ISRAEL JOURNAL OF MATHEMATICS **128** (2002), 93–134

INVARIANT MEASURES AND ASYMPTOTICS FOR SOME SKEW PRODUCTS

BY

JON AARONSON

School of Mathematical Sciences, Tel Aviv University
Ramat Aviv, 69978 Tel Aviv, Israel
e-mail: aaro@math.tau.ac.il

AND

HITOSHI NAKADA

Department of Mathematics, Keio University
Hiyoshi 3-14-1 Kohoku, Yokohama 223, Japan
e-mail: nakada@math.keio.ac.jp

AND

OMRI SARIG

Mathematics Institute, University of Warwick
Coventry, CV4 7AL, UK
e-mail: sarig@maths.warwick.ac.uk

AND

RITA SOLOMYAK

Department of Mathematics, Box 354350, University of Washington
Seattle, WA 98195-4350, USA
e-mail: rsolom@math.washington.edu

ABSTRACT

For certain group extensions of uniquely ergodic transformations, we identify all locally finite, ergodic, invariant measures. These are Maharam type measures. We also establish the asymptotic behaviour for these group extensions proving logarithmic ergodic theorems, and bounded rational ergodicity.

Received March 30, 2000

§0. Introduction and general framework

Let $(X, \mathcal{B})$ be a standard measurable space, and let $\tau: X \to X$ be an invertible measurable map. Let $\mathbb{G}$ be a locally compact, Abelian, Polish (LCAP) topological group and let $\phi: X \to \mathbb{G}$ be measurable.

The **skew product** transformation $\tau_\phi: X \times \mathbb{G} \to X \times \mathbb{G}$ is defined by

$$\tau_\phi(x, y) := (\tau x, y + \phi(x)).$$

A measure m: $\mathcal{B} \otimes \mathcal{B}(\mathbb{G}) \to [0, \infty]$ is called **locally finite** if $m(X \times K) < \infty \ \forall\, K \subset \mathbb{G}$ compact.

Our program is to identify all τ_ϕ-invariant locally finite measures and study their asymptotic behaviour.

It is known ([Fu], [Pa]) that if τ is a uniquely ergodic homeomorphism of a compact metric space (with invariant probability p), $\mathbb{G}$ is compact (with Haar probability measure $m_{\mathbb{G}}$) and $\phi: X \to \mathbb{G}$ is continuous, then ergodicity of τ_ϕ with respect to the product $p \times m_{\mathbb{G}}$ is equivalent to the unique ergodicity of τ_ϕ.

For non-compact $\mathbb{G}$, it is well known that if τ is uniquely ergodic (with invariant probability p), and τ_ϕ is ergodic with respect to $p \times m_{\mathbb{G}}$, then there is no τ_ϕ-invariant probability on $X \times \mathbb{G}$ (see, e.g., [A1] chapter 8, or [Sc2]).

It is natural to ask (as in [Ve]) for τ_ϕ-invariant locally finite measures. There is a natural class of τ_ϕ-invariant locally finite measures: the **Maharam measures** which we proceed to describe.

Let $(X, \mathcal{B})$ and τ be as above and let $h: X \to \mathbb{R}_+$ be measurable. We call a probability $\mu \in \mathcal{P}(X, \mathcal{B})$ **(h, τ)-conformal** if $\mu \circ \tau \sim \mu$ and $d\mu \circ \tau / d\mu = h$ μ-a.e.

Now let ϕ: $X \to \mathbb{G}$ be measurable, and let α: $\mathbb{G} \to \mathbb{R}$ be a continuous homomorphism. Let $\mu = \mu_\alpha$ be a $(e^{\alpha\circ\phi}, \tau)$-conformal probability on $(X, \mathcal{B})$. The associated **Maharam measure** is m_α: $\mathcal{B} \otimes \mathcal{B}(\mathbb{G}) \to [0, \infty]$ defined by $dm_\alpha(x, y) := e^{-\alpha(y)} d\mu(x) dy$ (where dy denotes Haar measure on $\mathbb{G}$). The reason for this terminology is that Maharam measures were first considered for $\mathbb{G} = \mathbb{R}$ in [Mah].

A Maharam measure is easily seen to be τ_ϕ-invariant, the dilation from the first coordinate being canceled by the translation in the second.

For the transformations τ_ϕ considered in this paper, we show the following properties:

Unique conformal probabilities: For each continuous homomorphism $\alpha : \mathbb{G} \to \mathbb{R}$, there is a unique $(e^{\alpha\circ\phi}, \tau)$-conformal probability $\mu = \mu_\alpha$ on $(X, \mathcal{B})$;

Maharam measures are ergodic: For each continuous homomorphism α: $\mathbb{G} \to \mathbb{R}$, the Maharam measure m_α is ergodic (for τ_ϕ);

ERGODIC MEASURES ARE MAHARAM: The only ergodic τ_ϕ-invariant locally finite measures are those proportional to Maharam measures.

Remarks:

(1) For $\mathbb{G}$ compact, the only continuous homomorphism α: $\mathbb{G} \to \mathbb{R}$ is $\alpha \equiv 0$, the only Maharam measures are of form $m \times m_{\mathbb{G}}$, and the above properties for τ_ϕ are equivalent to its unique ergodicity.

(2) As shown in [Sc2], there are abundances of $(e^{\alpha\circ\phi}, \tau)$-conformal infinite measures, and of non-locally finite, τ_ϕ-invariant, σ-finite measures.

We attempt our program in two cases. In §1, we treat the so-called **cylinder flow** $R_{\alpha,\chi}$: $\mathbb{T} \times \mathbb{R} \to \mathbb{T} \times \mathbb{R}$ defined by $R_{\alpha,\chi}(x,y) := (x+\alpha, y+\chi(x))$ where $\alpha \in \mathbb{T} \setminus \mathbb{Q}$ and where

$$\chi(x) = (\beta+1)\cdot 1_{[0,\frac{\beta}{\beta+1})} - \beta \quad \text{(some } \beta > 0\text{)},$$

the rest of the paper being devoted to certain group extensions of adic transformations by symmetric cocycles (see below).

Let S be a finite, ordered set, let A: $S\times S \to \{0,1\}$ be an irreducible, aperiodic matrix and let $\Sigma = \Sigma_A \subset S^{\mathbb{N}}$ be the corresponding (topologically mixing) subshift of finite type (SFT).

Let V be the adding machine on $S^{\mathbb{N}}$. The **adic transformation** on Σ is the induced transformation of V on Σ defined (in §2) for all except countably many points $x \in \Sigma$ by $\tau(x) = V^{\min\{n\geq 1:\ V^n(x)\in\Sigma\}}(x)$.

For f: $\Sigma \to \mathbb{G}$, we consider the **symmetric cocycle** ϕ_f: $\Sigma \to \mathbb{G}$ defined by $\phi_f(x) := \sum_0^\infty (f(T^i x) - f(T^i(\tau x)))$ where T: $\Sigma \to \Sigma$ is the shift, the sum terminating as $T^i(x) = T^i(\tau x)$ $\forall$ large $i \geq 1$.

In §2 we show that the class of τ_{ϕ_f}-invariant, locally finite measures for f aperiodic having finite memory is the collection of measures which are proportional to mixtures of the canonical Maharam measures (Theorems 2.1 and 2.2).

In §3 and §4, we consider the asymptotic properties of τ_{ϕ_f} with respect to Maharam measures, where f: $\Sigma \to \mathbb{R}^d$ is an aperiodic Hölder continuous function.

For $\alpha \in \mathbb{R}^d$, consider the Maharam measure m_α: $\mathcal{B}(\Sigma \times \mathbb{R}^d) \to [0,\infty]$ defined by $dm_\alpha(x,y) = e^{-\alpha\cdot y}d\mu(x)dy$ where $\mu = \mu_\alpha$ is the $(e^{\alpha\cdot f}, \tau)$-conformal measure. In §4, we show that τ_{ϕ_f} is boundedly rationally ergodic with return sequence $a(n) \asymp n/(\log n)^{d/2}$ (see [A2], and/or §4) with respect to m_0. Bounded rational ergodicity is a strong form of rational ergodicity, and so this entails a kind of absolutely normalized ergodic theorem:

$$\frac{S_n(f)}{a(n)} \rightsquigarrow \int_X f dm_0 \ \forall\ f \in L^1(m_0)$$

where $f_n \rightsquigarrow f$ if $\forall\, m_\ell \uparrow \infty\ \exists\, n_k = m_{\ell_k} \uparrow \infty$ such that $\forall\, p_j = n_{k_j} \uparrow \infty$, we have $\frac{1}{N}\sum_{j=1}^{N} f_{p_j} \to f$ a.e. as $N \to \infty$ (see [A1]).

For $\alpha \neq 0$, τ_{ϕ_f} is squashable with respect to m_α (see [A1]) and there is no such kind of ergodic theorem. Nevertheless, we show in §3 that the logarithmic ergodic theorem holds:

$$\frac{\log \sum_{k=0}^{n-1} F \circ \tau_{\phi_f}^k}{\log n} \longrightarrow \frac{h_{p_\alpha}(T)}{h_{top}(T)} \quad m_\alpha\text{-a.e. as } n \to \infty\ \forall F \in L^1(m_\alpha)_+$$

where p_α is the equilibrium measure of $\alpha \cdot f$ (see [Bo]).

There is some relation between the results of §2 and results in [P-S] remarked at the end of §2. The program in §3 and §4 has been previously carried out in full in [A-W] for $\Sigma = \{0,1\}^{\mathbb{N}}$, $f(x) = x_1$. Bounded rational ergodicity of certain of the cylinder flows was established in [A-K].

Horocycle flows on Abelian covers of compact, hyperbolic surfaces can be considered as "smooth analogues" of the skew products considered here. Ergodic, Maharam measures for these horocycle flows were introduced, and their asymptotics considered in [B-L].

We conclude this introduction with a Basic Lemma, to be used in §1 and §2.

For $a \in \mathbb{G}$, define Q_a: $X \times \mathbb{G} \to X \times \mathbb{G}$ by $Q_a(x,y) := (x, y+a)$, then $\tau_\phi \circ Q_a = Q_a \circ \tau_\phi$. If m is an ergodic τ_ϕ-invariant locally finite measure, then so is $m \circ Q_a$ ($a \in \mathbb{G}$) whence, as is well known, either $m \circ Q_a \perp m$ or $m \circ Q_a = cm$ for some $c \in \mathbb{R}_+$.

For m an ergodic τ_ϕ-invariant locally finite measure, set

$$H = H_m := \{a \in \mathbb{G}\colon m \circ Q_a \sim m\}.$$

0.1 Basic Lemma:

(i) *H is closed;*

(ii) *If $H = \mathbb{G}$, then m is proportional to a Maharam measure.*

Proof: (i) By unicity of absolutely continuous invariant measures, $\exists$ a multiplicative homomorphism $\Delta : H \to \mathbb{R}_+$ such that

$$\int_{X\times\mathbb{G}} f \circ Q_a dm = \Delta(a) \int_{X\times\mathbb{G}} f dm \quad \forall\, a \in H,\ f \in L^1(m).$$

For f: $X \times \mathbb{G} \to \mathbb{R}$ continuous with compact support, we have that $f \circ Q_{a_n} \to f \circ Q_a$ uniformly as $a_n \to a$ in $\mathbb{G}$. Suppose that $a_n \in H$, $a_n \to a \notin H$. This forces

$\Delta(a_n) \to 0$, since by the local finiteness assumption, $\forall\ \epsilon > 0 \exists\ f\colon X \times \mathbb{G} \to \mathbb{R}_+$ continuous with compact support such that

$$\int_{X\times\mathbb{G}} f dm = 1, \int_{X\times\mathbb{G}} f \circ Q_a dm < \epsilon$$

whence

$$\epsilon > \int_{X\times\mathbb{G}} f \circ Q_a dm \leftarrow \int_{X\times\mathbb{G}} f \circ Q_{a_n} dm = \Delta(a_n).$$

On the other hand $\exists\ f\colon X \times \mathbb{G} \to \mathbb{R}$ continuous, everywhere positive, and absolutely integrable. Clearly $f \circ Q_a > 0$ and $\int_{X\times\mathbb{G}} f \circ Q_a dm > 0$, contradicting $\Delta(a_n) \to 0$ and showing that $a \in H$.

(ii) There is a measurable (hence continuous) homomorphism $\alpha\colon \mathbb{G} \to \mathbb{R}$ such that $m \circ Q_a = e^{-\alpha(a)} m$. Define the measure $\overline{m}\colon \mathcal{B}(X \times \mathbb{G}) \to [0,\infty]$ by $d\overline{m}(x,y) := e^{\alpha(y)} dm(x,y)$. It follows that $\overline{m} \circ Q_a = \overline{m}$. For $A \in \mathcal{B}(X)$, $B \in \mathcal{B}(\mathbb{G})$ and $a \in \mathbb{G}$, we have

$$\overline{m}(A \times (B + a)) = \overline{m} \circ Q_a(A \times B) = \overline{m}(A \times B).$$

Since the Haar measure on $\mathbb{G}$ is unique up to a constant, $\forall\ A \in \mathcal{B}(X)$, $\exists\ \mu(A) \in \mathbb{R}_+$ such that

$$\overline{m}(A \times B) = \mu(A) m_{\mathbb{G}}(B) \quad (B \in \mathcal{B}(\mathbb{G})).$$

It follows that μ is a finite measure on X, and that

$$dm(x,y) = e^{-\alpha(y)} d\mu(x) dy.$$

The τ_ϕ-invariance of m now implies that $\mu \circ \tau \sim \mu$ with $d\mu \circ \tau / d\mu = e^{\alpha \circ \phi}$ (it being necessary to cancel the dilation due to translation of the second coordinate by dilation of the first). ∎

§1. Cylinder flows

Let $\mathbb{T} := \mathbb{R}/\mathbb{Z} \cong [0,1)$ denote the additive circle (the multiplicative circle being $\mathbb{S}^1 := e^{2\pi i \mathbb{T}} \subset \mathbb{C}$) and let $R_\alpha(x) := x + \alpha \mod 1$. The natural distance function on $\mathbb{T}$ is given by the norm $\|x\| := \min_{n\in\mathbb{Z}} |x + n|$.

For $\beta > 0$, let $\mathbb{G}_\beta \subset \mathbb{R}$ be the closed subgroup generated by 1 and β. Note that $\mathbb{G}_\beta = \beta\mathbb{Z}$ if $\beta \in \mathbb{Q}$ and $\mathbb{G}_\beta = \mathbb{R}$ if $\beta \notin \mathbb{Q}$. Consider, for $\beta > 0$, the function $\chi : \mathbb{T} \to \mathbb{G}_\beta$ defined by

$$\chi = \chi^{(\beta)} := 1_{[0,\frac{\beta}{\beta+1})} - \beta 1_{[\frac{\beta}{\beta+1},1)} = (\beta+1) 1_{[0,\frac{\beta}{\beta+1})} - \beta$$

and the skew products (or cylinder flows) $R_{\alpha,\chi^{(\beta)}} : \mathbb{T} \times \mathbb{G}_\beta \to \mathbb{T} \times \mathbb{G}_\beta$ defined by $R_{\alpha,\chi^{(\beta)}}(x,y) = (x+\alpha, y+\chi^{(\beta)}(x))$ for $\alpha \notin \mathbb{Q}$, $\beta > 0$.

The goal here is to identify all the locally finite, σ-finite, $R_{\alpha,\chi^{(\beta)}}$-invariant measures. Write $\chi_n^{(\beta)} := \sum_{k=0}^{n-1} \chi^{(\beta)} \circ R_\alpha^k$.

We recall some information about the continued fraction expansion

$$\alpha = \cfrac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}}$$

of $\alpha \in [0,1) \setminus \mathbb{Q}$. This can be found in [Kh].

The positive integers a_n are called the **partial quotients** of α.

Define $p_n, q_n \in \mathbb{Z}_+$, $\gcd(p_n, q_n) = 1$ by

$$\frac{p_n}{q_n} := \cfrac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\cdots + 1/a_n}}}$$

then

$$q_0 = 1,\ q_1 = a_1,\ q_{n+1} = a_{n+1}q_n + q_{n-1};$$
$$p_0 = 0,\ p_1 = 1,\ p_{n+1} = a_{n+1}p_n + p_{n-1};$$
$$\frac{p_{2n}}{q_{2n}} < \alpha < \frac{p_{2n+1}}{q_{2n+1}} \quad \text{and} \quad \frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}} = \frac{(-1)^{n+1}}{q_n q_{n+1}}.$$

The rationals p_n/q_n are called the **convergents** of α, and the numbers q_n are called (principal) **denominators** of α.

Recall the Denjoy–Koksma inequality, that $\|F_{q_n}\|_\infty \le \bigvee_{\mathbb{T}} F$ for any function F: $\mathbb{T} \to \mathbb{R}$ of bounded variation ($\bigvee_{\mathbb{T}} F < \infty$) such that $\int_{\mathbb{T}} F(t)dt = 0$. In particular, $|\chi_{q_n}^{(\beta)}| \le 2(\beta+1)$.

1.1 PROPOSITION: *$\forall\ \alpha \notin \mathbb{Q}$, $\beta > 0$ and $\eta > 0$, $\exists$ a unique $(\eta^{\chi^{(\beta)}}, R_\alpha)$-conformal probability measure $\mu = \mu_{\alpha,\beta,\eta} \in \mathcal{P}(\mathbb{T})$.*

Proposition 1.1 follows from a more general "folklore theorem" (pointed out to the authors by J-P. Conze and K. Schmidt):

THEOREM: *Let $\alpha \notin \mathbb{Q}$ and suppose that h: $\mathbb{T} \to \mathbb{R}$ has bounded variation and $\int_{\mathbb{T}} h(x)dx = 0$. There exists a unique (e^h, R_α)-conformal $\mu \in \mathcal{P}(\mathbb{T})$. Moreover, μ is non-atomic.*

Proof: We first prove existence.

Let Γ be the (countable) set of discontinuities of h and let $\Gamma_\infty := \bigcup_{n\in\mathbb{Z}} R_\alpha^n \Gamma$. As shown in [Ke]:

$\exists$ X a compact metric space, $T\colon X \to X$ a homeomorphism, $\pi\colon X \to [0,1)$ continuous and finite to one, $H\colon X \to \mathbb{R}$ continuous such that

(i) $\pi \circ T = R_\alpha \circ \pi$, (ii) $\forall\, x \notin \Gamma_\infty$, $|\pi^{-1}\{x\}| = 1$ and, $H(\pi^{-1}x) = h(x)$.

It follows from the Denjoy–Koksma inequality that

(iii) $|H_{q_n}(x)| \le \bigvee_{\mathbb{T}} h \;\forall\, x \in X \setminus \pi^{-1}\Gamma_\infty$ and hence (by continuity) $\forall\, x \in X$.

By theorem 4.1 in [Sc2], $\exists\, \mu \in \mathcal{P}(X)$ and $c \in \mathbb{R}$ such that $\mu \circ T \sim \mu$ and $d\mu \circ T/d\mu = e^{H+c}$. Since

$$1 = \mu(T^{q_n}X) = \int_X e^{H_{q_n}+cq_n}\,d\mu \asymp e^{cq_n}$$

as $n \to \infty$, we must have $c = 0$.

We claim that μ is nonatomic. Otherwise $\exists\, x \in X$ with $\mu(\{x\}) > 0$ whence $\exists\, \nu \in \mathcal{P}(X)$, $\nu \ll \mu$ with $\nu = \sum_{n\in\mathbb{Z}} a_n \delta_{T^n x}$ where $a_n > 0$. By $d\mu \circ T/d\mu = e^H$, $a_n = ce^{H_n(x)}$ for some $c > 0$ entailing $\nu(X) \ge c\sum_{n\in\mathbb{Z}} e^{H_{q_n}(x)} = \infty$ and contradicting $\nu \in \mathcal{P}(X)$.

Now define $\nu \in \mathcal{P}(\mathbb{T})$ by $\nu = \mu \circ \pi^{-1}$. It follows that ν is nonatomic, whence $\nu(\Gamma_\infty) = 0$ and $\nu \circ R_\alpha \sim \nu$ and $d\nu \circ R_\alpha/d\nu = e^h$ ν-a.e..

Existence and nonatomicity are now established and we turn to the proof of unicity.

We prove that if $\nu \circ R_\alpha \sim \nu$ and $d\nu \circ R_\alpha/d\nu = e^h$ ν-a.e., then R_α is ν-ergodic. This suffices since nonunicity implies existence of ρ with $\rho \circ R_\alpha \sim \rho$ and $d\rho \circ R_\alpha/d\rho = e^h$ ρ-a.e., and R_α not ρ-ergodic.

As above, ν is non-atomic, and by minimality of R_α, $\nu(J) > 0$ $\forall$ intervals J. Thus if $\pi\colon [0,1) \to [0,1)$ is defined by $\pi(x) := \nu((0,x))$ then π is an orientation preserving homeomorphism of $\mathbb{T}$, and $\nu \circ \pi^{-1} =$ Lebesgue measure. It follows that $S = \pi \circ R_\alpha \circ \pi^{-1}$ is absolutely continuous with $S' = e^{h\circ\pi}$ and by theorem 2b in [dM-vS] S is ergodic with respect to the Lebesgue measure. It follows that R_α is ergodic (ν). ∎

Remark: The $(\eta^{\chi^{(\beta)}}, R_\alpha)$-conformal $\mu = \mu_{\alpha,\beta,\eta} \in \mathcal{P}(\mathbb{T})$ can also be obtained using the methods of [Her] (as in [N1] and [N2]):

Define the continuous $f = f_{\eta,\beta}\colon \mathbb{R} \to \mathbb{R}$ by

$$f_{\eta,\beta}(x) = \begin{cases} \eta \cdot x & x \in [0, a(\eta,\beta)) \\ \eta^{-\beta}(x - a(\eta,\beta)) + a(\eta,\beta) & x \in [a(\eta,\beta), 1) \end{cases}$$

where $a(\eta,\beta) := (\eta^\beta - 1)/(\eta^{\beta+1} - 1)$ (this value of a is forced by the slopes, and continuity of $f_{\eta,\beta}$).

By the theory of rotation numbers, $\exists\ 0 < b < 1$, and an orientation preserving homeomorphism $\xi\colon \mathbb{T} \to \mathbb{T}$ with $\xi(0) = 0$, $\xi(1) = 1$ such that $\xi^{-1} \circ f_\alpha \circ \xi = R_\alpha$ where $f_\alpha := f_{\eta,\beta} + b$.

It can be shown that if $\mu := m \circ \xi$, then $d\mu \circ R_\alpha / d\mu = \eta^\chi$.

INVARIANT MEASURES FOR THE CYLINDER FLOW $R_{\alpha,\chi^{(1)}}$. Recall that $q \in \mathbb{N}$ is called a **Legendre** denominator for α if $\exists\ p \in \mathbb{N}$ such that $|\alpha - p/q| < 1/2q^2$. This is because of Legendre's theorem that a Legendre denominator for α is a principal denominator for α.

1.2 SUBLEMMA: *Suppose that q is an odd Legendre denominator for α, then* $|\chi_q^{(1)}| \equiv 1$.

Proof: in case $|\alpha - p/q| < 1/2q^2$.

Firstly $\{kp/q \mod 1\colon 0 \le k \le q-1\} =: \{0 = a_1 < a_2 < \cdots < a_q < 1\}$ with $a_i := k_i p/q$; and

$$\{kp/q + 1/2 \mod 1\colon 0 \le k \le q-1\} =: \{0 = b_1 < b_2 < \cdots < b_q < 1\}$$

satisfy $a_1 < b_1 < a_2 < b_2 < \cdots < a_q < b_q < 1$ with $b_i - a_i = a_{i+1} - b_i = 1/2q$.

Now let k_i, ℓ_i $(0 \le i \le q-1)$ be such that $a_i = k_i p/q \mod 1$ and $b_i = \ell_i p/q \mod 1$. Set $\bar{a}_i := k_i \alpha \mod 1$ and $\bar{b}_i = \ell_i \alpha \mod 1$.

We claim that $\bar{a}_1 < \bar{b}_1 < \bar{a}_2 < \bar{b}_2 < \cdots < \bar{a}_q < \bar{b}_q < 1$. The reason for this is that $|k\alpha - kp/q| < 1/2q$ $(0 \le k \le q-1)$ whence in case $\alpha > p/q$,

$$a_i < \bar{a}_i < a_i + \frac{1}{2q} = b_i < \bar{b}_i < b_i + \frac{1}{2q} = a_{i+1} < \cdots,$$

and in case $\alpha < p/q$,

$$a_{i+1} > \bar{a}_{i+1} > a_{i+1} - \frac{1}{2q} = b_i > \bar{b}_i > b_i - \frac{1}{2q} = a_i > \cdots.$$

Now $\chi_q^{(1)}$ is a step function with points of discontinuity $1 - \bar{a}_1 > 1 - \bar{b}_1 > 1 - \bar{a}_2 > 1 - \bar{b}_2 > \ldots > 1 - \bar{a}_q > 1 - \bar{b}_q \ge 0$, and jumps of $+2$ at $1 - \bar{a}_i$ $(1 \le i \le q)$ and -2 at $1 - \bar{a}_i$ $(1 \le i \le q)$. The values of $\chi_q^{(1)}$ are of the form $\{v, v+2\}$ for some $v \in \mathbb{Z}$. The only $v \in \mathbb{Z}$ permitted by the condition $\int_{\mathbb{T}} \chi_q^{(1)}(t)dt = 0$ is $v = -1$. Thus $|\chi_q^{(1)}| \equiv 1$. ∎

This subsection is based on the following lemma, which is obtained from sublemma 1.2 and the well known fact that there are infinitely many odd Legendre denominators for any $\alpha \notin \mathbb{Q}$:

1.3 LEMMA: *$\exists\ n_k \to \infty$ such that $|\chi^{(1)}_{q_{n_k}}| \equiv 1\ \forall\ k \geq 1$.*

Remark: Sublemma 1.2 can be strengthened: $|\chi^{(1)}_q| \equiv 1$ whenever q is an odd principal denominator for α. This is shown in [N1].

For $\eta > 0$, $\alpha \in \mathbb{T} \setminus \mathbb{Q}$ define the $R_{\alpha,\chi^{(1)}}$-invariant, Maharam measure $m_{\alpha,\eta}$ on $\mathcal{B}(\mathbb{T} \times \mathbb{Z})$ by

$$m_{\alpha,\eta}(A \times \{n\}) := \eta^{-n}\mu_{\alpha,1,\eta}(A).$$

1.4 THEOREM: *(1) $\forall\ \alpha \notin \mathbb{Q}$ and $\eta > 0$, $(\mathbb{T} \times \mathbb{Z}, \mathcal{B}(\mathbb{T} \times \mathbb{Z}), m_{\alpha,\eta}, R_{\alpha,\chi^{(1)}})$ is a conservative, ergodic measure preserving transformation.*

(2) If m is a locally finite measure on $\mathbb{T}\times\mathbb{Z}$ such that $(\mathbb{T}\times\mathbb{Z}, \mathcal{B}(\mathbb{T}\times\mathbb{Z}), m, R_{\alpha,\chi^{(1)}})$ is ergodic and measure preserving, then $\exists\ \eta, c > 0$ such that $m = cm_{\alpha,\eta}$.

Proof: The ergodicity of $(\mathbb{T}\times\mathbb{Z}, \mathcal{B}(\mathbb{T}\times\mathbb{Z}), m_{\alpha,\eta}, R_{\alpha,\chi^{(1)}})$ was established in [N1] (see [C-K] and also [A-K] for the Lebesgue case $\eta = 1$) and is standard using [Sc1] and Lemma 1.3:

$\exists\ n_k \to \infty$ (odd Legendre denominators) such that $|\chi^{(1)}_{n_k}| \equiv 1$ and

$$\mu_{\alpha,1,\eta}(R^{n_k}_\alpha A \triangle A) \to 0 \quad \forall\ A \in \mathcal{B}(\mathbb{T}).$$

We prove (2). Let m be an $R_{\alpha,\chi^{(1)}}$-ergodic locally finite measure on $\mathbb{T}\times\mathbb{Z}$. We claim that $m = cm_{\alpha,\eta}$ for some $c, \eta > 0$. By the Basic Lemma and Proposition 1.1, it suffices to prove that $H := \{n \in \mathbb{Z}\colon m \circ Q_n \sim m\} = \mathbb{Z}$.

Suppose that $H \neq \mathbb{Z}$, and write $m_k(A) := m(A\times\{k\})$. Then $\overline{m} := m_{-1}+m_1 \perp m_0$. $\exists\ U \subset \mathbb{T}$ open, such that $m_0(U) = 1$ and $\overline{m}(U) < 1/5$, whence $\exists\ I \subset \mathbb{T}$, an open interval such that $\overline{m}(I) < m_0(I)/5$.

Given $0 < p < 1$ and an open interval $L = (a - r, a + r)$, denote by L_p the subinterval $(a - pr, a + pr)$. Note that if $x \in L_p$ and $|y| < (1 - p)|L|/2$ then $x + y \in L$.

$\exists\ 0 < p < 1$ such that $m_0(I_p) > m_0(I)/2$. By Lemma 1.3, $\exists\ k \geq 1$ such that $\|q_{n_k}\alpha\| < (1 - p)|I|/2$ and $|\chi^{(1)}_{q_{n_k}}| \equiv 1$.

It follows that

$$R^{q_{n_k}}_{\alpha,\chi^{(1)}}(I_p \times \{0\}) \subset I \times \{-1, 1\}$$

whence

$$\begin{aligned} m_0(I)/2 < m_0(I_p) &= m(I_p \times \{0\}) \\ &= m(R^{q_{n_k}}_{\alpha,\chi^{(1)}}(I_p \times \{0\})) \leq m(I \times \{-1, 1\}) \\ &= \overline{m}(I) < m_0(I)/5. \end{aligned}$$

The contradiction shows the impossibility of $H \neq \mathbb{R}$, and thus proves (2). ∎

INVARIANT MEASURES FOR THE CYLINDER FLOW $R_{\alpha,\chi^{(\beta)}}$. For $\eta,\beta>0$, $\alpha\in\mathbb{T}\setminus\mathbb{Q}$ define the locally finite measure $m_{\alpha,\beta,\eta}$ on $\mathcal{B}(\mathbb{T}\times\mathbb{R})$ by

$$dm_{\alpha,\beta,\eta}(x,y):=\eta^{-y}d\mu_{\alpha,\beta,\eta}(x)dy.$$

Evidently $m_{\alpha,\beta,\eta}\circ R_{\alpha,\chi^{(\beta)}}=m_{\alpha,\beta,\eta}$.

Fix $\alpha\in\mathbb{T}\setminus\mathbb{Q}$. For $t\in\mathbb{R}$, consider the set

$$L(t)=L_\alpha(t):=\{a\in[0,1)\colon\exists\; n_k\to\infty,\; q_{n_k}t \mod 1\to a\}$$

(where $\{q_n\colon n\geq 1\}$ are the denominators of α).

Theorem 4.1 in [Ku-Ni] implies that $L(t)=[0,1)$ for Lebesgue-a.e. $t\in\mathbb{R}$. Moreover, it is shown in [Kr-Li] that for $\alpha\notin\mathbb{Q}$ with bounded partial quotients and $t\in\mathbb{R}$, $L(t)$ is finite iff $t\in\mathbb{Q}+\alpha\mathbb{Q}$.

1.5 LEMMA: *If $a\in L(\beta/\beta+1)$ is positive and $q_{n_k}\beta/(\beta+1)\mod 1\to a$, then $\forall\, x\in\mathbb{T}$, all limit points of $\{\chi^{(\beta)}_{q_{n_k}}(x)\}_{k\geq 1}$ are contained in*

$$\{(\beta+1)(N-a):\; N=-1,0,1,2\}.$$

Proof: Let $\epsilon>0$, $N\in\mathbb{Z}$ and suppose that $|q_n\frac{\beta}{\beta+1}-N-a|<\epsilon$; then $q_n\beta=(\beta+1)(N+a\pm\epsilon)$, whence

$$\chi^{(\beta)}_{q_n}=(\beta+1)(1_{[0,\frac{\beta}{\beta+1})})_{q_n}-q_n\beta=(\beta+1)(L-a\pm\epsilon)$$

where $L:=(1_{[0,\frac{\beta}{\beta+1})})_{q_n}-N\in\mathbb{Z}$.

Recalling that $|\chi^{(\beta)}_{q_n}|\leq 2(\beta+1)$ we see that $-2+a-\epsilon\leq L\leq 2+a+\epsilon$. It follows that for $a\in(0,1)$ and sufficiently small $\epsilon>0$: $L=-1,0,1,2$. ∎

1.6 THEOREM: *Suppose that $\alpha\notin\mathbb{Q}$, $\beta>0$ are such that $L(\frac{\beta}{\beta+1})$ is infinite; then:*

(1) *For each $\eta>0$, $(\mathbb{T}\times\mathbb{R},\mathcal{B}(\mathbb{T}\times\mathbb{R}),m_{\alpha,\beta,\eta},R_{\alpha,\chi^{(\beta)}})$ is a conservative, ergodic measure preserving transformation.*

(2) *If m is a locally finite measure on $\mathbb{T}\times\mathbb{R}$ such that*

$$(\mathbb{T}\times\mathbb{R},\mathcal{B}(\mathbb{T}\times\mathbb{R}),m,R_{\alpha,\chi^{(\beta)}})$$

is ergodic and measure preserving, then $\exists\;\eta,c>0$ such that $m=cm_{\alpha,\beta,\eta}$.

Proof: The ergodicity of $(\mathbb{T}\times\mathbb{R},\mathcal{B}(\mathbb{T}\times\mathbb{R}),m_{\alpha,\beta,\eta},R_{\alpha,\chi^{(\beta)}})$ was established in [N2] and in [St] for $\eta=1$ (Lebesgue measure).

We prove (2). Let m be an $R_{\alpha,\chi^{(\beta)}}$-ergodic locally finite measure on $\mathbb{T}\times\mathbb{R}$. We claim that $m = cm_{\alpha,\beta,\eta}$ for some $c,\eta > 0$. By the Basic Lemma and Proposition 1.1, it suffices to prove that $H := \{a \in \mathbb{R}\colon m\circ Q_a \sim m\} = \mathbb{R}$.

Suppose otherwise; then $H \neq \mathbb{R}$ and $\exists\, q \geq 0$ such that $H = q\mathbb{Z}$. It follows that $\exists\, a \in L(\frac{\beta}{\beta+1})$ with $(\beta+1)(N-a) \notin H\ \forall\ N = -1,0,1,2$.

(Else $L(\frac{\beta}{\beta+1}) \subseteq [0,1]\cap\bigcup_{N=-1,0,1,2}(N+\frac{1}{\beta+1}H)$, whence since $H = q\mathbb{Z}$, $L(\frac{\beta}{\beta+1})$ is finite in contradiction to our assumptions).

Fix such an a and set $E := \{(\beta+1)(N-a)\colon N = -1,0,1,2\}$; then $E \subset \mathbb{R}\setminus H$. Set $\overline{m} := \sum_{j\in E} m\circ Q_j$; then $\overline{m} \perp m$ and $\exists\ K \subset \mathbb{T}\times\mathbb{R}$ compact such that $m(K) > 0$, $\overline{m}(K) = 0$. $\exists\ U \subset \mathbb{T}\times\mathbb{R}$ open and precompact, such that $K \subset U$ and $\overline{m}(U) < m(K)/5\mathfrak{n}$ where $\mathfrak{n}$ is the Besicovitch covering constant for $\mathbb{R}^2$ ($\mathfrak{n} \leq 16$, see [W-Z]).

For each $z = (x,y) \in K$ $\exists$ an open rectangle $R(z)$ with diameter less than $\frac{1}{2}\min\{|j-j'|\colon j,j' \in E,\ j \neq j'\}$ such that $z \in R(z) \subset U$. $\exists$ a finite set $\Gamma \subset K$ such that $K \subset V := \bigcup_{z\in\Gamma} R(z)$ and $\sum_{z\in\Gamma} 1_{R(z)} \leq \mathfrak{n}$. Evidently $\overline{m}(V) < m(K)/5\mathfrak{n}$.

We claim that (at least) one of the rectangles $R = R(z)$ ($z \in \Gamma$) has the property that $\overline{m}(R) < m(R)/5$, else

$$\overline{m}(V) \geq \frac{1}{\mathfrak{n}}\sum_{z\in\Gamma}\overline{m}(R(z)) \geq \frac{1}{5\mathfrak{n}}\sum_{z\in\Gamma} m(R(z)) \geq \frac{1}{5\mathfrak{n}} m(K).$$

It follows from the restriction on the diameter of R that $\{Q_jR\colon j \in E\}$ is a disjoint collection, whence, if $S := \bigcup_{j\in E} Q_jR$, then

$$m(S) = \overline{m}(R) < m(R)/5.$$

Write $R = I\times J$ where $I \subset (0,1)$ and $J \subset \mathbb{R}$ are open intervals. Given $0 < p < 1$ and an open interval $L = (a-r, a+r)$, denote by L_p the subinterval $(a-pr, a+pr)$. Note that if $x \in L_p$ and $|y| < (1-p)|L|/2$ then $x+y \in L$.

$\exists\ 0 < p < 1$ such that $m(I_p\times J_p) > m(R)/2$. By Lemma 1.5, $\exists\ k \geq 1$ and $A \subset I_p$ such that

$$\|q_{n_k}\alpha\| < (1-p)|I|/2, \quad m(A\times J_p) > m(R)/3$$

and

$$\min_{j\in E}|\chi^{(\beta)}_{q_{n_k}}(x) - j| < (1-p)|J|/2 \quad \forall\ x \in A.$$

It follows that

$$R^{q_{n_k}}_{\alpha,\chi^{(\beta)}}(A\times J_p) \subset S$$

whence

$$m(R)/3 < m(A \times J_p) = m(R^{q_{n_k}}_{\alpha,\chi^{(\beta)}}(A \times J_p)) \geq m(S) < m(R)/5.$$

The contradiction shows the impossibility of $H \neq \mathbb{R}$. ∎

§2. Locally finite invariant measures for tail relations of skew products

Set $S := \{0, \ldots, s-1\}$ where $s \geq 2$, let $A := (A_{ij})_{S \times S}$ be a matrix of zeroes and ones such that $\forall j \exists i$ s.t. $A_{ij} = 1$ and $\forall i \exists j$ s.t. $A_{ij} = 1$. Set

$$\Sigma = \Sigma_A := \{x = (x_1, x_2, \ldots) \in S^{\mathbb{N}} : \forall i \ \ A_{x_i x_{i+1}} = 1\}.$$

We topologize Σ by considering the base of **cylinder sets**, sets of the form

$$[\epsilon_1, \ldots, \epsilon_n] = \{x \in \Sigma : \ x_k = \epsilon_k \ \forall \ 1 \leq k \leq n\}$$

where $\epsilon_1, \ldots, \epsilon_n \in S$.

Let T: $\Sigma \to \Sigma$ be the left shift, $T(x_1, x_2, \ldots) = (x_2, x_3, \ldots)$. The topological dynamical system (Σ, T) is called a **subshift of finite type**. Henceforth we assume that it is topologically mixing. This is equivalent to the existence of some $M \in \mathbb{N}$ such that all the entries of the matrix A^M are positive (see, e.g., [Bo]).

An **admissible word** (of length n) is an element $(\epsilon_1, \ldots, \epsilon_n) \in S^n$ (or **word**) satisfying $A_{\epsilon_j, \epsilon_{j+1}} = 1 \ \forall \ 1 \leq j \leq n-1$. Note that a cylinder $[\epsilon_1, \ldots, \epsilon_n]$ is nonempty iff its corresponding word $(\epsilon_1, \ldots, \epsilon_n)$ is admissible. We denote the collection of admissible words of length n, or **paths** of length $n-1$ (the number of steps), by $\mathcal{W}_n$, and set $\mathcal{W} := \bigcup_n \mathcal{W}_n$.

Consider T's **tail relation**

$$\mathfrak{T} = \mathfrak{T}(T) := \{(x, y) \in \Sigma^2 : \exists \ n \geq 0, \ T^n x = T^n y\}.$$

Consider the reverse lexicographic order on $\mathfrak{T}(T)$-equivalence classes:

$$x \prec y \text{ iff } \exists n_0 \text{ s.t. } x_{n_0} < y_{n_0} \text{ and } x_n = y_n \text{ for any } n > n_0.$$

It is easy to see that for any fixed $x \prec y$ there are finitely many z such that $x \prec z \prec y$, so the type of ordering in each equivalence class is either $\mathbb{Z}$, or $\mathbb{Z}^+$, or $\mathbb{Z}^-$. Let Σ_∞, $\Sigma_{-\infty}$ be the set of maximal and minimal elements of $(\Sigma, \prec)$, respectively. To characterize these elements, introduce functions $P_{\max}$, $P_{\min}$: $S \to S$

$$P_{\max}(a) = \max\{i \in S : A_{i,a} = 1\} \quad \text{and} \quad P_{\min}(a) = \min\{i :\in S : A_{i,a} = 1\}.$$

Note that

$$x \in \Sigma_\infty \Longrightarrow x_{n-1} = P_{\max}(x_n) \quad \text{for all } n.$$

It follows that there are at most s maximal points (similarly, at most s minimal points) and all of them are periodic.

The **adic transformation** τ: $\Sigma \setminus \Sigma_\infty \to \Sigma \setminus \Sigma_{-\infty}$ assigns to each x the smallest y strictly greater than x. Specifically, given $x \in \Sigma \setminus \Sigma_\infty$, $\exists\, \ell \geq 1$ such that

$$x_j = P_{\max}(x_{j+1})\ \forall\ 1 \leq j \leq \ell - 1 \text{ and } x_\ell < P_{\max}(x_{\ell+1}),$$

and we set $\tau(x) := (y_1, y_2, \ldots)$ where the y_k's are defined reverse-inductively:

$$y_k = \begin{cases} x_k, & k \geq \ell + 1, \\ \min\{i \in S:\ i > x_\ell,\ A_{i,x_{\ell+1}} = 1\}, & k = \ell, \\ P_{\min}(y_{k+1}), & 1 \leq k \leq \ell - 1. \end{cases}$$

It is convenient to restrict τ to $\Sigma_0 := \Sigma \setminus \bigcup_{j\geq 0} \tau^j \Sigma_{-\infty} \setminus \bigcup_{j \leq 0} \tau^j \Sigma_\infty$.

Remarks: (1) It is possible to visualize Σ as the space of infinite paths in the directed graph Γ with vertex set $S \times \mathbb{N}$ and edges connecting (b, n) to $(c, n+1)$ iff $A_{b,c} = 1$.

(2) If $\Omega = S^{\mathbb{N}}$ is the full shift, and V is the adding machine, then τ is the induced transformation V_{Σ_0} in the sense that $\tau(x) = V^{F(x)}(x)$ $(x \in \Sigma_0)$ where $F(x) := \min\{n \geq 1 \colon V^n(x) \in \Sigma_0\}$.

(3) Adic transformations were introduced in [V1] (see also [V2] and [V3]) in the more general setting of non-stationary Markov chains.

Let $\mathbb{G}$ be a locally compact, Abelian, Polish topological group. For $f\colon \Sigma \to \mathbb{G}$, consider the skew product transformation $T_f\colon \Sigma \times \mathbb{G} \to \Sigma \times \mathbb{G}$ defined by $T_f(x, y) := (Tx, y + f(x))$.

Now T_f's tail relation is

$$\begin{aligned} \mathfrak{T}(T_f) &:= \{((x,y),(x',y')) \in (\Sigma \times \mathbb{G})^2 \colon \exists\, n \geq 0,\ T_f^n(x,y) = T_f^n(x',y')\} \\ &= \{((x,y),(x',y')) \in (\Sigma \times \mathbb{G})^2 \colon (x,x') \in \mathfrak{T}(T),\ y' - y = \psi_f(x,x')\} \end{aligned}$$

where the **symmetric (or tail) cocycle** $\psi_f\colon \mathfrak{T} \to \mathbb{G}$ is defined by

$$\psi_f(x, x') := \sum_{k=0}^{\infty} (f(T^k x) - f(T^k x')).$$

Consider $\tau_{\phi_f}\colon \Sigma_0 \times \mathbb{G} \to \Sigma_0 \times \mathbb{G}$ defined by

$$\tau_{\phi_f}(x, y) := (\tau x, y + \phi_f(x)),$$

where $\phi_f(x) = \psi_f(x, \tau x)$. It is easy to see that the orbits of τ_{ϕ_f} are exactly the equivalence classes of $\mathfrak{T}(T_f) \cap (\Sigma_0 \times \mathbb{G})^2$.

In this section we identify the τ_{ϕ_f}-invariant locally finite measures for certain $f\colon \Sigma \to \mathbb{G}$ which we now proceed to describe. For $f\colon \Sigma \to \mathbb{S}^1$ and $k \geq 1$, let

$$v_k(f) := \sup\{|f(x) - f(y)| : x, y \in \Sigma,\ x_j = y_j\ \forall\ 1 \leq j \leq k\}.$$

The collection of $\mathbb{S}^1$-valued **Hölder continuous** functions on Σ is

$$\mathcal{H}_{\mathbb{S}^1} := \{f\colon \Sigma \to \mathbb{S}^1\colon \exists\ 0 \leq \theta < 1,\ v_n(f) = O(\theta^n) \text{ as } n \to \infty\}$$

and the collection of $\mathbb{S}^1$-valued functions on Σ **with summable variations** is

$$\mathcal{F}_{\mathbb{S}^1} := \{f\colon \Sigma \to \mathbb{S}^1\colon \sum_{k=1}^{\infty} v_k(f) < \infty\}.$$

The collection of $\mathbb{G}$-valued **Hölder continuous** functions on Σ is

$$\mathcal{H}_{\mathbb{G}} := \{f\colon \Sigma \to \mathbb{G}\colon \forall \gamma \in \widehat{\mathbb{G}}, \gamma \circ f \in \mathcal{H}_{\mathbb{S}^1}\},$$

and the collection of $\mathbb{G}$-valued function on Σ with **summable variations** is

$$\mathcal{F}_{\mathbb{G}} := \{f\colon \Sigma \to \mathbb{G}\colon \forall \gamma \in \widehat{\mathbb{G}}, \gamma \circ f \in \mathcal{F}_{\mathbb{S}^1}\}.$$

Finally, a function $f\colon \Sigma \to \mathbb{G}$ is said to **have finite memory** if $\exists\ N \geq 1$ such that $f(x) = f(x_1, \ldots, x_N)$.

These notions coincide with the usual notions of Hölder continuity, summable variations, and finite memory of $\mathbb{R}$-valued functions in the case $\mathbb{G} = \mathbb{R}$. Clearly, every $f\colon \Sigma \to \mathbb{R}$ which is Hölder (respectively with summable variations, finite memory) in the usual sense is also Hölder (respectively with summable variations, finite memory) according to the definition above. To see the other direction, note that if $f \in \mathcal{F}_{\mathbb{R}}$ then f is continuous, because $\gamma \circ f$ is continuous for every $\gamma \in \widehat{\mathbb{R}}$. Therefore $\|f\|_\infty < \infty$. Now consider the character $\gamma(x) = e^{2\pi i x/10\|f\|_\infty}$ to see that $v_n(f) \asymp v_n(\gamma \circ f)$. It follows that if f is Hölder continuous (respectively with summable variations, finite memory) in the above sense, then it is Hölder (respectively with summable variations, finite memory) in the ordinary sense.

Clearly, if $\alpha\colon \mathbb{G} \to \mathbb{R}$ is a continuous homomorphism, then $\alpha \circ \mathcal{H}_{\mathbb{G}} \subset \mathcal{H}_{\mathbb{R}}$ and $\alpha \circ \mathcal{F}_{\mathbb{G}} \subset \mathcal{F}_{\mathbb{R}}$.

A measurable function $f\colon \Sigma \to \mathbb{G}$ is called **periodic** if $\exists\ \gamma \in \widehat{\mathbb{G}}$, $z \in \mathbb{S}^1$ and $g\colon \Sigma \to \mathbb{S}^1$ measurable, not constant, such that $\gamma \circ f = z\overline{g}g \circ T$. It is known that in the case $f \in \mathcal{H}_{\mathbb{G}}$, g is necessarily in $\mathcal{H}_{S^1}$. The function f is called **aperiodic** if it is not periodic.

2.1 THEOREM: *Suppose that Σ_A is topologically mixing, and that $f \in \mathcal{H}_{\mathbb{G}}$ is aperiodic. For every continuous homomorphism α: $\mathbb{G} \to \mathbb{R}$:*

(1) *there is a unique $(e^{-\alpha(\phi_f)}, \tau)$-conformal probability $\mu_\alpha \in \mathcal{P}(\Sigma_0)$;*

(2) *μ_α is non-atomic;*

(3) *τ_{ϕ_f} is ergodic with respect to the Maharam measure on $\Sigma_0 \times \mathbb{G}$ defined by $dm_\alpha(x, y) = e^{-\alpha(y)} d\mu_\alpha(x) dy$.*

Theorem 2.1 is essentially known (although we indicate the proof). Our main result in this section is

2.2 THEOREM: *Suppose that f: $\Sigma \to \mathbb{G}$ is aperiodic and has finite memory.*

If m is an ergodic, τ_{ϕ_f}-invariant locally finite measure on $\Sigma_0 \times \mathbb{G}$, then $m = cm_\alpha$ for some $c > 0$ and some continuous homomorphism α: $\mathbb{G} \to \mathbb{R}$.

The collection of all locally finite, τ_ϕ-invariant measures on $\Sigma_0 \times \mathbb{G}$ is identified by Theorems 2.1 and 2.2 as the collection of mixtures of Maharam measures. This is because by the ergodic decomposition (see, e.g., [A1]), any locally finite, τ_ϕ-invariant measure is a mixture of ergodic ones.

Conditions for aperiodicity based on [Kow] were given in §3 of [A-D1]. We'll say that a topologically mixing subshift of finite type (Σ, T) is **almost onto** if $\forall\ a, b \in S,\ \exists\ n \geq 1,\ a = s_0, a_1, \ldots, s_n = b \in S$ such that $T[s_k] \cap T[s_{k+1}] \neq \emptyset$ $(0 \leq k \leq n-1)$.

2.3 PROPOSITION: *Suppose that Σ is mixing and almost onto, and that ϕ: $\Sigma \to \mathbb{G}$ satisfies $\phi(x) = \phi(x_0)$; then either ϕ is aperiodic, or $\exists\ \gamma \in \widehat{G},\ \lambda \in S^1$ such that $\gamma \circ \phi \equiv \lambda$. In particular, if $\overline{\mathrm{Group}(\phi(\Sigma) - \phi(\Sigma))} = \mathbb{G}$, then ϕ is aperiodic.*

Some of the proofs use the theory of non-singular equivalence relations and we provide some background.

Let $(X, \mathcal{B})$ be the standard Borel space. An equivalence relation $R \subset X \times X$ is called **standard**, if R is a Borel subset of $X \times X$, that is R is in the product σ-field $\mathcal{B} \times \mathcal{B}$. For any $x \in X$, $R(x) := \{y : (x, y) \in R\}$ is the **equivalence class** of x, and for a subset $A \subset X$, $R(A) = \cup\{R(x) : x \in A)\}$ is called the **saturation** of A. The standard equivalence relation R is called **countable** if $R(x)$ is countable for any x.

For a countable, standard relation R, $A \in \mathcal{B} \implies R(A) \in \mathcal{B}$. If G is a countable group of automorphisms of X then $R_G = \{(x, g(x)) :\ x \in X,\ g \in G\}$ is a countable, standard equivalence relation, and conversely, any countable standard relation R is generated in this way by a countable group of automorphisms (see theorem 1 in [F-M]). A σ-finite measure μ is called **non-singular** for R

if $\mu(R(A)) = 0$ whenever $\mu(A) = 0$; it is called **ergodic** if, in addition, either $\mu(R(A)) = 0$ or $\mu(X \setminus R(A)) = 0$ for every $A \in \mathcal{B}$.

By a **holonomy** we mean a Borel automorphism $\phi: A \to \phi(A)$ (some $A \in \mathcal{B}$) whose graph $\Gamma(\phi) := \{(x, \phi(x)) : x \in A\}$ is a subset of R. A σ-finite measure which is non-singular with respect to R is called **invariant** for R if $\mu(A) = \mu(\phi A)$ for any holonomy ϕ. By corollary 1 in [FM], μ is invariant under R iff μ is invariant for the action of any G with $R_G = R$.

The following proposition appears in [P-S] (see also [B-M]). We use the notation $a = M^{\pm 1} b$ for the double inequality $M^{-1}a \le b \le Mb$.

2.4 PROPOSITION: *Suppose that Σ_A is topologically mixing, and $f \in \mathcal{F}_{\mathbb{R}}$. There is a unique $(e^{-\phi_f}, \tau)$-conformal probability $\mu \in \mathcal{P}(\Sigma_0)$, and there exists $M > 1$ such that*

$$(\diamond) \qquad \mu([x_1, \ldots, x_n]) = M^{\pm 1} e^{-Pn + \sum_{k=0}^{n-1} f(T^k x)} \quad \forall\, x \in \Sigma,\ n \ge 1$$

where $P = \max\{h_p(T) + \int_\Sigma f dp : p \in \mathcal{P}(\Sigma),\ p \circ T^{-1} = p\}$.

The property $(\diamond)$ is known as the **Gibbs property**. A T-invariant probability with the Gibbs property is known as a **Gibbs measure**.

As is shown in [Bo] and [R1]:

- $\exists$ a unique probability $\mu_f \in \mathcal{P}(\Sigma)$ such that $d\mu_f \circ T / d\mu_f = \lambda e^{-f}$ for some $\lambda > 0$;
- T is exact (whence τ is ergodic) with respect to μ_f;
- $\exists$ a T-invariant probability $p_f \sim \mu_f$ such that $\|\log(dp_f)/(d\mu_f)\|_\infty < \infty$; and
- $\exists$ $M > 1$ such that

$$p_f([x_1, \ldots, x_n]),\ \mu_f([x_1, \ldots, x_n]) = M^{\pm 1} e^{-Pn + \sum_{k=0}^{n-1} f(T^k x)} \quad \forall\, x \in \Sigma,\ n \ge 1$$

where P is the **topological pressure** of f given by the **variational principle**

$$P := \max\{h_p(T) + \int_\Sigma f dp :\ p \in \mathcal{P}(\Sigma),\ p \circ T^{-1} = p\} = h_{p_f}(T) + \int_\Sigma f dp_f.$$

The probability p_f is known as the **equilibrium measure** of f (being the unique maximizing T-invariant probability) and is a Gibbs measure.

Proof of Proposition 2.4: For every admissible word $c = (c_1, \ldots, c_n)$ and $x \in \Sigma$ such that $A_{c_n x_1} = 1$ let (c, x) denote the concatenation $(c_1, \ldots, c_n; x_1, x_2, \ldots)$. The proof relies on the characterization of $(e^{-\phi_f}, \tau)$-conformal measures as those measures μ for which

$$d\mu \circ \kappa / d\mu = e^{-\psi_f((a,x),(b,x))}$$

whenever $a = [a_1, \ldots, a_n]$, $b = [b_1, \ldots, b_n]$ are nonempty with $a_n = b_n$, and κ: $a \to b$ is defined by $\kappa(a, x) := (b, x)$. To check this characterization, suppose $x \in \Sigma_0$ and set $y = \tau(x)$. By the definition of the Adic Transformation, there exists some n_0 such that for every $z \in [x_1, \ldots, x_{n_0}]$,

$$\tau(z) = (y_1, \ldots, y_{n_0}; z_{n_0+1}, z_{n_0+2}, \ldots).$$

Equivalently, $\tau|_{[x_1,\ldots,x_{n_0}]} = \kappa$ where κ: $a \to b$ is defined as before with $a = (x_1, \ldots, x_{n_0})$ and $b = (y_1, \ldots, y_{n_0})$. For $z = (a, w)$, the conformality condition now reads

$$\frac{d\mu \circ \kappa}{d\mu}(a, w) = e^{-\phi_f(a,w)} = e^{-\psi_f((a,w),\tau(a,w))} = e^{-\psi_f((a,w),(b,w))}.$$

Existence

We claim that μ_f is $(e^{-\phi_f}, \tau)$-conformal. To establish this, suppose that a, b and κ are as in the above. We show that

$$\frac{d\mu_f \circ \kappa}{d\,\mu_f}(a, x) = e^{-\psi_f((a,x),(b,x))}.$$

For v_a: $T[a_n] \to a$ defined by $v_a(x) := (a, x)$ we have that $v_a^{-1} = T^n : a \to T[a_n]$ whence

$$\frac{d\mu_f \circ v_a}{d\,\mu_f}(x) = \left(\frac{d\mu_f \circ T^n}{d\mu_f}(a, x)\right)^{-1} = \lambda^{-n} e^{\sum_{k=0}^{n-1} f\circ T^k(a,x)},$$

and, since $\kappa = v_b \circ v_a^{-1}$,

$$\begin{aligned}\frac{d\mu_f \circ \kappa}{d\,\mu_f}(a, x) &= \frac{d\mu_f \circ v_b}{d\,\mu_f}(T^n(a, x))\frac{d\mu_f \circ T^n}{d\,\mu_f}(a, x)\\ &= \frac{d\mu_f \circ v_b}{d\,\mu_f}(b, x)\frac{d\mu_f \circ T^n}{d\,\mu_f}(a, x)\\ &= e^{-\psi_f((a,x),(b,x))}.\end{aligned}$$

Uniqueness

Suppose that $\nu \in \mathcal{P}(\Sigma_0)$ is $(e^{-\phi_f}, \tau)$-conformal. It follows that if

$$a = [a_1, \ldots, a_n] \quad \text{and} \quad b = [b_1, \ldots, b_n]$$

are both nonempty with $a_n = b_n$, and κ: $a \to b$ is defined by $\kappa(a, x) := (b, x)$ then

$$\frac{d\nu \circ \kappa}{d\,\nu}(a, x) = e^{-\psi_f((a,x),(b,x))},$$

whence $\exists\ M>1,\ K_n(s)>0\ (n\ge 1,\ s\in S)$ such that

$$\nu([x_1,\dots,x_n]) = M^{\pm 1}K_n(x_n)e^{\sum_{k=0}^{n-1} f(T^k x)} \quad \forall\ n\ge 1,\ x\in\Sigma_0.$$

But

$$e^{\sum_{k=0}^{n-1} f(T^k x)} = M^{\pm 1}e^{Pn}p_f([x_1,\dots,x_n])$$

and so

$$\nu([x_1,\dots,x_n]) = M^{\pm 2}K_n(x_n)e^{Pn}p_f([x_1,\dots,x_n]).$$

It follows that

$$\nu(T^{-n}[s]) = M^{\pm 2}K_n(s)e^{Pn}p_f([s])$$

whence $\sum_{s\in S} K_n(s) \asymp e^{-Pn}$, $\nu([x_1,\dots,x_n]) \le M' p_f([x_1,\dots,x_n])$, and $\nu \ll \mu_f$.

Writing $F := d\nu/d\mu_f$, we see from $d\nu\circ\tau/d\nu = d\mu_f\circ\tau/d\mu_f$ that $F\circ\tau = F \mod \mu_f$, whence by ergodicity $F\equiv 1$ and $\nu=\mu_f$. ∎

Proof of Theorem 2.1: Let $\alpha\colon \mathbb{G}\to\mathbb{R}$ be a continuous homomorphism. By Proposition 2.4 and its proof, there is a unique $(e^{-\alpha(\phi_f)},\tau)$-conformal probability $\mu_\alpha\in\mathcal{P}(\Sigma_0)$, and this measure is equivalent to the (invariant) equilibrium probability measure $p_{\alpha(\phi_f)}$.

It is shown in [G] (see also [A-D2]) that if $f\in\mathcal{H}_{\mathbb{G}}$ is aperiodic then T_f is exact with respect to $m = p\times m_{\mathbb{G}}$ where p is some equilibrium measure on Σ. In particular, T_f is exact with respect to $m_\alpha \sim p_{\alpha(\phi_f)}\times m_{\mathbb{G}}$, whence τ_{ϕ_f} is ergodic with respect to m_α. ∎

Now let $f\colon \Sigma\to\mathbb{G}$ be measurable. If $\exists$ a globally supported, σ-finite T_f-nonsingular measure m on $\Sigma\times\mathbb{G}$ such that $(\Sigma\times\mathbb{G},\mathcal{B}(\Sigma\times\mathbb{G}),m,T_f)$ is exact, then f is aperiodic.

To see this, suppose otherwise, that $\exists\ \gamma\in\widehat{\mathbb{G}}$, $z\in\mathbb{S}^1$ and $g\colon\Sigma\to\mathbb{S}^1$ Hölder continuous, not constant, such that $\gamma\circ f = z\overline{g}g\circ T$. Consider $G\in L^\infty(\Sigma\times\mathbb{G})$ defined by $G(x,y) := \overline{g}(x)\gamma(y)$; then G is not m-a.e. constant and $G\circ T_f = zG$. Thus T_f is not weakly mixing and hence not exact (in particular, G is $T_f^{-n}\mathcal{B}$-measurable $\forall\ n\ge 0$).

2.5 PROPOSITION: *Let $f\in\mathcal{F}_{\mathbb{G}}$. Any τ_{ϕ_f}-invariant, ergodic locally finite measure m on $\Sigma\times\mathbb{G}$ with $H_m=\mathbb{G}$ is proportional to a Maharam measure, and the existence of such implies that f is aperiodic.*

Proof: Let m be a τ_{ϕ_f}-invariant, ergodic locally finite measure on $\Sigma\times\mathbb{G}$ with $H_m=\mathbb{G}$. By the Basic Lemma, m has the form $dm(x,y) = e^{\alpha(y)}d\mu(x)dy$ where $\alpha\colon\mathbb{G}\to\mathbb{R}$ is a continuous homomorphism and μ is $(e^{\alpha\circ\phi_f},\tau)$-conformal, whence

$(e^{\phi_\alpha \circ f}, \tau)$-conformal. Proposition 2.4 shows that the (unique) conformal measure has the Gibbs property $(\diamond)$, and is therefore globally supported on Σ. It follows that m is globally supported and so, as shown above, f is aperiodic. ∎

By possibly changing the state space, we may assume that $f(x) = g(x_1, x_2)$ in the assumptions of Theorem 2.2. The proof of Theorem 2.2 uses Lemma 2.6 below.

For $u: \Sigma \to \mathbb{S}^1$ and $\ell \geq 1$, set $u_\ell(x) := \prod_{j=0}^{\ell-1} u(T^j x)$.

2.6 Lemma: *Assume $u: \Sigma \to \mathbb{S}^1$ is Hölder continuous, then either:*
(1) $\exists\, z \in \mathbb{S}^1$, $g: \Sigma \to \mathbb{S}^1$ *Hölder continuous, such that* $u = z\bar{g}g \circ T$;
or
(2) $\exists\, \epsilon > 0,\ \ell_0 \geq 1$ *such that* $\forall\, \ell \geq \ell_0,\ x \in \Sigma, \exists\, y \in \Sigma$ *satisfying*

$$x_1 = y_1,\ T^\ell y = T^\ell x \quad \text{and} \quad |u_\ell(y) - u_\ell(x)| \geq \epsilon.$$

Proof: Let $L: C(X) \to C(X)$ be the operator $(Lf)(x) = \sum_{Ty=x} f(y)$. Ruelle's Perron–Frobenius theorem implies that $\exists \lambda > 0$, a Borel probability measure ν and a positive continuous function h such that $L^*\nu = \lambda\nu$, $Lh = \lambda h$, $\int h d\nu = 1$. Moreover, ν and h are uniquely determined up to a multiplicative constant, and $\forall f \in C(X)$, $\lambda^{-n} L^n f \to h \int f d\nu$ uniformly on Σ. Let P be the operator

$$Pf = \lambda^{-1} L\Big(\frac{h}{h \circ T} f\Big).$$

It is not difficult to check that $P1 = 1$ and that if $\varphi: \Sigma \to \mathbb{S}^1$ is continuous and $P\varphi \equiv 1$, then $\varphi \equiv 1$.

Let P_u be the perturbed operator $P_u f := P(uf)$. One checks that for every n,

$$P_u^n f = P^n(u_n f) = \lambda^{-n} L^n \Big(\frac{h}{h \circ T^n} u_n f\Big).$$

In [G-H] it is shown that either $\exists z \in \mathbb{S}^1$ and $\exists g: \Sigma \to \mathbb{S}^1$ Hölder continuous such that $P_u(g) = zg$, or $\|P_u^n f\|_\infty \to 0$ for every $f \in C(\Sigma)$.

We show that if (2) fails, then $\|P_u^n f\|_\infty \not\to 0$ for some $f \in C(\Sigma)$. This proves the lemma, because it implies that $\exists z \in \mathbb{S}^1$ and $\exists g: \Sigma \to \mathbb{S}^1$ Hölder continuous such that $P_u(g) = zg$, and $P_u(g) = zg$ implies that $P(\frac{g}{zg \circ T} u) = 1$, whence $u = z\frac{g \circ T}{g}$. $hd\nu$ is known to be ergodic and globally supported (see, e.g., [R2]). Therefore $|g| \equiv 1$ and (1) follows.

If (2) fails, $\forall \epsilon > 0$ there are $x^{(k)} \in \Sigma$, $1 \leq \ell_k \uparrow \infty$ such that if

$$y \in \Sigma,\ k \geq 1,\ x_1^{(k)} = y_1 \text{ and } T^{\ell_k} x^{(k)} = T^{\ell_k} y$$

then

$$|u_{\ell_k}(x^{(k)}) - u_{\ell_k}(y)| < \epsilon.$$

By possibly passing to a subsequence, we can ensure that $\exists a \in S \forall k \geq 1, x_1^{(k)} = a$. Set

$$\gamma_0 := \min\left\{\frac{h(x)}{h(y)} : x, y \in \Sigma\right\}.$$

Since Σ is compact, $\gamma_0 > 0$ and

$$\begin{aligned}
\|P_u^{\ell_k} 1_{[a]}\|_\infty &\geq |(P_u^{\ell_k} 1_{[a]})(T^{\ell_k} x^{(k)})| \\
&= \lambda^{-\ell_k} \left| \sum_{y \in \Sigma,\ T^{\ell_k} y = T^{\ell_k} x^{(k)}} \frac{h(y)}{h(T^{\ell_k} y)} u_{\ell_k}(y) 1_{[a]}(y) \right| \\
&\geq \gamma_0 \lambda^{-\ell_k} \sum_{y \in \Sigma,\ T^{\ell_k} y = T^{\ell_k} x^{(k)}} (1 - |u_{\ell_k}(y) - u_{\ell_k}(x^{(k)})|) 1_{[a]}(y) \\
&\geq \gamma_0 (1-\epsilon) \lambda^{-\ell_k} L^{\ell_k} 1_{[a]}(x^{(k)})
\end{aligned}$$

Since $\lambda^{-n}(L^n 1_{[a]})(x)$ tends uniformly to $h(x)\nu[a]$, we have that

$$\liminf_{n \to \infty} \|P_u^n 1_{[a]}\|_\infty \geq \gamma_0 (1 - \epsilon) \min_{x \in \Sigma} h(x) > 0$$

as required. ∎

If u: $\mathcal{W}_2(\Sigma) \to \mathbb{S}^1$, $u(x) = u(x_1, x_2)$ and $a \in \mathcal{W}_{n+1}$ is a path $a = (a_1, \ldots, a_{n+1})$ of length n, then u_n is constant on a. We denote

$$u_n(a) := u_n|_a = \prod_{i=1}^{n} u(a_i, a_{i+1}).$$

In Lemma 2.6, when $u(x) = u(x_1, x_2)$, (2) has the combinatorial form:
(2′) $\exists\ \ell_0$ such that $\forall\ \ell \geq \ell_0$, paths $a = (a_1, \ldots, a_{\ell+1}) \in \mathcal{W}_\ell$, $\exists$ a path $b = (b_1, \ldots, b_{\ell+1}) \in \mathcal{W}_\ell$ such that $a_1 = b_1$, $a_{\ell+1} = b_{\ell+1}$ and $u_\ell(a) \neq u_\ell(b)$.

Proof of Theorem 2.2: By the Basic Lemma and Proposition 2.4, it suffices to show that $H_m = \mathbb{G}$.

Suppose otherwise that $H \neq \mathbb{G}$; then $\exists\ \gamma \in \widehat{\mathbb{G}}$, $\gamma \neq 1$ such that $\gamma|_H \equiv 1$.

Since m is τ_{ϕ_f}-invariant, it is also $\mathfrak{T}(T_f)$-invariant and if κ: $A \to \kappa(A)$ ($A \in \mathcal{B}(\Sigma \times \mathbb{G})$ is a $\mathfrak{T}(T_f)$-holonomy, then $m(\kappa(A)) = m(A)$.

Using aperiodicity and Lemma 2.6, we fix $\ell \geq 1$ so large that $\forall$ paths $a = (a_1, \ldots, a_{\ell+1}) \in P_\ell$, $\exists$ a path $b = b_a = (b_1, \ldots, b_{\ell+1}) \in P_\ell$ such that $a_1 = b_1$, $a_{\ell+1} = b_{\ell+1}$ and $\gamma \circ f_\ell(a) \neq \gamma \circ f_\ell(b)$, equivalently $f_\ell(a) - f_\ell(b) \notin H$.

Set $J := \{f_\ell(a) - f_\ell(b_a) : a \in P_\ell\}$; then $J \subset \mathbb{G} \setminus H$ and J is finite. Set $\overline{m} := \sum_{j\in J} m \circ Q_j$; then $\overline{m} \perp m$ and $\exists\, K \subset \Sigma \times \mathbb{G}$ compact such that $m(K) > 0$, $\overline{m}(K) = 0$.

Set $M = |W_\ell|$. Approximating K by larger precompact open sets, we see that $\exists\, U \subset \Sigma \times \mathbb{G}$ open, $\overline{U}$ compact such that $K \subset U$ and $\overline{m}(U) < m(K)/2M$.

For each $z = (x, y) \in K$ $\exists$ a set $W(z) = C(z) \times V(z)$ of form cylinder $\times$ open such that $z \in W(z) \subset U$. By compactness of K $\exists\, z_1, \ldots, z_N$ such that $K \subset V := \bigcup_{k=1}^N W(z_k)$. We claim that V is a disjoint union of sets of form cylinder $\times$ open. To see this, let L be the maximum length of the cylinders $C(z_1), \ldots, C(z_N)$; then $V = \bigcup_{k=1}^N W(z_k) = \bigcup_{k=1}^N \bigcup_{c\in W_L,\ c\subset C(z_k)} c \times V(z_k)$ — a disjoint union. Thus $K \subset V$ and $\overline{m}(V) < m(V)/2M$.

It follows that $\exists$ a set $C \times W$ of form cylinder $\times$ open such that $m(C \times W) > 0$ and $\overline{m}(C \times W) < m(C \times W)/2M$, otherwise V would not have these properties.

Since $C \times W = \bigcup_{a\in W_\ell} (C, a) \times W$, $\exists\, a \in W_\ell$ such that $m((C, a) \times W) \geq m(C \times W)/M$.

Next, $\exists\, b = (b_1, \ldots, b_{\ell+1}) \in W_\ell$ such that $a_1 = b_1$, $a_{\ell+1} = b_{\ell+1}$ and $f_\ell(a) - f_\ell(b) \in J$.

Define τ: $(C, a) \times W \to C \times \mathbb{G}$ by $\tau((C, a, x), y) := ((C, b, x), y + f_\ell(b) - f_\ell(a))$. Evidently τ is a $\mathfrak{T}(T_f)$-holonomy and so by assumption, $m(\tau((C, a) \times W)) = m((C, a) \times W) \geq m(C \times W)/M$.

On the other hand, $\tau((C, a) \times W)) \subset Q_{f_\ell(b)-f_\ell(a)}C \times W$ whence

$$\begin{aligned}\frac{m(C \times W)}{M} &\leq m(\tau(C, a) \times W) \leq m(Q_{f_\ell(b)-f_\ell(a)}C \times W)\\ &\leq \overline{m}(C \times W) < \frac{m(C \times W)}{2M}\end{aligned}$$

and $\frac{1}{2} > 1$. This contradiction establishes Theorem 2.2. ∎

Remark: The proof of theorem 2.2 establishes the (stronger) statement:

Suppose that $f\colon \Sigma \to \mathbb{G}$ *is aperiodic and has finite memory.*

If m *is an ergodic,* $\mathfrak{T}(T_f)$*-invariant locally finite measure on* $\Sigma \times \mathbb{G}$*, then* $m = cm_\alpha$ *for some continuous homomorphism* $\alpha\colon \mathbb{G} \to \mathbb{R}$ *and some* $c > 0$.

We conclude this section with an application of Theorem 2.2 to the "Markov–Pascal-adic" transformations considered in [P-S].

Let $\Sigma = \Sigma_A$ be a mixing subshift of finite type and let $f\colon \Sigma \to \mathbb{G}$. We use the notation

$$x_i^j := (x_i, x_{i+1}, \ldots, x_j),\ x_i^\infty = (x_i, x_{i+1}, \ldots) \quad (x \in \Sigma).$$

Recall from [P-S], the equivalence relations: $S_A^+ \subset \Sigma_A \times \Sigma_A$ defined by

$$S_A^+ = \{(x,y) \in \Sigma_A \times \Sigma_A : \exists\, n \geq 1,\ x_n^\infty = y_n^\infty,$$
$$(y_1, \ldots, y_n) \text{ a permutation of } (x_1, \ldots, x_n)\};$$

and $S_A^f \subset \Sigma_A \times \Sigma_A$ defined by

$$S_A^f := \{(x,y) \in \Sigma_A \times \Sigma_A : \exists\, n \geq 1,\ x_n^\infty = y_n^\infty,\ f_n(x) = f_n(y)\}.$$

Evidently $S_A^+ = S_A^{F^\#}$ where $F^\# : \Sigma \to \mathbb{Z}^S$ is defined by $F^\#(x_1, x_2, \ldots)_i := \delta_{i,x_1}$ $(i \in S)$.

Suppose that $\mathbb{G}$ is discrete. Evidently if $f \colon \Sigma \to \mathbb{G}$ then

$$(x,y) \in S_A^f \iff ((x,0),(y,0)) \in \mathfrak{T}(T_f)$$

whence

$$(x,y) \in S_A^f \cap \Sigma_0^2 \iff \exists\, n \in \mathbb{Z},\ (y,0) = \tau_{\phi_f}^n(x,0)$$

and $S_A^f \cap \Sigma_0^2$ is generated by the induced transformation $(\tau_{\phi_f})_{\Sigma_0 \times \{0\}}$.

We claim (as in [P-S]) that if f has finite memory and $\alpha \colon \mathbb{G} \to \mathbb{R}$ is a homomorphism, then μ_α is S_A^f-invariant, ergodic.

To see this, recall from Theorem 2.1 that m_α is τ_{ϕ_f}-invariant, ergodic; whence $m_\alpha|_{\Sigma_0 \times \{0\}}$ is $(\tau_{\phi_f})_{\Sigma_0 \times \{0\}}$-invariant, ergodic; whence our claim (since $m_\alpha(A \times \{0\}) = \mu_\alpha(A)$).

2.7 COROLLARY: *Suppose that $f \colon \Sigma \to \mathbb{Z}^d$ $(d \geq 1)$ is aperiodic and has finite memory.*

If $\nu \in \mathcal{P}(\Sigma)$ is S_A^f-invariant and ergodic, then $\nu = \mu_\alpha$ for some homomorphism $\alpha \colon \mathbb{Z}^d \to \mathbb{R}$.

Proof: We'll deduce this from Theorem 2.2. To do this, we show first that $\nu(\Sigma \setminus \Sigma_0) = 0$.

We claim that all S_A^f-equivalence classes are infinite (this implies that ν is non-atomic, whence $\nu(\Sigma \setminus \Sigma_0) = 0$ as this set is countable).

To see this we'll need the **symmetrization** F of f defined on the mixing SFT $\Sigma \times \Sigma$ by $F(x,y) = f(x) - f(y)$ $(F \colon \Sigma \times \Sigma \to \mathbb{Z}^d)$. Evidently F has finite memory.

We claim that F is aperiodic. If not, then

$$e^{2\pi i q(f(x)-f(y))} = z\frac{g(Tx,Ty)}{g(x,y)} \quad (x, y \in \Sigma)$$

for some $q \in \mathbb{Z}$, $q \neq 0$, $z \in \mathbb{S}^1$, $g \colon \Sigma \times \Sigma \to \mathbb{S}^1$ and then

$$e^{2\pi i q(f_N(x)-f_N(y))} = z^N \frac{g(T^N x, T^N y)}{g(x,y)} \quad \forall\ N \geq 1,\ x, y \in \Sigma.$$

Choosing $N \geq 1$ and periodic points $y = T^N y, y' = T^{N+1} y'$, we have, for all $x \in \Sigma_0$,

$$\begin{aligned} e^{2\pi i q f_N(Tx)} &= e^{2\pi i q f_N(y)} z^N \frac{g(T^{N+1}x, y)}{g(Tx, y)}, \\ e^{2\pi i q f_{N+1}(x)} &= e^{2\pi i q f_{N+1}(y')} z^{N+1} \frac{g(T^{N+1}x, y')}{g(x, y')}, \end{aligned}$$

whence (!)

$$e^{2\pi i q f(x)} = Z \frac{G(Tx)}{G(x)}$$

contradicting the aperiodicity of f.

Let μ be the measure of maximal entropy on Σ and let $P \colon L^1(\mu \times \mu) \to L^1(\mu \times \mu)$ be the transfer operator. By the local limit theorem of [G-H], $\exists\ c > 0$ such that $\forall$ cylinders a, $b \subset \Sigma$,

$$n^{d/2} P^n(1_{(a \times b) \cap [F_n = 0]})(x, y) \to c\mu(a)\mu(b) \text{ uniformly on } \Sigma \times \Sigma \text{ as } n \to \infty.$$

Now fix $x \in \Sigma$ and $N \geq 1$; then $\exists\ n_N$ such that

$$n^{d/2} P^n(1_{([a] \times [b]) \cap [F_n = 0]})(T^n x, T^n x) \geq \frac{c}{2} \mu([a])\mu([b]) \ \forall\ a,\ b \in \mathcal{W}_N,\ n \geq n_N$$

whence

$$\begin{aligned} |\{y \in X \colon (x, y) \in S_A^f\}| &\geq |\{y \in X \colon T^{n_N} y = T^{n_N} x,\ F_{n_N}(x, y) = 0\}| \\ &\geq |\mathcal{W}_N| \to \infty \end{aligned}$$

as $N \to \infty$ and establishing our claim.

As mentioned above, $\nu(\Sigma \setminus \Sigma_0) = 0$ and the probability $\overline{\nu}$ on $\Sigma_0 \times \{0\}$ defined by $\overline{\nu}(A \times \{0\}) = \nu(A)$ is $(\tau_{\phi_f})_{\Sigma \times \{0\}}$-invariant and ergodic. Define the measure m on $\Sigma_0 \times \mathbb{Z}^d$ by

$$m(A) := \int_{\Sigma_0} \sum_{k=0}^{\varphi - 1} 1_A \circ \tau_{\phi_f}^k \, d\overline{\nu}.$$

The measure m is evidently locally finite. By Kac's formula, it is τ_{ϕ_f}-invariant, and by Kakutani's tower theorem it is τ_{ϕ_f}-ergodic (see, e.g., [A1]). Thus, by Theorem 2.2, $m = m_\alpha$ for some homomorphism $\alpha \colon \mathbb{Z}^d \to \mathbb{R}$. It follows that $\nu = \mu_\alpha$. ∎

2.8 COROLLARY: *Suppose Σ is a mixing, almost onto SFT.*

If $\nu \in \mathcal{P}(\Sigma)$ is S_A^+-invariant and ergodic, then $\nu = \mu_\alpha$ for some homomorphism $\alpha: \mathbb{Z}^d \to \mathbb{R}$.

Proof: As mentioned above, $S_A^+ = S_A^{F^\#}$ where $F^\#: \Sigma \to \mathbb{Z}^S$ is defined by $F^\#(x)_i := \delta_{i,x_1}$ $(i \in S)$. Since evidently $\text{Group}(F^\#(\Sigma) - F^\#(\Sigma)) = \mathbb{Z}^S$, $F^\#$ is aperiodic by Proposition 2.3. The result follows from Corollary 2.7. ∎

Remark: Theorems 2.9 and 2.11 in [P-S] both follow from Corollary 2.8. In both cases, $S = \{0,1\}$, $d = 1$ and Σ is almost onto.

§3. A logarithmic ergodic theorem

As in §2, let $S = \{0, 1, \ldots, s-1\}$ where $s \in \mathbb{N}$, let $A: S \times S \to \{0,1\}$ be an irreducible and aperiodic matrix and let $\Sigma = \Sigma_A^+ \subset S^{\mathbb{N}}$ be the corresponding (topologically mixing) subshift of finite type. Recall that $T: \Sigma \to \Sigma$ is the left shift, $\tau: \Sigma_0 \to \Sigma_0$ is the induced adding machine, where Σ_0 is obtained from Σ as in §2.

In this section, we consider the asymptotic properties of τ_{ϕ_f}, where $f: \Sigma \to \mathbb{R}^d$ is an aperiodic Hölder continuous function, with respect to Maharam measures. It will be convenient to use the supremum norm on $\mathbb{R}^d$, $\|(x_1, \ldots, x_d)\| := \max_{1 \le k \le d} |x_k|$.

Fix some $\alpha \in \mathbb{R}^d$ and consider the Maharam measure $m_\alpha: \mathcal{B}(\Sigma \times \mathbb{R}^d) \to [0, \infty]$ defined by $dm_\alpha(x,y) = e^{-\alpha \cdot y} d\mu(x) dy$ where $\mu = \mu_\alpha$ is the $(e^{\alpha \cdot f}, \tau)$-conformal measure.

As mentioned above, the aperiodicity of f implies that T_f is exact with respect to m_α. It follows that τ_{ϕ_f} is ergodic with respect to m_α (generating the the tail relation for T_f) and also conservative (being invertible, ergodic and preserving a non-atomic measure).

We prove the

LOGARITHMIC ERGODIC THEOREM:

$$(\ddagger) \qquad \frac{\log \sum_{k=0}^{n-1} F \circ \tau_{\phi_f}^k}{\log n} \longrightarrow \frac{h_{p_\alpha}(T)}{h_{top}(T)} \quad m_\alpha\text{-a.e. as } n \to \infty$$

$\forall F \in L^1(m_\alpha)_+$ where p_α is the equilibrium measure of $\alpha \cdot f$.

It will sometimes be convenient to denote

$$S_n(F) = S_n^{(\tau_{\phi_f})}(F) := \sum_{k=0}^{n-1} F \circ \tau_{\phi_f}^k.$$

The proof of the logarithmic ergodic theorem is based on the following two reductions:

Firstly, it is sufficient to establish (‡) for a single $F_0 \in L^1(m_\alpha)_+$ since then, by the ratio ergodic theorem,

$$\frac{S_n(F)}{S_n(F_0)} \to \frac{\int_X F dm}{\int_X F_0 dm} \quad \text{a.e.},$$

whence $\log S_n(F) \sim \log S_n(F_0)$ a.e.

Secondly, in order to establish (‡) for $F_0 \in L^1(m_\alpha)_+$, it is sufficient to find:

- sets A, $B \in \mathcal{B}(\Sigma \times \mathbb{R}^d)$ with $m_\alpha(A)$, $m_\alpha(B) > 0$ and
- (random) subsequences M_k: $A \to \mathbb{N}$, k: $B \to \mathbb{N}$ such that M_k, $N_k \uparrow \infty$, $\log M_k \sim \log M_{k+1}$, $\log N_k \sim \log N_{k+1}$ as $k \to \infty$;

satisfying

$$(\bar{\ddagger}) \qquad \limsup_{k\to\infty} \frac{\log S_{M_k}(F_0)}{\log M_k} \le \frac{h_{p_\alpha}(T)}{h_{top}(T)} \quad \text{on } A,$$

and

$$(\underline{\ddagger}) \qquad \liminf_{k\to\infty} \frac{\log S_{N_k}(F_0)}{\log N_k} \ge \frac{h_{p_\alpha}(T)}{h_{top}(T)} \quad \text{on } B.$$

To see this, note that $\forall\ n$ large $\exists\ k = k_n \ge 1$ such that $M_k \le n \le M_{k+1}$, whence

$$\frac{\log S_n(F_0)}{\log n} \le \frac{\log S_{M_{k+1}}(F_0)}{\log M_k}$$

and it follows from $\log M_k \sim \log M_{k+1}$ that

$$\limsup_{n\to\infty} \frac{\log S_n(F_0)}{\log n} \equiv \limsup_{k\to\infty} \frac{\log S_{M_k}(F_0)}{\log M_k}.$$

Similarly

$$\liminf_{n\to\infty} \frac{\log S_n(F_0)}{\log n} \equiv \liminf_{k\to\infty} \frac{\log S_{N_k}(F_0)}{\log N_k}.$$

The functions

$$\limsup_{n\to\infty} \frac{\log S_n(F_0)}{\log n} \quad \text{and} \quad \liminf_{n\to\infty} \frac{\log S_n(F_0)}{\log n}$$

are τ_{ϕ_f}-invariant, whence so are the sets

$$\overline{A} := \left[\limsup_{n\to\infty} \frac{\log S_n(F_0)}{\log n} \le \frac{h_{p_\alpha}(T)}{h_{top}(T)}\right], \quad \underline{B} := \left[\liminf_{n\to\infty} \frac{\log S_n(F_0)}{\log n} \ge \frac{h_{p_\alpha}(T)}{h_{top}(T)}\right].$$

By ergodicity, both sets (containing sets of positive measure by ($\bar{\ddagger}$) and ($\underline{\ddagger}$)) are of full measure and ($\ddagger$) is established for F_0.

In the Main Lemma (below), we'll establish ($\bar{\ddagger}$) and ($\underline{\ddagger}$) for $F_0 = 1_{\Sigma\times B_M(0)}$ and $A = B = \Sigma \times B_{M'}(0)$ (for some $M, M' > 0$ where $B_M(0) := \{y \in \mathbb{R}^d : \|y\| \le M\}$) using the local limit theorem of [G-H] and large deviations techniques.

The subsequences M_k, N_k are related to some counting functions, which we proceed to define.

We define the **counting functions** $\Lambda_n\colon \Sigma_A \to \mathbb{N}$ by

$$\Lambda_n(x) := \min\{N \ge 1\colon \{(\tau^k x)_1^n\colon 0 \le k \le N-1\} = \mathcal{W}_n\}$$

where $\mathcal{W}_n$ denotes the collection of admissible words of length n (as in §2). The reader may easily verify that in case Σ is a full shift, $\Lambda_n \equiv s^n = |\mathcal{W}_n|$ and consequently $k \mapsto (\tau^k x)_1^n$ defines a bijection $\{0,1,\ldots,s^n-1\} \leftrightarrow \mathcal{W}_n\ \forall\ x \in \Sigma$. In other words, τ generates $\mathfrak{T}$-equivalence classes efficiently. For a mixing topological Markov shift, as shown by the counting proposition below, the situation is analogous.

3.1 Counting Proposition: *Suppose that Σ_A is a mixing topological Markov shift, and that $L \ge 1$ is such that all entries of A^L are positive; then for $x \in \Sigma_0$:*

$$|\mathcal{W}_n| \le \Lambda_n(x) < 3|\mathcal{W}_{n+L}|.$$

Proof: The left hand inequality follows directly from the definition of $\Lambda_n(x)$. To see the right side, assume by way of contradiction that $\Lambda_n(x) \ge 3|\mathcal{W}_{n+L}|$; then there is a word $\underline{a} \in \mathcal{W}_{n+L}$ and $0 \le k_1 < k_2 < k_3 \le \Lambda_n(x) - 1$ such that $\tau^{k_j} x \in [\underline{a}]$ for $k = 1,2,3$. Set $\tau^{k_j} x = (\underline{a}, z^{(j)})$; then $z^{(1)} \prec z^{(2)} \prec z^{(3)}$. For every $\underline{\varepsilon} \in \mathcal{W}_n$ choose some point of the form $x(\underline{\varepsilon}) = (\underline{\varepsilon}, w_0^{L-1}, z^{(2)})$ where w_0^{L-1} is some word which makes $x(\underline{\varepsilon})$ admissible. Clearly, $\tau^{k_1} x \prec x(\underline{\varepsilon}) \prec \tau^{k_3} x$. Thus $\mathcal{W}_n$ is spanned by $\tau^j x$ for $0 \le j < k_3$ in contradiction to the minimality of $\Lambda_n(x)$. The right hand inequality is thus proved. ∎

Set $\lambda := \exp h_{top}(\Sigma)$ and assume without loss of generality that $L > 2$, where L is as in Proposition 3.1. For every $x \in \Sigma_0$ and n large enough set

$$u_n(x) := \min\{u > n + L\colon x_{u-1} < P_{\max}(x_u)\}, \quad u_n' := u_n - L,$$
$$\ell_n(x) := \max\{\ell < n + L\colon x_{\ell-1} < P_{\max}(x_\ell)\}, \quad \ell_n' := \ell_n - L,$$

where $P_{\max}$ is as in §2. By possibly adding a vector of constants to f, we may assume that $\int f dp_\alpha = (0,\ldots,0)$ (note that neither ϕ_f nor p_α change when a constant is added to f).

Set

$$\rho_n := (n+L) - \ell_n, \quad \sigma_n := u_n - (n+L).$$

3.2 LEMMA: *$\exists\, M_0 \in \mathbb{R}_+$ such that*

$$\limsup_{n\to\infty} \frac{\rho_n}{\log n},\ \limsup_{n\to\infty} \frac{\sigma_n}{\log n} \le M_0 \ \text{ a.e.}$$

Proof: We prove this only for σ_n, the proof for ρ_n being essentially the same. Set $P := P_{top}(\alpha \cdot f)$. Recall that Σ_∞ consists of at most s points, all of which are periodic. Set $\Sigma_\infty = \{x^{(1)}, x^{(2)}, \ldots, x^{(r)}\}$ and let p be the least common multiple of the periods of $x^{(i)}$; then $r \le s$ and, for every $x \in \Sigma_\infty$, $T^p x = x$. Define by induction $P^{k+1}_{\max} = P_{\max} \circ P^k_{\max}$. By the definition of σ_n, if $\sigma_n(x) > b$ then

$$T^{n+L}x \in [P^b_{\max}(x_{n+b+L}), \ldots, P_{\max}(x_{n+b+L}), x_{n+b+L}].$$

For $b > s$ the word $(P^b_{\max}(x_{n+b+L}), \ldots, P^{b-s}_{\max}(x_{n+b+L}))$ is made of a repeating period, hence is the prefix of a maximal point. Applying this argument to $b_n := \lfloor M_0 \log n \rfloor$, using the invariance of p_α and the structure of Σ_∞, we have

$$p_\alpha[\sigma_n > b_n] \le \sum_{i=1}^{r} p_\alpha\left[x^{(i)}_0, \ldots, x^{(i)}_{b_n - s}\right].$$

Since p_α is a Gibbs measure and since for every i, $T^p x^{(i)} = x^{(i)}$,

$$p_\alpha\left[x^{(i)}_0, \ldots, x^{(i)}_{b_n-s}\right] = O\left(e^{\alpha\cdot f_{b_n}(x^{(i)}) - b_n P}\right) = O\left(e^{\frac{b_n}{p}\,\alpha\cdot f_p(x^{(i)}) - b_n P}\right)$$

whence

$$p_\alpha[\sigma_n > M_0 \log n] = O\left(\sum_{i=1}^{r} n^{M_0\left(\frac{\alpha\cdot f_p(x^{(i)})}{p} - P\right)}\right). \tag{1}$$

It follows from the unicity of the equilibrium measure that $\alpha \cdot f_p\left(x^{(i)}\right)/p < P$. Thus, the exponents in (1) are all negative and for M_0 large enough,

$$\sum_{n=1}^{\infty} p_\alpha[\sigma_n > M_0 \log n] < \infty.$$

The result follows. ∎

The next lemma is the main lemma, being the version of ($\bar{\ddagger}$) and ($\underline{\ddagger}$)) that we prove. Let

$$B := 2L\|f\| + \sum_{k=1}^{\infty} v_k(\alpha \cdot f)$$

where $v_k(\alpha \cdot f) = \sup\{|\alpha\cdot\varphi(x) - \alpha\cdot\varphi(y)| : x_0^{k-1} = y_0^{k-1}\}$.

3.3 MAIN LEMMA: *There exists $M > 2B$ for which*

(2) $$\limsup_{n\to\infty} \frac{1}{n} \log S_{\Lambda_{u'_n}-1}(1_{\Sigma\times B_M(0)}) \le h_{p_\alpha}(T) \quad m_\alpha\text{-a.e. on: } \Sigma \times B_{M/2}(0),$$

(3) $$\liminf_{n\to\infty} \frac{1}{n} \log S_{\Lambda_{\ell'_n}-1}(1_{\Sigma\times B_M(0)}) \ge h_{p_\alpha}(T) \quad m_\alpha\text{-a.e. on: } \Sigma \times B_{M/2}(0).$$

The rest of this section is devoted to the proof of the main lemma. Set

$$\begin{aligned} U_N(x,M) &:= \{\underline{\varepsilon} \in \mathcal{W}_N : \forall y \in [\underline{\varepsilon}]\ \|f_N(y) - f_N(x)\| < M\}, \\ V_N(x,M) &:= \left\{y \in \Sigma_0 : \forall z \in \left[y_0^{N-1}\right]\ \|f_N(z) - f_N(x)\| < M\right\} \\ &= \bigcup_{\underline{\varepsilon}\in U_n(x,M)} [\underline{\varepsilon}]. \end{aligned}$$

3.4 LEMMA: *For each $M > 2B$, $\exists M_1, M_2 > 0$ such that for all $(x,t) \in \Sigma_0 \times B_{M/2}(0)$ and n large enough,*

$$|U_{\ell'_n}(x, M_2)| \le \sum_{j=0}^{\Lambda_{\ell'_n}-1} 1_{\Sigma\times B_M(0)}\left(\tau^j_{\phi_f}(x,t)\right)$$

and

$$\sum_{j=0}^{\Lambda_{u'_n}-1} 1_{\Sigma\times B_M(0)}\left(\tau^j_{\phi_f}(x,t)\right) \le |U_{u_n}(x, M_1)|.$$

Proof: Fix some $x \in \Sigma_0$ and $t \in \mathbb{R}^d$. We estimate

$$A_N := \sum_{j=0}^{\Lambda_N-1} 1_{\Sigma\times B_M(0)}\left(\tau^j_{\phi_f}(x,t)\right) \quad \text{for } N = u'_n, \ell'_n.$$

It follows from the minimality of Λ_n that $\forall 0 \le j \le \Lambda_N - 1$, $T^{N+L}\left(\tau^j x\right) = T^{N+L}(x)$, because all the entries of A^L are positive, so $\forall \underline{\varepsilon} \in \mathcal{W}_n$ there exists $\underline{c} \in \mathcal{W}_{L-1}$ such that $(\underline{\varepsilon}, \underline{c}, P_{\max}(x_{N+L}), x_{N+L}^\infty)$ is admissible and strictly larger than x. Thus $\sum_{k=0}^{j-1} \phi_f\left(\tau^k x\right) = f_{N+L}(x) - f_{N+L}\left(\tau^j x\right)$, whence

$$A_N = \sharp\left\{0 \le j \le \Lambda_N - 1 : \left\|f_{N+L}\left(\tau^j x\right) - f_{N+L}(x) - t\right\| \le M\right\}.$$

Since for $j < \Lambda_N$, $(\tau^j x)_{N+L}^\infty = x_{N+L}^\infty$, the map $j \mapsto (\tau^j x)_0^{N+L-1}$ is 1-1, so $A_N = |B_N|$ where

$$B_N = \left\{\left(\tau^j x\right)_0^{N+L-1} : \left\|f_{N+L}\left(\tau^j x\right) - f_{N+L}(x) - t\right\| \le M;\ 0 \le j < \Lambda_N\right\}.$$

We now prove the required inequalities. Setting $N = u'_n$ in the above inequality we have $\forall (x,t) \in \Sigma_0 \times B_{M/2}(0)$

$$\begin{aligned} A_{u'_n} =& |\{(\tau^j x)_0^{u_n-1} : \|f_{u_n}(\tau^j x) - f_{u_n}(x) - t\| \le M;\ 0 \le j < \Lambda_{u'_n}\}| \\ \le& |\{\underline{\varepsilon} \in \mathcal{W}_{u_n} : \forall y \in [\underline{\varepsilon}] \|f_{u_n}(y) - f_{u_n}(x)\| \le \frac{3}{2}M + B\}| \end{aligned}$$

and the upper inequality follows with $M_1 := B + 3M/2$.

Using the same argument for $N = \ell'_n$ one shows that for all $(x,t) \in \Sigma_0 \times B_{M/2}(0)$ and n large enough so that ℓ'_n is well defined,

$$A_{\ell'_n} \ge \left|\left\{ (\tau^j x)_0^{\ell_n - 1} : \forall y \in \left[(\tau^j x)_0^{\ell_n-1}\right] \|f_{\ell'_n}(y) - f_{\ell'_n}(x)\| < \frac{M}{2} - B\ 0 \le j < \Lambda_{\ell'_n} \right\}\right|.$$

Since $\left\{(\tau^j x)_0^{\ell'_n - 1} : 0 \le j \le \Lambda_{\ell'_n} - 1\right\} = \mathcal{W}_{\ell'_n}$,

$$A_{\ell'_n} \ge \left|\left\{\underline{\varepsilon} \in \mathcal{W}_{\ell'_n} : \forall y \in [\underline{\varepsilon}] \|f_{\ell'_n}(y) - f_{\ell'_n}(x)\| < \frac{M}{2} - B\right\}\right|$$

and this is the lower inequality for $M_2 := M/2 - B$. ∎

The following lemma provides, together with Lemma 3.4, the upper estimation (2) in the Main Lemma.

3.5 Lemma: $\forall M > 0\ \overline{\lim}_{n\to\infty} \frac{1}{n} \log |U_n(x,M)| \le h_{p_\alpha}(T)$ m_α-a.e.

Proof: Since p_α is the Gibbs measure for $\alpha \cdot f$, there exists some constant K such that for all $y \in \left[\varepsilon_0^{n-1}\right]$,

$$K^{-1} e^{\alpha\cdot f_n(y) - nP(\alpha\cdot f)} \le p_\alpha\left[\varepsilon_0^{n-1}\right] \le K e^{\alpha\cdot f_n(y) - nP(\alpha\cdot f)}.$$

By the definition of U_n, for every $\varepsilon_0^{n-1} \in U_n(x,M)$ and $y \in \left[\varepsilon_0^{n-1}\right]$,

$$p_\alpha\left[\varepsilon_0^{n-1}\right] \asymp e^{\alpha\cdot f_n(y) - nP(\alpha\cdot f)} \asymp e^{\alpha\cdot f_n(x) - nP(\alpha\cdot f)}$$

whence

$$|U_n(x,M)| \asymp \frac{p_\alpha(V_n(x,M))}{e^{\alpha\cdot f_n(x) - nP(\alpha\cdot f)}}.$$

Thus, $|U_n(x,M)| = O\left(e^{nP(\alpha\cdot f) - \alpha\cdot f_n(x)}\right)$. Recall that according to our assumptions, $\int \alpha \cdot f dp_\alpha = 0$, so $P(\alpha \cdot f) = h_{p_\alpha}(T)$. The lemma follows since by the ergodicity of p_α, for almost all $x \in \Sigma_0$, $\alpha \cdot f_n(x) = o(n)$. ∎

We now turn to the lower estimation (3) in the Main Lemma.

For every $N \in \mathbb{N}$ and $\delta > 0$ set

$$E_N(\delta) := \left\{ y \in \Sigma_0 : p_\alpha \left[y_0^{N-1} \right] > e^{-N(h_{p_\alpha}(T)-\delta)} \right\}.$$

By the definition of $U_N(x, M)$, $\forall M > 0, x \in \Sigma_0$ and $N > 0$,

$$|U_N(x,M)| \geq e^{N(h_{p_\alpha}(T)-\delta)} \left[p_\alpha(V_N(x,M)) - p_\alpha(E_N(\delta)) \right]. \tag{4}$$

We prove that

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log p_\alpha(E_n(\delta)) < 0 \quad p_\alpha\text{-a.e.},$$

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log p_\alpha(V_n(x,M)) = 0 \quad p_\alpha\text{-a.e.}$$

Since for almost all $x \in \Sigma_0$, $\ell'_n(x) \sim n$, (3) will follow from this, (4) and Lemma 3.4.

3.6 LEMMA: $\overline{\lim}_{n\to\infty} \frac{1}{n} \log p_\alpha(E_n(\delta)) < 0$ p_α-a.e.

Proof: p_α is a Gibbs measure, so $\exists K$ such that $\forall n\ \forall\ y$

$$p_\alpha\left[y_0^{n-1}\right] < K e^{\alpha\cdot f_n(y) - nP(\alpha\cdot f)}$$

whence

$$E_n(\delta) \subseteq \left\{ y \in \Sigma : K e^{\alpha\cdot f_n(y) - nP(\alpha\cdot f)} > e^{n\delta - nh_{p_\alpha}(T)} \right\}.$$

Since $p_\alpha(\alpha\cdot f) = 0$, $P(\alpha\cdot f) = h_{p_\alpha}(T)$. Thus, for n large enough

$$E_n(\delta) \subseteq \{ y \in \Sigma : \alpha\cdot f_n(y) > n\delta/2 \}.$$

We will prove that

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log p_\alpha \{ y \in \Sigma : \alpha\cdot f_n(y) > n\delta/2 \} < 0$$

using large deviations theory for the p_α-distributions of $\alpha\cdot f_n$.

Using the Hölder continuity of f and the Gibbs property of p_α, it is not difficult to prove that the following limit exists for $q \in \mathbb{R}$ (see [Bo]):

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbf{E}_{p_\alpha}\left(e^{q\alpha\cdot f_n}\right) = P(\alpha\cdot f + q\alpha\cdot f) - P(\alpha\cdot f) =: c(q)$$

where $P(\cdot)$ denotes topological pressure and $\mathbf{E}_{p_\alpha}$ denotes expectation with respect to p_α.

By standard large deviations theory (see, e.g., theorem II.6.1 of [El]):

$$\limsup_{n\to\infty} \frac{1}{n} \log p_\alpha \{ y \in \Sigma : \alpha\cdot f_n(y) \geq n\delta/2 \} \leq -\inf_{p\geq\delta/2} I(p)$$

where $I(p)$ is the Legendre–Fenchel transform of $c(q)$ defined by

$$I(p) := \sup_q \{pq - c(q)\}.$$

We outline the (standard) proof that $\inf_{p\geq\delta/2} I(p) > 0$.

By theorem 5.26 in [R1], $c(q)$ is C^2 in $\mathbb{R}$ (see also [G-H]). By aperiodicity, $\alpha\cdot f$ is not cohomologous to a constant and therefore (see [G-H])

$$c'(q) = p_q(\alpha\cdot f) \quad \text{and} \quad c''(q) > 0$$

where p_q is the equilibrium measure of $(1+q)\alpha\cdot f$. It follows that $I(p) = q_0p - c(q_0)$ where q_0 is the maximum point for $q \mapsto qp - c(q)$ satisfying

$$0 = p - c'(q_0) = p - p_{q_0}(\alpha\cdot f)$$

whence

$$I(p) = q_0 p_{q_0}(\alpha\cdot f) - P[(1+q_0)\alpha\cdot f] + P(\alpha\cdot f).$$

By the variational principle,

$$P[(1+q_0)\alpha\cdot f] = h_{p_{q_0}}(T) + p_{q_0}(\alpha\cdot f + q_0\alpha\cdot f).$$

Thus,

$$I(p) = P(\alpha\cdot f) - \left(h_{p_{q_0}}(T) + p_{q_0}(\alpha\cdot f)\right) > 0$$

for $p \neq 0$, because then $p_{q_0} \neq p_\alpha$ (since $p_{q_0}(\alpha\cdot f) = c'(q_0) = p \neq 0 = p_\alpha(\alpha\cdot f)$). Since I is finite and convex (being the the Legendre–Fenchel transform of the convex function c), it is continuous, whence $\inf_{p\geq\delta/2} I(p) > 0$. ∎

3.7 LEMMA: *There exists $M_3 > 0$ such that $\forall\delta > 0$, for p_α-a.e. $x \in \Sigma_0$, $\exists\, N_1 \in \mathbb{N}$ such that $\forall n > N_1\ \exists\, n' < \delta n$, $\underline{\varepsilon} \in \mathcal{W}_{n'}$ satisfying*

$$\|f_{n'}(y) - f_n(x)\| < M_3\ \forall y \in [\underline{\varepsilon}].$$

Proof: Fix some $\delta' > 0$ (to be determined later). By the Ergodic Theorem, for p_α-almost all $x \in \Sigma$, $\|f_n(x)\| = o(n)$, so there exists $N_1 = N_1(x,\delta')$ such that $\forall n > N_1$, $\|f_n(x)\| < \delta' n$. Since f is aperiodic and $p_\alpha(f) = 0$, $\left\{f\circ T^k\right\}_{k=1}^{\infty}$ satisfy a local limit theorem with respect to p_α (see [G-H]). Thus, $\exists k_0 \in \mathbb{N}$ and $c > 0$ such that $\forall\ (\omega_1,\ldots,\omega_d) \in \{+1,-1\}^d$, $k \geq k_0$,

$$p_\alpha[\forall i\ 3B < \omega_i(f_k)_i < 4B] \geq c/k^{d/2}$$

where $(f_k)_i$ denotes the i-th coordinate of that vector. In particular, for every $\omega = (\omega_1, \ldots, \omega_d) \in \{+1, -1\}^d$, there exists $\underline{u}(\omega) \in \mathcal{W}_{k_0}$ such that

$$\forall z \in [\underline{u}(\omega)] \ \forall i \ 2B < \omega_i f_{k_0}(z)_i < 5B. \tag{5}$$

It follows that for every $\underline{c} \in \mathcal{W}_L$ such that $\underline{u}(\omega)\underline{c} \in \mathcal{W}$ and $\forall z \in [\underline{u}(\omega)\underline{c}]$ and $\forall i$

$$B < \omega_i f_{k_0+L}(z)_i < 6B.$$

We use $\underline{u}(\omega)$ to construct $\underline{\varepsilon}$. Fix some $n > N_1$ and $1 \leq i \leq d$. We begin by constructing words $\underline{\varepsilon}^i \in \mathcal{W}_{n'_i}$ such that $|\underline{\varepsilon}^i| < \delta' n$ and such that for $N = |\underline{\varepsilon}^i|$ and all $z \in [\underline{\varepsilon}^i]$

$$|f_N(z)_j| < 7B \quad \text{for } j \neq i, \tag{6}$$

$$|f_N(z)_j - f_n(x)_j| < 7B \quad \text{for} j = i. \tag{7}$$

We construct by induction sign vectors $\omega^k = (\omega_1^k, \ldots, \omega_d^k)$ and words $\underline{c}^k \in \mathcal{W}_L$ such that for all k, $\underline{v}^k := (\underline{u}(\omega^1), \underline{c}^1, \underline{u}(\omega^2), \ldots, \underline{c}^{k-1}, \underline{u}(\omega^k))$ is admissible and such that (6) holds for all $z \in [\underline{v}^k]$ with $N = N_k := |\underline{v}^k|$. Choose $\omega^1 = (\omega_1^1, \ldots, \omega_d^1)$ by $\omega_i^1 = sgn f_n(x)_i$. Assume $\underline{v}^k$ has been chosen and choose some $z \in [\underline{u}^k]$. Define ω^k as follows: if $|f_{N_k}(z)_i - f_n(x)_i| < 7B$, stop and set $\underline{\varepsilon}^i := \underline{v}^k$; else set for $j = i$ $\omega_j^{k+1} := \operatorname{sgn}(f_n(x)_j - f_{N_k}(z)_j)$, and for $j \neq i$, $\omega_j^{k+1} := -\operatorname{sgn} f_{N_k}(z)_j$. Now set $\underline{v}^{k+1} := (\underline{v}^k, \underline{c}^{k+1}, \underline{u}(\omega^{k+1}))$ where $\underline{c}^{k+1} \in \mathcal{W}_L$ is some word which makes $\underline{v}^{k+1}$ admissible. Since at each step we get nearer to $f_n(x)_i$ in steps bounded from below by B, this procedure will stop after less than $\|f_n(x)\|/B \leq \delta' n/B$ steps, so $|\underline{\varepsilon}^i| \leq \delta' n(k_0 + L)/B$. It can be easily verified that $\underline{\varepsilon}^i$ satisfies (6) and (7) for $N = |\underline{\varepsilon}^i|$. Now consider

$$\underline{\varepsilon} := (\underline{\varepsilon}^1, \underline{c}^1, \underline{\varepsilon}^2, \ldots, \underline{c}^{d-1}, \underline{\varepsilon}^d)$$

where $\underline{c}^j \in \mathcal{W}_L$ make the above word admissible. The length of $\underline{\varepsilon}$ is less than $Ld + d(\delta' n(k_0 + L)/B)$ so, by choosing δ' small enough and n large enough (i.e., N_1 large enough), we can make this length smaller than δn as required. Also, it follows from the construction of $\underline{\varepsilon}^i$ that for all $z \in [\underline{\varepsilon}]$,

$$\|f_{|\underline{\varepsilon}|}(z) - f_n(x)\| < 8Bd.$$

The lemma is thus proved for $M_3 := 8Bd$. ∎

3.8 Lemma: *$\exists\, c > 0$, $N_2 \in \mathbb{N}$ such that $\forall n > N_2$*

$$p_\alpha \left\{ y \in \Sigma : \forall z \in \left[y_0^{n-1}\right] \ \|f_n(z)\| < 2B \right\} \geq c/n^{d/2}.$$

Proof: The probability in question is bounded from below by $p_\alpha[\|f_n\| < B]$, and this in turn is bounded below by the local limit theorem. ∎

3.9 LEMMA: *There exists $M_4 > 2B$ such that for almost all $x \in \Sigma_0$*

$$\underline{\lim}_{n\to\infty}\frac{1}{n}\log p_\alpha\left(V_n\left(x, M_4\right)\right) = 0 \quad p_\alpha\text{-a.e.}$$

Proof: Fix some arbitrary $\delta > 0$. Fix

$$N_4 > \max\left\{N_1, (N_2 + L)/(1-\delta), (N_3 + L)/(1-\delta)\right\}$$

where N_1 and N_2 are given by Lemma 3.7 and Lemma 3.8; and N_3 is large enough to ensure that $e^{-\delta n} < c/n^{\delta/2}$ for $n > N_3$.

Assume $n > N_4$. For almost all $x \in \Sigma_0$ and all $t \in \mathbb{R}$ $\exists \underline{\varepsilon} = \underline{\varepsilon}(x) \in \mathcal{W}_{n'}$ such that $n' < \delta n$ and

$$\forall z \in [\underline{\varepsilon}] \quad \|f_{n'}(z) - f_n(x)\| < M_3,$$

$$p_\alpha\left(\left\{y : \forall z \in [y_0^{n-(L+n')-1}] \quad \|f_{n-(L+n')}(z)\| < 2B\right\}\right) > e^{-\delta(n-(L+n'))} > e^{-\delta n}.$$

Set

$$W := \left\{y : \forall z \in [y_0^{n-(L+n')-1}] \quad \|f_{n-(L+n')}(z)\| < 2B\right\}.$$

Consider the set

$$V_n' := \bigcup\left\{[\underline{\varepsilon}; \underline{c}; y_0^{n-(n'+L)-1}] : y \in W \quad \text{and} \quad \underline{c} \in \mathcal{W}_L\right\}.$$

One checks that $V_n' \subseteq V_n(x, M_4)$ where $M_4 = M_3 + 3B$. We estimate the measure of V_n'. Since p_α is a Gibbs measure, there exist a constant $K_1 > 1$ such that $[\underline{a}], [\underline{b}], [\underline{a}, \underline{b}] \neq \emptyset \Rightarrow p_\alpha[\underline{a}, \underline{b}] > K_1^{-1} p_\alpha[\underline{a}] p_\alpha[\underline{b}]$ and there is a constant K_2 such that $\forall \underline{a} \in \mathcal{W}_N$ $p_\alpha[\underline{a}] > K_2^{-N}$. Set

$$W' := \left\{\left[y_0^{n-(n'+L)-1}\right] : y \in W\right\};$$

then

$$p_\alpha(V_n) > K_1^{-1} K_2^{-(n'+L)} \sum_{[\underline{a}] \in W'} p_\alpha[\underline{a}] \geq K_1^{-1} K_2^{-(n'+L)} p_\alpha(W).$$

Thus, $p_\alpha(V_n) > K_1^{-1} K_2^{-L} K_2^{-\delta n} e^{-\delta n}$. Since the above is true for all n such that $n > N_4$,

$$\underline{\lim}_{n\to\infty}\frac{1}{n}\log V_n(x, M_4) \geq -\delta(1 + \log K_2).$$

Since $\delta > 0$ is arbitrary, the lemma is proved. ∎

As mentioned above, Lemma 3.6, Lemma 3.9 imply via (4) that $\exists M > 2B$ such that

$$\liminf_{n\to\infty} \frac{1}{n} \log |U_n(x, M)| \geq h_{p_\alpha}(T) \text{ a.e.},$$

whence (using Lemma 3.4) we have (3). This proves the Main Lemma, and the logarithmic ergodic theorem. ∎

§4. Bounded rational ergodicity

Recall from [A2] that a conservative, ergodic, measure preserving transformation $(X, \mathcal{B}, m, T)$ is called **boundedly rationally ergodic** if there is a set $A \in \mathcal{B}$, $0 < m(A) < \infty$ such that $\exists M > 0$ such that for all $n \geq 1$,

$$(\star) \qquad \left\| \sum_{k=0}^{n-1} 1_A \circ T^k \right\|_{L^\infty(A)} \leq M \int_A \left(\sum_{k=0}^{n-1} 1_A \circ T^k \right) dm.$$

The rate of growth of the sequence

$$a_n = \frac{1}{m(A)^2} \int_A \sum_{k=0}^{n-1} 1_A \circ T^k dm$$

does not depend on the set $A \in \mathcal{B}$, $0 < m(A) < \infty$ satisfying $(\star)$. This sequence is known as the **return sequence** of T and denoted $a_n(T)$ (see [A1]). In this section we prove the following theorem:

THEOREM 4.1: *Let Σ be a topologically mixing subshift of finite type, let μ be the $(1, \tau)$-conformal measure and let $f \in \mathcal{H}_{\mathbb{R}^d}$ be aperiodic; then τ_{ϕ_f} is boundedly rationally ergodic with respect to $m_0 = \mu \times m_{\mathbb{R}^d}$ and*

$$a_n(\tau_{\phi_f}) \asymp \frac{n}{(\log n)^{d/2}}.$$

To prove Theorem 4.1, we show that for $A = \Sigma \times B_M(0)$, M large, $\exists\ 0 < c < C < \infty$ such that

$$\frac{cn}{(\log n)^{d/2}} \leq \int_A S_n(1_A) dm_0, \quad \|S_n(1_A)\|_{L^\infty(A)} \leq \frac{Cn}{(\log n)^{d/2}}.$$

As before, these estimations are first carried out along counting function sequences using the local limit theorem. We begin with the upper estimation.

Let p_0 be the measure of maximal entropy on Σ. It is known that $dp_0 = h_0 d\mu$ where h_0 is bounded away from zero and infinity. Since ϕ_f is invariant under addition of constants to f, we can and do assume that $\mathbf{E}_{p_0}(f) = (0, \ldots, 0)$.

LEMMA 4.2: *$\forall\ M > 0$, $\exists\ A(M) > 0$ such that*

$$p_0[\|f_n(\cdot) - b\| \le M] \le A(M)n^{-d/2} \quad \forall\ b \in \mathbb{R}^d,\ n \in \mathbb{N}.$$

Proof: Set $F := [\|y\| \le M] \subseteq \mathbb{R}^d$ and fix some $a = a(M) \in (0,1)$ such that

$$1_F(y_1,\ldots,y_d) \le 2\prod_{i=1}^{d}\left(\frac{\sin ay_i}{ay_i}\right)^2 = \hat{\gamma}(y)$$

where $\hat{\gamma}$ is the Fourier transform of

$$\gamma(t) := 2\Big(\frac{\pi}{2a^2}\Big)^{d/2} 1_{[\|t\|\le 2a]}(t)\prod_{i=1}^{d}\Big(1 - |\frac{t_i}{2a}|\Big).$$

It follows that

$$\begin{aligned}
p_0[\|f_n - b\| \le M] =&\mathbf{E}_{p_0}(1_F(f_n - b))\\
\le&\mathbf{E}_{p_0}(\hat{\gamma}(f_n - b))\\
=&\frac{1}{(2\pi)^{d/2}}\int_{[\|t\|\le 2a]} e^{ib\cdot t}\mathbf{E}_{p_0}(e^{-it\cdot f_n})\gamma(t)dt\\
\le&\frac{1}{(2\pi)^{d/2}}\int_{[\|t\|\le 2a]} \left|\mathbf{E}_{p_0}(e^{-it\cdot f_n})\right|\gamma(t)dt =: A_n(M).
\end{aligned}$$

Note that the last term, $A_n(M)$ does not depend on b.

As shown in [G-H], there exist $\varepsilon > 0$ and λ: $[\|\cdot\| < \varepsilon] \to \mathbb{C}$ such that $\lambda(t) = 1 - ct^2 + o(\|t\|^2)$ as $t \to 0$; and that for some $0 < \theta < 1$,

$$\mathbf{E}_{p_0}(e^{-it\cdot f_n}) = \begin{cases} \lambda(t)^n + O(\theta^n), & \|t\| \le \varepsilon,\\ O(\theta^n), & \|t\| \in [\varepsilon, 2a]. \end{cases}$$

Making ε smaller if necessary, we assume that for all $\|t\| \le \varepsilon$,

$$|\lambda(t)| \le 1 - \frac{1}{2}ct^2 \le e^{-ct^2}.$$

Using the above to estimate $A_n(M)$, we have that for some $K > 0$,

$$\begin{aligned}
A_n(M) \propto &\int_{[\|t\|\le 2a]} \left|\mathbf{E}_{p_0}(e^{-it\cdot f_n})\right|\gamma(t)dt\\
\le&2\int_{[\|t\|\le\varepsilon]} |\lambda(t)|^n\gamma(t)dt + (4a)^dK\theta^n\\
\le&\frac{2}{n^{d/2}}\int_{[\|\tau\|\le\varepsilon\sqrt{n}]} \left|\lambda(\frac{\tau}{\sqrt{n}})\right|^n \gamma\Big(\frac{\tau}{\sqrt{n}}\Big)d\tau + (4a)^dK\theta^n\\
\le&\frac{2}{n^{d/2}}\int_{[\|\tau\|\le\varepsilon\sqrt{n}]} e^{-c\tau^2}\gamma\Big(\frac{\tau}{\sqrt{n}}\Big)d\tau + (4a)^dK\theta^n\\
\sim&\frac{2\gamma(0)}{n^{d/2}}\int_{\mathbb{R}^d} e^{-c\tau^2}d\tau.
\end{aligned}$$

The lemma follows from this. ∎

Set $B := L\|f\|_\infty + \sum_{k>0} v_k(f)$ where L, as usual, is some number such that all the entries of A^L are positive. Fix some $M > 4B$, set

$$A := \Sigma_0 \times [\|t\| \le M]$$

and $\varphi(x,t) := 1_A$.

LEMMA 4.3: *There is some $C_1 > 0$ such that for almost all (x,t),*

$$|S_{\Lambda_n(x)}(1_A)(x,t)| \le C_1 \frac{\lambda^n}{n^{d/2}}.$$

Proof: Let s be the number of states of Σ, set $L_0 := L + s + 2$, and define

$$u_n(x) := \inf\{u \ge n + L_0 \colon x_{u-1} < P_{\max}(x_u)\},$$
$$\ell_n(x) := \sup\{\ell \le n + L_0 \colon x_{\ell-1} < P_{\max}(x_\ell)\}.$$

For p_0-almost all $x \in \Sigma$ these are finite. For such x we have the following representation:

$$x = (x_0^{\ell_n-1}, P_{\max}^{u_n-\ell_n-1}(x_{u_n-1}), \dots, P_{\max}(x_{u_n-1}), x_{u_n-1}, x_{u_n}^\infty).$$

Define $k_n(x) \in \mathbb{N}$ by the equation

$$\tau^{k_n(x)}(x) = (P_{\max}^{u_n-1}(x_{u_n-1}), \dots, P_{\max}(x_{u_n-1}), x_{u_n-1}, x_{u_n}^\infty).$$

If $b > x_{u_n-1}$ is the minimal state such that bx_{u_n} is admissible, then

$$\tau^{k_n(x)+1}(x) = (P_{\min}^{u_n-1}(b), \dots, P_{\min}(b), b, x_{u_n}^\infty).$$

We estimate $S_{\Lambda_n} 1_A$ by breaking it into two members:

$$\begin{aligned} S_{\Lambda_n(x)}(1_A)(x,t) &= S_{k_n(x)}(1_A)(x,t) + S_{\Lambda_n(x)-k_n(x)}(1_A)(\tau_{\phi_f}^{k_n(x)}(x,t)) \\ &\le S_{k_n(x)}(1_A)(x,t) + S_{\Lambda_n(\tau^{k_n(x)+1}x)}(1_A)(\tau_{\phi_f}^{k_n(x)+1}(x,t)) + 1 \\ &=: I + II + 1. \end{aligned}$$

The inequality follows from the minimality of $\Lambda_n(x)$ as

$$\{(\tau^j x)_0^{n-1} \colon 0 \le j \le k_n(x) + 1 + \Lambda_n(\tau^{k_n(x)+1}x)\} = \mathcal{W}_n.$$

To estimate I, we begin by noting that the map $j \mapsto (\tau^j x)_0^{\ell_n-1}$ is 1-1 for $0 \le j \le k_n - 1$. To see this, note that for such j, $x \prec \tau^j x \prec \tau^{k_n} x$ in the reverse lexicographic order whence

$$x_{\ell_n}^\infty = (\tau^{k_n} x)_{\ell_n}^\infty = (P_{\max}^{u_n-\ell_n-1}(x_{u_n-1}), \dots, P_{\max}(x_{u_n-1}), x_{u_n-1}, x_{u_n}^\infty).$$

Thus the difference between the $\tau^j x$'s must be reflected in the first ℓ_n coordinates. Since $\ell_n \le n + L_0$,

$$\begin{aligned} S_{k_n}(1_A)(x,t) =& |\{0 \le j \le k_n - 1\colon \|t + (\phi_f)_j(x)\| \le M\}| \\ =& |\{0 \le j \le k_n - 1\colon \|f_{n+L_0}(\tau^j x) - f_{n+L_0}(x) - t\| \le M\} \\ \le& |\{\underline{\varepsilon} \in \mathcal{W}_{n+L_0}\colon \forall y \in [\underline{\varepsilon}] \|f_{n+L_0}(y) - f_{n+L_0}(x) - t\| \le M + B\}|. \end{aligned}$$

Since p_0, being the measure of maximal entropy, is the Gibbs measure for the zero potential, there is some constant K such that for every $\underline{a} \in \mathcal{W}_n$, $K^{-1}\lambda^n < p_0[\underline{a}] \le K\lambda^n$. In particular, cylinders of the same length are of comparable sizes whence

$$|S_{k_n(x)}(1_A)(x,t)| \le K\lambda^{n+L_0} p_0[\|f_{n+L_0}(\cdot) - f_{n+L_0}(x) - t\| \le M + B].$$

Lemma 4.2 now implies that $I = O(\lambda^n n^{-d/2})$ uniformly on A.

We now estimate II. Set $(x',t') := \tau_{\phi_f}^{k_n(x)+1}(x,t)$. We have to estimate $S_{\Lambda_n(x')}(1_A)(x',t')$. We do this by showing that

$$\Lambda_n(x') \le k_n(x'), \tag{8}$$

thus reducing the problem to that which was discussed in the previous step.

There exists $n + L + 1 < u_n' < u_n(x')$ such that $P_{\min}(x_{u_n'}) < P_{\max}(x_{u_n'})$, since otherwise, there would be an admissible word $[a_1, \ldots, a_r]$ for some $r \le s+1$ with $a_1 = a_r$ and $P_{\max}(a_j) = P_{\min}(a_j)$ $(1 \le j \le r)$. This contradicts the aperiodicity of A.

Now consider

$$\begin{aligned} x' =& (P_{\min}^{u_n'}(x'_{u_n'}), \ldots, P_{\min}(x'_{u_n'}), (x')_{u_n'}^{u_n-1}, x_{u_n}^{\infty}), \\ y :=& (P_{\max}^{u_n'}(x'_{u_n'}), \ldots, P_{\max}(x'_{u_n'}), (x')_{u_n'}^{u_n-1}, x_{u_n}^{\infty}), \\ \tau^{k_n(x')}x' =& (P_{\max}^{u_n-1}(x_{u_n}), \ldots, P_{\max}(x_{u_n-1}), x_{u_n-1}^{\infty}). \end{aligned}$$

Since $u_n' > n + L + 1$, for every $\underline{\varepsilon} \in \mathcal{W}_n$ there is some w_0^{L-1} such that $x(\underline{\varepsilon}) := (\underline{\varepsilon}, w_0^{L-1}, y_{n+L}^{\infty})$ is admissible and since $u_n' < u_n$, $x' \prec x(\underline{\varepsilon}) \prec \tau^{k_n(x')+1}x'$. This shows that $\mathcal{W}_n$ is spanned by $(\tau^j(x'))_0^{n-1}$ for $j = 1, \ldots, k_n(x') - 1$, whence (8). ∎

This completes the upper estimation, and we now address the lower estimation.

LEMMA 4.4: *There exists n_0 such that for all x, $\exists 0 \le i_1 < i_2 \le \Lambda_{n+n_0}(x) - 1$ such that for every $i_1 \le j \le i_2$, $(\tau^j x)_{n+L}^{\infty}$ is the same, and*

$$\{(\tau^j x)_0^{n-1}\colon i_1 \le j \le i_2\} = \mathcal{W}_n.$$

Proof: Let L be large enough such that $A^L > 0$ and set $n_0 := L + n_1$ where $|\mathcal{W}_{n_1}| \geq 3$. Choose three different $\underline{a}_j \in \mathcal{W}_{n_1}$. There are $0 \leq k_1, k_2, k_3 \leq \Lambda_{n+n_0} - 1$ such that $z^{(j)} := T^{n+L}(\tau^{k_j} x) \in [\underline{a}_j]$. In particular, $z^{(j)}$ are different. Without loss of generality, $z^{(1)} \prec z^{(2)} \prec z^{(3)}$. For every $\underline{\varepsilon} \in \mathcal{W}_n$, construct an admissible word of the form $x(\underline{\varepsilon}) = (\underline{\varepsilon}, w_0^{L-1}, z^{(2)})$. Let x^- and x^+ be the minimal and maximal points among the $x(\underline{\varepsilon})$. Clearly, $\tau^{k_1}x \prec x^- \prec x^+ \prec \tau^{k_3}x$, whence $\exists 0 \leq i_1 < i_2 \leq \Lambda_{n+n_0}(x) - 1$ such that $x^- = \tau^{i_1}x$ and $x^+ = \tau^{i_2}x$. It follows that $\mathcal{W}_n$ is spanned by $\tau^j x$ for $j = i_1, \ldots, i_2$. Since $(x^-)_{n+L}^{\infty} = (x^+)_{n+L}^{\infty} = z^{(2)}$, $(\tau^j x)_{n+L}^{\infty}$ is constant for $j = i_1, \ldots, i_2$. ∎

LEMMA 4.5: *There exists $C_2 > 0$ such that for n large enough,*

$$\int_A S_{\Lambda_n(x)}(1_A)(x,t)dm(x,t) \geq C_2 \frac{\lambda^n}{n^{d/2}}.$$

Proof: It is enough to prove that for some C_3 and all $\|t\| \leq B$,

$$\int_\Sigma S_{\Lambda_n(x)}(\varphi)(x,t)dp_0(x) \geq C_3 \frac{\lambda^n}{n^{d/2}}$$

(the lemma will then follow by integration dt over $[\|t\| \leq B]$).

By Lemma 4.4 for some n_0, for every $x \in \Sigma$ and $n \in \mathbb{N}$ there are $0 \leq i_1 < i_2 \leq \Lambda_{n+n_0}(x) - 1$ such that $(\tau^j x)_{n+L}^{\infty}$ is constant for $j = i_1, \ldots, i_2$ and such that $\mathcal{W}_n = \{(\tau^j x)_0^{n-1} \colon j = i_1, \ldots, i_2\}$. It follows that

$$\begin{aligned}
&S_{\Lambda_{n+n_0}(x)}(1_A)(x,t) \\
&\qquad \geq \sum_{j=i_1}^{i_2} (1_A \circ \tau_{\phi_f}^j)(x,t) \\
&\qquad = |\{i_1 \leq j \leq i_2 \colon \|f_{n+L}(\tau^j x) - f_{n+L}(x) - t\| < M\}| \\
&\qquad = |\{(\tau^j x)_0^{n+L-1} \colon j \in [i_1, i_2], \|f_{n+L}(\tau^j x) - f_{n+L}(x) - t\| < M\}| \\
&\qquad \geq |\{\underline{\varepsilon} \in \mathcal{W}_n \colon \exists y \in [\underline{\varepsilon}], \|f_n(y) - f_n(x)\| \leq M - 4B\}| \\
&\qquad \geq K^{-1}\lambda^n p_0[\|F_n(x, \cdot)\| \leq M - 4B]
\end{aligned}$$

where $F \colon \Sigma \times \Sigma \to \mathbb{R}^d$ is the symmetrization of f (as in the proof of Corollary 2.7) given by $F(x,y) = f(x) - f(y)$, and $F_n(x,y) := \sum_{i=0}^{n-1} F(T^i x, T^i x)$. Integrating with respect to $dp_0(x)$ we have, for all $\|t\| < B$,

$$\int_\Sigma S_{\Lambda_{n+n_0}(x)}(1_A)(x,t) \geq K^{-1}\lambda^n (p_0 \times p_0)[\|F_n\| \leq M - 4B].$$

As in the proof of Corollary 2.7, $(\Sigma \times \Sigma, T \times T)$ is a subshift of finite type, $F\colon \Sigma \times \Sigma \to \mathbb{R}^d$ is Hölder continuous, and F is aperiodic. Therefore, F_n satisfy a local limit theorem ([G-H]):

$$(p_0 \times p_0)[[\|F_n\| \leq M - 4B|] \propto 1/n^{d/2},$$

whence the lemma. ∎

Proof of Theorem 4.1: We prove that for $M > 4B$, $A := \Sigma \times \{t\colon \|t\| < M\}$ satisfies that

$$\|1_A S_N 1_A\|_\infty = O(\|1_A S_N 1_A\|_{L^1(M_0)}) \quad (N \to \infty).$$

By the counting proposition, uniformly in x, $\Lambda_n(x) \asymp |\mathcal{W}_n| \asymp \lambda^n$, where $\lambda = e^{h_{top}(\Sigma)}$. Therefore, there exists $c \in \mathbb{N}$ such that for all $x \in \Sigma_0$ and n, $\lambda^{n-c+1} \leq \Lambda_n(x) \leq \lambda^{n+c}$. Fix $N > \lambda^{1+c}$ and choose the n such that $\lambda^n \leq N < \lambda^{n+1}$. The last estimations imply that for every $x \in \Sigma_0$,

$$\Lambda_{n-c}(x) \leq N < \Lambda_{n+c}(x)$$

whence, by the preceding lemmas, for almost all $(x,t) \in A$ and N large enough,

$$S_N(1_A)(x,t) \leq \frac{C_1 \lambda^{n+c}}{(n+c)^{d/2}},$$
$$\int_A S_N(1_A)dm \geq \frac{C_2 \lambda^{n-c}}{(n-c)^{d/2}}.$$

The theorem follows from this. ∎

References

[A1] J. Aaronson, *An Introduction to Infinite Ergodic Theory*, Mathematical Surveys and Monographs **50**, American Mathematical Society, Providence, R.I, 1997.

[A2] J. Aaronson, *Rational ergodicity, bounded rational ergodicity and some continuous measures on the circle*, Israel Journal of Mathematics **33** (1979), 181–197.

[A-D1] J. Aaronson and M. Denker, *Local limit theorems for partial sums of stationary sequences generated by Gibbs–Markov maps*, Stochastic Dynamics **1** (2001), no. 2 193–237.

[A-D2] J. Aaronson and M. Denker, *Group extensions of Gibbs–Markov maps*, to appear in Probability and Related Fields.

[A-K] J. Aaronson and M. Keane, *The visits to zero of some deterministic random walks,* Proceedings of the London Mathematical Society **44** (1982), 535–553.

[A-W] J. Aaronson and B. Weiss, *On the asymptotics of a 1-parameter family of infinite measure preserving transformations,* Boletim da Sociedade Brasileira de Matemática (N.S.) **29** (1998), 181–193.

[B-L] M. Babillot and F. Ledrappier, *Geodesic paths and horocycle flow on abelian covers,* in *Lie Groups and Ergodic Theory (Mumbai, 1996)*, Tata Institute of Fundamental Research Studies in Mathematics **14**, Tata Institute of Fundamental Research, Bombay, 1998, pp. 1–32.

[B-M] R. Bowen and B. Marcus, *Unique ergodicity for horocycle foliations,* Israel Journal of Mathematics **26** (1977), 43–67.

[Bo] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms,* Lecture Notes in Mathematics **470**, Springer-Verlag, Berlin–New York, 1975.

[C-K] J.-P. Conze and M. Keane, *Ergodicité d'un flot cylindrique,* Séminaire de Probabilités **I**, Univeristé Rennes, Rennes, 1976.

[El] R. Ellis, *Entropy, large deviations and statistical mechanics,* Grundlehren der Mathematischen Wissenschaften **271**, Springer-Verlag, New York–Berlin, 1985.

[F-M] J. Feldman and C. C. Moore, *Ergodic equivalence relations. cohomology, and von Neumann algebras. I,* Transactions of the American Mathematical Society **234** (1977), 289–324.

[Fu] H. Furstenberg, *Strict ergodicity and transformation of the torus,* American Journal of Mathematics **83** (1961), 573–601.

[G] Y. Guivarc'h, *Propriétés ergodiques, en mesure infinie, de certains systèmes dynamiques fibrés,* Ergodic Theory and Dynamical Systems **9** (1989), 433–453.

[G-H] Y. Guivarc'h and J. Hardy, *Théorèmes limites pour une class de chaînes de Markov et applications aux difféomorphisms d'Anosov,* Annales de l'Institut Henri Poincaré, Probabilites et Statistiques **24** (1988), 73–98.

[H-I-K] A. Hajian, Y. Ito and S. Kakutani, *Invariant measures and orbits of dissipative transformations,* Advances in Mathematics **9** (1972), 52–65.

[Her] M. Herman, *Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations,* Publications Mathématiques de l'Institut des Hautes Études Scientifiques **49** (1979), 5–233.

[Ka] V. Kaimanovich, *Ergodic properties of the horocycle flow and classification of Fuchsian groups,* Journal of Dynamical and Control Systems **6** (2000), 21–56.

[Ke] M. Keane, *Interval exchange transformations,* Mathematische Zeitschrift **141** (1975), 25–31.

[Kh] A. Ya. Khinchin, *Continued Fractions*, University of Chicago Press, Chicago and London, 1964.

[Kow] Z. S. Kowalski, *Quasi markovian transformations*, Ergodic Theory and Dynamical Systems **17** (1997), 885–898.

[Kr-Li] C. Kraaikamp and P. Liardet, *Good approximations and continued fractions*, Proceedings of the American Mathematical Society **112** (1991), 303–309.

[Ku-Ni] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974.

[Mah] D. Maharam, *Incompressible transformations*, Forum Mathematicum **56** (1964), 35–50.

[dM-vS] W. de Melo and S. van Strien, *One-dimensional Dynamics*, Ergebnisse Math. u. Grenzgeb. (3) **25**, Springer-Verlag, Berlin, 1993.

[N1] H. Nakada, *Piecewise linear homeomorphisms of type* III *and the ergodicity of cylinder flows*, Kodai Mathematical Seminar Reports **7** (1982), 29–40.

[N2] H. Nakada, *On a classification of PL-homeomorphisms of a circle*, in *Probability Theory and Mathematical Statistics (Tblisi 1982)*, Lecture Notes in Mathematics, **1021**, Springer, Berlin, Heidelberg, New York, 1983, pp. 474–480.

[Pa] W. Parry, *Compact abelian group extensions of discrete dynamical systems*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **13** (1969), 95–113.

[P-S] K. Petersen and K. Schmidt, *Symmetric Gibbs measures*, Transactions of the American Mathematical Society **349** (1997), 2775–2811.

[Po] M. Pollicott, *$\mathbb{Z}^d$-covers of horosphere foliations*, Discrete Continuous Dynamical Systems **6** (2000), 147–154.

[R1] D. Ruelle, *Thermodynamic Formalism (The Mathematical Structures of Classical Equilibrium Statistical Mechanics)*, in *Encyclopedia of Mathematics and its Applications* **5**, Addison-Wesley, Reading, Mass., 1978.

[R2] D. Ruelle, *Statistical mechanics of a one-dimensional lattice gas*, Communications in Mathematical Physics **9** (1968), 267–278.

[Sc1] K. Schmidt, *Cocycles of Ergodic Transformation Groups*, Lecture Notes in Mathematics **1**, MacMillan Co. of India, 1977.

[Sc2] K. Schmidt, *Unique ergodicity and related problems*, in *Ergodic Theory, Proceedings Oberwolfach 1978*, Lecture Notes in Mathematics **729**, Springer-Verlag, Berlin-New York, 1979, pp. 188–198.

[St] M. Stewart, *Irregularities of uniform distribution*, Acta Mathematica Academiae Scientiarum Hungaricae **37** (1981), 185–221.

[Ve] W. A. Veech, *Topological dynamics*, Bulletin of the American Mathematical Society **83** (1977), 775–830.

[V1] A. M. Vershik, *A theorem on Markov periodic approximation in ergodic theory* (Russian), in *Boundary Value Problems of Mathematical Physics and Related Questions in the Theory of Functions*, 14, Zapiski Nauchnykh Seminarov Leningrad. Otdel., Matematicheskii Institut Steklov (LOMI) **115** (1982), 72–82.

[V2] A. M. Vershik, *A theorem on the periodic Markov approximation in ergodic theory,* Journal of Soviet Mathematics **28** (1985), 667–674.

[V3] A. M. Vershik, *A new model of the ergodic transformations,* in *Dynamical Systems and Ergodic Theory (Warsaw, 1986)*, Banach Center Publications, **23**, PWN, Warsaw, 1989, pp. 381–384.

[W-Z] R. L. Wheeden and A. Zygmund, *Measure and Integral. An Introduction to Real Analysis,* Pure and Applied Mathematics, **43**, Marcel Dekker, Inc., New York–Basel, 1977.