

# THE VISITS TO ZERO OF A RANDOM WALK DRIVEN BY AN IRRATIONAL ROTATION

A. AVILA, D. DOLGOPYAT, E. DURIEV, AND O. SARIG

ABSTRACT. We give detailed analysis of the returns to zero of the “deterministic random walk”  $S_n(x) = \sum_{k=0}^{n-1} f(x+k\alpha)$  where  $\alpha$  is a quadratic irrational,  $f(x) = 1_{[\frac{1}{2}, 1)}(\{x\}) - 1_{[0, \frac{1}{2})}(\{x\})$ , and  $x$  is sampled uniformly in  $[0, 1]$ .

The method is to find the asymptotic behavior of the ergodic sums of  $L^1$  functions for linear flows on the infinite staircase surface.

Our methods also provide a new proof of J. Beck’s central limit theorem for  $S_n(0)$  where  $n \in \{1, \dots, N\}$  is uniform and  $N \rightarrow \infty$ , and they allow us to determine the generic points for certain infinite measure preserving skew products (“cylinder maps”).

## 1. INTRODUCTION AND OVERVIEW

The simple random walk (SRW) can be generated from a dynamical system as follows. Pick  $x$  in  $\mathbb{T} := \mathbb{R}/\mathbb{Z}$  uniformly, and iterate the *angle doubling map*

$$\tau : \mathbb{T} \rightarrow \mathbb{T}, \quad \tau(x) = 2x \pmod{1}.$$

Place a “walker” at  $0 \in \mathbb{Z}$ . At time step  $k$  ( $k \geq 1$ ), ask the walker to make one step to the left if  $\tau^k(x) \in [0, \frac{1}{2})$ , and one step to the right if  $\tau^k(x) \in [\frac{1}{2}, 1)$ . This procedure generates the the simple random walk, because the  $k$ -th step is  $(-1)^{x_k+1}$  where  $0.x_1x_2x_3 \dots$  is the binary expansion of  $x$ , and if  $x \in \mathbb{T}$  is chosen uniformly, then  $x_i$  are i.i.d. equal to  $+1$  or  $-1$  with probability  $\frac{1}{2}$ .

The angle doubling map is the standard example of a “chaotic” map. It is natural to ask what happens if we replace it by an (ergodic) “non-chaotic” map, such as the *irrational rotation*

$$R_\alpha : \mathbb{T} \rightarrow \mathbb{T}, \quad R_\alpha(x) = x + \alpha \pmod{1} \quad (\alpha \in \mathbb{R} \setminus \mathbb{Q} \text{ fixed}).$$

Whereas  $\tau$  is mixing, has positive entropy, and has countable Lebesgue spectrum, properties associated with “chaos”,  $R_\alpha$  is non-mixing, has zero entropy, and has discrete spectrum, properties associated with “determinism” (see [Pet] page 245).

If we replace  $\tau$  by  $R_\alpha$  we obtain the a stochastic process called the *deterministic random walk* [AK]. To define it formally, let  $\mathcal{B}$  denote the Borel  $\sigma$ -algebra of  $\mathbb{T}$ , let  $m_{\mathbb{T}}$  be the normalized Lebesgue measure on  $\mathbb{T}$  thought of as the unit interval mod 1, and define  $f : \mathbb{T} \rightarrow \mathbb{Z}$  by

$$f(x) = \begin{cases} -1 & x \in [0, \frac{1}{2}) \\ +1 & x \in [\frac{1}{2}, 1). \end{cases}$$

---

*Date:* July 27, 2013.

*Key words and phrases.* thermodynamic formalism, countable Markov shifts, spectral gap.

A.A. was partially supported by the ERC Starting Grant “Quasiperiodic” and by the Balzan project of Jacob Palis. D.D. was partially supported by the NSF grant DMS 1101635. O.S. was partially supported by ERC Starting Grant “ErgodicNonCompact”.

The *deterministic random walk (DRW)* with angle  $\alpha \in \mathbb{R} \setminus \mathbb{Q}$  is the sequence  $S_0 = 0$ ,  $S_n = \sum_{k=1}^n f \circ R_\alpha^k$  ( $n \geq 1$ ) on the probability space  $(\mathbb{T}, \mathcal{B}, m_{\mathbb{T}})$ . The following table compares it with the simple random walk:

	Simple RW	Deterministic RW
<b>Recurrence</b>	A.e. orbit returns to zero. The set of exceptions has positive Hausdorff dim. [BS]	A.e. orbit returns to zero [At] The set of exceptions is finite [Ra].
<b>Trace</b> $\{S_n(x) : n \geq 0\}$	$\mathbb{Z}$ for a.e. $x$ .	$\mathbb{Z}$ for a.e. $x$ [C],[CK],[Sch], but not for all $x$ [BR],[Pe].
<b>Drift</b> $\lim_{n \rightarrow \infty} S_n(x)/n$	A.e. orbit has zero drift. The set of exceptions has positive Hausdorff dim. [BS]	All orbits have zero drift by Weyl's Theorem ([KN, chapter 1])
<b>Central Limit Theorem</b>	Yes	No, by the Denjoy–Koksma Ineq. [Her, page 73]. Other choices of $f$ may have CLT [BD],[V]. See also [Hu],[B1]

In this paper we contribute to the study of the visits to zero of the deterministic random walk:  $N_n(x) = N_n(x; \alpha) := (\#\{0 \leq k \leq n-1 : S_k(x) = 0\})$ .

For the simple random walk, if the number of visits to zero up to time  $n$  is  $\widehat{N}_n$ , then  $\mathbb{E}(\widehat{N}_n) \sim \sqrt{2n/\pi}$  (de Moivre–Laplace Theorem), and  $\frac{1}{\sqrt{n}} \widehat{N}_n \xrightarrow[n \rightarrow \infty]{\text{dist}} \Theta(1)$ , where  $\Theta(t)$  is Brownian local time [Bor]. The deterministic random walk behaves differently. Aaronson and Keane showed in [AK] that if  $\alpha$  is a quadratic irrational, then there are constants  $c_1, c_2 > 0$  s.t.  $c_1(\frac{n}{\sqrt{\ln n}}) \leq \mathbb{E}(N_n) \leq c_2(\frac{n}{\sqrt{\ln n}})$ .

We show, among other things, that if  $\alpha$  is a quadratic irrational, then there is a positive constant  $c(\alpha)$  with the following properties:

- (1)  $\mathbb{E}(N_n) \sim c(\alpha)(\frac{n}{\sqrt{\ln n}}) =: a_n(\alpha)$  as  $n \rightarrow \infty$ .
- (2)  $\frac{1}{a_n(\alpha)} N_n \xrightarrow[n \rightarrow \infty]{\text{dist}} \sqrt{2} \exp[-\frac{1}{2}\chi^2]$ , where  $\chi$  has the standard normal distribution.
- (3)  $\frac{1}{\ln \ln n} \sum_{k=0}^{n-1} \frac{1}{n \ln n} \left( \frac{1}{a_n(\alpha)} N_n \right) \xrightarrow[n \rightarrow \infty]{} 1$  almost surely.
- (4)  $c(\alpha) = \sqrt{\frac{|\ln \lambda|}{4\pi\sigma^2}}$  where  $\lambda, \sigma^2 \in \mathbb{Q}[\alpha]$  can be calculated explicitly. For example,  $c(\sqrt{2}) = \sqrt{\frac{\sqrt{2}}{3\pi} \ln(17 + 12\sqrt{2})}$ .

These results should be contrasted with Kesten's work [Kes] which says that if  $\alpha$  is also randomized (i.e.  $(x, \alpha)$  chosen uniformly in  $[0, 1]^2$ ), then the right scaling for  $N_n(x; \alpha)$  is  $n/\ln n$ .

Our main tool is the *cylinder map*  $T_\alpha(x, \xi) = (x + \alpha \bmod 1, \xi + f(x))$  on  $\mathbb{T} \times \mathbb{Z}$ , together with the (infinite) invariant measure  $m_0 := m_{\mathbb{T}} \times m_{\mathbb{Z}}$  ( $m_{\mathbb{T}}$ =normalized Lebesgue measure on  $\mathbb{T}$ ,  $m_{\mathbb{Z}}$ =counting measure on  $\mathbb{Z}$ ). To see the connection to the DRW, write  $T_\alpha^n := T_\alpha \circ \dots \circ T_\alpha$  ( $n$  times) and observe by direct calculation that

- $S_n(x)$  is the second coordinate of  $T_\alpha^n(x, 0)$ , and
- $N_n(x) = \sum_{k=0}^{n-1} 1_{\mathbb{T} \times \{0\}}(T_\alpha^k(x, 0))$ , where  $1_E$  is the indicator function of  $E$ .

We will analyze the asymptotic behavior of the Birkhoff sums  $\sum_{k=0}^{n-1} G \circ T_\alpha^k$  for general non-negative functions  $G \in L^1(\mathbb{T} \times \mathbb{Z})$ . The results for  $N_n$  follow from the special case  $G = 1_{\mathbb{T} \times \{0\}}$ .



The function  $r : \mathfrak{S} \rightarrow (0, \infty)$  and the map  $T : \mathfrak{S} \rightarrow \mathfrak{S}$ ,  $T(p) = \varphi_\theta^{r(p)}(p)$ , are called the *roof function* and *Poincaré map* of  $\mathfrak{S}$ .

**Lemma 1.1.** *For  $\theta \neq \pm \frac{\pi}{2} + 2\pi k$ , the union of the horizontal sides of the horizontal rectangles in figure 1 is a Poincaré section for  $\varphi_\theta$  with constant roof function. The Poincaré map is isomorphic to the cylinder map  $T_\alpha$  where  $\alpha = \frac{1}{2} \tan \theta + \frac{1}{2}$ .*

The isomorphism is very simple: Divide  $\text{St}$  into horizontal rectangles, call one of them “rectangle zero” and tag the remaining rectangles by  $\xi \in \mathbb{Z}$  in such a way that the rectangle directly above rectangle  $\xi$  is rectangle  $\xi + 1$ . The point  $(x, \xi) \in \mathbb{T} \times \mathbb{Z}$  corresponds to the point  $\omega(x)$  on the top horizontal side of rectangle  $\xi$ , and located  $2x$  units of distance away from the left end.

Since  $T_\alpha$  is a Poincaré map for  $\varphi_\theta$  with constant roof function, there is a standard way to reduce the study of the Birkhoff sums of  $T_\alpha$  to the analysis of the Birkhoff integrals of  $\varphi_\theta$ . This is what we will do.

The gain is that  $\text{St}$  has many symmetries, and for special directions  $\theta$ , it is possible to find a “nice” automorphisms  $\psi : \text{St} \rightarrow \text{St}$  s.t.

$$\psi \circ \varphi_\theta^t = \varphi_\theta^{\lambda t} \circ \psi \quad (*)$$

for some  $0 < \lambda < 1$ . This is what happens for the  $\theta$  whose corresponding  $\alpha$  is a quadratic irrational. (\*) is the key to the asymptotic behavior of the Birkhoff sums of  $\varphi_\theta$  and  $T_\alpha$ , and therefore also to the asymptotic behavior of  $N_n$ .

## 2. THE INFINITE STAIRCASE AND ITS AUTOMORPHISMS

**$\mathbb{Z}$ -cover.**  $\text{St}^*$  is a regular  $\mathbb{Z}$ -cover of a finite area surface  $\text{St}_0^*$  (figure 1). Let

$$\pi : \text{St}^* \rightarrow \text{St}_0^*$$

be the covering map.  $\text{St}_0^*$  is a twice punctured torus (Figure 1(c)). Let  $\text{St}_0$  denote its completion. It is a torus, and  $\pi : \text{St}^* \rightarrow \text{St}_0^*$  extends continuously to a map  $\pi : \text{St} \rightarrow \text{St}_0$ . The extension is two-to-one on the singularities of  $\text{St}$  and infinite-to-one elsewhere.

The group of deck transformations of the covering is generated by an obvious translation. We denote it by

$$D : \text{St} \rightarrow \text{St}.$$

$D^2$  fixes the singularities.

**$\mathbb{Z}$ -coordinate.** Choose a bounded connected set  $R \subset \text{St}^*$  s.t.  $\text{St}^* = \bigsqcup_{k \in \mathbb{Z}} D^k(R)$ , for example one of the horizontal rectangles in figure 1 with the vertices and bottom horizontal side removed. The  $\mathbb{Z}$ -coordinate of  $p \in \text{St}^*$  (relative to  $R$ ) is

$$\xi(p) := \text{unique } k \text{ s.t. } p \in D^k(R).$$

Notice that  $\xi \circ D = \xi + 1$ . This definition depends on the choice of  $R$ . We will refer to this as “choosing a  $\mathbb{Z}$ -coordinate.”

**Homogeneous automorphisms.**  $\text{St}$  has an obvious atlas of charts whose change of coordinates transformations are euclidean translations. This allows us to identify the tangent spaces of  $\text{St}$  at different points with  $\mathbb{R}^2$  (and therefore with each other) consistently. We will use the convention that direction “up” in figure 1 is  $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \in \mathbb{R}^2$ .

Once we have identified the tangent spaces at different points with  $\mathbb{R}^2$ , we can view the differential  $d\psi_p : T_p(\text{St}) \rightarrow T_{\psi(p)}(\text{St})$  of a smooth map  $\psi : \text{St} \rightarrow \text{St}$

( $p \in \text{St}^*$ ) as a linear map  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ . The matrix representing this map is called the *derivative* at  $p$ .

A map  $\psi : \text{St} \rightarrow \text{St}$  is an *automorphism*, if  $\psi : \text{St} \rightarrow \text{St}$  is a homeomorphism;  $\psi(\text{St}^*) = \text{St}^*$ ;  $\psi : \text{St}^* \rightarrow \text{St}^*$  is differentiable;  $\psi : \text{St}^* \rightarrow \text{St}^*$  is orientation preserving; and  $\psi$  has constant derivative.

An automorphism  $\psi : \text{St} \rightarrow \text{St}$  is called *homogeneous*, if it commutes with  $D$  and preserves the  $D$ -orbits of the singularities of  $\text{St}$ . Notice that in this case,  $\psi^2$  fixes the singularities.

Every homogeneous automorphism  $\psi : \text{St} \rightarrow \text{St}$  is the lift of a toral automorphism  $\psi_0 : \text{St}_0 \rightarrow \text{St}_0$ : Take

$$\psi_0(p) = \pi[\psi(\tilde{p})] \text{ for some (any) } \tilde{p} \in \pi^{-1}(p).$$

The definition is proper, because  $\pi^{-1}(p)$  is an orbit of  $D$ ,  $\psi \circ D = D \circ \psi$  and  $\pi \circ D = \pi$ . It is easy to see that  $\psi_0$  is invertible and has constant derivative.

We call  $\psi_0$  the *projection of  $\psi$* . The assumption that  $\psi$  preserves the  $D$ -orbits of singularities means that  $\psi_0$  fixes the punctures of  $\text{St}_0^*$ .

**Proposition 2.1.** *All homogeneous automorphisms are area preserving.*

*Proof.* Let  $J, J_0$  denote the Jacobian functions of a homogeneous automorphism  $\psi$  and its projection  $\psi_0$ . Since  $\pi \circ \psi = \psi_0 \circ \pi$  and  $\pi$  is a local isometry,  $J = J_0 \circ \pi$ . The Jacobian of  $\psi$  is constant ( $\psi$  has constant derivative), therefore the Jacobian of  $\psi_0$  is constant. Since  $\psi_0$  is a self-bijection of a surface of finite area, this constant equals one. So  $J = J_0 \circ \pi \equiv 1$ , and  $\psi$  is area preserving.  $\square$

**Frobenius functions and drifts.** Fix some  $\mathbb{Z}$ -coordinate  $\xi : \text{St} \rightarrow \mathbb{Z}$ . The *Frobenius function* of a homogeneous automorphism  $\psi : \text{St} \rightarrow \text{St}$  is

$$F_\psi : \text{St}_0^* \rightarrow \mathbb{Z}, \quad F_\psi(p) = \xi[\psi(\tilde{p})] - \xi[\tilde{p}] \text{ for some (any) } \tilde{p} \in \pi^{-1}(p).$$

The definition is proper because  $\pi^{-1}(p) = \{D^n(p) : n \in \mathbb{Z}\}$ ,  $\varphi \circ D = D \circ \varphi$ , and  $\xi \circ D = \xi + 1$ .  $F_\psi$  depends on the choice of the  $\mathbb{Z}$ -coordinate. If we change the  $\mathbb{Z}$ -coordinate,  $F_\psi$  changes by a coboundary of  $\psi_0$ , see below.

The *average drift* (or just *drift*) of a homogenous automorphism  $\psi : \text{St} \rightarrow \text{St}$  is

$$\delta(\psi) := \frac{1}{\text{area}(\text{St}_0)} \int_{\text{St}_0} F_\psi(p) dp, \quad (dp = \text{area measure}).$$

We will see later that  $\delta(\psi)$  is the drift of a certain random walk associated to  $\psi$ .

**Lemma 2.2.** *The average drift is independent of the choice of the  $\mathbb{Z}$ -coordinate, and  $\delta(\psi \circ \phi) = \delta(\psi) + \delta(\phi)$  for any homogeneous automorphisms  $\psi, \phi$ .*

*Proof.* Let  $\psi_0 : \text{St}_0 \rightarrow \text{St}_0$  be the projection of  $\psi$ , and suppose  $\xi, \eta$  are two choices of  $\mathbb{Z}$ -coordinates with Frobenius functions  $F_\psi^\xi, F_\psi^\eta$ . We claim that  $\int F_\psi^\xi = \int F_\psi^\eta$ .

Define  $\Delta : \text{St}_0 \rightarrow \mathbb{Z}$ ,  $\Delta(p) = \xi(\tilde{p}) - \eta(\tilde{p})$  for some (any)  $\tilde{p} \in \pi^{-1}(p)$ . The definition is proper since  $\pi^{-1}(p)$  is a  $D$ -orbit, and  $(\xi - \eta) \circ D = (\xi + 1) - (\eta + 1) = \xi - \eta$ . A simple calculation shows that  $F_\psi^\xi - F_\psi^\eta = \Delta \circ \psi_0 - \Delta$ . Since  $\psi_0$  is measure preserving,  $\int (F_\psi^\xi - F_\psi^\eta) = \int (\Delta \circ \psi_0 - \Delta) = 0$ , and  $\int F_\psi^\xi = \int F_\psi^\eta$ .

Next suppose  $\psi, \phi$  are two homogeneous automorphisms. It is easy to see that  $\psi \circ \phi$  is a homogeneous automorphism, and for every  $p \in \text{St}_0$  and  $\tilde{p} \in \pi^{-1}(p)$ ,

$$\begin{aligned} F_{\psi \circ \phi}(p) &= \xi[\psi(\phi(\tilde{p}))] - \xi[\tilde{p}] = \xi[\psi(\phi(\tilde{p}))] - \xi[\phi(\tilde{p})] + \xi[\phi(\tilde{p})] - \xi[\tilde{p}] \\ &= (F_\psi \circ \phi_0)(p) + F_\phi(p), \text{ where } \phi_0 \text{ is the projection of } \phi. \end{aligned}$$

Since  $\phi_0$  is area preserving,  $\delta(\psi \circ \phi) = \int F_\psi \circ \phi_0 + \int F_\phi = \delta(\psi) + \delta(\phi)$ .  $\square$

By [HHW] the set of derivatives of homogeneous automorphisms equals

$$\Gamma = \{A \in \mathrm{SL}(2, \mathbb{Z}) : A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \pmod{2}\}$$

Here is a refinement of this statement. The proof is in the appendix.

**Proposition 2.3** (Classification of homogeneous automorphisms).

- (1) If  $A \in \mathrm{SL}(2, \mathbb{Z})$ ,  $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{2}$  and  $\delta_0 \in \mathbb{Z}$ , then there is a unique homogeneous automorphism with derivative  $A$  and drift  $\delta_0$ .
- (2) If  $A \in \mathrm{SL}(2, \mathbb{Z})$ ,  $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \pmod{2}$  and  $\delta_0 \in \frac{1}{2} + \mathbb{Z}$ , then there is a unique homogeneous automorphism with derivative  $A$  and drift  $\delta_0$ .
- (3) No other homogeneous automorphisms exist.

**Renormalizing hyperbolic automorphisms.** A homogeneous automorphism of  $\mathrm{St}$  is called *hyperbolic* if its derivative matrix has two real eigenvalues,  $\lambda, \lambda^{-1}$ , where  $0 < |\lambda| < 1$ .

**Definition 2.4.** A hyperbolic homogeneous automorphism  $\psi$  renormalizes  $\alpha \in \mathbb{R}$ , if  $\alpha = \frac{1}{2} + \frac{1}{2} \tan \theta \pmod{1}$  where  $\begin{pmatrix} \sin \theta \\ \cos \theta \end{pmatrix}$  is an eigenvector of the derivative of  $\psi$ . In this case we say that  $\alpha$  is renormalized by  $\psi$ .

The motivation is that if  $\alpha = \frac{1}{2} + \frac{1}{2} \tan \theta \pmod{1}$ , then  $T_\alpha$  is the Poincaré map of the linear flow in direction  $\theta$ ,  $\varphi_\theta : \mathrm{St} \rightarrow \mathrm{St}$ , and

$$\psi \circ \varphi_\theta^t = \varphi_\theta^{\lambda t} \circ \psi$$

where  $\lambda$  is the eigenvalue of  $\begin{pmatrix} \sin \theta \\ \cos \theta \end{pmatrix}$ .

There is no loss of generality in assuming that (a) the eigenvalues are positive, (b)  $\psi$  fixes the singularities of  $\mathrm{St}$ , (c)  $\psi$  has zero drift, and (d)  $0 < \lambda < 1$ : We saw above that homogeneous automorphism  $\psi$  has drift in  $\frac{1}{2}\mathbb{Z}$ , so  $2\delta(\psi)$  is always an integer. One of the automorphisms  $D^{-4\delta(\psi)}\psi^4, D^{4\delta(\psi)}\psi^{-4}$  satisfies (a),(b),(c),(d).

We characterize the irrational numbers  $\alpha$  which possess renormalizing automorphisms. Recall that a *quadratic irrational* is an irrational  $\alpha$  s.t.  $a\alpha^2 + b\alpha + c = 0$  for some  $a, b, c \in \mathbb{Z}$  not all equal to zero.

**Proposition 2.5.**  $\alpha$  is renormalized by a hyperbolic homogeneous automorphism iff it is a quadratic irrational.

*Proof.* The derivative of a hyperbolic homogeneous automorphism belongs to  $\mathrm{SL}(2, \mathbb{Z})$ . The eigenvalues of such matrices are quadratic irrationals, and the slopes of the eigenvectors of such matrices are quadratic irrationals. It follows that all irrationals with renormalizing hyperbolic automorphisms are quadratic.

For the converse suppose that  $\alpha$  is a quadratic irrational. We prove that a renormalizing automorphism exists. Let  $\alpha' := 2\alpha - 1$ . This is also a quadratic irrational.

By Lagrange's Theorem, the continued fraction expansion of  $\alpha'$  is eventually periodic. So there is a map  $\varphi(z) = \frac{a'z+b'}{c'z+d'}$  with  $\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z})$  s.t. the continued fraction expansion of  $\varphi(\alpha')$  is (completely) periodic:

$$\varphi(\alpha') = [a_0, \dots, a_{n-1}, a_0, \dots, a_{n-1}, \dots]. \quad (2.1)$$

Let  $\frac{p_k}{q_k}$  denote the principal convergents of  $\beta := \varphi(\alpha')$ . By the theory of continued fractions,  $\det \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} = (-1)^{n+1}$ , and  $\beta$  is a fixed point of  $\psi(z) = \frac{p_n z + p_{n-1}}{q_n z + q_{n-1}}$ . So  $\psi[\varphi(\alpha')] = \varphi(\alpha')$ , whence  $(\varphi^{-1}\psi\varphi)(\alpha') = \alpha'$ .

Let  $\phi := \varphi^{-1}\psi\varphi$ , then  $\phi^N(\alpha') = \alpha'$  for all  $N$ . We claim that for some  $N$ ,  $\phi^N$  is a Möbius transformation with matrix belonging to

$$\Gamma(2) := \{A \in \mathrm{SL}(2, \mathbb{Z}) : A \equiv \mathrm{Id} \pmod{2}\}.$$

Let  $A$  be the matrix which represents  $\phi^2$ . Obviously,  $\phi^2 \in \mathrm{SL}(2, \mathbb{Z})$ . Let  $[A]_2 \in \mathrm{SL}(2, \mathbb{Z}_2)$  denote the residue class of  $A \pmod{2}$ . The group  $\mathrm{SL}(2, \mathbb{Z}_2)$  is finite, therefore  $[A^N]_2 = ([A]_2)^N = \mathrm{Id}$  for some  $N$ . So  $\phi^{2N}$  is represented by a matrix in  $\Gamma(2)$ , proving the claim.

Write  $\phi^{2N}(z) = \frac{c+dz}{a+bz}$  for  $\begin{pmatrix} d & c \\ b & a \end{pmatrix} \in \Gamma(2)$ , then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ \alpha' \end{pmatrix} = \begin{pmatrix} a + b\alpha' \\ c + d\alpha' \end{pmatrix} = \frac{1}{a + b\alpha'} \begin{pmatrix} 1 \\ \phi^{2N}(\alpha') \end{pmatrix} = \frac{1}{a + b\alpha'} \begin{pmatrix} 1 \\ \alpha' \end{pmatrix}, \quad (2.2)$$

proving that  $\begin{pmatrix} 1 \\ \alpha' \end{pmatrix}$  is an eigenvector of  $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(2)$ . This matrix is hyperbolic, because its trace is bigger than two:  $a + d = \mathrm{tr}[\phi^{2N}] = \mathrm{tr} \left[ \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix}^{2N} \right]$ , and every  $2 \times 2$  matrix with determinant one and all of whose entries are positive integers, has trace bigger than two.

By (2.2), the homogeneous automorphism with zero drift and derivative  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  renormalizes  $\alpha = \frac{1}{2} + \frac{1}{2}\alpha'$ .  $\square$

The previous proof is constructive, but it does not provide a convenient tool for calculating renormalizing automorphisms. This is the purpose of the next result.

**Proposition 2.6.** *Any quadratic irrational  $\alpha$  equals  $\frac{1}{2} + \frac{k + \sqrt{q(q+1)}}{2n} \pmod{1}$  for some  $k, q, n \in \mathbb{Z}$  satisfying  $q(q+1) \neq 0$  and  $n|k^2 - q(q+1)$ . In this case there is a renormalizing homogeneous automorphism  $\psi$  with zero drift and derivative*

$$d\psi = \begin{pmatrix} 2(q-k) + 1 & 2 \cdot \frac{k^2 - q(q+1)}{n} \\ -2n & 2(q+k) + 1 \end{pmatrix}. \quad (2.3)$$

**Example:** For  $\alpha = \sqrt{2}$ , we can take  $k = n = 3, q = 8$ , and get the homogeneous automorphism with zero drift and derivative  $\begin{pmatrix} 11 & -42 \\ -6 & 23 \end{pmatrix}$ .

Similar formulas can be obtained for  $\sqrt{3}$  ( $k = n = 1, q = 3$ ),  $\sqrt{5}$  ( $k = n = 1, q = 4$ ),  $\sqrt{7}$  ( $k = n = 12, q = 63$ ) etc.

*Proof.* Since  $\alpha$  is a quadratic irrational, it has a hyperbolic renormalizing automorphism with zero drift. Let  $A$  be the derivative. By proposition 2.3,  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  with  $a, d$  odd and  $b, c$  even, and  $ad - bc = 1$ .

We claim that  $\mathrm{tr}(A) \equiv 2 \pmod{4}$ . Since  $a, d$  are odd, they are equal to  $\pm 1 \pmod{4}$ . Write  $a = 4\alpha + \varepsilon, d = 4\beta + \eta, c = 2\gamma, c = 2\delta$  with  $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$  and  $\varepsilon, \eta = \pm 1$ . Since  $1 = \det(A) \equiv \varepsilon\eta \pmod{4}$ ,  $\varepsilon = \eta$ . It follows that  $a \equiv d \pmod{4}$  and  $\mathrm{tr} A = a + d = 4(\alpha + \beta) \pm 2 \in 4\mathbb{Z} + 2$ .

Write  $\mathrm{tr}(A) = 4q + 2$  with some  $q \in \mathbb{Z}$ . Since  $a, d$  are odd and  $a + d = 4q + 2$ , we can put  $a, d$  in the form  $a = 2(q - k) + 1$  and  $d = 2(q + k) + 1$  with  $k \in \mathbb{Z}$ .

Since  $c$  is even,  $c = -2n$  with some  $n \in \mathbb{Z}$ . Since  $ad - bc = 1$ , either  $n = 0$  and  $A = \text{Id}$ , or  $n \neq 0$  and  $b = 2 \cdot \frac{k^2 - q(q+1)}{n}$ . So  $A = \begin{pmatrix} 2(q-k) + 1 & 2 \cdot \frac{k^2 - q(q+1)}{n} \\ -2n & 2(q+k) + 1 \end{pmatrix}$  with  $q, n, k \in \mathbb{Z}$  s.t.  $n \neq 0$  and  $n|k^2 - q(q+1)$ . Every choice of  $k, q, n$  like that determines a matrix in  $\text{SL}(2, \mathbb{Z})$  equal to  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{2}$ .

The characteristic polynomial of  $A$  is  $x^2 - x \text{tr} A + \det A = x^2 - (4q+2)x + 1$ . The eigenvalues are  $(2q+1) \pm 2\sqrt{q(q+1)}$ .  $A$  is hyperbolic iff  $q(q+1) \neq 0$ . The eigenvectors are proportional to  $(\frac{k \pm \sqrt{q(q+1)}}{n}, 1)$ , so the automorphism with derivative  $A$  renormalizes  $\alpha := \frac{1}{2} + \frac{k \pm \sqrt{q(q+1)}}{n}$ . Playing with the signs of  $k, n$  we see that there is no loss in taking  $\alpha := \frac{1}{2} + \frac{k + \sqrt{q(q+1)}}{n}$ .  $\square$

**Markov partitions and symbolic dynamics.** Every hyperbolic homogeneous automorphism  $\psi : \text{St} \rightarrow \text{St}$  covers a hyperbolic toral automorphism  $\psi_0 : \text{St}_0 \rightarrow \text{St}_0$ . Adler and Weiss introduced in [AW] a technique for coding  $\psi_0 : \text{St}_0 \rightarrow \text{St}_0$  as the action of the left shift map on the collection of two sided infinite paths on a finite directed graph. This is done using *Markov partitions*. The purpose of this section is to describe this method.

The original work of Adler & Weiss applies to general hyperbolic automorphisms. It is important to our purposes to carry out the Adler–Weiss construction in a way which respects that fact that  $\psi_0$  fixes the punctures of  $\text{St}_0$  and has derivative matrix

$$A \in \Gamma(2) := \{A \in \text{SL}(2, \mathbb{Z}) : A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{2}\}.$$

We will assume for simplicity that  $A$  has positive eigenvalues,  $0 < \lambda < 1$  and  $\lambda^{-1} > 1$ . Then there are vectors  $\underline{v} = \begin{pmatrix} 1 \\ v \end{pmatrix}$ ,  $\underline{w} = \begin{pmatrix} 1 \\ w \end{pmatrix}$  such that  $A\underline{v} = \lambda^{-1}\underline{v}$  and  $A\underline{w} = \lambda\underline{w}$ . Since  $A \in \Gamma(2)$ ,  $v, w$  are irrational. We call  $\underline{w}$  the *stable* direction and  $\underline{v}$  the *unstable* direction (of  $\psi_0$ ).

The first step in the Adler–Weiss construction is to divide the torus into two parallelograms  $Q_1, Q_2$  with sides parallel to  $\underline{v}, \underline{w}$ . They cut the torus along two line segments emanating from a single fixed point. We prefer to use one segment passing through the first puncture, and the other passing through the second puncture: this simplifies the analysis of the coded Frobenius function, see §6 below.

Suppose first that  $-1 < w < 0, v > 1$  (case 1), or  $0 < w < 1, v > 1$  (case 2). Then  $Q_1, Q_2$  can be constructed as in Figure 2. Notice that one of the parallelograms (which we call  $Q_1$ ) contains one or no punctures in its closure, and the other (which we call  $Q_2$ ) contains both punctures in its closure.

The general case can be reduced to case 1 or 2 by working with  $\theta \circ \psi_0 \circ \theta^{-1}$  or  $\theta \circ \psi_0^{-1} \circ \theta^{-1}$  for a suitable toral automorphism  $\theta : \text{St}_0 \rightarrow \text{St}_0$  which fixes the punctures. The derivative matrix of  $\theta$  is produced from the following lemma, applied to the irrational numbers  $\xi = v^{-1}, \eta = w^{-1}$  (see the appendix for proof):

**Lemma 2.7.** *For every  $\xi, \eta \in \mathbb{R} \setminus \mathbb{Q}$  s.t.  $\xi \neq \eta$  there exists  $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(2)$  such that  $s_1 := \frac{a\xi + b}{c\xi + d}, s_2 := \frac{a\eta + b}{c\eta + d}$  satisfy one of the following: One of  $s_1, s_2$  is in  $(0, 1)$  and the other is in  $(1, \infty)$ ; Or one of  $s_1, s_2$  is in  $(-1, 0)$  and the other is in  $(1, \infty)$ .*

$\theta$  itself can be produced using Proposition 2.3 by projecting the homogeneous automorphism with zero drift and derivative  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  to  $\text{St}_0$ . Homogeneity guarantees that  $\theta$  fixes the punctures.



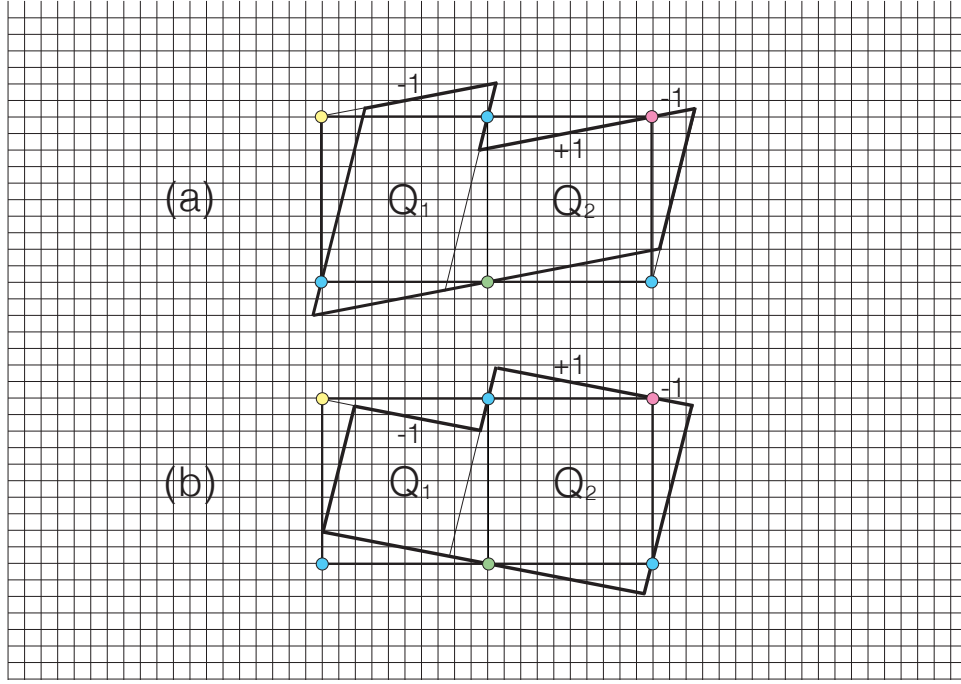


FIGURE 2. Partition of the torus

We call  $R := Q_1 \cup Q_2$  the *fundamental polygon* of  $\psi_0$ . The sides of  $R$  in direction  $\underline{w}$  (resp.  $\underline{v}$ ) are called *stable* (resp. *unstable*). Let  $\partial_s R :=$ union of stable sides of  $R$  and  $\partial_u R :=$ union of unstable sides of  $R$ . Since  $\partial_s R, \partial_u R$  are linear segments containing fixed points of  $\psi_0$  and in the direction of eigenvectors of  $d\psi_0$ , we have  $\psi_0(\partial_s R) \subset \partial_s R$  and  $\psi_0^{-1}(\partial_u R) \subset \partial_u R$ .

A *u-fibre* is a linear segment in direction  $\underline{v}$  with endpoints in  $\partial_s R$ . Since  $\psi_0(\partial_s R) \subset \partial_s R$  and  $A\underline{v} = \lambda\underline{v}$ , the  $\psi_0$ -image of a *u-fibre* is a finite union of *u-fibres*. Similarly, an *s-fibre* is a linear segment in direction  $\underline{w}$  and endpoints in  $\partial_u R$ . The  $\psi_0$ -image of an *s-fibre* is a subset of an *s-fibre*. We orient *u/s-fibres* in the direction of  $\underline{v}, \underline{w}$ .

Thus  $\psi_0(Q_i)$  is a finite union of non-overlapping parallelograms  $Q_{i1}, \dots, Q_{iN_i} \subset R$  with sides in the stable and unstable directions, and with *s*-sides contained in  $\partial_s R$ . We use the following convention for the order  $Q_{i1}, \dots, Q_{iN_i}$  ( $i = 1, 2$ ): Recall that *u-fibres* are oriented in the direction of  $\underline{u}$ , then every parallelogram  $Q_{ij}$  has a *bottom s-side*, and a *top s-side*. The ordering is done so that the top side of  $Q_{ij}$  is identified with the bottom side of  $Q_{i,j+1}$  ( $j = 1, \dots, N_i - 1$ ).

The interior of  $Q_{ij}$  is completely contained in the interiors of  $Q_k$  for  $k = 1$  or  $2$ . Otherwise,  $\psi_0(\text{int}(Q_{ij}))$  intersects  $\partial_u R$ , in contradiction to  $\psi_0^{-1}(\partial_u R) \subset \partial_u R$ .

Let  $\mathfrak{P} := \{Q_{ij} : i = 1, 2; 1 \leq j \leq N_i\}$ . Since  $\psi_0$  is bijective,  $\mathfrak{P}$  is a partition of  $\text{St}_0$ . By the previous paragraph,  $\mathfrak{P}$  is a refinement of  $\{Q_1, Q_2\}$ .  $\mathfrak{P}$  is the *Adler-Weiss Markov partition*.

The *dynamical graph* of  $\mathfrak{P}$  is the directed graph  $\mathcal{G}$  with set of vertices  $\mathfrak{P}$  and edges  $P_i \rightarrow P_j$  for any pair of  $P_i, P_j \in \mathfrak{P}$  s.t.  $\text{int}(P_i \cap \psi_0^{-1}(P_j)) \neq \emptyset$ . Let  $\Sigma(\mathcal{G})$

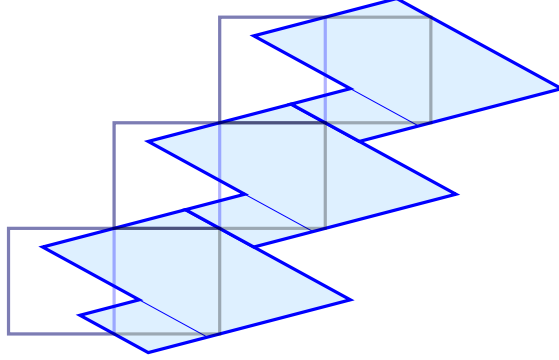


FIGURE 3. The  $\mathbb{Z}$ -coordinate associated to the canonical renormalizing automorphism of  $\sqrt{2}$

denote the collection of bi-infinite paths on  $\mathcal{G}$ :

$$\Sigma(\mathcal{G}) := \{(P_k)_{k \in \mathbb{Z}} \in \mathfrak{P}^{\mathbb{Z}} : P_k \rightarrow P_{k+1} \text{ for every } k \in \mathbb{Z}\}.$$

Equip  $\Sigma(\mathcal{G})$  with the metric  $d(\underline{x}, \underline{y}) := \exp(-\min\{|k| : x_k \neq y_k\})$ . Let  $\sigma : \Sigma(\mathcal{G}) \rightarrow \Sigma(\mathcal{G})$  denote the *left shift map*,  $\sigma : (x_i)_{i \in \mathbb{Z}} \mapsto (x_{i+1})_{i \in \mathbb{Z}}$ .

**Theorem 2.8** (Adler and Weiss). *For every  $(P_i)_{i \in \mathbb{Z}} \in \Sigma(\mathcal{G})$ , there is a unique point  $\pi_0[(P_i)_{i \in \mathbb{Z}}] \in \bigcap_{i \in \mathbb{Z}} \psi_0^{-i}(\overline{P_i})$ , and  $\pi_0 : \Sigma(\mathcal{G}) \rightarrow \text{St}_0$  has the following properties:*

- (1)  $\pi_0 : \Sigma(\mathcal{G}) \rightarrow \text{St}_0$  is onto and  $|\pi_0^{-1}(p)| = 1$  for Lebesgue almost every  $p \in \text{St}_0$
- (2)  $\pi_0$  is Hölder continuous and  $\pi_0 \circ \sigma = \psi_0 \circ \pi_0$
- (3) Let  $m_0$  denote the normalized Lebesgue measure on  $\text{St}_0$ , then  $m_0 \circ \pi_0^{-1}$  is a mixing stationary Markov measure on  $\Sigma$ .
- (4)  $m_0 \circ \pi_0^{-1}$  is the measure of maximal entropy for  $\sigma : \Sigma(\mathcal{G}) \rightarrow \Sigma(\mathcal{G})$ .

See [AW] for proof. Additional information on the combinatorial structure of  $\mathcal{G}$  can be found in §6.

Let  $\tilde{R}$  denote a connected lift of the fundamental polygon  $Q_1 \cup Q_2$  to  $\text{St}$ . The corresponding  $\mathbb{Z}$ -coordinate  $\xi : \text{St} \rightarrow \mathbb{Z}$  is called the  *$\mathbb{Z}$ -coordinate associated to the automorphism  $\psi$* , see figure 3.

The main advantage of the associated  $\mathbb{Z}$ -coordinate is the following fact, whose proof we defer for reasons of exposition to §6 (Lemma 6.8): If  $F_\psi$  is the Frobenius function of  $\psi$  with respect to the associated  $\mathbb{Z}$ -coordinate of  $\psi$ , then

$$F_\psi \text{ is } \mathfrak{P} \vee \psi_0^{-1}(\mathfrak{P})\text{-measurable or } \mathfrak{P} \vee \psi_0(\mathfrak{P})\text{-measurable.} \quad (2.4)$$

This means there exists a function  $g : \mathfrak{P} \times \mathfrak{P} \rightarrow \mathbb{Z}$  s.t. the *coded Frobenius function*

$$F := F_\psi \circ \pi_0 : \Sigma(\mathcal{G}) \rightarrow \mathbb{Z}$$

takes the form  $F[\underline{P}] = g(P_0, P_1)$  or  $F[(P_i)_{i \in \mathbb{Z}}] = g(P_{-1}, P_0)$  ( $\underline{P} = (P_i)_{i \in \mathbb{Z}} \in \Sigma(\mathcal{G})$ ).

The following additional property of  $F$  is proved in the appendix.

**Lemma 2.9** (Aperiodicity Lemma). *If  $e^{itF} = zh/h \circ \sigma$  where  $|z| = 1$ ,  $t \in \mathbb{R}$ , and  $h : \Sigma(\mathcal{G}) \rightarrow \mathbb{C}$  is continuous, then  $z = 1$ ,  $t \in 2\pi\mathbb{Z}$ , and  $h = \text{const}$ .*

This is called the *aperiodicity condition* in [GH], and should be viewed as a strong way of saying that  $F$  does not take values in a set of the form  $a + b\mathbb{Z}$  “up to a coboundary.” The aperiodicity condition is used in §3.

**The twist at a singularity.** The contents of this section are only used in §5.

Suppose  $\psi$  is a homogeneous hyperbolic automorphism of the infinite staircase, and let  $p$  denote one of the four singularities of St. Recall that  $D^2(p) = p$  and  $\psi^2(p) = p$ .

Let  $\underline{w}$  be some non-zero vector. There are infinitely many rays emanating from  $p$  in direction  $\underline{w}$ : one for each horizontal rectangle with vertex congruent to  $p$  such that the vector  $\underline{w}$  based at  $p$  points into the rectangle. Let  $L_i(p, \underline{w})$  denote the ray which starts at horizontal rectangle number  $i$ . So  $D(L_i(p, \underline{w})) = L_{i+1}(D(p), \underline{w})$ .

Now suppose  $\underline{w}$  is an eigenvector of  $d\psi^n$  for some  $n$ . Then  $d\psi^{2n}(\underline{w}) = \lambda\underline{w}$  with  $\lambda > 0$ , and  $\psi^{2n}(p) = p$ . It follows that  $\psi^{2n}[L_i(p, \underline{w})] = L_j(p, \underline{w})$  for some  $j = j(i)$ . It is not difficult to see that  $(j - i)/2n$  is independent of the choice of  $i$  and  $n$ .

**Definition 2.10.** The twist of  $\underline{w}$  at  $p$  is  $\tau_\psi(p, \underline{w}) := \frac{1}{2n}(j - i)$ .

**Lemma 2.11.**  $\tau_\psi(p, \underline{w}) \in \frac{1}{2}\mathbb{Z}$ . If  $\psi$  is hyperbolic with positive eigenvalues, then  $\tau_\psi(p, \underline{w}) \in \mathbb{Z}$ .

*Proof.* The first statement is because if  $d\psi(\underline{w}) = \lambda\underline{w}$  then we can take  $n = 1$ .

Now suppose in addition that  $\lambda > 0$ . If  $\psi(p) = p$ , then  $\psi[L_i(p, \underline{w})] = L_{i+k}(p, \underline{w})$  for some integer  $k$ , and therefore  $\psi^2[L_i(p, \underline{w})] = L_{i+2k}(p, \underline{w})$  and  $\tau_\psi(p, \underline{w}) = k \in \mathbb{Z}$ . If  $\psi(p) \neq p$ , then by homogeneity,  $\phi := D \circ \psi$  fixes  $p$ , and by the previous paragraph  $\tau_\phi(p, \underline{w}) \in \mathbb{Z}$ . So  $\tau_\psi(p, \underline{w}) = \tau_\phi(p, \underline{w}) - 1 \in \mathbb{Z}$ .  $\square$

*Example 1.* Let  $\psi$  be the homogeneous automorphism with derivative  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$  and drift  $\frac{1}{2}$  (see figure 5). Let  $p :=$ lower left corner of horizontal rectangle #0. For any vector  $\underline{w}$  with positive coordinates,  $\psi^2[L_i(p, \underline{w})] = L_{i+1}(p, -\underline{w})$ . So  $\tau_\psi(p, \underline{w}) = \frac{1}{2}$ .

*Example 2.* Let  $\psi$  denote the renormalizing automorphism of  $\sqrt{2}$  with with zero drift and derivative  $\begin{pmatrix} 11 & -42 \\ -6 & 23 \end{pmatrix}$ , and let  $\underline{w} := \begin{pmatrix} 1+2\sqrt{2} \\ 1 \end{pmatrix}$  be its contracted eigenvector. Let  $p$  be one of the singularities at the bottom left corner of one of the horizontal rectangles, say rectangle #0. We show below (Theorem 6.3) that  $\tau_\psi(p, \underline{w}) = 1$ .

**Lemma 2.12.** Suppose  $\psi$  is a hyperbolic homogeneous automorphism with zero drift and positive eigenvalues. Let  $p$  be a singularity, and  $\underline{w}$  an eigenvector of  $d\psi$ , then  $\tau_\psi(p, \underline{w}) =$ minus the drift of  $\phi$ , where  $\phi$  is the unique homogeneous automorphism which fixes  $L_0(p, \underline{w})$ , and which has the same derivative as  $\psi$ .

*Proof.* As in the proof of the previous lemma, there exist  $k \in \mathbb{Z}$  and  $\ell = 0, 1$  s.t.  $\psi[L_i(p, \underline{w})] = D^\ell[L_{i+k}(p, \underline{w})]$ . Let  $\phi := D^{-(k+\ell)} \circ \psi$ , then  $\phi$  fixes  $L_i(p, \underline{w})$  and has the same derivative as  $\psi$ . This determines  $\phi$  uniquely, because every other homogeneous automorphism with the same derivative has the form  $D^n \circ \phi$  with  $n \neq 0$ .

By the definition of  $k$  and  $\ell$ ,  $\tau_\psi(p, \underline{w}) = k + \ell$ . By the definition of  $\phi$  and Proposition 2.2,  $\delta(\phi) = \delta(\psi) - (k + \ell) = -(k + \ell)$ .  $\square$

## 3. ESTIMATES OF BIRKHOFF SUMS

In this section we find pointwise asymptotic estimates for the Birkhoff sums of the cylinder map  $T_\alpha : \mathbb{T} \times \mathbb{Z} \rightarrow \mathbb{T} \times \mathbb{Z}$

$$T_\alpha(x, t) = (x + \alpha(\bmod 1), t + f(x)),$$

where  $\alpha$  is a quadratic irrational, and  $f = 1_{[\frac{1}{2}, 1)} - 1_{[0, \frac{1}{2})}$ .

By proposition 2.5, there is a hyperbolic homogeneous automorphism  $\psi$  with zero drift s.t.  $\alpha = \frac{1}{2} + \frac{1}{2} \tan \theta(\bmod 1)$ , where  $\begin{pmatrix} \sin \theta \\ \cos \theta \end{pmatrix}$  is an eigenvector of the derivative of  $\psi$ , with eigenvalue  $0 < \lambda < 1$ .

Recall that the infinite staircase is made from a  $\mathbb{Z}$ -array of  $2 \times 1$  horizontal rectangles. Declare one of these rectangles to be “rectangle zero” and let  $\omega : \mathbb{T} \rightarrow \text{St}$  be the function which associates to  $\omega(x)$  the unique point on the top horizontal side of rectangle zero, located  $2x$  units of distance from its lower left corner.

In what follows  $\log^* := \log_{\lambda^{-1}}$ ,  $m_0 = m_{\mathbb{T}} \times m_{\mathbb{Z}}$ , and  $\xi : \text{St} \rightarrow \mathbb{Z}$  is some (any)  $\mathbb{Z}$ -coordinate on the infinite staircase.  $C_c(\cdot)$  denotes the space of continuous functions with compact support.

**Theorem 3.1.** *There exists  $\sigma^2 > 0$  such that for every  $(x, \ell) \in \mathbb{T} \times \mathbb{Z}$  for which  $\frac{1}{k} \xi[\psi^k(\omega(x))]$   $\xrightarrow[k \rightarrow \infty]{} 0$ , for every non-negative  $G \in C_c(\mathbb{T} \times \mathbb{Z})$ ,*

$$\sum_{i=0}^{n-1} (G \circ T^i)(x, \ell) = \frac{[1 + o(1)]n \int G dm_0}{2\sigma \sqrt{\pi \log^* n}} \cdot \sqrt{2} \exp \left[ -\frac{1 + o(1)}{2\sigma^2} \left( \frac{\xi[\psi^{\lfloor \log^* n \rfloor}(\omega(x))]}{\sqrt{\log^* n}} \right)^2 \right].$$

*The following uniformity in  $x$  is observed:  $\forall \varepsilon > 0 \exists \delta, N > 0$  (which depend on  $G$  but not  $x$ ) s.t. if  $|\frac{1}{k} \xi[\psi^k(\omega(x))]| < \delta$  and  $n > N$ , then the  $o(1)$  terms are in  $[-\varepsilon, \varepsilon]$ .*

Notice that the right hand side is oscillatory, and that oscillatory term (the exponential) does not depend on  $G$ .

We will show in the next section that  $\frac{1}{k} \xi[\psi^k(\omega(x))]$   $\xrightarrow[k \rightarrow \infty]{} 0$  almost everywhere in  $\mathbb{T} \times \mathbb{Z}$ . Thus Theorem (3.1) describes the almost sure behavior of Birkhoff sums for non-negative  $G \in C_c(\mathbb{T} \times \mathbb{Z})$ . By the ratio ergodic theorem, this is the almost sure behavior of every  $L^1$  function with non-zero integral.

The constant  $\sigma^2$  has meaning. We will see in §4 that  $\frac{\xi[\psi^k(\omega(x))]}{\sqrt{k}}$  converges to the centered normal distribution with variance  $\sigma^2$ . Thus  $\sum_{i=0}^{n-1} G \circ T^i$  grows a.e. like a constant times  $\frac{n}{\sqrt{\log^* n}}$ , but if we normalize by this growth rate, then we get oscillatory behavior. The oscillations are driven by the renormalizing automorphism. Similar results were proved for horocycle flows on  $\mathbb{Z}^d$  covers of hyperbolic surfaces of finite area in [LS1], [LS2], and for Hajian-Ito-Kakutani skew products in [AS].

We will obtain Theorem 3.1 from a study of the Birkhoff integrals of the linear flow in direction  $\theta$  on the infinite staircase. Denote this flow by  $\varphi_\theta$ . We will show:

**Theorem 3.2.** *There exists  $\sigma^2 > 0$  s.t. for every  $\omega \in \text{St}$  s.t.  $\frac{1}{k} \xi[\psi^k(\omega)]$   $\xrightarrow[k \rightarrow \infty]{} 0$ , and for every  $G \in C_c(\text{St})$  such that  $\int G dm > 0$  ( $m = \text{area measure}$ ),*

$$\int_0^n G[\varphi_\theta^t(\omega)] dt = \frac{1}{2} \cdot \frac{[1 + o(1)]n \int G dm}{2\sigma \sqrt{\pi \log^* n}} \cdot \sqrt{2} \exp \left[ -\frac{1 + o(1)}{2\sigma^2} \left( \frac{\xi[\psi^{\lfloor \log^* n \rfloor}(\omega)]}{\sqrt{\log^* n}} \right)^2 \right].$$

The extra  $\frac{1}{2}$  is because  $m([0, 2] \times [0, 1]) = 2$  whereas  $m_0(\mathbb{T} \times \{0\}) = 1$ .

**Notation.** Let  $\psi_0 : \text{St}_0 \rightarrow \text{St}_0$  denote the projection of  $\psi : \text{St} \rightarrow \text{St}$  to the covered torus  $\text{St}_0$ , and let  $\mathfrak{P}$  denote the Adler–Weiss Markov partition of  $\psi_0$ .

Let  $\underline{v} = \begin{pmatrix} 1 \\ v \end{pmatrix}$  and  $\underline{w} = \begin{pmatrix} 1 \\ w \end{pmatrix}$  denote eigenvectors of the derivative of  $\psi$  with eigenvalues  $\lambda$  and  $\lambda^{-1}$ . They define the *stable* and *unstable* directions.

Linear segments in  $\text{St}$  or  $\text{St}_0$  the direction of  $\pm \underline{w}$  will be called *stable*. For example,  $\{\varphi_\theta^t(\omega_0) : a < t < b\}$  is a stable linear segment in  $\text{St}$ .

The following is a particularly useful way to generate stable linear segments. Let  $\pi_0 : \Sigma(\mathcal{G}) \rightarrow \text{St}_0$  denote the Adler Weiss coding map given by Theorem 2.8. Let  $\Sigma^+ := \{(x_0, x_1, \dots) \in \mathfrak{P}^{\mathbb{N}} : x_i \rightarrow x_{i+1} \text{ for all } i \geq 0\}$ . For every  $\underline{x} \in \Sigma^+$ , let

$$W^s(\underline{x}) := \pi_0\{\underline{y} \in \Sigma(\mathcal{G}) : y_i = x_i \ (i \geq 0)\}.$$

**Lemma 3.3.**  *$W^s(\underline{x})$  is a stable linear segment. It is the  $s$ -fibre through  $\pi_0(\underline{x})$  in rectangle  $x_0 \in \mathfrak{P}$ . Let  $h(\underline{x}) := \ell^s(x_0)$  be its length, then  $\sum_{\sigma(\underline{y})=\underline{x}} h(\underline{y}) = \lambda^{-1}h(\underline{x})$ , where the sum ranges over  $\underline{y} \in \Sigma^+$  and  $\sigma(y_0, y_1, \dots) := (y_1, y_2, \dots)$ .*

*Proof.* By the Markov property of  $\psi_0$ ,  $C_n := x_0 \cap \psi_0^{-1}(x_1) \cap \dots \cap \psi_0^{-(n-1)}(x_{n-1})$  is a decreasing intersection of compact parallelograms with  $s$ -side of length  $\ell^s(x_0)$  and  $u$ -side of length  $O(\lambda^n)$ . The intersection is necessarily an  $s$ -fibre in  $x_0$ , which passes through  $\pi_0(\underline{x})$ . This is a stable linear segment.

The Markov property also implies that  $\psi_0^{-1}[W^s(\underline{x})] = \bigcup_{\sigma(\underline{y})=\underline{x}} W^s(\underline{y})$ . Since  $\psi_0$  contracts  $s$ -fibres linearly by factor  $\lambda$ ,  $\sum_{\sigma(\underline{y})=\underline{x}} h(\underline{y}) = \lambda^{-1}h(\underline{x})$ .  $\square$

Let  $W^s(\underline{x}, k) := \text{lift of } W^s(\underline{x}) \text{ to } \text{St} \text{ so that } \pi_0(\underline{x}) \text{ lifts to a point in } \xi = k$ . This is a stable segment in  $\text{St}$ , and it has length  $h(x_0)$ .  $W^s(\underline{x}, k) \subseteq [\xi = k]$ , because  $W^s(\underline{x})$  lies completely inside an element of  $\mathfrak{P}$ , and such sets lift in their entirety to subsets of  $D^i(F)$  ( $i \in \mathbb{Z}$ ) where  $F$  is the fundamental polygon of  $\psi$ .

**Proof of Theorem 3.2.** We begin with some reductions.

Any two  $\mathbb{Z}$ -coordinates are within uniformly bounded distance from one another, therefore if the theorem holds with one choice of a  $\mathbb{Z}$ -coordinate, then it holds with all other possible choices. We will work with the  $\mathbb{Z}$ -coordinate associated to  $\psi$ .

With this choice of  $\xi$ , the Frobenius function is either  $\mathfrak{P} \vee \psi_0^{-1}(\mathfrak{P})$ -measurable, or  $\mathfrak{P} \vee \psi_0(\mathfrak{P})$ -measurable. We will carry out the proof in the first case, and leave to the reader the (routine) modifications needed for the second case.

A *cylinder* is a set of the form  ${}_\ell[P_\ell, \dots, P_{\ell'}] := \bigcap_{i=\ell}^{\ell'} \psi_0^{-i}(P_i)$ , where  $P_i \in \mathfrak{P}$ . This is a parallelogram with sides parallel to  $\underline{v}$  and  $\underline{w}$ . Cylinders are subsets of  $\text{St}_0$ .

Instead of working with  $G \in C_c(\text{St})$ , we will work with indicators of lifts of cylinders to  $\text{St}_0$ . Any non-negative continuous function with compact support can be sandwiched between linear combinations of such functions, so this suffices for our purposes.

Here is the precise definition of the sets which we will work with:

$${}_\ell[P_\ell, \dots, P_{\ell'}]^k := \text{lift to } \{\xi = k\} \text{ of } {}_\ell[P_\ell, \dots, P_{\ell'}] := \bigcap_{i=\ell}^{\ell'} \psi_0^{-i}(P_i).$$

Here  $\ell' < \ell$  and  $P_{\ell'}, \dots, P_\ell \in \mathfrak{P}$  are arbitrary. In fact most of our calculations will be done in the special case  $\ell = k = 0$ . This is enough, because  $\exists i, j$  s.t.  ${}_\ell[P_\ell, \dots, P_{\ell'}]^k = (D^i \circ \psi^j)_0({}_0[P_\ell, \dots, P_{\ell'}]^0)$  where  $D$  is a deck transformation, and  $D^i \circ \psi^j$  preserves the area measure and does not affect the asymptotic drift  $\lim \xi[\psi^n(\omega)]/n$ .

Similarly we may assume without loss of generality that  $\xi(\omega) = 0$ . From now on, fix  $\omega \in \text{St}$  s.t.  $\xi(\omega) = 0$  and  $\frac{\xi_n(\omega)}{n} \rightarrow 0$ , and let

$$E := {}_0[P_0, \dots, P_{\ell-1}]^0.$$

We analyze  $\int_0^n 1_E[\varphi_\theta^t(\omega)]dt$ .

In what follows  $\ell[\cdot]$  is the euclidean length measure, and  $n_0 \in \mathbb{N}$  is a free parameter that will be calibrated at the end of the proof. For every  $n$ , let

$$n^* := \lfloor \log_{\lambda^{-1}}(n/n_0) \rfloor.$$

Notice that  $\lambda^{n^*} \cdot n \in [\lambda n_0, n_0]$ .

Let  $A_n(\omega) := \{\varphi_\theta^t(\omega) : 0 < t < n\}$ . This a stable linear segment, and we are interested in  $\int_0^n 1_E[\varphi_\theta^t(\omega)]dt = \ell[A_n(\omega) \cap E]$ .

Let  $B_n(\omega) := \psi^{n^*}[A_n(\omega)]$ . Since  $\psi$  contracts stable linear segments by factor  $\lambda$ ,  $B_n(\omega)$  is a stable linear segment with length  $\ell[B_n(\omega)] \in [\lambda n_0, n_0]$ . Break  $B_n(\omega)$  into a finite union of lifted  $s$ -fibres  $W^s(\underline{x}^{(1)}, \xi_1^*), \dots, W^s(\underline{x}^{(n_1)}, \xi_{n_1}^*)$  plus two pieces of stable fibres  $W^s(\underline{x}^{(0)}, \xi_0^*), W^s(\underline{x}^{(n_1+1)}, \xi_{n_1+1}^*)$  to take care of edge effects:

$$\biguplus_{i=1}^{n_1} W^s(\underline{x}^{(i)}, \xi_i^*) \subseteq B_n(\omega) \subseteq \biguplus_{i=0}^{n_1+1} W^s(\underline{x}^{(i)}, \xi_i^*). \quad (3.1)$$

Even though  $n_1, \underline{x}^{(i)}$  and  $\xi_i^*$  depend on  $n$ , some uniformities are observed:

- (1)  $\frac{\lambda n_0}{\max h} - 2 \leq n_1 \leq \frac{n_0}{\lambda \min h}$  ( $\because \lambda n_0 \leq \ell[B_n] \leq n_0$ ,  $\ell[W^s(\underline{x}^{(i)}, \xi_i^*)] = h(x_0^{(i)})$ ).
- (2)  $|\xi_i^* - \xi(\psi^{n^*}(\omega))| < \frac{n_0}{\min h}$  for all  $i$ , because  $\xi_0^*, \dots, \xi_{n_1+1}^*, \xi(\psi^{n^*}(\omega))$  are  $\mathbb{Z}$  coordinates of points in  $B_n(\omega)$ ,  $\ell[B_n(\omega)] \leq n_0$ , and because it takes at least  $\min h$  units of distance to cross the fundamental polygon of  $\xi$  when moving in the stable direction.

By the definition of  $B_n(\omega)$ ,  $\int_0^n 1_E(\varphi_\theta^t(\omega))dt = \ell[E \cap \psi^{-n^*}(B_n(\omega))]$ , so (3.1) translates to

$$\sum_{i=1}^{n_1} J(\underline{x}^{(i)}, \xi_i^*) \leq \int_0^n 1_E[\varphi_\theta^t(\omega)]dt \leq \sum_{i=0}^{n_1+1} J(\underline{x}^{(i)}, \xi_i^*), \quad (3.2)$$

where  $J_n(\underline{x}^{(i)}, \xi_i^*) := \ell[E \cap \psi^{-n^*}(W^s(\underline{x}^{(i)}, \xi_i^*))]$ . The remainder of the proof is dedicated to the analysis of  $J_n(\underline{x}^{(i)}, \xi_i^*)$ .

We start by asking when does a point  $\omega' \in W^s(\underline{x}^{(i)}, \xi_i^*)$  belong to  $\psi^{n^*}(E)$ . We claim that  $\psi^{-n^*}(\omega') \in E$  iff  $\psi_0^{-n^*}[\pi(\omega')] \in \pi(E)$  and  $(\sum_{j=0}^{n^*-1} F_\psi \circ \psi_0^j)[\psi_0^{-n^*}(\pi(\omega'))] = \xi_i^*$ , where  $\pi$  is the covering  $\text{St} \rightarrow \text{St}_0$ .

Explanation: By the definition of  $F_\psi$ , if  $\omega' \in W^s(\underline{x}^{(i)}, \xi_i^*)$ , then  $\xi_i^* - \xi[\psi^{-n^*}(\omega')] = \xi(\omega') - \xi[\psi^{-n^*}(\omega')] \equiv F_\psi[\psi_0^{-n^*}(\pi(\omega'))] + \dots + F_\psi[\psi_0^{-1}(\pi(\omega'))]$ . It follows that  $\xi[\psi^{-n^*}(\omega')] = 0 \Leftrightarrow (\sum_{j=0}^{n^*-1} F_\psi \circ \psi_0^j)[\psi_0^{-n^*}(\pi(\omega'))] = \xi_i^*$ .

Writing  $\omega'' := \psi_0^{-n^*}[\pi(\omega')]$  (a point in  $\text{St}_0$ ), we see that

$$J_n(\underline{x}^{(i)}, \xi_i^*) = \ell\{\omega'' \in {}_0[P_0, \dots, P_{\ell-1}]^0 : \psi_0^{n^*}(\omega'') \in W^s(\underline{x}^{(i)}), \text{ and } \sum_{j=0}^{n^*-1} F_\psi[\psi_0^j(\omega'')] = \xi_i^*\}. \quad (3.3)$$

We write this in more convenient form. Let  $\sigma : \Sigma^+ \rightarrow \Sigma^+$  denote the one-sided shift defined before Lemma 3.3. The assumption that  $F_{\psi}$  is  $\mathfrak{A} \vee \psi_0^{-1}\mathfrak{A}$ -measurable allows us to view  $F := F_{\psi} \circ \pi_0$  as a function on  $\Sigma^+$ ,  $F(\underline{x}) = g(x_0, x_1)$ . By the Markov property,  $\psi_0^{-n^*} [W^s(\underline{x}^{(i)})] = \biguplus_{\sigma^{n^*}(\underline{y})=\underline{x}^{(i)}} W^s(\underline{y})$ , and since  $F(\underline{y}) = g(y_0, y_1)$ ,

$$F_{n^*}(\underline{y}) := F(\underline{y}) + F(\sigma(\underline{y})) + \cdots + F(\sigma^{n^*-1}(\underline{y}))$$

is constant on  $W^s(\underline{y})$ . It follows that

$$J_n(\underline{x}^{(i)}, \xi_i^*) = \sum_{\sigma^n(\underline{y})=\underline{x}^{(i)}} h(y_0) 1_{[P]}(\underline{y}) \delta_0(F_{n^*}(\underline{y}) - \xi_i^*).$$

Here  $h(y_0)$  is the length of the stable side of the parallelogram  $y_0$ ,  $1_{[P]}(\underline{y})$  equals one when  $(y_0, \dots, y_{\ell-1}) = (P_0, \dots, P_{\ell-1})$  and zero otherwise, and  $\delta_0(k)$  equals one if  $k = 0$  and zero otherwise.

We will use the methods of Babillot & Ledrappier [BL1],[BL2] to estimate this sum.

For every  $w \in \mathbb{T}$ ,  $u \in \mathbb{R}$  define the operator  $(L_{u+iw}\varphi)(\underline{x}) = \sum_{\sigma(\underline{y})=\underline{x}} e^{(u+iw)F(\underline{y})}\varphi(\underline{y})$  on  $\mathcal{L} := \{\varphi : \Sigma^+ \rightarrow \mathbb{C} : \|\varphi\| := \|\varphi\|_{\infty} + \text{Lip}(\varphi)\}$ . No matter which  $u$  we choose,

$$\begin{aligned} J_n(\underline{x}^{(i)}, \xi_i^*) &= h(P_0) \sum_{\sigma^n(\underline{y})=\underline{x}^{(i)}} 1_{[P]}(\underline{y}) \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{(u+iw)(F_{n^*}(\underline{y})-\xi_i^*)} dw \\ &= \frac{h(p_0)}{2\pi} \int_{-\pi}^{\pi} e^{-(u+iw)\xi_i^*} (L_{u+iw}^{n^*} 1_{[P]})(\underline{x}^{(i)}) dw. \end{aligned}$$

The parameter  $u$  does not affect the value of the integral, but a judicious choice  $u = u(\xi_i^*, n^*)$  will facilitate the analysis of the integrand.

$L_z : \mathcal{L} \rightarrow \mathcal{L}$  ( $z = u + iw$ ) has the following properties ([PP] chapter 4):

- (1)  $L_0$  has leading eigenvalue  $\lambda^{-1}$ , with eigenprojection  $P\varphi = h\nu(\varphi)$  where  $h$  is given by Lemma 3.3, and  $\nu$  satisfies  $h d\nu = m_0 \circ \pi_0^{-1}$  (cf. Theorem 2.8).
- (2) The eigenvalue  $\lambda^{-1}$  is simple and isolated. All other eigenvalues are strictly smaller in absolute value.
- (3) For all  $u$  real,  $L_u$  has spectral radius  $\exp p(u)$  where

$$p(u) = P_{top}(uF) := \sup\{h_{\mu}(\sigma) + u \int F d\mu : \mu \text{ is a } \sigma\text{-inv. prob. measure}\}.$$

- (4) For all  $u, w$  real,  $w \notin 2\pi\mathbb{Z}$ ,  $L_{u+iw}$  has spectral radius strictly smaller than  $\exp p(u)$ . This uses the Aperiodicity Lemma (Lemma 2.9).
- (5) There is  $\varepsilon_{pert} > 0$  such that for every  $|z| < \varepsilon_{pert}$ ,  $L_z = \lambda(z)[P(z) + N(z)]$  where  $\lambda(z) \in \mathbb{C}$ ,  $P(z)$  is a projection with one-dimensional image, and  $N(z)$  is an operator with spectral radius strictly less than one s.t.  $PN = NP = 0$ . The maps  $z \mapsto \lambda(z), P(z), N(z)$  are analytic.
- (6)  $p(z) := \log \lambda(z)$  is an analytic extension of  $p(u)$  to  $U = \{z : |z| < \varepsilon_{pert}\}$ . On  $U$ ,  $p(z) = -\log \lambda + \frac{1}{2}\sigma^2 z^2 + o(z^2)$ , where  $\sigma > 0$ . This also uses the Aperiodicity Lemma.

Part (6) implies that the image of  $p'(\cdot)$  is a neighborhood of zero. Suppose  $\frac{\xi_i^*}{n^*}$  belongs to this neighborhood, and choose  $u$  s.t.  $p'(u) = \frac{\xi_i^*}{n^*}$ . The closer  $\frac{\xi_i^*}{n^*}$  is to

zero, the closer  $u$  is to zero. Since, by construction,  $\xi_i^* = \xi(\psi^{n^*}(\omega)) + O(1)$ , there exists  $\varepsilon_0 > 0$  so small and  $n_0$  so large that for all  $n^* > n_0$

$$\left| \frac{\xi(\psi^{n^*}(\omega))}{n^*} \right| < \varepsilon_0 \implies |u| < \varepsilon_{pert}.$$

The condition will be satisfied for all  $n$  large enough, because of the assumption that  $\xi[\psi^k(\omega)]/k \xrightarrow{k \rightarrow \infty} 0$ . Henceforth we assume that  $\left| \frac{\xi(\psi^{n^*}(\omega))}{n^*} \right| < \varepsilon_0$  and take  $|u| < \varepsilon_{pert}$  s.t.  $p'(u) = \frac{\xi_i^*}{n^*}$ .

Let  $\rho(L_{u+iw})$  denote the spectral radius of  $L_{u+iw}$ . Since  $u + iw \mapsto L_{u+iw}$  is continuous,  $u + iw \mapsto \rho(u + iw)$  is upper semi-continuous. Therefore, by part (4), there exists  $0 < \kappa < 1$  s.t.  $\sup\{e^{-p(u)}\rho(L_{u+iw}) : \text{dist}(w, 2\pi\mathbb{Z}) > \varepsilon_{pert}\} < \kappa$ .

Similar reasoning gives (perhaps for a slightly larger  $0 < \kappa < 1$ )

$$\sup\{|e^{-p(u)}\rho(N(u+iw))| : |u+iw| \leq \varepsilon_{pert}\} < \kappa.$$

It is not difficult to see, using the spectral radius formula and the continuity of  $z \mapsto L_z$ , that  $\|L_{u+iw}^{n^*} 1_{[P]}\| = O(e^{n^*p(u)}\kappa^{n^*})$  uniformly on  $\{w \in (-\pi, \pi) : |w| \geq \varepsilon_{pert}\}$ , and  $\|N(u+iw)^{n^*} 1_{[P]}\| = O(e^{n^*p(u)}\kappa^{n^*})$  uniformly on  $(-\varepsilon_{pert}, \varepsilon_{pert})$ .

If we split the domain of integration in the integral which defines  $J_n(\underline{x}^{(i)}, \xi_i^*)$  into  $(-\varepsilon_{pert}, \varepsilon_{pert})$  and its complement and then substitute  $L = \lambda(P + N)$  into the first piece, then we get the following (where  $J_n = J_n(\underline{x}^{(i)}, \xi_i^*)$ ,  $\underline{x} = \underline{x}^{(i)}$ ,  $\xi^* = \xi_i^*$ ):

$$J_n = \frac{h(P_0)}{2\pi} \int_{-\varepsilon_{pert}}^{\varepsilon_{pert}} e^{-(u+iw)\xi^*} \left[ \lambda(u+iw)^{n^*} (P(u+iw)1_{[P]})(\underline{x}) \right] dw + O(e^{n^*p(u)-u\xi^*}\kappa^{n^*}).$$

The error bound can be simplified using the Legendre transform. Let  $H(v)$  denote minus the Legendre transform of  $p(u)$ , namely  $H(v) := p(u) - up'(u)$  for the  $u = u(v)$  s.t.  $p'(u) = v$ . By the choice of  $u$ ,  $n^*p(u) - u\xi^* = n^*H(\xi^*/n^*)$ , whence

$$J_n = \frac{h(P_0)}{2\pi} \int_{-\varepsilon_{pert}}^{\varepsilon_{pert}} e^{-(u+iw)\xi^* + n^*p(u+iw)} (P(u+iw)1_{[P]})(\underline{x}) dw + O(e^{n^*H(\frac{\xi^*}{n^*})}\kappa^{n^*}).$$

The next step is to use the Taylor expansion of  $p(z)$  at  $z = u$  to see that the exponential term in the integrand equals

$$e^{n^*[p(u)-u\frac{\xi^*}{n^*}]} \cdot e^{in^*w[p'(u)-\frac{\xi^*}{n^*}]} \cdot e^{n^*[-\frac{1}{2}p''(u)w^2 + O(w^3)]}.$$

The first term is  $\exp[n^*H(\frac{\xi^*}{n^*})]$ , and the second term is 1 by the choice of  $u$ . So

$$\begin{aligned} J_n &= \frac{e^{n^*H(\frac{\xi^*}{n^*})}h(P_0)}{2\pi} \left[ \int_{-\varepsilon_{pert}}^{\varepsilon_{pert}} e^{-n^*[-\frac{1}{2}p''(u)w^2 + O(w^3)]} (P(u+iw)1_{[P]})(\underline{x}) dw + O(\kappa^{n^*}) \right] \\ &= \frac{e^{n^*H(\frac{\xi^*}{n^*})}h(P_0)}{2\pi} \left[ \int_{-\varepsilon_{pert}\sqrt{n^*}}^{\varepsilon_{pert}\sqrt{n^*}} e^{-\frac{1}{2}p''(u)v^2 + O(\frac{v^3}{\sqrt{n^*}})} (P(u + \frac{iv}{\sqrt{n^*}})1_{[P]})(\underline{x}) \frac{dv}{\sqrt{n^*}} + O(\kappa^{n^*}) \right]. \end{aligned}$$



We discuss the asymptotic behavior of this expression as  $n^* \rightarrow \infty$ , subject to the assumption that  $\frac{1}{n^*} \xi[\psi^{n^*}(\omega)] \rightarrow 0$ . Since  $\xi^* \equiv \xi_i^* = \xi[\psi^{n^*}(\omega)] + O(1)$ ,

$$\frac{\xi^*}{n^*} \xrightarrow{n^* \rightarrow \infty} 0, \text{ and therefore } u \xrightarrow{n^* \rightarrow \infty} 0.$$

Since  $\|P(z) - P\| \xrightarrow{|z| \rightarrow 0} 0$  and  $P1_{[P]} = h\nu[P]$  is bounded away from zero,

$$(P(u + i\frac{v}{\sqrt{n^*}})1_{[P]})(\underline{x}) = [1 + o(1)]h(\underline{x})\nu[P] = [1 + o(1)]\ell[W^s(\underline{x})]\nu[P] \text{ unif. as } n^* \rightarrow \infty.$$

(But caution!  $\underline{x} = \underline{x}^{(i)}$  varies as  $n^* \rightarrow \infty$  so the term on the right side fluctuates.)

If  $\varepsilon_{pert}$  and  $|u|$  are small enough then  $|p''(u)| > \frac{1}{2}p''(0) = \frac{1}{2}\sigma^2$  and  $|O(w^3)| \leq \frac{1}{8}\sigma^2|w|^2$  for  $|w| < \varepsilon_{pert}$ . We see that the exponential term is bounded by  $\text{const} \cdot e^{-\frac{1}{8}v^2}$ . By the dominated convergence theorem,

$$\begin{aligned} J_n &= [1 + o(1)] \frac{e^{n^* H(\frac{\xi^*}{n^*})}}{2\pi} h(P_0)\nu[P]\ell[W^s(\underline{x}^i)] \left[ \frac{1}{\sqrt{n^*}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\sigma^2 v^2} dv + O(\kappa^{n^*}) \right] \\ &= [1 + o(1)] \frac{e^{n^* H(\frac{\xi^*}{n^*})}}{\sqrt{2\pi\sigma^2 n^*}} h(P_0)\nu[P]\ell[W^s(\underline{x}^i)] = [1 + o(1)] \frac{e^{n^* H(\frac{\xi^*}{n^*})}}{\sqrt{2\pi\sigma^2 n^*}} m(E)\ell[W^s(\underline{x}^i)]. \end{aligned}$$

Notice that  $h(P_0)\nu[P] = (m_0 \circ \pi_0^{-1})(E) = \frac{1}{2}m(E)$ , where  $m$  is the (non-normalized) area measure.

Next we analyze the exponent. Since  $H(\cdot)$  is minus the Legendre transform of  $p(\cdot)$  and  $p(z) = -\log \lambda + \frac{1}{2}\sigma^2 z^2 + o(z^2)$ ,  $H(v) = -\log \lambda - \frac{v^2}{2\sigma^2} + o(v^2)$ . In particular  $H'(0) = 0$  and  $H''(0) = -\frac{1}{\sigma^2}$ . Recalling that  $\xi^* = \xi[\psi^{n^*}(\omega)] + O(1)$  and expanding  $H(u)$  around  $u_0 = \frac{\xi[\psi^{n^*}(\omega)]}{n^*}$ , we obtain

$$\begin{aligned} n^* H\left(\frac{\xi^*}{n^*}\right) &= n^* \left[ H\left(\frac{\xi[\psi^{n^*}(\omega)]}{n^*}\right) + H'\left(\frac{\xi[\psi^{n^*}(\omega)]}{n^*}\right) \frac{\xi^* - \xi[\psi^{n^*}(\omega)]}{n^*} + o\left(\frac{\xi^* - \xi[\psi^{n^*}(\omega)]}{n^*}\right) \right] \\ &= n^* H\left(\frac{\xi[\psi^{n^*}(\omega)]}{n^*}\right) + n^* [H'(0) + o(1)] \frac{O(1)}{n^*} + n^* o\left(\frac{O(1)}{n^*}\right) \quad (\because \frac{\xi[\psi^{n^*}(\omega)]}{n^*} \rightarrow 0) \\ &= n^* H\left(\frac{\xi[\psi^{n^*}(\omega)]}{n^*}\right) + o(1) \quad (\because H'(0) = 0). \end{aligned}$$

Now we expand  $H$  around zero to see that

$$n^* H\left(\frac{\xi^*}{n^*}\right) = -n^* \log \lambda - \frac{1}{2\sigma^2} [1 + o(1)] \left( \frac{\xi[\psi^{n^*}(\omega)]}{\sqrt{n^*}} \right)^2 + o(1).$$

This and the definition of  $n^*$  give

$$\begin{aligned} J_n(\underline{x}^{(i)}, \xi_i^*) &= [1 + o(1)] \frac{\lambda^{-n^*} m(E)\ell[W^s(\underline{x}^i)]}{2\sqrt{\log_{\lambda^{-1}} n}} \times \\ &\quad \times \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{1}{2\sigma^2} [1 + o(1)] \left( \frac{\xi[\psi^{n^*}(\omega)]}{\sqrt{n^*}} \right)^2 \right]. \end{aligned}$$

By (3.2), the sum of these expressions over  $i = 1, \dots, n_1$  gives a lower bound for  $\int_0^n 1_E[\varphi_\theta^t(\omega)] dt$ , and the sum over  $0, \dots, n_1 + 1$  gives an upper bound. The only term which depends on  $i$  is  $\ell[W^s(\underline{x}^{(i)})]$ . Since by (3.1),

$$\ell[B_n(\omega)] - 2 \max h \leq \sum_{i=1}^{n_1} \ell[W^s(\underline{x}^{(i)})] \leq \sum_{i=0}^{n_1+1} \ell[W^s(\underline{x}^{(i)})] \leq \ell[B_n(\omega)] + 2 \max h,$$

and since both sides are  $\ell[B_n(\omega)][1 + O(\frac{1}{n_0})] = \lambda^{n^*} n [1 + O(\frac{1}{n_0})]$ , we have

$$\int_0^n 1_E[\varphi_\theta^t(\omega)] dt \leq \frac{n(1+o(1))(1+O(\frac{1}{n_0}))}{2\sqrt{\log_{\lambda^{-1}} n}} m(E) \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1+o(1)}{2\sigma^2} \left(\frac{\xi[\psi^{n^*}(\omega)]}{\sqrt{n^*}}\right)^2},$$

$$\int_0^n 1_E[\varphi_\theta^t(\omega)] dt \geq \frac{n(1+o(1))(1+O(\frac{1}{n_0}))}{2\sqrt{\log_{\lambda^{-1}} n}} m(E) \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1+o(1)}{2\sigma^2} \left(\frac{\xi[\psi^{n^*}(\omega)]}{\sqrt{n^*}}\right)^2}.$$

We now remember that  $n_0$  is a free parameter, and can be chosen arbitrarily large. The asymptotic expansion of the theorem follows in the case  $G = 1_E$ .

The case  $G \in C_c(S)$  is treated by decomposing  $G = G^+ - G^-$  with  $G^\pm \in C_c(S)$  non-negative, and approximating  $G^\pm$  from above and below by linear combinations of indicators of cylinders.  $\square$

**Proof of Theorem 3.1.** It is enough to prove the asymptotic statement for functions of the form  $G(x, k) = \gamma(x)1_{\mathbb{T} \times \{0\}}(x, k)$  with  $\gamma \in C(\mathbb{T})$  s.t.  $\int \gamma(t) dt > 0$ . The case of  $G \in C_c(\mathbb{T} \times \mathbb{Z})$  then follows by approximation.

The infinite staircase case be decomposed into an infinite collection of horizontal  $2 \times 1$  rectangles. Fix one of them, calling it ‘‘rectangle zero’’, and identify it with  $[0, 2] \times [0, 1]$ . Define  $\tilde{G}$  on rectangle zero by

$$\tilde{G}(x', y') = \pi \cos \theta \cdot \gamma\left(\frac{1}{2}(x' - y' \tan \theta)\right) \cdot \sin(\pi y').$$

This is designed so that  $\int_0^{1/\cos \theta} (\tilde{G} \circ \varphi_\theta^t)(\omega(x)) dt = G(x, 0)$ . The upper limit  $1/\cos \theta$  is the time it takes  $\varphi_\theta^t(\omega(x))$  to reach the upper side of  $[0, 2] \times [0, 1]$ .

Extend  $\tilde{G}$  to the rest of the infinite staircase surface by setting it equal to zero outside rectangle zero. Since  $\tilde{G}(x' + 2, y') = \tilde{G}(x', y')$  and  $\tilde{G}(*, 0) = \tilde{G}(*, 1) = 0$ , this is a continuous function. A calculation shows that  $\int \tilde{G} dm = 2 \cos \theta \int_{\mathbb{T} \times \mathbb{Z}} G dm_0$ , where  $m$  is the (non-normalized) area measure on  $\text{St}$  and  $m_0 = m_{\mathbb{T}} \times m_{\mathbb{Z}}$ .

The orbit  $\{\varphi_\theta^t(\omega(x)) : 0 < t < n/\cos \theta\}$  can be split into segments of length  $1/\cos \theta$  which go across horizontal rectangles. The  $j$ -th segment enters the bottom side of rectangle  $\sum_{i=0}^{j-1} f(x + i\alpha)$  at distance  $2x + 2j\alpha \pmod 2$  from the left endpoint. Only the segments s.t.  $\sum_{i=0}^{j-1} f(x + i\alpha) = 0$  contribute to  $\int_0^{n/\cos \theta} \tilde{G}[\varphi_\theta^t(\omega(x))] dt$ . The contribution is  $G(x + j\alpha, 0) = (G \circ T_\alpha^j)(x, 0)$ .

It follows that  $\int_0^{n/\cos \theta} \tilde{G}(\varphi_\theta^t(\omega(x))) dt = \sum_{j=0}^{n-1} G(x + j\alpha, \sum_{i=1}^{j-1} f(x + i\alpha)) = \sum_{j=0}^{n-1} (G \circ T^j)(x, 0)$ . The theorem now follows from Theorem 3.2.  $\square$

#### 4. STOCHASTIC PROPERTIES OF BIRKHOFF SUMS

Theorem 3.1 expresses the Birkhoff sums of the cylinder map  $T_\alpha$  asymptotically in terms of  $\frac{1}{\sqrt{k}}(\Xi_k(x))$  where  $\Xi_k(x) := \xi[\psi^k(\omega(x))]$ ,  $\psi$  is a renormalizing automorphism of  $\alpha$  with zero drift,  $\xi$  is its associated  $\mathbb{Z}$ -coordinate, and  $\omega : \mathbb{T} \rightarrow \text{St}$  is the map which associates to  $x \in \mathbb{T}$  the point on the top side of a (fixed) horizontal rectangle at distance  $2x$  from its left endpoint.

To determine the stochastic properties of the Birkhoff sums of the cylinder map, it is sufficient to understand the stochastic process  $\{\Xi_k(x)\}_{k \geq 1}$ , when  $x \sim U[0, 1]$ . In this section we prove the following.

**Theorem 4.1.** *Choose  $x \in [0, 1]$  uniformly, then*

$$(1) \quad \Xi_k/k \xrightarrow[k \rightarrow \infty]{} 0 \text{ a.e.}$$

- (2)  $\forall \varepsilon > 0 \exists I(\varepsilon) > 0$  s.t.  $\mathbb{P}[|\Xi_k/k| > \varepsilon] = O(e^{-kI(\varepsilon)})$  ( $k \rightarrow \infty$ ).
- (3)  $\Xi_k/\sqrt{k} \xrightarrow[k \rightarrow \infty]{\text{dist}} N(0, \sigma^2)$ . Moreover, there is a probability space  $(\Omega, \mathcal{F}, \mu)$  equipped with two continuous time stochastic processes  $\tilde{\Xi}_t, \tilde{B}_t : \Omega \rightarrow \mathbb{R}$  s.t.  $\{\tilde{\Xi}_n\}_{n \geq 1} \stackrel{\text{dist}}{=} \{\Xi_n\}_{n \geq 1}$ ,  $\{\tilde{B}_t\}_{t \geq 0} \stackrel{\text{dist}}{=} \text{standard Brownian motion}$ , and for some  $0 < \delta < \frac{1}{2}$ ,  $|\tilde{\Xi}_t - \sigma \tilde{B}_t| = o(t^\delta)$  a.s. as  $t \rightarrow \infty$ .
- (4) If  $f, \hat{f} \in L^1(\mathbb{R})$ , then  $\lim_{n \rightarrow \infty} \frac{1}{\ln n} \sum_{k=1}^n \frac{1}{k} f(\Xi_k/\sqrt{k}) = \mathbb{E}[f(N)]$  almost surely, where  $N$  is the standard gaussian, and  $\hat{f}$  is the Fourier transform of  $f$ .

**Lemma 4.2.** *There are a stationary mixing Markov chain  $\{X_i\}_{i=1}^\infty$  with finite set of states  $S$ ,  $g : S \times S \rightarrow \mathbb{R}$  s.t.  $\mathbb{E}[g(X_0, X_1)] = 0$ , and a uniformly bounded sequence of random variables  $\varepsilon_k$  s.t.  $\Xi_k \stackrel{\text{dist}}{=} g(X_0, X_1) + \dots + g(X_{k-1}, X_k) + \varepsilon_k$  (equality of stochastic processes). The function  $g$  is not of the form  $H(x) - H(y) + \text{const}$  for any Borel  $H : S \rightarrow \mathbb{R}$ .*

*Proof.* Define  $\xi_k : \text{St}_0 \rightarrow \mathbb{Z}$  as follows: given  $p \in \text{St}_0$ ,

$$\xi_k(p) := \xi[\psi^k(\tilde{p})] - \xi(\tilde{p}) \text{ for some (all) } \tilde{p} \in \pi^{-1}(p).$$

This can be easily seen to be independent of the choice of  $\tilde{p}$ .

Next define  $x : \text{St}_0 \rightarrow [0, 1]$  as follows: given  $p \in \text{St}_0$ , lift  $p$  to a point  $\tilde{p} \in \text{St}$  in rectangle  $\#0$ , and project  $\tilde{p}$  to the top side of this rectangle in the stable direction. The result has the form  $\omega(x)$  for some unique  $x = x(p) \in [0, 1]$ .

*Claim.* If  $p$  is chosen uniformly in  $\text{St}_0$ , then  $x(p)$  is distributed uniformly in  $[0, 1]$ , and  $\varepsilon_k(p) := \xi_k(p) - \xi[\psi^k(\omega(x(p)))]$  are uniformly bounded on  $\text{St}_0$ .

The first statement is because rectangle zero can be identified with the parallelogram with a horizontal side of length 2 and a side in the stable direction. The second statement is because  $\tilde{p} - \omega(x) \propto \underline{w}$  where  $\underline{w}$  is in the stable direction of the derivative of  $\psi$ , so  $\text{dist}(\psi^k(\tilde{p}), \psi^k[\omega(x)]) \leq \lambda^k \sqrt{1 + \tan^2 \theta} \leq 1/\cos \theta$ .

It follows that  $\xi[\psi^k(\omega(x))] \stackrel{\text{dist}}{=} \xi_k + \varepsilon_k$ , where  $|\varepsilon_k| \leq 1/\cos \theta$  and  $\xi_k$  is the stochastic process

$$\xi_k(p) := \xi[\psi^k(p)], \text{ where } p \text{ is distributed uniformly in } \text{St}_0,$$

We will use the Adler–Weiss Theorem to represent  $\xi_k$  as a random walk driven by a Markov chain.

Let  $\mathfrak{P}$  denote the Adler–Weiss Markov partition, and  $\mathcal{G}$  the dynamical graph of  $\mathfrak{P}$ , see §2. Let  $\pi_0 : \Sigma(\mathcal{G}) \rightarrow \text{St}_0$  denote the symbolic coding of the projected automorphism  $\psi_0$ , given by Theorem 2.8, then  $\mathbb{P} := m_0 \circ \pi_0^{-1}$  is a Markov measure. It follows that  $X_k : \Sigma(\mathcal{G}) \rightarrow \mathfrak{P}$ ,  $X_k[\{P_i\}_{i \in \mathbb{Z}}] = P_i$  is a finite state Markov chain. Since hyperbolic toral automorphisms are mixing area preserving maps,  $\mathbb{P}$  is a mixing shift invariant measure. So  $\{X_k\}_{k \in \mathbb{Z}}$  is stationary and mixing.

By the definition of the Frobenius function,

$$\begin{aligned} \xi_k(p) &= \xi[\psi^k(\tilde{p})] - \xi(\tilde{p}) \text{ for some (all) } \tilde{p} \in \pi^{-1}(p) \\ &= \sum_{j=0}^{k-1} \xi[\psi^{j+1}(\tilde{p})] - \xi[\psi^j(\tilde{p})] = \sum_{j=0}^{k-1} \xi[\psi(\tilde{p}_j)] - \xi[\tilde{p}_j], \text{ where } \tilde{p}_j \in \pi^{-1}[\psi_0^j(p)]. \end{aligned}$$

So  $\xi_k(p) = \sum_{j=0}^{k-1} (F_\psi \circ \psi_0^j)(p)$ .

At the end of §2, we saw that  $F := F_\psi \circ \pi_0$  can be expressed in the form  $g(X_0, X_1)$  or  $g(X_{-1}, X_0)$  for some function  $g : \mathfrak{X} \times \mathfrak{X} \rightarrow \mathbb{Z}$ . Since  $m_0 \circ \pi_0^{-1} = \mathbb{P}$ ,

$$\sum_{j=0}^{k-1} F_\psi \circ \psi_0^j \stackrel{\text{dist}}{=} \sum_{j=0}^{k-1} F \circ \sigma^j = \sum_{j=0}^{k-1} g(X_j, X_{j-1}) \text{ or } \sum_{j=0}^{k-1} g(X_{j-1}, X_j).$$

Since  $\{X_j\}_{j \in \mathbb{Z}}$  is stationary,  $\xi_k(p) \stackrel{\text{dist}}{=} g(X_0, X_1) + \cdots + g(X_{k-1}, X_k)$  as required.

$\mathbb{E}[g(X_0, X_1)] = \int F_\psi dm_0 = 0$ , because  $\psi$  has zero drift. There is no function  $H : \mathfrak{X} \rightarrow \mathbb{R}$  s.t.  $g(X_0, X_1) = H(X_0) - H(X_1) + \text{const}$ , because of Lemma 2.9.  $\square$

**Proof of Theorem 4.1.** Let  $S_k g := g(X_0, X_1) + \cdots + g(X_{k-1}, X_k)$ .

- (1) By the ergodic theorem,  $S_k g / k \xrightarrow[k \rightarrow \infty]{} \mathbb{E}[g(X_0, X_1)] = 0$  a.s.
- (2) By the Gärtner–Ellis Theorem,  $\mathbb{P}[|S_k g / k| > \varepsilon] = O(e^{-kI(\varepsilon)})$  as  $k \rightarrow \infty$  where  $I(\cdot)$  is the Legendre transform of  $\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[\exp(u S_n g)] = p(u) =$  topological pressure of  $uF$ . Since  $g \neq H(X_0) - H(X_1) + \text{const}$ ,  $p(t)$  is analytic and strictly convex. So  $I(\varepsilon)$  is strictly convex. Since  $p'(0) = \mathbb{E}(g) = 0$ ,  $I(\varepsilon) > 0$  for all  $\varepsilon > 0$ . See §6 for a calculation of  $p(u)$  in a special case.
- (3) By the central limit theorem for finite state Markov chains,  $\frac{1}{\sqrt{k}} S_k g \xrightarrow[k \rightarrow \infty]{\text{dist}} N(0, \sigma_0^2)$  for  $\sigma_0^2 := \lim_{n \rightarrow \infty} \frac{1}{n} \text{Var}[S_n g]$ . Since  $g \neq H(X_0) - H(X_1) + \text{const}$ ,  $\sigma_0 \neq 0$  (Leonov’s Theorem). By Philipp & Stout’s Almost Sure Invariance Principle ([PS], chapter 4), there is a probability space  $(\Omega, \mathcal{F}, \mu)$  equipped with two continuous time stochastic processes  $\tilde{\Xi}_t, \tilde{B}_t : \Omega \rightarrow \mathbb{R}$  s.t.  $\{\tilde{\Xi}_n\}_{n \geq 1} \stackrel{\text{dist}}{=} \{S_n g + \varepsilon_n\}_{n \geq 1}$ ,  $\{\tilde{B}_t\}_{t \geq 0} \stackrel{\text{dist}}{=} \text{standard Brownian motion}$ , such that for some  $0 < \delta < \frac{1}{2}$ ,  $|\tilde{\Xi}_t - \sigma_0 \tilde{B}_t| = o(t^\delta)$  a.s. as  $t \rightarrow \infty$ .

By Theorem 4.13 in [PP],  $\sigma_0^2 = p''(0)$ . It follows that  $\sigma_0 = \sigma$  where  $\sigma$  is the constant appearing in Theorems 3.1 and 3.2.

- (4) If  $f, \hat{f} \in L^1(\mathbb{R})$ , then  $\lim_{n \rightarrow \infty} \frac{1}{\ln n} \sum_{k=1}^n \frac{1}{k} f(S_k g / \sqrt{k}) = \mathbb{E}[f(N)]$  almost surely, where  $N$  is the standard gaussian, and  $\hat{f}$  is the Fourier transform of  $f$ . This follows from (3) as in Lemma 2 in [LS1] (see also [Fi1]).

The theorem follows, since  $\{\Xi_k\}_{k \geq 1} \stackrel{\text{dist}}{=} \{S_k g + \varepsilon_k\}_{k \geq 1}$  with  $\varepsilon_k = O(1)$ .  $\square$

**Application to the Cylinder map.** Theorems 3.1 and 4.1 combine to give the following statement. Let  $\chi$  be a standard gaussian random variable.

**Theorem 4.3.** *Suppose  $\alpha$  is a quadratic irrational. There are  $\sigma^2 > 0$  and  $0 < \lambda < 1$  s.t. if  $a_n := \sqrt{\frac{|\ln \lambda|}{4\pi\sigma^2}} \left(\frac{n}{\sqrt{\ln n}}\right)$ , then for every  $G \in L^1(\mathbb{T} \times \mathbb{Z})$  s.t.  $\int G dm_0 = 1$ ,*

- (1)  $\frac{1}{a_n} \sum_{k=0}^{n-1} G \circ T_\alpha^k \xrightarrow[n \rightarrow \infty]{\text{dist}} \sqrt{2} \exp(-\frac{1}{2}\chi^2)$
- (2)  $\lim_{N \rightarrow \infty} \frac{1}{\ln N} \sum_{n=2}^N \frac{1}{n \ln n} \left(\frac{1}{a_n} \sum_{k=1}^n G \circ T_\alpha^k\right) = 1$  a.s.
- (3) *if in addition  $G \in C_c(\mathbb{T} \times \mathbb{Z})$ , then  $\int_{\mathbb{T} \times \{0\}} \left(\sum_{j=0}^{n-1} G \circ T_\alpha^j\right) dm = [1 + o(1)]a_n$ .*

Part 2 is a “higher order ergodic theorem” in the sense of A. Fisher [Fi1],[Fi2],[ADF].

*Proof.* (1) and (2) are immediate.

For (3) let  $A_\delta(n) := \{(x, 0) : |\Xi_n(x)/n| \leq \delta\}$ ,  $B_\delta(n) := \{(x, 0) : |\Xi_n(x)/n| > \delta\}$ . We break the integral into the main part  $\int_{A_\delta(n)}$  and the remainder  $\int_{B_\delta(n)}$ . The remainder is  $O(ne^{-nI(\delta)}) = o(a_n)$ , because of Theorem 4.1(2) and the boundedness of  $G$ . The main term is sandwiched between two bounds of the form

$$(1 + \varepsilon(\delta))a_n \cdot \sqrt{2}\mathbb{E}[\exp(-\frac{1-\varepsilon(\delta)}{2\sigma^2}(\Xi_{\lceil \log^* n \rceil}/\lceil \log^* n \rceil)^2)]$$

$$(1 - \varepsilon(\delta))a_n \cdot \sqrt{2}\mathbb{E}[\exp(-\frac{1+\varepsilon(\delta)}{2\sigma^2}(\Xi_{\lceil \log^* n \rceil}/\lceil \log^* n \rceil)^2)]$$

with  $\varepsilon(\delta) \xrightarrow{\delta \rightarrow 0^+} 0$  (this is a consequence of the uniformity in  $x$  in Theorem 3.1).

Since  $\Xi_k/\sqrt{k} \xrightarrow[k \rightarrow \infty]{\text{dist}} N(0, \sigma^2)$ , these bounds converge to  $(1 \pm \varepsilon(\delta))\sqrt{2}\mathbb{E}[e^{-\frac{1 \mp \varepsilon(\delta)}{2}\chi^2}]$  as  $n \rightarrow \infty$ . Since  $\mathbb{E}[e^{-\frac{1 \mp \varepsilon(\delta)}{2}\chi^2}] \xrightarrow{\delta \rightarrow 0} \mathbb{E}(e^{-\frac{1}{2}\chi^2}) = 2^{-\frac{1}{2}}$ , the main term is  $[1 + o(1)]a_n$ . Part (3) follows.  $\square$

### Application to the deterministic random walk.

**Theorem 4.4.** *Suppose  $\alpha$  is a quadratic irrational, and  $N_n$  is the number of visits of the DRW to zero up to (and not including) time  $n - 1$ , then*

- (1)  $\mathbb{E}(N_n) = [1 + o(1)]a_n$ , where  $a_n = \sqrt{\frac{|\ln \lambda|}{4\pi\sigma^2}}(\frac{n}{\sqrt{\ln n}})$ .
- (2)  $\frac{1}{a_n}N_n \xrightarrow[n \rightarrow \infty]{\text{dist}} \sqrt{2}\exp(-\frac{1}{2}\chi^2)$ , where  $\chi$  is a standard gaussian.
- (3)  $\lim_{N \rightarrow \infty} \frac{1}{\ln \ln N} \sum_{n=2}^N \frac{1}{n \ln n} (\frac{1}{a_n}N_n) = 1$  a.s.
- (4)  $\lambda$  is an eigenvalue of the renormalizing automorphism  $\psi$ , and  $\sigma^2$  is the asymptotic variance in  $\frac{1}{\sqrt{k}}\Xi_k \xrightarrow[k \rightarrow \infty]{\text{dist}} N(0, \sigma^2)$ .

This follows from the previous theorem and the identity  $N_n = \sum_{k=0}^{n-1} 1_{\mathbb{T} \times \{0\}} \circ T_\alpha^k$ .

**Stochastic interpretation of twists.** Theorem 4.1 and Lemma 4.2 extend trivially to automorphisms  $\psi$  with non-zero drift. One just needs to replace  $\Xi_k$  by  $\Xi_k - k\delta(\psi)$  where  $\delta(\psi)$  is the drift of  $\psi$ . The Markov chain and the function  $g$  Lemma 4.2 are defined as before, except that now  $\mathbb{E}(g) = \delta(\psi) \neq 0$ .

We can use this simple observation to calculate twists. Suppose  $\psi$  is a hyperbolic homogeneous automorphism with positive eigenvalues, and let  $\underline{w}$  be an eigenvector of its derivative. Recall from Lemma 2.12 that there is a unique homogeneous automorphism  $\phi$  with the same derivative as  $\psi$ , and which fixes the rays  $L_i(p, \underline{w})$ . The drift of  $\phi$  equals minus  $\tau_\psi(p, \underline{w})$ . Consequently,

**Corollary 4.5.** *Let  $\widehat{\Xi}_k := \xi[\phi^k(z)]$ , where  $z$  is distributed uniformly in horizontal rectangle zero, then  $\frac{1}{n}\widehat{\Xi}_n \rightarrow -\tau_\phi(p, \underline{w})$  a.s., and  $\frac{1}{\sqrt{n}}(\widehat{\Xi}_n + n\tau_\phi(p, \underline{w})) \xrightarrow[n \rightarrow \infty]{\text{dist}} N(0, \sigma^2)$ .*

## 5. APPLICATION TO A RESULT OF J. BECK

In this section we explain how to use the machinery developed in sections 2 and 4 to prove the following theorem of J. Beck [B1, B2]. Fix an irrational  $\alpha$  and let

$$Z_\alpha^*(n) := \#\{1 \leq k \leq n : \{k\alpha\} \in [0, \frac{1}{2}]\} - \frac{1}{2}n \equiv -\frac{1}{2} \sum_{k=1}^n f(\{k\alpha\}).$$

**Theorem 5.1** (Beck). *If  $\alpha$  is a quadratic irrational, then there are (explicit) constants  $C_1, C_2$  depending on  $\alpha$  s.t. for all  $a < b$  real*

$$\frac{1}{N} \#\{1 \leq n \leq N : \frac{Z_\alpha^*(n) - C_1 \ln N}{C_2 \sqrt{\ln N}} \in [a, b]\} \xrightarrow{N \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-u^2/2} du.$$

This is a reformulation of Theorem 1.1 in [B1] in the special case of the interval  $[0, \frac{1}{2})$ , but without the estimate of the rate of convergence provided there. The constants are calculated in [B2] using algebraic number theory and harmonic analysis.

First we explain how to translate Beck's theorem into a statement on linear flows on the infinite staircase.

In what follows  $\xi$  denotes a  $\mathbb{Z}$ -coordinate induced by the natural partition of the infinite staircase into horizontal rectangles, and  $p_0$  denotes the singularity at the bottom left corner of rectangle zero.

We wish to define  $\varphi_\theta^t(p_0)$  for  $t > 0$  for an irrational direction  $\theta$ . There is an element of choice here, because  $p_0$  is a singularity, and there are infinitely many rays in direction  $\theta$  emanating from  $p_0$  — one for each horizontal cylinder  $C$  s.t. that the vector  $(\frac{\sin \theta}{\cos \theta})$  based in  $p$  points inside  $R$ .

We define  $\varphi_\theta^t(p_0)$  to be the movement in unit speed along the ray  $L_0(p_0, (\frac{\sin \theta}{\cos \theta}))$  emanating from  $p_0$  in direction  $\theta$  which begins at rectangle zero.

**Lemma 5.2.** *Let  $\theta = \tan^{-1}(2\alpha - 1)$  and  $c := \sqrt{1 + \tan^2 \theta}$ , then*

$$Z_\alpha^*(n) = \frac{1}{2} \xi(\varphi_\theta^t(p_0)) \text{ for all } cn < t < c(n+1), \quad (5.1)$$

*Proof.* The constant  $c$  is exactly the time it takes  $\varphi_\theta$  to cross a horizontal cylinder in the vertical direction, so  $\varphi_\theta^{cn}(p_0)$  lies on the bottom horizontal side of a unique horizontal rectangle  $R_n$ , and  $\xi[\varphi_\theta^{cn}(p_0)] = \text{const}$  for  $cn < t < c(n+1)$ .

Let  $\xi_n$  denote the  $\mathbb{Z}$ -coordinate of  $R_n$ , and let  $x_n$  denote the distance of  $\varphi_\theta^{cn}(p_0)$  from the bottom left corner of  $R_n$ . We show by induction that  $x_n = 2n\alpha \pmod 2$  and  $\xi[\varphi_\theta^t(p_0)] = 2Z_n^*(n)$  for  $cn < t < c(n+1)$ .

At time zero, the flow is at  $p_0$ , so  $x_0 = 0$ , and by the definition of  $\varphi_\theta^t(p_0)$ ,  $\xi[\varphi_\theta^t(p_0)] = 0 = 2Z_\alpha^*(0)$  for all  $0 < t < c$ .

Suppose by induction that  $x_n = 2n\alpha \pmod 2$  and  $\xi[\varphi_\theta^t(p)] = 2Z_\alpha^*(n)$  for  $cn < t < c(n+1)$ . By the definition of St, the following implications hold:

- If  $x_n + \tan \theta \in [0, 1) + 2\mathbb{Z}$ , then  $\xi_{n+1} = \xi_n - 1$  and  $x_{n+1} = x_n + \tan \theta + 1 \pmod 2$ .
- If  $x_n + \tan \theta \in [1, 2) + 2\mathbb{Z}$ , then  $\xi_{n+1} = \xi_n + 1$  and  $x_{n+1} = x_n + \tan \theta - 1 \pmod 2$ .

We see that  $x_{n+1} = x_n + 2\alpha \pmod 2 = 2(n+1)\alpha \pmod 2$ , and

$$\begin{aligned} \xi_{n+1} &= \xi_n + 1_{[1,2)+2\mathbb{Z}}(x_n + \tan \theta) - 1_{[0,1)+2\mathbb{Z}}(x_n + \tan \theta) \\ &= \xi_n + 1_{[0,1)+2\mathbb{Z}}(x_n + 2\alpha) - 1_{[1,2)+2\mathbb{Z}}(x_n + 2\alpha) \\ &= 2Z_n^*(n) + 1_{[0,\frac{1}{2})}(\{(n+1)\alpha\}) - 1_{[\frac{1}{2},1)}(\{(n+1)\alpha\}) \\ &= 2 \left( Z_n^*(n) + 1_{[0,\frac{1}{2})}(\{(n+1)\alpha\}) - \frac{1}{2} \right) = 2Z_{n+1}^*(n+1). \quad \square \end{aligned}$$

**Proof of Beck's Theorem.** Since  $\alpha$  is a quadratic irrational, there are hyperbolic homogeneous automorphisms which renormalize  $\alpha$ .

Let  $\psi$  denote a hyperbolic homogeneous automorphism which renormalizes  $\alpha$ , has zero drift, and such that the eigenvalues of  $d\psi$  are positive. Let  $0 < \lambda < 1$

denote the contracting eigenvalue, and let  $\underline{w}$  denote a contracted eigenvector in direction  $\theta = \tan^{-1}(2\alpha - 1)$ . Let

$$C_1 := \frac{\tau_\psi(p_0, \underline{w})}{2|\ln \lambda|} \text{ and } C_2 := \frac{\sigma}{2\sqrt{|\ln \lambda|}} \quad (5.2)$$

where  $\tau_\psi(p_0, \underline{w})$  is the twist (Definition 2.10), and  $\sigma^2$  is the asymptotic variance mentioned in the previous sections. We will show that

$$D_N(a, b) := \frac{1}{N} \#\{1 \leq n \leq N : \frac{Z_\alpha^*(n) - C_1 \ln N}{C_2 \sqrt{\ln N}} \in [a, b]\} \xrightarrow{N \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-u^2/2} du.$$

Let  $\Gamma_N := \{\varphi_\theta^t(p_0) : c < t < c(N+1)\}$ , and  $\ell_{\Gamma_N}(\cdot)$  denote the length (Lebesgue) measure on  $\Gamma_N$ . By (5.1),

$$D_N(a, b) = \frac{1}{\ell_{\Gamma_N}(\Gamma_N)} \ell\{q \in \Gamma_N : \frac{\xi(q) - 2C_1 \ln N}{2C_2 \sqrt{\ln N}} \in [a, b]\} \quad (5.3)$$

Let  $N^* := \lfloor \log_{\lambda^{-1}} N \rfloor$ , and  $\gamma_N := \psi^{N^*}(\Gamma_N)$ . Since  $\psi \circ \varphi_\theta^t = \varphi_\theta^{\lambda t} \circ \psi$ ,  $\gamma_N$  is linear segment with bounded length in direction  $\theta$ .

By the definition of the twist,  $\psi^{N^*}(\Gamma_N) \subset D^{k_{N^*}}[L_0(p_0, \underline{w})]$ , where  $k_{N^*} = N^* \tau_\psi(p_0, \underline{w}) + O(1) = 2C_1 \ln N + O(1)$ . It follows that

$$\frac{1}{2} \xi(\cdot) = C_1 \ln N + O(1) \text{ uniformly on } \gamma_N.$$

By (5.3) and the identity  $\ell_{\Gamma_N} = \lambda^{-N^*} \ell_{\gamma_N} \circ \psi^{N^*}|_{\Gamma_N}$  where  $\ell_{\gamma_N}$  = Lebesgue measure on  $\gamma_N$ ,  $D_N(a, b) = \frac{1}{\lambda^{N^*} \ell_{\Gamma_N}(\Gamma_N)} \ell_{\gamma_N} \{\psi^{N^*}(q) : q \in \Gamma_N, \frac{\xi(q) - 2C_1 \ln N}{2C_2 \sqrt{\ln N}} \in [a, b]\}$ . From now on we set  $\ell := \ell_{\gamma_N}$ , then setting  $z = \psi^{-N^*}(q)$  we get

$$\begin{aligned} D_N(a, b) &= \frac{1}{\ell(\gamma_N)} \ell\{z \in \gamma_N : \frac{\xi(\psi^{-N^*}(z)) - 2C_1 \ln N}{2C_2 \sqrt{\ln N}} \in [a, b]\} \\ &= \frac{1}{\ell(\gamma_N)} \ell\{z \in \gamma_N : \frac{\xi(\psi^{-N^*}(z)) - \xi(z) + O(1)}{\sigma \sqrt{N^*} + O(1)} \in [a, b]\} \\ &= \frac{1}{\ell(\gamma_N)} \ell\{z \in \gamma_N : \frac{\xi(\psi^{-N^*}(z)) - \xi(z)}{\sigma \sqrt{N^*}} \in [a + O(\frac{1}{\sqrt{N^*}}), b + O(\frac{1}{\sqrt{N^*}})]\} \\ &= \frac{1}{\ell(\gamma_N)} \ell_{\widehat{\gamma}_N} \{z \in \widehat{\gamma}_N : \frac{\xi(\psi^{-N^*}(z))}{\sigma \sqrt{N^*}} \in [a + O(\frac{1}{\sqrt{N^*}}), b + O(\frac{1}{\sqrt{N^*}})]\}, \end{aligned}$$

where  $\widehat{\gamma}_N := D^{-k_{N^*}}(\gamma_N)$ , and  $\ell_{\widehat{\gamma}_N}$  is the Lebesgue measure on  $\widehat{\gamma}_N$ .

The advantage in passing to  $\widehat{\gamma}_N$ , apart from canceling the  $\xi(z)$  term up to bounded error, is that the family  $\{\widehat{\gamma}_N\}_{N \geq 1}$  is precompact. This is because the beginning point of  $\widehat{\gamma}_N$  is at distance  $c\lambda^{-N^*}$  from  $p_0$  on  $L_0(p_0, \underline{w})$ , and  $\ell(\widehat{\gamma}_N)$  is bounded away from zero and infinity. It follows that every sequence has a subsequence  $N_k \uparrow \infty$  along which  $\widehat{\gamma}_{N_k} \xrightarrow{k \rightarrow \infty} \widehat{\gamma}$ , where  $\widehat{\gamma}$  is a bounded linear segment in direction  $\theta$ , emanating from  $p_0$ , and beginning in rectangle zero. It is enough to prove that  $D_{N_k}(a, b) \rightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-u^2/2} du$  along such sequences.

Suppose  $N_k \uparrow \infty$  and  $\widehat{\gamma}_{N_k} \rightarrow \widehat{\gamma}$  as above. Let  $c_0 := \text{length of } \widehat{\gamma}$ . Fix  $\varepsilon$  much smaller than  $c_0$ , so small that  $\frac{c_0 \pm \varepsilon}{c_0 - \varepsilon} \in [e^{-\varepsilon}, e^\varepsilon]$ . Let  $\widehat{\gamma}^-$  and  $\widehat{\gamma}^+$  denote two linear segments in rectangle zero, in direction  $\theta$ , emanating from  $p_0$ , and with lengths  $c_0(1 - \varepsilon)$  and

$c_0(1 + \varepsilon)$  respectively. Then  $\hat{\gamma}^- \subset \hat{\gamma} \subset \hat{\gamma}^+$ , and  $D_{N_k}(a, b)$  is sandwiched between  $D_{N_k}^+(a, b), D_{N_k}^-(a, b)$ , where

$$D_{N_k}^+(a, b) := \frac{e^\varepsilon}{\ell(\hat{\gamma}^+)} \ell\{z \in \hat{\gamma}^+ : \frac{\xi(\psi^{-N^*}(z))}{\sigma\sqrt{N^*}} \in [a + O(\frac{1}{\sqrt{N^*}}), b + O(\frac{1}{\sqrt{N^*}})]\}$$

$$D_{N_k}^-(a, b) := \frac{1}{e^\varepsilon \ell(\hat{\gamma}^-)} \ell\{z \in \hat{\gamma}^- : \frac{\xi(\psi^{-N^*}(z))}{\sigma\sqrt{N^*}} \in [a + O(\frac{1}{\sqrt{N^*}}), b + O(\frac{1}{\sqrt{N^*}})]\},$$

The linear segments  $\hat{\gamma}^\pm$  are in the unstable (expanding) direction of  $\psi^{-1}$ . Let  $Q^\pm$  denote a thickening of these segments in the stable direction (the inside of a parallelogram with one side equal to  $\hat{\gamma}^\pm$  and the other side a segment in the stable (contracting) direction of  $\psi^{-1}$ ). For the same reasons explained in the proof of Lemma 4.2,

$$D_N^+(a, b) = \frac{e^\varepsilon}{m(Q^+)} m\{z \in Q^+ : \frac{\xi(\psi^{-N^*}(z))}{\sigma\sqrt{N^*}} \in [a + O(\frac{1}{N^*}), b + O(\frac{1}{N^*})]\} + o(1)$$

$$D_N^-(a, b) := \frac{e^{-\varepsilon}}{m(Q^-)} m\{z \in Q^- : \frac{\xi(\psi^{-N^*}(z))}{\sigma\sqrt{N^*}} \in [a + O(\frac{1}{N^*}), b + O(\frac{1}{N^*})]\} + o(1),$$

where  $m$  is the area measure.

We saw in the previous section that if  $z$  is chosen uniformly in rectangle number zero, then  $\frac{\xi(\psi^{-N^*}(z))}{\sigma\sqrt{N^*}} \xrightarrow[N^* \rightarrow \infty]{dist} N(0, 1)$ , because of the central limit theorem for finite state mixing Markov chains. The same is true for obvious reasons when  $z$  is sampled uniformly in a finite union of such rectangles. In  $D_{N_k}^\pm$  we are sampling  $z$  from a finite union of rectangles with respect to an absolutely continuous measure (Lebesgue times the density function  $\frac{1}{m(Q^\pm)} 1_{Q^\pm}$ ). By Eagleson's Theorem [E], the central limit theorem still holds, whence  $D_N^\pm(a, b) \xrightarrow[N \rightarrow \infty]{} \frac{e^{\pm\varepsilon}}{\sqrt{2\pi}} \int_a^b e^{-u^2/2} du$ . It follows that  $D_N(a, b) \xrightarrow[N \rightarrow \infty]{} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-u^2/2} du$ .  $\square$

It follows from the proof that  $C_1$  and  $C_2$  in Beck's theorem are given by (5.2). For example, if  $\alpha = \sqrt{2}$  then the calculations done in the next section give for a suitable automorphism  $\lambda = 17 - 12\sqrt{2} = (1 + \sqrt{2})^{-4}$ ,  $\tau_\psi(p_0, \underline{w}) = 1$ , and  $\sigma^2 = \frac{3}{8}\sqrt{2}$ . Thus  $C_1 = \frac{1}{8 \ln(1 + \sqrt{2})}$  and  $C_2 = \frac{1}{8} (\frac{3}{\sqrt{2} \ln(1 + \sqrt{2})})^{\frac{1}{2}}$ , in agreement with [B1], [B2].

## 6. CALCULATION OF CONSTANTS

The purpose of this section is to prove:

**Theorem 6.1.** *If  $\alpha$  is a quadratic irrational with renormalizing automorphism  $\psi$ , then  $\sigma^2 \in \mathbb{Q}[\alpha]$  and the twists of eigenvectors at singularities are in  $\frac{1}{2}\mathbb{Z}$ .*

**Theorem 6.2.** *In the special case when  $\alpha = \sqrt{2}$  and  $\psi$  is a renormalizing automorphism with derivative  $\begin{pmatrix} 11 & -42 \\ -6 & 23 \end{pmatrix}$  and zero drift,  $\sigma^2 = \frac{3}{8}\sqrt{2}$ .*

**Theorem 6.3.** *Let  $\psi$  be the automorphism in the previous theorem. If  $p_0$  is one of the singularities in the bottom left corner of a horizontal rectangle, and  $\underline{w}$  is the contracted eigenvector, then  $\tau_\psi(p_0, \underline{w}) = 1$ .*



To prove these results we recall that at the end of section 4 we showed the existence of a stationary mixing finite Markov chain  $\{X_i\}_{i \in \mathbb{Z}}$  and a function  $g$  s.t.

$$\mathbb{E}[g(X_0, X_1)] = -\tau_\phi(p_0, \underline{w}), \quad \frac{\sum_{i=0}^{k-1} g(X_i, X_{i+1}) + k\tau_\psi(p_0, \underline{w})}{\sqrt{k}} \xrightarrow[k \rightarrow \infty]{\text{dist}} N(0, \sigma^2). \quad (6.1)$$

We will calculate  $\{X_k\}_{k \in \mathbb{Z}}$  and  $g$  explicitly, and then find  $\mathbb{E}[g(X_0, X_1)]$  and  $\sigma^2$  using the theory of Markov chains.

Along the way we will prove (2.4), as promised in §2.

In what follows  $\psi$  is a hyperbolic homogeneous automorphism *with zero drift* which renormalizes  $\alpha$ . We assume without loss of generality that  $\psi$  has positive eigenvalues and that  $\psi$  fixes the singularities of  $\text{St}$  (otherwise we pass to  $\psi^2$ , and note that  $\sigma^2(\psi^2) = 2\sigma^2(\psi)$  and  $\tau_{\psi^2}(p_0, \underline{w}) = 2\tau_\psi(p_0, \underline{w})$ ).

Next we assume that  $\phi$  is the unique hyperbolic homogeneous automorphism with the same derivative as  $\psi$  and which fixes the rays  $L_i(p_0, \underline{w})$ , see Lemma 2.12. The drift of  $\phi$  equals  $-\tau_\psi(p_0, \underline{w})$ .

Both  $\psi$  and  $\phi$  project to the same toral automorphism, which we denote by  $\psi_0 : \text{St}_0 \rightarrow \text{St}_0$ .

**The Markov chain  $\{X_k\}_{k \in \mathbb{Z}}$ .** Let  $\mathfrak{P}$  denote the Adler–Weiss Markov partition of  $\psi_0$ , with dynamical graph  $\mathcal{G}$ . As always,  $A$  is the derivative of  $\psi$ ,  $\lambda$  is the eigenvalue of  $A$  in  $(0, 1)$ , and  $\underline{w}, \underline{v}$  are eigenvectors of  $\lambda, \lambda^{-1}$ .

Recall that  $\{X_k\}_{k \in \mathbb{Z}}$  is the Markov chain with the with set of states  $\mathfrak{P}$ , allowed transitions  $P \rightarrow Q$  iff  $\psi_0(\text{int}(P)) \cap \text{int}(Q) \neq \emptyset$ , and the transition matrix and stationary probability vector which generates of the measure of maximal entropy on  $\Sigma(\mathcal{G})$ . We calculate this data in terms of the fundamental polygon of  $\psi$ .

We begin with the cardinality of  $\mathfrak{P}$ . Recall that

$$\mathfrak{P} = \{Q_{ij} : i = 1, 2; j = 1, \dots, N_i\}$$

where  $\psi_0(Q_i) = \bigcup_{j=1}^{N_i} Q_{ij}$ .  $Q_{ij}$  are ordered so that the top  $s$ -side of  $Q_{i,j+1}$  is identified with the bottom  $s$ -side of  $Q_{i,j}$ . Since  $\mathfrak{P}$  is a refinement of  $\{Q_1, Q_2\}$ , some of the  $Q_{ij}$  are contained in  $Q_1$  and some are contained in  $Q_2$ . Let

$$N_{ik} := \#\{1 \leq j \leq N_i : Q_{ij} \subset Q_k\},$$

then  $N_i = N_{i1} + N_{i2}$  and  $|\mathfrak{P}| = \sum_{i,j} N_{ij}$ . The following lemma determines  $N_{ij}$ :

**Lemma 6.4.** *Let  $\ell^u(Q_i)$  denote the length of the unstable fibres in  $Q_i$ ,  $i = 1, 2$ , then  $(N_{ij})_{2 \times 2}$  is the unique solution in  $\mathbb{Z}$  to*

$$\begin{aligned} N_{11}\ell^u(Q_1) + N_{12}\ell^u(Q_2) &= \lambda^{-1}\ell^u(Q_1) \\ N_{21}\ell^u(Q_1) + N_{22}\ell^u(Q_2) &= \lambda^{-1}\ell^u(Q_2) \end{aligned} \quad (6.2)$$

*Proof.* If  $W^u$  is a  $u$ -fibre in  $Q_i$ , then  $\psi_0(W^u)$  can be partitioned into  $N_{i1}$   $u$ -fibres in  $Q_1$  and  $N_{i2}$   $u$ -fibres in  $Q_2$  (one for each  $Q_{ij}$ ). The sum of the lengths of these  $u$ -fibres must equal  $\ell[\psi_0(W^u)] = \lambda^{-1}\ell(W^u) = \lambda^{-1}\ell^u(Q_i)$ , so  $N_{ij}$  solve (6.2).

The existence of a solution of (6.2) in  $\mathbb{Z}$  implies that  $\ell^u(Q_1), \ell^u(Q_2)$  are linearly independent over  $\mathbb{Q}$ : Otherwise  $\lambda^{-1}$  is rational, which is never the case for an eigenvalue of a hyperbolic matrix in  $\text{SL}(2, \mathbb{Z})$ . It follows that  $(N_{ij})$  is the unique solution of (6.2) in integers.  $\square$

Next we calculate *incidence matrix* of  $\mathcal{G}$ ,  $T = (t_{PQ})_{\mathfrak{P} \times \mathfrak{P}}$ , where

$$t_{PQ} = \begin{cases} 1 & P \rightarrow Q, \\ 0 & \text{otherwise} \end{cases} \quad (P, Q \in \mathfrak{P}).$$

**Lemma 6.5** (Adler & Weiss).  $t_{Q_{ij}Q_{k\ell}} = 1 \Leftrightarrow Q_{ij} \subset Q_k$ . Thus,  $\text{rank}(T) = 2$  and every  $\mathfrak{P}$ -element in  $Q_i$  connects to  $N_{ik}$   $\mathfrak{P}$ -elements in  $Q_k$ .

*Proof.* Suppose  $Q_{ij} \in \mathfrak{P}$ . Since  $\mathfrak{P}$  refines  $\{Q_1, Q_2\}$ ,  $Q_{ij} \subset Q_k$  for  $k = 1$  or  $2$ . By construction,  $\psi_0(Q_k) = \bigcup Q_{k\ell}$ , so if  $Q_{ij} \rightarrow Q$  then  $\text{int}(Q) \subset \bigcup Q_{k\ell}$ , which means that  $\text{int}(Q)$  intersects  $\text{int}(Q_{k\ell})$  for some  $\ell$ . Since  $\{\text{int}(Q_{k\ell})\}$  are pairwise disjoint,  $Q = Q_{k\ell}$ , which proves the ( $\Rightarrow$ ) direction.

The ( $\Leftarrow$ ) direction is also true, otherwise  $\psi_0(\text{int}(Q_k))$  intersects  $\partial^u Q_{kj}$ . This is false, because  $\partial^u Q_{k\ell} \subset \psi_0(\partial^u Q_k) \subset \psi_0(\text{int}(Q_k))^c$ . So  $t_{Q_{ij}Q_{k\ell}} = 1 \Leftrightarrow Q_{ij} \subset Q_k$ .

We see that the incidence matrix  $T$  has two types of rows: those of  $\mathfrak{P}$ -elements  $P \subset Q_1$ , and those of  $\mathfrak{P}$ -elements in  $P \subset Q_2$ . These rows are different, because

- if  $P \in \mathfrak{P}, P \subset Q_1$ , then  $\#\{Q \in \mathfrak{P} : Q \subset Q_k, t_{PQ} = 1\} = N_{1k}$ ,
- if  $P \in \mathfrak{P}, P \subset Q_2$ , then  $\#\{Q \in \mathfrak{P} : Q \subset Q_k, t_{PQ} = 1\} = N_{2k}$ ,
- $\begin{pmatrix} N_{11} \\ N_{12} \end{pmatrix} \neq \begin{pmatrix} N_{21} \\ N_{22} \end{pmatrix}$ , otherwise by Lemma 6.4  $\ell^u(Q_1) = \ell^u(Q_2)$  and  $\lambda$  is rational.

Since different rows of zeroes and ones are linearly independent,  $\text{rank}(T) = 2$ .  $\square$

Next we determine the *transition matrix* of the Markov chain  $\{X_k\}_{k \in \mathbb{Z}}$ : the matrix  $(p_{PQ})_{P, Q \in \mathfrak{P}}$  s.t.  $p_{PQ} = \mathbb{P}(X_1 = Q | X_0 = P)$ .

**Lemma 6.6** (Adler & Weiss).  $(p_{PQ})_{P, Q \in \mathfrak{P}} = \lambda M^{-1} T M$  where  $M$  is the diagonal matrix with diagonal entries  $M_{PP} = \ell^u(P)$ .

*Proof.* By the Adler–Weiss Theorem,  $\mathbb{P} = m_0 \circ \pi_0^{-1}$ , where  $m_0$  is the normalized area measure and  $\mathbb{P}$  is the joint distribution measure given by

$$\mathbb{P}(E) := \mathbb{P}[(X_k)_{k \in \mathbb{Z}} \in E], \quad (E \subset \Sigma(\mathcal{G}) \text{ Borel}).$$

Therefore, if  $P = Q_{ij}, Q = Q_{k\ell}$ , then  $p_{P,Q} = m_0[P \cap \psi_0^{-1}(Q)]/m_0(P)$ .

$P, Q$ , and  $P \cap \psi_0^{-1}(Q)$  are parallelograms with sides in the stable and unstable directions. Let  $\ell^u(\cdot), \ell^s(\cdot)$  denote the lengths of these sides, then  $\ell^s(P) = \lambda \ell^s(Q_i)$ ,  $\ell^s[P \cap \psi_0^{-1}(Q)] = \ell^s(P) = \lambda \ell^s(Q_i)$ , and  $\ell^u[P \cap \psi_0^{-1}(Q)] = \lambda \ell^u(Q)$ . Denoting the angle between the stable and unstable directions by  $\beta$ , we see that  $p_{PQ} = t_{PQ} \frac{\lambda^2 \ell^s(Q_i) \ell^u(Q) \sin \beta}{\lambda \ell^s(Q_i) \ell^u(P) \sin \beta} = \lambda \ell^u(P)^{-1} t_{PQ} \ell^u(Q)$ .  $\square$

**Proof of (2.4) and calculation of  $g$ .** Let  $\xi$  denote the associate  $\mathbb{Z}$ -coordinate of  $\psi$ . The definition of  $g$  is based on (2.4), which says that the Frobenius function  $F_\psi$  of  $\xi$  is either (a)  $\mathfrak{P} \vee \psi_0^{-1}(\mathfrak{P})$ -measurable or (b)  $\mathfrak{P} \vee \psi_0(\mathfrak{P})$ -measurable. In case (a),  $g(P, Q)$  is the value of  $F_\psi$  on  $\text{int}(P) \cap \psi_0^{-1}[\text{int}(Q)]$ . In case (b),  $g(P, Q)$  is the value of  $F_\psi$  on  $\psi_0^{-1}[\text{int}(P)] \cap \text{int}(Q)$ . In this section we prove (2.4), and give an explicit formula for  $g$ .

Recall from §2 that  $\psi$  or  $\psi^{-1}$  has a fundamental polygon of the form  $R = \theta_0(R_0)$ , where  $\theta_0 : \text{St}_0 \rightarrow \text{St}_0$  is a toral automorphism which fixes the punctures of  $\text{St}_0^*$  and  $R_0$  is one of the shapes in figure 2.

We will limit ourselves to the case when  $\psi$  has such a fundamental polygon. The case of  $\psi^{-1}$  can be handled by the identity  $F_{\psi^{-1}} = -F_\psi \circ \psi_0$ .

Suppose  $W^u$  is a  $u$ -fibre in  $R$ . The  $\mathbb{Z}$ -displacement of  $W^u$  is defined by

$$\phi(W^u) := \xi(\text{endpoint of } \widetilde{W}^u) - \xi(\text{beginning point of } \widetilde{W}^u)$$

for some (any) lift  $\widetilde{W}^u$  of  $W^u$  to  $\text{St}_0$ .  $\phi(W^u)$  is independent of the lift, and can be easily determined from the endpoint of  $W^u$  as follows: Let  $H = (2, 1)$ , then referring to figure 2,

- the  $\mathbb{Z}$ -displacement of  $u$ -fibres terminating at  $\theta_0(AF)$  is  $(-1)$ ;
- the  $\mathbb{Z}$ -displacement of  $u$ -fibres terminating at  $\theta_0(EH)$  is  $(+1)$ ;
- the  $\mathbb{Z}$ -displacement of  $u$ -fibres terminating at  $\theta_0(HD)$  is  $(-1)$ .

The unique  $\mathfrak{P}$ -element which has a  $u$ -fibre terminating at  $H$  is called *critical*. The non-critical elements of  $\mathfrak{P}$  have the virtue that all their  $u$ -fibres have the same  $\mathbb{Z}$ -displacement. Let

$\phi(Q_{ij}) :=$  the value of the  $\mathbb{Z}$ -displacement of  $u$ -fibres in  $Q_{ij}$  ( $Q_{ij}$  non-critical).

**Lemma 6.7.** *The critical element is  $Q_{2N_2}$ .*

*Proof.* Call the critical element  $Q_{k\ell}$ .

$H$  lies on the top  $s$ -side of  $Q_2$ ,  $\psi_0(H) = H$ , and  $\psi_0(\partial^s Q_2) \subset \partial^s Q_2$ , therefore  $H \in \text{int}[\partial^s \psi_0(Q_2)]$ . It follows that  $k = 2$  and  $1 \leq \ell \leq N_2$ . Let  $W^u$  be the  $u$ -fibre in  $Q_{2\ell}$  whose closure contains  $H$ . If  $\ell < N_2$  then  $H = \psi_0^{-1}(H) \in \psi_0^{-1}[W^u] \subset \text{int}(Q_2) \cup \text{bottom } s\text{-side of } Q_2$ . This is false, so  $\ell = N_2$ .  $\square$

Let  $\tilde{q}_0$  denote one of the singularities in the middle of the horizontal side of one of the horizontal rectangles. Since, by assumption,  $\psi$  fixes  $\tilde{q}_0$  and has positive eigenvalues, there is a constant  $\tau$  s.t.  $\psi[L_i(\tilde{q}_0, \underline{w})] = L_{i+\tau}(\tilde{q}_0, \underline{w})$  ( $\tau = \tau_\psi(\tilde{q}_0, \vec{v})$ ). A continuity argument shows that necessarily  $\psi[L_i(\tilde{q}_0, -\underline{w})] = L_{i+\tau}(\tilde{q}_0, -\underline{w})$

**Lemma 6.8.** *If  $\psi$  has a fundamental domain of the form  $\theta(R_0)$  with  $R_0$  as in figure 2, then  $F_\psi$  is  $\mathfrak{P} \vee \psi_0^{-1}(\mathfrak{P})$ -measurable, and  $g(Q_{ij}, Q_{k\ell}) = \tau + \sum_{s=1}^{\ell-1} \phi(Q_{ks})$ .*

(The last expression makes sense because  $Q_{ks}$  is non-critical when  $s \leq \ell - 1$ .)

*Proof.* Let  $P := Q_{ij}$ ,  $Q := Q_{k\ell}$ , and suppose  $p \in \text{int}(P) \cap \psi_0^{-1}[\text{int}(Q)]$ . By the definition of the Frobenius function,  $F_\psi(p) = \xi(\psi(\tilde{p})) - \xi(\tilde{p})$  for some (any)  $\tilde{p} \in \pi^{-1}(p)$ . We choose the  $\tilde{p}$  s.t.  $\xi(\tilde{p}) = 0$ , then  $F_\psi(p) = \xi(\psi(\tilde{p}))$ .

To calculate this we construct a path  $\gamma$  in  $\text{St}_0$  from the fixed point  $q_0 = (1, 0)$  to  $p$  and analyze the lift of  $\psi_0[\gamma]$  to the infinite staircase. Let  $q$  denote the intersection of  $W^u(p)$  and  $W^s(q_0)$  (the  $u$  and  $s$  fibres of  $p$  and  $q_0$ ). The path  $\gamma$  we use is the concatenation of  $[q_0, q] \subset W^s(p_0)$  and  $[q, p] \subset W^u(p)$ .

The curve  $\gamma$  begins with a piece of a ray emanating from  $q_0$  in direction  $\underline{w}$ . Let  $\tilde{\gamma}$  denote its unique lift to  $\text{St}$  which begins with a segment in  $L_0(\tilde{q}_0, \underline{w})$ . Since  $\gamma$  does not cross  $\partial R$ , all points in  $\tilde{\gamma}$ , in particular its end point, have  $\mathbb{Z}$ -coordinate equal to zero. It follows that  $\tilde{\gamma}$  ends at  $\tilde{p}$ .

Let  $\tilde{\zeta} := \psi[\tilde{\gamma}]$ . This curve ends at  $\psi(\tilde{p})$ , so  $\xi[\text{end of } \tilde{\zeta}] = \xi[\psi(\tilde{p})]$ .

As for its beginning, since  $\psi[L_0(\tilde{q}_0, \underline{w})] \subset L_\tau(\tilde{q}_0, \underline{w})$ ,  $\tilde{\zeta} := \psi[\tilde{\gamma}]$  begins with a segment in  $L_\tau(\tilde{q}_0, \underline{w})$ . It follows that  $\xi[\text{beginning of } \tilde{\zeta}] = \tau$ .

The curve  $\tilde{\zeta}$  projects to  $\psi_0[\gamma]$ . To calculate  $\psi_0[\gamma]$ , we first recall that  $Q_{ij} \rightarrow Q_{k\ell}$ , and therefore  $P \subset Q_k = \bigcup_{s=1}^{N_k} \psi_0^{-1}(Q_{ks})$ .  $\{Q_{ks}\}_{s=1}^{N_k}$  are ordered so that  $[q, p] = \bigcup_{s=1}^{\ell-1} [q_s, q_{s+1}] \cup [q_\ell, p]$ , where  $[q_s, q_{s+1}] = W^u(p) \cap \psi_0^{-1}(Q_{ks})$  and  $[q_\ell, p] \subset \psi_0^{-1}(Q_{k\ell})$ . So  $\psi_0[\gamma]$  is a concatenation of

- $[q_0, \psi(q)]$  (a subsegment of  $W^s(q_0)$ ), followed by
- a  $u$ -fibre in  $Q_{k1}$  ( $\psi_0[q_1, q_2]$ ), followed by
- a  $u$ -fibre in  $Q_{k2}$  ( $\psi_0[q_2, q_3]$ ), followed by

- ...
- a  $u$ -fibre in  $Q_{k,\ell-1}$  ( $\psi_0[q_{\ell-1}, q_\ell]$ ), followed by
- the beginning of a  $u$ -fibre in  $Q_{k\ell}$ , which terminates at  $\psi_0(p)$ .

It follows that  $\xi[\text{end of } \tilde{\zeta}] - \xi[\text{beginning of } \tilde{\zeta}] = \sum_{s=1}^{\ell-1} \phi(Q_{ks})$ . Substituting the values of  $\xi$  at the endpoint of  $\tilde{\zeta}$ , we find that  $F_\psi(p) = \tau + \sum_{s=1}^{\ell-1} \phi(Q_{ks})$ .  $\square$

**Proof of Theorem 6.1.** That the twists always belong to  $\frac{1}{2}\mathbb{Z}$  was proved in Lemma 2.11, so we focus on the value of  $\sigma^2$ . There is no loss of generality in assuming that  $\psi$  has zero drift.

Let  $g$  be the function found in Lemma 6.8, and define a family of  $\mathfrak{P} \times \mathfrak{P}$  matrices  $\Phi(\theta)$  by

$$\Phi_{P,Q}(\theta) := p_{P,Q} \exp[\theta g(P, Q)] \quad (P, Q \in \mathfrak{P}).$$

These are a positive matrices, and the mixing of  $\sigma : \Sigma(\mathcal{G}) \rightarrow \Sigma(\mathcal{G})$  (Adler–Weiss Theorem) guarantees that they are primitive. By the Perron–Frobenius Theorem,  $\Phi(\theta)$  has a simple positive eigenvalue  $\lambda(\theta)$  such that  $\lambda(\theta)$  is larger than the modulus of all other eigenvalues. When  $\theta = 0$ ,  $\Phi$  is stochastic, and  $\lambda(0) = 1$ .

Since  $\Phi(\theta)$  depends analytically on  $\theta$ ,  $\lambda(\theta)$  is analytic on some interval  $(-\varepsilon, \varepsilon)$ . It is known that

$$\lambda(0) = 1, \quad \left. \frac{d}{d\theta} \ln \lambda(\theta) \right|_{\theta=0} = \mathbb{E}[g(X_0, X_1)] \quad \text{and} \quad \left. \frac{d^2}{d\theta^2} \ln \lambda(\theta) \right|_{\theta=0} = \sigma^2. \quad (6.3)$$

See Doeblin [D], Nagaev [Ngv], or chapter 4 in [PP].

We use this formula to show that  $\sigma^2 \in \mathbb{Q}[\alpha]$ , where  $\alpha$  is the angle normalized by  $\psi$ . Let  $A$  denote the derivative of  $\psi$  and let  $\lambda$  denote the eigenvalue of  $A$  in  $(0, 1)$ . Since  $\psi$  renormalizes  $\alpha$ ,  $\alpha = \frac{1}{2} + \frac{1}{2} \tan \theta \pmod{1}$  where  $A \begin{pmatrix} 1 \\ \tan \theta \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ \tan \theta \end{pmatrix}$ . Since  $A$  is a matrix of integers,  $\lambda \in \mathbb{Q}[\tan \theta] = \mathbb{Q}[\alpha]$ . We'll show that  $\sigma^2 \in \mathbb{Q}[\lambda]$ , and deduce that  $\sigma^2 \in \mathbb{Q}[\alpha]$ .

We need the following claim. Let  $\mathcal{A}$  denote the collection of functions of the form  $\varphi(\theta) = \sum_{k=-n}^n a_k e^{k\theta}$  with arbitrary  $n \in \mathbb{N}$  and  $a_k \in \mathbb{N} \cup \{0\}$ , and set  $\mu(\theta) := \lambda(\theta)/\lambda$ .

*Claim.* There are  $\beta_{ij}(\theta) \in \mathcal{A}$  s.t.  $\mu(\theta)$  is the largest eigenvalue of  $\begin{pmatrix} \beta_{11}(\theta) & \beta_{12}(\theta) \\ \beta_{21}(\theta) & \beta_{22}(\theta) \end{pmatrix}$ , for all  $\theta \in (-\varepsilon, \varepsilon)$ .

*Proof of the claim.* Let  $\Psi(\theta)$  denote the  $\mathfrak{P} \times \mathfrak{P}$  matrix  $(t_{PQ} \exp[\theta g(P, Q)])_{P,Q \in \mathfrak{P}}$ . By lemma 6.6,  $\Psi = \lambda^{-1} M \Phi M^{-1}$ , so  $\mu(\theta)$  is the leading eigenvalue of  $\Psi(\theta)$ .

As our formulas for  $t_{Q_{ij}, Q_{k\ell}}$  and  $g(Q_{ij}, Q_{k\ell})$  show, if  $P, Q \in \mathfrak{P}$  are both included in the same  $Q_k$ , then the  $P$ -row of  $\Psi(\theta)$  is equal to the  $Q$ -row of  $\Psi(\theta)$ , and if  $P, Q$  are not included in the same  $Q_k$  then the  $P$ -row and the  $Q$ -row are linearly independent. In particular,  $\text{rank}[\Psi(\theta)] = 2$ .

We think of  $\Psi(\theta)$  as of the linear transformation  $\underline{u} \mapsto \underline{u}\Psi(\theta)$  on  $\mathbb{R}^{\mathfrak{P}}$ . Let  $V_\theta := \text{Im}[\Psi(\theta)]$ . Then  $\dim V_\theta = 2$  and

$$V_\theta = \text{Span}\{\underline{e}_P \Psi(\theta), \underline{e}_Q \Psi(\theta)\}$$

when  $P, Q \in \mathfrak{P}$ ,  $P \subset Q_1$ ,  $Q \subset Q_2$ , and  $\underline{e}_P, \underline{e}_Q$  are the row vectors  $(\underline{e}_P)_R = \delta_{PR}$ ,  $(\underline{e}_Q)_R = \delta_{QR}$  where  $\delta_{PQ}$  is the Kronecker symbol.

$\Psi(\theta)$  preserves  $V_\theta$ , and since  $V_\theta$  contains all the (left) eigenvectors of  $\Psi(\theta)$ ,  $\mu(\theta)$  is the leading eigenvalue of  $\Psi(\theta)|_{V_\theta} : V_\theta \rightarrow V_\theta$ . We represent  $\Psi(\theta)|_{V_\theta} : V_\theta \rightarrow V_\theta$  in

the basis  $\{\underline{e}_P\Psi(\theta), \underline{e}_Q\Psi(\theta)\}$ . For every  $S \in \mathfrak{P}$ ,

$$\begin{aligned} ((\underline{e}_P\Psi)\Psi)_S &= (\underline{e}_P\Psi^2)_S = (\Psi^2)_{PS} = \sum_{R \in \mathfrak{P}, R \subset Q_1} \Psi_{P,R}\Psi_{R,S} + \sum_{R \in \mathfrak{P}, R \subset Q_2} \Psi_{P,R}\Psi_{R,S} \\ &= \sum_{R \in \mathfrak{P}, R \subset Q_1} \Psi_{P,R}\Psi_{P,S} + \sum_{R \in \mathfrak{P}, R \subset Q_2} \Psi_{P,R}\Psi_{Q,S} \\ &= \left( \sum_{R \in \mathfrak{P}, R \subset Q_1} \Psi_{P,R} \right) (\underline{e}_P\Psi)_S + \left( \sum_{R \in \mathfrak{P}, R \subset Q_2} \Psi_{P,R} \right) (\underline{e}_Q\Psi)_S \\ &= \left( \sum_{R \in \mathfrak{P}, R \subset Q_1} t_{PR}e^{\theta g(P,R)} \right) (\underline{e}_P\Psi)_S + \left( \sum_{R \in \mathfrak{P}, R \subset Q_2} t_{PR}e^{\theta g(P,R)} \right) (\underline{e}_Q\Psi)_S. \end{aligned}$$

The terms in the brackets belong to  $\mathcal{A}$ . A similar formula holds for  $(\underline{e}_Q\Psi)\Psi$ . So  $\Psi(\theta) : V_\theta \rightarrow V_\theta$  is represented by a  $2 \times 2$ -matrix with entries in  $\mathcal{A}$ , and  $\mu(\theta)$  is the leading eigenvalue of that matrix, as claimed.

Call the matrix in the claim  $B_\theta$ , and let  $f_\theta(t) = t^2 - a(\theta)t - b(\theta)$ , be the characteristic polynomial of  $B_\theta$ , then  $a(\theta) = \text{tr}(B_\theta) \in \mathcal{A}$  and  $b(\theta) = -\det(B_\theta) \in \mathcal{A} - \mathcal{A}$ . It follows that  $a^{(k)}(0), b^{(k)}(0) \in \mathbb{Z}$  for all  $k \geq 0$ .

The eigenvalues of  $B_\theta$  are zeroes of  $f_\theta$ , therefore  $f_\theta(\mu(\theta)) = 0$ . We differentiate this identity twice with respect to  $\theta$  and then substitute  $\theta = 0$ , noting that  $\mu'(0) = \lambda'(0) = \mathbb{E}(g) = \text{drift of } \psi = 0$ . Rearranging terms, we obtain

$$\mu''(0) = \frac{a''(0)\mu(0) + b''(0)}{2\mu(0) - a(0)}.$$

Similarly,  $(\ln \lambda)''(0) = (\ln \mu)''(0) = \frac{\mu''(0)}{\mu(0)}$ , so  $\sigma^2 = (\ln \lambda)''(0) = \frac{a''(0)\mu(0) + b''(0)}{2\mu(0)^2 - a(0)\mu(0)}$ . Since  $\mu(0)^2 - a(0)\mu(0) - b(0) = 0$ , we obtain

$$\sigma^2 = \frac{a''(0)\mu(0) + b''(0)}{a(0)\mu(0) + 2b(0)} \in \mathbb{Q}[\mu(0)].$$

It remains to recall that  $\lambda(0) = 1$ , therefore  $\mu(0) = 1/\lambda$ , so  $\mathbb{Q}[\mu(0)] \subset \mathbb{Q}[\alpha]$ .  $\square$

**Proof of Theorem 6.2.** We now specialize to the case of  $\alpha = \sqrt{2}$ .

Let  $\phi$  denote the homogeneous automorphism  $\phi$  with derivative  $\begin{pmatrix} 11 & -42 \\ -6 & 23 \end{pmatrix}$ , and which fixes  $L_0[\tilde{p}_0, \underline{w}]$  ( $\tilde{p}_0$  =singularity in the bottom left/right corner of horizontal rectangle zero,  $\underline{w}$  =contracted eigenvector). We will find the function  $g$  which drives the random walk of  $\phi$ , and then use (6.1) to calculate  $\sigma^2$  and the drift of  $\phi$ . There are four steps:

- (1) Finding the fundamental polygon of  $\phi$
- (2) Calculating the Markov partition and the transition matrix  $(t_{PQ})_{\mathfrak{P} \times \mathfrak{P}}$
- (3) Calculating  $g$  and  $\Psi(\theta)$
- (4) Finding a closed form for the leading eigenvalue  $\mu(\theta)$  of  $\Psi(\theta)$ , and using the identity  $(\ln \mu)''(0) = \sigma^2$ .

All the calculations can be done in closed form, but  $\mathfrak{P}$  is too large to do this reliably by hand (it's a  $58 \times 58$  matrix). We will supply alternative formulas for  $t_{PQ}$  and  $g(P, Q)$  which can be easily implemented on a computer, and which provably give the result with absolute precision.

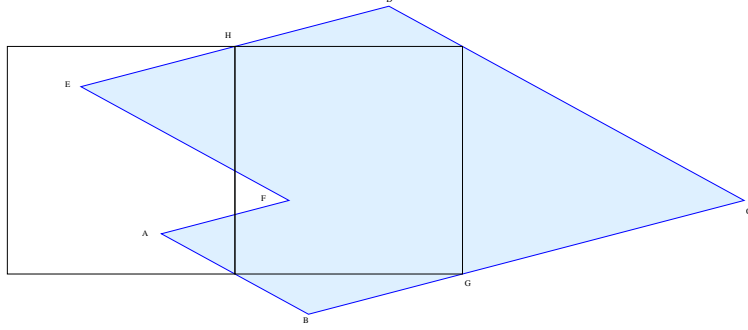


FIGURE 4. Fundamental domain for the renormalizing automorphism of  $\sqrt{2}$

*Step 1. The fundamental polygon.* The derivative matrix is  $\begin{pmatrix} 11 & -42 \\ -6 & 23 \end{pmatrix}$ . The eigenvalues are  $\lambda = 17 - 12\sqrt{2}$ ,  $\lambda^{-1} = 17 + 12\sqrt{2}$ , and the eigenvectors are  $\underline{v} = \begin{pmatrix} 1-2\sqrt{2} \\ 1 \end{pmatrix}$  and  $\underline{w} = \begin{pmatrix} 1+2\sqrt{2} \\ 1 \end{pmatrix}$ .<sup>1</sup>

This places us outside of the cases covered in figure 2. Instead of looking for a reduction to one these cases by means of Lemma 2.7, we find a fundamental polygon directly. It is given in figure 4. It is a simple calculation to find the vertices explicitly:  $A(\frac{4+\sqrt{2}}{8}, \frac{1}{4\sqrt{2}})$ ,  $B(\frac{12-\sqrt{2}}{8}, -\frac{1}{4\sqrt{2}})$ ,  $C(2 + \frac{7\sqrt{2}}{8}, \frac{4-\sqrt{2}}{8})$ ,  $D(\frac{12+\sqrt{2}}{8}, \frac{8+\sqrt{2}}{8})$ ,  $E(\frac{4-\sqrt{2}}{8}, \frac{8-\sqrt{2}}{8})$ ,  $F(\frac{7\sqrt{2}}{8}, \frac{4-\sqrt{2}}{8})$  (figure 4).

*Step 2. The Markov partition and the incidence matrix*

$\mathfrak{P} = \{Q_{11}, \dots, Q_{1N_1}; Q_{21}, \dots, Q_{2N_2}\}$ . We can find  $N_1$  and  $N_2$  using Lemma 6.4. The first step is find  $\ell^u(Q_1) = |AB|$  and  $\ell^u(Q_2) = |CD|$ . A direct calculation with the coordinates found in step 1 leads to nested roots, but this can be avoided by first expressing the coordinates of  $A, B, C, D$  in the form “puncture +  $t\underline{v}$ .” This leads to the presentation

$$\ell^u(Q_1) = \frac{\sqrt{2}}{4} \|\underline{v}\|, \quad \ell^u(Q_2) = \frac{2 + \sqrt{2}}{4} \|\underline{v}\|,$$

which makes the equations in Lemma 6.4 easy to solve.

The solution is  $N_{11} = 5$ ,  $N_{12} = 12$ ,  $N_{21} = 12$ ,  $N_{22} = 29$ . Thus  $N_1 = N_{11} + N_{12} = 17$  and  $N_2 = N_{21} + N_{22} = 41$ . As a result,  $\mathfrak{P}$  contains 58 elements, of which  $N_{11} + N_{21} = 17$  are in  $Q_1$  and  $N_{12} + N_{22} = 41$  are in  $Q_2$ . Every  $\mathfrak{P}$ -element in  $Q_1$  connects to 5 elements in  $Q_1$  and 12 elements in  $Q_2$ ; and every  $\mathfrak{P}$ -element in  $Q_2$  connects to 12 elements in  $Q_1$  and 29 elements in  $Q_2$ .

So far so good. But to find the incidence matrix we also need to know which of the  $Q_{ij}$  fall in  $Q_1$  and which fall in  $Q_2$ . A calculation by hand or “by inspection” is possible in principle, but not very reliable due to the size of the problem. We look for a method for calculating the position of  $Q_{ij}$  using a computer.

Let  $\ell^s(Q_i)$  denote the lengths of the stable sides of  $Q_i$ , then side  $BC$  of our fundamental polygon has length  $\ell^s(R) := \ell^s(Q_1) + \ell^s(Q_2)$ , and this side contains  $p_0 = (2, 0)$  (figure 2).

<sup>1</sup>Here we deviate from our convention to choose the eigenvectors in the form  $\begin{pmatrix} 1 \\ * \end{pmatrix}$ .

We parametrize this side by  $L^s := [0, 1]$ , representing a point by its normalized distance from the left endpoint. The puncture  $p_0$ , for example, is parametrized by  $\widehat{p}_0 := \frac{\sqrt{2}}{4}$  (nested roots can be avoided as above); the  $s$ -sides of  $Q_1$  and  $Q_2$  are parametrized by  $L_1^s := [0, \frac{\ell^s(Q_1)}{\ell^s(R)}] = [0, 1 - \frac{\sqrt{2}}{2})$  and  $L_2^s := [\frac{\ell^s(Q_1)}{\ell^s(R)}, 1] = [1 - \frac{\sqrt{2}}{2}, 1)$ .

The key observation is that with this parametrization, a  $u$ -fibre which begins at  $\tau \in [0, 1]$ , ends at  $\tau - \frac{\ell^s(Q_1)}{\ell^s(R)} \bmod 1 = \tau + \frac{\sqrt{2}}{2} \bmod 1$ . We can use this to keep track of the position of  $Q_{ij}$ , by following the image of a  $u$ -fibre in the interior of  $Q_i$ .

Suppose first that  $i = 1$ .  $Q_1$  contains the  $u$ -fibre  $W_1^u$  which starts at  $\frac{\ell^s(Q_1)}{2\ell^s(R)}$ . Since  $\psi_0$  contracts  $L^s$  towards  $p_0$  by a factor  $\lambda$ , it maps  $\frac{\ell^s(Q_1)}{2\ell^s(R)}$  to  $\widehat{q}_1 := \widehat{p}_0 - \lambda \left( \widehat{p}_0 - \frac{\ell^s(Q_1)}{2\ell^s(R)} \right) = \frac{2-57\sqrt{2}}{4} \bmod 1$ , so  $\psi_0(W_1^u)$  can be broken to  $u$ -fibres starting at

$$\begin{aligned} \tau_{1j} &:= \widehat{q}_1 - (j-1) \frac{\ell^s(Q_1)}{\ell^s(R)} \bmod 1 \\ &= \frac{2-57\sqrt{2}}{4} + (j-1) \frac{\sqrt{2}}{2} \bmod 1 \quad (j = 1, \dots, 17). \end{aligned} \quad (6.4)$$

$Q_{1j}$  is the parallelogram which contains the  $u$ -fibre which starts at  $\tau_{1j}$ . Similarly,  $Q_2$  contains the  $u$ -fibre  $W_2^u$  which starts at the fixed point  $p_0$ , so  $Q_{2j}$  is the parallelogram which contains the  $u$ -fibre which starts at

$$\begin{aligned} \tau_{2j} &:= \widehat{p}_0 - (j-1) \frac{\ell^s(Q_1)}{\ell^s(R)} \bmod 1 \\ &= \frac{\sqrt{2}}{4} + (j-1) \frac{\sqrt{2}}{2} \quad (j = 1, \dots, 41). \end{aligned} \quad (6.5)$$

It follows that  $Q_{ij} \subset Q_k$  iff  $L_k^s \ni \tau_{ij}$ , and therefore, by Lemma 6.5 we have the following explicit formula for the incidence matrix:

$$t_{Q_{ij}Q_{k\ell}} = 1_{L_k^s}(\tau_{ij}). \quad (6.6)$$

This can be calculated easily on a computer, *provided the precision of the calculation is smaller than the distance between  $\tau_{ij}$  and the endpoints of  $L_k^s$ .*

We estimate the precision we need. The endpoints of  $L_k^s$  are  $a \in \{0, 1, 1 - \frac{\sqrt{2}}{2}\}$ . Since  $\text{dist}(\tau_{ij}, a + \mathbb{Z}) \geq \min\{\frac{1}{4} \text{dist}(4\tau_{ij}, 4a + \mathbb{Z}), \frac{1}{4}\}$ , we have the (generous) lower bound  $\text{dist}(\tau_{ij}, \{0, 1 - \frac{\sqrt{2}}{2}, 1\} + \mathbb{Z}) \geq \min\{\text{dist}(k\sqrt{2}, \mathbb{Z}) : k = 1, \dots, 100\}$ . The last quantity is bounded below by  $2 \cdot 10^{-4}$ , as can be seen from the sixth principal convergent of  $\sqrt{2}$ ,  $\frac{239}{169}$ . So the precision we need for the calculation is just  $10^{-4}$ , which is easily available on a standard machine.

*Step 3. Calculating  $g(Q_{ij}, Q_{k\ell})$ .*

We use Lemma 6.8. Note that by choice of  $\phi$ ,  $\tau = 0$ , and the calculation of  $g$  boils down to finding the  $\mathbb{Z}$ -displacement of suitable  $u$ -fibres.

The  $\mathbb{Z}$ -displacement of a  $u$ -fibre can be determined from the location of its endpoint, see figure 4. This in turn can be determined from the location of the beginning point as follows. Suppose a  $u$ -fibre starts at  $\tau \in L^s = [0, 1]$ .

- If  $\tau \in [0, \frac{\ell^s(Q_1)}{\ell^s(R)}] = [0, 1 - \frac{\sqrt{2}}{2})$ , then the endpoint is in  $AF$  and the  $\mathbb{Z}$ -displacement is 0.
- If  $\tau \in [\frac{\ell^s(Q_1)}{\ell^s(R)}, \frac{\ell^s(Q_1)}{\ell^s(R)} + \frac{\ell^s(Q_2)}{2\ell^s(R)}] = [1 - \frac{\sqrt{2}}{2}, 1 - \frac{\sqrt{2}}{4})$ , then the endpoint is between  $E$  and  $(1, 1)$  and the  $\mathbb{Z}$ -displacement is  $(-1)$

- If  $\tau \in [\frac{\ell^s(Q_1)}{\ell^s(R)} + \frac{\ell^s(Q_2)}{2\ell^s(R)}, 1) = [1 - \frac{\sqrt{2}}{4}, 1)$ , then the endpoint is between  $(1, 1)$  and  $D$  and the  $\mathbb{Z}$ -displacement is  $(+1)$

In summary, a  $u$ -fibre which starts at  $\tau \in L^s$  has  $\mathbb{Z}$ -displacement

$$\gamma(\tau) = 1_{[1-\frac{\sqrt{2}}{4}, 1)}(\tau) - 1_{[1-\frac{\sqrt{2}}{2}, 1-\frac{\sqrt{2}}{4})}(\tau).$$

It follows that  $g(Q_{ij}, Q_{k\ell}) = \sum_{m=1}^{\ell-1} \gamma(\tau_{im})$ , and by step 2,

$$\Psi_{Q_{ij}, Q_{k\ell}}(\theta) = 1_{L_k^s}(\tau_{ij}) \exp \left( \theta \sum_{m=1}^{\ell-1} [1_{[1-\frac{\sqrt{2}}{4}, 1)}(\tau_{im}) - 1_{[1-\frac{\sqrt{2}}{2}, 1-\frac{\sqrt{2}}{4})}(\tau_{im})] \right),$$

where  $\tau_{ij}, t_{im}$  are given by (6.4) and (6.5),  $L_1^s = [0, 1 - \frac{\sqrt{2}}{2})$ , and  $L_2^s = [1 - \frac{\sqrt{2}}{2}, 1)$ .

As in the case of the incidence matrix, this can be calculated with complete precision on a standard computer.

*Step 4. Calculation of  $\sigma^2$ .* We implemented the formulas in the previous step on **Mathematica**, and found  $\Psi(\theta)$ .

As predicted by the general theory,  $\text{rank}[\Psi(\theta)] = 2$ , so the characteristic polynomial of  $\Psi(\theta)$  takes the form  $t^{|\mathfrak{B}|-2}[t^2 + b(\theta)t + c(\theta)]$ . The largest eigenvalue can therefore be found in closed form. We did this using **Mathematica** and got

$$\mu(\theta) = \frac{1}{2}e^{-2\theta} \left( 9 + 16e^\theta + 9e^{2\theta} + 3\sqrt{(1+e^\theta)^2(9+14e^\theta+9e^{2\theta})} \right). \quad (6.7)$$

It follows that  $\sigma^2 = (\log \mu)''(0) = \frac{3}{8}\sqrt{2}$ .  $\square$

**Proof of Theorem 6.3.** The calculations in the previous proof were done for the renormalizing automorphism  $\phi$  which fixes the rays  $L_0(p_0, \underline{w})$ . By Lemma 2.12,  $\tau_\psi(p_0, \underline{w}) = -\delta(\phi)$ . The drift of  $\phi$  is  $\mathbb{E}(g)$ , and by (6.3)  $\mathbb{E}(g) = (\log \mu)'(0)$ . It follows that  $\tau_\psi(p_0, \underline{w}) = -(\log \mu)'(0)$ . By (6.7),  $\tau_\psi(p_0, \underline{w}) = 1$ .

Here is another proof that  $\tau_\psi(p_0, \underline{w}) = 1$ . The first step is to express the derivative of  $\psi$  through generators of  $\Gamma(2)$ :

$$\begin{pmatrix} 11 & -42 \\ -6 & 23 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 6 & 1 \end{pmatrix} \begin{pmatrix} 1 & -4 \\ 0 & 1 \end{pmatrix},$$

Let  $\psi_i$  denote the unique homogeneous automorphism with zero drift and derivative  $\begin{pmatrix} 1 & -4 \\ 0 & 1 \end{pmatrix}$  ( $i=1$ ),  $\begin{pmatrix} 1 & 0 \\ 6 & 1 \end{pmatrix}$  ( $i=2$ ),  $\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$  ( $i=3$ ), and  $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$  ( $i=4$ ). By the uniqueness of homogeneous automorphisms with zero drift and given derivative,  $\psi = \psi_4 \circ \psi_3 \circ \psi_2 \circ \psi_1$ .

The automorphisms  $\psi_i$  are known explicitly (see appendix). To describe them, note that  $\text{St}$  has two canonical cylinder decompositions of  $\text{St}$ , one into horizontal cylinders and the other into vertical cylinders (Figure 5).

- (1)  $\psi_1$  acts on every horizontal cylinder in by  $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x-4y \pmod{2} \\ y \pmod{1} \end{pmatrix}$ , where  $(x, y)$  are measured from the bottom left corner of the corresponding horizontal rectangle. So  $\psi_1[L_0(p, \underline{w})] = L_0(p', \begin{pmatrix} -3+2\sqrt{2} \\ 1 \end{pmatrix})$ , where  $p' :=$ bottom right corner of horizontal rectangle  $\#0$  ( $p'$  is congruent to  $p$ )
- (2)  $\psi_2$  acts on every vertical cylinder by  $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \pmod{1} \\ y+6x \pmod{2} \end{pmatrix}$ , where  $(x, y)$  are measured from the bottom left corner of the corresponding vertical rectangle. So  $\psi_2[L_0(p', \begin{pmatrix} -3+2\sqrt{2} \\ 1 \end{pmatrix})] = L_0(p', \begin{pmatrix} 3-2\sqrt{2} \\ 17-12\sqrt{2} \end{pmatrix}) = L_0(p', \begin{pmatrix} -(3+2\sqrt{2}) \\ 1 \end{pmatrix}) = L_1(q, \begin{pmatrix} -(3+2\sqrt{2}) \\ 1 \end{pmatrix})$ , where  $q$  is the singularity at middle of the top side of horizontal rectangle  $\#1$  ( $q$



is congruent to  $p$  and  $p'$ ). We had to move one horizontal rectangle up, because  $-\binom{3+2\sqrt{2}}{1}$  based at  $p'$  points outside of horizontal rectangle #0.

(3)  $\psi_3$  acts every horizontal cylinder by  $\binom{x}{y} \mapsto \binom{x-2y \pmod{2}}{y \pmod{1}}$ , where  $(x, y)$  are measured from the bottom left corner of the corresponding horizontal rectangle. So  $\psi_3[L_1(q, -\binom{3+2\sqrt{2}}{1})] = L_1(q, -\underline{w})$ .

(4)  $\psi_4 = D^{-1} \circ \phi^2$ , where  $\phi$  maps horizontal cylinders into vertical cylinders by rotating the horizontal rectangles  $90^\circ$  counterclockwise around the midpoint of the top side of the corresponding horizontal rectangle. Now

(i)  $\phi[L_1(q, -\underline{w})] = L_1(q, \binom{1}{-(1+2\sqrt{2})})$ ;

(ii)  $\phi[L_1(q, \binom{1}{-(1+2\sqrt{2})})] = L_1(q, \underline{w}) = L_2(q, \underline{w}) = L_2(D^2(p), \underline{w})$  (we moved up, because  $\underline{w}$  based at  $q$ , points outside of rectangle #1);

(iii)  $D^{-1}[L_2(D^2(p), \underline{w})] = L_1(D(p), \underline{w})$ .

Consequently,  $\psi_4[L_1(q, -\underline{w})] = L_1(D(p), \underline{w})$ .

In summary,  $\psi[L_0(p, \underline{w})] = L_1(D(p), \underline{w}) = D[L_0(p, \underline{w})]$ . So  $\tau_\psi(p, \underline{w}) = 1$ .  $\square$

## 7. CHARACTERIZATION OF GENERIC POINTS

In this section we leave the study of the deterministic random walk, and turn to a different problem: The description of the generic points of the cylinder map.

The cylinder map  $T_\alpha$  is ergodic and conservative with respect to the infinite invariant measure  $m_0 = m_{\mathbb{T}} \times m_{\mathbb{Z}}$ , for every  $\alpha$  irrational ( $[\mathbf{C}]$ ,  $[\mathbf{CK}]$ ,  $[\mathbf{Sch}]$ ,  $[\mathbf{AK}]$ ). By Hopf's ratio ergodic theorem, for every  $F, G \in L^1(\mathbb{T} \times \mathbb{Z})$  s.t.  $\int G dm_0 > 0$ ,

$$\frac{\sum_{j=0}^{n-1} (F \circ T_\alpha^j)(x, k)}{\sum_{j=0}^{n-1} (G \circ T_\alpha^j)(x, k)} \xrightarrow{n \rightarrow \infty} \frac{\int F dm_0}{\int G dm_0} \quad (7.1)$$

for  $m_0$ -almost every  $(x, k) \in \mathbb{T} \times \mathbb{Z}$ .

But (7.1) does not hold for every  $(x, k) \in \mathbb{T} \times \mathbb{Z}$ . This is because  $T_\alpha$  admits other ergodic conservative locally finite measures  $[\mathbf{Nkd}]$ . If  $\mu$  is one of the other measures, then the limit in (7.1) is  $\int F d\mu / \int G d\mu$   $\mu$ -almost everywhere, and not  $\int F dm_0 / \int G dm_0$ .

This raises the question what exactly is the domain of validity of (7.1). To state the problem in a meaningful way, we need the following definition.

**Definition 7.1.** *A point  $(x, k)$  is called generic (for  $T_\alpha$  and  $m_0$ ), if it satisfies (7.1) for every  $F, G \in C_c(\mathbb{T} \times \mathbb{Z})$  s.t.  $\int G dm_0 > 0$ .*

By the discussion above almost every point is generic, but some points are not generic. It was asked in  $[\mathbf{Sa}]$  what are the generic points of  $T_\alpha$ .

In this section we give the answer in the special case when  $\alpha$  is a quadratic irrational. Let  $\omega : \mathbb{T} \rightarrow \text{St}$  be as in Theorem 3.1.

**Theorem 7.2.** *Suppose  $\alpha$  is a quadratic irrational, with renormalizing automorphism  $\psi$ , then  $(x, k)$  is generic for  $T_\alpha$  and  $m_0 = m_{\mathbb{T}} \times m_{\mathbb{Z}}$  iff  $\frac{1}{k} \xi[\psi^k(\omega(x))] \xrightarrow{k \rightarrow \infty} 0$ .*

Let  $\binom{\sin \theta}{\cos \theta}$  denote the stable direction of  $\psi$ , and  $\varphi_\theta$  the linear flow in direction  $\theta$  on the infinite staircase. A point  $\omega \in \text{St}$  is called generic for  $\varphi_\theta$  and the area measure  $m$ , if  $\int_0^n F[\varphi_\theta^t(\omega)] dt / \int_0^n G[\varphi_\theta^t(\omega)] dt \xrightarrow{n \rightarrow \infty} \int F dm / \int G dm$  for every  $F, G \in C_c(\text{St})$  s.t.  $\int G dm > 0$ . We will obtain Theorem 7.2 from the following result.

**Theorem 7.3.**  $\omega \in \text{St}$  is generic for  $\varphi_\theta$  and  $m$  iff  $\frac{1}{k}\xi[\psi^k(\omega)] \xrightarrow[k \rightarrow \infty]{} 0$ .

Theorem 7.2 follows from Theorem 7.3 in the same way Theorem 3.1 follows from Theorem 3.2.

**Proof of Theorem 7.3.** If the theorem holds for one choice of a  $\mathbb{Z}$ -coordinate, then it works for all choices of  $\mathbb{Z}$ -coordinates, therefore we are free to use the  $\mathbb{Z}$ -coordinate of our choice. We choose the  $\mathbb{Z}$ -coordinate associated to  $\psi$ .

Suppose  $\xi[\psi^k(\omega_0)]/k \xrightarrow[k \rightarrow \infty]{} 0$ , then  $\omega_0$  is generic because of Theorem 3.2 : the fluctuating exponential term cancels out upon division.

The remainder of the proof deals with the implication “genericity  $\Rightarrow$  zero drift.” We use the strategy of [SS].

Fix a generic point  $\omega_0 \in \text{St}$ , let  $A_T := \{\varphi_t^i(\omega_0) : 0 \leq t \leq T\}$ , and define  $\lambda_T$  to be the normalized length measure on  $A_T \cap [\xi = 0]$ . Since  $\omega_0$  is generic,  $\lambda_T$  converges weak star to the normalized Lebesgue measure on  $[\xi = 0]$ .

*Construction:* Fix some  $N \geq 1$  to be chosen later, and define for every  $k \geq 0$

$$X_k^N := \xi \circ \psi^{(k+1)N} - \xi \circ \psi^{kN} = \left( \sum_{j=kN}^{(k+1)N-1} F_\psi \circ \psi^j \right) \circ \pi.$$

We think of  $X_k^N$  as of bounded random variables on  $(A_T, \mathcal{B}(A_T), \lambda_T)$ . The bound is  $|X_k^N| \leq N \max |F_\psi|$ .

Let  $\log^* := \log_{\lambda^{-1}}$ , where  $\lambda$  is the eigenvalue of the derivative of  $\psi$  in  $(0, 1)$ . Since  $\xi = 0$   $\lambda_T$ -a.e. and  $F_\psi$  is uniformly bounded,

$$\begin{aligned} \sum_{k=0}^{\frac{1}{N}[\log^* T]-2} X_k^N &= \sum_{k=0}^{\frac{1}{N}[\log^* T]-1} X_k^N + O(N) = \xi \circ \psi^{[\log^* T]} - \xi + O(N) \\ &= \xi \circ \psi^{[\log^* T]} + O(N) \text{ uniformly on } \text{supp}(\lambda_T). \end{aligned}$$

The right hand side is nearly constant  $\lambda_T$ -a.s., because  $\psi^{[\log^* T]}$  contracts the support of  $\lambda_T$  (a subset of  $A_T$ ) to a set of diameter less than  $\lambda^{-1}$ , and the  $\mathbb{Z}$ -coordinates of points in such a set must be uniformly bounded away from one another. It follows that for  $\lambda_T$ -a.e.  $\omega \in A_T$ ,

$$\sum_{k=0}^{\frac{1}{N}[\log^* T]-2} X_k^N(\omega) = \xi[\psi^{[\log^* T]}(\omega_0)] + O(N) \text{ uniformly in } T.$$

Taking expectations with respect to  $\lambda_T$  and dividing by  $[\log^* T]$ , we obtain

$$\frac{\xi[\psi^{[\log^* T]}(\omega_0)]}{[\log^* T]} = \mathbb{E}_{\lambda_T} \left( \frac{1}{[\log^* T]} \sum_{k=0}^{\frac{1}{N}[\log^* T]-2} X_k^N \right) + o(1), \text{ as } T \rightarrow \infty. \quad (7.2)$$

The expectation of  $\frac{1}{[\log^* T]} \sum_{k=0}^{\frac{1}{N}[\log^* T]-2} X_k^N$  with respect to the *normalized Lebesgue's measure* on  $[\xi = 0]$  is zero (because  $\psi$  is an automorphism with zero drift). We will use the genericity (in the form  $\lambda_T \xrightarrow[T \rightarrow \infty]{w^*} \text{Normalized Lebesgue}_{[\xi=0]}$ ) to show that something close happens to the  $\lambda_T$ -expectation. More precisely:

*Claim.* If  $\omega_0$  is generic then for every  $\varepsilon$  there exists  $N$  s.t. for all  $T$  large enough

$$\max\{|\mathbb{E}_{\lambda_T}(\frac{1}{N}X_k^N)| : 0 \leq k \leq \frac{1}{N}[\log^* T] - 2\} < \varepsilon. \quad (7.3)$$

Together with (7.2), this implies that  $\frac{1}{k}\xi_k(\omega_0) \xrightarrow[k \rightarrow \infty]{} 0$  and finishes the proof.

We begin the proof of (7.3). Fix  $\omega_0$  generic and  $\varepsilon > 0$ , and let  $C, \delta_0, N_0$  be some parameters that will be calibrated at the end of the proof. Let  $\lambda_T^C$  denote the length measure on  $A_T \cap [|\xi| \leq C]$ , normalized so that  $\lambda_T^C[\xi = 0] = 1$  (this is not a probability measure).

*Step 1.* Choosing  $N > N_0$  and  $\tau$  s.t.  $\lambda_T^C[|\frac{1}{N}X_0^N| > \delta_0] < \delta_0$  for all  $T > \tau$ .

*Proof.* Fix  $0 < \delta < \delta_0$ . Since  $\frac{1}{k}\xi \circ \psi^k \xrightarrow[k \rightarrow \infty]{} 0$  almost everywhere with respect to Lebesgue's measure, we can use Egoroff's theorem to find  $N = N(\delta, C) > N_0$  s.t.  $\Lambda_0 := \{\omega \in \text{St} : |\xi(\omega)| \leq C, |\frac{1}{N}X_0^N| > \delta_0\}$  satisfies  $m(\Lambda_0) < \delta$ , where  $m$  is the area measure on  $\text{St}$ .

Since  $\omega_0$  is generic,  $\lambda_T^C$  converges  $w^*$  to the Lebesgue measure on  $[|\xi| \leq C]$ . The indicator functions of  $\Lambda_0$  and  $[\xi = 0]$  are discontinuous, but all discontinuities lie on the boundaries of the parallelograms of the Markov partition, and their images under  $\psi^{-N}$ .  $N$  is fixed, therefore the closure of the singular set has measure zero. Thus there is no problem to show using a standard approximation argument that  $\lambda_T^C(\Lambda_0) \xrightarrow[T \rightarrow \infty]{} m(\Lambda_0) < \delta$ . It follows that there exists  $\tau = \tau(\delta, C)$  s.t. for all  $T > \tau(\delta, C)$ ,  $\lambda_T^C(\Lambda_0) < \delta$ , proving step 1.

Step 1 allows us to bound  $\mathbb{E}_{\lambda_T}(\frac{1}{N}X_0^N)$  as follows. Choose  $\delta_0 < \varepsilon/(1 + \max|F_\psi|)$ , then for all  $T > \tau$ ,

$$|\mathbb{E}_{\lambda_T}(\frac{1}{N}X_0^N)| \leq \delta_0 + \max|\frac{1}{N}X_0^N| \cdot \lambda_T[|\frac{1}{N}X_0^N| \geq \delta_0] \leq \delta_0(1 + \max|F_\psi|) < \varepsilon.$$

It is tempting to try to bound  $\mathbb{E}_{\lambda_T}(\frac{1}{N}X_k^N)$  for  $k \neq 0$  in the same way. Unfortunately the methods of step 1 can only be used for bounded  $k$ , whereas (7.3) calls for a uniform bound for  $0 \leq k \leq [\log^* T]$ , as  $T \rightarrow \infty$ .

We will take an indirect approach. Imagine we were able to construct self maps  $\theta_k : A_T \rightarrow A_T$  ( $0 \leq k \leq \frac{1}{N}[\log^* T] - 2$ ) with the following properties:

- $X_k^N = X_0^N \circ \theta_k + \text{error}$ , uniformly bounded by  $E_0$
- $\theta_k$  is Borel, one-to-one, and  $\theta_k(A_T \cap [\xi = 0]) \subset A_T \cap [|\xi| \leq C]$
- $C_r^{-1} \leq \frac{d\ell \circ \theta_k}{d\ell} \leq C_r$  where  $d\ell$  is the (Lebesgue) length measure and  $C_r$  is a global constant, independent of  $N, T, k$ , and  $\varepsilon$

Then it would follow that  $\lambda_T \leq C_r \lambda_T^C \circ \theta_k|_{[\xi=0]}$ , and if  $E_0/N_0 < \delta_0$  then

$$\begin{aligned} \lambda_T[|\frac{1}{N}X_k^N| > 2\delta_0] &\leq C_r(\lambda_T^C \circ \theta_k)(\{|\frac{1}{N}X_k^N| > 2\delta_0\} \cap [\xi = 0]) \\ &\leq C_r \lambda_T^C[|\frac{1}{N}X_k^N \circ \theta_k^{-1}| > \delta_0] \leq C_r \lambda_T^C[|\frac{1}{N}X_0^N| > \delta_0] < C_r \delta_0. \end{aligned}$$

This, and the fact that  $\sup|X_k^N|$  are uniformly bounded, is sufficient to bound  $\mathbb{E}_{\lambda_T}(|\frac{1}{N}X_k^N|)$  uniformly and prove (7.3).

In reality we do not know how construct  $\theta_k$  like that, because of edge effects at the endpoints of  $A_T$ . Luckily these edge effects can be controlled well enough to push this argument through with minor modifications. The details follow.

*Step 2.* Breaking  $A_T$  into the ‘‘interior’’ and ‘‘edge’’  $s$ -fibres.

We use the Adler–Weiss coding of Theorem 2.8 and Lemma 3.3. Recall that an  $s$ -fibre is a set of the form  $W^s(\underline{x}, k) := \text{lift to } [\xi = k] \text{ of } W^s(\underline{x}) := \{\pi_0(\underline{y}) : y_0^\infty = x_0^\infty\}$ . This is a stable linear segment, with length  $h(x_0)$ .

We define an  $n_0$ -stable block to be the lift to  $[\xi = k]$  of a set of the form  $\{\pi_0(\underline{y}) : y_{-n_0}^\infty = x_{-n_0}^\infty\}$ . These are closed stable linear segments, and their length is  $\lambda^{n_0} h(x_{-n_0}) \asymp \lambda^{n_0} \xrightarrow{n_0 \rightarrow \infty} 0$ . ( $A_n \asymp B_n$  means  $C^{-1} < A_n/B_n < C$  for all  $n$  large.)

Different  $n_0$ -stable blocks are disjoint, or they meet at one or two endpoints. Every point belongs to at most two  $n_0$ -stable blocks.

$B_T := \psi^{[\log^* T]}(A_T)$  is a stable linear segment with length in  $[1, \lambda^{-1}]$ . The  $n_0$ -stable blocks which intersect the relative interior of  $B_T$  fall into two groups:

- two or less “edge”  $n_0$ -stable blocks which cover the endpoints of  $B_T$ ;
- $(1 - 4\lambda^{n_0} \max h)/(\lambda^{n_0} \min h)$  or more “interior”  $n_0$ -stable blocks which lie completely inside  $B_T$ .

Since the number of interior  $n_0$ -blocks tends to infinity as  $n_0 \rightarrow \infty$ , it is possible to fix once and for all  $n_0$  in such a way that there is at least one interior block.

This choice is independent of  $N_0$ , therefore it is possible to assume without loss of generality that  $N_0 > n_0$ .

If  $[\log^* T] \gg n_0$ , then the decomposition of  $B_T = \psi^{[\log^* T]}(A_T)$  into interior and boundary  $n_0$ -blocks induces a decomposition of  $A_T$  into interior and edge  $s$ -fibres. Here and throughout,

- An *edge  $s$ -fibre* of  $A_T$  is an  $s$ -fibre in  $\psi^{-[\log^* T]}(\text{edge } n_0\text{-block of } B_T)$ .
- An *interior  $s$ -fibre* of  $A_T$  is a stable fibre in  $\psi^{-[\log^* T]}(\text{interior } n_0\text{-block of } B_T)$ .

Let  $\mathcal{W}_{int}, \mathcal{W}_{bnd}$  be the collection of interior and edge stable fibres in  $A_T$ . The *interior* of  $A_T$  is  $A_T(int) := \bigcup \mathcal{W}_{int}$ , and the *boundary* of  $A_T$  is  $A_T(bnd) := \bigcup \mathcal{W}_{bnd}$ .

*Step 3.* Defining  $\theta_k$  on the interior of  $A_T$  ( $0 \leq k \leq \frac{1}{N}[\log^* T] - 2$ ).

Recall that  $\mathcal{G}$  is the dynamical graph of the Markov partition  $\mathfrak{P}$ . An *admissible word* is a finite word  $w_0 \cdots w_n \in \mathfrak{P}^n$  s.t.  $w_0 \rightarrow \cdots \rightarrow w_n$  is a path on  $\mathcal{G}$ .

Since  $\Sigma(\mathcal{G})$  is topologically mixing, there exists a constant  $M_{br}$  such that for every pair of  $a, b \in \mathfrak{P}$  there is a path  $\underline{w}_{ab}$  on  $\mathcal{G}$  of length  $M_{br}$  s.t.  $a\underline{w}_{ab}b$  is admissible. Fix for such  $a, b$  a word  $\underline{w}_{ab}$ , and call it the *bridge* from  $a$  to  $b$ . In what follows

$$(a, \text{bridge}, b) := (a, \underline{w}_{a_{\text{last}} b_{\text{first}}}, b).$$

We define  $\theta_k : A_T(int) \rightarrow A_T$  as follows. Suppose  $\omega \in A_T(int)$ . Fix a stable fibre  $W^s(\underline{x}, \eta) \subset A_T(int)$ . All but countably many points in  $W^s(\underline{x}, \eta)$  can be uniquely represented in the form  $(\underline{y}, \eta)$  where

$$\underline{y} \in \Sigma(\mathcal{G}), y_0^\infty = x_0^\infty, \xi(\omega) = \eta, \text{ and } \pi(\omega) = \pi_0(\underline{y}).$$

Write  $\underline{y} = (y_{-\infty}^{-1} | \underline{B}_0, \underline{B}_1, \dots, \underline{B}_k, y_{kN}^\infty)$ , where  $\underline{B}_i$  are words of length of  $N$ , and the zeroth coordinate is to the immediate right of  $|$ . We let  $\theta_k(\underline{y}, \eta) := (\underline{z}, \eta')$  where

$$\begin{aligned} \underline{z} &= \sigma^{4M_{br}}(y_{-\infty}^{-1} | \text{bridge}, \underline{B}_k, \text{bridge}, \underline{B}_1, \dots, \underline{B}_{k-1}, \text{bridge}, \underline{B}_0, \text{bridge}, y_{(k+1)N}^\infty) \\ \eta' &= \eta + \sum_{j=0}^{[\log^* T]-1} F(\sigma^j \underline{y}) - F(\sigma^j \underline{z}), \text{ where } F := F_\psi \circ \pi_0. \end{aligned}$$

What we have done here is to exchange block zero with block  $k$ , plug-in bridge words to ensure admissibility, and at the end apply the shift and modify  $\eta$  to ensure that we remain inside  $A_T$  (see below). Here are some properties of  $\theta_k$ .

(i)  $X_0^N \circ \theta_k = X_k^N + \text{bounded error}$ :  $\theta_k$  exchanges the zeroth block with the  $k$ -th block, and therefore  $|X_0^N \circ \theta_k - X_k^N| \leq 6M_{br} \max |F_\psi|$ . Here we are using (2.4).

(ii)  $\theta_k[A_T(int)] \subset A_T(int)$  for all  $0 \leq k \leq \frac{1}{N}[\log^* T] - 2$ : We abuse notation and identify a point in  $\text{St}$  with its symbolic coding in  $\Sigma(\mathcal{G}) \times \mathbb{Z}$ . Suppose  $(\underline{y}, \eta)$  belongs to an interior  $n_0$ -block and  $0 \leq k \leq \frac{1}{N}[\log^* T] - 2$ , then

$$\begin{aligned} (\sigma^{[\log^* T]} \underline{z})_{-N}^\infty &= \sigma^{[\log^* T]}(\underline{y})_{-N}^\infty, \text{ because of the } \sigma^{4M_{br}} \text{ in the definition of } \underline{z}, \text{ and} \\ \psi^{[\log^* T]}(\underline{z}, \eta') &= (\sigma^{[\log^* T]}(\underline{z}), \eta' + \sum_{j=0}^{[\log^* T]-1} F(\sigma^j \underline{z})) \\ &= (\sigma^{[\log^* T]}(\underline{z}), \eta' + \sum_{j=0}^{[\log^* T]-1} F(\sigma^j \underline{y})), \text{ by the definition of } \eta'. \end{aligned}$$

Since  $N > n_0$ ,  $\psi^{[\log^* T]}(\underline{z}, \eta')$  belongs to the same  $n_0$ -block which contains  $\psi^{[\log^* T]}(\underline{y}, \eta)$ . This is an interior  $n_0$ -block of  $B_T = \psi^{[\log^* T]}(A_T)$ . So  $(\underline{z}, \eta') \in A_T$ .

(iii)  $\theta_k[\text{supp}(\lambda_T)] \subset \text{supp}(\lambda_T^{C'})$  for  $C' := 100[M_{br} + 1] \max |F_\psi|$ : It is enough to check that  $|\eta' - \eta| = \left| \sum_{j=0}^{[\log^* T]-1} F(\sigma^j \underline{z}) - \sum_{j=0}^{[\log^* T]-1} F(\sigma^j \underline{y}) \right| \leq C'$ . To check this we recall that  $F$  is constant on 2-cylinders, therefore the difference between the two sums can only come from the following sources:

- the effect of the shift by  $\sigma^{4M_{br}}$ , bounded by  $2 \cdot 4M_{br} \max |F_\psi|$
- the sum over the bridge words, bounded by  $4M_{br} \max |F_\psi|$
- the value of  $F(\sigma^j \underline{z})$  for the  $j$  at the end of  $\underline{B}_k$ ,  $\underline{B}_{k-1}$ , and  $\underline{B}_0$ , with a total effect bounded by  $3 \max |F_\psi|$

This gives the bound above with  $C$  much smaller than claimed.

(iv)  $\theta_k$  is one-to-one on  $A_T(int)$ : To reconstruct  $\underline{y}$  from  $\underline{z}$  one just needs to erase the bridge words (whose position is always the same), and then exchange the  $k$ -th block with the zeroth block of what remains. Once  $\underline{y}$  is known,  $\eta$  can be easily calculated from  $\eta'$  and  $\underline{z}$ .

(v)  $d\ell \circ \theta_k / d\ell$  is uniformly bounded away from 0,  $\infty$ , where  $\ell$  is the length measure on  $A_T$ : For every  $\underline{a} \in \Sigma(\mathcal{G})$ , the length of the stable linear segment

$$[a_{-n}^\infty] := \{\omega \in A_T(int) \cap [\xi = 0] : \pi(\omega) = \pi_0(\underline{y}), y_j = a_j \ (j \geq -n)\}$$

is  $\lambda^n h(a_{-n})$ , because  $\psi^{-n}[a_{-n}^\infty] = W^s(\sigma^{-n}\underline{a})$ ,  $\ell[W^s(\sigma^{-n}\underline{a})] = h(a_{-n})$ , and  $\psi^{-n}$  expands linear stable segments by a factor of  $\lambda^{-n}$ . With this formula at hand, it is easy to see that  $\frac{d\ell \circ \theta_k}{d\ell} \in [\lambda^{4M_{br}}(\frac{\min h}{\max h}), \lambda^{4M_{br}}(\frac{\max h}{\min h})]$ .

*Step 4.* Defining  $\theta_k$  on  $A(bnd)$ .

Fix once and for all an interior stable fibre  $W^s(\underline{x}, \eta)$  of  $A_T$ . We will define  $\theta_k$  on an edge stable fibre by first mapping it into  $W^s(\underline{x}, \eta)$ , and then applying  $\theta_k|_{A(int)}$  as defined in step 3.

The resulting transformation is  $\theta_k : (y, \eta) \mapsto (z, \eta')$  where

$$z = \sigma^{5M_{br}}(y_{-\infty}^{-1} | \text{bridge}, \underline{B}_k, \text{bridge}, \underline{B}_1, \dots, \underline{B}_{k-1}, \text{bridge}, \underline{B}_0, \text{bridge}, y_{kN}^{[\log^* T]^{-1}}, \text{bridge}, x_{[\log^* T]}^{\infty})$$

$$\eta' = \eta + \sum_{j=0}^{[\log^* T]} F_{\psi}(\sigma^j \underline{y}) - F_{\psi}(\sigma^j \underline{z})$$

The following properties can be verified as in the previous step:

- (i)  $X_0^N \circ \theta_k = X_k^N + \text{bounded error}$ . The error is bounded by  $8M_{br} \max |F_{\psi}|$ .
- (ii)  $\theta_k[A(bnd)] \subset A_T(int)$  and  $\theta_k[A(bnd)] \cap [\xi = 0] \subset A_T(bnd) \cap [|\xi| \leq C']$  with  $C'$  as above.
- (iii)  $\theta_k$  is one-to-one on each boundary stable fibre, and its Radon-Nikodym derivative takes values in  $[\lambda^{5M_{br}}(\frac{\min h}{\max h}), \lambda^{5M_{br}}(\frac{\max h}{\min h})]$ .

*Step 5. Proof of (7.3).*

We estimate  $\mathbb{E}_{\lambda_T}(\frac{1}{N} X_k^N)$  for  $0 \leq k \leq \frac{1}{N}[\log_{\lambda^{-1}} T] - 2$ . Since  $|X_k^N| \leq N \max |F_{\psi}|$ ,

$$|\mathbb{E}_{\lambda_T}(\frac{1}{N} X_k^N)| \leq 2\delta_0 + \max |F_{\psi}| \cdot \lambda_T(\Lambda_k), \quad (7.4)$$

where  $\Lambda_k := \{\omega \in S : \xi(\omega) = 0, |\frac{1}{N} X_k^N| > 2\delta_0\}$ . We will bound  $\lambda_T(\Lambda_k)$ .

First we choose  $C, N_0, \delta_0$ :  $C := C' = 100(M_{br} + 1) \max |F_{\psi}|$ ,  $N_0$  so large that  $\frac{10M_{br} \max |F_{\psi}|}{N_0} < \delta_0$ , and  $\delta_0 := \varepsilon / (2 + 3\lambda^{-5M_{br}}(\frac{\max h}{\min h}) \max |F_{\psi}|)$ .

By construction,  $|X_k^N - X_0^N \circ \theta_k| \leq 8M_{br} \max |F_{\psi}|$ , so if  $|\frac{1}{N} X_k^N(\omega)| > 2\delta_0$  then  $|\frac{1}{N} X_0^N(\theta_k(\omega))| > 2\delta_0 - \frac{10M_{br} \max |F_{\psi}|}{N_0} > \delta_0$ . We see that

$$\theta_k(\Lambda_k) \subset \{\omega \in A_T : |\xi(\omega)| \leq C, |\frac{1}{N} X_0^N| > \delta_0\}.$$

Applying  $\theta_k^{-1}$  to both sides and recalling that  $\theta_k$  is piecewise invertible, at worst three-to-one (edge effects), and  $\frac{d\xi \circ \theta_k}{d\xi} \geq \lambda^{5M_{br}}(\frac{\min h}{\max h})$  we find that

$$\lambda_T(\Lambda_k) \leq 3\lambda^{-5M_{br}}(\frac{\max h}{\min h}) \lambda_T^C[|\frac{1}{N} X_0^N| > \delta_0].$$

By step 1,  $\lambda_T(\Lambda_k) \leq 3\lambda^{-5M_{br}}(\frac{\max h}{\min h}) \delta_0$ . Substituting this at (7.4) we find that  $|\mathbb{E}_{\lambda_T}(\frac{1}{N} X_k^N)| \leq \delta_0(2 + 3\lambda^{-5M_{br}}(\frac{\max h}{\min h}) \max |F_{\psi}|) < \varepsilon$  as required.  $\square$

#### APPENDIX A. PROOFS OF PROPOSITION 2.3, LEMMA 2.7, AND LEMMA 2.9

**Classification of homogeneous automorphisms.** We prove Proposition 2.3:

- (1) If  $A \in \text{SL}(2, \mathbb{Z})$ ,  $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{mod } 2$  and  $\delta_0 \in \mathbb{Z}$ , then there is a unique homogeneous automorphism with derivative  $A$  and drift  $\delta_0$ .
- (2) If  $A \in \text{SL}(2, \mathbb{Z})$ ,  $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \text{mod } 2$  and  $\delta_0 \in \frac{1}{2} + \mathbb{Z}$ , then there is a unique homogeneous automorphism with derivative  $A$  and drift  $\delta_0$ .
- (3) No other homogeneous automorphisms exist.

*Step 1.* Existence of homogeneous automorphisms as in parts 1 and 2.

*Proof.* Let  $\Gamma(2) := \{A \in \text{SL}(2, \mathbb{Z}) : A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{mod } 2\}$ ,  $\Gamma := \Gamma(2) \cup \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \Gamma(2)$ .  $\Gamma(2)$  is generated by  $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ ,  $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$ , and  $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ . We will construct

- A homogeneous automorphism with derivative  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  and drift  $+1$ ;
- A homogeneous automorphism with derivative  $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$  and drift  $0$ ;

- A homogeneous automorphism with derivative  $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$  and drift 0;
- A homogeneous automorphism with derivative  $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$  and drift 0;
- A homogeneous automorphism with derivative  $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$  and drift  $\frac{1}{2}$ .

The automorphisms in part (1) can be constructed from the first three automorphisms. The automorphisms in part (2) require an additional composition with the fourth automorphism.

*A homogeneous automorphism with derivative  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  and drift 1:  $D$*

*A homogeneous automorphism with derivative  $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$  and drift 0:* Divide St into horizontal cylinders, as indicated by the dashed lines in figure 5(a). Act on each cylinder by the map  $(x, y) \mapsto (x + 2y \bmod 2, y)$  up to identifications, where  $(x, y)$  are measured relative to the bottom left corner. These maps equal the identity on the boundary of the cylinder (the horizontal sides of the rectangle), therefore they glue to an automorphism  $\psi$ . The derivative of  $\psi$  is  $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ ,  $\psi$  commutes with  $D$ , and  $\psi$  fixes the singularities of St. It has zero drift, because the Frobenius function with respect to the horizontal rectangles vanishes.

*A homogeneous automorphism with derivative  $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$  and drift 0:* The same construction, but using the decomposition of St into vertical cylinders (figure 5(a)).

*A homogeneous automorphism with derivative  $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$  and drift  $\frac{1}{2}$ :* Decompose St into horizontal rectangles. Rotate every rectangle 90 degrees counterclockwise around the midpoint of its top side, turning it into a vertical rectangle. These maps glue continuously to an automorphism of St (figure 5(b)). Using the  $\mathbb{Z}$ -coordinate defined by the horizontal rectangles, one sees that the average drift is  $\frac{1}{2}$ .

*A homogeneous automorphism with derivative  $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$  and drift 0:* Suppose  $\psi$  is the automorphism with drift  $\frac{1}{2}$  and derivative  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$  constructed above, then  $\varphi := D^{-1} \circ \psi^2$  has zero drift and derivative  $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ .

*Step 2.* A homogeneous automorphism with derivative  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  is equal to  $D^k$  for some  $k \in \mathbb{Z}$ . Two homogeneous automorphisms with the same derivative and drift are equal.

*Proof.* Let  $\psi$  be a homogeneous automorphism with derivative  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ . Fix a horizontal rectangle  $R$  and let  $p_0$  denote the singularity at its lower right corner. Given  $p \in \text{int}(R)$ , let  $\gamma_p \subset R$  denote the linear segment from  $p_0$  to  $p$ . Since  $\psi$  fixes the  $D$ -orbit of  $p_0$ ,  $\psi[\gamma_p]$  is a linear segment from  $p_0$  or  $D(p_0)$  to  $\psi(p)$ . Since the derivative of  $\psi$  is  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ,  $\psi[\gamma_p]$  has the same slope, length, and direction as  $\gamma_p$ .

There are infinitely many such segments, one for every horizontal rectangle  $D^k(R)$ ,  $k \in \mathbb{Z}$ . By reasons of continuity there is some fixed  $k \in \mathbb{Z}$  such that  $\psi[\gamma_p] \subset D^k(R)$  for all  $p \in \text{int}(R)$ . It follows that  $\psi = D^k$  on  $R$ . Since  $\psi$  commutes with  $D$ ,  $\psi = D^k$  on St. This proves the first part of step 2.

For the second part, suppose  $\psi_1, \psi_2$  have the same derivative and drift, then  $\psi_1 \circ \psi_2^{-1}$  has derivative  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  and drift zero (Lemma 2.2). So  $\psi_1 \circ \psi_2^{-1} = D^k$  with  $k = \text{drift} = 0$ .

*Step 3.* For every homogeneous automorphism, either the derivative is  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bmod 2$  and the drift is in  $\mathbb{Z}$ , or the derivative is  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \bmod 2$  and the drift is in  $\frac{1}{2} + \mathbb{Z}$ .

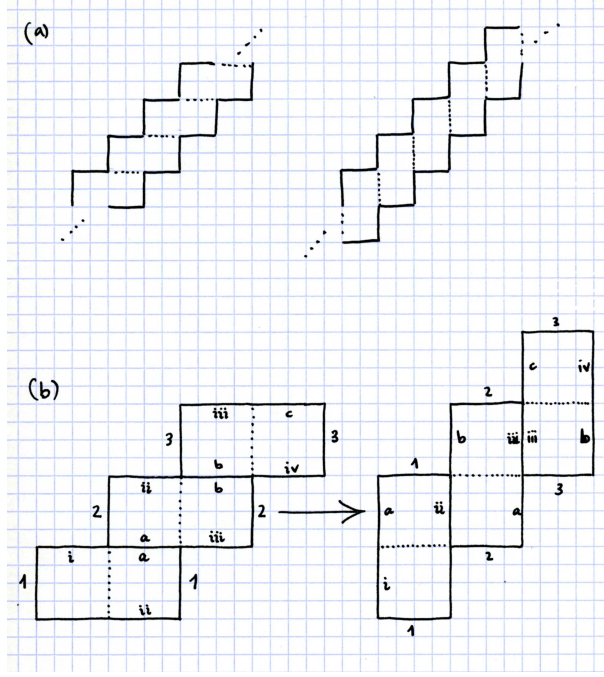


FIGURE 5. (a) decomposition of  $St$  into horizontal and vertical cylinders; (b) a homogeneous automorphism with drift  $\frac{1}{2}$

*Proof.* Suppose  $\psi$  is a homogeneous automorphism, and let  $\psi_0 : St_0 \rightarrow St_0$  be the projection of  $\psi$  to  $St_0$ .

$St_0 = \mathbb{R}^2/G$  where  $G$  is generated by the translations by  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$  and  $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ , see figure 1(c). The change of coordinates  $\Theta \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$  gives the identification  $St_0 \simeq \mathbb{R}^2/\sqrt{2}\mathbb{Z}^2$ . In these coordinates, the punctures are  $\sqrt{2}\mathbb{Z}^2$  and  $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \sqrt{2}\mathbb{Z}^2$ , and  $\psi_0$  is represented by  $\widehat{\psi}_0 := \Theta \circ \psi_0 \circ \Theta$ .

Let  $B$  denote the derivative of  $\widehat{\psi}_0$ . The linear segments in  $\mathbb{R}^2$  from  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  to  $\sqrt{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ,  $\sqrt{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$  project to closed curves  $\gamma_1, \gamma_2$  on  $\mathbb{R}^2/\sqrt{2}\mathbb{Z}^2$ .  $\widehat{\psi}_0$  fixes  $\sqrt{2}\mathbb{Z}^2$  (a singularity), therefore if we apply  $\widehat{\psi}_0$  to  $\gamma_1, \gamma_2$ , and lift the result to  $\mathbb{R}^2$  at  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ , then we get linear segments from  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  to  $B\begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}$  and  $B\begin{pmatrix} 0 \\ \sqrt{2} \end{pmatrix}$ . These segments project to the closed



curves  $\widehat{\psi}_0[\gamma_1], \widehat{\psi}_0[\gamma_2]$  on  $\mathbb{R}^2/\sqrt{2}\mathbb{Z}^2$ , so necessarily  $B(\sqrt{2}), B(\frac{0}{\sqrt{2}}) = \binom{0}{0} \bmod \sqrt{2}\mathbb{Z}^2$ . It follows that  $B \in \text{GL}(2, \mathbb{Z})$ .

Actually,  $B \in \text{SL}(2, \mathbb{Z})$ :  $|\det B| = 1$  because  $\widehat{\psi}_0$  is an orientation preserving self-bijection of a surface of finite area, and  $\det B > 0$  because  $\widehat{\psi}_0$  is orientation preserving.

Next we use the fact that  $\widehat{\psi}_0$  fixes  $\frac{1}{\sqrt{2}}\binom{1}{1} + \sqrt{2}\mathbb{Z}^2$ . Using the linear segment from  $\binom{0}{0}$  to  $\frac{1}{\sqrt{2}}\binom{1}{1}$  as above, we see that  $B(\frac{1}{\sqrt{2}}) \in (\frac{1}{\sqrt{2}}) + \sqrt{2}\mathbb{Z}^2$ . Multiplying by  $\sqrt{2}$  and considering the result modulo 2 we see that the rows of  $B$  have entries of different parity. So  $B \pmod{2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ , or  $\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \pmod{2}$ . Since  $\det B$  is odd,  $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  or  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \pmod{2}$ .

Returning to  $\psi_0 = \Theta \circ \widehat{\psi}_0 \circ \Theta$ , we see by direct calculation that the derivative of  $\psi$  also has entries with different parity at every row. As before, this means that the derivative of  $\psi$  is  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  or  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \pmod{2}$ .

Suppose the derivative of  $\psi$  is  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{2}$ . We saw in step one that there exists a homogeneous automorphism  $\phi$  with the same derivative and with drift zero. By step two,  $\psi \circ \phi^{-1} = D^k$  for some  $k \in \mathbb{Z}$ . It follows that  $\delta(\psi) = k \in \mathbb{Z}$ .

If the derivative of  $\psi$  is  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \pmod{2}$ , then there is a homogeneous automorphism  $\phi$  with the same derivative and with drift  $\frac{1}{2}$ . By step two,  $\psi \circ \phi^{-1} = D^k$  for some  $k \in \mathbb{Z}$ , and  $\delta(\psi) = k + \frac{1}{2}$ .  $\square$

**Proof of Lemma 2.7.** Suppose  $\xi, \eta \in \mathbb{R} \setminus \mathbb{Q}$  and  $\xi \neq \eta$ . We are asked to produce a matrix  $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(2)$  such that  $s_1 := \frac{a\xi+b}{c\xi+d}$  and  $s_2 := \frac{a\eta+b}{c\eta+d}$  satisfy one of the following: One of  $s_1, s_2$  is in  $(-1, 0)$  and the other is in  $(1, \infty)$  (“case 1”); Or one of  $s_1, s_2$  is in  $(0, 1)$  and the other is in  $(1, \infty)$  (“case 2”).

Let  $\mathbb{H} := \{z \in \mathbb{C} : \text{Im}(z) > 0\}$ , then  $\Gamma(2)$  acts on the upper half plane by  $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az+b}{cz+d}$ . It is well known that the hyperbolic polygon  $F$  with vertices  $-1, 0, 1, \infty$  is a fundamental domain for this action [Fo]. So  $\{g(F) : g \in \Gamma(2)\}$  is a tessellation of  $\mathbb{H}$ . Notice that the vertices of  $g(F)$  belong to  $\mathbb{Q} \cup \{\infty\}$ .

Label the sides of  $F$  on the inside by  $\{a, \bar{a}, b, \bar{b}\}$  as in figure 6. Notice that  $a$  is mapped to  $\bar{a}$  by  $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ , and  $b$  is mapped to  $\bar{b}$  by  $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$ . Extend the labeling to  $g(F)$  ( $g \in \Gamma(2)$ ) in the natural way. Now every side in the tessellation has two labels  $x, \bar{x}$ , one internal and the other external (which is which depends on the tile we use as reference).

Let  $\gamma$  denote the (open) upper half of the circle with diameter  $[\xi, \eta]$  or  $[\eta, \xi]$ . We think of  $\gamma$  as of a geodesic in the upper half plane, from  $\xi$  to  $\eta$ . Let  $\underline{x} = (\dots, x_{-1}, x_0, x_1, \dots)$  denote the ordered sequence of *internal* labels of the sides of the tiles  $\gamma$  enters. The position of the zeroth coordinate is not important. The following facts follow from the geometric structure of the tessellation:

- (1)  $\underline{x}$  is a doubly infinite (otherwise  $\xi$  or  $\eta$  is a vertex of  $g(F)$  for some  $g \in \Gamma(2)$ , in contradiction to the irrationality of  $\xi, \eta$ ).
- (2) For all  $i$ ,  $x_i \neq \bar{x}_{i+1}$  and  $\bar{x}_i \neq x_{i+1}$ .
- (3)  $\underline{x}$  does not begin or terminate with a constant ray (otherwise  $\xi$  or  $\eta$  is a vertex of  $g(F)$  for some  $g \in \Gamma(2)$ , in contradiction to their irrationality).

Suppose first that  $\underline{x}$  contains the symbol  $\bar{a}$ . Then it must contain  $\bar{a}b$  or  $\bar{a}\bar{b}$  (otherwise it terminates with the constant sequence  $\bar{a}\bar{a}\dots$ ). If  $\underline{x}$  does not contain

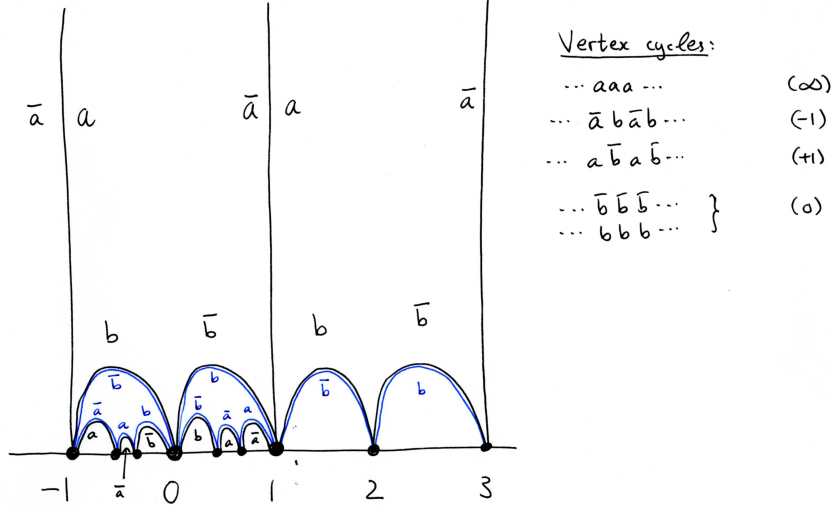


FIGURE 6. Fundamental domain of  $\Gamma(2)$

$\bar{a}$ , then it must contain  $b$  or  $\bar{b}$  (otherwise it equals  $\dots aaa\dots$ ). If  $\underline{x}$  contains  $b$  but not  $\bar{a}$ , then  $\underline{x}$  contains  $ba$  (otherwise it terminates with  $bbb\dots$ ). If  $\underline{x}$  contains  $\bar{b}$  but not  $\bar{a}$  then  $\underline{x}$  contains  $\bar{b}a$  (otherwise it terminates with  $\bar{b}\bar{b}\bar{b}\dots$ ). In summary,  $\underline{x}$  must contain at least one of the words  $\bar{a}b, \bar{b}a, \bar{a}\bar{b}, ba$ .

Notice that if the cutting sequence of the geodesic from  $\xi$  to  $\eta$  is  $(x_i)_{i \in \mathbb{Z}}$ , then the cutting sequence of the geodesic from  $\eta$  to  $\xi$  is  $(\bar{x}_{-i})_{i \in \mathbb{Z}}$ . Therefore we may assume without loss of generality that  $\underline{x}$  contains  $\bar{a}b$  or  $\bar{a}\bar{b}$ , otherwise exchange  $\xi \leftrightarrow \eta$ .

Suppose first that  $\underline{x}$  contains the word  $\bar{a}b$ , then  $\gamma$  enters some tile  $F^*$  through side  $\bar{a}$  and leaves it through side  $b$  (entering an adjacent tile with side  $b$ ). There is  $P \in \Gamma(2)$  which maps  $F^*$  onto  $F$ . Since  $P \cdot \gamma$  enters  $F$  through side  $\bar{a}$ ,  $P \cdot \xi \in (1, \infty)$ . Since  $P \cdot \gamma$  leaves  $F$  through side  $b$  (entering the adjacent tile through side  $b$ ),  $P \cdot \eta \in (0, 1)$ . This is case 1.

Next suppose  $\underline{x}$  contains the word  $\bar{a}\bar{b}$ , then  $\gamma$  enters some tile  $F^*$  through side  $\bar{a}$  and leaves it through side  $\bar{b}$  (entering an adjacent tile with side  $\bar{b}$ ). There is  $P \in \Gamma(2)$  which maps  $F^*$  onto  $F$ . Since  $P \cdot \gamma$  enters  $F$  through side  $\bar{a}$ ,  $P \cdot \xi \in (1, \infty)$ . Since  $P \cdot \gamma$  leaves  $F$  through side  $b$ ,  $P \cdot \eta \in (-1, 0)$ . This is case 2.  $\square$

**Proof of Lemma 2.9 (Aperiodicity Lemma).** Let  $F := F_\psi \circ \pi_0$ . We have to show that if  $e^{itF} = \lambda h/h \circ \sigma$  for some  $t \in \mathbb{R}$ ,  $\lambda \in S^1$ , and  $h : \Sigma(\mathcal{G}) \rightarrow S^1$  continuous, then  $t \in 2\pi\mathbb{Z}$ ,  $\lambda = 1$  and  $h = \text{const}$ .

*Proof.* We first consider the special case when  $\psi$  fixes all the singularities of  $\text{St}$ .

It is enough to find  $\underline{x}, \underline{y} \in \Sigma(\mathcal{G})$  such that for some  $n$ ,

$$\sigma^n(\underline{x}) = \underline{x}, \sigma^n(\underline{y}) = \underline{y}, \text{ and } F_n(\underline{x}) - F_n(\underline{y}) = \pm 1 \text{ where } F_n = \sum_{k=0}^{n-1} F \circ \sigma^k.$$

Indeed, if  $e^{itF} = \lambda h/h \circ \sigma$ , then  $e^{itF_n} = \lambda^n h/h \circ \sigma^n$ , whence  $e^{itF_n(\underline{x})} = \lambda^n$  and  $e^{itF_n(\underline{y})} = \lambda^n$ . Dividing, we find that  $e^{it} = 1$ , whence  $t \in 2\pi\mathbb{Z}$ . But if  $t \in 2\pi\mathbb{Z}$  then  $\lambda h/h \circ \sigma = e^{itF} = 1$ , whence  $h$  is a continuous eigenfunction of  $\sigma$ . Since  $\sigma : \Sigma(\mathcal{G}) \rightarrow \Sigma(\mathcal{G})$  is topologically mixing,  $\lambda = 1$  and  $h = \text{const}$ .

We will now construct such  $\underline{x}, \underline{y}$ . It is enough to do this in the case when  $\psi$  has a fundamental polygon of the form  $R = \theta_0(R_0)$ , where  $\theta_0 : \text{St}_0 \rightarrow \text{St}_0$  is a toral automorphism which fixes the punctures of  $\text{St}_0^*$ , and  $R_0$  is as in figure 2.

*Case (b).* The slope of the unstable direction is bigger than one, and the slope of the stable direction belongs to  $(-1, 0)$ .

Let  $\partial^s Q_i, \partial^u Q_i$  denote the stable and unstable boundaries of  $Q_i$ . In case (b),  $\partial^s Q_2 \ni (1, 0) =: p_0$  and  $\partial^u Q_2 \ni (2, 0) =: q_0$ . These are fixed points of  $\psi_0$ . Their lifts to  $\text{St}$  are fixed points of  $\psi$  (by assumption).

Recall that  $\psi_0(Q_2)$  is the union of parallelograms  $Q_{2,k}$ ,  $k = 1, \dots, N_2$ , where the bottom stable side of  $Q_{2,1}$  is part of the bottom stable side of  $Q_2$ , and the bottom stable side of  $Q_{2,k+1}$  is the top stable side of  $Q_{2,k}$ ,  $k = 1, \dots, N_2 - 1$ .

Using the fixed point  $p_0$  and the relation  $\psi_0(\partial^s Q_2) \subset Q_2$ , it is easy to see that  $Q_{2,1} \subset Q_2$ ,  $Q_{2,1} \rightarrow Q_{2,1}$ , and  $g(Q_{2,1}, Q_{2,1}) = 0$ . So  $\underline{x} = (\dots, Q_{2,1}, Q_{2,1}, Q_{2,1}, \dots)$  is a well defined point in  $\Sigma(\mathcal{G})$ ,  $\sigma^n(\underline{x}) = \underline{x}$ , and  $F_n(\underline{x}) = 0$  for all  $n$ . [Caution: for other  $k$ ,  $Q_{2,k}$  is not necessarily in  $Q_2$ .]

Similarly, some  $Q_{i,j} \subset Q_2$  contains  $q_0$  in its  $u$ -boundary, and  $Q_{i,j} \rightarrow Q_{i,j}$  with  $g(Q_{2,1}, Q_{i,j}) = 0$ . We claim that  $i = 2$  and  $1 < j < N_2$ :

- $q_0 = \psi_0^{-1}(q_0) \subset \psi_0^{-1}(\partial^u Q_{i,j}) \subset \partial^u Q_i$ . In case 1, this forces  $i = 2$ .
- The  $u$ -side of  $Q_{2,j} \cap \psi_0^{-1}(Q_{2,j})$  which contains  $q_0$  equals  $\psi_0^{-1}[W^u(q_0)]$ , where  $W^u(q_0)$  is the  $u$ -fibre of  $q_0$ . Since  $\psi_0$  is expanding on  $W^u(q_0)$ , the  $u$ -side of  $Q_{2,j} \cap \psi_0^{-1}(Q_{2,j})$  does not meet the endpoints of  $W^u(q_0)$ . So  $1 < j < N_2$ .

We see that  $Q_{2,j+1}$  is well defined, and that this is the parallelogram which follows the rightmost  $\mathfrak{P}$ -element in  $Q_2$ .

Looking at figure 2, we see that  $Q_{2,j+1} \subset Q_2$ ,  $g(Q_{2,1}, Q_{2,j+1}) = -1$ , and  $g(Q_{2,j+1}, Q_{2,1}) = 0$ .

We now define  $\underline{y} := (\dots, Q_{2,1}, Q_{2,j+1}; Q_{2,1}, Q_{2,j+1}; \dots)$ , then  $\underline{y} \in \Sigma(\mathcal{G})$ ,  $\sigma^2(\underline{y}) = \underline{y}$  and  $F_2(\underline{y}) = -1$ . Using  $\underline{x}, \underline{y}$  and  $n = 2$ , we get the aperiodicity of  $F$  in case 1.

*Case (a).* The slope of the unstable direction is bigger than one, and the slope of the stable direction is in  $(0, 1)$ .

Just like in case (b), the  $\mathfrak{P}$ -element in  $Q_2$  which contains  $p_0$  in its bottom  $s$ -side is  $Q_{2,1}$ , and  $Q_{2,1} \rightarrow Q_{2,1}$  with  $g(Q_{2,1}, Q_{2,1}) = 0$ . So  $\underline{x} = (\dots, Q_{2,1}, Q_{2,1}, Q_{2,1}, \dots)$  belongs to  $\Sigma$ ,  $\sigma(\underline{x}) = \underline{x}$ , and  $F(\underline{x}) = 0$ .

To construct  $\underline{y}$ , we separate cases according to whether  $Q_{2,2} \subset Q_1$  or  $Q_{2,2} \subset Q_2$ .

Suppose first that  $Q_{2,2} \subset Q_2$ . Looking at figure 2 and noting that  $p_0 \in \partial^s Q_2$ , we see that  $Q_{2,1} \subset Q_2$ , and that every  $\mathfrak{P}$ -element  $P$  in  $Q_2$  satisfies  $P \rightarrow Q_{2,1}$  and  $g(P, Q_{2,1}) = 0$ . In particular  $Q_{2,2} \rightarrow Q_{2,1}$  and  $g(Q_{2,2}, Q_{2,1}) = 0$ . Since  $Q_{2,1} \subset Q_2$ ,  $Q_{2,1} \rightarrow Q_{2,2}$ . Using the assumption that the slope of the unstable direction is

bigger than one, it is not difficult to see that  $g(Q_{2,1}, Q_{2,2}) = 1$ . We now set

$$\underline{y} = (\cdots; Q_{2,1}, Q_{2,2}; Q_{2,1}, Q_{2,2}; \cdots).$$

This is a point in  $\Sigma(\mathcal{G})$ ,  $\sigma^2(\underline{y}) = \underline{y}$ , and  $F_2(\underline{y}) = 1$ . Using  $\underline{x}$  and  $\underline{y}$  and  $n = 2$ , we get the aperiodicity of  $F$ , assuming that  $Q_{2,2} \subset Q_2$ .

Now consider the case that  $Q_{2,2} \subset Q_1$ . The following observations follow from figure 2 and the fact that  $p_0 \in \partial^s Q_2$ :

- As before,  $Q_{2,1} \rightarrow Q_{2,2}$ ,  $Q_{2,2} \subset Q_1$  and  $g(Q_{2,1}, Q_{2,2}) = 1$ .
- $Q_{1,1} \subset Q_2$ , and every  $\mathfrak{P}$ -element  $P$  in  $Q_1$  satisfies  $P \rightarrow Q_{1,1}$ ,  $g(P, Q_{1,1}) = 0$ . In particular,  $Q_{2,2} \rightarrow Q_{1,1}$ , and  $g(Q_{2,2}, Q_{1,1}) = 0$ .
- All  $\mathfrak{P}$ -elements  $P$  in  $Q_2$  satisfy  $P \rightarrow Q_{2,1}$  with  $g(P, Q_{2,1}) = 0$ . In particular,  $Q_{1,1} \rightarrow Q_{2,1}$  and  $g(Q_{1,1}, Q_{2,1}) = 0$ .

We now let  $\underline{y} := (\cdots; Q_{2,1}, Q_{2,2}, Q_{1,1}; Q_{2,1}, Q_{2,2}, Q_{1,1}; \cdots)$ . This is a point in  $\Sigma(\mathcal{G})$  s.t.  $\sigma^3(\underline{y}) = \underline{y}$  and  $F_3(\underline{y}) = 1$ . Using  $\underline{x}, \underline{y}$  and  $n = 3$ , we see that  $F$  is aperiodic.

This proves the lemma in case  $\psi$  fixes the singularities of  $\text{St}$ . The general case can be reduced to this case as follows.

Suppose  $\psi$  is a homogeneous automorphism, and assume that  $F := F_\psi \circ \pi_0$  satisfies  $e^{itF} = \lambda h/h \circ \sigma$  for some  $t \in \mathbb{R}$ ,  $\lambda \in S^1$ , and  $h : \Sigma(\mathcal{G}) \rightarrow S^1$  continuous.

Let  $G := F + F \circ \sigma$  and  $g := h(h \circ \sigma)$ , then  $e^{itG} = \lambda^2 g/g \circ \sigma$ . Observe that  $G = F_{\psi^2} \circ \pi_0$ , and that  $\psi^2$  fixes the singularities of  $\text{St}$  (this holds for any homogeneous automorphism, by virtue of the fact that it preserves the  $D$ -orbits of the singularities of  $\text{St}$ ). By the first part of the proof,  $t \in 2\pi\mathbb{Z}$ . It follows that  $\lambda h/h \circ \sigma = e^{itF} = 1$ , whence  $h$  is a continuous eigenfunction of  $\sigma$ . Since  $\sigma$  is topologically mixing,  $h = \text{const}$  and  $\lambda = 1$ .  $\square$

**Acknowledgements.** O.S. would like to thank Barak Weiss for useful discussions on the automorphisms of the infinite staircase surface. This work was initiated when O.S. and D.D. were visiting IMPA, and they would like that institution for its hospitality and excellent working conditions. Part of the work was carried out when E.D. was visiting the Weizmann Institute as part of the *Kupcinet-Getz* summer program.

## REFERENCES

- [ADF] J. Aaronson; M. Denker; A.M. Fisher: *Second order ergodic theorems for ergodic transformations of infinite measure spaces*. Proc. Amer. Math. Soc. **114** (1992), no. 1, 115–127.
- [AK] J. Aaronson; M. Keane: *The visits to zero of some deterministic random walks*. Proc. London Math. Soc. (3) **44** (1982), no. 3, 535–553.
- [ANSS] J. Aaronson; H. Nakada; O. Sarig; R. Solomyak: *Invariant measures and asymptotics for some skew products*. Israel J. Math. **128** (2002), 93–134.
- [AS] J. Aaronson and O. Sarig: *Exponential chi squared distributions in infinite ergodic theory*. To appear in Erg. Th. Dyn. Syst.
- [AW] R.L. Adler; B. Weiss: *Similarity of automorphisms of the torus*. Memoirs of the American Mathematical Society, No. **98**, American Mathematical Society, Providence, R.I. 1970 ii+43 pp.
- [At] G. Atkinson: *Recurrence of co-cycles and random walks*. J. London Math. Soc. (2) **13** (1976), no. 3, 486–488.
- [BL1] M. Babillot; F. Ledrappier: *Geodesic paths and horocycle flow on abelian covers*. Lie groups and ergodic theory (Mumbai, 1996), 1–32, Tata Inst. Fund. Res. Stud. Math., **14**, Tata Inst. Fund. Res., Bombay, 1998.

- [BL2] M. Babillot; F. Ledrappier: *Lalley's theorem on periodic orbits of hyperbolic flows*. Ergodic Theory Dynam. Systems **18** (1998), no. 1, 17–39.
- [BS] L. Barreira; J. Schmeling: *Sets of "non-typical" points have full topological entropy and full Hausdorff dimension*. Israel J. Math. **116** (2000), 29–70.
- [B1] J. Beck: *Randomness of the square root of 2 and the giant leap, part 1*, Periodica Math. Hungarica **60** (2010), 137–242.
- [B2] J. Beck: *Randomness of the square root of 2 and the giant leap, Part 2*. Periodica Math. Hungarica **62** (2011), 127–246.
- [Bor] A. N. Borodin: *On the distribution of random walk local time*. Ann. Inst. H. Poincaré Probab. Statist. **23** (1987), no. 1, 63–89.
- [BR] M. Boshernitzan; D. Ralston: *Continued fractions and heavy sequences*. Proc. Amer. Math. Soc. **137** (2009), no. 10, 3177–3185.
- [BD] R. Burton; M. Denker: *On the central limit theorem for dynamical systems*. Trans. Amer. Math. Soc. **302** (1987), no. 2, 715–726.
- [C] J.-P. Conze: *Equirépartition et ergodicité de transformations cylindriques*. (French) Séminaire de Probabilités, I (Univ. Rennes, Rennes, 1976), Exp. No. 2, 21 pp. Dépt. Math. Informat., Univ. Rennes, Rennes, 1976.
- [CK] J.-P. Conze; M. Keane: *Ergodicité d'un flot cylindrique*. (French) Séminaire de Probabilités, I (Univ. Rennes, Rennes, 1976), Exp. No. 5, 7 pp. Dépt. Math. Informat., Univ. Rennes, Rennes, 1976.
- [D] W. Doeblin: *Sur deux problèmes de M. Kolmogoroff concernant les chaînes dénombrables*. (French) Bull. Soc. Math. France **66** (1938), 210–220.
- [E] G. K. Eagleson: *Some simple conditions for limit theorems to be mixing*. (Russian) Teor. Veroyatnost. i Primenen. **21** (1976), no. 3, 653–660. *Engl. Transl. in Th. Probab. Appl.* **21** (1977), 637–642.
- [Fi1] A. M. Fisher: *Convex invariant means and a pathwise central limit theorem*, Adv. Math. **63** (1987), 213–246.
- [Fi2] A. M. Fisher: *Integer Cantor sets and an order-two ergodic theorem*. Ergodic Theory Dynam. Systems **13** (1993), no. 1, 45–64.
- [Fo] L.R. Ford: *Automorphic functions. Second edition, AMS Chelsea Publishing, American Mathematical Society, Providence, RI, 1957, xii + 333pp.*
- [GH] Y. Guivarc'h; J. Hardy: *Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov*. Ann. Inst. H. Poincaré Probab. Statist. **24** (1988), no. 1, 73–98.
- [Her] M.-R. Herman: *Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations*. (French) Inst. Hautes Études Sci. Publ. Math. No. **49** (1979), 5–233.
- [HHW] W.P. Hooper, P. Hubert and B. Weiss: *Dynamics on the infinite staircase surface*, Dis. Cont. Dyn. Sys. (series A) **33** (2013), 4341–4347.
- [HW] W.P. Hooper and B. Weiss: *Generalized staircases: recurrence and symmetry*. Ann. Inst. Fourier (Grenoble) **62** (2012), no. 4, 1581–1600.
- [Hu] F. Huveneers *Subdiffusive behavior generated by irrational rotations*, Erg Th. Dyn. Sys. **29** (2009) 1217–1233
- [Kel] G. Keller: *Generalized bounded variation and applications to piecewise monotonic transformations*, Z. Wahrsch. Verw. Gebiete **69** (1985), no. 3, 461–478.
- [Kes] H. Kesten: *Uniform distribution mod 1*. Ann. of Math. (2) **71** 1960 445–471.
- [Kh] A. Ya. Khinchin: *Continued fractions*. With a preface by B. V. Gnedenko. *Translated from the third (1961) Russian edition. Reprint of the 1964 translation. Dover Publications, Inc., Mineola, NY, 1997. xii+95 pp.*
- [KN] L. Kuipers and H. Niederreiter: *Uniform distribution of sequences*. *Dover Publ. Inc., Mineola, New York (2006)*.
- [LS1] F. Ledrappier; O. Sarig: *Unique ergodicity for non-uniquely ergodic horocycle flows*. Discrete Contin. Dyn. Syst. **16** (2006), no. 2, 411–433.
- [LS2] F. Ledrappier; O. Sarig: *Fluctuations of ergodic sums for horocycle flows on  $\mathbb{Z}^d$ -covers of finite volume surfaces*. Disc. Contin. Dyn. Syst. **22** (2008), no. 1-2, 247–325.
- [Ngv] S. V. Nagaev: *More exact limit theorems for homogeneous Markov chains*. (Russian) Teor. Veroyatnost. i Primenen. **6** (1961) 67–86.
- [Leo] V.P. Leonov: *On the dispersion of time means of a stationary stochastic process*. (Russian. English summary) Teor. Veroyatnost. i Primenen. **6** (1961) 93–101.

- [Nkd] H. Nakada: *Piecewise linear homeomorphisms of type III and the ergodicity of cylinder flows*. Keio Math. Sem. Rep. No. **7** (1982), 29–40.
- [PP] W. Parry; M. Pollicott: Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque* **187–188** (1990), *Soc. Math. de France*, 268pp.
- [Pe] Y. Peres: *A combinatorial application of the maximal ergodic theorem*, Bull. London Math. Soc. **20** (1988), 248–252.
- [Pet] K. Petersen: Ergodic theory. *Cambridge Studies in Advanced Mathematics, 2. Cambridge University Press, Cambridge, 1983. xii+329 pp. ISBN: 0-521-23632-0*
- [PS] W. Philipp; W. Stout: Almost sure invariance principles for partial sums of weakly dependent random variables. *Mem. Amer. Math. Soc.* **2** **161** iv+140pp.
- [Ra] D. Ralston:  $\frac{1}{2}$ -Heavy Sequences Driven By Rotation. Preprint arXiv:1106.0577.
- [Ru] D. Ruelle: Thermodynamic formalism. The mathematical structures of classical equilibrium statistical mechanics. With a foreword by Giovanni Gallavotti and Gian-Carlo Rota. *Encyclopedia of Mathematics and its Applications*, **5** Addison-Wesley Publishing Co., Reading, Mass., 1978. xix+183 pp.
- [Sa] O.M. Sarig: *Unique ergodicity for infinite measures*. Proceedings of the International Congress of Mathematicians. Volume III, 1777–1803, Hindustan Book Agency, New Delhi, 2010.
- [Sch] K. Schmidt: *A cylinder flow arising from irregularity of distribution*. *Compositio Math.* **36** (1978), no. 3, 225–232.
- [SS] O. Sarig; B. Schapira: *The generic points for the horocycle flow on a class of hyperbolic surfaces with infinite genus*. *Int. Math. Res. Not.* 2008, Art. ID rnn 086, 37 pp.
- [V] D. Volný: *Invariance principles and Gaussian approximation for strictly stationary processes*. *Trans. Amer. Math. Soc.* **351** (1999), no. 8, 3351–3371.

A. AVILA, CNRS U.M.R. 7586 INSTITUT DE MATHÉMATIQUES DE JUSSIEU - PARIS RIVE GAUCHE, BÂTIMENT SOPHIE GERMAIN, CASE 7012, 75205 PARIS CEDEX 13, FRANCE and IMPA ESTRADA DONA CASTORINA, 110 RIO DE JANEIRO, 22460-320 BRASIL  
*E-mail address:* [artur@math.jussieu.fr](mailto:artur@math.jussieu.fr)

D. DOLGOPYAT, 4417 MATHEMATICS BLDG, UNIVERSITY OF MARYLAND, COLLEGE PARK, MD 20742, USA  
*E-mail address:* [dmitry@math.umd.edu](mailto:dmitry@math.umd.edu)

E. DURIEV, DEPARTMENT OF MATHEMATICS, FACULTY OF ARTS AND SCIENCES, HARVARD UNIVERSITY, ONE OXFORD STREET CAMBRIDGE MA 02138 USA  
*E-mail address:* [edwardduriev@gmail.com](mailto:edwardduriev@gmail.com)

O. SARIG, FACULTY OF MATHEMATICS AND COMPUTER SCIENCE, WEIZMANN INSTITUTE OF SCIENCE, 234 HERZL STREET, REHOVOT, 7610001 ISRAEL  
*E-mail address:* [omri.sarig@weizmann.ac.il](mailto:omri.sarig@weizmann.ac.il)