

INVARIANT MEASURES FOR CERTAIN EXPANSIVE \mathbb{Z}^2 -ACTIONS

BY

KLAUS SCHMIDT*

Mathematics Institute, University of Warwick

Coventry CV4 7AL, UK

e-mail: ks@maths.warwick.ac.uk

ABSTRACT

Let $p > 1$ be prime, and let $Y \subset X = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{Z}^2}$ be an infinite, closed, shift-invariant subgroup with the following properties: the restriction to Y of the shift-action σ of \mathbb{Z}^2 on X is mixing with respect to the Haar measure λ_Y of Y , and every closed, shift-invariant subgroup $Z \subsetneq Y$ is finite. We prove that every sufficiently mixing, non-atomic, shift-invariant probability measure μ on Y is equal to λ_Y .

Let $G \neq \{0\}$ be a finite, abelian group, and let σ be the shift-action $(\sigma_{\mathbf{m}}(x))_{\mathbf{n}} = x_{\mathbf{m}+\mathbf{n}}$ of \mathbb{Z}^2 on $X = G^{\mathbb{Z}^2}$. Then there exist uncountably many distinct, σ -invariant probability measures on X which are Bernoulli—and hence mixing of every order—under σ : indeed, let ν be a probability measure on G , put $\mu = \nu^{\mathbb{Z}^2}$, and note that different measures ν, ν' lead to inequivalent Bernoulli measures $\nu^{\mathbb{Z}^2}$ and $\nu'^{\mathbb{Z}^2}$. However, if $Y \subset X$ is a closed, shift-invariant subgroup such that the restriction σ^Y of σ to Y is mixing (with respect to the normalized Haar measure λ_Y of Y), then λ_Y may be the only non-atomic, σ^Y -invariant, mixing probability measure, although there always exist many different invariant, non-atomic, ergodic, *non-mixing* probability measures on Y . In order to study this phenomenon we assume that $G = \mathbb{Z}/p\mathbb{Z}$ is a cyclic group of prime order p , and consider subgroups $Y \subset X^{(p)} = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{Z}^2}$ with the following properties.

- (a) Y is infinite, closed, and shift-invariant;

* The author would like to thank the Department of Mathematics, University of Vienna, for hospitality while this work was done.

Received May 31, 1993

- (b) Every closed, shift-invariant subgroup $Z \subsetneq Y$ is finite;
- (c) σ^Y is mixing with respect to λ_Y .

1. *Definition:* Let $Y \subset X^{(p)} = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{Z}^2}$ be a subgroup satisfying (a)–(c), and let μ be a shift-invariant probability measure on Y . A finite, non-empty set $E \subset \mathbb{Z}^2$ is μ -mixing if

$$\lim_{k \rightarrow \infty} \mu \left(\bigcap_{\mathbf{n} \in E} \sigma_{-k\mathbf{n}}^Y(B(\mathbf{n})) \right) = \prod_{\mathbf{n} \in E} \mu(B(\mathbf{n}))$$

for every map $\mathbf{n} \mapsto B(\mathbf{n})$ from E into the σ -algebra \mathfrak{B}_Y of Borel subsets of Y , and μ -non-mixing otherwise. A μ -non-mixing set $E \subset \mathbb{Z}^2$ is **minimal μ -non-mixing** if every set E' with $\emptyset \neq E' \subsetneq E$ is μ -mixing.

Every λ_Y -non-mixing set is also μ -non-mixing (cf. Lemma 4). If the converse is true, then the following theorem shows that μ has to be equal to λ_Y ; in other words, λ_Y is the unique ‘most mixing’ measure for σ^Y .

2. **THEOREM:** *Let $p > 1$ be a rational prime, $Y \subset X^{(p)}$ a subgroup satisfying (a)–(c), and let μ be a non-atomic, shift-invariant probability measure on Y . If there exists a minimal λ_Y -non-mixing set $E \subset \mathbb{Z}^2$ which is also minimal μ -non-mixing, then $\mu = \lambda_Y$.*

For the proof of Theorem 2 we need an explicit description of the subgroups $Y \subset X^{(p)}$ satisfying (a)–(c). Let $\mathfrak{R}_2^{(p)} = \mathbb{F}_p[u_1^{\pm 1}, u_2^{\pm 1}]$ be the ring of Laurent polynomials in the variables u_1, u_2 with coefficients in the prime field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, and write a typical element $f \in \mathfrak{R}_2^{(p)}$ as $f = \sum_{\mathbf{n} \in \mathbb{Z}^2} c_f(\mathbf{n})u^{\mathbf{n}}$, where $c_f(\mathbf{n}) \in \mathbb{F}_p$ and $u^{\mathbf{n}} = u_1^{n_1}u_2^{n_2}$ for every $\mathbf{n} = (n_1, n_2) \in \mathbb{Z}^2$. An element $f \in \mathfrak{R}_2^{(p)}$ is a generalized polynomial in a single variable if its support $\mathcal{S}(f) = \{\mathbf{n} \in \mathbb{Z}^2: c_f(\mathbf{n}) \neq 0\}$ is contained in the line $\{\mathbf{k} + m\mathbf{l}: m \in \mathbb{Z}\}$ for some $\mathbf{k}, \mathbf{l} \in \mathbb{Z}^2, \mathbf{l} \neq \mathbf{0}$. The dual group $\widehat{X^{(p)}}$ of $X^{(p)}$ can be identified with $\mathfrak{R}_2^{(p)}$ by setting $\langle f, x \rangle = e^{\frac{2\pi i}{p} \sum_{\mathbf{n} \in \mathbb{Z}^2} c_f(\mathbf{n})x_{\mathbf{n}}}$ for every $f \in \mathfrak{R}_2^{(p)}$ and $x = (x_{\mathbf{n}}) = (x_{\mathbf{n}}, \mathbf{n} \in \mathbb{Z}^2) \in X^{(p)}$, where the element $\sum_{\mathbf{n} \in \mathbb{Z}^2} c_f(\mathbf{n})x_{\mathbf{n}} \in \mathbb{F}_p$ is identified with the corresponding integer in $\{0, \dots, p-1\}$. If $Y \subset X^{(p)}$ is a closed, shift-invariant subgroup, then $\mathfrak{a} = Y^\perp \subset \mathfrak{R}_2^{(p)} = \widehat{X^{(p)}}$ is an ideal; conversely, if $\mathfrak{a} \subset \mathfrak{R}_2^{(p)}$ is an ideal, then $Y = \widehat{\mathfrak{R}_2^{(p)}/\mathfrak{a}}$ is a closed, shift-invariant subgroup of $X^{(p)}$. The following lemma is an immediate consequence of the Propositions 2.12–2.13 in [Kit-Sch1] and Theorem 3.5 in [Sch].

3. **LEMMA:** *Let $\{0\} \neq Y \subset X^{(p)}$ be a closed, shift-invariant subgroup, and let $\mathfrak{a} = Y^\perp \subset \widehat{X^{(p)}} = \mathfrak{R}_2^{(p)}$ be the annihilator of Y . Then Y satisfies (a)–(c) if and*

only if the ideal $\mathfrak{a} \subset \mathfrak{R}_2^{(p)}$ is non-zero, prime, and principal, and not generated by a generalized polynomial in a single variable.

Lemma 3 yields an abundance of subgroups $Y \subset X^{(p)}$ satisfying (a)–(c). Let $Y \subset X^{(p)}$ be such a subgroup, and let μ be a shift-invariant probability measure on Y . We put $\mathfrak{a} = Y^\perp \subset \mathfrak{R}_2^{(2)}$, $\mathfrak{M} = \mathfrak{R}_2^{(p)}/\mathfrak{a}$, observe that \mathfrak{M} is a module over the ring $\mathfrak{R}_2^{(p)}$, and write $\hat{\mu}: \hat{Y} = \mathfrak{M} \rightarrow \mathbb{C}$ for the Fourier transform of μ , defined by $\hat{\mu}(a) = \int \langle a, x \rangle d\mu(x)$ for every $a \in \hat{Y}$. Then

$$\hat{\mu}(u^{\mathbf{n}} \cdot a) = \hat{\mu}(a)$$

for every $a \in \hat{Y}$ and $\mathbf{n} \in \mathbb{Z}^2$, and

$$\lim_{k \rightarrow \infty} \hat{\mu} \left(\sum_{\mathbf{n} \in E} u^{k\mathbf{n}} \cdot a(\mathbf{n}) \right) = \prod_{\mathbf{n} \in E} \hat{\mu}(a(\mathbf{n}))$$

for every μ -mixing set $E \subset \mathbb{Z}^2$ and every map $\mathbf{n} \mapsto a(\mathbf{n})$ from E into \hat{Y} .

4. LEMMA: Let $E \subset \mathbb{Z}^2$ be a finite, non-empty set which is λ_Y -non-mixing. If μ is non-atomic, then E is μ -non-mixing.

Proof: If μ is non-atomic, and if $0 \neq a \in \mathfrak{M} = \mathfrak{R}_2^{(p)}/\mathfrak{a}$, then we claim that the character $\chi_a = \langle a, \cdot \rangle$ of Y defined by a is not μ -a.e. equal to a constant. Otherwise $|\hat{\mu}(a)| = 1$, and the shift-invariance of μ implies that the set $\mathfrak{N} = \{a \in \mathfrak{M}: |\hat{\mu}(a)| = 1\} \subset \mathfrak{M}$ is a non-zero submodule, and that $Z = \mathfrak{N}^\perp \subset Y$ is a proper, closed, shift-invariant subgroup. By assumption, Z is finite, so that $\hat{Z} = \mathfrak{M}/\mathfrak{N}$ is finite. We set $\bar{\mu}(B) = \mu(-B)$ for every $B \in \mathfrak{B}_Y$ and conclude that the convolution $|\mu|^2 = \mu * \bar{\mu}$, whose Fourier transform is given by $|\widehat{\mu}|^2 = |\hat{\mu}|^2$, is a probability measure with finite support. Hence μ has finite support, which is absurd.

Corollary 2.7 in [Kit-Sch2] implies that there exist elements $a_{\mathbf{n}} \in \mathfrak{M}$, $\mathbf{n} \in E$, not all equal to 0, such that $\sum_{\mathbf{n} \in E} u^{k\mathbf{n}} a_{\mathbf{n}} = 0$ for all k in an infinite subset $K \subset \mathbb{N}$. The first part of this proof shows that at last one of the characters $\chi_{a_{\mathbf{n}}}$ is non-constant μ -a.e. As $1 = \chi_{\sum_{\mathbf{n} \in E} u^{k\mathbf{n}} a_{\mathbf{n}}} = \prod_{\mathbf{n} \in E} \chi_{a_{\mathbf{n}}} \cdot \sigma_{k\mathbf{n}}^Y$ for every $k \in K$, the set E must be μ -non-mixing. ■

Proof of Theorem 2: Let $E \subset \mathbb{Z}^2$ be a minimal λ_Y -non-mixing set which is also minimal μ -non-mixing, and choose as in the the proof of Lemma 4 a map $a: E \rightarrow \mathfrak{M} \setminus \{0\}$ and an infinite subset $K \subset \mathbb{N}$ such that $\sum_{\mathbf{n} \in E} u^{k\mathbf{n}} a(\mathbf{n}) = 0$

for every $k \in K$. Then $u^{k\mathbf{n}} \cdot a(\mathbf{n}) = -\sum_{\mathbf{m} \in E \setminus \{\mathbf{n}\}} u^{k\mathbf{m}} a(\mathbf{m})$ for every $k \in K$ and $\mathbf{n} \in E$. For every $\mathbf{n} \in E$, the set $E \setminus \{\mathbf{n}\}$ is μ -mixing, and

$$\begin{aligned} \hat{\mu}(a(\mathbf{n})) &= \lim_{\substack{k \rightarrow \infty \\ k \in K}} \hat{\mu}(u^{k\mathbf{n}} \cdot a(\mathbf{n})) = \lim_{\substack{k \rightarrow \infty \\ k \in K}} \hat{\mu} \left(- \sum_{\mathbf{m} \in E \setminus \{\mathbf{n}\}} u^{k\mathbf{m}} \cdot a(\mathbf{m}) \right) \\ &= \prod_{\mathbf{m} \in E \setminus \{\mathbf{n}\}} \overline{\hat{\mu}(a(\mathbf{m}))}. \end{aligned}$$

By varying \mathbf{n} we obtain either that $|\hat{\mu}(a(\mathbf{n}))| = 1$ for every $\mathbf{n} \in E$, or that $\hat{\mu}(a(\mathbf{n})) = 0$ for every $\mathbf{n} \in E$.

As in the proof of Lemma 3 we set $\mathfrak{N} = \{a \in \mathfrak{M}: |\hat{\mu}(a)| = 1\}$ and observe that $\mathfrak{M}/\mathfrak{N}$ is finite and μ atomic whenever $\mathfrak{N} \neq \{0\}$. In particular, $\hat{\mu}(a(\mathbf{n})) = 0$ for every $\mathbf{n} \in E$, and by replacing the map $a: \mathbf{n} \mapsto a(\mathbf{n})$ from E to \mathfrak{M} with $a': \mathbf{n} \mapsto a(\mathbf{n})' = h \cdot a(\mathbf{n})$ for an arbitrary, but temporarily fixed element $h \in \mathfrak{R}_2^{(p)}$ we see that $\hat{\mu}(h \cdot a(\mathbf{n})) = 0$ whenever $\mathbf{n} \in E$, $h \in \mathfrak{R}_2^{(p)}$, and $h \cdot a(\mathbf{n}) \neq 0$. Fix $\mathbf{n} \in E$ for the moment, and put $\mathfrak{N}' = \{h \cdot a(\mathbf{n}): h \in \mathfrak{R}_2^{(p)}\}$ and $Z' = \mathfrak{N}'^\perp \subset Y$. Since $\hat{\mu}(a) = 0$ for every non-zero $a \in \mathfrak{N}$, the measure μ' induced by μ on Y/Z' is equal to the Haar measure on Y/Z' . Since σ^Y is topologically conjugate to a skew-product action of \mathbb{Z}^2 on $Y/Z' \times Z'$, the ergodicity of μ under σ^Y implies that $\mu = \lambda_Y$. ■

If we know a little more about the generator of the annihilator $Y^\perp = \mathfrak{a}$ of the subgroup Y , then we can weaken the assumptions of Theorem 2.

5. THEOREM: Let $0 \neq f \in \mathfrak{R}_2^{(p)}$ be an irreducible Laurent polynomial, $\mathfrak{a} = f\mathfrak{R}_2^{(p)}$, and assume that the shift-action σ^Y of \mathbb{Z}^2 on $Y = \mathfrak{a}^\perp \subset X^{(p)}$ is mixing, and that the support $\mathcal{S}(f)$ of f is a minimal λ_Y -non-mixing set. If μ is a non-atomic, σ^Y -invariant probability measure on Y such that $\mathcal{S}(f) \setminus \{\mathbf{n}\}$ is μ -mixing for some $\mathbf{n} \in \mathcal{S}(f)$, then $\mu = \lambda_Y$.

Proof: Since $\sum_{\mathbf{n} \in \mathcal{S}(f)} c_f(\mathbf{n})u^{p^k\mathbf{n}} \in \mathfrak{a}$ for every $k \geq 1$ we obtain that

$$\begin{aligned} \hat{\mu}(a) &= \hat{\mu}(u^{p^k\mathbf{n}} \cdot a) = \lim_{k \rightarrow \infty} \hat{\mu} \left(- \sum_{\mathbf{m} \in \mathcal{S}(f) \setminus \mathbf{n}} \frac{c_f(\mathbf{m})}{c_f(\mathbf{n})} u^{p^k\mathbf{m}} \cdot a \right) \\ &= \prod_{\mathbf{m} \in \mathcal{S}(f) \setminus \{\mathbf{n}\}} \overline{\hat{\mu} \left(\frac{c_f(\mathbf{m})}{c_f(\mathbf{n})} a \right)} \end{aligned}$$

for every $a \in \mathfrak{M} = \mathfrak{R}_2^{(p)}/\mathfrak{a}$. As \mathbb{F}_p is finite and μ is non-atomic, we conclude as in the proof of Theorem 2 that $\hat{\mu}(a) = 0$ whenever $0 \neq a \in \mathfrak{M}$, and that $\mu = \lambda_Y$.

■

6. *Examples:* In the following examples we consider irreducible Laurent polynomials $f \in \mathfrak{R}_2^{(2)}$ and set $\mathfrak{a} = f\mathfrak{R}_2^{(2)}$ and $Y = \mathfrak{a}^\perp \subset X^{(2)}$.

(1) (cf. [Kit-Sch2]) Let $f = 1 + u_1 + u_2 + u_1^2 + u_1u_2 + u_2^2$. Then $E = \{(0, 0), (1, 0), (0, 1)\}$ is a minimal λ_Y -non-mixing set, and Theorem 2 implies that every σ^Y -invariant probability measure μ on Y for which $\sigma_{\mathbf{n}}^Y$ is μ -mixing for every $\mathbf{n} \in \{(1, 0), (0, 1), (1, -1)\}$, is equal to λ_Y .

(2) (cf. [Led], [Kit-Sch2]) Let $f = 1 + u_1 + u_2$. Then $S(f) = \{(0, 0), (1, 0), (0, 1)\}$ is minimal λ_Y -non-mixing, and by letting \mathbf{n} vary in E we see from Theorem 5 that every probability measure μ on Y for which any of the transformations $\sigma_{(1,0)}^Y$, $\sigma_{(0,1)}^Y$, or $\sigma_{(-1,1)}^Y$ is mixing, must be equal to λ_Y . This example is of interest because of certain formal similarities between σ^Y and the \mathbb{N}^2 -action on \mathbb{T} generated by multiplication by 2 and 3 (cf. [Fur], [Rud]).

(3) Let $f = 1 + u_1 + u_1^{-1} + u_2$. Then $S(f) = \{(-1, 0), (0, 0), (1, 0), (0, 1)\}$ is minimal λ_Y -non-mixing, and by setting $\mathbf{n} = (0, 1)$ we see from Theorem 5 that every non-atomic, σ^Y -invariant probability measure μ on Y for which $\sigma_{(1,0)}^Y$ is three-mixing, is equal to λ_Y .

7. *Remark:* If $Y \subset X^{(p)}$ is a subgroup satisfying (a)–(c) then there exist shift-invariant, ergodic probability measures μ on Y which are different from λ_Y . A method for constructing such measures on the group Y in Example 6 (2) is described in [Kit-Sch1]. In general we put $\mathfrak{a} = Y^\perp$ and consider, for every $k \geq 1$, the subring $\mathcal{R}^{(k)} = \mathbb{F}_p[u_1^{\pm p^k}, u_2^{\pm p^k}] \subset \mathfrak{R}_2^{(p)}$. We write $\mathfrak{M}^{(k)}$ instead of $\mathfrak{M} = \mathfrak{R}_2^{(d)}/\mathfrak{a}$ in order to emphasize that $\mathfrak{M} = \mathfrak{M}^{(k)}$ is to be viewed as an $\mathcal{R}^{(k)}$ -module. Then $\mathfrak{M}^{(k)}$ has non-trivial $\mathcal{R}^{(k)}$ -submodules of infinite index. If $\mathfrak{N}^{(k)} \subset \mathfrak{M}^{(k)}$ is such a submodule, then $Y^{\mathfrak{N}^{(k)}} = (\mathfrak{N}^{(k)})^\perp \subset Y$ is an infinite, closed subgroup of Y which is invariant under the shifts $\{\sigma_{(mp^k, np^k)}: (m, n) \in \mathbb{Z}^2\}$. Although the Haar measure $\lambda^{\mathfrak{N}^{(k)}} = \lambda_{Y^{\mathfrak{N}^{(k)}}$ is not shift-invariant, its orbit average $\mu^{\mathfrak{N}^{(k)}} = \frac{1}{p^{2k}} \sum_{m=0}^{p^k-1} \sum_{n=0}^{p^k-1} \lambda^{\mathfrak{N}^{(k)}} \cdot \sigma_{(m,n)}$ is a non-atomic, shift-invariant, ergodic probability measure on Y which is obviously non-mixing under every $\sigma_{\mathbf{n}}^Y$. By choosing sequences $\mathfrak{N}^{(1)} \subset \mathfrak{N}^{(2)} \subset \dots$ of submodules $\mathfrak{N}^{(k)} \subset \mathfrak{M}^{(k)}$ we obtain sequences $(\mu^{\mathfrak{N}^{(k)}}, k \geq 1)$ of such measures, and every limit point μ of such a sequence is again shift-invariant. Note that these shift-invariant measures μ may

be chosen to have the property that the entropy $h_\mu(\sigma_{\mathbf{n}})$ is either positive or equal to 0 for some given non-zero $\mathbf{n} \in \mathbb{Z}^2$.

8. *Examples:* (1) Let $Y \subset X^{(p)}$ be a subgroup satisfying (a)–(c). For fixed $k \geq 0, l \geq 1$, let $f_{k,l} = 1 + u_1^{p^k} + u_1^{2p^k} + \dots + u_1^{p^k(p^l-1)}$, and put $\mathfrak{N} = f_{k,l}\mathcal{R}^{(k+l)}/\mathfrak{a} \subset \mathfrak{M}^{(k+l)}$. Then $Y^{\mathfrak{N}} = \{y = (y_{\mathbf{n}}) \in Y: y_{(mp^{k+l}, np^{k+l})} + y_{(mp^{k+l}+p^k, np^{k+l})} + \dots + y_{(mp^{k+l}+p^k(p^l-1), np^{k+l})} = 0 \pmod{p} \text{ for every } (m, n) \in \mathbb{Z}^2\}$, and the orbit average $\mu^{\mathfrak{N}}$ of $\lambda_{Y^{\mathfrak{N}}}$ is non-atomic, shift-invariant, ergodic, and not equal to λ_Y .

(2) Let $f = 1 + u_1 + u_2^2 \in \mathfrak{A}_2^{(2)}$, $\mathfrak{a} = f\mathfrak{A}_2^{(2)}$, and $Y = \mathfrak{a}^\perp \subset X^{(2)}$. Then $\mathfrak{M} = \mathfrak{A}_2^{(2)}/\mathfrak{a}$ is a module over the ring $\mathcal{R} = \mathbb{F}_2[u_1^{\pm 1}, u_2^{\pm 2}]$, and $\mathcal{L} = (1 + u_2)\mathcal{R}/\mathfrak{a} \subset \mathfrak{M}$ is an \mathcal{R} -submodule with infinite index. Put $Y^{\mathcal{L}} = \mathcal{L}^\perp = \{y = (y_{\mathbf{n}}, \mathbf{n} \in \mathbb{Z}^2): y_{(2m, n)} = y_{(2m+1, n)} \text{ for all } (m, n) \in \mathbb{Z}^2\}$, and note that $\mu = \frac{1}{2}(\lambda_{Y^{\mathcal{L}}} + \lambda_{Y^{\mathcal{L}}} \cdot \sigma_{(0,1)}^Y)$ is non-atomic, shift-invariant and ergodic, but non-ergodic under $\sigma_{(1,0)}$.

9. *Problem:* If $Y \subset X^{(p)}$ is a subgroup satisfying (a)–(c), can there exist a non-atomic, shift-invariant probability measure $\mu \neq \lambda_Y$ on Y which is mixing under some $\sigma_{\mathbf{n}}^Y, \mathbf{n} \in \mathbb{Z}^{2?}$

References

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

[Kit-Sch1] B. Kitchens and K. Schmidt, *Markov subgroups of $(\mathbb{Z}/2)^{\mathbb{Z}^2}$* , *Contemporary Mathematics* **135** (1992), 265–283.

[Kit-Sch2] B. Kitchens and K. Schmidt, *Mixing sets and relative entropies for higher dimensional Markov shifts*, *Ergodic Theory and Dynamical Systems*, to appear.

[Led] F. Ledrappier, *Un champ markovien peut être d’entropie nulle et mélangéant*, *Comptes Rendus de l’Academie des Sciences, Paris* **287** (1978), 561–562.

[Rud] D.J. Rudolph, *$\times 2$ and $\times 3$ invariant measures and entropy*, *Ergodic Theory and Dynamical Systems* **8** (1990), 395–406.

[Sch] K. Schmidt, *Automorphisms of compact abelian groups and affine varieties*, *Proceedings of the London Mathematical Society* **61** (1990), 480–496.