ISRAEL JOURNAL OF MATHEMATICS 122 (2001), 29–42

p-ADIC FOLIATION AND EQUIDISTRIBUTION

BY

ELON LINDENSTRAUSS*

School of Mathematics, Institute for Advanced Study
Olden lane, Princeton, NJ 08540, USA
and
Department of Mathematics, Stanford University
Stanford, CA 94305, USA
e-mail: elon@ias.edu

ABSTRACT

We show that if μ is a measure on $\mathbb{R}/\mathbb{Z}$ ergodic under the $\times m$ map with positive entropy, then μ-a.s. $\{a_n x\}$ is equidistributed, for a significantly larger collection of integer sequences a_n than was previously known. In particular, we show that μ-a.s. $\{r^n x\}$ is equidistributed as long as m does not divide any power of r (this was previously known only if r and m are relatively prime). The proof uses the p-adic analogue of results from the theory of smooth dynamical systems.

1. Introduction

The starting point of this paper are theorems by D. Rudolph and A. Johnson [15], [6], and a sharpening of Rudolph's theorem by B. Host [4]. For a more detailed history of this problem, see [13]. Some additional references can be found in the bibliography.

Let r, s be two integers. What D. Rudolph has shown for r, m relatively prime, and A. Johnson for r, m multiplicatively independent (i.e. there is no $s \in \mathbb{Z}$ so that both r and m are powers of s) is that any measure μ on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$

1. invariant both under the $\times m$ map and the $\times r$ map,
2. ergodic with respect to the $\mathbb{Z}_+^2$ action generated by these maps,

* The author acknowledges the support of NSF grant DMS 97-29992.
Received January 10, 2000

3. and so that μ has positive entropy with respect to the $\mathbb{Z}_+$ action generated by the $\times m$ map,

is Lebesgue measure. This gives partial confirmation to a conjecture of H. Furstenberg from 1967, namely that any measure satisfying only 1 and 2 above is either Lebesgue or is supported on a finite number of points.

Host sharpened Rudolph's (but not Johnson's) result, and while the theorems of Rudolph and Johnson show that there are no measures satisfying certain conditions other than Lebesgue measure, Host's theorem is more constructive in the sense that it proves that a large collection of measures satisfies a certain regularity property. In this paper we show how Johnson's theorem can be strengthened in the same way:

THEOREM 1.1: *Suppose that m does not divide any power of r, and let μ be a measure on $\mathbb{T}$ invariant and ergodic under the $\times m$ map. Then for μ a.e. $x \in \mathbb{T}$, $(\{r^n x\})_{n=1}^\infty$ is equidistributed in $\mathbb{T}$ (with respect to the usual Lebegue measure).*

The case m relatively prime to r was proved by Host [4].

While the condition on m and r in Theorem 1.1 is more restrictive than being multiplicatively independent, the general case of Johnson's theorem still follows, since if m and r are multiplicatively independent the semigroup generated by them contains an m' and an r' such that m' does not divide any power of r'. Another point worth mentioning is that since a $(\times r, \times m)$ ergodic measure on $\mathbb{T}$ need not be $\times m$ ergodic, some (rather simple) argument must be given to show that this theorem does indeed imply Johnson's theorem. This argument, however, is the same as the one showing that Host's theorem implies Rudolph's theorem.

For the special case of μ Bernoulli, i.e., that the distribution μ induced on the digits of the base m expansion is i.i.d. (or, more generally, if this process is weakly Bernoulli), Feldman and Smorodinsky [2] proved that a.s. $(\{r^n x\})_{n=1}^\infty$ is equidistributed in $\mathbb{T}$ whenever m and r are multiplicatively independent.

In this formulation the roles of r and m are decidedly different. Indeed, r enters only as a parameter that determines the sequence $(r^n)_{n=1}^\infty$. And so, one may ask for what other sequences $(a_n)_{n=1}^\infty$ is it true that for every such μ, for μ-a.e. x, $(\{a_n x\})_{n=1}^\infty$ is equidistributed. A criterion when one might apply Host's original proof was found by D. Meiri [13]; roughly the requirement is that as m-adic integers a_n are evenly distributed (we denote the m-adic integers by $\mathbb{Z}_m$). A slight variation of Meiri's condition is the notion we call having m-adic subexponential cells (Definition 3.1). Meiri also showed that many natural sequences (such as $r^n + s^n$ where r and s are relatively prime to m, or even r^{s^n})

satisfy these conditions.

Unfortunately, when r and m are not relatively prime the distribution of r^n in $\mathbb{Z}_m$ is not sufficiently regular for Host's proof to work. To deal with this sequence, we must take into account the common prime factors of r and m. Let $r = p_1^{a_1} \cdots p_n^{a_n}$ and $m = q_1^{b_1} \cdots q_\ell^{b_\ell}$ be the prime decomposition of r and m respectively, and take $m_1 = \prod_{q_i \nmid r} q_i^{b_i}$, $m_2 = \prod_{q_i | r} q_i^{b_i}$. Now, while the sequence $(r^n)_{n=1}^\infty$ is not sufficiently evenly distributed in $\mathbb{Z}_m = \mathbb{Z}_{m_1} \oplus \mathbb{Z}_{m_2}$, in $\mathbb{Z}_{m_1}$ this sequence is evenly distributed, and in $\mathbb{Z}_{m_2}$ the sequence $(r^n)_{n=1}^\infty$ tends very rapidly to zero. More generally, one can consider sequences a_n that, even though they are not necessarily evenly distributed in $\mathbb{Z}_m$, have the property that the conditional distribution of $a_n \in \mathbb{Z} \subset \mathbb{Z}_{m_1} \oplus \mathbb{Z}_{m_2}$ in $\mathbb{Z}_{m_1}$ given its projection to $\mathbb{Z}_{m_2}$ (more precisely, the distribution of $a_n \bmod m_1^k$ given $a_n \bmod m_2^k$) is sufficiently regular. The precise condition we need, of having (almost) m_1-adic subexponential cells rel m_2, is given in Definition 3.1. Naturally, this is satisfied in the case of $\{r^n\}$ where r, m, m_1, m_2 are as above, since then $a_n \bmod m_2^k$ is essentially constant — only a small finite number of the a_n are not congruent to $0 \bmod m_2^k$. Our main result is that as long as $m_1 \neq 0$, such sequences are m-Host:

THEOREM 1.2: *Let $\{a_i\}$ be a sequence with m_1-adic subexponential cells relative to m_2. Then for any measure μ invariant and ergodic under $\times m$ (with $m = m_1 m_2$) of positive entropy, for μ-a.e. x, $\{a_n x\}$ is equidistributed in $\mathbb{T}$ (with respect to Lebesgue measure).*

As was already observed by A. Katok and R. Spatzier in [8], one can consider the two sided extension $\bar{\mathbb{T}}_m$ of $(\mathbb{T}, \times m)$ as a foliated space, with two sub-foliations corresponding to m_1 and m_2 respectively. For example, the leaf of the m_1-foliation containing the point $(\ldots, a_{-2}, a_{-1}, a_0, a_1, a_2, \ldots) \in \bar{\mathbb{T}}_m$ (with each $a_i \in \{0, \ldots, m-1\}$) is the set of all $(\ldots, b_{-2}, b_{-1}, b_0, b_1, b_2, \ldots)$ such that:

1. $a_i = b_i$ for all i larger than some N;
2. for every $M < N$,

$$\sum_{i=M}^{N} m^{-i+N} a_i \equiv 0 \pmod{m_2^{N-M+1}}.$$

Note that each leaf is isomorphic in a natural way to $\mathbb{Q}_{m_1}$.

To be able to apply a modified version of Host's proof in this case, we need first to define the entropy related to the m_1-adic foliation, and then show that if the entropy of μ is positive the same is true for the part of entropy related to this foliation. The first step is proving the analog of a theorem of Ledrappier and Young [10] on the relationship between entropy, Hausdorff dimension and

Liapunov exponents for diffeomorphism (i.e. smooth dynamical systems). Our proof is similar in some respects to the proof of a special case of Ledrappier and Young's result by R. Kenyon and Y. Peres in [9]. Then one uses the specific geometry of the space to give additional constraints on how the entropy is divided according to these two sub-foliations.

An alternative approach to extending Host's theorem can be found in [11] via convolutions. This approach can be applied more generally than the method we develop here, but does not give the equidistribution of $\{a_n x\}$ but only a weaker average result (there are also other applications of the convolution results of [11] that are not related to Host's theorem).

2. p-adic foliation

Let $m = m_1 m_2$, with $(m_1, m_2) = 1$. If $x \in \mathbb{T}$, we will identify x with its expansion to base m. We will use the notation $x_{a\cdots b} = x_a \cdots x_b$ (considered either as a string of $b-a+1$ digits or an element of $\mathbb{Z}/m^{b-a+1}\mathbb{Z}$) and $x^{(i)}_{1\cdots b} = x_{1\cdots b} \bmod m_i^b$ for $i = 1, 2$. We set $\alpha_{a\cdots b}$ to be the σ-algebra generated by the function $x \mapsto x_{a\cdots b}$ and $\alpha^{(i)}_{a\cdots b}$ the σ-algebra generated by $x \mapsto x^{(i)}_{a\cdots b}$. Note that unless $b = \infty$, both $\alpha_{a\cdots b}$ and $\alpha^{(i)}_{a\cdots b}$ are finite algebras of sets. All these definitions can be equally well applied when $m = m_1 m_2 \cdots m_d$ with the m_ℓ pairwise relatively prime; in this case i can have the values $1, \ldots, d$. If α is a finite algebra of sets and β is a σ-algebra, we set

$$\mu(\alpha|\beta)(x) = \mathbf{E}(1_{\alpha(x)}(y)|\beta)(x),$$

where $\alpha(x)$ denotes the set in α containing x.

THEOREM 2.1: *Let μ be a m-invariant measure on $\mathbb{T}$, with $m = m_1 m_2$, $(m_1, m_2) = 1$. Then for a.e. x the following limits exist and are constant a.e.:*

$$h_1 = \lim \frac{-\log \mu(\alpha^{(1)}_{1\cdots n} \mid \alpha_{n+1\cdots\infty})}{n}, \tag{2.1}$$

$$h_2 = \lim \frac{-\log \mu(\alpha^{(2)}_{1\cdots n} \mid \alpha^{(1)}_{1\cdots n} \vee \alpha_{n+1\cdots\infty})}{n}. \tag{2.2}$$

These limits exist also in $L^1(\mu)$. Furthermore, if $h(\mu)$ is the entropy of μ with respect to the times m map, then $h(\mu) = h_1 + h_2$.

Remark: This theorem is in complete analogy to (one part of) a theorem of Ledrappier and Young on the relations between entropy, Lyapunov exponents and Hausdorff dimension for diffeomorphism. There one uses the sub-foliation of the stable or unstable manifolds to divide the entropy into parts corresponding to every positive (or every negative) Lyapunov exponent; here we divide the entropy into parts corresponding to the p-adic foliations modulo m_1 and m_2.

The following lemma is completely standard:

LEMMA 2.2: *Let $\mathcal{P}$ be a finite partition of a measure space $(X, \mathcal{B}, \mu)$. Then for any set $E \subset X$,*

$$-\int_E \log \mu(\mathcal{P}(x)) \leq -\mu(E) \log \mu(E) + \mu(E) \log |\mathcal{P}|.$$

Proof: Indeed, set $\mathcal{P}' = \mathcal{P} \vee \{E, E^C\}$. The lemma follows from the following calculation,

$$\begin{aligned}
-\int_E \log \mu(\mathcal{P}(x)) &\leq -\int_E \log \mu(\mathcal{P}'(x)) \\
&= -\sum_{P \in \mathcal{P}} \mu(P \cap E) \log \mu(P \cap E) \\
&= -\mu(E) \log \mu(E) - \mu(E) \sum_{P \in \mathcal{P}} \mu(P|E) \log \mu(P|E) \\
&\leq -\mu(E) \log \mu(E) + \mu(E) \log |\mathcal{P}|,
\end{aligned}$$

where $\mu(P|E)$ denotes

$$\mu(P|E) = \frac{\mu(P \cap E)}{\mu(E)}. \qquad ∎$$

Proof of Theorem 2.1: By a variant of the Shannon–McMillen–Brieman theorem,

$$\frac{-\log \mu(\alpha_{1\cdots n} \mid \alpha_{n+1\cdots\infty})}{n} \longrightarrow h(\mu)$$

both pointwise and in L^1. Furthermore,

$$\mu(\alpha_{1\cdots n} \mid \alpha_{n+1\cdots\infty}) = \mu(\alpha^{(1)}_{1\cdots n} \mid \alpha_{n+1\cdots\infty}) \mu(\alpha^{(2)}_{1\cdots n} \mid \alpha^{(1)}_{1\cdots n} \vee \alpha_{n+1\cdots\infty}),$$

and so it suffices to prove that the limit in (2.2) exists.

Consider now the solenoid $\bar{\mathbb{T}}_m$, i.e. the two sided extension of $(\mathbb{T}, \times m)$. Let T denote the extension of the $\times m$ map to $\bar{\mathbb{T}}_m$. The measure μ can be extended to a T invariant measure $\bar{\mu}$ on $\bar{\mathbb{T}}_m$. We can extend our notation $\bar{x}_{a\cdots b}$ also to $\bar{x} \in \bar{\mathbb{T}}_m$ (note that a and b may be negative), and again we set

$$\bar{x}^{(i)}_{a\cdots b} = \bar{x}_{a\cdots b} \bmod m_i^{b-a+1},$$

and let $\bar{\alpha}_{a\ldots b}$ and $\bar{\alpha}^{(i)}_{a\ldots b}$ be the corresponding σ-algebras. Notice that for $a < a' < b$,

$$\bar{x}^{(i)}_{a'\cdots b} = \bar{x}^{(i)}_{a\cdots b} \bmod m_i^{b-a'+1}.$$

Thus $\bar{\alpha}^{(i)}_{a\cdots b}$ refines $\bar{\alpha}^{(i)}_{a'\cdots b}$, and we define

$$\bar{\alpha}^{(i)}_{-\infty\cdots b} = \bigvee_{a=0}^{\infty} \bar{\alpha}^{(i)}_{-a\cdots b}.$$

Set

$$\rho(\bar{x}) = -\log \bar{\mu}(\bar{\alpha}^{(2)}_0 \mid \bar{\alpha}^{(1)}_{-\infty\cdots 0} \vee \bar{\alpha}_{1\cdots\infty})(\bar{x}).$$

Clearly $\rho(\bar{x})$ is non-negative and in L^1; indeed, $\int \rho(\bar{x}) d\bar{\mu}(\bar{x}) \leq \log m_2$. We note that by the Chinese remainder theorem

$$\bar{\alpha}^{(1)}_{-\infty\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty} \vee \bar{\alpha}^{(2)}_{k+1\cdots n} = \bar{\alpha}^{(1)}_{-\infty\cdots k} \vee \bar{\alpha}_{k+1\cdots\infty},$$

and so we see that

$$\begin{aligned}
-\log \bar{\mu}(\bar{\alpha}^{(2)}_{1\cdots n} \mid \bar{\alpha}^{(1)}_{-\infty\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(\bar{x}) = &-\log \bar{\mu}(\bar{\alpha}^{(2)}_{n} \mid \bar{\alpha}^{(1)}_{-\infty\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(\bar{x}) \\
&-\log \bar{\mu}(\bar{\alpha}^{(2)}_{n-1} \mid \bar{\alpha}^{(1)}_{-\infty\cdots n-1} \vee \bar{\alpha}_{n\cdots\infty})(\bar{x}) \\
&\;\vdots \\
&-\log \bar{\mu}(\bar{\alpha}^{(2)}_{1} \mid \bar{\alpha}^{(1)}_{-\infty\cdots 1} \vee \bar{\alpha}_{2\cdots\infty})(\bar{x}) \\
&= \sum_{i=1}^{n} \rho(T^i \bar{x}).
\end{aligned} \tag{2.3}$$

By the ergodic theorem the limit

$$L(\bar{x}) = \lim \frac{-\log \bar{\mu}(\bar{\alpha}^{(2)}_{1\cdots n} \mid \bar{\alpha}^{(1)}_{-\infty\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(\bar{x})}{n}$$

exists both pointwise and in L^1; furthermore, $L(\bar{x}) = L$ is a constant a.e. Thus the theorem will follow once we show that a.e.

$$\lim_{n\to\infty} \frac{1}{n} \left| \frac{\log \mu(\bar{\alpha}^{(2)}_{1\cdots n} \mid \bar{\alpha}^{(1)}_{1\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(\bar{x})}{\log \mu(\bar{\alpha}^{(2)}_{1\cdots n} \mid \bar{\alpha}^{(1)}_{-\infty\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(\bar{x})} \right|.$$

Let

$$Z_n = \frac{\mu(\bar{\alpha}^{(2)}_{1\cdots n} \mid \bar{\alpha}^{(1)}_{1\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(x)}{\mu(\bar{\alpha}^{(2)}_{1\cdots n} \mid \bar{\alpha}^{(1)}_{-\infty\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(x)}.$$

Then

$$\begin{aligned}\mathbf{E}(Z_n) &= \mathbf{E}(\mathbf{E}(Z_n \mid \bar{\alpha}^{(1)}_{-\infty\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})) \\ &= \mathbf{E}\Big(\sum_{a\in\mathbb{Z}/m_2^n} \mu(\{\bar{x}^{(2)}_{1\cdots n} = a\} \mid \bar{\alpha}^{(1)}_{1\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty}) \Bigm| \bar{\alpha}^{(1)}_{-\infty\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty}\Big) \\ &= 1.\end{aligned}$$

Therefore, for every $\epsilon > 0$, the series $\sum_{k=1}^{\infty} \mu(Z_k \geq e^{\epsilon k})$ converges and the Borel–Cantelli lemma implies that $\overline{\lim}_{k\to\infty}(1/k)\log Z_k \leq \epsilon$ a.s., and so since ϵ is arbitrary,

$$\underline{\lim}_{n\to\infty} \frac{-\log\mu(\bar{\alpha}^{(2)}_{1\cdots n} \mid \bar{\alpha}^{(1)}_{1\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(\bar{x})}{n} \geq \lim_{n\to\infty} \frac{-\log\mu(\bar{\alpha}^{(2)}_{1\cdots n} \mid \bar{\alpha}^{(1)}_{-\infty\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(\bar{x})}{n} = L.$$

To see the opposite inequality, consider

$$\rho_\ell(\bar{x}) = -\log\bar{\mu}(\bar{\alpha}^{(2)}_0 \mid \bar{\alpha}^{(1)}_{-\ell\cdots 0} \vee \bar{\alpha}_{1\cdots\infty}).$$

Since for any a the family of random variables $\mu(\{\bar{x}^{(2)}_0 = a\} \mid \bar{\alpha}^{(1)}_{-\ell\cdots 0} \vee \bar{\alpha}_{1\cdots\infty})$ for $\ell = 0, 1, \ldots$ is a martingale, by the martingale convergence theorem, $\rho_\ell(\bar{x}) \to \rho(\bar{x})$ as $\ell \to \infty$ for a.e. $\bar{x}$. To see that $\rho_\ell(\bar{x}) \to \rho(\bar{x})$ also in L^1, it suffices to show that the ρ_ℓ are uniformly integrable. Let Q be of small measure (say $\bar{\mu}(Q) = \delta$). We cover Q by two sets:

$$Q' = \left\{x\colon \mathbf{E}(1_Q \mid \bar{\alpha}^{(1)}_{-\ell\cdots 0} \vee \bar{\alpha}_{1\cdots\infty})(x) > \delta^{1/2}\right\},$$

and $Q'' = Q \setminus Q'$. Note that $\mu(Q')$ is at most $\delta^{1/2}$. We need to show that if δ is small then $\int_Q \rho_\ell(\bar{x})d\bar{\mu}(\bar{x})$ is small, uniformly in ℓ. This follows from

$$\begin{aligned}\int_{Q'} \rho_\ell(\bar{x})d\bar{\mu}(\bar{x}) = &- \int_{Q'}\Big[\sum_a \bar{\mu}(\bar{\alpha}^{(2)}_0 \mid \bar{\alpha}^{(1)}_{-\ell\cdots 0} \vee \bar{\alpha}_{1\cdots\infty})(x) \times \\ &\times \log\bar{\mu}(\bar{\alpha}^{(2)}_0 \mid \bar{\alpha}^{(1)}_{-\ell\cdots 0} \vee \bar{\alpha}_{1\cdots\infty})(x)\Big] d\bar{\mu}(x) \\ &\leq \bar{\mu}(Q')\log m_2 \leq \delta^{1/2}\log m_2,\end{aligned}$$

and, using Lemma 2.2,

$$\int_{Q''} \rho_\ell(x) \leq -\delta^{1/2}\log\delta^{1/2} + \delta^{1/2}\log m_2.$$

Thus both of these integrals tend to 0 as $\delta \to 0$, uniformly in ℓ.

Fix $\epsilon > 0$, and using the above discussion we find an ℓ such that $\|\rho_\ell - \rho\|_1 < \epsilon$. We now compare $-\log \mu(\bar{\alpha}^{(2)}_{1\cdots n} \mid \bar{\alpha}^{(1)}_{-\infty\ldots n} \vee \bar{\alpha}_{n+1\cdots\infty})(x)$ with $\sum_{i=1}^n \rho_\ell(T^i\bar{x})$.

First note that by the ergodic theorem,

$$\frac{1}{n}\sum_{i=1}^{n} \rho_\ell(T^i\bar{x}) \to \int \rho_\ell(\bar{y})\, d\bar{\mu}(\bar{y}),$$

which is within ϵ from L. Set

$$R_n = \frac{\prod_{i=1}^n \exp(\rho_\ell(T^i\bar{x}))}{\bar{\mu}(\bar{\alpha}^{(2)}_{1\cdots n} \mid \bar{\alpha}^{(1)}_{-\ell+1\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(\bar{x})}.$$

As before, it is easy to see that $\mathbf{E}(R_n) = 1$, and so $\overline{\lim}_{n\to\infty}(1/n)\log R_n \le 0$ a.s. Combined with $\|\rho_\ell - \rho\|_1 < \epsilon$ we get

$$\overline{\lim}\, \frac{-\log \bar{\mu}(\bar{\alpha}^{(2)}_{1\cdots n} \mid \bar{\alpha}^{(1)}_{-\ell+1\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(\bar{x})}{n} \le L + \epsilon.$$

To conclude the proof, we apply the same trick once again with

$$W_n = \frac{m_2^{-\ell}\bar{\mu}(\bar{\alpha}^{(2)}_{1\cdots n} \mid \bar{\alpha}^{(1)}_{-\ell+1\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(\bar{x})}{\bar{\mu}(\bar{\alpha}^{(2)}_{-\ell+1\cdots n} \mid \bar{\alpha}^{(1)}_{-\ell+1\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(\bar{x})}.$$

Again we see that $\mathbf{E}(W_n) = 1$ for which we deduce that a.s. for all n large enough,

$$-\frac{1}{n}\log \bar{\mu}(\bar{\alpha}^{(2)}_{1\cdots n} \mid \bar{\alpha}^{(1)}_{-\ell+1\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(\bar{x}) + \frac{\ell}{n}\log m_2 \ge$$
$$\ge -\frac{1}{n}\log \bar{\mu}(\bar{\alpha}^{(2)}_{-\ell+1\cdots n} \mid \bar{\alpha}^{(1)}_{1\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(\bar{x}) - \epsilon.$$

Thus for a.e. $\bar{x}$

$$\begin{aligned} L &\le \underline{\lim}_{n\to\infty} \frac{-\log \bar{\mu}(\bar{\alpha}^{(2)}_{-\ell+1\cdots n} \mid \bar{\alpha}^{(1)}_{-\ell+1\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(\bar{x})}{n} \\ &\le \overline{\lim}_{n\to\infty} \frac{-\log \bar{\mu}(\bar{\alpha}^{(2)}_{1\cdots n} \mid \bar{\alpha}^{(1)}_{-\ell+1\cdots n} \vee \bar{\alpha}_{n+1\cdots\infty})(\bar{x})}{n} + 2\epsilon \\ &\le L + 3\epsilon. \quad \blacksquare \end{aligned}$$

COROLLARY 2.3: *Let μ be a m-invariant measure on $\mathbb{T}$, with $m = m_1 m_2 \cdots m_d$ and the m_i pairwise relatively prime. Then for almost every x and $1 \le i \le d$ the following limit exists and is constant a.e.:*

$$h_i = \lim \frac{-\log \mu(\alpha^{(i)}_{1\cdots n} \mid \bigvee_{j=1}^{i-1} \alpha^{(j)}_{1\cdots n} \vee \alpha_{n+1\cdots\infty})}{n}.$$

Furthermore, $h(\mu) = \sum_{i=1}^{d} h_i$.

Proof: Let

$$I_n^{(i)}(x) = -\log \mu \left(\alpha_{1\cdots n}^{(i)} \mid \bigvee_{j=1}^{i-1} \alpha_{1\cdots n}^{(j)} \vee \alpha_{n+1\cdots\infty} \right).$$

Note that for $k \le d$

$$\sum_{i=1}^{k} I_n^{(i)}(x) = -\log \mu \left(\bigvee_{i=1}^{k} \alpha_{1\cdots n}^{(i)} \mid \alpha_{n+1\cdots\infty} \right). \tag{2.4}$$

In particular, for $d = n$ we get that

$$\frac{1}{n} \sum_{i=1}^{n} I_n^{(i)}(x) = -\frac{1}{n} \log \mu(\alpha_{1\cdots n} \mid \alpha_{n+1\cdots\infty}) \longrightarrow h(\mu),$$

so if we prove that the limits defining the h_i exist and are constant a.e., it follows that $\sum_{i=1}^{d} h_i = h(\mu)$.

Moreover, from Theorem 2.1 applied to $\bar{m}_1 = m_1 \cdots m_k$ and $\bar{m}_2 = m_{k+1} \cdots m_d$ it follows that the limit

$$\lim \frac{1}{n} \log \mu \left(\bigvee_{i=1}^{k} \alpha_{1\cdots n}^{(i)} \mid \alpha_{n+1\cdots\infty} \right)$$

exists and is constant a.e. Using equation (2.4) it follows that $(1/n)I_n^{(i)}(x)$, as the difference between two functions converging a.e. to constants, converges a.e. to a constant which is by definition h_i. ∎

THEOREM 2.4: *Let h_i be as in Corollary 2.3. Then for any $k \le d$,*

$$\sum_{i=k}^{d} h_i \le \frac{\sum_{i=k}^{d} \log m_i}{\log m} h(\mu).$$

In particular, since $\sum_{i=1}^{d} h_i = h(\mu)$, if $h(\mu) > 0$ so is h_1.

Proof: Set $\bar{m} = \prod_{i=k}^{d} m_i$. We first note that for any b, if $\ell = \lceil \frac{\log \bar{m}}{\log m} b \rceil$, then given $x_{b+1\cdots\infty}$ and $x_{1\cdots b}^{(i)}$ for all $i < k$, the value of $x_{1\cdots\ell}$ determines $x_{1\cdots b}^{(j)}$ for all $j \ge k$. Indeed, those $y \in \mathbb{T}$ for which $y_{1\cdots b}^{(i)} = x_{1\cdots b}^{(i)}$ for $i < k$ and $y_{b+1\cdots\infty} = x_{b+1\cdots\infty}$ are in a coset of the group

$$\left\{ \frac{k}{\bar{m}^b} : 0 \le k \le \bar{m}^b - 1 \right\}.$$

Since the distance between every two elements of this group is at least

$$\frac{1}{\bar{m}^b} \geq \frac{1}{m^\ell},$$

the first ℓ digits of y to base m uniquely determine it.

Now, letting $b \to \infty$, we conclude that

$$\begin{aligned}\sum_{j=k}^{d} h_j \longleftarrow & \frac{-1}{b} \int_{\mathbb{T}} \log \mu \left(\bigvee_{j=k}^{d} \alpha^{(j)}_{1\cdots b} \mid \bigvee_{i=1}^{k-1} \alpha^{(i)}_{1\cdots b} \vee \alpha_{b+1\cdots\infty} \right)(x)\, d\mu(x) \\ \leq & \frac{-1}{b} \int_{\mathbb{T}} \log \mu \left(\alpha_{1\cdots\ell} \mid \bigvee_{i=1}^{k-1} \alpha^{(i)}_{1\cdots b} \vee \alpha_{b+1\cdots\infty} \right)(x)\, d\mu(x) \\ \leq & \frac{-1}{b} \int_{\mathbb{T}} \log \mu(\alpha_{1\cdots\ell}) d\mu(x) \longrightarrow h(\mu) \frac{\log \bar{m}}{\log m},\end{aligned}$$

and so $\sum_{j=k}^{d} h_j \leq \frac{\log \bar{m}}{\log m} h(\mu)$. ∎

3. Equidistribution

Definition 3.1: Let $\{a_i\}$ be a sequence of integers, m_1 and m_2 relatively prime. Set

$$\begin{aligned} A(M,n,r) &= \{a_i\colon 1 \leq i \leq M \text{ and } a_i \equiv r \bmod m_1^n\}, \\ A(M,n,*;s) &= \{a_i\colon 1 \leq i \leq M \text{ and } a_i \equiv s \bmod m_2^n\}, \\ A(M,n,r;s) &= \{a_i\colon 1 \leq i \leq M,\ a_i \equiv r \bmod m_1^n \text{ and } a_i \equiv s \bmod m_2^n\}. \end{aligned}$$

The sequence a_i will be said to have **m_1-adic (almost) subexponential cells** if for every $\epsilon > 0$, for some increasing sequence M_n such that $M_{n+1}/M_n = O(1)$, for every n there is an $\hat{A}_n$ such that:

1. $|\hat{A}_n| < \epsilon |M_n|$,
2. $|A(M_n,n,r) \setminus \hat{A}_n| / |M_n| = O(m_1^{-(1-\epsilon)n})$ uniformly in r, n.

(We will omit the word 'almost' from now on.)

The sequence a_i will be said to have **m_1-adic subexponential cells rel m_2** if for every ϵ there are M_n and $\hat{A}_n$ as above such that

2′. $|A(M_n,n,r;s) \setminus \hat{A}_n| / |A(M_n,n,*;s)| = O(m_1^{-(1-\epsilon)n})$ uniformly in r, s, n.

Examples:

1. Suppose r is relatively prime to $m_1 m_2$. Then $a_i = r^i$ has m_1-adic subexponential cells, m_1-adic subexponential cells rel m_2, as well as $m_1 m_2$-adic

subexponential cells. In this case we can take $M_n = (m_1m_2)^n$. Note that, in general, each of these conditions implies the previous one.

2. Suppose r is relatively prime to m_1 but is divisible by every prime factor of m_2. Then $a_i = r^i$ has m_1-adic subexponential cells, and since practically all a_i, $1 \leq i \leq M_n$ are divisible by m_2^n this implies that a_i has m_1-adic subexponential cells rel m_2. However, a_i does not have m_1m_2-adic subexponential cells. This has been the motivating example for this work.
3. More generally, at least when m_1 is a power of some prime, whenever $a_i = \sum_{i=1}^N c_k r_k^i$ with $c_k \in \mathbb{Z}$, $|r_1| < |r_2| < \cdots$ with at least one r_k relatively prime to m_1 (and not equal to ± 1), then a_i has m_1-adic subexponential cells rel m_2. The restriction that m_1 is a power of some prime is not really a restriction from our point of view since we can still apply Theorem 3.2 on this sequence.
4. A discussion of recursion sequences, i.e. sequences such that $\sum_{i=0}^t c_i a_{n-i} = 0$ for every n ($c_1,\ldots, c_t \in \mathbb{Z}$), can be found in [11].

These examples all follow easily using the techniques of p-adic interpolation that are used to prove Theorem 3.2 of [13].

THEOREM 3.2: *Let $\{a_i\}$ be a sequence of m_1-adic subexponential cells relative to m_2. Then for any measure μ invariant and ergodic under $\times m$ (with $m = m_1m_2$) of positive entropy, for μ-a.e. x, $\{a_nx\}$ is equidistributed in $\mathbb{T}$ (with respect to Lebesgue measure).*

Proof: Let $v \in \mathbb{Z}\setminus\{0\}$, $0 < \epsilon < h(\mu)/10$, and let M_n and $\hat{A}_n$ be as in Definition 3.1, $M_{n-1} \leq M \leq M_n$. Set (for $r = 0,\ldots,m_1^n - 1$ or '$*$')

$$\tilde{A}(M,n,r;s) = A(M,n,r;s) \setminus \hat{A}_n.$$

Define

$$G_M(x) = \sum_{i=1}^{M} e_v(a_ix) \quad (\text{with } e_v(t) = e(t) = \exp(2\pi ivt)).$$

Since v will be fixed throughout, we suppress it in all notations. Set

$$G_{M,n,s}(x) = \sum_{a\in\tilde{A}(M,n,*;s)} e(ax);$$

thus

$$\left|G_M(x) - \sum_{s=0}^{m_2^n-1} G_{M,n,s}(x)\right| < C\epsilon M, \quad \text{where } C = \max(M_\ell/M_{\ell-1}).$$

STEP 1: For any $s,\, t \in \mathbb{Z}/m_2^n$

$$|G_{M,n,s}(x)| = |G_{M,n,s}(x + t/m_2^n)|.$$

Thus we can write $|G_{M,n,s}(x)| = |G_{M,n,s}|(x^{(1)}_{1\cdots n}; x_{n+1\cdots\infty})$.

This follows from the fact that for every $a,\ b \in \tilde{A}(M,n,*;s)$ we have that $m_2^n | a-b$, and so $e((a-b)(x+t/m_2^n)) = e((a-b)x)$. Thus

$$|G_{M,n,s}(x)|^2 = \sum_{a,b\in\tilde{A}(M,n,*;s)} e((a-b)x) = |G_{M,n,s}(x+t/m_2^n)|.$$

STEP 2:

$$\begin{aligned}\sum_{r=0}^{m_1^n-1} |G_{M,n,s}(x+r/m_1^n)|^2 &= \sum_{r=0}^{m_1^n-1} |G_{M,n,s}|(r; x_{n+1\cdots\infty})^2 \\ &= O_\epsilon(|A(M_n, n, *; s)|^2 m_1^{\epsilon n}).\end{aligned}$$

Indeed,

$$\begin{aligned}\sum_{r=0}^{m_1^n-1} |G_{M,n,s}(x+r/m_1^n)|^2 &= \sum_{r=0}^{m_1^n-1} \sum_{a,b\in\tilde{A}(M,n,*;s)} e((a-b)(x+r/m_1^n)) \\ &= m_1^n \sum_{r=0}^{m_1^n-1} |\tilde{A}(M_n,n,r;s)|^2 \\ &= O_\epsilon(|A(M_n,n,*;s)|^2 m_1^{\epsilon n}),\end{aligned}$$

where the last inequality is a consequence of $2'$ in Definition 3.1.

STEP 3: For **every** $x_{n+1\cdots\infty}$, there are at most $O_\epsilon(m_1^{\epsilon n})$ $x^{(1)}_{1\cdots n}$ so that

$$\sum_{s=0}^{m_2^n-1} |G_{M,n,s}|(x^{(1)}_{1\cdots n}; x_{n+1\cdots\infty}) > \epsilon M.$$

Indeed, denote the set of these $x^{(1)}_{1\cdots n}$ by $V_{M,n}(x_{n+1\cdots\infty})$. Then (in all summations below s runs from 0 to $m_2^n - 1$, $x^{(1)}_{1\cdots n}$ on all possible values)

$$\begin{aligned}&|V_{M,n}(x_{n+1\cdots\infty})| \\ &\quad\le (\epsilon M)^{-2} \sum_{x^{(1)}_{1\cdots n}} \left(\sum_s |G_{M,n,s}|(x^{(1)}_{1\cdots n}; x_{n+1\cdots\infty})\right)^2 \\ &\quad\le (\epsilon M)^{-2} \sum_{x^{(1)}_{1\cdots n}} \left(\sum_s |\tilde{A}(M_n,n,*;s)|\right)\left(\sum_s \frac{|G_{M,n,s}|(x^{(1)}_{1\cdots n}; x_{n+1\cdots\infty})^2}{|\tilde{A}(M_n,n,*;s)|}\right)\end{aligned}$$

$$\leq (\epsilon M)^{-2} M_n \sum_s \frac{1}{|\tilde{A}(M_n, n, *; s)|} \sum_{x^{(1)}_{1\cdots n}} |G_{M,n,s}|(x^{(1)}_{1\cdots n}; x_{n+1\cdots\infty})^2$$
$$= O_\epsilon\left(M^{-2} M_n \sum_s |\tilde{A}(M_n, n, *; s)| m_1^{\epsilon n}\right)$$
$$= O_\epsilon(m_1^{\epsilon n} M_n^2 / M^2) = O_\epsilon(m_1^{\epsilon n}).$$

Note that all estimates above are uniformly in n, x.

STEP 4: For a.e. x, if M is large enough,

$$\frac{1}{M} G_M(x) < 10(1+C)\epsilon. \tag{3.1}$$

Take $I_\epsilon = \{\lfloor (1+\epsilon)^\tau \rfloor\}_{\tau \in \mathbb{N}}$ and $I_{\epsilon,n} = I_\epsilon \cap [M_{n-1}, M_n]$. Set

$$V_n(x_{n+1\cdots\infty}) = \bigcup_{M \in I_{\epsilon,n}} V_{M,n}(x_{n+1\cdots\infty}).$$

Note first that in order to prove (3.1) it is clearly enough to show that for all large enough $M \in I_\epsilon$

$$\frac{1}{M} G_M(x) < 5(1+C)\epsilon.$$

Furthermore, $|V_n(x_{n+1\cdots\infty})| = O_\epsilon(m_1^{\epsilon n})$.

Let

$$B_n = \{y\colon \mu(\alpha^{(1)}_{1\cdots n} \mid \alpha_{n+1\cdots\infty})(y) < 2^{-(h-\epsilon)n}\}.$$

By Theorems 2.1 and 2.4, a.e. x is in B_n for all n large enough. But now clearly

$$\mu(x \in B_n \text{ and } x^{(1)}_{1\cdots n} \in V_n(x_{n+1\cdots\infty})) = O_\epsilon(m_1^{2\epsilon n} 2^{-hn}).$$

Since for small ϵ these probabilities are summable, by Borel–Cantelli, for all n large enough, either $x \notin B_n$ or $x^{(1)}_{1\cdots n} \notin V_n(x_{n+1\cdots\infty})$. The former is impossible by the theorems cited above; the latter implies that for all $M \in I_{\epsilon,n}$

$$\frac{1}{M} G_M(x) \leq C\epsilon + \frac{1}{M} \sum_s |G_{M,n,s}(x)| \leq (1+C)\epsilon.$$

From Step 4 the theorem follows by using Weyl's criterion for equidistribution. ∎

ACKNOWLEDGEMENT: The author would like to thank H. Furstenberg, D. Rudolph, B. Weiss, L-S. Young and especially D. Meiri and Y. Peres for many discussions on this and related questions.

References

[1] J. Feldman, *A generalization of a result of Lyons about measures in* [0, 1], Israel Journal of Mathematics **81** (1993), 281–287.

[2] J. Feldman and M. Smorodinsky, *Normal numbers from independent processes,* Ergodic Theory and Dynamical Systems **12** (1992), 707–712.

[3] H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation,* Mathematical Systems Theory **1** (1967), 1–49.

[4] B. Host, *Nombres normaux, entropie, translations,* Israel Journal of Mathematics **91** (1995), 419–428.

[5] B. Host, *Some results of uniform distribution in the multidimensional torus,* preprint.

[6] A. Johnson, *Measures on the circle invariant under multiplication by a nonlacunary subgroup of the integers,* Israel Journal of Mathematics **77** (1992), 211–240.

[7] A. Johnson and D. Rudolph, *Convergence under $\times_p$ of $\times_q$ invariant measures on the Circle,* Advances in Mathematics **115** (1995), 117–140.

[8] A. Katok and R. Spatzier, *Invariant measures for higher rank hyperbolic abelian actions,* Ergodic Theory and Dynamical Systems **16**, (1996), no. 4, 751–778 (correction in Ergodic Theory and Dynamical Systems **18** (1998), no. 2, 503–507).

[9] R. Kenyon and Y. Peres, *Measures of full dimension on affine-invariant sets,* Ergodic Theory and Dynamical Systems **16** (1996), 307–323.

[10] F. Ledrappier and L.-S. Young, *The metric entropy of diffeomorphisms — Part II: Relations between entropy, exponents and dimension,* Annals of Mathematics **122** (1985), 509–539.

[11] E. Lindenstrauss, D. Meiri and Y. Peres, *Entropy of convolutions on the circle,* Annals of Mathematics **149** (1999), 871–904.

[12] R. Lyons, *On measures simultaneously 2- and 3-invariant,* Israel Journal of Mathematics **61** (1988), 219–224.

[13] D. Meiri, *Entropy and uniform distribution of orbits in $\mathbb{T}^d$,* Israel Journal of Mathematics **105** (1998), 155–183.

[14] D. Meiri and Y. Peres, *Bi-invariant sets and measures have integer Hausdorff dimension,* Ergodic Theory and Dynamical Systems **19** (1999), 523–534.

[15] D. Rudolph, *$\times 2$ and $\times 3$ invariant measures and entropy,* Ergodic Theory and Dynamical systems **10** (1990), 395–406.