Ergod. Th. & Dynam. Sys. (1982), **2**, 49–68
Printed in Great Britain

Dynamical properties of quasihyperbolic toral automorphisms

D. A. LIND

Department of Mathematics, University of Washington, Seattle, Washington 98195

(*Received* 9 *October* 1981)

Abstract. We study the dynamical properties of ergodic toral autmorphisms that have some eigenvalues of modulus one. For such automorphisms, all sufficiently fine smooth partitions generate measurably, but never topologically, and are never weak Bernoulli. The points of period k become uniformly distributed exponentially fast, and Lipschitz functions mix exponentially fast. Every reasonably smooth compact null set has the property that there is a dense set of periodic points whose entire orbit misses the set, but this is false for general compact null sets. Katznelson's property of almost weak Bernoulli can be strengthened to a certain exponential rate of independence, but breaks down at a critical number. Finally, open sets have return times that decay exponentially fast.

1. *Introduction*

Hyperbolic toral automorphisms have inspired many developments in topological dynamics and ergodic theory. The important technique of Markov partitions originated in an analysis of automorphisms of the two-dimensional torus [**1**]. They also supplied the original motivation for introducing Anosov (and later Axiom A) diffeomorphisms, and gave the first example of a structurally stable diffeomorphism with a dense set of periodic points [**17**].

Less work has been done on those ergodic toral automorphisms that have eigenvalues of modulus one. Call such automorphisms *quasihyperbolic*. In 1971 Katznelson showed that ergodic toral automorphisms are measure isomorphic to Bernoulli shifts [**8**]. Hyperbolic automorphisms have Markov partitions, which help greatly in studying their dynamical properties. However, quasihyperbolic automorphisms lack Bowen's specification property [**9**], so never have Markov partitions. Enough hyperbolicity remains for Marcus [**11**] to prove that measures on periodic orbits are weakly dense in the space of invariant probability measures, answering a question in [**9**].

It is then natural to inquire about which properties of hyperbolic systems are shared by quasihyperbolic automorphisms, and also whether there are new phenomena. Some things go through (like Bernoullicity), while others do not (like Markov partitions and specification). This paper is devoted to studying this question. The basic tools are harmonic analysis, a lemma of Katznelson on Diophantine approximation, the ergodic theorem, and in certain places a deep result of Gelfond on the approximation of logarithms of algebraic integers by rationals.

After a review of the geometry of quasihyperbolic automorphisms in § 2, we show in § 3 that all sufficiently fine partitions measure theoretically generate. However, even smooth fine partitions cannot topologically generate (§ 4). This is proved by showing that for a fixed smooth partition there is a dense set of periodic points whose orbit misses the boundary of the partition. A possible strengthening is shown to be false by constructing a compact null set containing at least one point from each periodic orbit. The proof of the above fact about periodic points shows that they become uniformly distributed exponentially fast, and shows that Bowen's result on the periodic point measures converging to the measure of maximal entropy [**2**] remains true here. The same technique also shows (§ 7) that Lipschitz functions mix exponentially fast.

Bowen [**4**] proved that smooth partitions for hyperbolic systems always obey an asymptotic independence property called weak Bernoulli. Surprisingly, fine measurable partitions for quasihyperbolic automorphisms are never weak Bernoulli (§ 5), although they must obey a weaker condition called very weak Bernoulli. A corollary of this, using a result of del Junco and Rahe [**5**], is that conditional entropy for fine partitions must converge quite slowly. In his proof that toral automorphisms are Bernoulli, Katznelson introduced the notion of almost weak Bernoulli between weak and very weak Bernoulli. We give in § 6 a quantitative reason why he was forced to do this, and show that almost weak Bernoulli breaks down at a critical number that may yield an interesting invariant of toral automorphisms.

Many of the standard examples of systems known to be measure isomorphic to Bernoulli shifts are also 'almost topologically' or 'finitarily' Bernoulli as well. For example, Keane and Smorodinsky have shown there is measurable isomorphism between a hyperbolic toral automorphism and a Bernoulli shift that is a homeomorphism off an invariant null set. A natural question is whether quasihyperbolic automorphisms are finitarily Bernoulli. In this connection, Smorodinsky has observed that a finitarily Bernoulli transformation necessarily has exponentially decaying return probabilities on each open set. In § 8 we check that quasihyperbolic automorphisms have this necessary condition, lending support to the conjecture that they are finitarily Bernoulli.

Finally, some questions and conjectures are gathered in § 9.

The author thanks the Mathematics Department at Stanford University for their generous hospitality during a visit in which part of this work was done.

2. *Geometry of quasihyperbolic automorphisms*

We begin with a description of the essential geometric features of toral automorphisms.

Let S be an ergodic automorphism of $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$. Thus $S \in \mathrm{GL}\,(n, \mathbb{Z})$ is given by an $n \times n$ matrix with integer entries and determinant ± 1. Considering S as a linear transformation of $\mathbb{R}^n$, there is an S-invariant decomposition

$$\mathbb{R}^n = E^s \oplus E^c \oplus E^u,$$

where E^s corresponds to those eigenvalues less than 1 in modulus, E^c to those of modulus 1, and E^u to the rest. Ergodicity of S is equivalent to no eigenvalue being a root of unity.

It will be convenient to equip $\mathbb{R}^n$ with a norm adapted to S. Let λ generically denote an eigenvalue of S. Choose numbers ρ and ξ such that

$$\max\{|\lambda|: |\lambda|<1\}<\rho^{-1}<1<\rho<\min\{|\lambda|: |\lambda|>1\},$$
$$1<\max\{|\lambda|, |\lambda|^{-1}\}<\xi.$$

We construct norms $\| \ \|_s$, $\| \ \|_u$ on E^s, E^u such that

$$\|Sx\|_s \le \rho^{-1}\|x\|_s, \qquad \|S^{-1}x\|_s \le \xi\|x\|_s \quad (x \in E^s), \tag{2.1}$$

$$\|S^{-1}x\|_u \le \rho^{-1}\|x\|_u, \qquad \|Sx\|_u \le \xi\|x\|_u \quad (x \in E^u). \tag{2.2}$$

To obtain $\| \ \|_u$, begin with an arbitrary norm $\| \ \|'$ on E^u. First adapt $\| \ \|'$ for S as follows. Put

$$\|x\|'_u = \sum_{j=0}^{\infty} \xi^{-j}\|S^j x\|'.$$

It is easily checked that since $|\lambda|<\xi$, this series converges, defines a norm on E^u, and has

$$\|Sx\|'_u \le \xi\|x\|'_u.$$

Now adapt $\| \ \|'_u$ to S^{-1} by putting

$$\|x\|_u = \sum_{k=0}^{\infty} \rho^k \|S^{-k}x\|'_u.$$

Since $\rho<\min\{|\lambda|: |\lambda|>1\}$, the series again converges, defines a norm, and has

$$\|S^{-1}x\|_u \le \rho^{-1}\|x\|_u.$$

The property $\|Sx\|'_u \le \xi\|x\|'_u$ persists to $\| \ \|_u$. This gives (2.2). Replace S by S^{-1} to obtain (2.1).

Call λ's with $|\lambda|=1$ *unitary eigenvalues*. Since they are complex, the real Jordan form for S on E^c is a block matrix

$$\operatorname{diag}[J(R_1, n_1), \dots, J(R_q, n_q)],$$

where the Jordan block

$$J(R_i, n_i) = \begin{pmatrix} R_i & I & & \\ & R_i & I & \\ & & \ddots & I \\ & & & R_i \end{pmatrix} \tag{2.3}$$

has n_i copies of a 2×2 rotation matrix R_i, and 2×2 identity matrices I as the only other non-zero entries.

Thus E^c decomposes into a direct sum $E_1 \oplus \cdots \oplus E_q$ of S-invariant subspaces, with

$$E_i = E_{i1} \oplus \cdots \oplus E_{in_i} \quad \text{and} \quad \dim E_{ij} = 2.$$

Norm E_{ij} with $\| \ \|_{ij}$ so the rotation R_i is an isometry. If π_{ij} is projection to E_{ij} along the complementary subspace, then $R_i = \pi_{ij} S \pi_{ij}$, and so

$$\|\pi_{ij} S x\|_{ij} = \|x\|_{ij} \quad \text{for } x \in E_{ij}.$$

Let $\pi_s:\mathbb{R}^n \to E^s$, $\pi_u:\mathbb{R}^n \to E^u$ be projections along complementary subspaces. Define $\| \ \|$ on $\mathbb{R}^n$ by

$$\|x\| = \max\{\|\pi_s x\|_s, \|\pi_u x\|_u, \|\pi_{ij} x\|_{ij}\}.$$

With respect to this norm, $\|J(R_i, n_i)\| \le 2$, so $\|S\| \le \max(\xi, 2) \le 2\xi$.

Let $p:\mathbb{R}^n \to \mathbb{T}^n$ be the quotient map. For $x, y \in \mathbb{T}^n$ put

$$d(x, y) = \min\{\|\tilde{x} - \tilde{y}\|: p\tilde{x} = x, p\tilde{y} = y\}.$$

Then d is a translation invariant metric on $\mathbb{T}^n$. If

$$\delta_0 = \tfrac{1}{2}\min\{\|z\|: z \in \mathbb{Z}^n, z \neq 0\},$$

$$\tilde{B}(\delta_0) = \{\tilde{x} \in \mathbb{R}^n: \|\tilde{x}\| < \delta_0\}, \quad \text{and} \quad B(\delta_0) = \{x \in \mathbb{T}^n: d(0, x) < \delta_0\},$$

then $p:\tilde{B}(\delta_0) \to B(\delta_0)$ is an isometry. Also note that if $\|\tilde{w}\| < \delta_0$ and $p\tilde{w} = w$, then $d(x, x + w) = \|\tilde{w}\|$.

Since the intersection of each subspace E^s, E_{ij}, E^u with $\mathbb{Z}^n$ is $\{0\}$, it is not dangerous to identify each with its image under p in $\mathbb{T}^n$. In particular, we think of $\tilde{B}(\delta_0)$ as lying in $\mathbb{T}^n$. Generally, $\| \ \|$ will be applied to points thought of as being in $\mathbb{R}^n$, while d is applied to elements of $\mathbb{T}^n$.

To localize the action of S, we formulate a notion of fineness for partitions. Put $\phi = \delta_0/10\xi$, where δ_0 and ξ are as above. Let

$$\alpha = \{A_1, \ldots, A_r\}$$

be a partition of $\mathbb{T}^n$ into measurable sets. Call α *fine* (or *S-fine*) if

$$\max\{\operatorname{diam} A_j\} < \phi.$$

3. *Generators*

A partition α *generates* under S if the completed σ-algebra generated by the $S^j\alpha$ ($j \in \mathbb{Z}$) coincides with the whole σ-algebra off a null set.

THEOREM 1. *Let S be an ergodic automorphism of $\mathbb{T}^n$. Then every S-fine measurable partition generates under S.*

Proof. We use the fact that α generates iff it separates points, i.e. there is a null set $N \subset \mathbb{T}^n$ such that for distinct points $x, y \in \mathbb{T}^n \backslash N$ there is an integer k with $S^k x$ and $S^k y$ in different atoms of α (see [**14**]).

If S is hyperbolic this is easy. Suppose

$$y = x + w_s + w_u \quad \text{with } w_s \in E^s, w_u \in E^u, \|w_s\| < \phi, \|w_u\| < \phi.$$

Replacing S by S^{-1} if necessary, we can assume $w_u \neq 0$. Choose k maximal so that

$$\|S^j w_u\| < \phi, \qquad 0 \le j \le k.$$

Since $\|S^j w_u\| \ge \rho^j \|w_u\|$ where $\rho > 1$, k is finite. Now

$$\|S^{k+1} w_u\| \le \|S\| \|S^k w_u\| \le 2\xi\phi < \tfrac{1}{4}\delta_0 \quad \text{and} \quad \|S^{k+1} w_s\| < \phi,$$

so

$$d(S^{k+1}x, S^{k+1}y) = \|S^{k+1}(w_s + w_u)\| \ge \|S^{k+1} w_u\| \ge \phi$$

by definition of k. Thus $S^{k+1}x$ and $S^{k+1}y$ are in different atoms of α.

For hyperbolic S there is no need to remove a null set to separate points. In fact, a fine partition α generates topologically in the sense that for every choice of i_j,

$$\bigcap_{-\infty}^{\infty} S^j \bar{A}_{i_j}$$

contains at most one point. On the other hand, for quasihyperbolic S it is at least necessary to remove a dense set of lower dimensional disks to separate points (see § 4).

Suppose now that $E^c \neq 0$. Call S *central spin* if each Jordan block in (2.3) is two-dimensional. For such automorphisms there are no off-diagonal I's, and they are isometries on E^c. If S has off-diagonal I's in some Jordan block for E^c, call it *central skew*. There is a different argument for each type.

If $y = x + w_s + w_c + w_u$ with either $w_s \neq 0$ or $w_u \neq 0$, the argument above for the hyperbolic case applies unchanged. Thus we may suppose $y = x + w$ where $w \in E^c$ with $\|w\| < \phi$.

First suppose S is central spin. Since

$$E^c = E_1 \oplus \cdots \oplus E_q,$$

it suffices to assume $w \in E_i$ for some i. We then remove a null set for each E_i, and will have point separation on the rest of $\mathbb{T}^n$.

Simplify notation by letting $E_i = E$, and let F be its invariant complement. The matrix of S on E is

$$R_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

with respect to an orthonormal basis $\{e_1, e_2\}$. Let $H = F \oplus \mathbb{R}e_2$, $L = \mathbb{R}e_1$, and π_L be projection to L (identified with $\mathbb{R}$) along H.

Fix $A \in \alpha$. Choose $\tilde{A} \subset \mathbb{R}^n$ with diam $\tilde{A} < \phi$, so that $p: \tilde{A} \to A$ is an isometry. Let μ denote Haar measure on $\mathbb{R}^n$ or $\mathbb{T}^n$, the distinction being clear from context, normalized so that $\mu(\tilde{A}) = \mu(p\tilde{A})$. If

$$\tilde{A}(t) = \{x \in \tilde{A} : \pi_L x \le t\},$$

then $\mu(\tilde{A}(t))$ is 0 for large negative t and is $\mu(\tilde{A})$ for large positive t. Thus

$$t_0 = \inf\{t : \mu(\tilde{A}(t)) = \mu(\tilde{A})\}$$

is finite. Since $\mu(\tilde{A} \backslash \tilde{A}(t_0)) = 0$, a harmless modification of A by a null set allows us to assume $\tilde{A} = \tilde{A}(t_0)$.

Let $\tilde{A}_m = \tilde{A} \backslash \tilde{A}(t_0 - 1/m)$, and put $A_m = p\tilde{A}_m$. Then $\mu(A_m) > 0$. If $\tilde{x} \in \tilde{A}_m$ and $w \in E$ with $\|w\| < \phi$ and $\pi_L w > 1/m$, then

$$\pi_L(\tilde{x} + w) > (t_0 - 1/m) + 1/m = t_0,$$

so $\tilde{x} + w \notin \tilde{A}$. Applying p shows that if $x \in A_m$ and w is as above, then $x + w \notin A$. This is the property we will use to separate points.

Define $U_\theta : \mathbb{T} \to \mathbb{T}$ by $U_\theta(\psi) = \psi + \theta$. Since θ is irrational by ergodicity of S, U_θ is ergodic with respect to Haar measure ν on $\mathbb{T}$. The product of a mixing and an ergodic transformation is ergodic, so $S\times U_\theta$ on $(\mathbb{T}^n\times\mathbb{T},\mu\times\nu)$ is ergodic. Let

$$I=(-\tfrac{1}{12},\tfrac{1}{12})\subset\mathbb{T},\qquad K=(-\tfrac{1}{6},\tfrac{1}{6}).$$

Since $(\mu\times\nu)(A_m\times I)>0$, the ergodic theorem applied to $S\times U_\theta$ shows that

$$(\mu\times\nu)\{(x,\psi)\in\mathbb{T}^n\times\mathbb{T}\colon (S^kx,\psi+k\theta)\in A_m\times I\quad\text{for some }k\}=1.$$

By Fubini's theorem, there is a null set $N_m\subset\mathbb{T}^n$ such that for $x\in\mathbb{T}^n\backslash N_m$ we have

$$\nu\{\psi\in\mathbb{T}\colon (S^kx,\psi+k\theta)\in A_m\times I\quad\text{for some }k\}=1.$$

A set of full measure in $\mathbb{T}$ is dense, so for every $x\notin N_m$ and every $\psi\in\mathbb{T}$ there is a k such that

$$(S^kx,\psi+k\theta)\in A_m\times K.$$

Let $N=\bigcup_{m=1}^{\infty}N_m$, so $\mu(N)=0$. We show that if $y=x+w$ with $x\notin N$, $0\neq w\in E$, and $\|w\|<\phi$, then x and y are separated by α under S. It helps to think of S^kx as a sun moving about $\mathbb{T}^n$ and $S^ky=S^kx+S^kw$ as a planet revolving around x.

Choose m so that $\|w\|=b>2/m$. Write $w=bR_\psi e_1$. Since $x\notin N$, there is a k so that $S^kx\in A_m$ and $\psi+k\theta\in K$. Now

$$S^ky-S^kx=S^kw=bR_\theta^kR_\psi e_1=bR_{\psi+k\theta}e_1,$$

a vector in E of norm $b<\phi$. Also, since π_L on E is orthogonal projection to L,

$$\pi_L(S^kw)=\pi_L(bR_{\psi+k\theta}e_1)=b\cos 2\pi(\psi+k\theta)$$

$$>b\cos 2\pi/6>\left(\frac{2}{m}\right)\left(\frac{1}{2}\right)=\frac{1}{m}.$$

Thus $S^kx\in A_m\subset A$, while by the above $S^ky\notin A$. This completes the central spin argument.

Now suppose S is central skew. The proof is not as delicate because of a polynomial drift in the non-diagonal Jordon blocks.

Recall the S-invariant decomposition of E^c into

$$E_1\oplus\cdots\oplus E_q,$$

where each $E_i=E_{i1}\oplus\cdots\oplus E_{in_i}$ on which S has matrix (2.3). The central spin argument does not apply directly here since the E_{ij} are not S-invariant.

Let $y=x+w$ with $0\neq w\in E^c$ and $\|w\|<\phi$. Put

$$E_0=E_{11}\oplus E_{21}\oplus\cdots\oplus E_{q1}$$

on which S acts isometrically. If $w\in E_0$, the previous argument does apply to separate x and y not in N.

Next we claim that if $w\in E^c\backslash E_0$, then $\|S^kw\|\to\infty$ as $k\to\infty$. For such a w, there are i and $j\geq 2$ such that $\pi_{ij}w\neq 0$. Fix such an i, and choose j maximal so that $\pi_{ij}w\neq 0$. Then

$$\pi_{i,j-1}(S^kw)=R_i^k(\pi_{i,j-1}w)+kR_i^{k-1}(\pi_{ij}w),$$

so

$$\|S^kw\|\geq\|\pi_{i,j-1}(S^kw)\|\geq k\|R_i^{k-1}(\pi_{ij}w)\|-\|\pi_{i,j-1}w\|\to\infty$$

as $k\to\infty$. If $y=x+w$ with $w\in E^c\backslash E_0$, since $\|S^kw\|\to\infty$ the argument from the hyperbolic case applies to show that x and y are separated, completing the proof. □

4. *Periodic points*

Call a partition of $\mathbb{T}^n$ *smooth* if its atoms all have a piecewise smooth boundary. Suppose α is a smooth fine partition for a quasihyperbolic S. Although α is a measure theoretic generator (§ 3), we show here that α never topologically generates. In fact, there will always exist a dense collection of lower dimensional disks such that each pair of points in each disk is not separated by α under S. This is proved by showing that if $\partial\alpha$ denotes the union of the boundaries of the atoms of α, then a dense set of periodic points have orbits that miss $\partial\alpha$. Indeed, in a precise sense, exponentially most periodic points have this property. The proof uses a Diophantine approximation lemma of Katznelson employed in the original proof of the Bernoullicity of toral automorphisms. A strong form of this property requires a deep result of Gelfond on the approximation of logarithms of algebraic numbers. As a byproduct, we obtain that the periodic point measures converge to Haar measure exponentially fast on Lipschitz functions.

Let $P_k = \{t \in \mathbb{T}^n : S^k t = t\}$, and μ_k be the probability measure equidistributed on P_k. The cardinality $|P_k|$ of P_k grows exponentially fast. For it is easily shown that

$$|P_k| = |\det(S^k - I)| = \prod |\lambda^k - 1|.$$

If $h = h(S) = \sum \{\log|\lambda| : |\lambda| > 1\}$ denotes the topological entropy of S, then for those k for which λ^k is not too close to 1, $|P_k|$ grows like e^{hk}. The existence of the limit

$$\lim_{k\to\infty} \frac{1}{k} \log |P_k| = h$$

is equivalent to the validity of Gelfond's theorem that for unitary λ, $|\lambda^k - 1| > e^{-\varepsilon k}$ only finitely often for each $\varepsilon > 0$ [**6**]. This has been noticed independently by Peter Walters about three years ago in answer to a question by Brian Marcus. We discuss this in more detail later.

Even for relatively small periods, P_k is exceedingly numerous. For

$$S = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

the points of period 20 number 228 826 125, and those with strict period 20 only about 0.006% fewer (use Möbius inversion as in (4.2) of [**17**]).

Since P_k is a finite subgroup of $\mathbb{T}^n$, namely the kernel of $S^k - I$, duality shows that the Fourier transform $\hat{\mu}_k$ is the indicator function of $(S^k - I)\mathbb{Z}^n$. The key result is that, except for 0, the support $(S^k - I)\mathbb{Z}^n$ of $\hat{\mu}_k$ misses a ball whose radius grows exponentially in k. This will show that μ_k mixes uniformly for trigonometric polynomials whose degree is exponential in k.

We continue to use the norm on $\mathbb{R}^n$ introduced in § 2, and use $B(r)$ for the ball of radius r.

For several of the following results, there is a weak form involving a subset or subsequences of k's. These weak forms are proved using only Katznelson's lemma, and are sufficient to obtain corollaries 4.2 and 4.3, our main objectives. However, by invoking Gelfond's result, we show that the results are actually true for all large enough k. We feel it is important to keep clear what can be proved by relatively elementary means.

LEMMA 4.1. *There is an* $r>1$ *such that*

$$(S^k - I)\mathbb{Z}^n \cap B(r^k) = \{0\}$$

for all sufficiently large k *(weak form: infinitely many* k*).*

Proof. Suppose $z \in \mathbb{Z}^n$ and $\|(S^k - I)z\| < r^k$. Then

$$z = z_s + z_c + z_u \in E^s \oplus E^c \oplus E^u,$$

and

(i) $\|(S^k - I)z_s\| < r^k$,
(ii) $\|(S^k - I)z_c\| < r^k$,
(iii) $\|(S^k - I)z_u\| < r^k$.

The basic idea is that if r is close enough to 1, these conditions will force z to be very close to $E^s \oplus E^c$, which is ruled out by Katznelson's lemma.

First note that (i) implies

$$r^k > \|z_s - S^k z_s\| \geq (1 - \rho^{-k})\|z_s\|,$$

so for large k

$$\|z_s\| < r^k (1 - \rho^{-k})^{-1} < 2r^k.$$

Also, (iii) shows that

$$r^k > \|S^k z_u - z_u\| \geq (\rho^k - 1)\|z_u\|,$$

so for large k

$$\|z_u\| < r^k (\rho^k - 1)^{-1} < 2(r/\rho)^k.$$

Finally, (ii) implies

$$\|z_c\| = \|(S^k - I)^{-1}(S^k - I)z_c\| \leq \|(S^k - I)^{-1}\| r^k,$$

so it becomes necessary to estimate the norm of $(S^k - I)^{-1}$ on E^c. By § 2, this is bounded by the maximum over i of

$$\|(J(R_i, n_i)^k - I)^{-1}\|.$$

To estimate these, assume R is an isometry and $J(R, m)$ acts on $E_1 \oplus \cdots \oplus E_m$. Then $J(R, m)^k$ has (i, j)th entry 0 if $i < j$, and

$$\binom{k}{i-j} R^{i-j} \quad \text{for } i \geq j.$$

Thus the norm of each entry of $J(R, m)^k$ is bounded by k^m. The cofactor expansion of $(J(R, m)^k - I)^{-1}$ shows that its norm is bounded by

$$\|(R^k - I)^{-1}\|^m (m!\, k^{m^2}).$$

Since $m < n$, we conclude there is a constant C such that

$$\|z_c\| \leq C \max_{|\lambda|=1} |\lambda^k - 1|^{-2n} k^{n^2} r^k.$$

Since $\lambda \neq 1$, there is an infinite subset of k's such that $|\lambda^k - 1| > \frac{1}{10}$ holds for each unitary λ. Thus, given $\varepsilon > 0$, for all large enough k in this subset (the weak form)

$$\|z_c\| \leq C_1 k^{n^2} r^k < C_2 r^k e^{\varepsilon k}. \tag{4.1}$$

To obtain (4.1) for all large k, we appeal to a result of Gelfond [**6**]. The reason for this added difficulty is that, although $\lambda = e^{2\pi i\theta}$ is algebraic,

$$\theta = \frac{1}{2\pi i} \log \lambda$$

is transcendental, and so elementary methods to estimate the fractional part of $k\theta$ fail. Gelfond's work deals in part with rational approximations to logarithms of algebraic numbers. One consequence is that for unitary λ, given $\varepsilon > 0$ there is a constant $C(\lambda, \varepsilon)$ such that for all k

$$|\lambda^k - 1| > C(\lambda, \varepsilon)\, e^{-\varepsilon k}. \tag{4.2}$$

Recent work of Feldman (see [**19**: Ch. 9]) shows that there are effectively computable numbers $N = N(\lambda)$ and $C = C(\lambda)$ such that

$$|\lambda^k - 1| > Ck^{-N}$$

for all k. Since this refined estimate gives no better information in our case than (4.2), we will not use it further.

Clearly, assuming (4.2) and given $\varepsilon > 0$, we obtain (4.1) for all sufficiently large k (the strong form).

Let $E = E^s \oplus E^c$. Katznelson's lemma [**8**] shows there is a constant C_3 such that

$$\text{dist}\,(z, E) > C_3 \|z\|^{-\dim E} \geq C_3 \|z\|^{-n}$$

for $0 \neq z \in \mathbb{Z}^n$. Since $\text{dist}\,(z, E) = \|z_u\|$, we obtain from $\|z\| < C_2 r^k e^{\varepsilon k}$ that

$$2(r/\rho)^k > \|z_u\| > C_3 \|z\|^{-n} > C_4 (re^{\varepsilon})^{-nk}.$$

Thus

$$(r^{n+1})^k > \tfrac{1}{2} C_4 (\rho e^{-\varepsilon})^k.$$

This is incompatible for large k if $\rho e^{-\varepsilon} > 1$ and r is close enough to 1. Thus choose $\varepsilon > 0$ so that $e^{\varepsilon} < \rho$ to complete the proof. In fact, every r with $1 < r < \rho^{1/(n+1)}$ will work. □

COROLLARY 4.1. *The periodic point measures μ_k converge weakly to μ (weak form: along a subsequence).*

Proof. $\hat{\mu}_k$ converges pointwise to $\hat{\mu}$. □

Remark. This shows the result of Bowen [**2**] extends to quasihyperbolic automorphisms.

We now turn to a smooth partition α. Its boundary $\partial\alpha$ is a compact set such that if $B(\partial\alpha, \varepsilon)$ denotes $\partial\alpha + B(\varepsilon)$, then

$$\mu(B(\partial\alpha, \varepsilon)) = O(\varepsilon).$$

Call a measurable set $K \subset \mathbb{T}^n$ *Lipschitz* if $\mu(B(K, \varepsilon)) = O(\varepsilon)$. Such sets are null and are contained in compact null Lipschitz sets.

THEOREM 2. *Let K be a Lipschitz set in $\mathbb{T}^n$. Define Q_k to be those points in P_k whose orbit hits K. Then $|Q_k|/|P_k| \to 0$ exponentially fast (weak form: on a subsequence of k's).*

Before proving this, we deduce some consequences. Since P_k becomes uniformly distributed (corollary 4.1), the following is immediate.

COROLLARY 4.2. *If S is an ergodic automorphism and K is Lipschitz, then there is a dense set of periodic points whose orbits miss K.*

Since the weak form of corollary 4.1 and of theorem 2 both depend on that of lemma 4.1, they are both valid for the same subset of k's, and so even the weak forms prove corollary 4.2. Even in the hyperbolic case it would be nice to have a more geometric proof of this corollary.

COROLLARY 4.3. *If S is a quasihyperbolic automorphism and α is a smooth fine partition, then there is a dense set $\{D_i\}$ of two-dimensional disks in $\mathbb{T}^n$ such that each pair of distinct points in each disk is not separated by α under S.*

Proof. Since $\partial\alpha$ is Lipschitz, by corollary 4.2 there is a dense set $\{t_i\}$ of periodic points whose orbits miss $\partial\alpha$. For each t_i there is an $\varepsilon_i > 0$ such that if D_i is the ε_i disk around t_i in the E_{11} isometric eigenspace, then $S^j D_i$ is completely contained in one atom of α for each j. □

Proof of theorem 2. We may suppose that K is compact Lipschitz. The idea is to dominate the indicator function of an exponentially small neighbourhood of K by trigonometric polynomials whose degree does not grow too fast in k, integrate against μ_k, and apply lemma 4.1 to show

$$|P_k \cap K|/|P_k| \to 0$$

exponentially fast. Then trivially $|Q_k| \le k|P_k \cap K|$, yielding the result.

Let $K_m(t) = (2m+1)^{-1} \sum_{-m}^{m} e^{2\pi i j t}$ be the Fejér kernel. A simple estimate (see [**7**: I (3.10)]) shows there is a constant $C_1 > 0$ such that

$$\int_{-\delta}^{\delta} K_m(t)\, dt > 1 - C_1/m\delta.$$

Define the n-dimensional kernel F_m on $\mathbb{T}^n$ by

$$F_m(t_1, \ldots, t_n) = K_m(t_1) \cdots K_m(t_n).$$

Then $\hat{F}_m$ has support $\operatorname{supp} \hat{F}_m = \{-m, \ldots, m\}^n = H_m$, say, and there is a constant $C_2 > 0$ such that

$$\int_{B(\delta)} F_m\, d\mu > 1 - C_2/m\delta. \tag{4.3}$$

Choose s with $1 < s < \sqrt{r}$, where r is as in lemma 4.1, and let $m = [s^{2k}]$. For k large enough (weak form: for infinitely many k), $H_m \subset B(r^k)$. Let g_k be the indicator function of $B(K, s^{-k})$. For $t \in K$,

$$B(t, s^{-k}) \subset B(K, s^{-k}),$$

so from (4.3) we obtain for large k that

$$F_m * g_k > (1 - C_2/ms^{-k})\chi_K > (1 - C_2 s^{-k})\chi_K > \tfrac{1}{2}\chi_K.$$

Hence

$$\frac{|P_k \cap K|}{|P_k|} = \int \chi_K \, d\mu_k < 2 \int F_m * g_k \, d\mu_k. \tag{4.4}$$

Now supp $(F_m * g_k)^\wedge \subset H_m \subset B(r^k)$, while supp $\hat{\mu}_k \cap B(r^k) = \{0\}$. Thus the integral on the right side of (4.4) has value

$$\hat{F}_m(0)\hat{g}_k(0)\hat{\mu}_k(0) = \hat{g}_k(0) = \int g_k \, d\mu = \mu(B(K, s^{-k})),$$

so

$$\frac{|P_k \cap K|}{|P_k|} < 2\mu(B(K, s^{-k})).$$

Since $Q_k = \bigcup_{j=0}^{k-1} (S^j K \cap P_k) = \bigcup_{j=0}^{k-1} S^j(K \cap P_k)$, $|Q_k| \le k|P_k \cap K|$. Thus

$$\frac{|Q_k|}{|P_k|} < 2k\mu(B(K, s^{-k})), \tag{4.5}$$

which converges to 0 exponentially fast for Lipschitz K. □

We remark that the theorem is true for compact null K for which

$$k\mu(B(K, s^{-k})) \to 0.$$

However, some restriction on K more stringent than compact null is necessary, as the following observation shows.

PROPOSITION. *Let S be a homeomorphism of a compact metric space (X, d) preserving a non-atomic Borel probability measure μ whose support is K. Assume S is ergodic on (X, μ). Then there is a compact null set that contains at least one point from each periodic orbit of S.*

Proof. Pick $x_0 \in X$ and let $V_j = B(x_0, 1/j)$. Since supp $\mu = X$, $\mu(V_j) > 0$. Ergodicity of S shows that

$$K_j = X \backslash \left(\bigcup_{-\infty}^{\infty} S^i V_j \right)$$

is null, and it is clearly compact. Let

$$K = \{x_0\} \cup \bigcup_{j=1}^{\infty} (\bar{V}_j \cap K_{j+1}).$$

Since $\bar{V}_j \searrow \{x_0\}$, K is compact, and clearly null. Suppose x is periodic under S with orbit

$$P = \{x, Sx, \ldots, S^m x\}.$$

Since $x_0 \in K$, we may assume $x_0 \notin P$. Hence there is a largest j such that $P \cap \bar{V}_j \ne \varnothing$. Then $S^k x \in \bar{V}_j$ for some k, while maximality of j means that $S^i x \notin \bar{V}_{j+1}$ for every i, so $S^k x \in K_{j+1}$. Thus

$$S^k x \in \bar{V}_j \cap K_{j+1} \subset K,$$

finishing the proof. □

Let Lip^γ $(0<\gamma<1)$ denote the space of functions $f:\mathbb{T}^n\to\mathbb{R}$ such that

$$|f(x)-f(y)|\le Cd(x,y)^\gamma$$

for some constant C.

THEOREM 3. *If $f\in\mathrm{Lip}^\gamma$, then $\int f\,d\mu_k$ converges to $\int f\,d\mu$ exponentially fast.*

Proof. Using the same notation as in the proof of theorem 2, it is standard that if $f\in\mathrm{Lip}^\gamma$, then

$$\|F_m*f-f\|_\infty<Cm^{-\gamma}$$

for a constant $C=C(f)$ (see [**7**: p. 21]). Put $m=[s^k]$ for fixed $s<r$, where r comes from lemma 4.1. Then

$$\operatorname{supp}\hat{F}_m\cap\operatorname{supp}\hat{\mu}_k=\{0\}\quad\text{for large } k$$

by lemma 4.1, and so

$$\int F_m*f\,d\mu_k=\hat{f}(0)=\int f\,d\mu.$$

Thus

$$\left|\int f\,d\mu_k-\int f\,d\mu\right|=\left|\int(f-F_m*f)\,d\mu_k\right|\le\|f-F_m*f\|_\infty<Cm^{-\gamma}<C_1s^{-\gamma k}.\qquad\square$$

5. *The failure of weak Bernoulli*

Bowen [**4**] showed that smooth partitions are weak Bernoulli for hyperbolic systems, in particular for hyperbolic toral automorphisms. This means the following.

Two finite measurable partitions α and β are *ε-independent* if

$$\sum_{A\in\alpha,B\in\beta}|\mu(A\cap B)-\mu(A)\mu(B)|<\varepsilon.\tag{5.1}$$

Let S be a measure-preserving transformation. For typographical convenience let $\alpha[r,t]$ denote the common refinement of

$$\{S^{-j}\alpha: r\le j\le t\}.$$

A partition α is *weak Bernoulli* for S if for every $\varepsilon>0$ there is a $K_0(\varepsilon)$ such that for every $M>K>K_0$ and every $N>0$, the partitions $\alpha[-N,0]$ and $\alpha[K,M]$ are ε-independent. The importance of this notion is that it is a sufficient condition for S to be measurably isomorphic to a Bernoulli shift on the σ-algebra generated by α under S [**12**]. In particular, one can show that S is Bernoulli by finding a weak Bernoulli generator. This is the way that Friedman and Ornstein first proved that mixing finite-state Markov chains are Bernoulli.

It turns out that every partition in a Bernoulli shift obeys a weaker asymptotic independence property called very weak Bernoulli (see [**12**]). Smorodinsky [**18**] constructed an example of a partition in a Bernoulli shift that is not weak Bernoulli. Shields [**16**] found a more natural example using a skew product of the 2-shift with rotations of the circle. We show here that fine partitions are never weak Bernoulli for quasihyperbolic automorphisms (cf. Bowen's result above).

The geometrical idea is roughly the same as in Shields [**16**]. Past fibres are pieces of unstable manifold, while future fibres are pieces of stable manifold. Because of the extra central direction, stable and unstable manifolds intersect almost nowhere, and cannot be ε-independent. The approximate version of this with thickened fibres, established via the martingale theorem, suffices for the proof.

THEOREM 4. *Let S be a quasihyperbolic automorphism. Then no fine measurable partition is weak Bernoulli for S.*

Proof. It is convenient to pass to a limit on one side and use the measurable partition machinery of Rohlin [**14**]. Recall that a partition ζ of a Lebesgue space (X, μ) into fibres C is *measurable* if the quotient X/ζ is also a Lebesgue space. Rohlin shows that a.e. C carries a Lebesgue space measure μ_C, and X/ζ carries the natural quotient measure μ_ζ. These are related via a Fubini theorem as follows. If $M \subset X$ is measurable, then $C \to \mu_C(M \cap C)$ is μ_ζ-measurable, and

$$\mu(M) = \int_{X/\zeta} \mu_C(M \cap C)\, d\mu_\zeta(C).$$

We now reformulate the weak Bernoulli criterion by allowing an infinite past. To say that $\alpha[-N, 0]$ is ε-independent of $\alpha[K, M]$ means that

$$\sum_{C \in \alpha[-N,0]} \sum_{D \in \alpha[K,M]} |\mu_C(D) - \mu(D)|\mu(C) < \varepsilon, \tag{5.2}$$

where $\mu_C(D) = \mu(C \cap D)/\mu(C)$ if $\mu(C) > 0$ and 0 otherwise. The first sum is really an integral over the atomic Lebesgue space $X/\alpha[-N, 0]$ assigning measure $\mu(C)$ to the fibre C. Put $\alpha^- = \alpha[-\infty, 0]$ and $\alpha^+ = \alpha[0, \infty]$. Letting $N \to \infty$ and applying the martingale theorem to (5.2) (see [**16**] for details), we get that weak Bernoulli is equivalent to: for every $\varepsilon > 0$ there is a $K_0(\varepsilon)$ such that for every $M > K > K_0$ we have

$$\int_{X/\alpha^-} \sum_{D \in \alpha[K,M]} |\mu_C(D) - \mu(D)|\, d\mu_{\alpha^-}(C) < \varepsilon. \tag{5.3}$$

What do past fibres $C \in \alpha^-$ look like? Suppose x and y are in the same atom of $S^{-j}\alpha$ for $-\infty < j \le 0$. The argument in the proof of theorem 1 shows that x and y can only differ in the E^u direction. Thus C is a translate of a piece of E^u with diameter $<\phi$. Similarly, each $D \in \alpha^+$ is a translate of a piece of E^s with diameter $<\phi$.

We need a finite approximation of the last statement. The proof of theorem 1 shows that given $\delta > 0$, there is an $L_0(\delta)$ such that for $L > L_0$, a collection $\mathscr{D}$ of atoms D in $\alpha[0, L]$ has $\mu(\bigcup \mathscr{D}) > 1 - \delta$ and each $D \in \mathscr{D}$ is contained in a translate of

$$B = B^s(\phi) \oplus B^c(\delta) \oplus B^u(\rho^{-L}\phi).$$

Note the thinness of B in the E^c direction.

Lebesgue measure μ on $\mathbb{R}^n$ splits into $\mu_s \times \mu_c \times \mu_u$ on $E^s \oplus E^c \oplus E^u$. Lift B to $\tilde{B}$ in $\mathbb{R}^n$. Suppose $C \in \alpha^-$ has $C \cap S^{-K}D \neq \varnothing$, where $D \in \mathscr{D}$ is contained in $p\tilde{t} + B$. Then

$$C \subset p(S^{-K}\tilde{t} + S^{-K}\tilde{B} + \tilde{B}^u(\varphi)).$$

Now

$$\begin{aligned}\mu(S^{-K}\tilde{B}+\tilde{B}^u(\phi)) &\le \mu(S^{-K}\tilde{B}^s(\phi)\oplus S^{-K}\tilde{B}^c(\delta)\oplus \tilde{B}^u(2\phi))\\ &\le \mu_s(\tilde{B}^s(\xi^K\phi))\mu_c(\tilde{B}^c(K^n\delta))\mu_u(\tilde{B}^u(2\phi))\\ &\le \Delta\xi^{Kn}(K^n\delta)^n,\end{aligned}$$

where Δ is a constant independent of K and δ. For fixed K, by choosing δ small enough we get this less than $\frac{1}{10}$.

For $D\in\mathcal{D}$ let $D^*=\bigcup\{C\in\alpha^-: C\cap S^{-K}D=\emptyset\}$. Hence for $C\subset D^*$ we have $\mu_C(S^{-K}D)=0$. Letting $M=K+L$, a typical atom D' in $\alpha[K,M]$ has the form $S^{-K}D$, $D\in\alpha[0,L]$. The above shows that

$$\mu(D^*)>1-\mu(S^{-K}\tilde{B}+\tilde{B}^u(\phi))>\tfrac{9}{10}$$

for $D\in\mathcal{D}$ if $L>L_0(\delta)$. Hence

$$\begin{aligned}&\int_{X/\alpha^-}\sum_{D'\in\alpha[K,M]}|\mu_C(D')-\mu(D')|\,d\mu_{\alpha^-}(C)\\ &\quad>\sum_{D\in\mathcal{D}}\int_{X/\alpha^-}|\mu_C(S^{-K}D)-\mu(D)|\,d\mu_{\alpha^-}(C)-2\delta\\ &\quad>\sum_{D\in\mathcal{D}}\int_{D^*/\alpha^-}|\mu_C(S^{-K}D)-\mu(D)|\,d\mu_{\alpha^-}(C)-2\delta-\tfrac{1}{10}\\ &\quad=\sum_{D\in\mathcal{D}}\int_{D^*/\alpha^-}\mu(D)\,d\mu_{\alpha^-}(C)-2\delta-\tfrac{1}{10}\\ &\quad>\tfrac{9}{10}(1-\delta)-2\delta-\tfrac{1}{10}>\tfrac{1}{2}.\end{aligned}$$

Thus for every K there is an $M=K+L$ such that $\alpha[K,M]$ is not $\frac{1}{2}$-independent of α^-, completing the proof. □

COROLLARY 5.1. *If S is a quasihyperbolic automorphism and α is a fine partition, then*

$$\sum_{k=1}^{\infty}|h(\alpha|\alpha[1,k])-h(S,\alpha)|=\infty.$$

Proof. Del Junco and Rahe [**5**] have shown that if conditional entropy $h(\alpha|\alpha[1,k])$ converges to $h(S,\alpha)$ fast enough to make the discrepancy summable, then α is weak Bernoulli for S. □

6. *The limits of almost weak Bernoulli*

In his proof that ergodic toral automorphisms are Bernoulli, Katznelson introduced the notion of an almost weak Bernoulli partition, a property lying between weak and very weak Bernoulli. The previous section explains why he was forced to a property weaker than weak Bernoulli. Below is a quantitative analysis of how far almost weak Bernoulli can be pushed.

A partition α is *almost weak Bernoulli* for S if for every $\varepsilon>0$ there is a $K_0(\varepsilon)$ such that for every $K>K_0$ and $N>0$ we have that $\alpha[-N,0]$ is ε-independent of $\alpha[K,K^2]$. Let dep $(\alpha^-,\alpha[K,M])$ denote the quantity in (5.3). As before, letting $N\to\infty$ shows that an equivalent formulation is

$$\text{dep}\,(\alpha^-,\alpha[K,K^2])\to 0\quad\text{as } K\to\infty.$$

This property easily implies very weak Bernoulli. Katznelson showed that smooth partitions are almost weak Bernoulli, so they generate Bernoulli shifts. A careful look at his proof shows that it actually yields more. There is an $r > 1$ such that

$$\text{dep}\,(\alpha^-, \alpha[K, r^K]) \to 0 \quad \text{as } K \to \infty,$$

so the amount of future allowed can grow exponentially in the gap size. However, for central skew automorphisms, this fails for all large enough r.

THEOREM 5. *Suppose S is central skew. Then there are numbers $1 < r_1 < r_2$ such that if α is a smooth fine partition, then*

$$\lim_{K\to\infty} \text{dep}\,(\alpha^-, \alpha[K, r^K]) = 0 \quad (1 < r < r_1), \tag{6.1}$$

$$\liminf_{K\to\infty} \text{dep}\,(\alpha^-, \alpha[K, r^K]) > \tfrac{1}{2} \quad (r_2 < r). \tag{6.2}$$

Proof. To prove (6.1), note that if K^2 is replaced by r^K in the proof of theorem 1 of [**8**], the inequality (19) there becomes

$$\|v_1(a)\| > C_2 \rho^K r^{-n(n+l+1)K},$$

where $\|v_1(a)\| \to 0$ as $K \to \infty$. This still leads to an inconsistency if

$$1 < r < r_1 = \rho^{1/n(n+l+1)}.$$

This argument applies to all ergodic automorphisms.

For simplicity, we prove (6.2) for the lowest dimensional central skew automorphism. The same idea works for all of them.

Let S_0 be the companion matrix of

$$x^4 + 4x^3 - 2x^2 + 4x + 1,$$

and $S = J(S_0, 2)$, a central skew automorphism of $\mathbb{T}^8$ (see [**9**]). Here $E^c = E_{11} \oplus E_{12}$, and $S|E^c$ has matrix $J(R, 2)$, where R is an isometry. The idea is that the shearing of S in E^c cuts down the measure of intersections at a predictable rate (unlike the ergodic theorem used in § 5), giving an explicit bound for the $L_0(\delta)$ in § 5.

As above, each $C \in \alpha^-$ is a translate of a piece of E^u of diameter $< \varphi$. The argument in [**9**] shows that each $D \in \alpha[0, L]$ is contained in a translate of

$$B^s(\phi) \oplus B^c_{11}(\phi) \oplus B^c_{12}(2\phi/L) \oplus B^u(\phi\rho^{-L}).$$

Note that the size of $B^c_{12}(2\phi/L)$ decreases like L^{-2}.

For $D \in \alpha[0, r^K]$ put $D^* = \bigcup\{C \in \alpha^- : C \cap S^{-K}D = \varnothing\}$. As before,

$$\begin{aligned}\mu(D^*) &< \mu_s(B^s(\xi^K\phi))\mu_{11}(B^c_{11}(2\phi))\mu_{12}(B^c_{12}(2\phi r^{-K}))\mu_u(B^u(\phi)) \\ &< \Delta\xi^{2K} r^{-2K}\end{aligned}$$

where Δ depends only on S. The same estimate as in the proof of theorem 4 shows that if $r > \xi$, then

$$\text{dep}\,(\alpha^-, \alpha[K, r^K]) > \tfrac{1}{2}$$

for large enough K. Thus we can let $r_2 = \xi$. □

For the automorphism of $\mathbb{T}^8$ described above, this shows that for every smooth fine α,

$$\lim_{K\to\infty} \operatorname{dep}(\alpha^-, \alpha[K, (1.013)^K]) = 0,$$

while

$$\liminf_{K\to\infty} \operatorname{dep}(\alpha^-, \alpha[K, (4.612)^K]) > \tfrac{1}{2}.$$

Since ε-independence can only be improved by lumping atoms, there is a critical number $r(S, \alpha)$ such that (6.1) holds if $r < r(S, \alpha)$ while it fails if $r > r(S, \alpha)$. If S is central skew, then $r(S, \alpha)$ is uniformly bounded by r_2 for smooth fine α, while if S is hyperbolic then $r(S, \alpha) = \infty$ by weak Bernoullicity. It is unknown to us whether $r(S, \alpha) < \infty$ if S is central spin. We conjecture that $r(S, \alpha)$ has the same value for all smooth fine α. The common number $r(S)$ would perhaps be an interesting invariant for quasihyperbolic automorphsims. What is its value?

7. *Exponential mixing*

Anosov diffeomorphisms preserving a smooth measure mix Lipschitz functions exponentially fast. On the combinatorial level, this is described in Bowen [**3**: 1.26]. This carries over to quasihyperbolic automorphisms.

LEMMA 7.1. *Let* $1 < r < \rho^{1/(n+1)}$, *and let*

$$H_k = \{j \in \mathbb{Z} : |j| \le r^k\}^n.$$

Then for sufficiently large k we have $H_k \cap S^{-k} H_k = \{0\}$.

Proof. There is a constant $b > 0$ such that

$$[-r^k, r^k]^n \subset B(br^k).$$

By Katznelson's lemma, there is a constant $C > 0$ such that if $0 \ne z \in \mathbb{Z}^n$, then

$$\operatorname{dist}(z, E^c \oplus E^u) > C\|z\|^{-n}.$$

Suppose $z \in H_k$ with $S^k z \in H_k$. Then $\|S^k z\| < br^k$, $\|\pi_s z\| < br^k$. Hence

$$\operatorname{dist}(S^k z, E^c \oplus E^u) > C\|S^k z\|^{-n} > Cb^{-n} r^{-nk}$$

while

$$\operatorname{dist}(S^k z, E^c \oplus E^u) = \|S^k \pi_s z\| < \rho^{-k} br^k.$$

Comparison shows

$$\left(\frac{\rho}{r^{n+1}}\right)^k < \frac{b^{n+1}}{C},$$

which fails for large k. □

Recall the definition of Lip^γ from §4. We abbreviate $\int f\,d\mu$ to $\mu(f)$.

THEOREM 6. *For $f, g \in \mathrm{Lip}^\gamma$, $\mu(f \cdot S^k g)$ converges to $\mu(f)\mu(g)$ exponentially fast.*

Proof. Use Fejér approximations as in the proof of Theorem 2. There are trigonometric polynomials f_k and g_k with frequencies in H_k such that

$$\mu(f_k) = \mu(f), \qquad \mu(g_k) = \mu(g),$$

and

$$\|f-f_k\|_\infty, \|g-g_k\|_\infty < C_1(r^k)^{-\gamma}.$$

Now

$$\mu(f\cdot S^k g)=\mu[(f-f_k)\cdot S^k g]+\mu[f_k\cdot S^k(g-g_k)]+\mu(f_k\cdot S^k g_k).$$

The first two terms are $O(r^{-\gamma k})$, while the third is eventually

$$\hat{f}_k(0)\hat{g}_k(0)=\mu(f)\mu(g)$$

since

$$\operatorname{supp}\hat{f}_k\cap\operatorname{supp}(S^k g_k)^\wedge\subset H_k\cap S^{-k}H_k=\{0\}. \qquad \square$$

Remarks. (1) For $f, g\in C^\infty(\Pi^n)$, lemma 7.1 plus standard approximations show that $\mu(f\cdot S^k g)$ converges to $\mu(f)\mu(g)$ faster than any exponential.

(2) Using lacunary absolutely convergent Fourier series, it is easy to construct continuous f and g for which $\mu(f\cdot S^k g)$ converges more slowly than a preassigned rate.

8. *Exponential decay of return probabilities*

A mapping between two measure spaces that also carry topologies is called *finitary* if it is continuous after removing appropriate invariant null sets from each space. Smorodinsky observed that if a transformation T of (Y, ν) is a finitary image of a Bernoulli shift, then T must necessarily have the following property. Let $V\subset Y$ contain an open set (up to a null set), and for $y\in V$ define

$$r_V(y)=\min\{j>0: T^j y\in V\}.$$

The property is that every such V has exponentially decaying return probabilities, i.e.

$$\nu\{y: r_V(y)=k\}\to 0$$

exponentially fast. Smorodinsky next constructed a mixing countable state Markov chain without this property for a particular choice of V, thereby giving a mixing Markov chain of finite entropy that is measurably but not finitarily isomorphic to a Bernoulli shift. Constrast this with the Keane–Smorodinsky theorem that all finite state Markov chains are finitarily isomorphic to Bernoulli shifts. Recently, Dan Rudolph has applied his criterion for finitarily Bernoulli [**15**] to show that a mixing countable state Markov chain is finitarily Bernoulli if and only if it has exponentially decaying return probabilities.

No one knows whether quasihyperbolic automorphisms are finitarily Bernoulli. One thing to check is the character of return probabilities. We show that these satisfy Smorodinsky's necessary condition, lending further support to the conjecture that ergodic group automorphisms are finitarily Bernoulli.

THEOREM 7. *An ergodic toral automorphism has exponentially decaying return probabilities on open sets.*

Proof. Let $V\subset\mathbb{T}^n$ contain (up to a null set) a cube whose complement we denote by U. Since

$$\{y\in V: r_V(y)=k\}=V\cap S^{-1}V^c\cap\cdots\cap S^{-k+1}V^c\cap S^{-k}V$$
$$\subset S^{-1}U\cap\cdots\cap S^{-k+1}U,$$

it suffices to prove that $\mu(\bigcap_{j=0}^k S^j U)\to 0$ exponentially fast. Using either the

Stone–Weierstrass theorem or Fejér approximation, it is easy to find a trigonometric polynomial g such that $g \geq \chi_U$ and $\int g\, d\mu = \eta < 1$. We interrupt the proof for a needed lemma.

Definition. A sequence $\{F_j\}_{-\infty}^{\infty}$ of subsets of an abelian group is *independent* if whenever $z_{j_i} \in F_{j_i}$ and $z_{j_1} + z_{j_2} + \cdots + z_{j_r} = 0$ with the j_i distinct, then each $z_{j_i} = 0$.

LEMMA 8.1. *Let S be an aperiodic automorphism of $\mathbb{Z}^n$, and suppose $F \subset \mathbb{Z}^n$ is finite. Then there is a positive integer m such that $\{S^{mj}F : j \in \mathbb{Z}\}$ is independent.*

Proof. We can assume $0 \in F$, and that F is symmetric. It is then enough to find m such that if $\sum_{j=0}^{N} S^{mj} z_j = 0$ with $z_j \in F$, then each $z_j = 0$. Let $\pi_s : \mathbb{R}^n \to E^s$ along $E^c \oplus E^u$. Since

$$(E^c \oplus E^u) \cap \mathbb{Z}^n = \{0\},$$

there are $\delta, \Delta > 0$ such that if $0 \neq z \in F$, then $\delta < \|\pi_s z\| < \Delta$. Pick m so that $\rho^{-m} < \delta/2\Delta$, where ρ is the same as from § 2.

Suppose $\sum_{j=0}^{N} S^{mj} z_j = 0$ with $z_0 \neq 0$. Then

$$\delta < \|\pi_s z_0\| = \left\| \sum_{j=1}^{N} S^{mj} \pi_s z_j \right\| \leq \sum_{j=1}^{\infty} \rho^{-mj} \Delta < 2\rho^{-m} \Delta,$$

contradicting the choice of m. □

We remark that Peters **[13]** has a more quantitative version of this kind of independence under powers of S, but for specific types of F.

Proof of theorem 7 (continued). Let $F = \operatorname{supp} \hat{g}$, a finite subset of $\mathbb{Z}^n$. Choose m as in lemma 8.1. Since

$$\{S^{mj}F : j \in \mathbb{Z}\}$$

is independent, the $S^{mj}g$ are uncorrelated random variables. Thus if $q = [k/m]$, then

$$\mu\left(\bigcap_{j=0}^{k} S^j U\right) \leq \mu\left(\bigcap_{i=0}^{q} S^{mi} U\right) = \int \prod_{i=0}^{q} S^{mi} \chi_U \, d\mu$$

$$\leq \int \prod_{i=0}^{q} S^{mi} g \, d\mu = \prod_{i=0}^{q} \int S^{mi} g \, d\mu = \eta^{q+1} < (\eta^{1/m})^k. \qquad □$$

Remark. Lemma 8.1 and theorem 7 are valid for ergodic automorphisms of general compact metrizable abelian groups. The proof would take us too far afield, and so is deferred to another paper.

9. *Questions and conjectures*

Here are some questions, and more or less likely answers, that cropped up during this work.

(1) Is $r(S, \alpha)$ of § 6 the same for all smooth fine α? If so, what is its value? Conjecture: Yes, with $r(S, \alpha) = h(S)$.

(2) Are there geometric proofs for the geometric statements proved here by harmonic analysis? In particular, it would be nice to have geometric proofs of theorem 2, corollary 4.2, and theorem 6.

(3) Livsic showed that a Lipschitz function g on a basic set for an Axiom A diffeomorphism is cohomologous to 0 (i.e. $f = g \circ S - g$ for some Lipschitz g) iff the sum of the values of f over every periodic orbit is 0. Is the same true for quasihyperbolic automorphisms?

(4) Marcus [**11**] proved that for quasihyperbolic S the periodic orbit measures are weakly dense in the set of all invariant probability measures. However, are 'most' periodic orbits nearly uniformly distributed in some sense? If ν is an S-invariant measure, is there a relationship between the asymptotic number of periodic orbit measures in P_k 'close' to ν and $h_\nu(S)$?

(5) The proposition in § 4 shows there is a compact null K containing at least one point from each periodic orbit. Is there such a K containing exactly one point from each periodic orbit? Can such a K be countable?

(6) In theorem 6, we conjecture that if $f, g \in \mathrm{Lip}^1$, then

$$\lim_{k\to\infty} -\frac{1}{k} \log |\mu(f \cdot S^k g) - \mu(f)\mu(g)| \le h(S),$$

and that $h(S)$ is best possible.

The research was partially supported by NSF Grant no MCS 7704915.

REFERENCES

[**1**] R. L. Adler & B. Weiss. *Similarity of Automorphisms of the Torus.* Memoirs Amer. Math. Soc. **98** (1970).
[**2**] Rufus Bowen. Periodic points and measures for Axiom A diffeomorphisms. *Trans. Amer. Math. Soc.* **154** (1971), 377–397.
[**3**] Rufus Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms.* Lecture Notes in Math. no. 470. Springer: Berlin, 1975.
[**4**] Rufus Bowen. Smooth partitions of Anosov diffeomorphisms are weak Bernoulli. *Israel J. Math.* **21** (1975), 95–100.
[**5**] A. del Junco & M. Rahe. Finitary codings and weak Bernoulli partitions. *Proc. Amer. Math. Soc.* **75** (1974), 259–264.
[**6**] A. O. Gelfond. *Transcendental and Algebraic Numbers.* Dover: New York, 1960.
[**7**] Yitzhak Katznelson. *An Introduction to Harmonic Analysis.* Wiley: New York, 1968.
[**8**] Yitzhak Katznelson. Ergodic automorphisms of T^n are Bernoulli shifts. *Israel J. Math.* **10** (1971), 186–195.
[**9**] D. A. Lind. *Ergodic Group Automorphisms and Specification.* Lecture Notes in Math. no. 729, pp. 93–104. Springer: Berlin, 1979.
[**10**] D. A. Lind. Finitarily splitting skew products. In *Ergodic Theory and Dynamical Systems I* (A. B. Katok, ed.), pp. 65–80. Birkhäuser: Boston, 1981.
[**11**] Brian Marcus. A note on periodic points of toral automorphisms. *Monatsh. Math.* **89** (1980), 121–129.
[**12**] Donald Ornstein. *Ergodic Theory, Randomness, and Dynamical Systems.* Yale Math. Monographs 5. Yale University Press: New Haven, 1974.
[**13**] Justin Peters. A discrete analogue of a theorem of Katznelson. *Israel J. Math.* **37** (1980), 251–255.
[**14**] V. A. Rohlin. On the fundamental ideas of measure theory. *Amer. Math. Soc. Translations*, Ser. 1, **10** (1962), 1–54.
[**15**] Daniel J. Rudolph. A characterization of those processes finitarily isomorphic to a Bernoulli shift. In *Ergodic Theory and Dynamical Systems I* (A. B. Katok, ed.), pp. 1–64. Birkhäuser: Boston, 1981.

[**16**] Paul C. Shields. Weak and very weak Bernoulli partitions. *Monatsh. Math.* **84** (1972), 133–142.
[**17**] S. Smale. Differentiable dynamical systems. *Bull. Amer. Math. Soc.* **73** (1967), 747–817.
[**18**] M. Smorodinsky. A partition on a Bernoulli shift which is not weak Bernoulli. *Math. Syst. Th.* **5** (1971), 201–203.
[**19**] Kenneth B. Stolarsky. *Algebraic Numbers and Diophantine Approximation.* Dekker: New York, 1974.