Rotation Sets and Entropy

A Dissertation Presented

by

Jaroslaw Marcin Kwapisz

to

The Graduate School

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Mathematics

State University of New York
at
Stony Brook

August 1995

State University of New York
at Stony Brook

The Graduate School

Jaroslaw Marcin Kwapisz

We, the dissertation committee for the above candidate for the Doctor of Philosophy degree, hereby recommend acceptance of the dissertation.

John Milnor
Distinguished Professor of Mathematics
Dissertation Director

Yair Minsky
Assistant Professor of Mathematics
Chairman of Defense

Mikhail Lyubich
Associate Professor of Mathematics

Folkert Tangerman
Visiting Assistant Professor
Department of Applied Mathematics and Statistics
Outside Member

This dissertation is accepted by the Graduate School.

Graduate School

Abstract of the Dissertation
Rotation Sets and Entropy

by

Jaroslaw Marcin Kwapisz

Doctor of Philosophy

in

Mathematics

State University of New York at Stony Brook

1995

This thesis is concerned with the dynamics generated on the two dimensional torus by a homeomorphism isotopic to the identity. For each rotation vector of such a map, we define *the topological entropy at the rotation vector*; this is a function having the rotation set of the map for its natural domain and the topological entropy of the map for its maximum. We show how the ideas of the thermodynamical formalism aid the evaluation of this function. For maps that are pseudo-Anosov (relative to a finite set), we demonstrate its strict convexity and real analyticity. We then describe those Gibbs states that realize the topological entropy at the rotation vector as their measure theoretic entropy.

In the general case, we show that the topological entropy is positive at the rotation vector that lies in the interior of the rotation set. We prove that any such vector is, in fact, the common rotation vector for points of a compact invariant set that carries positive topological entropy.

Presumably, the entropy on this set can be made arbitrarily close to the topological entropy at the corresponding rotation vector. We verify this for pseudo-Anosov maps and prove that, in general, both entropies are explicitly estimated from below in terms of the size of the rotation set and the relative distance of the vector from the boundary. A key tool here is the theory of quasi-conformal mappings.

Finally, we analyze the rotation sets arising in a model of a resonantly kicked charged particle in a constant magnetic field. This provides a physically interesting example to which our general theorems apply.

To Agnieszka; To the wind

Contents

List of Figures

Acknowledgements

I am indebted a great deal to my adviser John Milnor and to Philip Boyland for their support during my four years at Stony Brook.

I would also like to thank Michal Misiurewicz and Feliks Przytycki whose teaching and mentoring began my adventure with dynamical systems.

I acknowledge the Alfred Sloan Foundation for supporting this work through the doctoral dissertation fellowship.

I owe special thanks to Agnieszka Bielinska, Benjamin Hinkle, and Albert Fisher for many remarks that improved this text, and to Scott Sutherland and Brian Yarrington for their help with the computer typesetting.

Introduction

This work originated from an attempt to understand certain aspects of dynamics generated on a two-dimensional torus by a homeomorphism that is isotopic to the identity. Our main object of study is the interplay between the rotation set and the topological entropy. We put these two classical invariants in the same context by using the ideas of the *thermodynamical formalism*, and we find a qualitative relation between them by invoking the methods of the theory of quasi-conformal mappings.

In what follows, we introduce the basic concepts, advertise our main results, and show how they are a continuation of the work by others. We tend to be brief since a significant amount of expository material can be found in the subsequent chapters.

Our favorite model of a two-dimensional torus $\mathbf{T}^2$ is that obtained by identifying points of $\mathbf{R}^2$ which differ by a vector having integer coordinates. The resulting projection $\pi : \mathbf{R}^2 \to \mathbf{T}^2$ is a universal covering and we fix it once and for all. The space of all homeomorphisms of $\mathbf{T}^2$ that are isotopic to the identity will be denoted by $\mathcal{H}(\mathbf{T}^2)$. If $f : \mathbf{T}^2 \to \mathbf{T}^2$ is in $\mathcal{H}(\mathbf{T}^2)$ then it lifts to $F : \mathbf{R}^2 \to \mathbf{R}^2$ which commutes with the deck group, i.e. $F(\tilde{x} + v) = F(\tilde{x}) + v$

for all $\tilde{x} \in \mathbf{R}^2, v \in \mathbf{Z}^2$; we will write $\tilde{\mathcal{H}}(\mathbf{T}^2)$ for the space of all lifts of elements of $\mathcal{H}(\mathbf{T}^2)$.

Note that a map $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ is a C^0-bounded perturbation of the identity by $\tilde{\phi_F} = F - \mathrm{id} : \mathbf{R}^2 \to \mathbf{R}^2$ which, being invariant under precompositions with the deck translations, descends to a continuous function $\phi_F : \mathbf{T}^2 \to \mathbf{R}^2$ that is called the *displacement of* F. Thus, for $\tilde{x} \in \mathbf{R}^2$ and $x = \pi(\tilde{x})$, $F^n(\tilde{x}) - \tilde{x} = \phi_F(x) + ... + \phi_F(f^{n-1}(x))$ and the set of all limit points of the sequence $(\frac{1}{n}(F^n(\tilde{x}) - \tilde{x}))_{n \in \mathbf{N}}$ is bounded. This set is the *rotation set of* $\tilde{x}$ *under* F, denoted also by $\rho(F, \tilde{x})$.

For a point $x \in \mathbf{T}^2$, the set $\rho(F, \tilde{x})$ does not depend on the choice of $\tilde{x} \in \pi^{-1}(x)$, so we can also write $\rho(F, x)$. In the case when $\rho(F, x)$ consists of one point we call it the *rotation vector of* x. This happens, for example, when x is periodic. More generally, if μ is an ergodic probability measure for f and $x \in \mathbf{T}^2$ is typical for μ, then the Birkhoff ergodic theorem implies that $\rho(F, x) = \{\int \phi_F \, d\mu\}$.

The definition of $\rho(F, x)$ is in the spirit of Poincare's notion of the rotation number for a circle homeomorphism — for definitions and historical background one may consult [ALM93, Boy92, MZ89]. Much as in the case of degree-one circle maps ([NPT83]), the set $\rho(F, \tilde{x})$ generally changes as we vary $\tilde{x}$. Unfortunately, the union $\bigcup_{\tilde{x} \in \mathbf{R}^2} \rho(F, \tilde{x})$ often fails to be compact. For this reason, following [MZ89], we define the *rotation set of* F as

$$\rho(F) := \bigcap_{m>0} \mathrm{cl} \left(\bigcup_{n>m} \left\{ \frac{1}{n}(F^n(\tilde{x}) - \tilde{x}) : \tilde{x} \in \mathbf{R}^2 \right\} \right) \subset \mathbf{R}^2.$$

Without going into details, let us just note that other existing definitions of the rotation set yield sets that are contained in the above and may at most lack some points in the boundary. (These can not be the extremal points either.)

As the intermediate value theorem forces the rotation set for a degree-one circle map to be an interval ([NPT83]), the following result is ultimately a consequence of the Jordan simple curve theorem.

Theorem 0.0.1 ([MZ89]) *If $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ then $\rho(F)$ is convex.*

Let us note that, while the definition of $\rho(F)$ makes sense for non-invertible maps F, it renders then a set that is a continuum but often lacks convexity ([MZ89]).

Theorem 0.0.1 makes one ask: *Can any compact convex subset of $\mathbf{R}^2$ be realized as $\rho(F)$ for some $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$?* This question is still open. Even deciding which line segments are rotation sets is not yet resolved. Given any interval with rational slope passing through a point with rational coordinates it is easy to find a homeomorphism that has the interval as its rotation set. The construction in [FM90] implies that intervals with at least one rational endpoint and irrational slope can be realized as well. Based on their results for time-one maps of toral flows, Franks and Misiurewicz conjecture that these are the two only possibilities ([FM90]). In the case of nonempty interior, it is known that all polygons with vertices having rational coordinates can be realized as rotation sets ([Kwa92]). There is also an example with (countable) infinity of the extremal points ([Kwa]).

While the basic nature of the rotation set is not fully understood, there is a considerable number of results linking it to dynamical phenomena exhibited by the map. The following theorem of Franks is a pillar of the progress in this direction.

Theorem 0.0.2 ([Fra89]) *If $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ and $p \in \mathbf{Z}^2$, $q \in \mathbf{N}$ are such that $p/q \in \operatorname{int}(\rho(F))$, then $F^q(\tilde{x}) = \tilde{x} + p$ for some $\tilde{x} \in \mathbf{R}^2$. Moreover, if q and the two coordinates of p have no common divisor, then q is the least $j \in \mathbf{N}$ with the property that $F^j(\tilde{x}) - \tilde{x} \in \mathbf{Z}^2$.*

Remark. If v is an extremal point of $\rho(F)$, then $v = \int \phi_F \, d\mu$ for some ergodic measure μ ([MZ89]). Thus the assertion of the theorem is true also when p/q is extremal, by Theorem 3.5 in [Fra88].

In terms of the map $f \in \mathcal{H}(\mathbf{T}^2)$ that is a projection of F, the above theorem says that any vector v in the interior of $\rho(F)$ with rational coordinates is the rotation vector of a periodic point $x = \pi(\tilde{x}) \in \mathbf{T}^2$. Moreover, the period of this periodic point can be required to be equal to the smallest denominator q in the representation $v = p/q$, where $p \in \mathbf{Z}^2$, $q \in \mathbf{N}$. (Such periodic points are often called *primitive*.) Once again the assumption about invertibility of f is relevant; Barge and Walker have a non-invertible example with no periodic points and with the nonempty interior of the rotation set ([BW93]).

One reason why Franks's result is important is that it facilitates application of very powerful methods of the Thurston-Nilsen theory ([Thu88, FLP79, Boy94]) to $f \in \mathcal{H}(\mathbf{T}^2)$ with nonempty interior of the rotation set. Indeed, by puncturing $\mathbf{T}^2$ at the points of the orbits found by Theorem 0.0.2, one gets a surface of negative Euler characteristic with a homeomorphism that, not only does not have to be isotopic to the identity, but often is isotopic to a pseudo-Anosov map (cf. Theorem 1.2.1). This trick, going back to Bowen ([Bow78a]), was applied by Llibre and MacKay and subsequently by Misiurewicz and Ziemian to prove, among other things, the following theorems.

Theorem 0.0.3 ([LM91]) *If $f \in \mathcal{H}(\mathbf{T}^2)$, $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ is its lift, then, for any continuum $C \subset \mathrm{int}(\rho(F))$, there is a point $x \in \mathbf{T}^2$ with $\rho(F,x) = C$.*

Theorem 0.0.4 ([MZ91]) *If $f \in \mathcal{H}(\mathbf{T}^2)$ and $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ is its lift, then, for any $v \in \mathrm{int}(\rho(F))$, there is an invariant set $K \subset \mathbf{T}^2$ such that $\rho(F,x) = v$ for all $x \in K$.*

Thus nonempty interior of the rotation set implies uncountably many invariant sets each with a different non-collinear rotation vector. Analogous behavior for annulus homeomorphisms is exhibited by integrable twist maps; however, on the torus (as with degree-one circle maps), it is a sure sign of chaos.

Theorem 0.0.5 ([LM91]) *If $f \in \mathcal{H}(\mathbf{T}^2)$, $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ is its lift and $\mathrm{int}\rho(F) \neq \emptyset$, then the topological entropy $h_{\mathrm{top}}(f)$ of f is positive.*

For degree-one circle maps there exist explicit and sharp lower bounds for the topological entropy in terms of the rotation set ([ALM93]); this prompted Llibre and MacKay to ask whether there is such a relation in the case of $\mathbf{T}^2$. We give an answer to this question — a precise formulation of the result can be found in the beginning of Chapter 1. Below, we give a simplified version that is formally still a conjecture but approximates our theorem very well.

"Theorem" 1 (cf. Theorem 1.1.1) *There exist a universal constant* $C > 0$ *such that, if* $f \in \mathcal{H}(\mathbf{T}^2)$ *and* $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ *is its lift, then*

(i) $h_{\mathrm{top}}(f) \geq C \ln \mathrm{Area}(\rho(F))$, *for* $\rho(F)$ *that are not "small";*

(ii) $h_{\mathrm{top}}(f) \geq C\sqrt{\mathrm{Area}(\rho(F))}$, *for* $\rho(F)$ *that are not "large".*

Our estimate is not sharp. (Nevertheless, it captures the right asymptotics for "large" rotation sets — see the example in Chapter 1.) We think of it more as an evidence of the power of our method than an isolated result. The method that we advocate takes over where Llibre and Mackay left off and exploits the connection between the topological entropy and the quasi-conformal dilatation of the underlying pseudo-Anosov map. This approach has a lot of potential and ultimately should lead to better results.

The second main idea in this thesis is that of *topological entropy at the rotation vector*. Once again our motivation was Theorem 0.0.5 which made us ask: *If* $\mathrm{int}(\rho(F)) \neq \emptyset$, *how much topological entropy is contributed by the points with a particular rotation vector?* We make the first step towards answering this question. Namely, for a map $f \in \mathcal{H}(\mathbf{T}^2)$ and its lift $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$, we define for each $v \in \rho(F)$ the *topological entropy at the rotation vector*, denoted by $h_{\mathrm{top}}^{(v)}(F)$. We show that $h_{\mathrm{top}}^{(v)}(F)$ obeys a form of the *thermodynamical formalism*, which we then employ to give a complete analysis of the pseudo-Anosov case and, ultimately, strengthen "Theorem" 1 to the following:

"Theorem" 2 (cf. Theorem 3.1.1) *For some universal constant* $C > 0$, *if* $f \in \mathcal{H}(\mathbf{T}^2)$ *is a diffeomorphism,* $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ *is its lift and* $v \in \mathrm{int}(\rho(F))$, *then there exists a compact invariant set* $K \subset \mathbf{T}^2$ *such that* $\rho(F, x) = v$, *for all*

$x \in K$, and

$$h_{\mathrm{top}}^{(v)}(f) \geq h_{\mathrm{top}}(f|_K) \geq C \ln \mathrm{Area}(\rho(F)) \cdot \tau(v),$$

for $\rho(F)$ that are not "small",

$$h_{\mathrm{top}}^{(v)}(f) \geq h_{\mathrm{top}}(f|_K) \geq C \sqrt{\mathrm{Area}(\rho(F))} \cdot \tau(v),$$

for $\rho(F)$ that are not "large".

The quantity $\tau(v)$ is a certain measure of the "distance" of v from the boundary of $\rho(F)$.

Ultimately, one should be able to make the choice of the set K above so that $h_{\mathrm{top}}(f|_K)$ is arbitrarily close to $h_{\mathrm{top}}^{(v)}(f)$. We prove this only for maps that are *pseudo-Anosov relative to a finite set* (see [Boy94] for the definition). Another important issue, which we resolve here only for that class of maps, is the regularity of $h_{\mathrm{top}}^{(v)}$ as a function of v. We have the following theorem (summarizing Proposition 3.2.2 and Theorem 3.2.1).

Theorem 0.0.6 *If $f \in \mathcal{H}(\mathbf{T}^2)$ is pseudo-Anosov relative to a finite set and $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ is its lift, then*

(i) $h_{\mathrm{top}}^{(v)}(F)$ is real-analytic and strictly concave on $\mathrm{int}(\rho(F))$;

(ii) for any $v \in \mathrm{int}(\rho(F))$, there exists a fully supported ergodic measure $\mu^{(v)}$ such that the measure theoretic entropy $h_{\mu^{(v)}}(f) = h_{\mathrm{top}}^{(v)}(F)$ and $\int \phi_F \, d\mu^{(v)} = v$;

(iii) for any $v \in \mathrm{int}(\rho(F))$ and $\eta \in (0,1)$, there exists a compact invariant subset K of $\mathbf{T}^2$ such that $h_{\mathrm{top}}(F|_K) \geq \eta h_{\mathrm{top}}^{(v)}(F)$ and $\rho(F,x) = \{v\}$ for all $x \in K$.

Actually, there is more in our analysis than the above theorem spells out; for example, $h^{(v)}_{\text{top}}(F)$ can be explicitly calculated from the Markov partition.

The line of attack in proving the theorem is not original at all: we use symbolic dynamics to reduce it to facts about subshifts of finite type. Still, our discussion of subshifts of finite type contains some new results. In particular, the following theorem may be useful in many applications.

Theorem 0.0.7 (cf. Theorem 2.2.4) *Suppose $\sigma : \Lambda \to \Lambda$ is a transitive subshift of finite type and $\psi : \Lambda \to \mathbf{R}^d$ is locally constant. If $v = \int \psi \, d\mu$ for some invariant ergodic probability measure μ, then, for any $\eta \in (0,1)$, there exists a compact invariant set K such that $h_{\text{top}}(\sigma|_K) \geq \eta h_\mu(\sigma)$ and*

$$\sup_{n \in \mathbf{N}, x \in K} \left\| \sum_{i=0}^{n-1} \psi(f^i(x)) - nv \right\| < +\infty.$$

The above theorem for integer valued ψ and integer v coincides with the result in [MT92]. There, it is applied to calculate the minimal entropy of some sequences of patterns of periodic orbits for one-dimensional maps. The discussion in [MT92] yields concrete formulas only for symmetric patterns. Our methods (which are different) can be used to treat equally efficiently the general case. In particular, the *Gibbs states* that we discuss realize the variational principle in [MT92] thus answering the question posed there. [1]

As a tool for comprehending $h^{(v)}_{\text{top}}$ we develop (in Chapter 2) a weak but quite general version of the *thermodynamical formalism* for a dynamical system on a compact space with a continuous observable in $\mathbf{R}^d$. Theorem 2.1.1

[1] The relation of my work to that of [MT92] was brought to my attention only recently. I will give the details elsewhere.

plays the central role here by connecting the entropy $h_{\rm top}^{(v)}$ with the appropriate *pressure function*. Even though this theorem can be viewed as a consequence of the standard thermodynamical formalism ([Wal82]), the implication is not trivial; to establish it, we generalize the variational principle for topological entropy ([Wal82]) by making it sensitive to the averages of the observable (see Theorem 2.1.2). This material should be thought of as just another realization of the paradigm of the thermodynamical formalism ([Rue78, Bow75, Wal82]). We hope that our presentation will help to absorb these ideas into the theory of rotation sets (including the context of other surfaces than torus, see [Fra94, Boy94].)

Let us also note that in dealing with $h_{\rm top}^{(v)}$ one can get pretty far using elementary *large deviation* techniques. The inspiration for this approach comes from a book by Ellis ([Ell85]). In Appendix A we use it to give a proof of the *thermodynamical formalism* theorem (Theorem 2.1.1).

Finally, our last chapter is devoted to a concrete family of torus homeomorphisms that are isotopic to the identity. The purpose of this discussion is two-fold. For one, we wanted to give a concrete physical model that satisfies the assumptions of our theorems. The other reason is mathematical: these are perhaps the simplest and most accessible analytically mappings exhibiting nonempty interior of the rotation set; even though studied from other points of view for quite some time, they somehow avoided to this date theoretical analysis with the emphasis on their rotation sets ([LL93, ZZRSC86b, CSUZ87, ZZRSC86a]). We give the beginning of such an analysis. In particular, we prove monotonicity of the rotation set for a certain (physically) natural one-parameter family of torus maps. This may be interesting since such examples were not known before.

Chapter 1

An estimate of entropy for toroidal chaos

In this chapter we show that topological entropy can be estimated from below by looking at the "size" of the rotation set. This "size" can be roughly thought of as the area of the rotation set (but, in general, is not comparable with it).

1.1 Statement of the result

Consider $f \in \mathcal{H}(\mathbf{T}^2)$ and its lift $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$. Assume that $\mathrm{int}(\rho(F)) \neq \emptyset$. It was proved in [LM91] that f exhibits then *toroidal chaos*; in particular, f has positive topological entropy $h_{\mathrm{top}}(f) > 0$ (see Theorem 0.0.5). Our goal is to provide a lower bound on the entropy in terms of the *size of the rotation set*. To formulate the result, let us first explain how we measure the *size of the rotation set*.

Definition 1.1.1 *Let $K \subset \mathbf{R}^2$ be convex and compact. If* $\mathrm{int}(K)$ *contains three non-collinear points of $\mathbf{Z}^2$, then we define*

$$A(K) := \max\{\sqrt{r_1 r_2} : \ r_1, r_2 \in \mathbf{N} \text{ and there exist } \ p, v, w \in \mathbf{Z}^2 \text{ such that}$$
$$p, p + r_1 v, p + r_2 w \in \mathrm{int}(K) \text{ and are non-collinear}\},$$

$$I(K) := 1.$$

Otherwise, we set

$$A(K) := 0,$$

$$I(K) := \inf\{\lambda \in \mathbf{N} : \ \mathrm{int}(\lambda K) \text{ contains three non-collinear points of } \mathbf{Z}^2\}.$$

Note that, for $m \in \mathbf{N}$, if $A(K) > 0$, then $A(mK) \geq mA(K)$, and that

$$A(I(K) \cdot K) \geq 1. \tag{1.1}$$

Theorem 1.1.1 (Main Lower Bound ([Kwa93])) *There exist universal constants C, C' such that, for any $f \in \mathcal{H}(\mathbf{T}^2)$ and its lift $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$, we have*

$$h_{\mathrm{top}}(f) \geq C \ln_+ A(\rho(F)),$$

$$h_{\mathrm{top}}(f) \geq C' / I(\rho(F)).$$

We think of the first estimate as significant for "large" $\rho(F)$ and of the second as significant for "small" $\rho(F)$.

Remark. In fact, we will prove that if r_1 and r_2 realizing $A(\rho(F))$ are both greater than 5 then

$$h_{\mathrm{top}}(f) \geq \frac{1}{3} \ln \sqrt{(r_1 - 1)(r_2 - 1)} - \frac{1}{2} \ln 4.$$

By considering the eighth iterate of F, this yields $C = 1/27$ and $C' = \ln 2/54$ — see Claim 1.4.1 and the discussion opening Section 1.4.

As indicated by the remark, our statement of the theorem is a compromise between sharpness of the result and complexity of its formulation. It will be clear from the proof that we are far from providing optimal estimates, so we see no point in complicating the inequalities. Nevertheless, the theorem is asymptotically sharp (up to a multiplicative constant) for "large" rotation sets, as shown by the following example.

Example. For $n \in \mathbf{N}$, define $U_n, V_n \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ by $V_n(x, y) := (x+n\sin(2\pi y), y)$, $U_n(x, y) := (x, y + n\sin(2\pi x))$, $x, y \in \mathbf{R}$. Consider $F_n := V_n \circ U_n \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ and the corresponding $f_n \in \mathcal{H}(\mathbf{T}^2)$. The points $(0, 0), (\frac{1}{4}, 0), (0, \frac{1}{4})$ have rotation vectors $(0, 0), (0, n), (n, 0)$ respectively, so $A(\rho(F_n)) \geq n$. On the other hand, the maximal Lyapunov exponent dominates $h_{\text{top}}(f_n)$ and is trivially bounded by $\ln \sup \|\nabla f_n\| \leq \ln(2\pi n)$. In this way, we have

$$\limsup_{n\to\infty} \ln\left(h_{\text{top}}(f_n)/A(\rho(F_n))\right) \leq 1.$$

One would like to have an example of a family $\{f_\epsilon\}_{\epsilon>0}$, $f_\epsilon \in \mathcal{H}(\mathbf{T}^2)$, with lifts $\{F_\epsilon\}_{\epsilon>0}$ having $\rho(F_\epsilon)$ shrinking to a point, as $\epsilon \to 0$, and such that $h_{\text{top}}(f_\epsilon) \cdot I(\rho(F_\epsilon)) \leq \text{Const}$, for $\epsilon > 0$. This would mean that our estimate is asymptotically sharp for "small" rotation set. No such example is known to the author.

The rest of this chapter is devoted to the proof of the theorem. We isolate the main steps into separate sections. Section 1.2 describes reduction to the case of a pseudo-Anosov map. Section 1.3 explains how entropy is linked with the quasi-conformal dilatation of conformal structures on $\mathbf{T}^2$. The very estimates proving the theorem can be found in Section 1.4.

1.2 Comparison with pseudo-Anosov maps

We describe here a trick that allows one to pass from investigation of a general map in $\mathcal{H}(\mathbf{T}^2)$ with nonempty interior of the rotation set to a pseudo-Anosov map with approximately the same rotation set. Our version is only a trivial extension of the original one in [LM91]. We use the notion of a *pseudo-Anosov map* (of a surface) *rel a finite invariant set*. Referring the reader to [Boy94] for an exposition, we only say here that, if $f : S \to S$ is a *pseudo-Anosov map rel a finite invariant set* P, then it is a factor of a *pseudo-Anosov map* ([Thu88, FLP79]) on the surface obtained by compactifying the ideal boundary of $S \setminus P$ with circles. The factor map collapses the boundary circles to the corresponding points of P and is injective otherwise. Thus the dynamics of the two maps almost coincide. (In particular, they have the same topological entropy ([Bow78a]).)

Theorem 1.2.1 *Suppose that* $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ *and* $\mathrm{int}(\rho(F)) \neq \emptyset$. *Let* $f \in \mathcal{H}(\mathbf{T}^2)$ *be the map that lifts to* F. *If* $\rho' \subset \mathrm{int}(\rho(F))$ *is a convex non-degenerate polygon with vertices in* $\mathbf{Q}^2$, *then there exists a finite* f*-invariant set* $P \subset \mathbf{T}^2$ *and a pseudo-Anosov rel* P *map* $g \in \mathcal{H}(\mathbf{T}^2)$ *that is isotopic to* f *rel* P *and has a lift* $G \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ *with* $\rho' \subset \rho(G)$.

Remark. The above theorem is also true when ρ' has some vertices on the boundary of $\rho(F)$ as long as these are extremal points of $\rho(F)$ (cf. the remark after Theorem 0.0.2).

Pseudo-Anosov maps have minimal topological entropy in their isotopy class ([Thu88, FLP79]), so

$$h_{\text{top}}(f) \geq h_{\text{top}}(g) > 0. \tag{1.2}$$

Remark. Pseudo-Anosov maps not only minimize entropy but also have "the simplest dynamics" in their isotopy class ([Han85, Boy10]). In particular, we have $\rho(G) \subset \rho(F)$. (For details and proofs see Theorem 3.3.1 and the corollary after it.)

Proof of Theorem 1.2.1. Let $v_1, ..., v_m \in \mathbf{Q}^2$ be the vertices of ρ' (or, equally well, any finite set in $\mathbf{Q}^2$ such that $\rho' = \text{conv}\ \{v_1, ..., v_m\}$, where *conv* abbreviates convex hull). Write $v_i = p_i/q_i$, where $p_i \in \mathbf{Z}^2$, $q_i \in \mathbf{N}$ and the two components of p_i and q_i have no common divisor, all this for $i = 1, ..., m$. By Theorem 0.0.2, there exist $x_1, ..., x_m \in \mathbf{R}^2$ such that $F^{q_i}(x_i) = x_i + p_i$, $i = 1, ..., m$. Let P be the union of the projections on the torus of the orbits of $x_1, ..., x_m$. If $g \in \mathcal{H}(\mathbf{T}^2)$ is isotopic to f rel P and $G \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ is the lift of g equivariantly isotopic to F relative $\pi^{-1}(P)$, then $\rho' = \text{conv}\{v_1, ..., v_m\} \subset \rho(G)$. In particular, we can take for g the Thurston canonical representative of the isotopy class of f relative P (see e.g. Th. 7.1 in [Boy94]). Following [LM91], one can argue that g can not have any *reducing curves*, and so it is pseudo-Anosov. Indeed, since a reducing curve must be non-peripheral, we have a priori three possibilities: it is essential on $\mathbf{T}^2$, and then $\rho(G)$ would have to be contained in a line; it bounds a disk containing points of different orbits in P, which would contradict the fact that the rotation vectors of these orbits are different; it bounds a disk containing two or more points of one orbit in P, which would force p_i and q_i to have a common divisor for some i, contrary to the assumption. Q.E.D.

1.3 Entropy and conformal structures

It is clear from the previous section that, to prove the Main Lower Bound, we merely have to correlate the size of the rotation set with the topological entropy for pseudo-Anosov maps. A natural first impulse is to try linking the two via standard symbolic dynamics given by the Markov partitions. This however leaves one with a task of understanding how topology of the torus influences the structure of the corresponding subshifts of finite type. We propose here a different approach. Namely, we will take advantage of the fact that the the entropy of the pseudo-Anosov map g can be calculated by looking at its quasi-conformal dilatation with respect to a preferred family of conformal structures on $\mathbf{T}^2$. We explain this below. (Following [Alh66], we abbreviate *quasi-conformal* to q.c.)

Denote by $\mathcal{C}(\mathbf{T}^2 \setminus P)$ the set of all conformal structures on $\mathbf{T}^2 \setminus P$. If $\sigma \in \mathcal{C}(\mathbf{T}^2 \setminus P)$ and $\psi : \mathbf{T}^2 \setminus P \to \mathbf{T}^2 \setminus P$ is a σ-q.c. map, then we write $K_\sigma[\psi]$ for the (maximal) q.c. dilatation of ψ ([Alh66]). (For conformal maps $K = 1$.) According to Bers (Ch. III, [Abi80]), there exists $\sigma_g \in \mathcal{C}(\mathbf{T}^2 \setminus P)$ solving the following extremal problem

$$K_{\sigma_g}[g] = \inf\{K_\sigma[\psi] : \ \sigma \in \mathcal{C}(\mathbf{T}^2 \setminus P), \ \psi \text{ isotopic to } g\}. \tag{1.3}$$

Moreover, σ_g is modeled around each boundary component of $\mathbf{T}^2 \setminus P$ on a punctured disc (see Th. 3, §2, Ch. III, [Abi80]). Thus σ_g extends uniquely through the punctures to a conformal structure on $\mathbf{T}^2$ which, with no risk of confusion, is also denoted by σ_g.

For the sake of giving a more intuitive picture, let us sketch an alternative (and more geometrical) way of finding σ_g. From the definition of pseudo-Anosov map relative a finite set (see [Boy94]), there exists a pair of transverse measured singular foliations $\mathcal{F}^s$ and $\mathcal{F}^u$ such that $g(\mathcal{F}^s) = \lambda^{-1}\mathcal{F}^s$ and $g(\mathcal{F}^u) = \lambda\mathcal{F}^u$, where $\lambda = \exp h_{\text{top}}(g)$. The two foliations define a smooth structure together with a flat Riemannian metric on the complement of the set of singularities $Q \subset \mathbf{T}^2$ ($P \subset Q$). This Riemannian metric naturally determines a conformal structure on $\mathbf{T}^2 \setminus Q$. However, due to the form of $\mathcal{F}^s$ and $\mathcal{F}^u$ at the singularities that is postulated by the definition, this structure has removable singularities at points of Q and so determines a unique conformal structure σ_g on $\mathbf{T}^2$. From the action of g on $\mathcal{F}^s$ and $\mathcal{F}^u$ one sees the q.c. dilatation to be ([Abi80])

$$K_\sigma[g] = \lambda^2 = \exp(2h_{\text{top}}(g)).$$

Putting this together with inequality (1.2) we get

$$h(f) \geq (1/2) \ln K_{\sigma_g}[g]. \tag{1.4}$$

In our applications we will not use any a priori knowledge of σ_g, thus, weakening (1.4), we state the following theorem.

Theorem 1.3.1 *If* $f \in \mathcal{H}(\mathbf{T}^2)$ *and* P *is a collection of* primitive *periodic orbits with different non-collinear rotation vectors, then*

$$h_{\text{top}}(f) \geq \inf\{(1/2)\ln K_\sigma[\psi] : \ \sigma \in \mathcal{C}(\mathbf{T}^2),$$
$$\psi \ \sigma\text{-}q.c. \ map\ of\ \mathbf{T}^2 \ isotopic\ to\ f\ rel\ P\}.$$

Let us note here a curious problem related to the above theorem.

Open Problem. If S is a compact Riemann surface and $f : S \to S$ is a q.c. homeomorphism with q.c. dilatation $K[f]$, is $h_{\text{top}}(f) \leq (1/2)\ln K[f]$?

For a C^1-smooth map, the answer is yes. Indeed, by the variational principle for topological entropy ([Man87]), we just need to prove the following: if μ is an ergodic invariant measure, then $h_\mu(f) \leq \ln K[f]/2$. Consider the Lyapunov exponents $\chi^- \leq \chi^+$ of μ. By Ruelle's inequality ([Man87]), we can write $h_\mu(f) \leq \max\{0, \chi^+\}$ and $h_\mu(f) = h_\mu(f^{-1}) \leq \max\{0, -\chi^-\}$. Thus either $h_\mu(f) = 0$, and there is nothing to prove, or $h_\mu(f) \leq \min\{\chi^+, -\chi^-\} \leq (\chi^+ - \chi^-)/2$. By looking at the derivative $Df^n(x)$ at a typical point x and for large n, we get $\exp\chi^+/\exp\chi^- \leq K[f]$, which already implies our claim. In the non-smooth case, there are estimates by Gromov with a non-optimal multiplicative constant ([Gro87]).

1.4 Proof of the Main Lower Bound

We have two inequalities to prove

$$h_{\mathrm{top}}(f) \geq C \ln_+ A(\rho(F)), \tag{1.5}$$

and

$$h_{\mathrm{top}}(f) \geq C'/I(\rho(F)). \tag{1.6}$$

In fact, (1.6) follows from (1.5). For a proof, set $\lambda := I(\rho(F)) \in \mathbf{N}$. Using (1.1), we get $A(\rho(F^{2\lambda})) = A(2\lambda\rho(F)) \geq 2$. Application of (1.5) to $f^{2\lambda}$ yields $2\lambda\, h_{\mathrm{top}}(f) = h_{\mathrm{top}}(f^{2\lambda}) \geq C \ln 2$. The inequality (1.6) with $C' = (C \ln 2)/2$ follows.

In the rest of this section we tackle (1.5). Let $p, v, w \in \mathbf{Z}^2$ and $r_1, r_2 \in \mathbf{N}$ be as in the definition of $A(\rho(F))$. Passing perhaps to the sixth iterate of F, we may assume that $r_1, r_2 \geq 6$. Also, it is easy to see that we can require that v, w form a basis of $\mathbf{Z}^2$ as a $\mathbf{Z}$-module. By the theorem of Franks (Theorem 0.0.2), for each of $p, p + r_1 v, p + r_2 w \in \rho(F) \cap \mathbf{Z}^2$, there is a fixed point of f realizing it as the rotation vector. Let P be the union of those three fixed points. By Theorem 1.3.1, to estimate $h_{\mathrm{top}}(f)$, we need to consider an arbitrary conformal structure $\sigma \in \mathcal{C}(\mathbf{T}^2)$ and a σ-q.c. map $g : \mathbf{T}^2 \to \mathbf{T}^2$ which is isotopic to f rel P. By the uniformization theorem ([Abi80]), the conformal torus $(\mathbf{T}^2, \sigma)$ is biholomorphic to $\mathbf{C}/\Gamma$, where Γ is a rank-two lattice in $\mathbf{C}$. The map g lifts to a map $\Psi : \mathbf{C} \to \mathbf{C}$ which is q.c. with respect to the standard conformal structure on $\mathbf{C}$ with the q.c. dilatation $K[\Psi] = K_{\sigma_g}[g]$. Let $e_1, e_2 \in \Gamma$ be the vectors corresponding to the vectors v, w. After perhaps post-composing Ψ

with a deck transformation, the three fixed points in P lift to $p_0, p_1, p_2 \in \mathbf{C}$ such that

$$\Psi(p_0) = p_0,\ \ \Psi(p_1) = p_1 + r_1 e_1,\ \ \Psi(p_2) = p_2 + r_2 e_2. \tag{1.7}$$

Our task is to verify, from the above topological data on Ψ, the following claim which already implies (1.5).

Claim 1.4.1 (target estimate) *If $r_1, r_2 \geq 6$, then*

$$K[\Psi] \geq \frac{1}{4}(r_1 - 1)^{1/3}(r_2 - 1)^{1/3}. \tag{1.8}$$

The proof of the claim is rather technical. To better separate the ideas, after necessary definitions, we formulate the key observations as separate *Facts* and then prove them. All the pieces of the puzzle are put together in the final argument at the end of this section.

For any ordered pair (a, b) of orthogonal vectors in $\mathbf{C}$, we have the rectangle $R = R(a, b) = \{xa + yb :\ \ 0 \leq x, y \leq 1\}$. For any $s \in [0, 1]$, let $\alpha_s[R] : [0, 1] \to R$ be given by $\alpha_s[R](t) = ta + sb$, and $\beta_s[R] : [0, 1] \to R$ be given by $\beta_s[R](t) = sa + tb$. Thus we have two families of curves filling R: *a-family* $\mathcal{A} := \{\alpha_y[R]\}_{y \in [0,1]}$, and *b-family* $\mathcal{B} := \{\beta_x[R]\}_{x \in [0,1]}$. Clearly, $\partial R = \alpha_0[R] \cup \alpha_1[R] \cup \beta_0[R] \cup \beta_1[R]$. We call the rectangle R *fundamental for* Γ if every orbit under of Γ intersects its interior exactly once. All the above definitions extend in an obvious way to any translate of the rectangle R.

For a family $\mathcal{G}$ of rectifiable curves in $\mathbf{C}$, we denote the extremal length of $\mathcal{G}$ by $\lambda(\mathcal{G})$, i.e.

$$\lambda(\mathcal{G}) := \sup \left\{ L(\rho)^2 :\ \ \rho : \mathbf{C} \to \mathbf{R}^+ \text{ measurable with } \int\!\!\int \rho^2 dx dy = 1 \right\},$$

where

$$L(\rho) := \inf \left\{ \int_\gamma \rho |dz| : \ \gamma \in \mathcal{G} \right\}.$$

For the Euclidean length of curve γ we simply write $l(\gamma)$.

For the families $\mathcal{A}$ and $\mathcal{B}$ we have ([Alh66])

$$\lambda(\mathcal{A}) = |a|^2/|R|, \ \lambda(\mathcal{B}) = |b|^2/|R|, \tag{1.9}$$

where $|R|$ stands for the Euclidean area of R. Our estimation of $K[\Psi]$ will be based on the *quasi-invariance* of the extremal length, a property encapsulated by the following inequality ([Alh66])

$$K[\Psi] \geq \lambda(\Psi\mathcal{G})/\lambda(\mathcal{G}). \tag{1.10}$$

For $\mathcal{G}$ above we will take *a-families* and *b-families* of appropriate rectangles. The following simple lemma plays a key role in finding "good rectangles".

Lemma 1.4.1 *If $v, w \in \mathbf{C}$ are orthogonal, $v \in \Gamma$ and $R = R(v, w)$ is fundamental for Γ, then there exists $z \in \mathbf{C}$, for which $R' := R + z$ satisfies*

$$\lambda(\Psi\mathcal{A}) \geq l(\Psi\alpha_0)^2/|R'| = l(\Psi\alpha_1)^2/|R'|, \tag{1.11}$$

$$\lambda(\Psi\mathcal{B}) \geq l(\Psi\beta_0)^2/|R'| = l(\Psi\beta_1)^2/|R'|, \tag{1.12}$$

where $\mathcal{A} = \{\alpha_y\}$, $\mathcal{B} = \{\beta_x\}$ are the a-family and b-family for R'.

Proof. First slide R along the line $\mathbf{R}w$ to find t_0 such that

$$l(\Psi(\alpha_0[R] + t_0 w)) = \inf\{l(\Psi(\alpha_0[R] + tw)) : \ t \in \mathbf{R}\}.$$

This means that

$$l(\Psi\alpha_0[R+z]) = \inf\{l(\Psi\alpha):\ \alpha \in a\text{-}family\ of\ R+z\}, \qquad (1.13)$$

where $z = t_0 w$. Observe that (1.13) holds also for any $z = t_0 w + sv$, $s \in \mathbf{R}$, because for any $s_1, s_2 \in \mathbf{R}$ the curves $\alpha_y[R + t_0 + s_i v]$, $i = 1, 2$, project to the same loop on the torus and so their images under ψ have equal length. In this way, if sliding along $\mathbf{R}v$ we have found s_0 such that

$$l(\Psi\beta_0[R+z]) = \inf\{l(\Psi\beta):\ \beta \in b\text{-}family\ of\ R+z\},\ z = t_0 w + s_0 v, \quad (1.14)$$

we got both (1.13) and (1.14) satisfied for $z = t_0 w + s_0 v$. Set $R' := R + z$. The inequalities in (1.11) and (1.12) follow from the definition of the extremal length when one takes for ρ a multiple of the characteristic function of $\Psi R'$ while remembering that $|R'| = |\Psi R'|$. The equalities are trivial since α_0 and α_1, as well as β_0 and β_1, project to the same arcs on the torus. Q.E.D.

Let $\theta \in [0, \pi)$ be the measure of the convex angle between e_1 and e_2.

Fact 1.4.1 *For $i, j \in \{1, 2\}$, $i \neq j$, we have*

$$\begin{aligned} 8(|e_i|^2 + \sin^2\theta |e_j|^2) K[\Psi] &\geq (r_i - 1)^2 |e_i|^2 + (r_j - 1)^2 \sin^2\theta |e_j|^2 \\ &+ \max\{0, r_j |e_j| |\cos\theta| - |e_i|\}^2. \end{aligned}$$

Proof. Let $v := e_i$ and w be the component of e_j orthogonal to e_i. Let R' be as in Lemma 1.4.1. Due to (1.7), $\Psi(R')$ intersects R', $R' + r_i e_i$, $R' + r_j e_j$, so

$$l(\Psi(\partial R')) \geq \mathrm{dist}(R', R' + r_i e_i) + \mathrm{dist}(R', R' + r_j e_j),$$

and consequently, (using the Pythagorean theorem),

$$\begin{aligned} l(\Psi(\partial R')) &\geq (r_i - 1)|e_i| \\ &+ \left((r_j - 1)^2 |e_j|^2 \sin^2\theta + \max\{0, r_j|e_j||\cos\theta| - |e_i|\}^2\right)^{1/2}. \end{aligned} \tag{1.15}$$

Feeding (1.9), (1.11), (1.12) into (1.10), we obtain

$$\begin{aligned} (|e_i|^2 + \sin^2\theta |e_j|^2) K[\Psi] &\geq |e_i|^2 \cdot \lambda(\Psi\mathcal{A})/\lambda(\mathcal{A}) + |e_j|^2 \sin^2\theta \cdot \lambda(\Psi\mathcal{B})/\lambda(\mathcal{B}) \\ &\geq |e_i|^2 \frac{l(\Psi\alpha_0)^2}{|R|} \frac{|R|}{l(\alpha_0)^2} + |e_j|^2 \sin^2\theta \frac{l(\Psi\beta_0)^2}{|R|} \frac{|R|}{l(\beta_0)^2} \\ &= l(\Psi\alpha_0)^2 + l(\Psi\beta_0)^2 \\ &\geq \frac{1}{2}(l(\Psi\alpha_0) + l(\Psi\beta_0))^2 \\ &= \frac{1}{8} l(\Psi(\partial R'))^2. \end{aligned}$$

This inequality put together with (1.15) proves Fact 1.4.1. Q.E.D.

Note that, if the lengths of e_1 and e_2 were "comparable", Fact 1.4.1 would yield an estimate similar to (1.8) of Claim 1.4.1. The following fact exhibits one mechanism preventing the $|e_1|$ and $|e_2|$ from being too disproportional.

Fact 1.4.2 *If $d \in (0,1)$ and $e_1 = 1$, $e_2 = d\sqrt{-1}$, $\Gamma = \mathbf{Z} + d\sqrt{-1}\,\mathbf{Z}$, then*

$$d^2 K[\Psi] \geq (r_1 - 1)^2.$$

Proof. Let R' be the rectangle obtained by the application of Lemma 1.4.1 with $v := 1$ and $w := d\sqrt{-1}$. Set $\gamma := \bigcup\{\beta_1 + kd\sqrt{-1} : \ k \in \mathbf{Z}\}$. Clearly, γ is a straight line cutting $\mathbf{C}$ in two half-planes. Denote by H^- the half-plane containing R' and by H^+ the one containing $R' + 1$. The rectangle R'

is fundamental, so, by (1.7), there are $q_0, q_1 \in R'$ such that $\Psi(q_0) = q_0$ and $\Psi(q_1) = q_1 + r_1$. Observe that, because $q_0 + 1 = \Psi(q_0 + 1) \in \Psi(H^+)$, we have

$$q_0 = \Psi(q_0) \in \Psi(H^-) \quad \text{and} \quad q_0 + 1 = \Psi(q_0 + 1) \notin \Psi(H^-),$$

and, because $q_1 + r_1 + 1 = \Psi(q_1 + 1) \in \Psi(H^+)$, also

$$q_1 + r_1 = \Psi(q_1) \in \Psi(H^-) \quad \text{and} \quad q_1 + r_1 + 1 \notin \Psi(H^-).$$

Thus straight line segments in $\mathbf{C}$, $[q_0, q_0 + 1]$ and $[q_1 + r_1, q_1 + r_1 + 1]$, both intersect $\Psi(\gamma) = \partial(\Psi(H^-))$. Consequently, if pr_x is the orthogonal projection on the real axis $\mathbf{R} \subset \mathbf{C}$, then $\mathrm{diam}(\mathrm{pr}_x(\Psi\gamma)) \geq r_1 - 1$. This, using Γ-equivariance of Ψ, yields

$$l(\Psi\beta_1) \geq \mathrm{diam}(\mathrm{pr}_x(\Psi\beta_1)) = \mathrm{diam}(\mathrm{pr}_x(\Psi\gamma)) \geq r_1 - 1.$$

Now we apply (1.9), (1.10) and (1.12) to get

$$K[\Psi] \geq \lambda(\Psi\mathcal{B})/\lambda(\mathcal{B}) \geq l(\Psi\beta_1)^2/d^2 \geq (r_1 - 1)^2/d^2,$$

and we are done. Q.E.D.

Fact 1.4.3 *Suppose that* $e_1 = 1$, $e_2 = l\cos\theta + l\sin\theta\sqrt{-1}$, $l \in (0, 1)$. *If*

$$\sin^2\theta \geq (r_1 - 1)^2/(32K[\Psi]),$$

then

$$K[\Psi] \geq \frac{1}{4}(r_1 - 1)^{1/3}(r_2 - 1)^{1/3}.$$

Proof. Put $d := l\sin\theta$. First assume that $\theta = \pi/2$. Fact 1.4.1, applied with $i = 1$, $j = 2$ (i.e. $e_i = 1$ and $e_j = d\sqrt{-1}$), gives us

$$8(1+d^2)K[\Psi] \geq (r_1-1)^2 + (r_2-1)^2 d^2 \geq (r_2-1)^2 d^2.$$

Inserting in the above inequality the estimate of d given by Fact 1.4.2, we derive

$$8 \cdot 2 \cdot K[\Psi] \geq (r_2-1)^2 (r_1-1)^2 / K[\Psi],$$

and finally,

$$K[\Psi] \geq \frac{1}{4}(r_1-1)(r_2-1). \tag{1.16}$$

In general, $\theta \neq \pi/2$ and then we consider $\Phi := L_\theta \circ \Psi \circ L_\theta^{-1}$, where $L_\theta : \mathbf{C} \to \mathbf{C}$ is given by $L_\theta(x + y\sqrt{-1}) := x + \cot\theta\, y + y\sqrt{-1}$ and transforms the lattice $\mathbf{Z} + d\sqrt{-1}\mathbf{Z}$ to $\mathbf{Z} + (l\cos\theta + l\sin\theta\sqrt{-1})\mathbf{Z}$. The map L_θ is area-preserving, so $K[L_\theta]$ equals the largest eigenvalue of $L_\theta^* L_\theta$, where L_θ^* is the adjoint of L_θ. The trace of $L_\theta^* L_\theta$ is $2 + \cot^2\theta$, so we have

$$K[L_\theta] \leq 2 + \cot^2\theta.$$

Observe that (1.16) holds once Ψ is replaced by Φ. Since by sub-multiplicativity of q.c. dilatation

$$K[\Psi] \geq K[\Phi]/K[L_\theta]^2,$$

we get

$$K[\Psi] \geq \frac{1}{4}(r_1-1)(r_2-1)/(2+\cot^2\theta)^2. \tag{1.17}$$

Our hypothesis on θ can be written as

$$\cot^2\theta \leq \sin^{-2}\theta \leq 32K[\Psi]/(r_1-1)^2. \tag{1.18}$$

If $\cot^2\theta \le 1$, then, by (1.17) and using $r_1, r_2 \ge 6$, we see that

$$\begin{aligned} K[\Psi] &\ge \frac{1}{4\cdot 3^2}(r_1-1)(r_2-1) \ge \frac{5^{2/3}5^{2/3}}{4\cdot 9}(r_1-1)^{1/3}(r_2-1)^{1/3} \\ &\ge \frac{1}{4}(r_1-1)^{1/3}(r_2-1)^{1/3}. \end{aligned}$$

If $\cot^2\theta \ge 1$, then, by (1.17) and (1.18), we obtain

$$K[\Psi] \ge \frac{1}{4}(r_1-1)(r_2-1)/(3\cot^2\theta)^2 \ge \frac{1}{4\cdot 3^2\cdot 32^2}(r_1-1)^5(r_2-1)/K[\Psi]^2,$$

which yields

$$\begin{aligned} K[\Psi] &\ge \left(\frac{(r_1-1)^4}{4\cdot 3^2\cdot 32^2}\right)^{1/3}(r_1-1)^{1/3}(r_2-1)^{1/3} \\ &\ge \left(\frac{5^4}{4\cdot 3^2\cdot 32^2}\right)^{1/3}(r_1-1)^{1/3}(r_2-1)^{1/3} \ge \frac{1}{4}(r_1-1)^{1/3}(r_2-1)^{1/3}. \end{aligned}$$

Q.E.D.

Conclusion of the proof of Claim 1.4.1. We may assume that $|e_1| \ge |e_2|$, because otherwise one can switch the two. With the aid of a linear conformal map we can further adjust e_1, e_2 to $e_1 := 1$ and $e_2 := l\cos\theta + l\sin\theta\sqrt{-1}$, where $l \in (0,1]$. For "large" θ, namely when

$$\sin^2\theta \ge (r_1-1)^2/(32K[\Psi]),$$

we are done by Fact 1.4.3. Assume then that θ is "small", that is

$$\sin^2\theta \le (r_1-1)^2/(32K[\Psi]).$$

Apply first Fact 1.4.1 with $i=1$ and $j=2$ to see that

$$16K[\Psi] \ge 8(1+\sin^2\theta\, l^2)K[\Psi] \ge (r_1-1)^2,$$

and so $\sin^2\theta \leq 1/2$. Thus, using $r_1 \geq 6$, we verify that

$$r_1|e_1||\cos\theta| - |e_2| \geq 6\sqrt{1-\sin^2\theta} - l \geq 6\sqrt{1/2} - 1 > 0,$$

and application of Fact 1.4.1, with $i = 2$, $j = 1$, yields

$$\begin{aligned} 8(l^2+\sin^2\theta)K[\Psi] &\geq (r_2-1)^2l^2 + (r_1-1)^2\sin^2\theta + (r_1|\cos\theta| - l)^2 \\ &\geq (r_2-1)^2l^2 + (r_1-1)^2 - 2r_1|\cos\theta| \qquad (1.19) \\ &\geq (r_2-1)^2l^2 + (r_1-1)^2/2, \end{aligned}$$

where for the last inequality we used $r_1 \geq 6$ again. Thus, dropping $(r_2-1)^2l^2$ and isolating l^2, we get

$$l^2 \geq (r_1-1)^2/(2\cdot 8K[\Psi]) - \sin^2\theta \geq (r_1-1)^2/(32K[\Psi]) \geq \sin^2\theta,$$

and, applying (1.19) again, we derive

$$\begin{aligned} K[\psi] &\geq (r_2-1)^2l^2/(16l^2) + (r_1-1)^2/(32l^2) \\ &\geq (r_2-1)^2/16 + (r_1-1)^2/32 \\ &\geq \frac{1}{16}(r_1-1)(r_2-1) \\ &\geq \frac{1}{4}(r_1-1)^{1/3}(r_2-1)^{1/3}. \end{aligned}$$

Q.E.D.

Chapter 2

Topological entropy at the asymptotic average

Our ultimate goal in this work is to discuss topological entropy relative to a vector in the rotation set. A large part of our discussion naturally fits in a more general context, so we stray from the main course in this chapter. In the first section we define *topological entropy at the asymptotic average* for a map of a compact metrizable space with a continuous observable. We prove the corresponding *variational principle* and use it for a rather weak version of the *thermodynamical formalism*, which holds in this generality. Section 2.2 deals with subshifts of finite type. Invoking the standard *transfer matrix* trick, we give a real-analytic formula for the topological entropy at the asymptotic average. More importantly, we develop a technique that enables us to construct, for any given asymptotic average, compact invariant sets "almost realizing" the topological entropy at this asymptotic average.

2.1 Relative variational principle and thermodynamical formalism

Consider a compact metrizable topological space X with metric d and a dynamical system generated by a homeomorphism $f : X \to X$. Moreover, assume that a continuous function (called an *observable*) $\phi : X \to \mathbf{R}^d$ is given. Our primary example will be a torus homeomorphism isotopic to the identity with the displacement function of one of its lifts as the observable. The rotation set becomes then a specialization of *the set of asymptotic averages* $\rho(f,\phi)$ defined by

$$\rho(f,\phi) := \bigcap_{m>0} \mathrm{cl}\left(\bigcup_{n>m} \{(1/n)S_n(f,\phi)(x) : x \in X\}\right) \subset \mathbf{R}^d, \qquad (2.1)$$

where $S_n(f,\phi) := \phi(x) + ... + \phi(f^{n-1}(x))$ is the Birkhoff sum (cf. [Zie95]).

The set $\rho(f,\phi)$ is compact. Abstracting the proof of connectedness of the rotation set in [MZ89], one can see that $\rho(f,\phi)$ is connected for connected X.

The complexity of the dynamics detected by an asymptotic average in $\rho(f,\phi)$ may be measured by the associated topological entropy. In defining this entropy one may follow a few paths leading to different invariants which nevertheless coincide for the important examples of our primary concern (e.g. for pseudo-Anosov homeomorphisms). Let us start with the most elementary approach building directly on Bowen's definition of topological entropy (see [Wal82]).

Given $S \subset X$, we say that S is (ϵ, n)-*separated* if, for every two different $x, y \in S$, we have $d(f^k(x), f^k(y)) \geq \epsilon$, for some $k \in \{0, 1, ..., n-1\}$. For a set

$E \subset \mathbf{R}^d$ and $\epsilon > 0$, we define

$$s^E(f, \phi, \epsilon, n) := \max\{\#S : S \subset X \ (\epsilon, n)\text{-separated}, \forall_{x \in S} \ (1/n)S_n(f, \phi)(x) \in E\},$$

and, under the convention that $\ln 0 = -\infty$, we set

$$h^E(f, \phi, \epsilon) := \limsup_{n \to \infty} \frac{1}{n} \ln s^E(f, \phi, \epsilon, n) \in \{-\infty\} \cup [0, \infty].$$

The *topological entropy associated with E* is then

$$h^E_{\text{top}}(f, \phi) := \lim_{\epsilon \to 0} h^E(f, \phi, \epsilon).$$

Notice that $h^E(f, \phi, \epsilon)$, as a function of ϵ, is decreasing, so the limit actually exists. Also, as suggested by the notation, $h^E_{\text{top}}(f, \phi)$ is a topological invariant; it does not depend on the metric, nor is altered by replacing the pair (f, ϕ) with a $(h^{-1} \circ f \circ h, \phi \circ h)$, where h is a homeomorphism of X (cf. Proposition 2.1.2). Actually, one can easily make h^E_{top} manifestly topologically invariant by rephrasing its definition in terms of open covers, as it is done for the usual topological entropy (see [Wal82]).

Also, $h^E_{\text{top}}(f, \phi, \epsilon)$ is monotonic in E and, for two sets $E, F \in \mathbf{R}^d$, we have

$$h^{E \cup F}_{\text{top}}(f, \phi) = \max\left\{h^E_{\text{top}}(f, \phi), h^F_{\text{top}}(f, \phi)\right\}. \tag{2.2}$$

Thus a bulk of information carried by $h^E_{\text{top}}(f, \phi, \epsilon)$ is still recoverable from *topological entropy at the asymptotic average*, which is defined as the following function of $v \in \mathbf{R}^d$

$$h^{(v)}_{\text{top}}(f, \phi) := \lim_{r \to 0} h^{B_r(v)}_{\text{top}}(f, \phi),$$

where $B_r(v)$ is the open ball of radius r centered at v.

Actually, if we regularize $E \mapsto h^E(f,\phi)$ by setting

$$h^{(E)}_{\text{top}}(f,\phi) := \lim_{r\to 0} h^{B_r(E)}_{\text{top}}(f,\phi),$$

then $v \mapsto h^{(v)}_{\text{top}}(f,\phi)$ is characterized by the relation $h^{(E)}_{\text{top}}(f,\phi) = \max\{h^{(v)}_{\text{top}}(f,\phi) : v \in E\}$, where E is a closed subset of $\mathbf{R}^d$ (cf. Lemma A.0.1).

The following is a summary of basic properties of $h^{(v)}_{\text{top}}(f,\phi)$ that can be easily derived from the definition — see also the appendix. In fact, part (ii) below expresses the compatibility of $h^{(v)}_{\text{top}}(f,\phi)$ and $\rho(f,\phi)$ that was one of the axioms lying at the origin of $h^{(v)}_{\text{top}}(f,\phi)$.

Proposition 2.1.1 (basics) *With the above definitions, we have*

(i) $-\infty \le h^{(v)}_{\text{top}}(f,\phi) \le h_{\text{top}}(f)$, $v \in \mathbf{R}^d$;

(ii) $\rho(f,\phi) = \{v \in \mathbf{R}^d : h^{(v)}_{\text{top}}(f,\phi) \ge 0\}$;

(iii) there exists $v_* \in \mathbf{R}^d$ *such that* $h^{(v_*)}_{\text{top}}(f,\phi) = h_{\text{top}}(f)$;

(iv) $h^{(v)}_{\text{top}}(f,\phi)$ *is upper semi-continuous in* v, *i.e.*

$$\limsup_{x\to v} h^{(v)}_{\text{top}}(f,\phi) = h^{(v)}_{\text{top}}(f,\phi).$$

Proof of Proposition 2.1.1. Parts (i), (ii), (iv) require no comment. To see (iii), note that (2.2) implies that, for any finite cover of $\rho(f,\phi)$, there exists an element E of the cover with $h^{E}_{\text{top}}(f,\phi) = h_{\text{top}}(f)$. To nail down v_* simply consider finer and finer covers (cf. Lemma A.0.1). Q.E.D.

There are many properties of $h^{(v)}_{\text{top}}(f,\phi)$ that are directly analogous to those of the topological entropy, and the proofs are typically just simple modifications of the standard ones. We mention here only those properties that we will need in the further development.

Proposition 2.1.2 (Quotient Rule) *Let (X, d) and (Y, e) be compact metric spaces and $\phi : X \to \mathbf{R}^d$. If $f : X \to X$ is a factor of $g : Y \to Y$ via $h : Y \to X$ (i.e. h is continuous surjective and $f \circ h = h \circ g$) then, for any $E \subset \mathbf{R}^d$, $h^E_{\text{top}}(f, \phi) \le h^E_{\text{top}}(g, \phi \circ h)$. Consequently, $h^{(v)}_{\text{top}}(f, \phi) \le h^{(v)}_{\text{top}}(g, \phi \circ h)$, $v \in \mathbf{R}^d$.*

Proof. Fix an arbitrary $\epsilon > 0$. By uniform continuity of h, there is $\delta > 0$ such that $e(y, y') < \delta$ implies $d(h(y), h(y')) < \epsilon$ for all $y, y' \in Y$. Now, given $S \subset X$ which is (ϵ, n)-separated with $S_n(f, \phi)(x) \in E$ for $x \in S$, form $S' \subset Y$ by choosing a point from $h^{-1}(x)$ for each $x \in S$. This set is (δ, n)-separated for g and, for $y \in S'$, $S_n(g, \phi \circ h)(y) = S_n(f, \phi)(h(y)) \in E$. Since $\#S' = \#S$, we get $h^E(f, \phi, \epsilon) \le h^E(g, \phi \circ h, \delta) \le h^E_{\text{top}}(g, \phi \circ h)$, which finishes the proof because of arbitrariness of ϵ. Q.E.D.

Proposition 2.1.3 (Power Rule) *For any $m \in \mathbf{N}$ and $E \subset \mathbf{R}^d$, we have*

$$h^E_{\text{top}}(f^m, (1/m)S_m(\phi)) = m \cdot h^E_{\text{top}}(f, \phi).$$

In particular, for $v \in \mathbf{R}^d$,

$$h^{(v)}_{\text{top}}(f^m, (1/m)S_m(\phi)) = m \cdot h^{(v)}_{\text{top}}(f, \phi).$$

Proof. This is a straightforward upgrade of the corresponding argument in [Wal82]. We trace it for the sake of completeness. Fix $E \subset \mathbf{R}^d$. We will prove that $m \cdot h^E_{\text{top}}(f, \phi) = h^E_{\text{top}}(f^m, (1/m)S_m(f, \phi))$. If S is an (ϵ, n)-separated set for f^m, then it is (ϵ, nm)-separated for f. Moreover,

$$(1/n)S_n(f^m, (1/m)S_m(f, \phi))(x) = (1/(mn))S_{mn}(f, \phi)(x), \qquad (2.3)$$

for any $x \in X$. It follows that $h^E_{\text{top}}(f^m, (1/m)S_m(\phi)) \leq m \cdot h^E_{\text{top}}(f, \phi)$. On the other hand, for $\epsilon > 0$ there is $\delta > 0$ such that $d(x, y) \leq \delta$ forces $d(f^i(x), f^i(y)) < \epsilon$, for $0 \leq i \leq m$. Let S be a (δ, n)-separated set for f^m maximal among those that satisfy $(1/n)S_n(f^m, (1/m)S_m(f, \phi))(x) \in E$ for all $x \in S$. By the maximality and (2.3), for any $y \in X$ with $(1/(mn))S_{mn}(f, \phi)(y) \in E$, there is $x \in S$ so that $d(f^{km}(x), f^{km}(y)) < \delta$, $0 \leq k \leq n-1$, and so $d(f^i(x), f^i(y)) < \epsilon$, $0 \leq i \leq mn-1$. Consequently, if S' is an arbitrary $(2\epsilon, mn)$-separated set such that $(1/(mn))S_{mn}(f, \phi)(x) \in E$, $x \in S'$, then it has at most $\#S$ elements. By arbitrariness of S', one gets $h^E(f^m, (1/m)S_m(\phi), \delta) \geq m \cdot h^E(f, \phi, 2\epsilon)$. It follows that $h^E_{\text{top}}(f^m, (1/m)S_m(\phi)) \geq m \cdot h^E_{\text{top}}(f, \phi)$. Q.E.D.

The next theorem, in the spirit of *thermodynamical formalism* ([Rue78]), is a powerful tool for calculating $h^{(v)}_{\text{top}}(f, \phi)$. The central role is played by the restriction of the topological pressure functional $\mathcal{P}$ ([Wal82]) to the linear space of functions $\{\langle t, \phi(\cdot)\rangle\}_{t \in \mathbf{R}^d}$, where $\langle \cdot, \cdot \rangle$ is a scalar product in $\mathbf{R}^d$. Namely, for every $t \in \mathbf{R}^d$, we consider

$$p^{(t)}_{\text{top}}(f, \phi) :=$$

$$\lim_{\epsilon \to 0} \limsup_{n \to \infty} \quad (1/n) \ln \max \left\{ \textstyle\sum_{x \in S} \exp\langle S_n(f, \phi)(x), t\rangle : \ S \ (\epsilon, n)\text{-separated} \right\}.$$

Note that $p^{(0)}_{\text{top}}(f, \phi) = h_{\text{top}}(f)$ and that $h_{\text{top}}(f) = +\infty$ implies $p^{(t)}_{\text{top}}(f, \phi) = +\infty$ for all $t \in \mathbf{R}^d$. However, if $h_{\text{top}}(f) < +\infty$, then $p^{(t)}_{\text{top}}(f, \phi) < +\infty$ for all $t \in \mathbf{R}^d$; this is the situation we are mainly interested in.

Since $\mathcal{P}$ is a convex functional, $p^{(t)}_{\text{top}}(f, \phi)$ is a convex function of t. Convexity will play a significant role in our considerations, so let us fix some definitions before we go further. Our main reference is [Roc72]. For a convex subset $D \subset \mathbf{R}^d$, denote by $\mathrm{aff}(D)$ its *affine hull*, i.e. the smallest affine subspace of $\mathbf{R}^d$ containing it. The interior of D with respect to the natural topology on $\mathrm{aff}(D)$ will be called *relative interior* and denoted $\mathrm{ri}(D)$. If $g : \mathbf{R}^d \to (-\infty, +\infty]$ is an arbitrary function, $\mathrm{epi}(g) := \{(x, y) : y \geq g(x)\}$ is the *epigraph of* g, and the set $\mathrm{dom}(g) := \{x \in \mathbf{R}^d : g(x) < +\infty\}$ is called the *effective domain of* g. The *closure of* g is the function $\mathrm{cl}(g)$ defined by $\mathrm{cl}(g)(y) := \liminf_{x \to y} g(x)$; it is the greatest lower semi-continuous function majorized by g, and the epigraph of $\mathrm{cl}(g)$ is the closure of the epigraph of g. The *convex conjugate* $\hat{g} : \mathbf{R}^d \to [-\infty, +\infty]$ is defined by $\hat{g}(y) := \sup\{\langle x, y\rangle - g(x) : \ x \in \mathbf{R}^d\}$. The convex conjugate is a lower semi-continuous convex function. If g is a convex function, we have $\hat{g}(y) = \sup\{\langle x, y\rangle - g(x) : \ x \in \mathrm{ri}(\mathrm{dom}(g))\}$ (see [Roc72] p.104). Convex conjugacy has an involutive property; namely, $(\hat{g})\hat{} = \mathrm{cl}(\text{convex hull}(g))$, where *convex hull* (g) is the pointwise supremum of all convex functions bounding g from below. The epigraph of convex $\mathrm{hull}(g)$ does not have to be equal to the convex hull of that of g. However, this is the case when g is lower semi-continuous and has bounded effective domain (see Proposition A.0.3 and its proof). Then also $h :=$ convex $\mathrm{hull}(g)$ is lower semi-continuous and h coincides with g at the extremal points of the epigraph of h. In particular, $g(x) = h(x)$ at the ordinates x of those extremal points that lie over $\mathrm{ri}(\mathrm{dom}(h))$ — those values of x are referred to as *points of strict convexity of* h. Finally, if we assume that $h_{\mathrm{top}}(f) < \infty$, then all the convex functions in the following discussion are *proper*, i.e. they take values in $(-\infty, \infty]$.

Theorem 2.1.1 (thermodynamical formalism) *If $h_{\text{top}}(f) < \infty$, then the convex hull of $v \mapsto -h^{(v)}_{\text{top}}(f,\phi)$ is the convex conjugate $v \mapsto \hat{p}^{(v)}_{\text{top}}(f,\phi)$ of $t \mapsto p^{(t)}_{\text{top}}(f,\phi)$. In particular,*

$$-h^{(v)}_{\text{top}}(f,\phi) \geq \sup_{t \in \mathbf{R}^d} \left(\langle t, v \rangle - p^{(t)}_{\text{top}}(f,\phi) \right),$$

and the equality holds at all points v where the right side is strictly convex.

Using (ii) of Proposition 2.1.1, we get the following straightforward corollary.

Corollary 2.1.1 *If $h_{\text{top}}(f) < \infty$, then the convex hull of the set of asymptotic averages $\rho(f,\phi)$ is equal to the essential domain $\text{dom}(\hat{p}^{(v)}_{\text{top}}(f,\phi))$ of the convex conjugate of $p^{(t)}_{\text{top}}(f,\phi)$.*

Theorem 2.1.1 may be viewed as a *large deviations* type result (see [Ell85]). It was inspired by Theorem II 6.1 in [Ell85]. In the appendix we give a broader exposition together with an elementary proof. In this chapter we show Theorem 2.1.1 as a byproduct of connecting $h^{(v)}_{\text{top}}(f,\phi)$ with the main body of ergodic theory. The major link in the connection is the following improvement of the variational principle for topological entropy ([Wal82]).

Denote by $\mathcal{M}(X,f)$ the set of all Borel probability measures invariant under f and by $\mathcal{E}(X,f)$ the set of all those measures in $\mathcal{M}(X,f)$ that are ergodic.

Theorem 2.1.2 (Relative Variational Principle) *We have:*

(i) for any closed convex set $V \subset \mathbf{R}^d$,

$$h^{V}_{\text{top}}(f,\phi) \leq \sup \left\{ h_\mu(f) : \mu \in \mathcal{M}(X,f),\ \int \phi \, d\mu \in V \right\};$$

(ii) for any open set $U \subset \mathbf{R}^d$,

$$h^U_{\text{top}}(f,\phi) \geq \sup\left\{h_\mu(f) : \mu \in \mathcal{E}(X,f),\ \int \phi \, d\mu \in U\right\}.$$

Thus, if we define (using the convention that $\sup \emptyset = -\infty$)

$$h^v_{\text{erg}}(f,\phi) := \sup\left\{h_\mu(f) : \mu \in \mathcal{E}(X,f),\ \int \phi \, d\mu = v\right\}, \tag{2.4}$$

and

$$h^v_{\text{meas}}(f,\phi) := \sup\left\{h_\mu(f) : \mu \in \mathcal{M}(X,f),\ \int \phi \, d\mu = v\right\}, \tag{2.5}$$

then we have the following corollary.

Corollary 2.1.2 *For $v \in \mathbf{R}^d$, we have*

$$\limsup_{x \to v} h^x_{\text{erg}}(f,\phi) \leq h^{(v)}_{\text{top}}(f,\phi) \leq \limsup_{x \to v} h^x_{\text{meas}}(f,\phi).$$

In particular,

$$\text{cl}\left(-h^v_{\text{meas}}(f,\phi)\right) := -\limsup_{x \to v} h^x_{\text{meas}}(f,\phi) = \text{convex hull}\left(-h^{(v)}_{\text{top}}(f,\phi)\right).$$

Proof of Corollary 2.1.2. The inequalities follow immediately from the definitions. To justify the second assertion, note that, since the metric entropy $h_\mu(f)$ is affine as the function of μ, $-h^v_{\text{meas}}(f,\phi)$ is a convex function of v. From the ergodic decomposition theorem, $-h^v_{\text{meas}}(f,\phi)$ is the convex hull of $-h^v_{\text{erg}}(f,\phi)$. We are done because $h^{(v)}_{\text{top}}(f,\phi)$ is pinched between the two. Q.E.D.

Also, by its convexity, $-h^v_{\text{meas}}(f,\phi)$ is continuous at all points v in the relative interior of its effective domain, $\text{ri}(\rho(f,\phi))$; so, for those points, the rightmost "lim sup" in Corollary 2.1.2 is superfluous. Another instance when we do not need that "lim sup" is when $h_\mu(f)$ is upper semi-continuous in μ because then $h^v_{\text{meas}}(f,\phi)$ is upper semi-continuous in v. This happens for expansive f ([Wal82]) and also for any f that is C^∞-smooth ([New89]).

The following example shows that one can indeed have $\limsup_{x\to v} h^x_{\text{erg}}(f,\phi) < h^{(v)}_{\text{top}}(f,\phi)$.

Example. For $(x_i)_{i\in\mathbf{Z}} \in \{1,2,3,4\}^{\mathbf{Z}}$, $(x_k,...,x_l)$ is a *parity block* if $x_k,...,x_l$ are all even or all odd. We call it *maximal* if x_{k-1}, x_{l+1} have the parity opposite to that of x_i, $i = k,...,l$. Consider a subshift (Λ,σ) where

$$\Lambda = \{x \in \{1,2,3,4\}^{\mathbf{Z}} : x \text{ has at most two maximal parity blocks}\}$$

and σ shifts to the left: $(\sigma((x_i)_{i\in\mathbf{Z}}))_i := x_{i+1}$. The space Λ is compact and invariant under σ. Let the observable $\phi : \Lambda \to \mathbf{R}$ be given by $\phi(x) := (-1)^{x_0}$. One can easily verify that $\rho(\sigma,\phi) = [-1,1]$. All ergodic invariant measures are supported on the union of $\Lambda_{even} := \{x \in \Lambda : \ x_i \text{ even for all } i\}$ and $\Lambda_{odd} := \{x \in \Lambda : \ x_i \text{ odd for all } i\}$ because any recurrent point is contained in one of the two sets. Nevertheless, $h^{(0)}_{\text{top}}(\sigma,\phi) \geq \ln 2$. Indeed, for any $n \in 2\mathbf{N}$, consider points of Λ with maximal parity blocks of the form $(..., x_{n/2-1})$ and $(x_{n/2}, ...)$. They all have vanishing $S_n(\sigma,\phi)$, and there are $2 \cdot 2^n$ such points with pairwise different blocks of the first n symbols.

Perhaps a more interesting example is $\Lambda := \{x \in \{1,2,3,4\}^{\mathbf{Z}} :$
x has no two finite maximum parity blocks of equal length$\}$.
We leave analysis of this example to the reader.

Corollary 2.1.3 *If μ_0 is an ergodic equilibrium state for $\langle t_0, \phi\rangle$, that is ([Rue78, Bow75, Wal82])*

$$p_{\rm top}^{(t_0)}(f,\phi) = h_{\mu_0}(f) + \Big\langle t_0, \int \phi \, d\mu_0 \Big\rangle,$$

then, for $v_0 := \int \phi \, d\mu_0$, we have $h_{\rm top}^{(v_0)}(f,\phi) = h_{\rm maes}^{v_0}(f,\phi) = h_{\rm erg}^{v_0}(f,\phi) = h_{\mu_0}(f)$.

In particular, if μ_* is an ergodic measure of maximal entropy, then $v_* := \int \phi \, d\mu_*$ satisfies $h_{\rm top}^{(v_*)}(f,\phi) = h_{\rm top}(f)$ realizing thus the maximum of $h_{\rm top}^{(v_*)}(f,\phi)$. Such a measure exists for C^∞ maps due to the result by Newhouse ([New89]).

Proof of Theorem 2.1.1 from Theorem 2.1.2. From the variational principle for topological pressure ([Wal82]) we have

$$\begin{aligned} p_{\rm top}^{(t)}(f,\phi) &= \sup\Big\{ h_\mu(f) + \int \langle \phi, t\rangle \, d\mu : \ \mu \in \mathcal{M}(X,f)\Big\} = \\ &\sup\left\{\langle t, v\rangle + h_{\rm meas}^v(f,\phi) : \ v \in {\rm dom}\,(-h_{\rm meas}^v(f,\phi))\right\}. \end{aligned}$$

Thus $p_{\rm top}^{(t)}(f,\phi) = (-h_{\rm meas}^v(f,\phi))\hat{}$, so, by the involutive property of "ˆ" and by Corollary 2.1.2, we have

$$(p_{\rm top}^{(t)}(f,\phi))\hat{} = (-h_{\rm meas}^v(f,\phi))\hat{}\hat{} = \ {\rm cl}(-h_{\rm meas}^v(f,\phi)) = \text{convex hull}\left(-h_{\rm top}^{(v)}(f,\phi)\right).$$

(We skipped "cl" before convex hull $\left(-h_{\rm top}^{(v)}(f,\phi)\right)$ because this function is already lower semi-continuous — see Proposition A.0.3.) Q.E.D.

Proof of Corollary 2.1.3. In the previous proof we got

$$\hat{p}_{\rm top}^{(v)}(f,\phi) = \ {\rm cl}(-h_{\rm meas}^v(f,\phi)), \ v \in \mathbf{R}^d. \tag{2.6}$$

Thus, combining trivial inequalities, we can write

$$\operatorname{cl}\left(-h^{v_0}_{\mathrm{maes}}(f,\phi)\right) \le -h^{v_0}_{\mathrm{maes}}(f,\phi) \le -h^{v_0}_{\mathrm{erg}}(f,\phi) \le -h_{\mu_0}(f) =$$
$$\left\langle t_0, \int \phi\, d\mu_0 \right\rangle - p^{(t_0)}_{\mathrm{top}}(f,\phi) \le \sup_{t\in\mathbf{R}^d}\left\{\langle t, v\rangle - p^{(t)}_{\mathrm{top}}(f,\phi)\right\} = \hat{p}^{(v)}_{\mathrm{top}}(f,\phi) =$$
$$\operatorname{cl}\left(-h^{v_0}_{\mathrm{maes}}(f,\phi)\right).$$

In this way, $\operatorname{cl}\left(-h^{v_0}_{\mathrm{maes}}(f,\phi)\right) = -h^{v_0}_{\mathrm{erg}}(f,\phi)$. By Corollary 2.1.2, $-h^{(v)}_{\mathrm{top}}(f,\phi)$ is pinched between the two, which finishes the proof. Q.E.D.

Remark. If $p^{(t)}_{\mathrm{top}}(f,\phi)$ is differentiable at t_0, then Corollary 2.1.3 is true without the hypothesis that μ_0 is ergodic. In fact, almost all measures ν in the ergodic decomposition of μ_0 satisfy then $h_\nu(f) = h_{\mu_0}(f)$ and $\int \phi\, d\nu = \int \phi\, d\mu_0$. Indeed, affine functions $\xi_\nu : t \mapsto h_\nu(f) + \langle t, \int \phi\, d\nu\rangle$ are all majorized by $p^{(t)}_{\mathrm{top}}(f,\phi)$ and, integrated with respect to ν, yield ξ_{μ_0} which is tangent to $p^{(t)}_{\mathrm{top}}(f,\phi)$ at t_0. Since there is a unique tangent at t_0, almost all ξ_ν's must be equal to ξ_{μ_0} which implies our claim.

The following fact will be used numerous times ahead.

Proposition 2.1.4 (Quotient Rule) *Let (X,d) and (Y,e) be compact metric spaces and $\phi : X \to \mathbf{R}^d$ be continuous. If $f : X \to X$ is a factor of $g : Y \to Y$ via $h : Y \to X$ (i.e. h is continuous surjective and $f\circ h = h \circ g$) then, $h^v_{\mathrm{erg}}(f,\phi) \le h^v_{\mathrm{erg}}(g, \phi\circ h)$, $v \in \mathbf{R}^d$.*

Proof. Let $\mathcal{M}(Y)$, $\mathcal{M}(X)$ be the spaces of the Borel probability measures on Y and X respectively, both equipped with the weak*-topology. The map h induces the *push forward* operator on measures, $h_* : \mathcal{M}(Y) \to \mathcal{M}(X)$. Since, for $\nu \in \mathcal{M}(Y,g)$, $h_{h_*(\nu)}(f) \le h_\nu(G)$, it is enough to prove that h_* maps $\mathcal{E}(Y,g)$ surjectively onto $\mathcal{E}(X,f)$. This can be derived (via ergodic decomposition) from the lemma 8.3 in [Man87]. We give a more elementary argument due to Milnor. Fix an arbitrary $\mu \in \mathcal{E}(X,f)$. Choose measures $\mu_n \in \mathcal{M}(X)$ supported on finite sets and converging to μ. By surjectivity of h, each atom of μ_n has a preimage in Y; and, by taking those preimages with the obvious weights, one easily gets $\nu_n \in \mathcal{M}(Y)$ such that $h_*(\nu_n) = \mu_n$. Passing perhaps to a subsequence, we can assume that ν_n's converge to a measure $\nu \in \mathcal{M}(Y)$. Since h_* is continuous, $h_*(\nu) = \mu$. To make ν invariant under g, replace ν with any limit point of the sequence of time averages $\left((1/n)\sum_{i=1}^{n-1} g_*^i(\nu)\right)_{n\in\mathbf{N}}$ — those averaged measures also map under h_* to μ. Thus we got $\nu \in \mathcal{M}(Y,g)$ such that $h_*(\nu) = \mu$. This shows that a compact convex set $C := h_*^{-1}(\mu) \cap \mathcal{M}(Y,g)$ is not empty; therefore, it has an extremal point $\eta \in \mathcal{M}(Y,g)$. The measure η must be ergodic. Indeed, if it were a nontrivial convex combination of two other measures $\eta_1, \eta_2 \in \mathcal{M}(Y,g)$, then, by extremality of η in C, $h_*(\eta_i) \neq \mu$, $i = 1,2$. This contradicts ergodicity of μ because $\mu = h_*(\nu)$ is a convex combination of $h_*(\eta_1)$ and $h_*(\eta_2)$. Q.E.D.

Proof of Theorem 2.1.2. We build on Misiurewicz's proof of the variational principle for topological entropy as found in [Wal82]. Basically, we need to augment it with arguments that keep track of the averages of the observable. This is trivial for the proof of (i). Then the natural measures (i.e. limits of those equidistributed on the appropriate (ϵ, n)-separated sets) have the right averages, and we just quote [Wal82] claiming that their entropy is as required.

Part (ii) is more involved; we are forced to interweave the proof from [Wal82] with several arguments including application of Egoroff's and Birkhoff's theorems.

Proof of (i). We will show that, for any $\gamma > 0$, there is $\mu \in \mathcal{M}(X, f)$ with $\int \phi \, d\mu \in V$ and $h_\mu(f) > h^V_{\text{top}}(f, \phi) - \gamma$. Choose $\epsilon > 0$ so that $h^V(f, \phi, \epsilon) > h^V_{\text{top}}(f, \phi) - \gamma$. From the definition of $h^V(f, \phi, \epsilon)$, there exists a sequence $n_k \in \mathbf{N}$, $n_k \to \infty$ and (ϵ, n_k)-separated sets E_k such that $\lim_{k\to\infty}(1/n_k) \ln \#E_k = h^V(f, \phi, \epsilon)$ and $(1/n_k)S_{n_k}(f, \phi)(x) \in V$, for $x \in E_k$. Consider the uniform atomic measures on E_k, call them δ_{E_k}, and set $\mu_k := (1/n_k) \sum_{j=0}^{n_k-1} \delta_{E_k} \circ f^{-j}$. Note that, by convexity and closedness of V, $\int \phi \, d\mu_k = \int (1/n_k)S_{n_k}(f, \phi) \, d\delta_{E_k} \in V$. Passing perhaps to the subsequence of (n_k), we can assure that μ_k's weak*-converge to some probability measure μ. Clearly $\int \phi \, d\mu \in V$. More importantly, Misiurewicz proves that $h_\mu(f) \geq \lim_{k\to\infty} 1/n_k \ln \#E_k$ (see part (2) of the proof of Th.8.6., p. 189 in [Wal82]), so $h_\mu(f) \geq h^V(f, \phi, \epsilon) > h^V_{\text{top}}(f, \phi) - \gamma$. Q.E.D.

Proof of (ii). We will show that: if $h^U_{\text{top}}(f, \phi) > 2$ and $\mu \in \mathcal{M}(X, f)$ is such that

$$\lim_{n\to\infty} (1/n)S_n(f, \phi)(x) = \overline{\phi} := \int \phi \, d\mu \in U, \tag{2.7}$$

for μ-almost all x in X, then $h_\mu(f) \leq h^U_{\text{top}}(f, \phi) + \ln 2 + 1$. Clearly, (2.7) is satisfied by ergodic μ with $\int \phi \, d\mu \in U$; however, unlike ergodicity, (2.7) is invariant under replacing f and ϕ by the iterate f^m and by $(1/m)S_m(f, \phi)$ respectively. By considering arbitrarily high iterates and using the power rule (Proposition 2.1.3), we conclude that $h_\mu(f) \leq h^U_{\text{top}}(f, \phi)$.

Let $\mathcal{A} = A_1, ..., A_k$ be an arbitrary finite Borel partition of X. Choose $\epsilon > 0$ so that $\epsilon < 1/(k \ln k)$. Using Egoroff's theorem, we may find compact sets $B_j \subset A_j$, with $\mu(A_j \backslash B_j) < \epsilon$, $1 \leq j \leq k$, such that $(1/n)S_n(f, \phi)$ converges uniformly to $\overline{\phi}$ on every B_j. Now, because of this uniform convergence and openness of U and since $\overline{\phi} \in U$, there is $0 < d < 1$ such that

$$(1/m)S_m(f, \phi)(x) + 2d \cdot \{\phi(x) : x \in X\} \subset U, \tag{2.8}$$

for any $x \in \bigcup_{j=1}^{k} B_j$, and sufficiently large $m \in \mathbf{N}$.

Put $B_0 := X \setminus \bigcup\{B_j : 1 \leq j \leq k\}$ and introduce a partition $\mathcal{B} := \{B_0, B_1, ..., B_k\}$. Now, for $1 \leq j \leq k$, we have $B_j \subset A_j$, so the entropy $H_\mu(\mathcal{A}/B_j)$ of $\mathcal{A}$ restricted to B_j vanishes for those j's. Also, by the trivial estimate (see Corollary 4.2.1 p. 80 in [Wal82]), $H_\mu(\mathcal{A}/B_0) \leq \ln k$. Thus we get conditional entropy $H_\mu(\mathcal{A}/\mathcal{B}) = \sum_{0 \leq j \leq k} \mu(B_j) H_\mu(\mathcal{A}/B_j) \leq \mu(B_0) \ln k \leq \epsilon k \ln k < 1$. Consequently, (by (iv) of Th.4.12, p. 89 in [Wal82]) we have

$$h_\mu(f, \mathcal{A}) \leq h_\mu(f, \mathcal{B}) + H_\mu(\mathcal{A}/\mathcal{B}) < h_\mu(f, \mathcal{B}) + 1. \tag{2.9}$$

For each n, let us split $\mathcal{B}^n := \bigvee_{j=0}^{n-1} f^{-j}\mathcal{B}$ into $\mathcal{B}_1^n := \{B \in \mathcal{B}^n : B \subset B_0 \cap ... \cap f^{-\lfloor dn \rfloor} B_0\}$ and $\mathcal{B}_2^n := \mathcal{B}^n \setminus \mathcal{B}_1^n$. Note that $\#\mathcal{B}_1^n \leq (\#\mathcal{B})^{n - \lfloor dn \rfloor}$, thus, for sufficiently large n, we have

$$\#\mathcal{B}_2^n \geq \#\mathcal{B}^n - (\#\mathcal{B})^{n - \lfloor dn \rfloor} \geq (1/2)\#\mathcal{B}^n. \tag{2.10}$$

Put $K_n := \{x \in X : (1/m)S_m(f, \phi)(x) \in U \text{ for all } m \geq n\}$. We claim that the union $\bigcup \mathcal{B}_2^n$ is contained in K_n for sufficiently large n. Indeed, if $x \in A \in \mathcal{B}_2^n$, then there is s, $0 \leq s \leq \lfloor dn \rfloor$, such that $f^s(x) \in B_j$ with $j \neq 0$.

Now, a simple estimate gives

$$\begin{aligned}&\|(1/m)S_m(f,\phi)(x) - (1/m)S_m(f,\phi)(f^s(x))\| \\ \leq\ & 2\lfloor dn \rfloor \sup_{z\in X}\|\phi(z)\|/m \leq 2d \sup_{z\in X}\|\phi(z)\|,\end{aligned}$$

so, by the choice of d (see (2.8)), we have $(1/m)S_m(f,\phi)(x) \in U$.

Adopting Misiurewicz's trick, consider $\mathcal{C} := \{B_0 \cup B_1, ..., B_0 \cup B_k\}$. This is an open cover of K_n (also X). Write $N(\mathcal{C}^n, K_n)$ for the minimal cardinality of a subfamily of $\mathcal{C}^n := \bigvee_{j=0}^{n-1} f^{-j}\mathcal{C}$ that covers K_n, and let $\mathcal{C}_n$ be such a subfamily with $\#\mathcal{C}_n = N(\mathcal{C}^n, K_n)$. By our claim, each $B \in \mathcal{B}_2^n$ lies in K_n; therefore, there exists $C \in \mathcal{C}_n$ containing B. It is also clear from the definition of $\mathcal{C}$ that at most 2^n elements of $\mathcal{B}_2^n$ can fit into any fixed element of $\mathcal{C}_n$. We conclude that

$$\#\mathcal{B}_2^n \leq N(\mathcal{C}^n, K_n)2^n. \tag{2.11}$$

By Th. 7.7 in [Wal82], if δ is the Lebesgue number of $\mathcal{C}$, then $N(\mathcal{C}^n, K_n) \leq s(\delta/2, n, K_n)$, (where $s(\delta/2, n, K_n)$ is the maximal cardinality of an $(\delta/2, n)$-separated subset of K_n). Note that every $(\delta/2, n)$-separated subset of K_n has the averages $(1/n)S_n(f,\phi)(x)$ in U, for large n. Thus, for those n, we get $N(\mathcal{C}^n, K_n) \leq s^U(f, \phi, \delta/2, n)$.

Now, combining (2.10), (2.11) and the last inequality, we get, for large n,

$$H_\mu(\mathcal{B}^n) \leq \ln \#\mathcal{B}^n \leq \ln(2\#\mathcal{B}_2^n) \leq \ln\left(2^{n+1}s^U(f,\phi,\delta/2,n)\right).$$

Passing to the limit, first with $n \to \infty$ then with $\delta \to 0$, we obtain

$$h_\mu(f, \mathcal{B}) \leq \ln 2 + h_{\text{top}}^U(f, \phi).$$

From (2.9), $h_\mu(f,\mathcal{A}) \le \ln 2 + h^U_{\text{top}}(f) + 1$, and by arbitrariness of $\mathcal{A}$ we get $h_\mu(f) \le \ln 2 + h^U_{\text{top}}(f,\phi) + 1$, as promised. Q.E.D.

2.2 Topological entropy at the asymptotic average for subshifts of finite type

As an example of utilizing Theorem 2.1.1 to calculate $h^{(v)}_{\text{top}}(f,\phi)$, we consider the case of a transitive subshift of finite type with locally constant observable. This setting is exactly what we will need to analyze the case of pseudo-Anosov maps in the next chapter.

In the first subsection, using standard techniques, we identify $p^{(t)}_{\text{top}}(f,\phi)$ with the leading positive eigenvalue of the appropriate *transfer matrix*. This gives an explicit real analytic formula for $h^{(v)}_{\text{top}}(f,\phi)$. We also note that $h^{(v)}_{\text{top}}(f,\phi)$ is realized as the metric entropy for the corresponding *Gibbs state*.

The second subsection is devoted to showing that, apart from having the associated ergodic *Gibbs state*, each $v \in \text{ri}(\rho(f,\phi))$ is the *asymptotic average* for a compact invariant set with the topological entropy arbitrarily close to $h^{(v)}_{\text{top}}(f,\phi)$.

In most of our formulations we restrict attention to v in $\text{ri}(\rho(\sigma,\psi))$ since ultimately (in the case of torus homeomorphisms) only such rotation vectors will be considered. However, the case of v in the boundary of $\rho(f,\phi)$ can be reduced to that of $v \in \text{ri}(\rho(\sigma,\psi))$ by passing to a smaller subshift of finite type using the *scaffold construction* (taken from [MT91]). We briefly describe this trick in the last subsection and show how it relates to $h^{(v)}_{\text{top}}(f, \phi)$.

For subshifts of finite type, the entropy $h^{(v)}_{\text{top}}(f, \phi)$ contains much more refined information that what is suggested by its definition. It can actually be used to count the number of periodic orbits with a prescribed average of the observable ([MT91]). This belongs to a topic that parallels the classical theorems on the distribution of the prime numbers via the formalism of Zeta functions. There is a large amount of published material on the subject, see for example [PP90] and the references within.

2.2.1 Transfer matrix, Gibbs states, and explicit formulas

Consider a transitive subshift of finite type $\sigma : \Lambda \to \Lambda$ with a locally constant observable function $\psi : \Lambda \to \mathbf{R}^d$. Without any loss of generality, we can assume that there is an *irreducible* ([Wal82]) binary matrix $A = (a_{i,j})_{1\le i,j\le N}$ and a $\mathbf{R}^d$-valued matrix $(\psi_{i,j})_{1\le i,j\le N}$ such that $\Lambda = \{x \in \{1, ..., N\}^{\mathbf{Z}} :\ a_{x_i,x_{i+1}} = 1,\ i \in \mathbf{Z}\}$ and $\psi(x) = \psi_{x_0,x_1}$ for $x \in \Lambda$. The metric d on Λ, that we shall use, is given by $d((x_k), (y_k)) := 2^{-|k|}$, where k is the closest to 0^{th} position at which x_k and y_k differ.

A convenient representation of this setting is achieved by interpreting points of Λ as bi-infinite paths in the *transition graph* $G \subset \{1, ..., N\}^2$ with vertices $\{1, ..., N\}$ and an edge (i, j) from i to j whenever $a_{i,j} = 1$. By a *path* α in G we understand a map $\alpha : \{k, ..., l-1\} \to G$, where $k, l \in \mathbf{Z}\cup\{-\infty, +\infty\}$, and $\alpha(i + 1)$ emanates from the endpoint vertex of $\alpha(i)$, $k \le i < l - 1$. The number $l-k$ is the length of the path, denoted by $l(\alpha)$. A *loop* is a path with the endpoint of $\alpha(l-1)$ coinciding with the starting point of $\alpha(k)$. The loops visiting any vertex of G at most once are called *elementary*. Clearly, there is a finite number of the elementary loops in G. Each periodic orbit of σ determines uniquely a loop with the length equal to the period. To keep track of the averages of ψ, each edge $(i,j) \in G$ is tagged with the *vector weight* $\psi_{i,j}$. By $\sum_\alpha \psi$ we denote the sum of the weights along the edges of the path α. If α is a loop, we write $\rho(\alpha)$ for the average $(1/l(\alpha)) \sum_\alpha \psi$. Transitivity of σ guarantees that any pair of vertices $i, j \in G$ can be connected by a path $\tau_{i,j}$ in G. We will refer to the paths in a fixed collection $\{\tau_{i,j}\}_{(i,j)\in G}$ as *connecting paths* and call $t := \max\{l(\tau_{i,j}) : i, j\}$ a *connecting length of* G.

The shift map σ is expansive with the expansiveness constant $\epsilon = 1$ (see [Wal82]); therefore, in an analogous way to the usual topological entropy, one can easily verify that $h^{(v)}_{\text{top}}(\sigma, \psi)$ can be calculated on $(1,n)$-separated sets. If $S \subset \Lambda$ is $(1,n)$-separated, then it determines a collections of $\#S$ different paths in G of length n — for a point $(x_i) \in S$, the corresponding path α passes through $x_0, ..., x_n$ and $S_n(\sigma, \psi)(x) = \sum_\alpha \psi$. Clearly, every collection of different paths of length n arises from a $(1,n)$-separated set in this way. Thus we can write

$$h^{(v)}_{\text{top}}(\sigma, \psi) = \lim_{r\to 0} \limsup_{n\to\infty} (1/n) \ln s^{B_r(v)}(\sigma, \psi, 1, n) = \tag{2.12}$$

$$\lim_{r\to 0} \limsup_{n\to\infty} (1/n) \ln \#\{\alpha : \ \alpha \text{ path in } G, \ l(\alpha) = n, \ (1/n) \sum_\alpha \psi \in B_r(v)\}.$$

Since every path can be extended to a loop by adding an appropriate *connecting path*, we have

$$h_{\text{top}}^{(v)}(\sigma,\psi) = \lim_{r\to 0}\limsup_{n\to\infty}(1/n)\ln \#\{\alpha :\ \alpha \text{ loop in } G,\ l(\alpha)=n,\ \rho(\alpha)\in B_r(v)\}. \tag{2.13}$$

One can even insist that all the loops above depart from an arbitrarily fixed vertex. The "lim sup" in (2.13) can not be replaced with "lim inf" unless σ is mixing. This (slight) inconvenience is offset by the following simple fact (cf. the example in Appendix A).

Fact 2.2.1 *In the formula (2.12), one can replace* lim sup *with* lim inf.

Proof. Take an arbitrary $0<\eta<1$. We will show that with "lim inf" we get at least $\eta h_{\text{top}}^{(v)}(\sigma,\psi)$. Fix $v\in\mathbf{R}^d$ and $r>0$. By (2.13), there exist arbitrarily large $n_0\in\mathbf{N}$ such that there is a collection Γ of different loops of length n_0 with cardinality $\#\Gamma \geq \exp\left(h_{\text{top}}^{(v)}(\sigma,\psi)n_0\eta\right)$ and $(1/n_0)\sum_\gamma \phi \in B_{r/2}(v)$, $\gamma\in\Gamma$. As noted above, we can assume that all these loops depart from a fixed vertex of G so that we can concatenate them. For any $n\in\mathbf{N}$, set $k := \lfloor n/n_0 \rfloor$. By considering all possible $(\#\Gamma)^k$ concatenations of k-tuples from Γ^k, we conclude that

$$s^{B_r(v)}(\sigma,\psi,1,n) \geq s^{B_{r/2}(v)}(\sigma,\psi,1,kn_0) \geq (\#\Gamma)^k \geq \exp\left(h_{\text{top}}^{(v)}(\sigma,\psi)kn_0\eta\right)$$

(where for the first inequality n must be sufficiently large). Taking the logarithms and "lim inf" with $n\to\infty$ finishes the proof. Q.E.D.

Combination of the results in the preceding section of this chapter with the known facts about subshifts of finite type gives us the following theorem

— very much in the spirit of [Bow75, Rue78]. (A reader not familiar with the *Gibbs states* may want to consult [Bow75].)

Theorem 2.2.1 *Suppose that (as above) $\sigma : \Lambda \to \Lambda$ is a transitive subshift of finite type with observable $\psi : \Lambda \to \mathbf{R}^d$ that depends only on the first two symbols. If $A = (a_{i,j})_{1\leq i,j\leq N}$ is the corresponding binary matrix and G is the corresponding transition graph with weights $(\psi_{i,j})_{1\leq i,j\leq N}$, then:*
(i) $\rho(\sigma, \phi)$ is a convex polyhedron with vertices in

$$\{\rho(\alpha) : \quad \alpha \textit{ is an elementary loop in } G\};$$

(ii) $-h^{(v)}_{\text{top}}(\sigma, \psi)$ is convex continuous on $\rho(\sigma, \psi)$ and strictly convex and real analytic on $\text{ri}(\rho(\sigma, \psi))$;
(iii) $p^{(t)}_{\text{top}}(\sigma, \psi) = \ln(\lambda(t))$, where $\lambda(t)$ is the leading positive eigenvalue of the transfer matrix $A(t) = (a_{i,j} \exp\langle t, \psi_{i,j}\rangle)_{1\leq i,j\leq N}$, $t \in \mathbf{R}^d$;
(iv) $-h^{(v)}_{\text{top}}(\sigma, \psi) = \hat{p}^{(v)}_{\text{top}}(\sigma, \psi)$, for $v \in \text{ri}(\rho(\sigma, \phi))$;
(v) $h^{(v)}_{\text{top}}(\sigma, \psi) = h^v_{\text{erg}}(\sigma, \psi) = h^v_{\text{meas}}(\sigma, \psi) = h_{\mu^{(v)}}(\sigma)$, where $v \in \text{ri}(\rho(\sigma, \psi))$ and $\mu^{(v)}$ is the Gibbs measure *corresponding to the function on Λ given by $x \mapsto \langle t, \psi(x)\rangle$, with $t \in \mathbf{R}^d$ uniquely determined by $v = \nabla_t p^{(t)}_{\text{top}}(\sigma, \psi)$ (where ∇_t is the gradient with respect to t).*

Remark. The dimension of $\rho(\sigma, \psi)$ may be lower than d. Thus real analyticity in (ii) is with respect to the affine hull of $\rho(\sigma, \psi)$. Nevertheless, even though we have chosen not to do so, one can always assume that $\rho(\sigma, \psi)$ has the dimension of the target space of ψ. This is achieved by replacing ψ with its projection onto the affine hull of $\rho(\sigma, \psi)$.

Remark. Let us add to (v) that, since ψ is locally constant, the Gibbs measure is a Markov measure. More specifically, if $u = (u_i)_{i=1,...N}$ and $v = (v_i)_{i=1,...N}$ are the eigenvectors, corresponding to the eigenvalue $\lambda(t)$, of $A(t)$ and of the transpose $A(t)^*$ respectively, then this Markov measure is determined by a unique stochastic vector proportional to $(u_i v_i)_{i=1,...N}$ and by the stochastic matrix $(\lambda^{-1} A(t)_{i,j} v_j / v_i)_{i,j=1,...N}$ (cf. p. 194 in [Wal82].) Also, $u_i, v_i > 0$, $i = 1, ..., N$, so this measure is supported on the whole Λ.

Remark. For a subshift of finite type that is not transitive, the entropy $h^{(v)}_{\text{top}}(\sigma, \psi)$ is the maximum of the entropies on the transitive components.

The terminology *transfer matrix* comes from statistical mechanics (see [Ell85, Rue78]). Part (iii) of Theorem 2.2.1 is the essence of the *transfer matrix method* (see [May91] for an exposition).

Proof of (i). This has been apparently realized by many; it is implicitly contained in [Fri82b], mentioned in [Kwa92], and proved in the full generality in [Zie95]. One way to argue is as follows. By transitivity, the rotation set is the closure of the ψ-averages $\rho(\beta)$ taken over all loops β in G. Thus it is enough to see that, for every loop β, the sum $\sum_\beta \psi$ can be represented as a sum of $\sum_\alpha \psi$'s for a number of elementary loops α. This can be done by induction with respect to the length of β. If β is not elementary, then certain vertex v repeats at least twice and, by taking the two segments of β between two different occurrences of v, one can "pinch" β to two <u>shorter</u> loops without changing the sum of weights. Q.E.D.

Proof of (iii). The topological pressure is given by

$$p_{\text{top}}^{(t)}(\sigma,\psi) = \lim_{n\to\infty}(1/n)\ln\left(\sum\left\{\exp\left\langle t, \sum_{\alpha}\psi\right\rangle :\ \alpha \text{ path in } G,\ l(\alpha)=n\right\}\right), \tag{2.14}$$

or, in terms of loops, we can write (cf. (2.13) and [May91])

$$p_{\text{top}}^{(t)}(\sigma,\psi) = \limsup_{n\to\infty}(1/n)\ln\left(\sum\left\{\exp\left\langle t, \sum_{\alpha}\psi\right\rangle :\ \alpha \text{ loop in } G,\ l(\alpha)=n\right\}\right).$$

(The convergence takes place for mixing σ.) The right-hand side can be written as

$$\limsup_{n\to\infty}(1/n)\ln(\operatorname{trace}(A(t)^n)) = \ln(\lambda(t)),$$

where the matrix $A(t) = (a_{i,j}\exp\langle t, \psi_{i,j}\rangle)_{1\le i,j\le N}$, and $\lambda(t)$ is its leading positive eigenvalue given by the Perron-Frobenius theorem ([Wal82]). Q.E.D.

Proof of (iv). Because $A(t)$ is a non-negative irreducible matrix, the leading positive eigenvalue $\lambda(t)$ is a simple root of the corresponding characteristic polynomial, and so it depends real-analytically on $t \in \mathbf{R}^d$. In this way $p_{\text{top}}^{(t)}(\sigma,\psi)$ is smooth; and consequently, its convex conjugate is strictly convex on the interior of the effective domain (see p. 253, [Roc72], cf. (2.15) below). We are done by Theorem 2.1.1. (The effective domain is equal to $\rho(\sigma,\psi)$ by Corollary 2.1.1.) Q.E.D.

Proof of (v). Since by (iii) $p_{\text{top}}^{(t)}(\sigma,\psi)$ is smooth, its convex conjugate $\hat{p}_{\text{top}}^{(v)}(\sigma,\psi)$ can be calculated for $v \in \operatorname{ri}\left(\operatorname{dom}\left(\hat{p}_{\text{top}}^{(v)}(\sigma,\psi)\right)\right)$ as follows (cf. Th. 23.5, [Roc72])

$$\hat{p}_{\text{top}}^{(v)}(\sigma,\psi)) = \langle v, t\rangle - p_{\text{top}}^{(t)}(\sigma,\psi), \tag{2.15}$$

where t is uniquely determined by $\nabla_t p^{(t)}_{\text{top}}(\sigma,\psi) = v$. Moreover, we observed while proving (iv) that $\rho(\sigma,\psi) = \text{dom}(\hat{p}^{(v)}_{\text{top}}(\sigma,\psi))$.

Let $\mu^{(v)}$ be the *Gibbs state* for $\langle t,\psi\rangle$. According to Ruelle (see Th.1.22 in [Bow75]) it is an (ergodic) *equilibrium state*, that is

$$p^{(t)}_{\text{top}}(\sigma,\psi) = h_{\mu^{(v)}}(\sigma) + \Big\langle t, \int \psi \, d\mu^{(v)} \Big\rangle.$$

The affine function $\xi : t' \mapsto h_{\mu^{(v)}}(\sigma) + \langle t', \int \psi \, d\mu^{(v)}\rangle$ is majorized by the convex function $p^{(t)}_{\text{top}}(\sigma,\psi)$ due to the variational principle, and, by the above equality, it is actually tangent to $p^{(t)}_{\text{top}}(\sigma,\psi)$. Comparison with (2.15) yields $\int \psi \, d\mu^{(v)} = v$ and $-h_{\mu^{(v)}}(\sigma) = \hat{p}^{(v)}_{\text{top}}(\sigma,\psi)$. We are done by (iv) and Corollary 2.1.3. Q.E.D.

Proof of (ii). While proving (iv) we observed strict convexity of $\hat{p}^{(v)}_{\text{top}}(\sigma,\psi) = -h^{(v)}_{\text{top}}(\sigma,\psi)$, for $v \in \text{ri}(\rho(\sigma,\psi))$. By (2.15) and real-analyticity of $p^{(t)}_{\text{top}}(\sigma,\psi)$, $v \mapsto h^{(v)}_{\text{top}}(\sigma,\psi)$ is real-analytic on $\text{ri}(\rho(\sigma,\psi))$.

Still, one needs to show convexity on the whole $\rho(\sigma,\psi)$. Then continuity follows; indeed, $-h^{(v)}_{\text{top}}(\sigma,\psi)$ is lower semi-continuous by (iv) of Proposition 2.1.1, and, by Th. 10.2, p.84 [Roc72], it is upper semi-continuous as a convex function on a polyhedron.

To demonstrate convexity, it is enough to show the following: if $v_1, v_2 \in \rho(\sigma,\psi)$ and $v = v_1/2 + v_2/2$, then

$$h^{(v)}_{\text{top}}(\sigma,\psi) \geq (1/2)h^{(v_1)}_{\text{top}}(\sigma,\psi) + (1/2)h^{(v_2)}_{\text{top}}(\sigma,\psi).$$

It is routine to extend this inequality to 2-adic convex combinations and then (by upper semi-continuity of $h^{(v)}_{\text{top}}(\sigma,\psi)$) to arbitrary convex combinations.

Fix an arbitrary $r > 0$. By (2.12) and Fact 2.2.1, there are arbitrarily large $n_0 \in \mathbf{N}$ for which one can find two families Γ_1 and Γ_2 of paths [1] in G, with length n_0, so that $\#\Gamma_i \geq \exp\left((h_{\text{top}}^{(v_i)}(\sigma, \psi) - r)n_0\right)$ and $\sum_\gamma \psi \in n_0 B_r(v_i)$ for $\gamma \in \Gamma_i$, $i = 1, 2$. Moreover, we may assume that $n_0 r$ exceeds the *connecting length* t.

For each pair of paths (γ_1, γ_2) in $\Gamma_1 \times \Gamma_2$ we can find a *connecting path* τ so that $\gamma := \gamma_1 \tau \gamma_2$ is a path in G. At least $\#\Gamma_1 \#\Gamma_2 / t$ of thus obtained paths γ will have the same length

$$n := l(\gamma) = 2n_0 + l \leq (1 + r)2n_0, \tag{2.16}$$

where $l := l(\tau)$. Let Γ be a collection of those γ's. Observe that, if $\gamma \in \Gamma$, then

$$(1/n) \sum_\gamma \psi \in B_{r+\delta}(v),$$

where δ accounts for the contribution introduced by τ and is smaller than r for large n_0. Thus we have

$$s^{B_{2r}}(\sigma, \psi, 1, n) \geq \#\Gamma \geq \exp\left((h_{\text{top}}^{(v_1)}(\sigma, \psi) - r)n_0\right) \cdot \exp\left((h_{\text{top}}^{(v_2)}(\sigma, \psi) - r)n_0\right)/t.$$

By passing with n_0 to infinity, while keeping in mind (2.16), we get

$$h_{\text{top}}^{B_{2r}(v)}(\sigma, \psi) \geq (2(1 + r))^{-1} \left(h_{\text{top}}^{(v_1)}(\sigma, \psi) + h_{\text{top}}^{(v_2)}(\sigma, \psi) - 2r\right).$$

We are done by the arbitrariness of r. Q.E.D.

[1] These can be actually chosen to be all loops through a fixed vertex.

2.2.2 Observable shadowing with the right entropy

Let σ, Λ, ψ, and G be as in the preceding subsection. By (v) of Theorem 2.2.1, we know that, for any v in $\mathrm{ri}(\rho(\sigma,\psi))$, there is an ergodic measure with the metric entropy equal to $h_{\mathrm{top}}^{(v)}(\sigma,\psi)$. That measure is supported on all of Λ (see the remark after Theorem 2.2.1), and the support could not possibly be smaller as shown by the following proposition.

Proposition 2.2.1 *Suppose that $v \in \mathrm{ri}(\rho(\sigma,\psi))$. If K is a compact σ-invariant proper subset of Λ with $\rho(\sigma|_K,\psi) = v$, then $h_{\mathrm{top}}^{(v)}(\sigma|_K) < h_{\mathrm{top}}^{(v)}(\sigma,\psi)$.*

The proposition also demonstrates that the following theorem is optimal.

Theorem 2.2.2 *For $v \in \mathrm{ri}(\rho(\sigma,\psi))$ and $0 < \eta < 1$, there exists a compact σ-invariant set $K \subset \Lambda$ such that $h_{\mathrm{top}}(\sigma|_K) \geq \eta h_{\mathrm{top}}^{(v)}(\sigma,\psi)$ and $\rho(\sigma|_K,\psi) = v$. In fact, K is such that, for some constant C, $\|S_n(\sigma,\psi)(x) - nv\| \leq C$, $x \in K$, $n \in \mathbf{N}$.*

This theorem, for the special case of ψ with values in $\mathbf{Z}$ and $v \in \mathbf{Z}$, is proved in [MT92]. Also, in [Zie95] it is shown how to find for each $v \in \rho(\sigma,\psi)$ a compact invariant set K with $\rho(\sigma|_K,\psi) = v$. Our methods, developed independently of [MT92, Zie95], are different. Let us also note that, using the *scaffold construction* of the next subsection, one can get a stronger result for h_{erg}^{v}, see Theorem 2.2.4.

The proposition is simple, so we prove it first. We need the following lemma.

Lemma 2.2.1 *If K is a compact, σ-invariant* <u>*proper*</u> *subset of Λ, then, for any $t \in \mathbf{R}^d$, we have $p_{\text{top}}^{(t)}(\sigma|_K, \psi) < p_{\text{top}}^{(t)}(\sigma, \psi)$.*

Proof of Proposition 2.2.1 from Lemma 2.2.1. Pick $v \in \text{ri}(\rho(\sigma, \psi))$. There is the corresponding $t_0 \in \mathbf{R}^d$ such that $-h_{\text{top}}^{(v)}(\sigma, \psi) = \langle v, t_0 \rangle - p_{\text{top}}^{(t_0)}(\sigma, \psi)$, (see (2.15)). Using Theorem 2.1.1 for the first inequality and the lemma for the third, we obtain

$$-h_{\text{top}}^{(v)}(\sigma|_K, \psi) \geq \sup_{t \in \mathbf{R}^2} \{\langle v, t \rangle - p_{\text{top}}^{(t)}(\sigma|_K, \psi)\} \geq$$
$$\langle v, t_0 \rangle - p_{\text{top}}^{(t_0)}(\sigma|_K, \psi) > \langle v, t_0 \rangle - p_{\text{top}}^{(t_0)}(\sigma, \psi) = -h_{\text{top}}^{(v)}(\sigma, \psi).$$

Q.E.D.

Proof of Lemma 2.2.1. Let μ be the *Gibbs measure* associated with $\langle t, \psi \rangle$. Its characteristic property is that there is a constant $C > 0$ such that, for any $x = (x_i)_{i \in \mathbf{Z}} \in \Lambda$ and $n \in \mathbf{N}$, ([Bow75])

$$C^{-1} \leq \mu([x]_0^n) / \exp\Big(\langle S_n(\sigma, \psi)(x), t \rangle - np_{\text{top}}^{(t)}(\sigma, \psi)\Big) \leq C, \tag{2.17}$$

where $[x]_0^n := \{y \in \Lambda : \ y_i = x_i, \ i = 0, ..., n-1\}$ is a n-*cylinder*. Since K is proper and μ is ergodic and supported on the whole Λ, we have $\mu(K) = 0$. We can contain K in a union of disjoint cylinders that total to arbitrarily small μ measure; more concretely, for sufficiently high $n_0 \in \mathbf{N}$, there exists a discrete subset E_{n_0} of Λ such that $K \subset \bigcup_{x \in E_{n_0}} [x]_0^{n_0}$ and $\mu(\bigcup_{x \in E_{n_0}} [x]_0^{n_0}) \leq (10C)^{-1}$. Now, for $n \in n_0\mathbf{N}$, consider any $(1, n)$-separated subset of $S \subset K$. Any n_0-cylinder that is determined by n_0 consecutive symbols of a point in S must be one of $[x]_0^{n_0}$, $x \in E_{n_0}$. This, together with (2.17), yields the following estimate

$$\sum_{x\in S} \exp\langle S_n(\sigma,\psi)(x), t\rangle \;\le\; \prod_{i=1}^{n/n_0} \sum_{x\in E_{n_0}} \exp\langle S_{n_0}(\sigma,\psi)(x), t\rangle \le$$

$$\prod_{i=1}^{n/n_0} \sum_{x\in E_{n_0}} C\mu([x]_0^{n_0}) \exp\left(np_{\text{top}}^{(t)}(\sigma,\psi)\right) \;\le\; \left(C(10C)^{-1} \exp\left(n_0 p_{\text{top}}^{(t)}(\sigma,\psi)\right)\right)^{n/n_0}.$$

Taking the logarithms and passing to the limit with $n \to \infty$, we get (by (2.14))

$$p_{\text{top}}^{(t)}(\sigma|_K, \psi) \le p_{\text{top}}^{(t)}(\sigma,\psi) + (1/n_0)\ln(1/10).$$

Q.E.D.

We turn our attention to Theorem 2.2.2 now. As most of proofs in this section our argument hinges on concatenation of paths in G, however this time we need an original approach to control the averages. Sacrificing the efficiency of exposition, we take an opportunity to present a general technique that easily yields trajectories in Λ that have an a priori prescribed behavior of the averages. (See Lemma 2.2.2, ahead.)

We start with a simple characterization of convex polyhedra that may be interesting by itself. (Below, for two sets $A, B \subset \mathbf{R}^d$, $A + B$ stands for the *Minkowski sum*, i.e. $A + B = \{a + b : \; a \in A, \; b \in B\}$.)

Theorem 2.2.3 (characterization of convex polyhedra) *For a compact convex $\rho' \subset \mathbf{R}^d$, the following are equivalent :*

(i) ρ' is a polyhedron;

(ii) there exist $m \in \mathbf{N}$ and $\omega_1, ..., \omega_m \in \rho'$ such that $\rho' + \rho' = \bigcup_{i=1}^m \omega_i + \rho'$. Moreover, the ω_i's can be chosen so that they are convex combinations with rational coefficients of the extremal points of ρ'.

Proof of Theorem 2.2.3. To prove that (ii) implies (i), we just notice that, for $x, y \in \rho'$: if the sum $x + y$ is an extremal point of $\rho' + \rho' = 2 \cdot \rho'$, then both x and y must be extremal and equal to each other. Thus m bounds from above the number of extremal points of ρ' what makes it a polyhedron.

To prove (ii) from (i), we first note that affine transformations of ρ' do not affect validity of (ii). Since every polygon with n vertices is an affine image of the standard simplex $\Delta^n = \{x \in \mathbf{R}^n : \; x_i \geq 0, \; \sum x_i = 1\} \subset \mathbf{R}^n$, we may consider only the case when $\rho' = \Delta^n$.

We claim that

$$2 \cdot \Delta^n = \bigcup \{w + \Delta^n : w = (w_1, ..., w_n) \in \Delta^n, \; w_i \in (1/n)\mathbf{N}, \; i = 1, ..., n\}.$$

To get this, fix any $x = (x_1, ..., x_n) \in 2 \cdot \Delta^n$. Clearly $\sum_{i=1}^n x_i = 2$, so we see that $\sum_{i=1}^n \lfloor nx_i \rfloor \geq \sum_{i=1}^n (nx_i - 1) = 2n - n = n$. Hence, we can choose $k_i \in \mathbf{N}$, $0 \leq k_i \leq \lfloor nx_i \rfloor$, $i = 1, ..., n$, such that $\sum_{i=1}^n k_i = n$. Now, by setting $w_i := k_i/n$ and $y_i := x_i - w_i$, $i = 1, ..., n$, we get $\sum_{i=1}^n w_i = n/n = 1$ and $\sum_{i=1}^n y_i = 2 - 1 = 1$, so both $w = (w_i)$ and $y = (y_i)$ belong to Δ^n. Clearly $x = w + y$. Q.E.D.

The following "shadowing type result" is the consequence of Theorem 2.2.3 that we are interested in.

Corollary 2.2.1 *Suppose that ρ' and ω_i's are as in Theorem 2.2.3. If $x : \mathbf{R}^+ \to \mathbf{R}^d$ is piecewise C^1 with $x(0) = 0$ and $x'(t) \in \rho'$, $t \in \mathbf{R}^+$, then there exists a sequence $\kappa \in \{1, .., m\}^{\mathbf{N}}$ such that $x(n) \in \omega_{\kappa_1} + ... + \omega_{\kappa_{n-1}} + \rho'$, for all $n \in \mathbf{N}$.*

Proof. We proceed by induction with respect to n. Since $x'(s) \in \rho'$, by convexity of ρ', $x(1) = \int_0^1 x'(s)ds \in \rho'$ taking care of the case $n = 1$. For the induction step, write $x(n+1) = x(n) + \int_n^{n+1} x'(s)ds \in x(n) + \rho'$ and use induction hypothesis to see that $x(n+1) \in \omega_{\kappa_1} + ... + \omega_{\kappa_{n-1}} + \rho' + \rho'$. By (ii) of Theorem 2.2.3, one can to find $1 \leq \kappa_n \leq m$ so that $x(n+1) \in \omega_{\kappa_1} + ... + \omega_{\kappa_n} + \rho'$. Q.E.D.

Let us now explain how Corollary 2.2.1 renders orbits in Λ (or equivalently paths in the graph G) with prescribed behavior of the Birkhoff averages of ψ.

By transitivity, we can find loops η_1, ... η_r, all emanating from a common arbitrarily chosen vertex i_0, such that $\rho' := \text{conv}\{\rho(\eta_1), ..., \rho(\eta_r)\}$ approximates $\rho(\sigma, \psi)$ as well as we wish. Now, let ω_i, $i = 1, ..., m$, be as in (ii) of Theorem 2.2.3. Since ω_i's are all rational convex combinations of the vertices, by taking appropriate concatenations of η_i's, we can obtain loops $\beta_1, ..., \beta_m$ so that $\omega_i := \rho(\beta_i)$, $i = 1, ..., m$. Moreover, replacing each loop β_i with its appropriate multiple we can make all β_i's of equal length l_0. One can concatenate β_i's at will: if $q \in \mathbf{N}$ and $\kappa \in \{1, ..., m\}^q$, then $\beta_\kappa := \beta_{\kappa_1} ... \beta_{\kappa_q}$ is a loop in G.

The following lemma is in fact much stronger than what we need for Theorem 2.2.2. For a path ξ in G, by ξ_m^n we understand its *restriction* to $\{m, ..., n-1\}$ (if it is well defined).

Lemma 2.2.2 *If $x : \mathbf{R} \to \mathbf{R}^d$ is piecewise C^1 with $x(0) = 0$ and $x'(t) \in \rho'$, $t \in \mathbf{R}$, then there exists a bi-infinite path ξ in G following x, i.e.*

$$\left\| \sum_{\xi_m^n} \psi - (x(n) - x(m)) \right\| \leq A := 4 l_0 \text{diam}(\rho') + 2 l_0 \sup_{x \in \Lambda} \|\psi(x)\|,$$

for all $n < m,\ n, m \in \mathbf{Z}$. *Moreover, for* $n \in l_0\mathbf{N}$, *the edge* $\xi(n)$ *ends at the vertex* i_0.

Clearly, the smoothness hypothesis is superficial.

Proof. By applying Corollary 2.2.1 to the rescaled path $t \mapsto (1/l_0)x(l_0 t)$ twice, once with time running forward ($t > 0$), once with time running backward ($t < 0$), we get $\kappa \in \{1, ..., M\}^{\mathbf{Z}}$ such that for any $k, l \in \mathbf{Z}$ with $k < l$,

$$\left\| \sum_{i=k}^{l-1} \omega_{\kappa_i} - (1/l_0)(x(l_0 l) - x(l_0 k)) \right\| \leq 2\,\mathrm{diam}(\rho').$$

Take for ξ the infinite concatenation $\xi := \beta_\kappa = ...\beta_{\kappa_{-1}}\beta_{\kappa_0}\beta_{\kappa_1}...$. By multiplying the above inequality by l_0, while remembering that $\omega_i = \rho(\beta_i) = (1/l_0)\sum_{\beta_i} \psi$, we see that

$$\left\| \sum_{\xi_{ll_0}^{kl_0}} \psi - (x(ll_0) - x(kl_0)) \right\| \leq 2l_0 \mathrm{diam}(\rho').$$

The above inequality for restrictions of ξ to intervals between multiples of l_0 yields the result because, for m, n with $|m-n| \leq l_0$, both $\| \sum_{\gamma_m^n} \psi\|$ and $\|x(n) - x(m)\|$ are trivially bounded by $l_0 \sup_{x\in\Lambda} \|\psi(x)\|$ and l_0 diam(ρ') respectively. Q.E.D.

Proof of Theorem 2.2.2. We will construct a sequence of natural numbers $n_i \to \infty$ and a sequence of $(1, n_i)$-separated sets $E_i \subset \Lambda$ such that

$$\#E_i \geq \exp(h_{\mathrm{top}}^{(v)}(\sigma, \psi)\eta n_i) \tag{2.18}$$

and, for $y \in E_i$,

$$\|S_k(\sigma, \psi)(y) - kv\| \leq C, \tag{2.19}$$

where $k \in \mathbf{N}$ and C is a constant independent on i. Then the set $K := \mathrm{cl}(\bigcup_{l\in\mathbf{N},i\in\mathbf{N}} \sigma^l(E_i))$ satisfies the assertion of Theorem 2.2.2. This is because (2.19) is satisfied on a closed set of $y \in \Lambda$ and, if it holds for y, then it holds for $\sigma(y), \sigma^2(y), ...$ with C replaced by $2C$.

Fix $v \in \mathrm{ri}(\rho(\sigma,\psi))$. In constructing E_i we will think of its points as infinite paths in G. As in the discussion preceding Lemma 2.2.2, choose loops $\beta_1, ...\beta_m$ so that $v \in \mathrm{ri}(\rho')$, where $\rho' = \mathrm{conv}\{\rho(\beta_1), ..., \rho(\beta_m)\}$. Recall that all β_i's have a common starting vertex i_0. Let $\delta > 0$ be rational and such that $1/(1+\delta) \geq \sqrt{\eta}$. Take $r > 0$ satisfying $v - B_{2r/\delta}(0) \subset \rho'$. From the definition of $h^{(v)}_{\mathrm{top}}(\sigma,\psi)$ rephrased as (2.13), find $n_0 \in \mathbf{N}$ and a collection S_0 of loops in G, with length n_0 and starting at i_0, such that $\#S_0 \geq \exp(h^{(v)}_{\mathrm{top}}(\sigma,\psi)\sqrt{\eta}n_0)$ and $(1/n_0)\sum_\gamma \psi \in B_r(v)$ for any $\gamma \in S_0$. Require also that n_0 is large enough so that

$$A = 4l_0\mathrm{diam}(\rho') + 2l_0 \sup_{x\in\Lambda} \|\psi(x)\| \leq rn_0. \tag{2.20}$$

With no loss of generality one can assume that $n_0\delta \in l_0\mathbf{N}$.

Set $n_i := in_0(1+\delta)$, $i = 1, 2, ...$.

Fix $i \in \mathbf{N}$, and choose $\gamma_1, ...\gamma_i \in S_0$ arbitrarily. With this data we will associate a path ζ of length n_i and of the form $\gamma_1\xi_1\gamma_2\xi_2....\gamma_i\xi_i$, where ξ_j's will be chosen so that

$$\left\| \sum_{\gamma_1\xi_1...\gamma_j\xi_j} \psi - n_j v \right\| \leq rn_0, \ j = 1, 2, ..., i. \tag{2.21}$$

We define ξ_j's inductively. We start with finding ξ_1. Observe that, since $\sum_{\gamma_1} \psi \in n_0 v + B_{n_0 r}(0)$,

$$\left(n_0(1+\delta)v - \sum_{\gamma_1} \psi \right) / (\delta n_0) \in (n_0 \delta v + B_{n_0 r}(0))/(\delta n_0) = v + B_{r/\delta}(0) \subset \rho'.$$

Consequently, there is a (linear) path $x(t) \in \mathbf{R}^d, 0 \le t \le \delta n_0$, such that $x'(t) \in \rho'$, $t \in [0, \delta n_0]$, $x(0) = 0$, $x(\delta n_0) = n_0(1+\delta)v - \sum_{\gamma_1} \psi$. From Lemma 2.2.2, we get a path ξ_1 such that $l(\xi_1) = \delta n_0$ and

$$\left\| \sum_{\xi_1} \psi - x(\delta n_0) \right\| \le A \le r n_0.$$

Note that ξ_1 is in fact a loop that starts (and ends) at the vertex i_0, and so is γ_1; thus we can concatenate the two. By the definition of x, $\sum_{\gamma_1} \psi + x(\delta n_0) - n_1 v = 0$, so from the above inequality we verify (2.21) for $j = 1$:

$$\left\| \sum_{\gamma_1 \xi_1} \psi - n_1 v \right\| \le \left\| \sum_{\gamma_1} \psi + x(\delta n_0) - n_1 v \right\| + \left\| \sum_{\xi_1} \psi - x(\delta n_0) \right\| \le r n_0. \quad (2.22)$$

To get ξ_{j+1}, having $\xi_1, ..., \xi_j$ already defined and satisfying (2.21), we repeat the above procedure with a tiny modification. Observe that, since $\sum_{\gamma_{j+1}} \psi \in n_0 v + B_{n_0 r}(0)$, using (2.21) as the induction hypothesis, we get

$$\left(n_{j+1} v - \sum_{\gamma_1 \xi_1 \gamma_2 \ldots \gamma_j \xi_j \gamma_{j+1}} \psi \right) / (\delta n_0) =$$
$$\left(n_0 \delta v + (n_0 v - \sum_{\gamma_{j+1}} \psi) + (n_j v - \sum_{\gamma_1 \xi_1 \gamma_2 \ldots \gamma_j \xi_j} \psi) \right) / (\delta n_0)$$
$$\in (\delta n_0 v + B_{2 n_0 r}(0))/(\delta n_0) = v - B_{2r/\delta}(0) \subset \rho'.$$

As before, we conclude that there is a (linear) path $x(t) \in \mathbf{R}^d, 0 \le t \le \delta n_0$, such that $x'(t) \in \rho'$, $t \in [0, \delta n_0]$, $x(0) = 0$, $x(\delta n_0) = n_{j+1} v - \sum_{\gamma_1 \xi_1 \ldots \xi_j \gamma_{j+1}} \psi$.

From Lemma 2.2.2, we get a path ξ_{j+1} such that $l(\xi_{j+1}) = \delta n_0$ and

$$\left\| \sum_{\xi_{j+1}} \psi - x(\delta n_0) \right\| \leq A \leq r n_0.$$

For the concatenation $\gamma_1 \xi_1 ... \gamma_{j+1} \xi_{j+1}$ we have

$$\left\| \sum_{\gamma_1 \xi_1 ..., \gamma_{j+1} \xi_{j+1}} \psi - n_{j+1} v \right\| \leq$$

$$\left\| \sum_{\gamma_1 \xi_1 ..., \gamma_{j+1}} \psi + x(\delta n_0) - n_{j+1} v \right\| + \left\| \sum_{\xi_{j+1}} \psi - x(\delta n_0) \right\| \leq 0 + r n_0.$$

This ends the induction step.

The above procedure yields $(\# S_0)^i$ different paths ζ of the form $\zeta = \gamma_1 \xi_1 ... \gamma_i \xi_i$, because we have that many choices of $\gamma_1, ..., \gamma_i \in S_0$. To get from each such ζ a point of E_i, just extend the path $\zeta = \gamma_1 \xi_1 ... \gamma_i \xi_i$ to a bi-infinite path. In this way, E_i is indeed a $(1, n_i)$-separated set with

$$\# E_i = (\# S_0)^i \geq \left(\exp(h^{(v)}_{\text{top}}(\sigma, \psi) \sqrt{\eta} n_0) \right)^i = \exp\left(h^{(v)}_{\text{top}}(\sigma, \psi) n_i \sqrt{\eta} / (1 + \delta) \right).$$

By the choice of δ, (2.18) follows. The (2.19), for $k \leq n_i$, is an immediate consequence of (2.21) (with C depending on n_0). To get it right for all k, one just has to exercise some care while extending ζ's to the bi-infinite paths. Namely, the semi-infinite paths that are concatenated to ζ must *follow* $x(t) := tv$, as in Lemma 2.2.2. Q.E.D.

2.2.3 The scaffold of subshifts of finite type

The following discussion (based on [MT91]) explains how to use the results of the previous sections to learn about $h^{(v)}_{\text{top}}(\sigma, \psi)$ and $h^{v}_{\text{erg}}(\sigma, \psi)$ even if

$v \notin \mathrm{ri}(\rho(\sigma,\psi))$. Since $\rho(\sigma,\psi)$ is a convex polyhedron, we can think of $\rho(\sigma,\psi)$ as the total space of a polyhedral complex with strata $\rho^{(i)}(\sigma,\psi)$, $i = 1, 2, \dots .d$, $\dim(\rho^{(i)}(\sigma,\psi)) = i$, where $\rho^{(d)}(\sigma,\psi) = \{\rho(\sigma,\psi)\}$ and, inductively, $\rho^{(i-1)}(\sigma,\psi)$ consists of all faces of polyhedrons in $\rho^{(i)}(\sigma,\psi)$. (By a face of a k-dimensional convex polyhedron we understand a maximal $(k-1)$-dimensional convex polyhedron contained in its relative boundary.) Fix a point $v \in \rho(\sigma,\psi)$. The *support of* v *in* $\rho(\sigma,\psi)$ is the unique polyhedron in the complex that contains v and is of least dimension among such, denote it by F. Observe that $v \in \mathrm{ri}(F)$. Let $G^{(F)}$ be the graph obtained by the following pruning of G: an edge e of G is left in $G^{(F)}$, if and only if there is a loop α in G passing through e that satisfies $\rho(\alpha) \in F$. For the corresponding subshift of finite type we write $\sigma^{(F)} : \Lambda^{(F)} \to \Lambda^{(F)}$. The collection of all thus obtained $\sigma^{(F)}$'s (as we vary v in $\rho(\sigma,\psi)$) together with the obvious inclusions between them forms what is called in [MT91] a *scaffold of subshifts of finite type*. Note that subshifts $\sigma^{(F)}$ need not to be transitive, but this adds only a trivial complication to the considerations (recall the remark after Theorem 2.2.1).

A simple argument in [MT91] shows that any loop α in $G^{(F)}$ must have $\rho(\alpha) \in F$. One can easily deduce then that

$$\rho(\sigma^{(F)},\psi) = F.$$

It is also observed in [MT91] that any ergodic measure μ for which $\int \psi \, d\mu \in F$ is supported on $\Lambda^{(F)}$; consequently,

$$h^v_{\mathrm{erg}}(\sigma,\psi) = h^v_{\mathrm{erg}}(\sigma^{(F)},\psi), \text{ for } v \in F. \tag{2.23}$$

As a corollary, we get the following extension of Theorem 2.2.2 to $v \notin \mathrm{ri}(\rho(\sigma,\psi))$.

Theorem 2.2.4 *Let $v \in \mathbf{R}^d$ be such that $h^v_{\mathrm{erg}}(\sigma,\psi) \geq 0$ (i.e. $v = \int \psi\, d\mu$, for some ergodic probability measure μ). For $0 < \eta < 1$, there exists a compact σ-invariant set $K \subset \Lambda$ such that $h_{\mathrm{top}}(\sigma|_K) \geq \eta h^v_{\mathrm{erg}}(\sigma,\psi)$ and $\rho(\sigma|_K,\psi) = v$. In fact, K is such that, for some constant C, $\|S_n(\sigma,\psi)(x) - nv\| \leq C$, $x \in K$, $n \in \mathbf{N}$.*

Proof. If $v = \int \psi\, d\mu$ for an ergodic measure μ, then v belongs to $\rho(\sigma,\psi)$. Let F be the support of v in $\rho(\sigma,\psi)$. If F is a point, then one can take for K the whole subshift $\Lambda^{(F)}$ because, by (2.23), $h^v_{\mathrm{erg}}(\sigma,\psi) = h^v_{\mathrm{erg}}(\sigma^{(F)},\psi)$. In general $v \in \mathrm{ri}(F)$, so, by (v) of Th. 2.2.1, $h^v_{\mathrm{erg}}(\sigma,\psi) = h^{(v)}_{\mathrm{top}}(\sigma^{(F)},\psi)$, and we find $K \subset \Lambda^{(F)} \subset \Lambda$ by applying Theorem 2.2.2 to the subshift $\sigma^{(F)}$. Q.E.D.

The behavior of $h^{(v)}_{\mathrm{top}}$ upon passing from σ to the subshift $\sigma^{(F)}$ is given by the following proposition.

Proposition 2.2.2 *For $v \in \rho(\sigma,\psi)$ with support $F \subset \rho(\sigma,\psi)$, we have*

$$- h^{(v)}_{\mathrm{top}}(\sigma,\psi) = \text{convex hull}\left(-h^{(v)}_{\mathrm{top}}(\sigma^{(F)},\psi)\right). \tag{2.24}$$

(If $\sigma^{(F)}$ is transitive, then the *convex hull* is superfluous.) We will give two proofs of the proposition.

Short proof. Recall that σ is expansive making $h^v_{\mathrm{meas}}(\sigma,\psi)$ upper semi-continuous (cf. [Wal82]). Thus, since $-h^{(v)}_{\mathrm{top}}(\sigma,\psi)$ is convex, Corollary 2.1.2 yields

$$-h^{(v)}_{\mathrm{top}}(\sigma,\psi) = -h^v_{\mathrm{meas}}(\sigma,\psi) = \text{convex hull}\left(-h^v_{\mathrm{erg}}(\sigma,\psi)\right).$$

Using (2.23), we can write

$$-h^{v}_{\mathrm{erg}}(\sigma,\psi) = -h^{v}_{\mathrm{erg}}(\sigma^{(F)},\psi) \geq -h^{(v)}_{\mathrm{top}}(\sigma^{(F)},\psi) \geq -h^{(v)}_{\mathrm{top}}(\sigma,\psi).$$

Taking the *convex hull* of each function above finishes the argument. Q.E.D.

Note that ultimately the above proof owes its efficiency to nontrivial results of ergodic theory. For a reader ready to fend off superficial technicalities, we supply the following argument unveiling elementary mechanism standing behind the proposition.

Long elementary proof. Clearly $h^{(v)}_{\mathrm{top}}(\sigma,\psi) \geq h^{(v)}_{\mathrm{top}}(\sigma^{(F)},\psi)$; and so, by convexity of $-h^{(v)}_{\mathrm{top}}(\sigma,\psi)$, we have $-h^{(v)}_{\mathrm{top}}(\sigma,\psi) \leq \text{convex hull}\left(-h^{(v)}_{\mathrm{top}}(\sigma^{(F)},\psi)\right)$. We will prove the opposite inequality by showing that, for arbitrary $0 < \eta < 1$,

$$-\eta h^{(v)}_{\mathrm{top}}(\sigma,\psi) \geq \text{convex hull}\left(-h^{(v)}_{\mathrm{top}}(\sigma^{(F)},\psi)\right).$$

It is enough to consider the case when F is a face of $\rho(\sigma,\psi)$. Let V be the affine space spanned by F. Set

$$d := \min\left\{\text{dist}\,(\rho(\alpha),V):\ \sum_{\alpha}\psi \notin V,\ \alpha \text{ elementary loop in } G\right\}.$$

When $h^{(v)}_{\mathrm{top}}(\sigma,\psi) = 0$, there is nothing to prove. Otherwise, observe that one can find $\delta > 0$ such that, for sufficiently large $n \in \mathbf{N}$,

$$n\delta N(n,\delta n)\#G^{\delta n}\exp\left(2\delta n\right) \leq \exp\left((1-\eta)h^{(v)}_{\mathrm{top}}(\sigma,\psi)n\right), \tag{2.25}$$

where $\mathrm{N}(k,m)$ is the binomial Newton symbol. Also, let $W \subset F$ be discrete and such that the balls $\{B_\delta(w)\}_{w\in W}$ cover $F = \rho(\sigma^{(F)},\psi)$.

Pick $0 < r_0 < \delta$ and $k_0 \in \mathbf{N}$ so that:

if γ' is a path in $G^{(F)}$ with $l(\gamma') \geq k_0$, then $(1/l(\gamma')) \sum_{\gamma'} \psi \in \bigcup_{w \in W} B_{r_0}(w)$, and, for $w \in W$,

$$s^{B_{r_0}(w)}(\sigma^{(F)}, \psi, 1, k) \leq \exp\left((h_{\text{top}}^{(w)}(\sigma^{(F)}, \psi) + \delta)k\right), \text{ for } k > k_0. \tag{2.26}$$

(See the page 30 for definition of $s^{B_r(w)}(\sigma^{(F)}, \psi, 1, k)$.)

Now, take $r > 0$ small enough to have

$$Nr(k_0 + 1)/d \leq \delta, \tag{2.27}$$

where N is the number of vertices in G.

Consider for a while a fixed loop γ in G of length $l(\gamma) = n$ and with $\rho(\gamma) = (1/l(\gamma)) \sum_\gamma \psi \in B_r(v)$. We will show that "most" of γ lies in $G^{(F)}$.

As described in the argument for (i) of Theorem 2.2.1, one can rearrange the edges of γ to get elementary loops $\alpha_1, ..., \alpha_p$ such that $\sum_\gamma \psi = \sum_{i=1}^p \sum_{\alpha_i} \psi$. Note that

$$\text{dist}\left(\sum_\gamma \psi, nV\right) = \sum_i l(\alpha_i) \text{dist}(\rho(\alpha_i), V) \geq \#\{i : \ \rho(\alpha_i) \notin V\}.$$

Since $v \in F \subset V$, $\sum_\gamma \psi \in B_{nr}(nv)$ implies that at most nr/d of α_i's have $\rho(\alpha_i) \notin V$. Since the length of an elementary loop is by definition less than N, at most Nnr/d of the edges in γ can belong to $G \setminus G^{(F)}$. After deleting these edges, the path γ splits into at most nNr/d paths in $G^{(F)}$, of which those longer than k_0 are denoted by γ_1, ... γ_q, $q \leq nNr/d$, in the order of appearance along γ. The combined length of γ_i's is greater than $n - nNr/d - nNrk_0/d = n - nNr(k_0 + 1)/d \geq (1 - \delta)n$, where we used (2.27) for the last inequality.

One of the consequences is the following basic estimate

$$\left\|(1/\sum_{j=1}^{q} l(\gamma_j))\sum_{i=1}^{q}\sum_{\gamma_i}\psi-(1/l(\gamma))\sum_{\gamma}\psi\right\|\leq\delta':=2\delta\sup_{x\in\Lambda}\|\psi(x)\|. \tag{2.28}$$

Furthermore, with each γ_i associate $w_i \in W$ chosen so that

$$\left\|(1/l(\gamma_i))\sum_{\gamma_i}\psi-w_i\right\|\leq r_0\leq\delta,\ i=1,...q. \tag{2.29}$$

This and (2.28) yield

$$(1/\sum_{j=1}^{q} l(\gamma_j))\sum_{i=1}^{q} l(\gamma_i)w_i\in B_{r+\delta+\delta'}(v). \tag{2.30}$$

We claim that relatively few loops γ of length n as considered above can share the same γ_i's (in the same order and positions along γ) and also have a coinciding vector of w_i's associated with γ_i's. Indeed, by basic combinatorics, the number of possibilities for the choice of the deleted elements is bounded by

$$n\delta\,\mathrm{N}(n,n\delta)(\#G)^{n\delta}$$

and the number of choices of w_i's by

$$\mathrm{N}(n,\#W)\leq\exp(n\delta),$$

where the right most inequality holds for large n.

By definition of $h_{\mathrm{top}}^{(v)}(\sigma,\psi)$ (see (2.13)), there are arbitrarily large $n\in\mathbf{N}$ with at least $\exp((h_{\mathrm{top}}^{(v)}(\sigma,\psi)-\delta)n)$ different γ's as considered above. The two combinatorial estimates and "the pigeon-hole principle" guarantee that some selection of w_i's and the deleted elements (and thus also of γ_i's) is shared by at least

$$\exp\left((h_{\mathrm{top}}^{(v)}(\sigma,\psi)-\delta)n\right)/n\delta\,\mathrm{N}(n,n\delta)(\#G)^{n\delta}\exp(\delta n)\geq\exp\left(h_{\mathrm{top}}^{(v)}(\sigma,\psi)\eta n\right)$$

different γ's, where the inequality is due to the choice of δ, see (2.25).

For each $i = 1, ...q$, denote by S_i a collection of all different γ_i's showing in those γ's. The above estimate implies that $\#S_1 ... \#S_q \geq \exp\left(\eta h_{\text{top}}^{(v)}(\sigma, \psi) n\right)$. Also, by the choice of r_0 and k_0 and using (2.29), we get

$$\#S_i \leq \exp\left((h_{\text{top}}^{(w_i)}(\sigma^{(F)}, \psi) + \delta) l(\gamma_i)\right), \ i = 1, ...q.$$

Consequently, after taking the logarithms and limsup with $n \to \infty$, we get

$$h_{\text{top}}^{(v)}(\sigma, \psi)\eta \leq \left(1 / \sum_{j=1}^{q} l(\gamma_j)\right) \sum_{i=1}^{q} l(\gamma_i) h_{\text{top}}^{(w_i)}(\sigma^{(F)}, \psi) + \delta$$

and thus, if $w := \left(1/\sum_{j=1}^{q} l(\gamma_j)\right) \sum_{j=1}^{q} l(\gamma_i) w_i$, then

$$-h_{\text{top}}^{(v)}(\sigma, \psi)\eta \geq \text{convex hull}\left(-h_{\text{top}}^{(w)}(\sigma^{(F)}, \psi)\right) - \delta.$$

By (2.30), when δ is shrunk to zero, w approaches v, and the desired inequality follows by lower semi-continuity of the right side above in the argument w. Q.E.D.

Chapter 3

Topological entropy at the rotation vector for torus maps

The rotation set, in the case of torus homeomorphisms, is very far from being a complete topological invariant; this makes sensible the search for finer invariants. While the presence of a particular vector v in the rotation set merely tells us that there is "some" dynamics with this asymptotic average displacement, one may try to measure the abundance of those dynamics. Realizing this philosophy, we associate with v the corresponding "topological entropy". We use quotation marks because, even though there is a well agreed on definition of the topological entropy of a map ([Wal82]), there are many ways of relativizing it to v. We will mainly use a simple topological concotion $h_{\text{top}}^{(v)}$ and its measure theoretic counterpart h_{erg}^{v}, both of which are extensively discussed in Chapter 2 and Appendix A. For the pseudo-Anosov maps, $h_{\text{top}}^{(v)}$ and h_{erg}^{v} coincide and are explicitly calculable once we know the Markov partition (Theorem 2.2.1). The pseudo-Anosov maps, our favorite examples, also provide a way to give lower bounds for $h_{\text{top}}^{(v)}$ for a general map.

In the first section, we define the relevant entropies and state the lower bound they obey. In the second section, we translate the results of Section 2.2 to the case of the pseudo-Anosov maps. In the third section, combining the main theorems of Chapter 1 and Chapter 2, we prove the lower bound. Last section is devoted to a concrete example of a calculation of $h_{\text{top}}^{(v)}$.

3.1 Definitions and the main result

Recall that for $f \in \mathcal{H}(\mathbf{T}^2)$ and its lift $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$, $F - \text{id}$ descends to the *displacement* function $\phi_F : \mathbf{T}^2 \to \mathbf{R}^2$. Note that $\rho(F) = \rho(f, \phi_F)$. More generally, given an invariant set $K \subset \mathbf{T}^2$, we will write $\rho(F|_K)$ for $\rho(f|_K, \phi_F)$.

Definition 3.1.1 *For $v \in \mathbf{R}^2$, the topological entropy at the rotation vector v is*

$$h_{\text{top}}^{(v)}(F) := h_{\text{top}}^{(v)}(f, \phi_F) \in \{-\infty\} \cup [0, +\infty].$$

Similarly, we have

$$h_{\text{erg}}^{v}(F) := h_{\text{erg}}^{v}(f, \phi_F) \in \{-\infty\} \cup [0, +\infty].$$

Our ultimate goal in this chapter is to bound entropies $h_{\text{erg}}^{v}(F)$ and $h_{\text{top}}^{(v)}(F)$ from below in terms of the "size" of the rotation set and the relative distance of v to its boundary.

To this end, we state the following definition. For a convex compact $K \subset \mathbf{R}^2$ with $\text{int}(K) \neq \emptyset$, we define the *tent function of K* by

$$\tau_K(v) := \inf \{\|a - v\|/\|a - b\| : \ v \in \text{conv}\{a, b\}, \ a, b \in \partial K\}, \ v \in \text{int}(K).$$

Before we say more about the properties of $\tau_K(v)$, we formulate our main result. (For the definitions of $A(\cdot)$ and $I(\cdot)$, used below, see Section 1.1.)

Theorem 3.1.1 (Main Lower Bound) *There are universal constants C and C' such that, if $f \in \mathcal{H}(\mathbf{T}^2)$ is a diffeomorphism, $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ is its lift and $v \in \operatorname{int}(\rho(F))$, then*

$$h_{\mathrm{top}}^{(v)}(F) \geq h_{\mathrm{erg}}^{v}(F) \geq \max\{C \ln_+ A(\rho(F)), C'/I(\rho(F))\} \cdot \tau_{\rho(F)}(v). \tag{3.1}$$

Moreover, there exists an invariant compact set $\Lambda \subset \mathbf{T}^2$ such that

$$\rho(F|_\Lambda) = \{v\}$$

and

$$h_{\mathrm{top}}(f|_\Lambda) \geq \max\{C \ln_+ A(\rho(F)), C'/I(\rho(F))\} \cdot \tau_{\rho(F)}(v). \tag{3.2}$$

The proof of the theorem is given in Section 3.3; however, Section 3.2 is a necessary prerequisite for it.

Let us now shed some light on $\tau_K(v)$. First of all, it depends only on the affine properties of K. There is however a trivial bound

$$\tau_K(v) \geq \operatorname{dist}(v, \partial K)/\operatorname{diam}(K).$$

The "infimum" in the definition of $\tau_K(v)$ is achieved when a and b are *antipodal*, meaning that there is a pair of parallel lines, one tangent to ∂K at a, and the other tangent to ∂K at b. Indeed, by compactness of ∂K and continuity of the norm, there are a and b such that $v \in \operatorname{conv}\{a, b\}$ and $\tau_K(v) = \|a - v\|/\|a - b\|$. Figure 3.1 below gives an idea of what would go wrong if a and b were not antipodal.

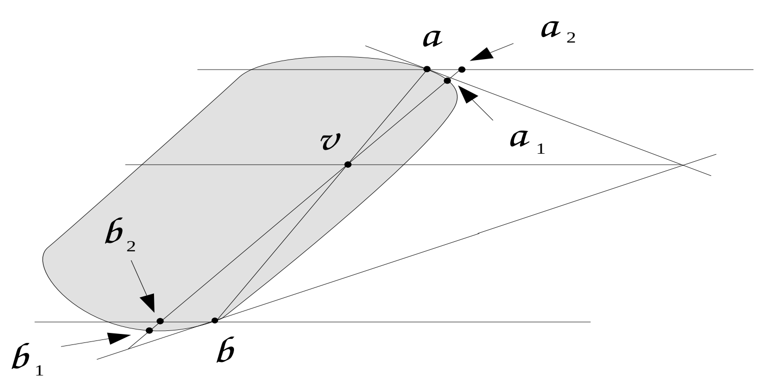

Figure 3.1: When a and b are not antipodal.

From Figure 3.1, note that $\|a_1 - v\| < \|a_2 - v\|$ and $\|v - b_1\| > \|v - b_2\|$, so $\|a_1 - v\|/\|v - b_1\| < \|a_2 - v\|/\|v - b_2\| = \|a - v\|/\|v - b\|$. Thus $\|a_1 - v\|/\|a_1 - b_1\| < \|a - v\|/\|a - b\|$, a contradiction.

Finally, observe that $\tau_K(v)$ is concave in v. In fact, we have

$$\tau_K(v) = \inf\left\{g(v) : \ g : K \to \mathbf{R}^+ \text{ concave with } \sup_{x\in K} g(x) = 1\right\}. \tag{3.3}$$

(This is why we call $\tau_K(v)$ *the tent function of* K.) For a proof, fix $v \in K$. If $g : K \to \mathbf{R}^+$ is concave with $\sup g = 1$, then one can take $a, b \in \partial K$ such that $\text{conv}\{a, b\}$ contains v and a point $w \in K$ where $g(w) = 1$. By concavity,

$$g(v) \geq \frac{\|v - w\|}{\|w - b\|} g(w) + \frac{\|v - b\|}{\|w - b\|} \cdot 0 = \|a - v\|/\|a - b\| \geq \tau_K(v).$$

On the other hand, if $\tau_K(v) = \|a - v\|/\|a - b\|$, then there is a g such that $g(v) \leq \|a - v\|/\|a - b\|$; simply consider the two parallel tangent lines through a and b and take g linear and constant on each of them.

3.2 The central example: pseudo-Anosov maps

Throughout this section let $f \in \mathcal{H}(\mathbf{T}^2)$ be a map that is *pseudo-Anosov rel a finite set* P and let $F : \mathbf{R}^2 \to \mathbf{R}^2$ be its lift to the universal cover. An abundance of rotation sets is exhibited by such examples. Indeed, we know that, given a convex polygon ρ' with vertices in $\mathbf{Q}^2$, there is $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ such that $\rho(F) = \rho'$ ([Kwa93]). By Theorem 1.2.1 and the subsequent remark, we have $G \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ that is a lift of a pseudo-Anosov rel a finite set and has $\rho(G) = \rho'$.

Effective treatment of $h^{(v)}_{\text{top}}(F)$ and $h^{v}_{\text{erg}}(F)$ in the pseudo-Anosov case is made possible by the method of symbolic dynamics.

Proposition 3.2.1 ([FLP79]) *There exists a Markov partition $\mathcal{R} = \{R_1, ..., R_N\}$ for f. If G is a graph with the set of vertices $\{1, ..., N\}$ that has an edge from i to j whenever $f(R_i) \cap \text{int}(R_j) \neq \emptyset$, then the subshift of finite type (Λ, σ) associated with G (see sec. 2.2) is mixing and factors onto f. More precisely, for any $x = (x_i)_{i \in \mathbf{Z}} \in \Lambda$ the intersection $\bigcap_{i \in \mathbf{Z}} f^i(R_{x_i})$ consists of a single point, denoted by $p(x)$, and thus defined map $p : \Lambda \to \mathbf{T}^2$ has the following properties:*

(i) $p \circ \sigma = f \circ p$,

(ii) p is continuous and surjective,

(iii) p is finite-to-one,

(iv) p is 1-1 onto a topologically residual set consisting of all the points which full orbits never hit the boundary of the Markov partition.

The above proposition gives a rather complete description of the dynamics of f in terms of the shift σ. To keep track of the averages of the *observable* ϕ

in the symbolic model, we need to associate with every edge (i,j) of the graph G a vector weight $\psi_{i,j}$, reflecting the displacement in the universal cover that is inflicted on lifts of points in R_i that are mapped into R_j. One way to do that (taken from [Kwa92]) is as follows. Select a point z_i in each Markov box R_i, $i = 1, ...N$. If $(i,j) \in G$, then $f(R_i) \cap \text{int}(R_j) \neq \emptyset$, so there is a pair of corresponding lifts such that $F(\tilde{R}_i) \cap \text{int}(\tilde{R}_j) \neq \emptyset$; define $\psi_{i,j} := \tilde{z}_i - \tilde{z}_j$, where $\tilde{z}_i \in \tilde{R}_i$ and $\tilde{z}_j \in \tilde{R}_j$ are the lifts of z_i and z_j respectively. This definition does not depend on the choice of the lift $\tilde{R}_i$. Also, recall that the weights $\psi_{i,j}$ can be thought of as an observable $\psi : \Lambda \to \mathbf{R}^2$ given by $\psi((x_i)_{i\in\mathbf{Z}}) := \psi_{x_0,x_1}$.

Claim 3.2.1 *For $x = (x_i)_{i\in\mathbf{Z}} \in \Lambda$ and any $n \in \mathbf{N}$, we have*

$$\|S_n(f, \phi_F)(p(x)) - S_n(\sigma, \psi)(x)\| \leq 2\,\text{diam}(\mathcal{R}).$$

In fact, $\phi_F \circ p$ and ψ are cohomologous.

Proof. Having fixed a lift $\tilde{R}_{x_0}$ of R_{x_0}, let $\tilde{R}_{x_1}$ be the lift of R_{x_1} such that $\tilde{z_1} - \tilde{z_0} = \psi_{x_0,x_1}$, let $\tilde{R}_{x_2}$ be the lift of R_{x_2} such that $\tilde{z_2} - \tilde{z_1} = \psi_{x_1,x_2}$, etc. Thus defined sequence $\tilde{R}_{x_j}$ $(j = 0, 1, 2...)$, by the definition of ψ, has the following property: if $\tilde{q} \in \tilde{R}_{x_0}$ is a lift of $q = p(x)$, then $F^i(\tilde{q}) \in \tilde{R}_{x_j}$, $j = 0, 1, 2....$ The inequality follows because $S_n(f, \phi_F)(q) = F^n(\tilde{q}) - \tilde{q}$ and $S_n(\sigma, \psi)(x) = \tilde{z}_{x_n} - \tilde{z}_{x_0}$, whereas $\tilde{z}_j \in \tilde{R}_{x_j}$, $j \in \mathbf{N}$. Moreover, if $u(x) := \tilde{q} - \tilde{z}_{x_0}$, then $\phi_F \circ p - \psi = u \circ \sigma - u$, making $\phi_F \circ p$ and ψ cohomologous. Q.E.D.

Proposition 3.2.2 *Suppose that $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ descends to $f \in \mathcal{H}(\mathbf{T}^2)$ that is pseudo-Anosov rel a finite set. If (Λ, σ) is the corresponding subshift of finite* *type and ψ is the corresponding observable on Λ, then*

$$\rho(F) = \rho(f, \phi_F) = \rho(\sigma, \psi),$$

and, for any $v \in \mathrm{int}(\rho(F))$,

$$h_{\mathrm{top}}^{(v)}(F) = h_{\mathrm{erg}}^{v}(F) = h_{\mathrm{top}}^{(v)}(\sigma, \psi) = h_{\mathrm{erg}}^{v}(\sigma, \psi). \tag{3.4}$$

Moreover,

(i) there is a fully supported ergodic probability measure $\mu^{(v)}$ such that $h_{\mu^{(v)}}(f) = h_{\mathrm{erg}}^{v}(F)$ and $\int \phi_F \, d\mu^{(v)} = v$, $v \in \mathrm{int}(\rho(F))$,

(ii) $h_{\mathrm{top}}^{(v)}(F)$ is strictly concave and real analytic in v on $\mathrm{int}(\rho(F))$.

Remark. In fact $h_{\mathrm{erg}}^{v}(F) = h_{\mathrm{erg}}^{v}(\sigma, \psi)$ for all $v \in \rho(F)$.

Proof Proposition 3.2.2. Part (ii) follows from the main assertion by (ii) of Theorem 2.2.1. The statement about the rotation sets follows easily from (i) and (ii) of Proposition 3.2.1 coupled with Claim 3.2.1. To deal with the entropies, consider an invariant σ-ergodic probability measure ν. It pushes forward to $\mu := p_*(\nu)$, which is ergodic with respect to f; and $\int \psi \, d\nu = \int \phi_F \, d\mu$ because $\psi \circ p$ is cohomologous to ϕ_F. Moreover, if ν is fully supported, then we claim that

$$h_\mu(f) = h_\nu(\sigma). \tag{3.5}$$

This is standard. Observe that μ is also fully supported; and, by ergodicity, the full measure μ is carried on the invariant residual set on which p is 1-1. Hence, the measure theoretical systems $(\mathbf{T}^2, f, \mu)$ and (Λ, σ, ν) are isomorphic via p.

Now fix $v \in \operatorname{int}(\rho(F))$. Taking for ν the *Gibbs state* averaging ψ to v as in (v) of Theorem 2.2.1, we see that (3.5) yields

$$h_{\text{top}}^{(v)}(f,\phi_F) \geq h_{\text{erg}}^{v}(f,\phi_F) \geq h_\mu(f) = h_\nu(\sigma) = h_{\text{erg}}^{v}(\sigma,\psi) = h_{\text{top}}^{(v)}(\sigma,\psi).$$

On the other hand, since f is a factor of σ, $h_{\text{top}}^{(v)}(f,\phi_F) \leq h_{\text{top}}^{(v)}(\sigma,\psi)$ (see Proposition 2.1.4), so all the quantities above must be equal. This proves (3.4), and we see that $\mu^{(v)} := \mu$ is as stipulated by (i). Q.E.D.

Proof of the remark. Since f is a factor of σ we have $h_{\text{erg}}^{v}(f,\phi_F) \leq h_{\text{erg}}^{v}(\sigma,\psi)$ (see Proposition 2.1.4). To prove the opposite inequality, take a σ-ergodic probability measure ν with $\int \psi \, d\nu = v$ and set $\mu := p_*(\nu)$. We need to show that $h_\mu(f) \geq h_\nu(\sigma)$. Let $E \subset \mathbf{T}^2$ be the full orbit of the union of the stable boundaries and the unstable boundaries of the Markov boxes, denoted by $\partial\mathcal{R}^s$ and $\partial\mathcal{R}^u$ respectively. If $\mu(E) = 0$, then $h_\mu(f) = h_\nu(\sigma)$ (as in the proof of the proposition). Otherwise, by ergodicity, $\mu(E) = 1$. Since $\partial\mathcal{R}^s$ is forward invariant under f and $\partial\mathcal{R}^u$ is backward invariant under f, we have either $\mu(\partial\mathcal{R}^s) = 1$ or $\mu(\partial\mathcal{R}^u) = 1$. Say, we have the first case (for the other use f^{-1}). Then, $h_\mu(f) \leq h_{\text{top}}(f|_{\partial\mathcal{R}^s}) = 0$, where the vanishing of the topological entropy follows from the fact that $\partial\mathcal{R}^s$ is a collection of smooth arcs (cf. [Bow78b]) with the length strictly contracted by f ([FLP79]). Since p is finite-to-one $h_{\text{top}}(\sigma|_{p^{-1}(\partial\mathcal{R}^s)}) = 0$ (see Th. 17 in [Bow71]); thus $h_\nu(\sigma) \leq h_{\text{top}}(\sigma|_{p^{-1}(\partial\mathcal{R}^s)}) = 0$. Our claim follows. Q.E.D.

Using the *scaffold construction* (see the subsection 2.2.3 and Theorem 2.2.2), one can verify the following: for $v \notin \operatorname{int}(\rho(F))$, if there are any measures realizing $h^{(v)}_{\text{top}}(\sigma,\psi)$, then they are the Gibbs states for the smaller subshift of finite type $\sigma^{(F)}$ that corresponds to the *support* F of v in $\rho(F)$. If ν is such a Gibbs state realizing $h^{v}_{\text{erg}}(\sigma,\psi)$, then the argument above shows that $p_*(\nu)$ realizes $h^{v}_{\text{erg}}(f,\phi_F)$.

As a corollary of Theorem 2.2.2 we get the following.

Theorem 3.2.1 *Suppose that $f \in \mathcal{H}(\mathbf{T}^2)$ is pseudo-Anosov rel a finite set, F is its lift and $v \in \text{int}(\rho(F))$. For any $0 < \eta < 1$, there is a compact set $Q \subset \mathbf{T}^2$ invariant under f such that $\rho(F|_Q) = \{v\}$ and $h_{\text{top}}(f|_Q) > \eta h^{(v)}_{\text{erg}}(F)$.*

Proof of Theorem 3.2.1. Take $Q = p(K)$, where $K \subset \Lambda$ is as in Theorem 2.2.2. The fact that $\rho(F|_Q) = \rho(\sigma|_K, \psi)$ follows from Claim 3.2.1. The map p is finite-to-one, so we have $h_{\text{top}}(\sigma|_K) = h_{\text{top}}(f|_Q)$ (by Th. 17 in [Bow71] and the fact that entropy on a finite set must be zero). Q.E.D.

Our focus on the case $v \in \text{int}(\rho(F))$ is partially justified by the following fact. [1]

Proposition 3.2.3 *If $f \in \mathcal{H}(\mathbf{T}^2)$ is pseudo-Anosov rel a finite set and F is its lift, then $\text{int}(\rho(F)) \neq \emptyset$.*

Proof of Proposition 3.2.3. Passing to an iterate of F and perhaps postcomposing it with a deck transformation, we can assume that F has a fixed point $x = F(x) \in \mathbf{R}^2$. Fix a Markov partition for f. Denote by R the lift of the Markov box containing x. We will show that there are three non-collinear

[1] This is essentially due to Fried, compare Theorem H in [Fri82a].

vectors in $\rho(F)$. Think of $\mathbf{T}^2_2 := \mathbf{R}^2/(2\mathbf{Z})^2$, a 4-to-1 covering of $\mathbf{T}^2 = \mathbf{R}^2/\mathbf{Z}^2$. The map F factors to $f_2 : \mathbf{T}^2_2 \to \mathbf{T}^2_2$, which is also pseudo-Anosov rel a finite set. By transitivity of the unstable foliation ([FLP79]), the unstable leaf through the projection of x to $\mathbf{T}^2_2$ intersects the projections of $\text{int}(R)$, $\text{int}(R) + (1, 0)$, and $\text{int}(R) + (0, 1)$. It follows that, for some large $n \in \mathbf{N}$, we have $F^n(R) \cap \text{int}(R) + k_0 \neq \emptyset$, $F^n(R) \cap \text{int}(R) + k_1 + (0, 1) \neq \emptyset$ and $F^n(R) \cap \text{int}(R) + k_0 + (1, 0) \neq \emptyset$, where $k_i \in (2\mathbf{Z})^2$, $i = 1, 2, 3$. This puts k_0/N, $(k_1 + (1, 0))/N$ and $(k_2 + (0, 1))/N$ in the rotation set of F. These three vectors are not collinear because the vectors $k_1 - k_0 + (1, 0)$ and $k_2 - k_0 + (0, 1)$ have the corresponding coordinates of opposite parity and so can not be linearly dependent. Q.E.D.

3.3 Proof of the main result

We have $f \in \mathcal{H}(\mathbf{T}^2)$ with a lift $F \in \tilde{\mathcal{H}}(\mathbf{T}^2)$. Assume that $\text{int}(\rho(F)) \neq \emptyset$; otherwise, there is nothing to prove. Fix $v \in \text{int}(\rho(F))$ and take arbitrary $\epsilon > 0$. One can choose a convex polygon $\rho' \subset \text{int}(\rho(F))$ with vertices in $\mathbf{Q}^2$ so that $A(\rho') = A(\rho(F))$, $I(\rho') = I(\rho(F))$ and $\tau_{\rho'}(v) \geq \tau_{\rho(F)}(v) - \epsilon$. By Theorem 1.2.1, there is a finite set P (a union of *primitive* periodic orbits) and a map $g \in \mathcal{H}(\mathbf{T}^2)$ such that: f and g are isotopic rel P, g is pseudo-Anosov rel P, g has a lift $G \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ with $\rho(G) \supset \rho'$.

The following theorem asserts that the dynamics of g are fully reflected by f. It is a variation on Handel's global shadowing result in [Han85] that is most closely related to Theorem 3.2 in [Boy10]. (Also, this is the only place where we are forced to use smoothness of f. One should be able to remove this obstacle.)

Theorem 3.3.1 *If $f \in \mathcal{H}(\mathbf{T}^2)$ is a diffeomorphism, P is a finite invariant set and $g : \mathbf{T}^2 \to \mathbf{T}^2$ is pseudo-Anosov rel P and isotopic to f rel P, then there is a compact invariant set $Y \subset \mathbf{T}^2$ and a surjection $h : Y \to \mathbf{T}^2$ that is homotopic to the inclusion map $Y \hookrightarrow \mathbf{T}^2$ and semi-conjugates $f|_Y$ to g (i.e. $h \circ f|_Y = g \circ h$).*

The map g and the factor map h from the theorem can be lifted to $G \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ and an equivariant $H : \pi^{-1}(Y) \to \mathbf{R}^2$ respectively, so that $H \circ F|_{\pi^{-1}(Y)} = G \circ H$. In particular, for $x \in Y$, we have

$$\|S_n(f, \phi_F)(x) - S_n(g, \psi_G \circ h)(x)\| \le D := 2 \sup_{y \in \pi^{-1}(Y)} \|H(y) - y\|, \; n \in \mathbf{N}. \quad (3.6)$$

Indeed, if $\tilde{x}$ is a lift of x, the left side is equal to

$$\|(F^n(\tilde{x}) - \tilde{x}) - (G^n \circ H(\tilde{x}) - H(\tilde{x}))\| = \|F^n(\tilde{x}) - H \circ F^n(\tilde{x}) + H(\tilde{x}) - \tilde{x}\|.$$

W need the following easy corollaries, which proof we postpone.

Corollary 3.3.1 *Under the assumptions of Theorem 3.3.1, we have:*

(i) for any compact invariant $Q \subset \mathbf{T}^2$, the set $W := h^{-1}(Q)$ is compact invariant, $\rho(F|_W) = \rho(G|_Q)$ and $h_{\text{top}}(f|_W) \ge h_{\text{top}}(g|_Q)$;

(ii) for any $\mu \in \mathcal{E}(\mathbf{T}^2, g)$, there is $\nu \in \mathcal{E}(\mathbf{T}^2, f)$, such that $\mu := h_*(\nu)$, $\int \phi_G \, d\mu = \int \phi_F \, d\nu$ and $h_\nu(f) \ge h_\mu(g)$;

(iii) for any $v \in \mathbf{R}^d$, $h^{(v)}_{\text{top}}(F) \ge h^{(v)}_{\text{top}}(G)$ and $h^v_{\text{erg}}(F) \ge h^v_{\text{erg}}(G)$;

(iv) $\rho(G) \subset \rho(F)$.

Proof of Theorem 3.1.1. To prove (3.1) of the theorem, observe that, by (ii) of Theorem 2.2.1, $h_{\rm top}^{(v)}(G)$ is a concave function of v; thus, from (3.3) and (iii) of Proposition 2.1.1, we have

$$h_{\rm top}^{(v)}(G) \geq \sup\{h_{\rm top}^{(x)}(G) :\ x \in \mathbf{R}^2\} \cdot \tau_{\rho(G)}(v) = h_{\rm top}(g) \cdot \tau_{\rho(G)}(v).$$

Using Theorem 1.1.1 and (iii) of Corollary 3.3.1, we obtain

$$h_{\rm top}^{(v)}(F) \geq h_{\rm top}^{(v)}(G) \geq \max\{C \ln_+ A(\rho(G)), C'/I(\rho(G))\} \cdot \tau_{\rho(G)}(v),$$

and the analogous inequality for $h_{\rm erg}^{v}$. Because $\rho' \subset \rho(G) \subset \rho(F)$ and $\tau_{\rho(G)}(v) \geq \tau_{\rho(F)}(v) - \epsilon$ (where ϵ is arbitrary), the inequality (3.1) follows. The assertion (3.2) is an immediate consequence of Theorem 3.2.1 via (i) of Corollary 3.3.1. Q.E.D.

Proof of Corollary 3.3.1. Part (i): the equality of the rotation sets follows from (3.6); the entropy assertion is standard ([Wal82]). Part (iii): use (3.6) to note that, in calculating $h_{\rm top}^{(v)}(F)$ and $h_{\rm erg}^{v}(F)$, one can use $\phi_G \circ h$ in the place of ϕ_F and then recall the quotient rules — Proposition 2.1.2 and Proposition 2.1.4. The part (ii) can be extracted from the proof of Proposition 2.1.4. The part (iv) is immediate from (3.6). (It also follows from (iii) by (ii) of Proposition 2.1.1.) Q.E.D.

Proof of Theorem 3.3.1. We will show how to reduce the result to Theorem 3.2 (and its proof) in [Boy10].

We start with a construction taken from [Bow78a] and first used in a similar context in [LM91]. Let M be the compact surface with boundary obtained by blowing up the points of P to small circular discs and then by removing its interiors. Denote by $c : M \to \mathbf{T}^2$ a continuous map that collapses the boundary circles of M to the corresponding points of P and is 1-1 elsewhere. Using the derivative of f at points of P, one constructs $f_1 : M \to M$ that factors to f via c, i.e. $c \circ f_1 = f \circ c$. Now, let $g_1 : M \to M$ be Boyland's *condensed homeomorphism* in the isotopy class of f_1 (see Case 5, p. 13 in [Boy10]). In our case, g_1 is just a pseudo-Anosov map with some standardized behavior at the boundary — it is called (in [Boy10]) *boundary-adjusted pA* map. In particular, g_1 factors to a map that is pseudo-Anosov rel P and (after perhaps some conjugation) coincides with g, i.e. $c \circ g_1 = g \circ c$. Theorem 3.2 in [Boy10] yields a compact invariant set $Y_1 \subset M$ and a surjective map $h_1 : Y \to M$ (homotopic to the inclusion) such that $h_1 \circ f_1 = g_1 \circ h_1$. The map h_1 needs not to be continuous, however $c \circ h_1$ is.

Put $Y = c(Y_1)$. Because the map f_1 was obtained as a blow up of f, it is a tautology to say that it is uniformly continuous on $\mathrm{int}(M)$ in the metric induced from $\mathbf{T}^2$ by $c : \mathrm{int}(M) \to \mathbf{T}^2$. Now, one has to trace word by word Boyland's proof of uniform continuity of $c \circ h_1$ on $\mathrm{int}(M)$ while substituting in the place of *the standard metric on* $\mathrm{int}(M)$ the one induced by c. As a result, we get uniform continuity of $c \circ h_1$ in this new metric. It follows that $h := c \circ h_1 \circ c^{-1} : Y \to \mathbf{T}^2$ is a well defined and continuous map. This map has all the required properties. Q.E.D.

3.4 Example

We calculate $h_{\text{top}}^{(v)}$ numerically for a concrete example of a pseudo-Anosov map. This map is derived from a map in $\mathcal{H}(\mathbf{T}^2)$ via the method from Section 1.2 coupled with the Bestvina-Handel algorithm ([BH92]). Our discussion also illustrates how results of Section 3.2 can directly bare on $h_{\text{top}}^{(v)}$ when dealing with a single explicitly defined map.

Let $H, V \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ be given by

$$H(x, y) := (x + |\sin(\pi y)|, y),$$

$$V(x, y) := (x, y + |\sin(\pi x)|),$$

and let h and v be the corresponding maps in $\mathcal{H}(\mathbf{T}^2)$. Consider the composition $f := v \circ h \in \tilde{\mathcal{H}}(\mathbf{T}^2)$ and its lift $F := V \circ H \in \tilde{\mathcal{H}}(\mathbf{T}^2)$. Note that $F(0,0) = (0,0)$, $F(1/2, 0) = (1/2, 1)$, $F(0, 1/2) = (1, 1/2)$, and $F(1/2, 1/2) = (3/2, 3/2)$. Thus the square $Q := \{(x, y) : \ 0 \leq x, y \leq 1\}$ is contained in $\rho(F)$. In fact, we have $\rho(F) = Q$ because, by writing $\phi_F = \phi_G \circ h + \phi_H$, one can verify that the range of the displacement function ϕ_F sits in Q.

Now, let $P \subset \mathbf{T}^2$ consist of the three fixed points of f that have $(0,0)$, $(1/2, 0)$, $(0, 1/2)$ for their lifts. The isotopy class of f rel P is represented by a map g that is pseudo-Anosov rel P (cf. Theorem 1.2.1). To find out about g, we use the algorithm by Bestvina and Handel.

The surface $\mathbf{T}^2 \backslash P$ deformation retracts to its *spine* consisting of the union of the oriented loops $\alpha, \beta, \gamma, \delta$, marked on Figure 3.2. To record the isotopy classes of h, v, f, we postcompose them with the retraction and look at the resulting transformation of the spine into itself. On Figure 3.2, the dotted lines indicate the images of γ and δ under the resulting action on the spine for h. Alternatively, we write $h_{\#} : \gamma \mapsto \overline{\beta}\gamma\alpha$, $\delta \mapsto \overline{\beta}\delta\alpha$, $\alpha \mapsto \beta$, $\beta \mapsto \beta$, where the bar over an oriented path switches the orientation.

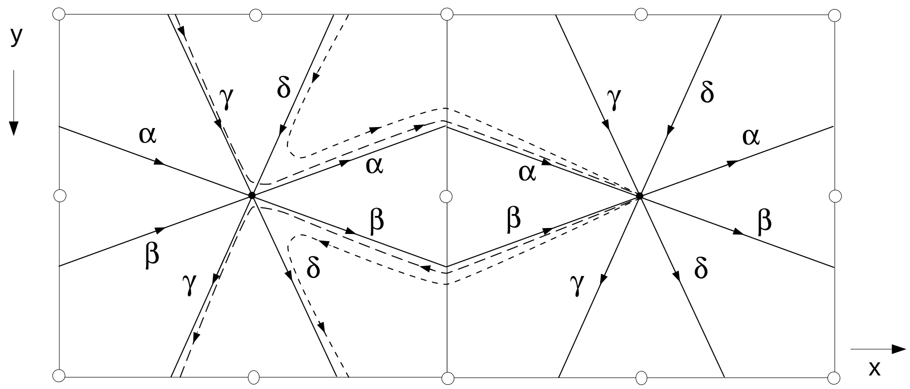

Figure 3.2: The map $h_{\#}$ on the spine.

Similarly, for v, we get $v_{\#} : \alpha \mapsto \overline{\delta}\alpha\gamma$, $\beta \mapsto \overline{\delta}\beta\gamma$, $\gamma \mapsto \gamma$, $\delta \mapsto \delta$. For the composition $f = v \circ h$, we get $f_{\#} : \alpha \mapsto \overline{\delta}\alpha\gamma$, $\beta \mapsto \overline{\delta}\beta\gamma$, $\gamma \mapsto \overline{\gamma}\overline{\beta}\delta\gamma\overline{\delta}\alpha\gamma$, $\delta \mapsto \overline{\gamma}\overline{\beta}\delta\alpha\gamma$.

It turns out that action of $f_{\#}$ on the spine is not *efficient*; this means that, as we apply $f_{\#}$ repeatedly to arcs making up the spine, we get paths that *backtrack* (see [BH92] for the definition). Bestvina and Handel describe *moves* that transform the map $f_{\#}$ and the spine leading to an *efficient* map in a finite number of steps. The implementation of the algorithm, in our case, involves nine steps and is more tedious then instructive, so we skip it. The result is a graph contained in $\mathbf{T}^2 \setminus P$ (see Figure 3.3), which is a deformation retract of $\mathbf{T}^2 \setminus P$, and a transformation g_1 of this graph that is *efficient.*

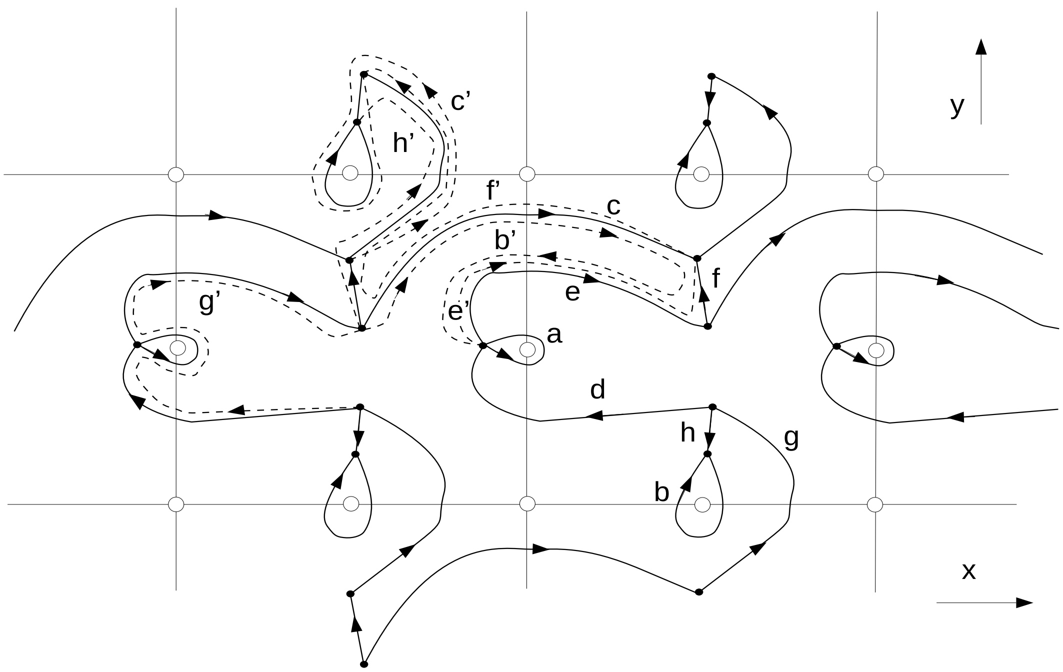

Figure 3.3: The efficient map g_1.

The marked arcs are mapped by g_1 as follows: $a \mapsto a' = a$, $b \mapsto b' = b$, $c \mapsto c' = ghb\overline{h}$, $d \mapsto d' = c\overline{f}\overline{e}$, $e \mapsto e' = ef$, $f \mapsto f' = \overline{c}fg$, $g \mapsto g' = dae$, $h \mapsto h' = fgh$. As described in [BH92], the graph can be smoothed into a train track that carries a pseudo-Anosov map g which *deforms* to g_1. The Markov boxes for g can be obtained by thickening the arcs c, d, e, f, g, h to rectangles.

Actually, Figure 3.3 depicts not only g but a certain lift G of g. To determine the rotation set $\rho(G)$, we read off the figure the displacements under G:

$$
\begin{aligned}
a &\mapsto a \\
b &\mapsto b + (-1,1) \\
c &\mapsto g + (-1,1)\ h + (-1,1)\ b + (-1,1)\ \overline{h} + (-1,1), \\
d &\mapsto c\ \overline{f}\ \overline{e},
\end{aligned}
$$

$$
\begin{aligned}
e &\mapsto e\ f, \\
f &\mapsto \overline{c}\ f + (-1,0)\ g + (-1,1), \\
g &\mapsto d + (-1,0)\ a + (-1,0)\ e + (-1,0), \\
h &\mapsto f + (-1,0)\ g + (-1,1)\ h + (-1,1).
\end{aligned}
$$

Tying this with the material in Section 2.2, we draw the weighted *transition graph* of the corresponding subshift of finite type — see Figure 3.4. (For clarity we skip weights $(0,0)$.)

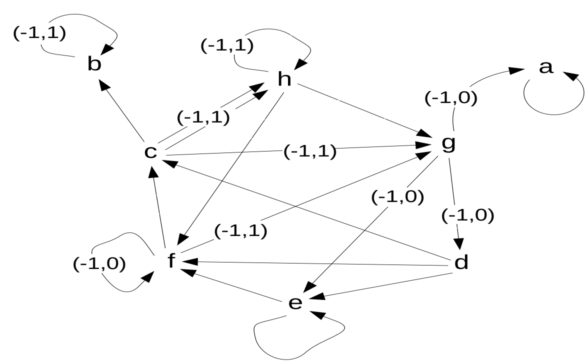

Figure 3.4: The transition graph.

It is apparent from this graph that (by Proposition 3.2.2 and (i) of Proposition 2.2.1)

$$\rho(G) = \mathrm{conv}\{(0,0), (-1,0), (-1,1)\}.$$

Note that the map F (that we started with) is equivariantly isotopic rel $\pi^{-1}(P)$ to $G + (1,0)$, not G. By Corollary 3.3.1, we have

$$h_{\mathrm{top}}^{(v)}(F) \geq h_{\mathrm{top}}^{(v+(1,0))}(G), \;\; v \in \mathbf{R}^2.$$

To calculate $h_{\text{top}}^{(v)}(G)$, we extract from the graph its *transfer matrix*, see below. (The columns and rows, counted starting at the upper left corner, correspond to c, d, e, f, g, h. We skip a and b because they do not correspond to any Markov boxes.)

$$A := \begin{pmatrix} 0 & 0 & 0 & 0 & xy & 2xy \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & x & xy & 0 \\ 0 & x & x & 0 & 0 & 0 \\ 0 & 0 & 0 & x & xy & xy \end{pmatrix}, \tag{3.7}$$

where $x = \exp(-s)$ and $y := \exp t$, $(s,t) \in \mathbf{R}^2$. The characteristic polynomial of A is $p(x,y,z) := \det(A - zI)$ and can be found to be equal to

$$\begin{aligned} p(x,y,z) &= x^4y^2 - x^3yz - x^3y^2z - x^4y^2z + x^2yz^2 + x^3yz^2 + x^3y^2z^2 \\ &- 6x^2yz^3 + xz^4 + xyz^4 + x^2yz^4 - z^5 - xz^5 - xyz^5 + z^6. \end{aligned}$$

Given a rotation vector $(u,v) \in \text{int}(\rho(G))$, a way to find $h_{\text{top}}^{(u,v)}(G)$ is through solving the following system of equations

$$p(x,y,z) = 0, \tag{3.8}$$

$$vz\frac{\partial p}{\partial z}(x,y,z) = x\frac{\partial p}{\partial x}(x,y,z), \tag{3.9}$$

$$uz\frac{\partial p}{\partial z}(x,y,z) \quad = \quad -y\frac{\partial p}{\partial y}(x,y,z). \tag{3.10}$$

Provided the right solution (x, y, z) is taken, we have

$$h_{\text{top}}^{(v,u)}(G) = \ln(z \cdot x^v \cdot y^{-u}). \tag{3.11}$$

Indeed, by (iii) of Theorem 2.2.1, $p_{\text{top}}^{(s,t)} = \ln z$, where z is the leading positive root of the first equation. By (iv) of Theorem 2.2.1,

$$-h_{\text{top}}^{(v,u)}(G) = \sup\{sv + tu - \ln z : \ s, t \in \mathbf{R}\}, \tag{3.12}$$

where the supremum is achieved when $v = \partial(\ln z)/\partial s = -x(\partial z/\partial x)/z$ and $u = \partial(\ln z)/\partial t = y(\partial z/\partial y)/z$. This was used to obtain the two last equations of the system by taking partial derivatives of the first one with respect to s, for one, and with respect to t, for the other.

The system (3.8) may have many solutions. Continuation of a known solution is a way to select the right one. For example, one can start with $x = 1$ and $y = 1$ finding the leading positive root of $p(1, 1, z)$ to be approximately 2.6103. Thus

$$h_{\text{top}}(g) = \max_{(v,u)} h_{\text{top}}^{(v,u)}(G) = \ln 2.6103.$$

The corresponding rotation vector, calculated from the measure of maximal entropy (see (v) of Theorem 2.2.1), is

$$(v_{max}, u_{max}) = (-2/3, 1/3).$$

Changing (u, v) away from (v_{max}, u_{max}), we may follow $h_{\text{top}}^{(v,u)}(G)$. In this way, using *Maple*, we obtained for example $h_{\text{top}}^{(-0.42,0.33333)} \approx 1.74690$.

Chapter 4

Rotation sets for the resonantly kicked linear oscillator

As we have seen in the previous chapters, one can infer a lot about dynamics of a torus map whose rotation set has nonempty interior. We also know that there is an abundance of such maps. (See the example in the section 1.1, Proposition 3.2.3, also [Kwa92].) However, our examples so far were constructed ad hoc to get the rotation set with specific properties. In this chapter we consider a model arising from plasma physics.

4.1 The physical context

Consider a particle of mass M and charge e moving in $\mathbf{R}^3$ filled with a uniform vertical magnetic field of induction B_0. The force exerted by the field on the particle is proportional to the vector product of its velocity and the field. Thus the trajectories are curled into spirals with circular projections on the horizontal plane. More precisely, if $x, y, z \in \mathbf{R}$ are the coordinates of the particle (z being the vertical one), then the Newtonian equations of motion are

$$\begin{aligned} M\ddot{x} &= \frac{eB_0}{c}\dot{y}, \\ M\ddot{y} &= -\frac{eB_0}{c}\dot{x}, \\ M\ddot{z} &= 0, \end{aligned}$$

where c is the speed of light. The vector of horizontal velocity $(\dot{x}, \dot{y})$ rotates with the *Larmor frequency* $\omega = eB_0/cM$. The three momenta: $p_z := M\dot{z}$, $p_y := M\dot{y} + M\omega x$, and $p_x := M\dot{x} - M\omega y$, are conserved. This simple behavior can change radically in the presence of an external electric field in the system. We will restrict our attention to the case when the electric field is aligned with the x-axis and does not depend on y, z coordinates (cf. [ZZRSC86b]). These assumptions reduce the number of essential degrees of freedom to one since the momenta p_z, p_y continue to be integrals of the motion. Furthermore, we will require that the field is a "tight" wave packet with amplitude $E(x,t) = \sum_{n=-\infty}^{\infty} TE(x)\delta(t - Tn)$, where $T > 0$ is fixed, $E(x)$ is $2\pi/k$-periodic in x, and $\delta(\cdot)$ is the Dirac delta function. The meaning of the formula is that the packet interacts with the particle by boosting its x-velocity by $\frac{e}{M}TE(x)$ every T seconds. Thus, if $(\dot{x}_n, \dot{y}_n)$ is the horizontal velocity prior to the kick at the time nT $(n \in \mathbf{Z})$, then

$$(\dot{x}_{n+1}, \dot{y}_{n+1}) = R_\alpha \left(\dot{x}_n + \frac{eT}{M} E\left(\frac{p_y - M\dot{y}_n}{M\omega}\right),\ \dot{y}_n\right) \tag{4.1}$$

where R_α is the rotation by the angle $\alpha = \omega T$. In this way, the analysis of the physical model reduces to investigation of dynamics on $\mathbf{R}^2$ generated by the mapping

$$F : (\dot{x}_n, \dot{y}_n) \mapsto (\dot{x}_{n+1}, \dot{y}_{n+1}).$$

This reduction was first carried out in [ZZRSC86b].

What makes the map F interesting is that (even for small amplitudes of E) numerical experiments and physical heuristic suggest that it exhibits trajectories escaping towards the infinity with a nonzero "average acceleration". A significant physical consequence is an unbounded diffusion of plasma particles called also *stochastical heating* ([ZZRSC86b]): the particles increase their kinetic energy (heat up) by exploiting the energy of the electric field.

In general, there are no rigorous arguments for the existence of the above behavior. A lot of attention is thus devoted to resonant cases when α is rational ([ZZRSC86b, CSUZ87, ZZRSC86a]). Among those, the cases of choice are $\alpha = \pi/2$, $\alpha = \pi/3$ or $\alpha = 2\pi/3$ — the only possibilities if we require that some positive iterate of F is a doubly-periodic map of $\mathbf{R}^2$, i.e. it commutes with some faithful $\mathbf{Z}^2$-action on $\mathbf{R}^2$. Moding out by the $\mathbf{Z}^2$-action yields a torus diffeomorphism isotopic to the identity. In terms of the torus map, the presence of *stochastical heating* corresponds to a nontrivial rotation set. Indeed, in view of (4.1), the rotation set can be interpreted as the totality of asymptotic average accelerations exhibited by the particles (with fixed momentum p_y).

In what follows we will analyze closer the case of $\alpha = 2\pi/3$. (The cases $\alpha = \pi/3, \pi/2$ can be treated analogously.) Our modest goal is to see that, for rather high values of nonlinearity, the rotation set is indeed nontrivial. We calculate it for special cases and provide crude estimate from below in the general case. In the case when ϕ is "saw-tooth-like", we prove monotonicity of the rotation set as a function of the amplitude of E (also for rather large amplitudes). The more delicate part of our discussion is greatly simplified by formulation of the problem in terms of a second order recurrence equation. This simple reduction has not been taken up in this context before. We close our discussion by mentioning a toy model in statistical mechanics governed by the recurrence equation.

4.2 The mapping and its rotation set

The rotation $R_{2\pi/3}$ is linearly conjugated to the map

$$L : (u, v) \mapsto (u, -u - v), \ \ u, v \in \mathbf{R}.$$

By choosing the conjugacy that sends $\dot{x}$-axis to u-axis and by rescaling the variables, one conjugates F (as prescribed by (4.1)) to G given by

$$G : (u, v) \mapsto L(u + \phi(v), \ v) = (v, -u - v - \phi(v)), \tag{4.2}$$

where $\phi(v) = eT/M\omega k \cdot E(-kv + \frac{p_y}{M\omega})$ is 2π-periodic. The third iterate of G commutes with the translations of $\mathbf{R}^2$ by vectors in $(2\pi\mathbf{Z})^2$, and so it factors to a torus map

$$g_3 : \mathbf{T}^2 \to \mathbf{T}^2.$$

Even though G and g_3 depend on the choice of ϕ, we suppress this dependence in the notation since it will be always clear which ϕ we are talking about. From now on, if we do not state otherwise, ϕ is an arbitrary fixed continuous function of period 2π. However, we will remember that ϕ naturally comes embedded in the two parameter family

$$\phi_{K,\eta}(x) := K\phi(x+\eta), \; x \in \mathbf{R}, \tag{4.3}$$

where $\eta \in \mathbf{R}$ and $K > 0$ correspond to the momentum p_y and the strength of the forcing respectively.

Fact 4.2.1 *For any $\eta \in \mathbf{R}$, replacing $\phi(\cdot)$ with a shifted function $\phi(\cdot + \eta) + 3\eta$ is equivalent to conjugating G by a translation $(u, v) \mapsto (u - \eta, v - \eta)$. In particular, the rotation set $\rho(G^3)$ of G^3 is unaltered by the shift.*

For $r \in \mathbf{R}$, we will denote by $\mathrm{Hex}(r)$ the hexagon with vertices $\pm(r, 0)$, $\pm L(r, 0) = \pm(0, -r)$, $\pm L^2(r, 0) = \pm(-r, r)$. The following are the most basic properties of the rotation set $\rho(G^3)$.

Proposition 4.2.1 *If G is defined by (4.2), then*

(i) $\rho(G^3)$ *has a three-fold symmetry, namely* $\rho(G^3) = L(\rho(G^3)) = L^2(\rho(G^3))$;

(ii) *if* ϕ *is odd, then* $\rho(G^3) = -\rho(G^3)$;

(iii) $\rho(G^3)$ *is contained in* $\mathrm{Hex}(\sup \phi - \inf \phi)$.

Proof of (i). Let $Q : \mathbf{R}^2 \to \mathbf{R}^2$ be given by

$$(u, v) \mapsto (u + \phi(v), \; v),$$

so that $G = L \circ Q$. The displacement of a point $p \in \mathbf{R}^2$ under $n+1$ iterates of G is

$$G^{n+1}(p) - p = L \circ Q \circ G^n(p) - L \circ G^n(p) + L(G^n(p) - p) + Lp - p.$$

The first and the last differences on the right side are bounded, so $(1/(n+1))$ $(G^{n+1}(p)-p)$ is asymptotically equal to $L\left((1/n)\left(G^n(p)-p\right)\right)$, as $n\to\infty$; and thus $\rho(G^3)=L(\rho(G^3))$. Q.E.D.

Proof of (ii). Just note that if ϕ is odd, then Q commutes with central symmetry $p\mapsto -p$, and so does the whole map G. Q.E.D.

Proof of (iii). Write the displacement of $p\in\mathbf{R}^2$ under G^3 as follows

$$G^3(p)-p=L\left(Q\circ G^2(p)-G^2(p)\right)+L^2\left(Q\circ G(p)-G(p)\right)+L^3\left(Q(p)-p\right).$$

The arguments of $L, L^2, L^3=\mathrm{id}$, above, belong to a segment I with endpoints $(\inf\phi,0)$ and $(\sup\phi,0)$, so $G^3(p)-p$ sits inside the algebraic (Minkowski) sum $L(I)+L^2(I)+I$, which equals $\mathrm{Hex}(\sup\phi-\inf\phi)$. Q.E.D.

Proposition 4.2.1, (ii) can be sharp, as it is shown by the example below. For convenience, we note the following first.

Fact 4.2.2 *If $(u,v)\in\mathbf{R}^2$ satisfies $\phi(-v-u-\phi(v))=\phi(v)$ and $\gamma:=\phi(v)-\phi(u)\in 2\pi\mathbf{Z}$, then (u,v) is a lift of a fixed point of g_3 and $\rho(G^3,(u,v))=(0,\gamma)$.*

Proof. We apply G to (u,v) three times

$$\begin{aligned}(u,v) &\mapsto (v,-u-v-\phi(v))\\ &\mapsto (-u-v-\phi(v),u+\phi(v)-\phi(-v-u-\phi(v)))\\ &= (-u-v-\phi(v),u)\mapsto(u,v+\phi(v)-\phi(u)).\end{aligned}$$

Q.E.D.

Actually, from the proof, one can see that the inverse of the fact is also true.

Example. Consider any ϕ normalized so that $\sup\phi = K$ and $\inf\phi = -K$ (see Fact 4.2.1). By Fact 4.2.2, if there are u_1, u_2, v_1, v_2 such that $\phi(v_1) = \phi(v_2) = \phi(-v_1 - u_1) = K$ and $\phi(u_1) = \phi(u_2) = \phi(-v_2 - u_2) = -K$, we conclude that both $\pm(0, 2K)$ are in the rotation set provided also $K = 0 \pmod{\pi}$. By (i) of Proposition 4.2.1, the rotation set is then actually equal to $\mathrm{Hex}(2K)$. It is not difficult to draw a graph of ϕ with the required placement of the maxima and the minima. (Use for example $u_1 = -\pi/2$, $u_2 = -\pi/6$, $v_1 = \pi/6$, $v_2 = \pi/2$.) (One can see that any suitable ϕ has to attain one of its global extrema at least three times per period.)

Another consequence of Fact 4.2.2 is the following very crude lower bound on the rotation set.

Proposition 4.2.2 *If $l := \sup\phi - \inf\phi \geq 24\pi$, then the triangle with vertices $L^i(0, l/2)$, $i = 1, 2, 3$, is contained in the rotation set $\rho(G^3)$.*

Proof. Using Fact 4.2.1, we can shift the graph of ϕ so that there are v_0, u_0 with $\phi(v_0) = \sup\phi \geq 12\pi$ and $\phi(u_0) = \inf\phi \leq -12\pi$. Now, as we push v continuously away from v_0 (in either direction), before $\phi(v)$ drops below $\sup\phi/2$, the expression $-2v - \phi(v)$ changes by at least $-4\pi + 6\pi = 2\pi$, thus sweeping the whole period. In particular, $\phi(-2v - \phi(v))$ sweeps, in the process, an interval containing $[\inf\phi, \inf\phi/2]$; and so, by the intermediate value theorem, there is v such that $l/2 = \sup\phi/2 - \inf\phi/2 \leq \phi(v) - \phi(-2v - \phi(v)) \in 2\pi\mathbf{Z}$. Set $u = -2v - \phi(v)$, and see that (u, v) satisfies the assumption of Fact 4.2.2, thus placing $(0, \phi(v) - \phi(u))$ in the rotation set. By (i) of Proposition 4.2.1, $L^i(0, \phi(v) - \phi(u))$, $i = 1, 2$, are in the rotation set as well. Q.E.D.

4.3 The recurrence, higher period orbits, monotonicity

A convenient way of viewing a trajectory of a point $(u, v) \in \mathbf{R}^2$ under G is to label it as a sequence $(x_k)_{k\in\mathbf{Z}}$ so that $x_0 = u$, $x_1 = v$ and $(x_k, x_{k+1}) = G(x_{k-1}, x_k)$, for $k \in \mathbf{Z}$. It is then characterized by the following recurrence equation

$$x_{k-1} + x_k + \phi(x_k) + x_{k+1} = 0, \ k \in \mathbf{Z}. \tag{4.4}$$

For $n \in 3\mathbf{N}$, a periodic orbit of G of period n thus corresponds to $x = (x_k)_{k=0}^{n-1} \in \mathbf{R}^n$ satisfying (4.4) with a cyclic indexing modulo n. For $\phi = 0$, this is when every point is fixed by G^3, the two-dimensional space of solutions to (4.4) is conveniently spanned by

$$\begin{aligned} e &= (1, -1, 0, 1, -1, 0, ..., 1, -1, 0), \\ Se &= (0, 1, -1, 0, 1, -1, ..., 0, 1, -1), \\ S^2e &= (-1, 0, 1, -1, 0, 1, ..., -1, 0, 1), \end{aligned}$$

where $S : \mathbf{R}^n \to \mathbf{R}^n$ shifts the coordinates cyclically to the right, and $S^3e = e$.

The orbits of G corresponding to periodic orbits of g_3 on the torus with nonzero rotation vector come from $(x_k)_{k\in\mathbf{Z}}$ satisfying (4.4) together with the periodicity condition $x_{k+n} - x_k = v_{k \pmod 3}$, where $v_1, v_2, v_3 \in 2\pi\mathbf{Z}$, $n \in 3\mathbf{N}$. To search for them, we may use the following scheme. (Below $\langle x, y\rangle := \sum_{i=1}^{n} x_i y_i$, for $x, y \in \mathbf{R}^n$.)

Lemma 4.3.1 *For $n \in 3\mathbf{N}$, if $z = (z_k)_{k=0}^{n-1} \in \mathbf{R}^n$ satisfies*

$$z_{k-1} + z_k + \phi(z_k) + z_{k+1} = -\xi_k, \ 0 \le k \le n-1, \tag{4.5}$$

where $\xi \in (2\pi\mathbf{Z})^n$ and the indexing is cyclic, then there is $\rho \in (2\pi\mathbf{Z})^{\mathbf{Z}}$ satisfying

$$\rho_{k-1} + \rho_k + \rho_{k+1} = \xi_{k \pmod n}, \ k \in \mathbf{Z}, \tag{4.6}$$

and the sequence $x_k := z_{k \pmod n} + \rho_k$ is a solution to (4.4). The displacement over a period is given by

$$x_{k+n} - x_n = \langle \xi, S^k e\rangle, \ k \in \mathbf{Z}, \tag{4.7}$$

so the corresponding rotation vector for G^3 is $(1/n)\big(\langle \xi, e\rangle, \langle \xi, Se\rangle\big)$. Moreover, one can obtain in this way all of $(x_k) \in \mathbf{R}^{\mathbf{Z}}$ corresponding to orbits of G that cover periodic orbits of g_3 on the torus.

Proof. Solving (4.6) is straightforward: fix $\rho_0, \rho_1 \in 2\pi\mathbf{Z}$ and figure out the rest of ρ_i's successively. That $x_k = z_{k \pmod n} + \rho_k$ satisfy (4.4) follows from putting together (4.5) and (4.6). To calculate the displacement, replace in (4.4) k with $k+1$ and subtract unaltered (4.4) to get

$$x_{k+3} - x_k = \phi(x_{k+1}) - \phi(x_{k+2}) = \phi(z_{k+1 \pmod n}) - \phi(z_{k+2 \pmod n}). \tag{4.8}$$

Sum over the range $k, k+3, ..., k+n$ to obtain

$$x_{k+n} - x_k = \Big\langle \big(\phi(z_{j \pmod n})\big)_{j=k+1}^{k+n}, S^k e \Big\rangle = \langle \xi, S^k e\rangle,$$

where we used the fact that $\left\langle (z_{j-1 \pmod n} + z_{j \pmod n} + z_{j+1 \pmod n})_{j=k+1}^{k+n}, S^k e \right\rangle = 0$. Now suppose that $(x_k)_{k \in \mathbf{Z}}$ covers a periodic orbit of period $n \in 3\mathbf{N}$. Then $x_{k+n} - x_k =: v_{k \pmod 3} \in 2\pi \mathbf{Z}$, and, because both x_k's and the shifted sequence x_{k+n}'s satisfy (4.4), the sequence $(v_{k \pmod 3})_{k \in \mathbf{Z}}$ obeys the linear equation $v_{k-1 \pmod 3} + v_{k \pmod 3} + v_{k+1 \pmod 3} = 0$. This makes $\xi_{k \pmod n} := -x_{k-1 \pmod n} - x_{k \pmod n} - \phi(x_{k \pmod n}) - x_{k+1 \pmod n}$, $k \in \mathbf{Z}$, well defined. To satisfy (4.5), set $z_k := x_k$, $k = 0, ..., n-1$. Q.E.D.

Solving (4.5) for large n is a daunting task regardless of our choice of ϕ. We thus have little hope that the rotation set can be calculated exactly for a continuum of values of the parameters K and η in (4.3). A way to somehow bridge the particular cases for which the rotation set is nontrivial leads through monotonicity results as the one below.

We will call ϕ *unimodal*, if it has only two intervals of monotonicity over its minimal period. We will refer to the endpoints of the maximal intervals of monotonicity as *critical points of* ϕ.

Proposition 4.3.1 *For a piecewise-smooth unimodal and odd* ϕ *with* $D := \inf_x |\phi'(x)| \geq 6$, *the rotation set of* G^3 *corresponding to* $K\phi$ *is nondecreasing in* K *for* $K \in [1, +\infty)$. *Actually, all periodic orbits of the torus map can be continued as* K *increases.*

The proof is based on the following "closing lemma".

Lemma 4.3.2 *Fix* $\epsilon > 0$ *and suppose that* $z = (z_k)_{k=0}^{n-1} \in \mathbf{R}^n$ *satisfies*

$$z_{k-1} + z_k + K\phi(z_k) + z_{k+1} = -\xi_k + \epsilon_k,$$

where $\epsilon_k \in (-\epsilon, \epsilon)$*. If, for some open intervals* I_k *that are free of the critical points of* ϕ *and contain* z_k*,* $k = 0, ..., n-1$*, we have*

$$d := \inf\{|1 + K\phi'(t)| : \ t \in I_k, \ 0 \le k \le n-1\} > 2, \tag{4.9}$$

and

$$B_{\frac{d}{d-2}\epsilon}(z_k + K\phi(z_k)) \subset \{z + K\phi(z) : \ z \in I_k\}, \tag{4.10}$$

then there is $z^* = (z_k^*) \in \mathbf{R}^n$ *such that* $z_k^* \in I_k$*, and*

$$z_{k-1}^* + z_k^* + K\phi(z_k^*) + z_{k+1}^* = -\xi_k, \ 0 \le k \le n-1. \tag{4.11}$$

Proof. We obtain z^* as the limit of a sequence of corrections of z, each diminishing the supremum of ϵ_k's by a definite multiplicative factor $\lambda < 1$. The hypothesis (4.9) guarantees that $z_k + K\phi(z_k)$ changes faster (in z_k) than the "coupling term" $z_{k-1} + z_{k+1}$, thus ϵ_k can be effectively decreased by manipulating z_k only. The hypothesis (4.10) secures enough "maneuvering space" to perform all infinitely many corrections. With this in mind, it will be convenient to think of $\frac{d}{d-2}\epsilon$ as follows: fix $\kappa \in (0,1)$ and let $\lambda = 1 - \kappa + 2\kappa/d < 1$, then

$$\frac{d}{d-2}\epsilon = \frac{\kappa}{\kappa - 2\kappa/d}\epsilon = \kappa\frac{1}{1-\lambda}\epsilon = \kappa\epsilon + \kappa\lambda\epsilon + \kappa\lambda^2\epsilon + ... \tag{4.12}$$

We describe the first correction. From (4.12), $\kappa\epsilon \le \frac{d}{d-2}\epsilon$, so, by (4.10),

$$B_{\kappa\epsilon}(z_k + K\phi(z_k)) \subset \{z + K\phi(z) : \ z \in I_k\}. \tag{4.13}$$

It follows that we can find $z_k' \in I_k$, $k = 0, ..., n-1$, so that

$$|z_{k-1} + z_k' + K\phi(z_k') + z_{k+1} + \xi_k| = (1-\kappa)|\epsilon_k|$$

and

$$|z'_k + K\phi(z'_k) - z_k - K\phi(z_k)| = \kappa|\epsilon_k| \leq \kappa\epsilon. \tag{4.14}$$

The hypothesis (4.9) implies then that

$$|z'_k - z_k| \leq \frac{\kappa}{d}|\epsilon_k|.$$

Putting the two inequalities together we get

$$z'_{k-1} + z'_k + K\phi(z'_k) + z'_{k+1} = -\xi_k + \epsilon'_k,$$

for some ϵ'_k such that

$$|\epsilon'_k| \leq (1-\kappa)|\epsilon_k| + \frac{\kappa}{d}|\epsilon_{k-1}| + \frac{\kappa}{d}|\epsilon_{k+1}| \leq \left(1 - \kappa + 2\frac{\kappa}{d}\right)\epsilon = \lambda\epsilon.$$

In this way

$$\epsilon' = \max_{k=0}^{n-1} |\epsilon'_k| \leq \lambda\epsilon. \tag{4.15}$$

To repeat the procedure leading from $z = (z_k) \in \mathbf{R}^n$ to $z' = (z'_k) \in \mathbf{R}^n$ with z' as our new z, we need to verify (4.13) with z_k and ϵ replaced by z'_k and ϵ' respectively. In view of (4.10) and (4.14), it suffices to check that

$$\kappa\epsilon + \kappa\epsilon' \leq \frac{d}{d-2}\epsilon.$$

This is however a consequence of (4.15) and (4.12). It should be clear now how the series in (4.12) unfolds as we need to guarantee the possibility of carrying out the infinitely many consecutive steps. Q.E.D.

Proof of Proposition 4.3.1. Fix $K > 1$. We want to prove that the rotation set of G^3 corresponding to $K\phi$ contains that corresponding to ϕ. Note that, by (i) of Proposition 4.2.1, the rotation set is either equal to $\{(0,0)\}$ or it has nonempty interior. In the latter case, by Theorem 0.0.2, the rotation set is the closure of the totality of the rotation vectors of periodic orbits. Consequently, it is enough to show the following: for each periodic orbit of the torus map generated by ϕ, there is one with the same rotation vector for $K\phi$. In view of Lemma 4.3.1, it suffices to find, for $z \in \mathbf{R}^n$ satisfying (4.5), a point $z^* \in \mathbf{R}^n$ satisfying (4.11).

Let $M := \sup \phi = -\inf \phi$. Clearly, ϕ and $K\phi$ have the same intervals of monotonicity, $K\phi$ has larger image, and

$$|K\phi(x) - \phi(x)| \leq (K-1)M, \text{ for all } x \in \mathbf{R}. \tag{4.16}$$

Thus, if $z \in \mathbf{R}^n$ satisfies (4.5) for some $\xi \in (2\pi\mathbf{Z})^n$, we can find for each z_k a point y_k in the same interval of monotonicity such that $K\phi(y_k) = \phi(z_k)$. By (4.16) and the assumption on the derivative

$$|z_k - y_k| \leq (K-1)M/(KD) \leq (K-1)M/(6K).$$

In this way

$$y_{k-1} + y_k + K\phi(y_k) + y_{k+1} = -\xi_k + \epsilon_k,$$

with $|\epsilon_k| \leq \epsilon := 3(K-1)M/(6K) = (K-1)M/(2K)$.

We will use Lemma 4.3.2 to find $z^* \in \mathbf{R}^n$ that obeys (4.11). The hypothesis (4.9) is satisfied because $d \geq D-1 \geq 5$. Let I_k be the monotonicity interval containing y_k, that is $I_k = [c_k, c'_k]$ where c_k and c'_k are the critical point of ϕ closest to y_k. To verify hypothesis (4.10), observe that image under $K\phi$ of any maximal interval of monotonicity extends exactly $(K-1)M$ beyond the end-points of the corresponding image for ϕ. This and the fact that $|K\phi'| \geq KD$ yield

$$|K\phi(c_k) + c_k - K\phi(y_k) - y_k| \geq |K\phi(c_k) - K\phi(y_k)| - |c_k - y_k| \geq$$
$$|K\phi(c_k) - K\phi(y_k)| \left(1 - \frac{1}{KD}\right) \geq (K-1)M\left(1 - \frac{1}{KD}\right),$$

and the analogous inequality for c'_k. Since $\frac{d}{d-2}\epsilon \leq \frac{D-1}{D-3}\frac{(K-1)M}{2K}$, to satisfy (4.10), we need $\frac{D-1}{D-3}\frac{(K-1)M}{2K} \leq (K-1)M(1 - 1/(KD))$. Dividing both sides by $(K-1)M$ and setting $D = 6$, we get a stronger inequality $\frac{5}{3\cdot 2K} \leq 1 - \frac{1}{6K}$, which is true because $K \geq 1$. Q.E.D.

4.4 A mechanical toy-model

Finally, let us point out a mechanical toy-model governed by (4.4). For $(z_k) \in \mathbf{R}^{\mathbf{Z}}$, interpret each z_k, $k \in \mathbf{Z}$, as the position of a unit mass particle along the real axis. Let the particles interact by putting a Hookean string between every two with consecutive indices. Also, subject the particles to a potential force with the potential $U(x) = -\frac{3}{2}x^2 - \Phi$, where Φ is the anti-derivative of ϕ. In the resulting mechanical model, the net force F_k exerted on the k^{th} particle is the sum of the potential force $-U'(z_k) = -3z_k - \phi(z_k)$ and the interaction forces $(z_{k+1} - z_k)$ and $(z_{k-1} - z_k)$ from the two neighbors. Thus we have

$$F_k = z_{k-1} + z_k + z_{k+1} + \phi(z_k),$$

and the solutions to (4.4) can be interpreted as the stationary configurations of this mechanical system. The construction is somewhat parallel to one relating twist maps on the annulus to the Aubry theory (see e.g. [Ban88, Ang90]). However, for now, we have no interesting theorems that stem from this approach.

Appendix A

Elementary large deviations for topological dynamics

We give a proof of Theorem 2.1.1 that is elementary, it makes no use of ergodic theory. Our original inspiration for this approach stems from Theorem II.6.1 in [Ell85]. However, we present here a point of view not taken up in [Ell85] that yields a very simple argument for Theorem 2.1.1. It can also be used to considerably simplify the proof of Theorem II.6.1 in [Ell85].

Recall that we have $f : X \to X$, where X is compact and f is continuous. Also, the observable $\phi : X \to \mathbf{R}^d$ is a continuous function. Since f and ϕ will be fixed thruought the rest of the appendix, we will suppress the dependence on them in the most of our notation. In particular, we will write $\phi_n(x)$ for $S_n(f, \phi)(x)$, that is $\phi_n(x) = \phi(x) + ... + \phi(f^{n-1}(x))$. For $s \in \mathbf{R}^d$ and $n \in \mathbf{N}$, we have the *Gibbs weight* ascribed to every point $x \in X$, namely $\mu_n^s(x) := \exp\langle \phi_n(x), s \rangle$. Given a finite set $S \subset X$, the Gibbs weight $\mu_n^s(\cdot)$ determines a measure μ_n^s supported on S and given by $\mu_n^s := \sum_{x \in S} \mu_n^s(x) \cdot \delta_x$, where δ_x is a unit mass at x.

We are interested in the limiting distribution of the averages of ϕ among (ϵ, n)-separated orbits counted with respect to the Gibbs weights. (Theorem 2.1.1 pertains to the case $s = 0$.) The corresponding pressure function $p^s : X \to \mathbf{R}$ (cf. *free energy* in [Ell85]) is given by

$$p^s(t) := \lim_{\epsilon \to 0} \limsup_{n \to \infty} (1/n) \ln \sup \left\{ \sum_{x \in S} \exp\langle \phi_n(x), t \rangle \mu_n^s(x) : S \ (n, \epsilon)\text{-separated} \right\}.$$

Observe that

$$p^s(t) = p^0(t+s) = p_{top}^{(t+s)}(f, \phi), \quad s, t \in \mathbf{R}^d, \tag{A.1}$$

and so, for the convex conjugate $\hat{p}^s$ of p^s, we have

$$\hat{p}^s(z) = \hat{p}^0(z) - \langle z, s \rangle, \quad s, z \in \mathbf{R}^d. \tag{A.2}$$

We should also note that replacing "lim sup" with "lim inf" in the definition of p^s yields the same quantity (see (viii) of Theorem 9.4 in [Wal82]).

For a set $E \subset \mathbf{R}^d$, we generalize $h_{top}^E(f, \phi)$ and introduce

$$H_+^s(E) := \lim_{\epsilon \to 0} \limsup_{n \to \infty} (1/n) \ln \sup\{\mu_n^s(S) : S \ (\epsilon, n)\text{-separated} \\ \text{with } \phi_n(x)/n \in E, \text{ for all } x \in S\ \},$$

$$H_-^s(E) := \lim_{\epsilon \to 0} \liminf_{n \to \infty} (1/n) \ln \sup\{\mu_n^s(S) : S \ (\epsilon, n)\text{-separated} \\ \text{with } \phi_n(x)/n \in E, \text{ for all } x \in S\}.$$

Both $H_+^s(E)$ and $H_-^s(E)$ take values in $\{-\infty\} \cup [0, +\infty]$. Clearly, if $E \subset F \subset \mathbf{R}^d$, then $H_{+/-}^s(E) \le H_{+/-}^s(F)$. (Here $H_{+/-}^s$ stands for either of H_+^s and H_-^s.) One can also easily verify that, for $E, F \subset \mathbf{R}^d$,

$$H_+^s(E \cup F) = \max\{H_+^s(E), H_+^s(F)\}, \tag{A.3}$$

and if $H^s_+(E) < H^s_-(E \cup F)$, then $H^s_-(F) = H^s_-(E \cup F)$, what implies

$$H^s_-(E) \leq H^s_-(E \cup F) \leq \max\{H^s_+(E), H^s_-(F)\}. \tag{A.4}$$

Mimicking the way $h^{(v)}_{top}(f,\phi)$ arises from $h^E_{top}(f,\phi)$, we associate with H^s_+ and H^s_- the following entropies at $z \in \mathbf{R}^d$:

$$h^s_+(z) := \lim_{r\to 0} H^s_+(B_r(z)),$$
$$h^s_-(z) := \lim_{r\to 0} H^s_-(B_r(z)).$$

Note that $h^0_+(z) = h^{(z)}_{\text{top}}(f,\phi)$, $z \in \mathbf{R}^d$. In general, $h^s_+ \neq h^s_-$ (see the example ahead and Proposition A.0.2). The function $z \mapsto h^s_{+/-}(z)$ is quite obviously upper semi-continuous. Also, note that if $D := \text{conv}\{\phi(x) : \ x \in X\}$, then $(1/n)\phi_n(x) \in D$, $x \in X$, $n \in \mathbf{N}$. Consequently, for $E \subset \mathbf{R}^d$, if $E \cap D = \emptyset$, then

$$H^s_{+/-}(E) = -\infty. \tag{A.5}$$

In particular, $h^s_{+/-}(z) = -\infty$, for $z \notin D$. (Actually, one can take here $D := \rho(f,\phi)$ — see Proposition 2.1.1.)

The dependence of $-h^s_{+/-}(z)$ on s is analogous to that of $\hat{p}^s$, namely

$$h^s_{+/-}(z) = \langle z, s\rangle + h^0_{+/-}(z), \quad z, s \in \mathbf{R}^d. \tag{A.6}$$

This is because, if $r > 0$ and $S \subset X$ is a (ϵ, n)-separated set with $(1/n)\phi_n(x) \in B_r(z)$ for $x \in S$, then

$$\exp(-nr) \leq |\mu^s_n(x)/\exp\langle nz, s\rangle| \leq \exp(nr), \ \ s \in \mathbf{R}^d;$$

and consequently,

$$\exp(-2nr) \leq |\exp\langle nz, s\rangle \mu^0_n(S)/\mu^s_n(S)| \leq \exp(2nr).$$

Lemma A.0.1 *For any* $s \in \mathbf{R}^d$,

(a) if $U \subset \mathbf{R}^d$ *is open, then*

$$H^s_{+/-}(U) \geq \sup\{h^s_{+/-}(z) : \; z \in U\},$$

(b) if $V \subset \mathbf{R}^d$ *is closed, then*

$$H^s_+(V) \leq \max\{h^s_+(z) : \; z \in V\}.$$

Proof. Part (a) is trivial. We prove (b) now. Let D be the convex hull of the image of ϕ. From (A.5), we see that $H^s_+(V) = H^s_+(V \cap D)$. Thus we can assume that V is compact. Now, set $V_0 := V$, and cover V_0 with a finite family $\mathcal{A}$ of its compact subsets of diameter less than 1. By (A.3), we have $H^s_+(V_0) = \max\{H^s_+(A) : \; A \in \mathcal{A}\}$; denote by V_1 an element of $\mathcal{A}$ such that $H^s_+(V_0) = H^s_+(V_1)$. Analogously, covering V_1 with its compact subsets of diameter less than $1/2$, we get $V_2 \subset V_1$ with $H^s_+(V_1) = H^s_+(V_2)$. By continuing this procedure, we get a nested sequence of compact sets $V_0 \supset V_1 \supset V_2 \supset V_3 \ldots$ with diameters $\mathrm{diam}(V_n) < 1/n$ and $H^s_+(V_n) = H^s_+(V)$, $n \in \mathbf{N}$. Thus $\bigcap_{n \in \mathbf{N}} V_n$ consists of a point, and one can easily verify that, if v is this point, then $H^s_+(v) \geq H^s_+(V)$. Q.E.D.

The following is a version of Theorem 2.1.1.

Proposition A.0.1 (equivalent of Theorem 2.1.1) *If* $h_{\mathrm{top}}(f) < +\infty$, *then*

(i) convex hull$(-h^0_+) = \hat{p}^0$,

(ii) $(-h^0_+)\hat{\ } = p^0$.

Remark. Suppose that $h_{\text{top}}(f) = +\infty$. Then $p^0(t) = +\infty$, $t \in \mathbf{R}^d$, so $\hat{p}^0(z) = -\infty$, $z \in \mathbf{R}^d$. Also $H^s_+(\mathbf{R}^d) = h_{\text{top}}(f) = +\infty$; and, from (b) of Lemma A.0.1, one concludes that convex hull$(-h^0_+)$ is *improper* (i.e. assumes the value $-\infty$). This implies that $(-h^0_+)\hat{}\,(t) = +\infty$, $t \in \mathbf{R}^d$ — (ii) is satisfied. In contrast, (i) fails — even though convex hull$(-h^0_+)(z) = -\infty$, for z in the relative interior of the essential domain (cf. Theorem 7.2 [Roc72]); it is $+\infty$ or finite outside of this interior.

Proof of Proposition A.0.1. First of all parts (i) and (ii) are equivalent: one can derive the other by taking convex conjugates of both sides. In particular, from (ii), after taking convex conjugates, we get (by the involutive property of $\hat{}$ — see the page 34) that cl(convex hull$(-h^0_+)) = \hat{p}^0$. Since $-h^0_+$ is lower semi-continuous and has bounded essential domain, we can drop "cl" to get (i) (see Proposition A.0.3). We prove (ii) now.

Clearly $p^s(0) = H^s_{+/-}(\mathbf{R}^d)$, and so we get from Lemma A.0.1 the following

(a) $p^s(0) \geq h^s_+(z)$, $s, z \in \mathbf{R}^d$,

(b) $p^s(0) \leq \sup\{h^s_+(z) :\ z \in \mathbf{R}^d\}$.

Making the dependence on s explicit (via (A.6) and (A.2)), write (a) as $p^0(s) \geq \langle z, s\rangle + h^0_+(z)$. Optimizing over all $z \in \mathbf{R}^d$, we get $p^0(s) \geq \sup\{\langle z, s\rangle + h^0_+(z) :\ z \in \mathbf{R}^d\} = (-h^0_+)\hat{}\,(z)$. Similarly, (b) yields $p^0(s) \leq \sup\{\langle s, z\rangle + h^0_+(z) :\ z \in \mathbf{R}^d\}$, i.e. $p^0(s) \leq \hat{h}^0_+(s)$. Hence, $\hat{h}^0_+(s) = p^0(s)$, what proves (ii). Q.E.D.

Simple considerations from convex analysis show that, in Proposition A.0.1, the entropy h^0_+ can be replaced with h^0_-. Recall that the set of subdifferentials of p^s at $t \in \mathbf{R}^d$ is given by

$$\partial p^s(t) := \{z \in \mathbf{R}^d : p^s(t') \geq p^s(t) + \langle z, t' - t\rangle \text{ for all } t' \in \mathbf{R}^d\}.$$

If $\partial p^s(t)$ consists of a single point, then p^s is differentiable at t and $\partial p(t) = \{\nabla p(t)\}$. The following is a convenient characterization (p. 218, [Roc72]):

$$\hat{p}^s(z) + p^s(t) = \langle z, t\rangle \equiv z \in \partial p^s(t) \equiv t \in \partial \hat{p}^s(z). \tag{A.7}$$

As a corollary, any gradient $v := \nabla p^s(t)$ is a point of strict convexity of $\hat{p^s}$. (Actually, gradients v are exactly the points at which the epigraph epi$(\hat{p^s})$ is *exposed*, i.e. there is a supporting hyperplane to epi$(\hat{p^s})$ that touches it only at $(v, \hat{p}^s(v))$ — see Corollary 25.1.3, p. 243 in [Roc72].) The following proposition explains the special role of the gradients in our context.

Proposition A.0.2 *If p^0 is differentiable at $s \in \mathbf{R}^d$ and $v := \nabla p^0(s)$, then*

$$-h^0_+(v) = -h^0_-(v) = \hat{p}^0(v). \tag{A.8}$$

By Th. 18.6 p. 167 in [Roc72], the exposed points are dense among all extremal points. This (using lower semicontinuity of $-h^0_-$ and $\hat{p}^0$) leads to the following corollary strengthening Proposition A.0.1.

Corollary A.0.1 *If $h_{\text{top}}(f) < +\infty$, then* convex hull$(-h^0_-) = \hat{p}^0$.

Proof of Proposition A.0.2. As we have already said, $\hat{p}^0$ is strictly convex at v, and so $-h^0_+(v) = \text{convex hull}(-h^0_+)(v) = \hat{p}^0(v)$. Now, by (A.7), $v = \nabla p^s(0)$ is the unique point v such that

$$p^s(0) = \sup\{\langle z, 0\rangle - \hat{p}^s(z) : \ z \in \mathbf{R}^d\} = -\hat{p}^s(v).$$

Thus, for any $r > 0$, $\max\{-\hat{p}^s_+(z) :\ z \notin B_r(v)\} < p^s(0)$. However, Lemma A.0.1 yields

$$H^s_+(\mathbf{R}^d \setminus B_r(v)) \leq \max\{h^s_+(z) :\ z \notin B_r(v)\},$$

and, since (by Proposition A.0.1) $h^s_+(z) \leq -\hat{p}^s_+(z)$, we get

$$H^s_+(\mathbf{R}^d \setminus B_r(v)) < p^s(0) = H^s_{+/-}(\mathbf{R}^d).$$

By (A.4), $H^s_-(B_r(v)) = H^s_-(\mathbf{R}^d) = p^s(0)$, and so $h^s_-(v) = -\hat{p}^s(v)$. Our claim follows via (A.6) and (A.2). Q.E.D.

Remark. For readers familiar with [Ell85], it may be worthwhile to note the following. When $p^0(t)$ is smooth, $h^0_{+/-}(z) = -\hat{p}^0(z)$ and Lemma A.0.1 (with $s = 0$) becomes an analogue of Th. II.6.1 in [Ell85]. Our proof leads along a different path than that of Ellis. We start with $h^s_{+/-}$, which are almost tautologically the right *entropy functions* (cf. Def. II.3.1, [Ell85]); and then we show their relation to the *free energy* $p^s(t)$ as a manifestation of the dual behavior of $h^s_{+/-}$ and $p^s(t)$ under variation of s (see (A.6), (A.2)). This approach modifies in a straightforward way to Ellis's setting yielding what some may find to be a simpler and more natural proof of Th. II.6.1.

The following example shows that, in general, $h^s_+ \neq h^s_-$.

Example (cf. the example in the first section of Chapter 2). For $(x_i)_{i\in\mathbf{Z}} \in \{1,2,3,4,5,6\}^{\mathbf{Z}}$, $(x_k, ..., x_l)$ is a *mod* 3 *block* if $x_k, ..., x_l$ are congruent modulo 3. We call it *maximal* if x_{k-1}, x_{l+1} are both not congruent modulo 3 to x_i, $i = k, ..., l$. The number $l - k + 1$ is referred to as the *length of the block*. Fix an arbitrary unbounded subset of natural numbers $\mathcal{N}$. Consider a subshift

(Λ, σ), where

$$\Lambda = \{x \in \{1,2,3,4,5,6\}^{\mathbf{Z}} : x \text{ has at most three maximal parity blocks each either infinite or with length in } \mathcal{N}\}$$

and σ shifts to the left: $(\sigma((x_i)_{i\in\mathbf{Z}}))_i := x_{i+1}$, $i \in \mathbf{Z}$. The space Λ is easily seen to be compact and σ invariant. Let the observable $\phi : \Lambda \to \mathbf{R}^3$ be given by

$$\phi(x) := \begin{cases} (1,0,0), \text{ if } x_0 = 0 \bmod 3 \\ (0,1,0), \text{ if } x_0 = 1 \bmod 3 \\ (0,0,1), \text{ if } x_0 = 2 \bmod 3 \end{cases}.$$

For $n \in 3\mathcal{N}$, one has a $(1,n)$-separated set S with cardinality 2^n and such that $(1/n)\phi_n(x) = (1/3, 1/3, 1/3)$, $x \in S$. Indeed, consider $(x_i) \in \Lambda$ that have $(x_0, ..., x_{n/3-1})$, $(x_{n/3}, ..., x_{2n/3-1})$, $(x_{2n/3}, ..., x_{n-1})$ as *mod* 3 *blocks*, where there are $2^{n/3}$ possibilities for each block given a residue class mod 3. Thus, $h^0_+(1/3, 1/3, 1/3) \geq \ln 2$. (Actually, we have equality since one can easily see that $h_{\text{top}}(\sigma) = \ln 2$.) On the other hand, we claim that if $\mathcal{N}$ is very sparse, say $\mathcal{N} := \{k!\}_{k\in\mathbf{N}}$, then $h^0_-(1/3, 1/3, 1/3) = -\infty$. Indeed, let $n = k! - 1$. For $x = (x_i) \in \Lambda$, any maximal *mod* 3 *block* contained in $(x_0, ..., x_{n-1})$ has length not exceeding $(k-1)!$. Thus at least one of the three coordinates of $(1/n)\phi_n(x)$ does not exceed $(k-1)!/n \approx 1/k$. This makes $(1/n)\phi_n(x)$ lie in a definite distance from $(1/3, 1/3, 1/3)$ for all $x \in \Lambda$ and infinitely many n of the form $n = k! - 1$, $k \in \mathbf{N}$. Our claim follows.

We finish with the following fact from convex analysis that we used numerous times.

Proposition A.0.3 *If* $f : \mathbf{R}^d \to (-\infty, +\infty]$ *is lower semi-continuous and* dom(f) *is bounded, then* convex hull(f) *is lower semi-continuous.*

Proof. By lower semi-continuity, f has a finite infimum $M \in \mathbf{R}$. By considering $f - M$ in place of f, we may assume that f is non-negative.

Let $K := \text{epi}(f) \subset \mathbf{R}^{d+1}$. The essential domain dom(f) is the projection of K on the $\mathbf{R}^d$ of the arguments. It is enough to demonstrate that conv(K) is closed, because then it is the epigraph of convex hull(f) (see the definitions on p. 36 in [Roc72]).

Suppose that $(z_n, y_n) \in \text{conv}(K)$ and $\lim_{n\to\infty}(z_n, y_n) = (z, y)$. We want to show that $(z, y) \in \text{conv}(K)$. From the Caratheodory theorem ([Roc72]), there are $(w_{n,i}, t_{n,i}) \in K$ and $\alpha_{n,i}, \lambda_n > 0$, $\sum_{i=1}^{d+1} \alpha_{n,i} = 1$ such that

$$(z_n, y_n) = \sum_{i=1}^{d+1} \alpha_{n,i} \cdot (w_{n,i}, t_{n,i}) + (0, \lambda_n). \tag{A.9}$$

Let $B := \{i : \ (t_{n,i})_{n\in\mathbf{N}} \text{ is unbounded}\}$. Passing perhaps to a subsequence, we can assume that, as $n \to \infty$, we have $t_{n,i} \to t_i \in \mathbf{R}$, $i \notin B$; and additionally, $z_n \to z \in \mathbf{R}^d$, $\lambda_n \to \lambda \in \mathbf{R}$, $w_{n,i} \to w_i \in \mathbf{R}^d$ and $\alpha_{n,i} \to \alpha_i$, $i = 1, ...n$. Note that there must be $\alpha_i = 0$ for $i \in B$; otherwise y_n's could not stay bounded. Thus the sum on the right side below

$$(z', y') := \sum_{i\notin B} \alpha_i \cdot (w_i, t_i) + (0, \lambda),$$

is a convex combination, and $z' = z$. Also, $y' \leq y$ because $t_{n,i} \geq 0$. Since K is closed, we have $(w_i, t_i) \in K$, so $(z', y') \in \mathrm{conv}(K)$. It follows that (z, y) is also contained in $\mathrm{conv}(K)$. Q.E.D.

Bibliography

[Abi80] W. Abikoff. *The Real Analytic Theory of Teichmuller Space.* Springer-Verlag, Berlin, 1980.

[Alh66] L. Alhfors. *Lectures on Quasiconformal Mappings.* van Nostrand, New York, 1966.

[ALM93] L. Alseda, J. Llibre, and M. Misiurewicz. *Combinatorial Dynamics and Entropy in Dimension One*, volume 5 of *Advanced series in Nonlinear Dynamics.* World Scientific, Singapore, 1993.

[Ang90] S.B. Angenent. Monotone recurrence relations, their Birkhoff orbits and topological entropy. *Erg. Th. & Dyn. Sys.*, 10:15–41, 1990.

[Ban88] V. Bangert. Mather Sets for Twist Maps and Geodesics on Tori. volume 1 of *Dynamics Reported.* Oxford University Press, Oxford, 1988.

[BH92] M. Bestvina and M. Handel. Train tracks for surface homeomorphisms. CUNY preprint, 1992.

[Bow71] R. Bowen. Entropy for group endomorphisms and homogenous spaces. *Trans. Amer. Math. Soc.*, 153:401–414, 1971.

[Bow75] R. Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphism*, volume 470 of *Springer Lecture Notes in Math.* Springer-Verlag, Berlin, 1975.

[Bow78a] R. Bowen. Entropy and the fundamental group. In *The Structure of Attractors in Dynamical Systems*, volume 668 of *Lecture Notes in Mathematics.* Springer, Berlin, 1978.

[Bow78b] R. Bowen. Markov partitions are not smooth. *Proc. Amer. Math. Soc.*, 71:130–132, 1978.

[Boy92] P. Boyland. The rotation set as a dynamical invariant. In *Proceedings of IMA Workshop on Twist Maps*, volume 44 of *IMA Volumes in Math. Appl.* Springer, Berlin, 1992.

[Boy94] P. Boyland. Topological methods in surface dynamics. *Topology and its Applications*, 58:223–298, 1994.

[Boy10] P. Boyland. Isotopy stability of dynamics on surfaces. IMS, Stony Brook preprint, 1993/10.

[BW93] M. Barge and R. Walker. Periodic point free maps of tori which have rotation sets with interior. *Nonlinearity*, 6 (3):481–489, 1993.

[CSUZ87] A.A. Chernikov, R.Z. Sagdeev, D.A. Usikov, and G.M. Zaslavsky. The Hamiltonian method for quasicrtystal symmetry. *Physics Letters A*, 125 (2):102–106, 1987.

[Ell85] R. S. Ellis. *Entropy, Large Deviations, and Statistical Mechanics.* Springer-Verlag, New York, 1985.

[FLP79] A. Fathi, F. Lauderbach, and V. Poenaru. *Travaux de Thurston sur les surfaces*, volume 66-67 of *Asterisque.* 1979.

[FM90] J. Franks and M. Misiurewicz. Rotation sets of toral flows. *Proc. Amer. Math. Soc.*, 109 no. 1:243–249, 1990.

[Fra88] J. Franks. Recurrence and fixed points of surface homeomorphisms. *Erg. Th. & Dyn. Sys.*, 8:99–107, 1988.

[Fra89] J. Franks. Realizing rotation vectors for torus homeomorphisms. *Trans. Amer. Math. Soc.*, 311:107–115, 1989.

[Fra94] J. Franks. Rotation vectors and fixed points of area preserving surface diffeomorphisms. Preprint, 1994.

[Fri82a] D. Fried. Flow equivalence, hyperbolic systems and a new zeta function for flows. *Comment. Math. Helv.*, 57:237–259, 1982.

[Fri82b] D. Fried. The geometry of cross sections to flows. *Topology*, 21:353–371, 1982.

[Gro87] M. Gromov. On the entropy of holomorphic maps. Preprint, 1987.

[Han85] M. Handel. Global shadowing of pseudo-anosov homeomorphisms. *Erg. Th. & Dyn. Sys.*, 5:373–377, 1985.

[Kwa] J. Kwapisz. A toral diffeomorphism with a non-polygonal rotation set. IMS, Stony Brook, preprint #1994/7 (to appear in Nonlinearity).

[Kwa92] J. Kwapisz. Every convex rational polygon is a rotation set. *Erg. Th. & Dyn. Sys.*, 12:333–339, 1992.

[Kwa93] J. Kwapisz. An estimate of entropy for toroidal chaos. *Erg. Th. & Dyn. Sys.*, 13:123–129, 1993.

[LL93] A.J. Lichtenberg and M.A. Lieberman. *Regular and Chaotic Dynamics.* Springer, Berlin, 1993.

[LM91] J. Llibre and R. MacKay. Rotation vectors and entropy for homeomorphisms of the torus isotopic to the identity. *Erg. Th. & Dyn. Sys.*, 11:115–128, 1991.

[Man87] R. Mane. *Ergodic Theory and Differentiable Dynamics.* Springer, Berlin, 1987.

[May91] D. H. Mayer. Continued fractions and related transformations. In *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces.* Oxford University Press, Oxford, 1991.

[MT91] B. Marcus and S. Tuncel. The weight-per-symbol polytope and scaffolds of invariants associated with markov chains. *Erg. Th.*

& Dyn. Sys., 11:129–180, 1991.

[MT92] M. Misiurewicz and J. Tolosa. Entropy of snakes and the restricted variational principle. *Erg. Th. & Dyn. Sys.*, 12:791–802, 1992.

[MZ89] M. Misiurewicz and K. Ziemian. Rotation sets for maps of tori. *J. London Math. Soc.*, 40:490–506, 1989.

[MZ91] M. Misiurewicz and K. Ziemian. Rotation sets and ergodic measures for torus homeomorphisms. *Fund. Math.*, 137:45–52, 1991.

[New89] S. Newhouse. Continuity properties of entropy. *Ann. of Math. (2)*, 129:215–235, 1989.

[NPT83] S. Newhouse, J. Palis, and F. Takens. Bifurcations and stability of families of diffeomorphisms. *Publ. Math. IHES*, 57:5–71, 1983.

[PP90] W. Parry and M. Pollicott. *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, volume 187-188 of *Astérisque*. 1990.

[Roc72] R. T. Rockafellar. *Convex Analysis*. Princeton Mathematical Series. Princeton University Press, Princeton, 1972.

[Rue78] D. Ruelle. *Thermodynamical formalism*. Encyclopedia of Mathematics and its Applications. Addison-Wesley, London, 1978.

[Thu88] W. Thurston. On the geometry and dynamics of diffeomorphisms of surfaces. *Bull. Amer. Math. Soc.*, 19:417–431, 1988.

[Wal82] P. Walters. *An Introduction to Ergodic Theory*, volume 79 of *Graduate texts in mathematics.* Springer-Verlag, New York, 1982.

[Zie95] K. Ziemian. Rotation sets for subshifts of finite type. *Fund. Math.*, 146:189–201, 1995.

[ZZRSC86a] G.M. Zaslavsky, Mu.YU Zakharov, D.A. Usikov R.Z. Sagdeev, and A.A. Chernikov. Generation of ordered structures with a symmetry axis from a Hamiltonian dynamics. *Pis'ma Zh. Eksp. Teor. Fiz.*, 44 (7):451–456, 1986.

[ZZRSC86b] G.M. Zaslavsky, Mu.YU Zakharov, D.A. Usikov R.Z. Sagdeev, and A.A. Chernikov. Stochastic web and diffusion of particles in a magnetic field. *Sov. Phys. JETP*, 64:294–303, 1986.

Index