal{M}_T$ is upper semi-continuous. For $t \in \mathbb{R}$, if m_t is an equilibrium measure for tf then (m_t) has at least one accumulation point $m \in \mathcal{M}_T$ as $t \to \infty$, and:*
(i) *m is an f-maximizing measure,*
(ii) *$h(m) = \max\{h(\mu) : \mu \in \mathcal{M}_{\max}(f)\}$,*
(iii) *$\lim_{t\to\infty} h(m_t) = h(m)$.*

In particular, under the hypotheses of Theorem 4.1, if $\mathcal{M}_{\max}(f) = \{m\}$ then $m_t \to m$ as $t \to \infty$, so Theorem 2.4 implies that for typical f the weak* limit $\lim_{t\to\infty} m_t$ exists, and is characterized as being the unique f-maximizing measure. A wider investigation of the nature of the set of accumulation points of (m_t), and of whether $\lim_{t\to\infty} m_t$ *always* exists, was initially focused on the case of (X, T) a subshift of finite type and f locally constant (hypotheses guaranteeing that the unique equilibrium measure m_{tf} is Markov); it was found [**48**, **83**, **140**] that limits $\lim_{t\to\infty} m_{tf}$ are not necessarily ergodic, nor necessarily the evenly weighted centroid of ergodic maximizing measures of maximal entropy. In this setting, the convergence question was resolved by Brémont [**37**], who showed† that the zero temperature limit *does* always exist, even when $\mathcal{M}_{\max}(f)$ is not a singleton.

THEOREM 4.2. [**37**] (Zero temperature convergence for locally constant functions) *For (X, T) a subshift of finite type, and $f : X \to \mathbb{R}$ locally constant, $\lim_{t\to\infty} m_{tf}$ exists; indeed $L_p(X) = \{\lim_{t\to\infty} m_{tf} : f \in C(X)$ depends on p coordinates$\}$ is finite for each $p \in \mathbb{N}$.*

For example, given (X, T) the full shift on two symbols, the set $L_2(X)$ has cardinality 7, and its elements can be listed explicitly (see [**37**]). For larger p, and for other subshifts of finite type (X, T), the set of possible limits $L_p(X)$ becomes harder to describe. Progress on this problem was made initially by Leplaideur [**108**], then by Chazottes, Gambaudo and Ugalde [**45**] and Garibaldi and Thieullen [**68**], using a variety of techniques, and can be summarized as the following.

THEOREM 4.3. (Description of zero temperature limit for locally constant functions) *If (X, T) is a subshift of finite type, and $f : X \to \mathbb{R}$ is locally constant, then $m = \lim_{t\to\infty} m_{tf}$ is concentrated on a certain subshift of finite type X_f which is itself a finite union of transitive subshifts of finite type. The finitely many ergodic components m_i of $m = \sum_{i=1}^{q} \alpha_i m_i$ are unique equilibrium measures of auxiliary potential functions; these potentials, and the weights α_i, can be constructed algorithmically.*

In the more general setting of Lipschitz functions on subshifts of finite type, the question of whether zero temperature limits always exist remained open for several years, being finally‡ settled negatively by Chazottes and Hochman [**46**].

† The paper [**37**] uses ideas from analytic geometry (semi-algebraic and sub-analytic maps) which are outside the standard toolkit of most ergodic theorists, and despite its elegant brevity, the approach of [**37**] has not subsequently been pursued.
‡ It was noted in [**46**] that van Enter and Ruszel [**58**] had already given an example of non-convergence in the zero temperature limit, albeit in a somewhat different context: a nearest neighbour potential model with the shift map acting on a subset of $(\mathbb{R}/\mathbb{Z})^{\mathbb{Z}}$, the significant difference being that the state space $\mathbb{R}/\mathbb{Z}$ is non-discrete.

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

THEOREM 4.4. [**46**] (Zero temperature non-convergence) *For (X, T) the full shift on two symbols, there exist Lipschitz functions $f : X \to \mathbb{R}$ for which $\lim_{t\to\infty} m_{tf}$ does not exist. Indeed such f may be defined as $f(x) = -\mathrm{dist}(x, Y)$, where $Y \subset X$ is a (carefully constructed) subshift.*

The flexibility of the approach in [**46**] allows the full shift in Theorem 4.4 to be replaced by any (one-sided or two-sided) mixing subshift of finite type, and allows the construction of subshifts Y such that the set of accumulation points of (m_{tf}) is e.g. non-convex, or only containing positive entropy measures, or not containing ergodic measures. Bissacot, Garibaldi and Thieullen [**19**] have shown that non-convergence in the zero temperature limit can arise for certain functions on the full two-shift which take only countably many values, and where the only ergodic maximizing measures are the Dirac measures at the two fixed points. Yet another approach to non-convergence in the zero temperature limit has been introduced by Coronel and Rivera-Letelier [**53**], partially based on the methods of [**58**], establishing a certain persistence of the non-convergence phenomenon.

THEOREM 4.5. [**53**] (Persistence of zero temperature non-convergence) *For (X, T) a full shift on a finite alphabet, there exists a Lipschitz function $f_0 : X \to \mathbb{R}$, and complementary open subsets U^+ and U^- of X, such that for any sequence of positive reals $t_i \to \infty$, there is an arbitrarily small Lipschitz perturbation f of f_0 such that the sequence $m_{t_i f}$ has an accumulation point whose support lies in U^+, and an accumulation point whose support lies in U^-.*

Temporarily widening our notion of a dynamical system to include *higher-dimensional* shifts† (i.e. G-actions on $X = F^G$, where $G = \mathbb{Z}^d$ or $\mathbb{N}^d$ for some integer $d \geq 2$, and F is finite), the following result‡ of [**46**] represents an interesting counterpoint to Theorems 4.2 and 4.4.

THEOREM 4.6. [**46**] (Zero temperature non-convergence for locally constant functions on higher-dimensional shifts) *For $d \geq 3$, there exist locally constant functions f on $\{0, 1\}^{\mathbb{Z}^d}$ such that for every family $(m_t)_{t>0}$, where m_t is an equilibrium measure for tf, the limit $\lim_{t\to\infty} m_t$ does not exist.*

For the case of (X, T) a countable alphabet subshift of finite type, where X is non-compact and the entropy map $\mu \mapsto h(\mu)$ is not upper semi-continuous, additional summability and boundedness hypotheses on the locally Hölder function $f : X \to \mathbb{R}$, together with primitivity assumptions on X, ensure the existence and uniqueness of the equilibrium measures m_{tf}, that the family (m_{tf}) does in fact have an accumulation point m, and that $h(m) = \lim_{t\to\infty} h(m_{tf}) = \max\{h(\mu) : \mu \in \mathcal{M}_{\max}(f)\}$ (see [**63**, **93**, **124**]), representing an analogue of Theorem 4.1. If in addition f is locally constant, Kempton [**102**] (see also [**63**]) has established the analogue of Theorem 4.2, guaranteeing the weak* convergence of (m_{tf}) as $t \to \infty$. Iommi and Yayama [**80**] consider almost

† Zero temperature non-convergence results for higher-dimensional shifts are also proved in [**53**].
‡ The proof of Theorem 4.6 in [**46**] relied on work of Hochman [**74**] establishing that certain one-dimensional subshifts can be simulated in *finite type* subshifts of dimension $d = 3$; this fact has now been generalized to dimension $d = 2$ (see [**7**, **57**]), suggesting that Theorem 4.6 is probably valid for all $d \geq 2$ (though certainly not for $d = 1$, in view of Theorem 4.2).

additive sequences $\mathcal{F}$ of continuous functions defined on appropriate countable alphabet subshifts of finite type, proving that the family of equilibrium measures $(m_{t\mathcal{F}})$ is tight (based on [**93**]), hence has a weak* accumulation point, and that any such accumulation point is a maximizing measure for $\mathcal{F}$ (see also [**47**, **64**, **151**, **161**] for general ergodic optimization in the context of sequences of functions $\mathcal{F}$).

Zero temperature limits have been analysed for certain specific families of functions: in [**83**] for T the doubling map and f a degree-one trigonometric polynomial, in [**13**] a specific class of functions defined on the full shift on two symbols and taking countably many values, in [**11**] a one-parameter family of functions defined on the full shift on three symbols, each sharing the same two ergodic maximizing measures, and in [**10**, **112**] for the XY model of statistical mechanics. Connections with large deviation theory have been studied in [**14**, **111**, **113**], and the role of the flatness of the potential function has been investigated in [**109**].

One source of interest in zero temperature limits of equilibrium measures is *multifractal analysis*, i.e. the study of level sets of the form $K_\gamma = \{x \in X : \lim_{n\to\infty} (1/n) S_n f(x) = \gamma\}$. Each K_γ is T-invariant, and the *entropy spectrum of Birkhoff averages*, i.e. the function $H : [\alpha(f), \beta(f)] \to \mathbb{R}_{\geq 0}$ defined by† $H(\gamma) = h_{\mathrm{top}}(K_\gamma)$, is in certain (hyperbolic) settings described by the family of equilibrium measures $(m_{tf})_{t\in\mathbb{R}}$, in the sense that $\Gamma : t \mapsto \int f\, dm_{tf}$ is a homeomorphism $\mathbb{R} \to (\alpha(f), \beta(f))$, and (see [**107**], and e.g. [**25**, **70**, **73**, **136**])

$$H(\gamma) = h(m_{\Gamma^{-1}(\gamma)f}) = \max\left\{h(\mu) : \mu \in \mathcal{M}_T,\ \int f\, d\mu = \gamma\right\} \quad \text{for all } \gamma \in (\alpha(f), \beta(f)).$$

The function H is concave, and extends continuously to the boundary of $[\alpha(f), \beta(f)]$, though the absence of equilibrium measures m_{tf} with $\int f\, dm_{tf}$ on the boundary prompted investigation of extremal measures (see [**48**, **81**, **83**, **140**]), and of the (typical) values $H(\alpha(f))$ and $H(\beta(f))$ (see [**146**]).

Finally, we note that zero temperature limits of equilibrium measures have been studied in a variety of other dynamical settings, including Frenkel–Kontorova models [**6**], quadratic-like holomorphic maps [**54**], multimodal interval maps [**79**] and Hénon-like maps [**152**].

5. *Revelations*

The fundamental problem of ergodic optimization is to say something about maximizing measures. A most satisfactory resolution is to explicitly identify the f-maximizing measure(s) for a given $(X, T, f) \in \mathfrak{C}$, though in some cases we may be content with an approximation to an f-maximizing measure, or a result asserting that $\mathcal{M}_{\max}(f)$ lies in some particular subset of $\mathcal{M}_T$. More generally, for a given $(X, T) \in \mathfrak{D}$ and a subset $U \subset C(X)$, we may hope to identify a subset $\mathcal{N} \subset \mathcal{M}_T$ such that $\mathcal{M}_{\max}(f) \subset \mathcal{N}$ for all $f \in U$, or instead $\mathcal{M}_{\max}(f) \subset \mathcal{N}$ for all f belonging to a large subset of U.

† The topological entropy $h_{\mathrm{top}}(K_\gamma)$ of the (in general non-compact) invariant set K_γ is as defined by Bowen [**34**], or equivalently by Pesin and Pitskel' [**137**].

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

In a variety of such settings, it has been noted that a key technical tool is a function we shall refer† to as a *revelation*, and an associated result we shall refer to as a *revelation theorem* (see §6). First we require the following concept, describing a situation where the ergodic optimization problem is easily solved.

Definition 5.1. Given $(X, T) \in \mathfrak{D}$, we say $f \in C(X)$ is *revealed* if its set of maxima $f^{-1}(\max f)$ contains a compact T-invariant set.

In the (rare) cases when the function f is revealed, it is clear that the maximum ergodic average $\beta(f)$ equals $\max f$, and that the set of f-maximizing measures is precisely the (non-empty) set of T-invariant measures whose support is contained in $f^{-1}(\max f)$.

More generally, if we can find $\psi \in C(X)$ satisfying

$$\int \psi \, d\mu = 0 \quad \text{for all } \mu \in \mathcal{M}_T, \tag{4}$$

and such that $f + \psi$ is revealed, then $\beta(f) = \beta(f + \psi)$ equals $\max(f + \psi)$, and $\mathcal{M}_{\max}(f) = \mathcal{M}_{\max}(f + \psi)$ is precisely the set of T-invariant measures whose support is contained in the set $(f + \psi)^{-1}(\max(f + \psi))$.

A natural choice of function ψ satisfying (4) is a *continuous coboundary*, i.e. $\psi = \varphi - \varphi \circ T$ for some $\varphi \in C(X)$, and the ergodic optimization literature has focused mainly (though not exclusively, see e.g. [**125**]) on such ψ, since for practical purposes it usually suffices. This motivates the following definition.

Definition 5.2. For $(X, T, f) \in \mathfrak{C}$, a continuous coboundary ψ is called a *revelation* if $f + \psi$ is a revealed function, i.e. if

$$(f + \psi)^{-1}(\max(f + \psi)) \text{ contains a compact } T\text{-invariant set.} \tag{5}$$

Formalizing the above discussion, we record the following.

PROPOSITION 5.3. *If ψ is a revelation for $(X, T, f) \in \mathfrak{C}$, then $\beta(f) = \max(f + \psi)$, and*

$$\mathcal{M}_{\max}(f) = \mathcal{M}_{\max}(f + \psi) = \{\mu \in \mathcal{M} : \operatorname{supp}(\mu) \subset (f + \psi)^{-1}(\max(f + \psi))\} \neq \emptyset.$$

A consequence of Proposition 5.3 is that if $(X, T, f) \in \mathfrak{C}$ has a revelation then it enjoys the following property (referred to in [**27**, **125**] as the *subordination principle*): if $\mu \in \mathcal{M}_T$ is f-maximizing, and if the support of $\nu \in \mathcal{M}_T$ is contained in the support of μ, then ν is also f-maximizing.

Example 5.4.

(a) If $T(x) = 2x \pmod 1$, the function $f(x) = (2\cos 2\pi x - 1)(\sin 2\pi x + 1)$ is not revealed, but can be written as $f = g - \psi$ where $g(x) = 2\cos 2\pi x - 1$ is revealed, and $\psi(x) = \sin 2\pi x - \sin 4\pi x$ is a continuous coboundary, hence a revelation for f. The unique f-maximizing measure is therefore the g-maximizing measure, namely the Dirac measure δ_0.

† The terms *revelation* and *revealed function* are introduced here, since despite the ubiquity of these concepts there is as yet no established consensus on terminology. The function $\psi = \varphi - \varphi \circ T$ we call a revelation has previously been referred to as a *sub-coboundary* or the solution of a *sub-cohomology equation*, and the function φ in this context has been called a *sub-action*, a *Barabanov function*, a *transfer function*, or a *maximizing function*. The notion of a revealed function has sometimes gone by the name of a *normal form*, or in the context of joint spectral radius problems corresponds to a *(maximizing) Barabanov norm*.

(b) If (X, T) is the full shift on the alphabet $\{0, 1\}$, and the functions $(f_\theta)_{\theta\in\mathbb{R}}$ are defined by $f_\theta(x) = f_\theta((x_i)_{i=1}^\infty) = \theta x_1 + x_2 - (\theta + 2)x_1x_2$, then f_θ is revealed if and only if $\theta = 1$. For all $c \in \mathbb{R}$, the function $\psi_c(x) = c(x_1 - x_2)$ is a coboundary, and $(f_\theta + \psi_c)(x) = (\theta + c)x_1 + (1 - c)x_2 - (\theta + 2)x_1x_2$. If $\theta > -1$ then $\psi_{(1-\theta)/2}$ is a revelation, and reveals the invariant measure supported on the period-2 orbit to be the unique f_θ-maximizing measure. If $\theta < -1$ then ψ_1 is a revelation, with unique f_θ-maximizing measure the Dirac measure concentrated on the fixed point $\overline{0}$. If $\theta = -1$ then $\psi_{(1-\theta)/2} = \psi_1$ is a revelation, and reveals that the f_{-1}-maximizing measures are those whose support is contained in the golden mean subshift of finite type.

In this article we have chosen to interpret optimization as maximization, while noting that the minimizing measures for f are the maximizing measures for $-f$, and that the minimum ergodic average $\alpha(f) = \min_{\mu\in\mathcal{M}_T} \int f\,d\mu$ is equal to $-\beta(-f) = -\max_{\mu\in\mathcal{M}_T} \int(-f)\,d\mu$. Occasionally there is interest in simultaneously considering the maximization and minimization problems; indeed the above discussion suggests the possibility of simultaneously revealing both the minimizing and maximizing measures, by a judicious choice of revelation. This possibility was considered by Bousch [**28**], who showed (see Theorem 5.5 below) that if the f-maximizing measures can be revealed, and if the f-minimizing measures can be revealed, then indeed it *is* possible to reveal both maximizing and minimizing measures simultaneously.

To make this precise, let us introduce the following terminology. For a given dynamical system $T : X \to X$, a *revelation* for f, in the sense of Definition 5.2, will also be called a *maximizing revelation*, while a revelation for $-f$ will be called a *minimizing revelation* for f (i.e. a minimizing revelation ψ is a continuous coboundary such that $(f + \psi)^{-1}(\min(f + \psi))$ contains a compact T-invariant set). We say that ψ is a *bilateral revelation* for f if it is both a minimizing revelation and a maximizing revelation.

THEOREM 5.5. [**28**] (Bilateralizing the maximizing and minimizing revelations) *For $(X, T, f) \in \mathfrak{C}$, if there exists both a minimizing and a maximizing revelation, then there exists a bilateral revelation (i.e. a continuous coboundary $\varphi - \varphi \circ T$ with $(f + \varphi - \varphi \circ T)(X) = [\alpha(f), \beta(f)]$).*

For example, revisiting the family $(f_\theta)_{\theta\in\mathbb{R}}$ from Example 5.4(b), if $-3 \le \theta \le -2$ then ψ_c is seen to be a bilateral revelation for f_θ, for all $c \in [-\theta - 1, 2]$.

6. *Revelation theorems*

By a *revelation theorem*† we mean a result of the following kind.

THEOREM 6.1. (Revelation theorem: model version) *For a given (type of) dynamical system $(X, T) \in \mathfrak{D}$, and a given (type of) function $f \in C(X)$, there exists a revelation $\varphi - \varphi \circ T$ (i.e. $f + \varphi - \varphi \circ T$ is a revealed function).*

† Our terminology is consistent with that of §5, as again there is no established consensus on how to describe such theorems: the revelation theorem has been variously called the *normal form theorem*, the *positive Livsic theorem*, *Mañé's lemma*, the *Bousch–Mañé cohomology lemma* and the *Mañé–Conze–Guivarc'h lemma*.

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

In a typical revelation theorem, (X, T) is assumed to enjoy some hyperbolicity, and there is some restriction on the modulus of continuity of f. This is reminiscent of Livsic's Theorem (see [**110**], or e.g. [**100**, **139**]), which asserts that if (X, T) is suitably hyperbolic, and f is suitably regular (e.g. Hölder continuous) such that $\int f\, d\mu = 0$ for all $\mu \in \mathcal{M}_T$, then f is a continuous coboundary. Indeed a Livsic-type theorem can be viewed as a special case of a revelation theorem, as it follows by applying an appropriate revelation theorem to both f and $-f$, then invoking Theorem 5.5.

Revelation theorems date back to the 1990s: Conze and Guivarc'h proved a version as part of [**52**], there are parallels with the work of Mañé on Lagrangian flows [**118**, **119**], while the first published revelation theorem resembling Theorem 6.1 was due to Savchenko† [**145**], for (X, T) a subshift of finite type and f Hölder continuous. Other pioneering papers containing revelation theorems were those of Bousch [**26**] and Contreras, Lopes and Thieullen [**51**].

Common features of these early revelation theorems are that f is Hölder or Lipschitz, and that T is expanding. The following revelation theorem for expanding maps is a particular case of a result in [**30**] (which is valid for more general *amphidynamical systems*), and recovers those in [**26**, **51**, **52**, **145**].

THEOREM 6.2. (Revelation theorem: expanding T, Lipschitz f) *For expanding $(X, T) \in \mathfrak{D}$, every Lipschitz function $f : X \to \mathbb{R}$ has a Lipschitz revelation.*

Since an α-Hölder function for the metric d is a Lipschitz function for the metric d_α defined by $d_\alpha(x, y) = d(x, y)^\alpha$, we deduce the following.

COROLLARY 6.3. (Revelation theorem: expanding T, Hölder f) *For expanding $(X, T) \in \mathfrak{D}$, every α-Hölder function $f : X \to \mathbb{R}$ has an α-Hölder revelation, for all $\alpha \in (0, 1]$.*

To prove Theorem 6.2, we claim that the function φ defined by

$$\varphi(x) = \sup_{n\geq 1} \sup_{y \in T^{-n}(x)} (S_n f(y) - n\beta(f)) \tag{6}$$

is such that $\varphi - \varphi \circ T$ is a Lipschitz revelation. Without loss of generality we may assume that $\beta(f) = 0$, so that (6) becomes

$$\varphi(x) = \sup_{n\geq 1} \sup_{y \in T^{-n}(x)} S_n f(y).$$

To show that $\varphi - \varphi \circ T$ is a revelation, we first claim‡ that

$$f + \varphi - \varphi \circ T \leq 0, \tag{7}$$

and note this immediately implies that $(f + \varphi - \varphi \circ T)^{-1}(0)$ contains a compact T-invariant set, since otherwise there could not be any $m \in \mathcal{M}_T$ satisfying $\int f\, dm = 0 = \beta(f)$. To prove (7), note that

$$\varphi(Tx) = \sup_{n\geq 1} \sup_{y \in T^{-n}T(x)} S_n f(y) \geq \sup_{n\geq 1} \sup_{y \in T^{-(n-1)}(x)} S_n f(y), \tag{8}$$

† Savchenko's three-page paper contains no discussion of why the revelation theorem is interesting or useful, though does include some comments on its genesis: the problem had been proposed in Anosov and Stepin's Moscow dynamical systems seminar in November 1995, and had also been conjectured by Bill Parry. Savchenko's proof relies on thermodynamic formalism, exhibiting φ as a sub-sequential limit of $(1/t) \log h_t$, where h_t is the eigenfunction for the dominant eigenvalue of the Ruelle operator with potential function tf.
‡ Note that the proof of this claim does not require that f is Lipschitz or that T is expanding.

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

because $T^{-(n-1)}(x) \subset T^{-n}(T(x))$. Now if $y \in T^{-(n-1)}(x)$ then $S_n f(y) = f(x) + S_{n-1} f(y)$ for all $n \ge 1$ (with the usual convention that $S_0 f \equiv 0$), so (8) gives

$$\varphi(Tx) \ge f(x) + \sup_{n\ge 1} \sup_{y\in T^{-(n-1)}(x)} S_{n-1} f(y). \tag{9}$$

However,

$$\sup_{n\ge 1} \sup_{y\in T^{-(n-1)}(x)} S_{n-1} f(y) = \sup_{N\ge 0} \sup_{y\in T^{-N}(x)} S_N f(y) \ge \sup_{N\ge 1} \sup_{y\in T^{-N}(x)} S_N f(y) = \varphi(x), \tag{10}$$

so combining (9) and (10) gives $\varphi(Tx) \ge f(x) + \varphi(x)$, which is the desired inequality (7).

To complete the proof of Theorem 6.2 it remains to show that φ is Lipschitz, i.e. that there exists $K > 0$ such that for all $x, x' \in X$,

$$\varphi(x) - \varphi(x') \le K\, d(x, x'). \tag{11}$$

Given $x, x' \in X$, for any $\varepsilon > 0$ there exists $N \ge 1$ and $y \in T^{-N}(x)$ such that

$$\varphi(x) \le S_N f(y) + \varepsilon. \tag{12}$$

Since T is expanding we may write $y = T_{i_1} \circ \cdots \circ T_{i_N}(x)$ where the T_{i_j} denote inverse branches of T (i.e. each $T \circ T_{i_j}$ is the identity map), and we now define $y' \in X$ by $y' := T_{i_1} \circ \cdots \circ T_{i_N}(x')$. In particular, $y \in T^{-N}(x')$, so $S_N f(y') \le \sup_{z\in T^{-N}(x')} S_N f(z)$, and therefore

$$S_N f(y') \le \sup_{n\ge 1} \sup_{z\in T^{-n}(x')} S_n f(z) = \varphi(x'). \tag{13}$$

If $\lambda > 1$ is an expanding constant for T, i.e. $d(T(z), T(z')) \ge \lambda\, d(z, z')$ for all z, z' sufficiently close to each other, then $\gamma = \lambda^{-1}$ is a Lipschitz constant for each of the inverse branches of T, so that if $0 \le j \le N-1$ then $d(T_{i_{j+1}} \circ \cdots \circ T_{i_N}(x), T_{i_{j+1}} \circ \cdots \circ T_{i_N}(x')) \le \gamma^{N-j}\, d(x, x')$, therefore

$$f(T_{i_{j+1}} \circ \cdots \circ T_{i_N}(x)) - f(T_{i_{j+1}} \circ \cdots \circ T_{i_N}(x')) \le \mathrm{Lip}(f)\gamma^{N-j}\, d(x, x'),$$

and hence

$$S_N f(y) - S_N f(y') \le \sum_{j=0}^{N-1} \mathrm{Lip}(f)\gamma^{N-j}\, d(x, x') < \frac{\gamma}{1-\gamma}\mathrm{Lip}(f)\, d(x, x'). \tag{14}$$

Combining (12), (13) and (14) then gives $\varphi(x) - \varphi(x') < (\gamma/(1-\gamma))\,\mathrm{Lip}(f)\, d(x, x') + \varepsilon$, but $\varepsilon > 0$ was arbitrary, so

$$\varphi(x) - \varphi(x') \le \frac{\gamma}{1-\gamma}\, \mathrm{Lip}(f)\, d(x, x'),$$

which is the desired Lipschitz condition (11) with $K = (\gamma/(1-\gamma))\,\mathrm{Lip}(f)$, so Theorem 6.2 is proved.

In fact there are various different routes to proving Theorem 6.2, stemming from other possible choices of φ (see e.g. [**66**, **67**] for further details), notably the choice

$$\varphi(x) = \limsup_{n\to\infty} \sup_{y\in T^{-n}(x)} (S_n f(y) - n\beta(f)),$$

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

which moreover (see [**26**, **51**]) satisfies the functional equation

$$\varphi(x) + \beta(f) = \sup_{y \in T^{-1}(x)} (f + \varphi)(y). \tag{15}$$

Similar functional equations arise in a number of related settings, for example weak KAM (Kolmogorov–Arnold–Moser) theory [**59–62**] and infinite horizon optimal control theory [**44**, Theorem 5.2]. Indeed (15) can be interpreted as an eigenequation for the operator defined by its right-hand side, with the maximum ergodic average $\beta(f)$ playing the role of its eigenvalue; the nonlinear operator may be viewed as an analogue of the classical Ruelle transfer operator (see e.g. [**9**, **139**, **144**]) with respect to the max-plus algebra (in which the max operation plays the role of addition, and addition plays the role of multiplication, see e.g. [**12**]).

There is a revelation theorem for maps T which satisfy a condition that is weaker than being expanding: Bousch [**27**] defined $T : X \to X$ to be *weakly expanding* if its inverse T^{-1} is 1-Lipschitz when acting on the set of compact subsets of X, equipped with the induced Hausdorff metric (i.e. for all $x, y \in X$, there exists $x' \in T^{-1}(y)$ such that $d(x, x') \le d(Tx, Tx')$). The focus of [**27**] was on functions which are *Walters* (the notion was introduced in [**156**]) for the map T: for all $\epsilon > 0$ there exists $\delta > 0$ such that for all $n \in \mathbb{N}$, $x, y \in X$, if $d(T^i x, T^i y) < \delta$ for $0 \le i < n$ then $|S_n f(x) - S_n f(y)| < \epsilon$.

THEOREM 6.4. [**27**] (Revelation theorem: weakly expanding T, Walters f) *If $(X, T) \in \mathfrak{D}$ is weakly expanding, then every Walters function $f : X \to \mathbb{R}$ has a revelation.*

The first revelation theorem in the setting of *invertible* hyperbolic systems appeared in [**27**], for maps T satisfying an abstract notion of hyperbolicity dubbed *weak local product structure*: for all $\epsilon > 0$ there exists $\eta > 0$ such that if orbits $(x_i)_{i \le 0}$ and $(y_i)_{i \ge 0}$ satisfy $d(x_0, y_0) \le \eta$ then there exists an orbit $(z_i)_{i \in \mathbb{Z}}$ with $d(x_i, z_i) \le \epsilon$ for $i \ge 0$, and $d(y_i, z_i) \le \epsilon$ for $i \ge 0$.

THEOREM 6.5. [**27**] (Revelation theorem: T with weak local product structure) *If $(X, T) \in \mathfrak{D}$ is transitive and has weak local product structure, then every Walters function has a revelation.*

In particular, a transitive Anosov diffeomorphism has weak local product structure, and in this case a Hölder continuous function is Walters, so Theorem 6.5 implies the existence of a revelation. The following stronger result confirms, as suggested by Livsic's theorem, that in this case the revelation is also Hölder.

THEOREM 6.6. (Revelation theorem: T Anosov, f Hölder) *If $(X, T) \in \mathfrak{D}$ is a transitive Anosov diffeomorphism, and f is α-Hölder, then there exists a revelation $\varphi - \varphi \circ T$, where φ is α-Hölder.*

A version of Theorem 6.6 was proved by Lopes and Thieullen [**115**], who showed that if f is α-Hölder then φ is β-Hölder for some $\beta < \alpha$; the fact that φ can be chosen with the same Hölder exponent as f was established by Bousch [**30**].

Morris [**127**] considered the existence and non-existence of revelations in the context of circle maps with an indifferent fixed point (improving on earlier work [**36**, **149**]).

Specifically, he considered expanding circle maps of Manneville–Pomeau type $\alpha \in (0, 1)$, generalizing the Manneville–Pomeau map $x \mapsto x + x^{1+\alpha} \pmod 1$, and proved the following.

THEOREM 6.7. [**127**] (Revelation theorem: T of Manneville–Pomeau type) *If T is an expanding circle map of Manneville–Pomeau type $\alpha \in (0, 1)$, then every Hölder function of exponent $\gamma > \alpha$ has a $(\gamma - \alpha)$-Hölder revelation; however there exist α-Hölder functions which do not have a revelation.*

The estimate on the Hölder exponent of the revelation in Theorem 6.7 is sharp: there exist γ-Hölder functions without any revelation of Hölder exponent strictly larger than $\gamma - \alpha$ (see [**127**]). Branco [**35**] has considered certain degree-two circle maps with a super-attracting fixed point, proving that if f is α-Hölder, and the super-attracting fixed point is not maximizing, then there exists an α-Hölder revelation.

If $T : X \to X$ and $f : X \to \mathbb{R}$ are continuous, but X is not compact, there is no guarantee that f-maximizing measures exist: the supremum $\sup_{\mu \in \mathcal{M}_T} \int f\, d\mu$ need not be attained by any $m \in \mathcal{M}_T$. One way of proving the existence of f-maximizing measures is to establish a revelation theorem for f, an approach developed in [**94**, **95**], with particular focus on the case where (X, T) is a subshift of finite type on the countable alphabet $\mathbb{N}$. In this setting, if a function with summable variations is such that its values on a given cylinder set† are sufficiently larger than its values 'at infinity', then a revelation exists, and in particular $\mathcal{M}_{\max}(f, T)$ is non-empty. A prototypical result of this kind (see [**18**, **94**]) is the following.

THEOREM 6.8. (Revelation theorem: non-compact subshift of finite type) *For (X, T) the one-sided full shift on the alphabet $\mathbb{N}$, if f is bounded above, of summable variations, and there exists $I \in \mathbb{N}$ with $\sum_{j=1}^{\infty} \mathrm{var}_j(f) < \inf f|_{[I]} - \sup f|_{[i]}$, for all sufficiently large i, then f has a revelation, and in particular has a maximizing measure.*

The inequality in Theorem 6.8 clearly holds whenever $\sup f|_{[i]} \to -\infty$, which in particular is the case if f satisfies the summability condition $\sum_{i=1}^{\infty} e^{\sup f|_{[i]}} < \infty$ familiar from thermodynamic formalism (see e.g. [**121**]). Note that [**17**] provides alternative criteria guaranteeing the existence of a maximizing measure for certain functions f defined on irreducible countable alphabet subshifts of finite type: the approach is more direct than in [**94**], and while it does not prove the existence of a revelation, it does establish the subordination principle.

Going beyond the setting of discrete dynamics, there has been some work on revelation theorems for flows‡: for T^t a smooth Anosov flow without fixed points, and f Hölder continuous, there exists a Hölder function φ satisfying $\int_0^s f(T^t(x))\, dt + \varphi(x) - \varphi(T^s(x)) \le s\beta(f)$ for all $x \in X$, $s \in \mathbb{R}^+$ (see [**116**, **141**]), which moreover is smooth in the flow direction (see [**116**]). An analogous result holds for certain expansive non-Anosov geodesic flows, see [**114**].

† Here we use $[i]$ to denote the cylinder set consisting of all sequences $(x_n)_{n=1}^{\infty}$ such that $x_1 = i$.
‡ Note that although the majority of work on ergodic optimization has been placed in the setting of discrete time, there have been various developments in the context of flows (see [**28**, **114**, **116**, **117**, **141**, **158**]).

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

7. *Typically periodic optimization (TPO)*

Given a dynamical system $(X, T) \in \mathfrak{D}$ of a particular kind (e.g. enjoying some appropriate hyperbolicity), we wish to establish properties of *typical* maximizing measures: for a topological vector space V of real-valued functions on X, we aim to show there exists $V' \subset V$ which is topologically large (e.g. containing an open dense subset of V) such that all $f \in V'$ have maximizing measure(s) with a certain specified property. The specified property we have in mind is that the maximizing measure be *periodic*, though first we note that a weaker property follows as a simple consequence of §6 (where for definiteness (X, T) is assumed to be expanding or Anosov, and $V = \mathrm{Lip}$, so Theorems 6.2 and 6.6 can be used).

THEOREM 7.1. (Typical maximizing measures are not fully supported) *Suppose* $(X, T) \in \mathfrak{D}$ *is either expanding or Anosov, and is transitive but not reduced to a single periodic orbit. The open dense set* Lip', *defined as the complement in* Lip *of the closed subspace* $\{c + \varphi - \varphi \circ T : c \in \mathbb{R}, \varphi \in \mathrm{Lip}\}$, *is such that if* $f \in \mathrm{Lip}'$ *then no* f*-maximizing measure is fully supported.*

The possibility of typical maximizing measures being *periodic* was suggested by the early work on ergodic optimization for finite-dimensional spaces of functions, as described in §3. We state this below as the (purposefully imprecise) Conjecture 7.3, but first require some notation.

Definition 7.2. For $(X, T) \in \mathfrak{D}$, and V a Banach space consisting of certain continuous real-valued functions on X, define V_{Per} to be the set of those $f \in V$ such that $\mathcal{M}_{\max}(f)$ contains at least one measure supported on a single periodic orbit.

Conjecture 7.3. (TPO conjecture) If $(X, T) \in \mathfrak{D}$ is a suitably hyperbolic dynamical system, and V is a Banach space consisting of suitably regular continuous functions, then V_{Per} contains an open dense subset of V.

The earliest published paper containing specific articulations of the TPO conjecture was that of Yuan and Hunt [**159**], where (X, T) was assumed to be either an expanding map or an Axiom A diffeomorphism; the analogue of the TPO conjecture was conjectured [**159**, Conjecture 1.1] for V a space of smooth (e.g. C^1) functions on X, though the case $V = \mathrm{Lip}$ was discussed in more detail. In subsequent years, this case $V = \mathrm{Lip}$ became a focus of attention among workers in ergodic optimization, culminating in its resolution (see Theorem 7.10 below) by Contreras [**50**], building on the work of [**126**, **142**, **159**].

The first proved (infinite-dimensional) version of the TPO conjecture was due to Contreras, Lopes and Thieullen [**51**], in a paper prepared at around the same time as [**159**]. In the context of (smooth) expanding maps (on the circle), they noted a significant consequence of the revelation theorem (a version of which they proved [**51**, Theorem 9]), which would also be exploited by subsequent authors: *if* it is known that (X, T, f) has a revelation ψ for every $f \in V$, then the revealed function $f + \psi$ is usually more amenable to analysis, and in particular it may be possible to exhibit small perturbations of $f + \psi$ which lie in the desired set V_{Per}, and thus deduce that f itself can be approximated by members of V_{Per}. Their choice of $V = V^\alpha$ was as a closed subspace of H^α, the space of

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

α-Hölder functions on X; the space V^α is defined to consist of the closure (in H^α) of those functions which are actually *better* than α-Hölder, i.e. they are β-Hölder for some $\beta > \alpha$ (so V^α is defined for $\alpha < 1$, but undefined in the Lipschitz class $\alpha = 1$). The superior approximation properties enjoyed by V^α yield the following.

THEOREM 7.4. [**51**] (TPO on a proper closed subspace of Hölder functions) *For* $\alpha \in (0, 1)$ *and* (X, T) *a circle expanding map,* V^α_{Per} *contains an open dense subset of* V^α.

Bousch [**27**, p. 305] was able to use his revelation theorem (Theorem 6.4) for the set W of Walters functions in the particular context of the one-sided full shift, where W can be given the structure of a Banach space, to prove the following.

THEOREM 7.5. [**27**] (TPO for Walters functions on a full shift) *For* (X, T) *a full shift,* W_{Per} *contains an open dense subset of* W.

An important ingredient in the proof of Theorem 7.5 is that locally constant functions are dense in W, and that for such functions f the set $\mathcal{M}_{\max}(f)$ is stable under perturbation (indeed $\mathcal{M}_{\max}(f)$ is the set of all invariant measures supported by some subshift of finite type, and such subshifts always contain at least one periodic orbit).

Quas and Siefken [**142**] also considered the setting of (X, T) a one-sided full shift, and spaces of functions which are Lipschitz with respect to non-standard metrics on X: for a sequence $A = (A_n)_{n=1}^\infty$ with $A_n \searrow 0$, define a metric d_A on X by $d_A(x, y) = A_n$ if x and y first differ in the nth position (i.e. $x_i = y_i$ for $1 \le i < n$, and $x_n \ne y_n$), and let $\mathrm{Lip}(A)$ denote the space of functions on X which are Lipschitz with respect to d_A, equipped with the induced Lipschitz norm. Quas and Siefken required the additional condition $\lim_{n\to 0} A_{n+1}/A_n = 0$ (in which case members of $\mathrm{Lip}(A)$ are referred to as *super-continuous* functions) and proved the following.

THEOREM 7.6. [**142**] (TPO for super-continuous functions) *For* (X, T) *a full shift, if* $\lim_{n\to 0} A_{n+1}/A_n = 0$ *then* $\mathrm{Lip}(A)_{\mathrm{Per}}$ *contains an open dense subset of* $\mathrm{Lip}(A)$.

In the same context of super-continuous functions on a one-sided full shift (X, T), Bochi and Zhang [**24**] found a more restrictive condition on the sequence A which suffices to guarantee that $\mathrm{Lip}(A)_{\mathrm{Per}}$ is a *prevalent*† subset of $\mathrm{Lip}(A)$.

THEOREM 7.7. [**24**] (Prevalent periodic optimization) *For* (X, T) *the one-sided full shift on two symbols, if* $A_{n+1}/A_n = O(2^{-2^{n+2}})$ *as* $n \to \infty$ *then* $\mathrm{Lip}(A)_{\mathrm{Per}}$ *is a prevalent subset of* $\mathrm{Lip}(A)$.

The proof in [**24**] uses Haar wavelets to reduce the problem to a finite-dimensional one with a graph-theoretic reformulation as a maximum cycle mean problem. Since the hard part of proving Theorem 7.6 is to show that $\mathrm{Lip}(A)_{\mathrm{Per}}$ contains a dense subset of $\mathrm{Lip}(A)$, and any prevalent subset is dense, we note that Theorem 7.7 constitutes a strengthening of Theorem 7.6 in the case that $A_{n+1}/A_n = O(2^{-2^{n+2}})$ as $n \to \infty$.

† Prevalence is a probabilistic notion of typicalness, introduced by Hunt, Sauer and Yorke [**77**], and in finite-dimensional spaces coincides with the property of being of full Lebesgue measure. Specifically, for V a complete metrizable topological vector space, a Borel set $S \subset V$ is called *shy* if there exists a compactly supported measure which gives mass zero to every translate of S, and a *prevalent* set in V is defined to be one whose complement is shy.

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

Prior to Contreras' proof of Theorem 7.10 below, a number of authors (notably [**29**, **38**, **159**]) had considered the case $V = \mathrm{Lip}$ in the TPO conjecture, and established partial and complementary results. The first of these was due to Yuan and Hunt [**159**].

THEOREM 7.8. [**159**] (Non-periodic measures are not robustly optimizing) *Let* $(X, T) \in \mathfrak{D}$ *be an expanding map. If* $f \in \mathrm{Lip}$ *has a non-periodic maximizing measure* μ*, then there exists* $g \in \mathrm{Lip}$*, arbitrarily close to* f *in the Lipschitz topology, such that* μ *is not* g*-maximizing.*

Bousch [**29**] gave an alternative proof of Theorem 7.8, in the more general setting of amphidynamical systems, making explicit the role of revelations, and quantifying the phenomenon of periodic orbits of low period being more stably maximizing than those of high period: if $f \in \mathrm{Lip}$ has a periodic maximizing measure μ of (large) period N, then there exist $O(1/N)$-perturbations of f in the Lipschitz norm for which μ is no longer maximizing. More precisely, the following proposition is given.

PROPOSITION 7.9. [**29**] (A bound on orbit-locking for Lipschitz functions) *If* $(X, T) \in \mathfrak{D}$ *is expanding, then there exists* $K_T > 0$ *such that if* $f \in \mathrm{Lip}$ *has* μ *as an* f*-maximizing measure, where* μ *is supported on a periodic orbit of period* $N > K_T$*, then there exists* $g \in \mathrm{Lip}$ *with* $\mathrm{Lip}(f - g) \le (N/K_T - 1)^{-1}$ *such that* μ *is not* g*-maximizing.*

The constant K_T in Proposition 7.9 can be chosen as $K_T = 6C_T L_T$, where L_T is a Lipschitz constant for T, and C_T is such that $\mathrm{Lip}(\varphi) \le C_T \,\mathrm{Lip}(f)$ whenever $\varphi - \varphi \circ T$ is a revelation for a Lipschitz function f, so for example in the particular case of $T(x) = 2x$ (mod 1) on the circle, we may take $L_T = 2$ and $C_T = 1$ (see [**26**, Lemme B]), so for any $f \in \mathrm{Lip}$ whose maximizing measure μ is supported on an orbit of period $N > 12$, there exists $g \in \mathrm{Lip}$ with $\mathrm{Lip}(f - g) \le 12/(N - 12)$ such that μ is not g-maximizing.

A proof of the TPO conjecture in the important case $V = \mathrm{Lip}$ was given by Contreras as the following.

THEOREM 7.10. [**50**] (TPO for Lipschitz functions) *For* $(X, T) \in \mathfrak{D}$ *an expanding map,* $\mathrm{Lip}_{\mathrm{Per}}$ *contains an open dense subset of* Lip.

To sketch† a proof of Theorem 7.10, we first note that if μ is any periodic orbit measure, it is relatively easily shown that $\{f \in \mathrm{Lip} : \mu \text{ is } f\text{-maximizing}\}$ is a closed set with non-empty interior, so it suffices to show that $\mathrm{Lip}_{\mathrm{Per}}$ is dense in Lip. Let us say that $\mu \in \mathcal{M}_T$ is a *Yuan–Hunt measure* if for all $x \in \mathrm{supp}(\mu)$, $Q > 0$, there exist integers $m, p \ge 0$ such that‡

$$\min\{d(T^i x, T^j x) : m \le i,\ j \le m + p,\ 0 < |i - j| < p\} > Q d(T^{m+p} x, T^m x). \quad (16)$$

If $\mathcal{M}_{\mathrm{YH}}$ denotes the set of Yuan–Hunt measures, and

$$\mathrm{Lip}_{\mathrm{YH}} = \{f \in \mathrm{Lip} : \mathcal{M}_{\max}(f) \cap \mathcal{M}_{\mathrm{YH}} \ne \emptyset\},$$

† This sketch follows the exposition of Bousch [**31**].
‡ Condition (16) was introduced by Yuan and Hunt [**159**, p. 1217], who called it the Class I condition. (A Class II condition was also introduced in [**159**], in terms of approximability by periodic orbit measures, which stimulated related work in [**42**, **49**].)

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

then clearly every invariant measure supported on a periodic orbit lies in $\mathcal{M}_{\mathrm{YH}}$, and thus $\mathrm{Lip}_{\mathrm{Per}} \subset \mathrm{Lip}_{\mathrm{YH}}$. Yuan and Hunt proved [**159**, Lemma 4.10] that $\mathrm{Lip}_{\mathrm{Per}}$ is dense in $\mathrm{Lip}_{\mathrm{YH}}$, so to prove Theorem 7.10 it suffices to show that $\mathrm{Lip}_{\mathrm{YH}}$ is dense in Lip. Contreras [**50**] showed, by estimating the lengths of pseudo-orbits, that if $\mu \in \mathcal{M}_T \setminus \mathcal{M}_{\mathrm{YH}}$ then μ has strictly positive entropy. However, a result of Morris [**126**] asserts that the set of Lipschitz functions with a positive entropy maximizing measure is of the first category; it follows that $\mathrm{Lip} \setminus \mathrm{Lip}_{\mathrm{YH}}$ is of the first category, and therefore $\mathrm{Lip}_{\mathrm{YH}}$ is dense in Lip, as required.

8. *Other typical properties of maximizing measures*

For a suitably hyperbolic dynamical system (X, T), the fact that typical properties of maximizing measures in $C(X)$ are rather different from those of more regular continuous functions discussed in §7 is illustrated by the following result.

THEOREM 8.1. [**128**] (Typical maximizing measures for continuous functions) *For $(X, T) \in \mathfrak{D}$ either expanding or Anosov, and transitive but not reduced to a single fixed point, there is a residual subset $C' \in C(X)$ such that if $f \in C'$ then $\mathcal{M}_{\max}(f)$ is a singleton containing a measure which is fully supported, has zero entropy, and is not strongly mixing.*

Parts of Theorem 8.1 had been proved elsewhere (see [**32**] for the fact that typical maximizing measures are fully supported, and [**40**] for the fact that typical maximizing measures have zero entropy), while Morris's proof in [**128**] was a natural consequence of his following more abstract result (together with results of Sigmund [**148**] on residual subsets of the set of invariant measures).

THEOREM 8.2. [**128**] (Maximizing measures inherit typical properties from $\mathcal{M}_T$) *Suppose $(X, T) \in \mathfrak{D}$ is such that the set of ergodic measures is weak* dense in $\mathcal{M}_T$. Then for typical continuous functions f, the f-maximizing measure inherits any property which is typical in $\mathcal{M}_T$. More precisely, if $\mathcal{M}'$ is a residual subset of $\mathcal{M}_T$, then $\{f \in C(X) : \mathcal{M}_{\max}(f) \subset \mathcal{M}'\}$ is a residual subset of $C(X)$.*

The most surprising aspect of Theorem 8.1 is that in $C(X)$ a typical maximizing measure is *fully supported*. Not only does this contrast with Theorem 7.1 and the typical *periodic* optimization results of §7, but it also contrasts with intuition; indeed an open problem is to exhibit constructively a continuous function $f : X \to \mathbb{R}$, and an expanding or Anosov dynamical system $(X, T) \in \mathfrak{D}$, such that the unique f-maximizing measure is fully supported. Clearly such a unique maximizing measure must be ergodic (since the set $\mathcal{M}_{\max}(f)$ is convex, and ergodic maximizing measures are precisely its extreme points), though it turns out that this is the only restriction, as described by the following theorem.

THEOREM 8.3. [**88**] (Every ergodic measure is uniquely maximizing) *If $(X, T) \in \mathfrak{D}$ then for any ergodic $\mu \in \mathcal{M}$, there exists $f \in C(X)$ such that μ is the unique f-maximizing measure.*

The above results, and those of §7, involve fixing the dynamical system $(X, T) \in \mathfrak{D}$, and enquiring about typical properties of f-maximizing measures for f lying in some Banach space V. More generally, one might view the triple $(X, T, f) \in \mathfrak{C}$ as varying in some given topological space, and again enquire about typical properties of $\mathcal{M}_{\max}(X, T, f)$;

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

alternatively we may fix the (compact metric) state space X, and view pairs (T, f) as elements of some topological space $\mathcal{P}$, and again enquire about typical properties of $\mathcal{M}_{\max}(T, f)$.

The existing literature has focused on three versions of this general problem, the first (and most studied) of which is the case $\mathcal{P} = \{T\} \times V$ described previously. A second case is when T varies within some topological space $\mathcal{T}$ of maps, and the function $f = f_T$ varies non-trivially with T. In fact the limited work on this second case has focused (see [**51**, **96**, **123**]) on *Lyapunov maximizing measures* for expanding maps T, i.e. where $f_T = \log |T'|$, so that $\max_{\mu \in \mathcal{M}_T} \int \log |T'| \, d\mu$ is the *maximum Lyapunov exponent*. If $\mathcal{T}$ is the space of C^1 expanding maps then a typical Lyapunov maximizing measure is shown (see [**96**] for X the circle, and [**123**] for X a more general manifold) to be unique, fully supported and of zero entropy (this can be considered an analogue of Theorem 8.1); by contrast, in the context of $C^{1+\alpha}$ expanding maps an analogue of Theorem 7.4 is established (see [**51**]), and indeed it is likely that other results in §7 have natural analogues in the context of Lyapunov maximizing measures.

The third version of the general problem involves fixing the continuous function $f : X \to \mathbb{R}$ (with possibly additional hypotheses on f) and varying the map T within some topological space $\mathcal{T}$. For example, if X is a compact connected manifold of dimension greater than 2, T is varied in the space $\mathcal{T}$ of homeomorphisms of X and $f \in C(X)$ is considered fixed, it can be shown (see [**153**]) that there is a *dense* subset $\mathcal{T}' \subset \mathcal{T}$ such that $\mathcal{M}_{\max}(X, T, f)$ contains a periodic orbit measure. However, provided f is non-constant when restricted to any non-empty open subset, it turns out that periodic maximization is *not* typical.

THEOREM 8.4. [**1**] (Typical optimization is not periodic) *Let X be a compact connected Riemannian manifold of dimension at least 2. If $f \in C(X)$ is non-constant when restricted to any non-empty open subset, then there is a residual subset $\mathcal{T}'_f \subset \mathrm{Homeo}(X)$ such that for every $T \in \mathcal{T}'_f$, the set $\mathcal{M}_{\max}(X, T, f)$ contains no periodic orbit measures.*

Analogous results have been established for the space $\mathcal{T}$ of endomorphisms (i.e. continuous surjections): for example in [**15**] (see also [**154**]) it is shown that for any compact Riemannian manifold X, and any $f \in C(X)$, there is a dense subset $\mathcal{T}' \subset \mathcal{T}$ such that $\mathcal{M}_{\max}(X, T, f)$ contains a periodic orbit measure for all $T \in \mathcal{T}'$; however if X is the circle it is known (see [**1**]) that such a $\mathcal{T}'$ is meagre unless the function f is monotone on some sub-interval (in which case $\mathcal{T}'$ has non-empty interior).

9. *Sturmian optimization and ergodic dominance*

In ergodic optimization, Sturmian measures were first observed in the context of the model problem described in §3, as maximizing measures for functions of the form $v_\theta(x) = \cos 2\pi(x - \theta)$, with underlying dynamical system $T(x) = 2x \pmod 1$ on the circle $\mathbb{R}/\mathbb{Z}$. As well as their definition in terms of rotations (see §3), Sturmian measures can be characterized as precisely those T-invariant probability measures whose support is contained in a sub-interval of the form $[\gamma, \gamma + 1/2]$ (i.e. a closed semi-circle), see e.g. [**33**, **43**]. In other words, the family of Sturmian measures can be defined as the maximizing measures for the family of characteristic functions $\chi_{[\gamma, \gamma+1/2]}$, $\gamma \in \mathbb{R}/\mathbb{Z}$. In view of this

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

definition, it is perhaps not so surprising that Sturmian measures arise as maximizing measures for certain naturally occurring functions, and indeed they have subsequently been identified as maximizing measures for functions other than the family v_θ. One such example (see [**4**]) is the family of functions $u_\theta(x) = -d(x, \theta)$, $\theta \in \mathbb{R}/\mathbb{Z}$, where d is the usual distance function on the circle. As for the family of functions v_θ, the Sturmian measures are *precisely* the maximizing measures for the functions u_θ, with each non-periodic Sturmian measure being maximizing for a single function u_θ, but periodic Sturmian measures corresponding to a positive length closed interval of parameters θ.

Moving beyond finite-dimensional families of functions, there exist infinite-dimensional function cones where Sturmian measures are guaranteed to be maximizing; by a function cone we mean a set K of functions on X which is closed under addition (i.e. $K + K \subset K$) and multiplication by non-negative reals (i.e. $\mathbb{R}_{\geq 0}K \subset K$). If $X = [0, 1]$ then the set of concave real-valued functions on X is a cone, and if T is the doubling map† on $[0, 1]$, with $T(1) = 1$ and $T(x) = 2x \pmod 1$ for $x < 1$, then the following theorem holds.

THEOREM 9.1. [**89**, **90**] (Sturmian maximizing measures for concave functions) *For the doubling map $T : [0, 1] \to [0, 1]$, if $f : [0, 1] \to \mathbb{R}$ is concave then it has a Sturmian maximizing measure. If f is strictly concave then its maximizing measure is unique and Sturmian.*

The set of increasing functions on $[0, 1]$ is also a function cone. For the doubling map on $[0, 1]$, the Dirac measure δ_1 is clearly f-maximizing for every increasing function $f : [0, 1] \to \mathbb{R}$. This simple fact has a more surprising generalization: if‡ $\beta \in (1, 2)$ and $T_\beta : [0, 1] \to [0, 1]$ is given by $T_\beta(x) = \beta x$ on $[0, 1/\beta]$ and $T_\beta(x) = \beta x - 1$ on $(1/\beta, 1]$, then for certain β (e.g. the golden mean $\beta = (1 + \sqrt{5})/2$) there exists a single T_β-invariant probability measure μ_β which is simultaneously maximizing for all increasing functions on $[0, 1]$, and in this case μ_β is once again Sturmian§.

THEOREM 9.2. [**5**] (Sturmian maximizing measure for all increasing functions) *For the map $T_\beta : [0, 1] \to [0, 1]$, if $\beta \in (1, 2)$ is the dominant root of $x^{ap+1} - \sum_{i=0}^{p} x^{ia}$ for some integers $a, p \geq 1$, then the point 1 generates a Sturmian periodic orbit, and the T_β-invariant measure on this orbit is f-maximizing for every increasing function $f : [0, 1] \to \mathbb{R}$.*

For any cone K such that $K - K$ is dense in $C(X)$, a partial order $\prec$ on Borel probability measures arises by declaring that $\mu \prec \nu$ if and only if $\int f\, d\mu \leq \int f\, d\nu$ for all $f \in K$. Both the cone of increasing functions and the cone of concave functions enjoy this property, and for these cases the associated partial order is known as a *stochastic dominance* order (see e.g. [**16**, **103**, **120**, **147**]). We therefore use the term *ergodic dominance* to refer to the study of the partially ordered set $(\mathcal{M}_T, \prec)$, and the identification

† Although the doubling map on $[0, 1]$ is not continuous, its set of invariant probability measures is nevertheless weak* compact, so $\mathcal{M}_{\max}(f) \neq \emptyset$ for all continuous f.
‡ For $\beta > 2$ there is a slightly different version of Theorem 9.2 for analogous maps T_β (see [**5**] for details).
§ Sturmian measures can, as before, be defined in terms of circle rotations; alternatively, in this context they are characterized by having support contained in a closed interval of length $1/\beta$.

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

of maximal and minimal elements in $(\mathcal{M}_T, K)$ may be viewed as a generalization of ergodic optimization. For K the cone of increasing functions, ergodic dominance in the context of the full shift on two symbols has been investigated in [**2**, **3**].

For K the cone of concave functions, ergodic dominance has been studied in [**98**] for orientation-reversing expanding maps $T : [0, 1] \to [0, 1]$, and in [**150**] for certain unimodal maps. A necessary condition for the comparability of two measures μ, ν is that their barycentres coincide, i.e. $\int x \, d\mu(x) = \int x \, d\nu(x)$, so if T is the doubling map then $(\mathcal{M}_T, \prec)$ cannot have a maximum element, though each of the sets $\mathcal{M}_{T,\varrho} = \{\mu \in \mathcal{M}_T : \int x \, d\mu(x) = \varrho\}$ does turn out to have such an element.

THEOREM 9.3. [**89**, **90**] (Sturmian measures as maximum elements in each $\mathcal{M}_{T,\varrho}$) *If* $T : [0, 1] \to [0, 1]$ *is the doubling map, and* $\prec$ *is the partial order induced by the cone of concave functions on* $[0, 1]$*, then the Sturmian measure of rotation number* ϱ *is the maximum element in* $\mathcal{M}_{T,\varrho}$*, for all* $\varrho \in [0, 1]$.

In fact Theorem 9.1 can be viewed as one of several corollaries to Theorem 9.3; others are that Sturmian measures have strictly smallest variance around their means, and that Sturmian periodic orbits have larger geometric mean than any other periodic orbits with the same arithmetic mean (see [**89**, **90**, **92**] for further details).

Underlying the results in this section is an idea of Bousch [**26**], which provides an approach to proving the Sturmian nature of maximizing measures: the existence of a revelation is guaranteed by Theorem 6.2, and it is potentially feasible to show that the corresponding revealed function takes its maximum value on a closed interval of length $1/\beta$, in which case the maximizing measure is Sturmian. This approach can also be used for more general expanding maps T (see e.g. [**86**, **91**, **97**]) where the closed interval in question has the property that T is injective when restricted to its interior, and for certain generalizations of Sturmian measures (see e.g. [**38**, **39**, **72**]).

The article [**97**] treats a problem concerning the joint spectral radius of pairs of matrices (see e.g. [**55**, **99**, **106**, **143**] for background to this area), which is reformulated as an ergodic optimization problem involving a one-parameter family of expanding maps, and a one-parameter family of functions, whose maximizing measures turn out to be precisely the family of Sturmian measures. The role of Sturmian measures (or orbits) in this context had previously been noted in [**21**, **33**, **71**, **133**]. More generally, joint spectral radius problems have a number of parallels with ergodic optimization, and the two fields enjoy a fruitful interaction, see e.g. [**22**, **23**, **33**, **129–132**].

Acknowledgement. This research was supported by EPSRC grant number EP/L02246X/1.

REFERENCES

[**1**] S. Addas-Zanata and F. A. Tal. Support of maximizing measures for typical C^0 dynamics on compact manifolds. *Discrete Contin. Dyn. Syst.* **26** (2010), 795–804.

[**2**] V. Anagnostopoulou. Sturmian measures and stochastic dominance in ergodic optimization. *PhD Thesis*, Queen Mary University of London, 2008.

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

[3] V. Anagnostopoulou. Stochastic dominance for shift-invariant measures. *Discrete Contin. Dyn. Syst.*, to appear.

[4] V. Anagnostopoulou, K. Díaz-Ordaz Avila, O. Jenkinson and C. Richard. Sturmian maximizing measures for the piecewise-linear cosine family. *Bull. Braz. Math. Soc. (N.S.)* **43** (2012), 285–302.

[5] V. Anagnostopoulou and O. Jenkinson. Which beta-shifts have a largest invariant measure? *J. Lond. Math. Soc.* **72** (2009), 445–464.

[6] N. Anantharaman. On the zero-temperature or vanishing viscosity limit for Markov processes arising from Lagrangian dynamics. *J. Eur. Math. Soc. (JEMS)* **6** (2004), 207–276.

[7] N. Aubrun and M. Sablik. Simulation of effective subshifts by two-dimensional subshifts of finite type. *Acta Appl. Math.* **126** (2013), 35–63.

[8] P. Avramidou. Optimization of ergodic averages along squares. *Dyn. Syst.* **25** (2010), 547–553.

[9] V. Baladi. *Positive Transfer Operators and Decay of Correlations (Advanced Series in Nonlinear Dynamics, 16)*. World Scientific, Singapore, 2000.

[10] A. Baraviera, L. Cioletti, A. O. Lopes, J. Mohr and R. Souza. On the general one-dimensional XY model: positive and zero temperature, selection and non-selection. *Rev. Math. Phys.* **23** (2011), 1063–1113.

[11] A. T. Baraviera, R. Leplaideur and A. O. Lopes. Selection of ground states in the zero temperature limit for a one-parameter family of potentials. *SIAM J. Appl. Dyn. Syst.* **11** (2012), 243–260.

[12] A. T. Baraviera, R. Leplaideur and A. O. Lopes. *Ergodic Optimization, Zero Temperature Limits and the Max-Plus Algebra, Publicações Matemáticas do IMPA (29^o Coloquio Brasileiro de Matemática)*. IMPA, Rio de Janeiro, 2013.

[13] A. T. Baraviera, A. O. Lopes and J. Mengue. On the selection of sub-action and measure for a subclass of potentials defined by P. Walters. *Ergod. Th. & Dynam. Sys.* **33** (2013), 1338–1362.

[14] A. T. Baraviera, A. O. Lopes and Ph. Thieullen. A large deviation principle for equilibrium states of Hölder potentials: the zero temperature case. *Stoch. Dyn.* **6** (2006), 77–96.

[15] T. Batista, J. Gonschorowski and F. Tal. Density of the set of endomorphisms with a maximizing measure supported on a periodic orbit. *Discrete Contin. Dyn.* **35** (2015), 3315–3326.

[16] V. S. Bawa. Stochastic dominance: a research bibliography. *Manag. Sci.* **28** (1982), 698–712.

[17] R. Bissacot and R. Freire Jr. On the existence of maximizing measures for irreducible countable Markov shifts: a dynamical proof. *Ergod. Th. & Dynam. Sys.* **34** (2014), 1103–1115.

[18] R. Bissacot and E. Garibaldi. Weak KAM methods and ergodic optimal problems for countable Markov shifts. *Bull. Braz. Math. Soc. (N.S.)* **41** (2010), 321–338.

[19] R. Bissacot, E. Garibaldi and Ph. Thieullen. Zero-temperature phase diagram for double-well type potentials in the summable variation class. *Ergod. Th. & Dynam. Sys.* doi:10.1017/etds.2016.57. Published online: 19 September 2016.

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

[21] V. D. Blondel, J. Theys and A. A. Vladimirov. An elementary counterexample to the finiteness conjecture. *SIAM J. Matrix Anal. Appl.* **24** (2003), 963–970.

[22] J. Bochi and I. D. Morris. Continuity properties of the lower spectral radius. *Proc. Lond. Math. Soc.* **110** (2015), 477–509.

[23] J. Bochi and M. Rams. The entropy of Lyapunov-optimizing measures of some matrix cocycles. *J. Mod. Dyn.* **10** (2016), 255–286.

[24] J. Bochi and Y. Zhang. Ergodic optimization of prevalent super-continuous functions. *Int. Math. Res. Not. IMRN* **19** (2016), 5988–6017.

[25] T. Bohr and D. Rand. The entropy function for characteristic exponents. *Physica D* **25** (1986), 387–398.

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

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

[28] T. Bousch. Un lemme de Mañé bilatéral. *C. R. Acad. Sci. Paris Sér. I* **335** (2002), 533–536.

[29] T. Bousch. Nouvelle preuve d'un théorème de Yuan et Hunt. *Bull. Soc. Math. France* **126** (2008), 227–242.

[30] T. Bousch. Le lemme de Mañé-Conze-Guivarc'h pour les systèmes amphi-dynamiques rectifiables. *Ann. Fac. Sci. Toulouse Math.* **20** (2011), 1–14.

[31] T. Bousch. Genericity of minimizing periodic orbits, after Contreras. *British Math. Colloq. talk*. (April 2014), QMUL.

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

[33] T. Bousch and J. Mairesse. Asymptotic height optimization for topical IFS, Tetris heaps, and the finiteness conjecture. *J. Amer. Math. Soc.* **15** (2002), 77–111.

[34] R. Bowen. Topological entropy for noncompact sets. *Trans. Amer. Math. Soc.* **184** (1973), 125–136.

- **[35]** F. Branco. Sub-actions and maximizing measures for one-dimensional transformations with a critical point. *Discrete Cont. Dyn. Syst.* **17** (2007), 271–280.
- **[36]** S. Branton. Sub-actions for Young towers. *Discrete Cont. Dyn. Syst.* **22** (2008), 541–556.
- **[37]** J. Brémont. On the behaviour of Gibbs measures at temperature zero. *Nonlinearity* **16** (2003), 419–426.
- **[38]** J. Brémont. Finite flowers and maximizing measures for generic Lipschitz functions on the circle. *Nonlinearity* **19** (2006), 813–828.
- **[39]** J. Brémont. Dynamics of injective quasi-contractions. *Ergod. Th. & Dynam. Sys.* **26** (2006), 19–44.
- **[40]** J. Brémont. Entropy and maximizing measures of generic continuous functions. *C. R. Math. Acad. Sci. Sér. I* **346** (2008), 199–201.
- **[41]** J. Brémont and Z. Buczolich. Maximizing points and coboundaries for an irrational rotation on a circle. *Ergod. Th. & Dynam. Sys.* **33** (2013), 24–48.
- **[42]** X. Bressaud and A. Quas. Rate of approximation of minimizing measures. *Nonlinearity* **20** (2007), 845–853.
- **[43]** 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.
- **[44]** D. A. Carlson, A. B. Haurie and A. Leizarowitz. *Infinite Horizon Optimal Control: Deterministic and Stochastic Systems*, 2nd edn. Springer, 1991.
- **[45]** J.-R. Chazottes, J.M. Gambaudo and E. Ugalde. Zero-temperature limit of one dimensional Gibbs states via renormalization: the case of locally constant potentials. *Ergod. Th. & Dynam. Sys.* **31** (2011), 1109–1161.
- **[46]** J.-R. Chazottes and M. Hochman. On the zero-temperature limit of Gibbs states. *Comm. Math. Phys.* **297** (2010), 265–281.
- **[47]** Y. Chen and Y. Zhao. Ergodic optimization for a sequence of continuous functions. *Chinese J. Contemp. Math.* **34** (2013), 351–360.
- **[48]** Z. N. Coelho. Entropy and ergodicity of skew-products over subshifts of finite type and central limit asymptotics. *PhD Thesis*, Warwick University, 1990.
- **[49]** D. Collier and I. D. Morris. Approximating the maximum ergodic average via periodic orbits. *Ergod. Th. & Dynam. Sys.* **28** (2008), 1081–1090.
- **[50]** G. Contreras. Ground states are generically a periodic orbit. *Invent. Math.* **205** (2016), 383–412.
- **[51]** G. Contreras, A. O. Lopes and Ph. Thieullen. Lyapunov minimizing measures for expanding maps of the circle. *Ergod. Th. & Dynam. Sys.* **21** (2001), 1379–1409.
- **[52]** J.-P. Conze and Y. Guivarc'h. Croissance des sommes ergodiques, *Manuscript*, circa 1993.
- **[53]** D. Coronel and J. Rivera-Letelier. Sensitive dependence of Gibbs measures at low temperatures. *J. Stat. Phys.* **160** (2015), 1658–1683.
- **[54]** D. Coronel and J. Rivera-Letelier. Sensitive dependence of geometric Gibbs states. *Preprint*, 2017, arXiv:1708.03965.
- **[55]** I. Daubechies and J. C. Lagarias. Sets of matrices all infinite products of which converge. *Linear Algebra Appl.* **162** (1992), 227–261.
- **[56]** A. Davie, M. Urbański and A. Zdunik. Maximizing measures of metrizable non-compact spaces. *Proc. Edinb. Math. Soc.* **50** (2007), 123–151.
- **[57]** B. Durand, A. Romashchenko and A. Shen. Fixed-point tile sets and their applications. *J. Comput. System Sci.* **78** (2012), 731–764.
- **[58]** A. C. D. van Enter and W. M. Ruszel. Chaotic temperature dependence at zero temperature. *J. Stat. Phys.* **127** (2007), 567–573.
- **[59]** A. Fathi. Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens. *C. R. Acad. Sci. Paris Sér. I* **324**(9) (1997), 1043–1046.
- **[60]** A. Fathi. Solutions KAM faibles conjuguées et barrières de Peierls. *C. R. Acad. Sci. Paris Sér. I* **325**(6) (1997), 649–652.
- **[61]** A. Fathi. Orbites hétéroclines et ensemble de Peierls. *C. R. Acad. Sci. Paris Sér. I* **326**(10) (1998), 1213–1216.
- **[62]** A. Fathi. Sur la convergence du semi-groupe de Lax-Oleinik. *C. R. Acad. Sci. Paris Sér. I* **327**(3) (1998), 267–270.
- **[63]** R. Freire and V. Vargas. Equilibrium states and zero temperature limit on topologically transitive countable Markov shifts. *Trans. Amer. Math. Soc. Preprint*, 2015, arXiv:1511.01527, to appear, doi:10.1090/tran/7291.
- **[64]** E. Garibaldi and J. T. A. Gomes. Aubry set for asymptotically sub-additive potentials. *Stoch. Dyn.* **16** (2016), 1660009.
- **[65]** E. Garibaldi and A. O. Lopes. Functions for relative maximization. *Dyn. Syst.* **22** (2007), 511–528.

[**66**] E. Garibaldi and A. O. Lopes. On Aubry–Mather theory for symbolic dynamics. *Ergod. Th. & Dynam. Sys.* **28** (2008), 791–815.

[**67**] E. Garibaldi, A. O. Lopes and Ph. Thieullen. On calibrated and separating sub-actions. *Bull. Braz. Math. Soc. (N.S.)* **40** (2009), 577–602.

[**68**] E. Garibaldi and Ph. Thieullen. Description of some ground states by Puiseux techniques. *J. Stat. Phys.* **146** (2012), 125–180.

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

[**70**] T. C. Halsey, M. H. Jensen, L. P. Kadanoff, I Procaccia and B. J. Shraiman. Fractal measures and their singularities: the characterization of strange sets. *Phys. Rev. A* **33** (1986), 1141–1151.

[**71**] K. G. Hare, I. D. Morris, N. Sidorov and J. Theys. An explicit counterexample to the Lagarias–Wang finiteness conjecture. *Adv. Math.* **226** (2011), 4667–4701.

[**72**] E. Harriss and O. Jenkinson. Flattening functions on flowers. *Ergod. Th. & Dynam. Sys.* **27** (2007), 1865–1886.

[**73**] H. Hentschel and I. Procaccia. The infinite number of generalized dimensions of fractals and strange attractors. *Physica D* **8** (1983), 435–444.

[**74**] M. Hochman. On the dynamics and recursive properties of multidimensional symbolic systems. *Invent. Math.* **176** (2009), 131–167.

[**75**] B. R. Hunt and E. Ott. Optimal periodic orbits of chaotic systems. *Phys. Rev. Lett.* **76** (1996), 2254–2257.

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

[**77**] B. R. Hunt, T. Sauer and J. A. Yorke. Prevalence: a translation-invariant 'almost every' on infinite-dimensional spaces. *Bull Amer. Math. Soc. (N.S.)* **27** (1992), 217–238.

[**78**] G. Iommi. Ergodic optimization for renewal type shifts. *Monatsh. Math.* **150** (2007), 91–95.

[**79**] G. Iommi and M. Todd. Natural equilibrium states for multimodal maps. *Comm. Math. Phys.* **300** (2010), 65–94.

[**80**] G. Iommi and Y. Yayama. Zero temperature limits of Gibbs states for almost-additive potentials. *J. Stat. Phys.* **155** (2014), 23–46.

[**81**] O. Jenkinson. Conjugacy rigidity, cohomological triviality, and barycentres of invariant measures. *PhD Thesis*, Warwick University, 1996.

[**82**] O. Jenkinson. Frequency locking on the boundary of the barycentre set. *Exp. Math.* **9** (2000), 309–317.

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

[**84**] O. Jenkinson. Directional entropy of rotation sets. *C. R. Acad. Sci. Paris Sér. I* **332** (2001), 921–926.

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

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

[**87**] O. Jenkinson. Ergodic optimization. *Discrete Contin. Dyn. Syst.* **15** (2006), 197–224.

[**88**] O. Jenkinson. Every ergodic measure is uniquely maximizing. *Discrete Contin. Dyn. Syst.* **16** (2006), 383–392.

[**89**] O. Jenkinson. Optimization and majorization of invariant measures. *Electron. Res. Announc. Amer. Math. Soc.* **13** (2007), 1–12.

[**90**] O. Jenkinson. A partial order on $\times 2$-invariant measures. *Math. Res. Lett.* **15** (2008), 893–900.

[**91**] O. Jenkinson. On sums of powers of inverse complete quotients. *Proc. Amer. Math. Soc.* **136** (2008), 1023–1027.

[**92**] O. Jenkinson. Balanced words and majorization. *Discrete Math. Algorithms Appl.* **1** (2009), 463–483.

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

[**94**] O. Jenkinson, R. D. Mauldin and M. Urbański. Ergodic optimization for countable alphabet subshifts of finite type. *Ergod. Th. & Dynam. Sys.* **26** (2006), 1791–1803.

[**95**] O. Jenkinson, R. D. Mauldin and M. Urbański. Ergodic optimization for non-compact dynamical systems. *Dyn. Sys.* **22** (2007), 379–388.

[**96**] O. Jenkinson and I. D. Morris. Lyapunov optimizing measures for C^1 expanding maps of the circle. *Ergod. Th. & Dynam. Sys.* **28** (2008), 1849–1860.

[**97**] O. Jenkinson and M. Pollicott. Joint spectral radius, Sturmian measures, and the finiteness conjecture. *Ergod. Th. & Dynam. Sys.* doi:10.1017/etds.2017.18. Published online: 02 May 2017.

[**98**] O. Jenkinson and J. Steel. Majorization of invariant measures for orientation-reversing maps. *Ergod. Th. & Dynam. Sys.* **30** (2010), 1471–1483.

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press

[**99**] R. Jungers. *The Joint Spectral Radius (Lecture Notes in Control and Information Sciences, 385)*. Springer, Berlin, 2009.

[**100**] A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*. Cambridge University Press, Cambridge, 1995.

[**101**] G. Keller. *Equilibrium States in Ergodic Theory*. Cambridge University Press, Cambridge, 1998.

[**102**] T. Kempton. Zero temperature limits of Gibbs equilibrium states for countable Markov shifts. *J. Stat. Phys.* **143** (2011), 795–806.

[**103**] R. P. Kertz and U. Rösler. Stochastic and convex orders and lattices of probability measures, with a Martingale interpretation. *Israel J. Math.* **77** (1992), 129–164.

[**104**] T. Kucherenko and C. Wolf. Geometry and entropy of generalized rotation sets. *Israel J. Math.* **199** (2014), 791–829.

[**105**] T. Kucherenko and C. Wolf. Ground states and zero-temperature measures at the boundary of rotation sets. *Ergod. Th. & Dynam. Sys.* doi:10.1017/etds.2017.27. Published online: 02 May 2017.

[**106**] J. C. Lagarias and Y. Wang. The finiteness conjecture for the generalized spectral radius of a set of matrices. *Linear Algebra Appl.* **214** (1995), 17–42.

[**107**] O. E. Lanford. *Entropy and Equilibrium States in Classical Statistical Mechanics (Springer Lecture Notes in Physics, 20)*. Ed. A. Lenard. Springer, Berlin, 1973, pp. 1–113.

[**108**] R. Leplaideur. A dynamical proof for convergence of Gibbs measures at temperature zero. *Nonlinearity* **18** (2005), 2847–2880.

[**109**] R. Leplaideur. Flatness is a criterion for selection of maximizing measures. *J. Stat. Phys.* **147** (2012), 728–757.

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

[**111**] A. O. Lopes and J. Mengue. Zeta measures and thermodynamic formalism for temperature zero. *Bull. Braz. Math. Soc. (N.S.)* **41** (2010), 321–338.

[**112**] A. O. Lopes and J. Mengue. Selection of measure and a large deviation principle for the general one-dimensional XY model. *Dynam. Syst.* **29** (2014), 24–39.

[**113**] A. O. Lopes, J. Mohr, R. Souza and Ph. Thieullen. Negative entropy, zero temperature and Markov chains on the interval. *Bull. Braz. Math. Soc. (N.S.)* **40** (2009), 1–52.

[**114**] A. O. Lopes, V. Rosas and R. Ruggiero. Cohomology and subcohomology problems for expansive, non Anosov geodesic flows. *Discrete Contin. Dyn. Syst.* **17** (2007), 403–422.

[**115**] A. O. Lopes and Ph. Thieullen. Sub-actions for Anosov diffeomorphisms. *Geometric Methods in Dynamics II. Astérisque* **287** (2003), 135–146.

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

[**117**] A. O. Lopes and Ph. Thieullen. Mather measures and the Bowen-Series transformation. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **23** (2006), 663–682.

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

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

[**120**] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and its Applications (Mathematics in Science and Engineering, 143)*. Academic Press, New York, 1979.

[**121**] R. D. Mauldin and M. Urbański. *Graph Directed Markov Systems: Geometry and Dynamics of Limit Sets*. Cambridge University Press, Cambridge, 2003.

[**122**] K. McGoff and A. B. Nobel. Optimal tracking for dynamical systems. *Preprint*, 2016, arXiv:1601.05033.

[**123**] T. Morita and Y. Tokunaga. Measures with maximum total exponent and generic properties of C^1 expanding maps. *Hiroshima Math. J.* **43** (2013), 351–370.

[**124**] I. D. Morris. Entropy for zero-temperature limits of Gibbs-equilibrium states for countable-alphabet subshifts of finite type. *J. Stat. Phys.* **126** (2007), 315–324.

[**125**] I. D. Morris. A sufficient condition for the subordination principle in ergodic optimization. *Bull. Lond. Math. Soc.* **39** (2007), 214–220.

[**126**] I. D. Morris. Maximizing measures of generic Hölder functions have zero entropy. *Nonlinearity* **21** (2008), 993–1000.

[**127**] I. D. Morris. The Mañé–Conze–Guivarc'h lemma for intermittent maps of the circle. *Ergod. Th. & Dynam. Sys.* **29** (2009), 1603–1611.

[**128**] I. D. Morris. Ergodic optimization for generic continuous functions. *Discrete Contin. Dyn. Syst.* **27** (2010), 383–388.

[**129**] I. D. Morris. Criteria for the stability of the finiteness property and for the uniqueness of Barabanov norms. *Linear Algebra Appl.* **443** (2010), 1301–1311.

[**130**] I. D. Morris. A rapidly-converging lower bound for the joint spectral radius via multiplicative ergodic theory. *Adv. Math.* **225** (2010), 3425–3445.

[**131**] I. D. Morris. The generalised Berger–Wang formula and the spectral radius of linear cocycles. *J. Funct. Anal.* **262** (2012), 811–824.
[**132**] I. D. Morris. Mather sets for sequences of matrices and applications to the study of joint spectral radii. *Proc. Lond. Math. Soc.* **107** (2013), 121–150.
[**133**] I. D. Morris and N. Sidorov. On a devil's staircase associated to the joint spectral radii of a family of pairs of matrices. *J. Eur. Math. Soc. (JEMS)* **15** (2013), 1747–1782.
[**134**] M. Morse and G. A. Hedlund. Symbolic dynamics II. Sturmian trajectories. *Amer. J. Math.* **62** (1940), 1–42.
[**135**] S. Newhouse. Continuity properties of entropy. *Ann. of Math. (2)* **129** (1989), 215–235.
[**136**] Ya. Pesin. *Dimension Theory in Dynamical Systems*. University of Chicago Press, Chicago, IL, 1997.
[**137**] Ya. Pesin and B. Pitskel'. Topological pressure and the variational principle for noncompact sets. *Funct. Anal. Appl.* **18** (1984), 307–318.
[**138**] W. Parry. Handwritten notes on zero temperature limits of equilibrium states, circa 1990.
[**139**] W. Parry and M. Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque* **187–188** (1990), 1–268.
[**140**] M. Pollicott and R. Sharp. Rates of recurrence for $\mathbb{Z}^q$ and $\mathbb{R}^q$ extensions of subshifts of finite type. *J. Lond. Math. Soc.* **49** (1994), 401–416.
[**141**] M. Pollicott and R. Sharp. Livsic theorems, maximizing measures and the stable norm. *Dyn. Syst.* **19** (2004), 75–88.
[**142**] A. Quas and J. Siefken. Ergodic optimization of supercontinuous functions on shift spaces. *Ergod. Th. & Dyn. Syst.* **32** (2012), 2071–2082.
[**143**] G.-C. Rota and G. Strang. A note on the joint spectral radius. *Indag. Math.* **22** (1960), 379–381.
[**144**] D. Ruelle. *Thermodynamic Formalism*. Addison-Wesley, Reading, MA, 1978.
[**145**] S. V. Savchenko. Homological inequalities for finite topological Markov chains. *Funct. Anal. Appl.* **33** (1999), 236–238.
[**146**] J. Schmeling. On the completeness of multifractal spectra. *Ergod. Th. & Dynam. Sys.* **19** (1999), 1595–1616.
[**147**] M. Shaked and J. G. Shanthikumar. *Stochastic Orders*. Springer, New York, 2007.
[**148**] K. Sigmund. Generic properties of invariant measures for Axiom A diffeomorphisms. *Invent. Math.* **11** (1970), 99–109.
[**149**] R. Souza. Sub-actions for weakly hyperbolic one-dimensional systems. *Dyn. Syst.* **18** (2003), 165–179.
[**150**] J. Steel. Concave unimodal maps have no majorisation relations between their ergodic measures. *Proc. Amer. Math. Soc.* **139** (2011), 2553–2558.
[**151**] R. Sturman and J. Stark. Semi-uniform ergodic theorems and applications to forced systems. *Nonlinearity* **13** (2000), 113–143.
[**152**] H. Takahasi. Equilibrium measures at temperature zero for Hénon-like maps at the first bifurcation. *SIAM J. Appl. Dyn. Syst.* **15** (2016), 106–124.
[**153**] F. A. Tal and S. Addas-Zanata. On maximizing measures of homeomorphisms on compact manifolds. *Fund. Math.* **200** (2008), 145–159.
[**154**] F. A. Tal and S. Addas-Zanata. Maximizing measures for endomorphisms of the circle. *Nonlinearity* **21** (2008), 2347–2359.
[**155**] P. Veerman. Symbolic dynamics of order-preserving orbits. *Physica D* **29** (1987), 191–201.
[**156**] P. Walters. Invariant measures and equilibrium states for some mappings which expand distances. *Trans. Amer. Math. Soc.* **236** (1978), 127–153.
[**157**] P. Walters. *An Introduction to Ergodic Theory*. Springer, New York, 1981.
[**158**] T.-H. Yang, B. R. Hunt and E. Ott. Optimal periodic orbits of continuous time chaotic systems. *Phys. Rev. E* **62** (2000), 1950–1959.
[**159**] G. Yuan and B. R. Hunt. Optimal orbits of hyperbolic systems. *Nonlinearity* **12** (1999), 1207–1224.
[**160**] K. Ziemian. Rotation sets for subshifts of finite type. *Fund. Math.* **146** (1995), 189–201.
[**161**] Y. Zhao. Conditional ergodic averages for asymptotically additive potentials. *Preprint*, 2014, arXiv:1405.1648.
[**162**] Y. Zhao. Maximal integral over observable measures. *Acta Math. Sinica* **32** (2016), 571–578.

https://doi.org/10.1017/etds.2017.142 Published online by Cambridge University Press