ALMOST BLOCK INDEPENDENCE FOR THE THREE DOT $\mathbb{Z}^2$ DYNAMICAL SYSTEM

Thomas B. Ward

Israel J. of Math., to appear.

Abstract. We show that the measure preserving action of $\mathbb{Z}^2$ dual to the action defined by the commuting automorphisms $\times x$ and $\times y$ on the discrete group $\mathbb{Z}[x^{\pm 1}, y^{\pm 1}]/\langle 1+x+y\rangle\mathbb{Z}[x^{\pm 1}, y^{\pm 1}]$ is measurably isomorphic to a $\mathbb{Z}^2$ Bernoulli shift. This was conjectured in recent work by Lind, Schmidt and the author, where it was shown that this action has completely positive entropy. An example is given of $\mathbb{Z}^2$ actions which are measurably isomorphic without being topologically conjugate.

§1. Introduction

Let

$$X = \{\mathbf{x} \in \mathbb{T}^{\mathbb{Z}^2} \mid x_{(n,m)} + x_{(n+1,m)} + x_{(n,m+1)} = 1 \text{ for all } n, m \in \mathbb{Z}\}. \tag{1.1}$$

Then X is a compact abelian group carrying a natural $\mathbb{Z}^2$ action $\alpha : \mathbb{Z}^2 \to Aut(X)$ given by the restriction of the shift action on $\mathbb{T}^{\mathbb{Z}^2}$ to the closed, shift–invariant subgroup X:

$$(\alpha_{(k,l)}\mathbf{x})_{(n,m)} = x_{(n+k,m+l)}. \tag{1.2}$$

This action is an example of a $\mathbb{Z}^d$ action on a compact abelian group, and these have been systematically studied in [7], [11], [18], and [19]. The action is mixing (by Theorem 11.2(4) of [7]), has a dense set of periodic points (Theorem 7.2 of [7]), and has trivial Pinsker algebra (Theorem 6.13 of [11]). It follows that the action is mixing of all orders (Corollary 6.7 of [11]). By Theorem 6.14 of [11], Haar measure is maximal for α, and the topological entropy of α is given in [11], Example 5.1:

$$h(\alpha) = \int_0^1 \int_0^1 \log|1 + e^{2\pi is} + e^{2\pi it}|\mathrm{d}s\mathrm{d}t = \frac{3\sqrt{3}}{4\pi} L(2, \chi_3) \tag{1.3}$$

Supported in part by N. S. F. grant No. DMS–88–02593 at the University of Maryland

where $L(., \chi_3)$ is the Dirichlet L–series with the character $\chi_3(3k) = 0$, $\chi_3(3k \pm 1) = \pm 1$.

The group $\mathbb{T}^{\mathbb{Z}^2}$ is dual to the ring of Laurent polynomials $\mathbb{Z}[x^{\pm 1}, y^{\pm 1}]$, and the closed subgroup X is the annihilator of the ideal $\langle 1+x+y \rangle$. The ideal $\langle 2, 1+x+y \rangle$ gives the $\mathbb{Z}^2$ action considered by Ledrappier, [8]: this example has zero entropy and is mixing, but is not mixing of all orders.

Our purpose here is to show that the $\mathbb{Z}^2$ action defined by $\langle 1+x+y \rangle$ is measurably isomorphic to a Bernoulli $\mathbb{Z}^2$ action; this is a special case of the more general conjecture (Conjecture 6.8 in [11]), that a $\mathbb{Z}^d$ action on a compact abelian group with the Descending Chain Condition (see [7]) is isomorphic to a Bernoulli action if it has completely positive entropy. This, in turn, is a special case of the following question, due to Thouvenot: is the K property equivalent to being measurably isomorphic to a Bernoulli shift for higher dimensional Markov shifts?

For $d = 1$, an ergodic automorphism of a compact abelian group was shown to be K in 1964 by Rokhlin [16]. An ergodic automorphism of $\mathbb{T}^n$, and the natural automorphic extension to a solenoid of an ergodic endomorphism of $\mathbb{T}^n$ were shown to be Bernoullian by Katznelson in 1971, [3]; this was extended to infinite dimensional tori by Lind, [9], and independently by Chu, [1]. The final result that an ergodic automorphism of a compact group is measurably isomorphic to a Bernoulli shift was shown by Lind in 1977, [10], and independently by Miles and Thomas, [12].

One dimensional mixing Markov shifts are isomorphic to Bernoulli shifts [13], and a partial result in the direction of Thouvenot's question has been shown by Rosenthal [17]: an ergodic Markov $\mathbb{Z}^2$ system with finite alphabet has the *weak Pinsker property*, which means it can be written as a direct product of a Bernoulli system and a system of arbitrarily small entropy.

The results of [7] and [10] applied to the action α show that for any $(a, b) \neq (0, 0)$, the automorphism $\alpha_{(a,b)}$ of X is measurably isomorphic to an infinite entropy Bernoulli shift. This property is clearly necessary but is not sufficient to guarantee that the action is measurably isomorphic to a Bernoulli $\mathbb{Z}^2$ action: the ideal generated by $\{x + 2, y + 2\}$ defines a $\mathbb{Z}^2$ action β with the property that $\beta_{(a,b)}$ is isomorphic to an infinite entropy Bernoulli shift for any $(a, b) \neq (0, 0)$, but the action β has zero entropy by [11]. In this connection, see also Example 5.7 of [11].

For any subset $F \subset \mathbb{Z}^2$, let $\pi^{(F)} : X \to \mathbb{T}^F$ denote projection onto the coordinates in F; the image of $\pi^{(F)}$ is a closed subgroup of $\mathbb{T}^F$. Let μ_F denote the normalised Haar measure on $\pi^{(F)}$.

Let $\mathbf{e}(t) = e^{2\pi i t}$. A Laurent polynomial $f_{\mathbf{a}}(u) = \sum a_k u^k \in \mathbb{Z}[u, u^{-1}]$ defines a character $\chi_{\mathbf{a}}$ of $\mathbb{T}^{\mathbb{Z}}$ by

$$\chi_{\mathbf{a}}(\mathbf{t}) = \prod \mathbf{e}(a_k t_k) = \mathbf{e}(\sum a_k t_k).$$

The mapping $f_{\mathbf{a}} \mapsto \chi_{\mathbf{a}}$ identifies the dual of $\mathbb{T}^{\mathbb{Z}}$ with $\mathbb{Z}[u, u^{-1}]$. It will be convenient to also note the further identification of $\mathbb{Z}[u, u^{-1}]$ with $\sum_{\mathbb{Z}} \mathbb{Z}$, effected by sending

$$f_{\mathbf{a}}(u) = \sum_{k=-n}^{m} a_k u^k$$

to the finitely supported infinite integer vector $\mathbf{a} = (\dots, 0, a_{-n}, \dots, a_m, 0, \dots)$.

Now let $S = \{1, \dots, r\}$ be a finite set. Given two probability measures μ and ν on $S^{\mathbb{Z}^2}$, and D a finite subset of $\mathbb{Z}^2$, define the space of joinings $J_D(\mu, \nu)$ as follows. Write μ^D, ν^D for the marginal measures induced by μ and ν on S^D. If λ is a probability measure on $S^D \times S^D$ then write λ^1, λ^2 for the two marginals of λ, and set

$$J_D(\mu, \nu) = \{\lambda \mid \lambda^1 = \mu^D, \lambda^2 = \nu^D\}.$$

The $\bar{d}$ distance between the probability measures μ and ν is defined as in [21]. Let d be the trivial metric on S given by $d(i,j) = 0$ if $i = j$, $d(i,j) = 1$ if $i \neq j$, and let $x = \{x_n\}$, $y = \{y_n\}$ be the processes with alphabet S defined by μ and ν respectively. Then define

$$\bar{d}_D(\mu, \nu) = \inf_{\lambda \in J_D(\mu,\nu)} \frac{1}{|D|} \sum_{d \in D} \int d(x_d, y_d) \mathrm{d}\lambda,$$

and

$$\bar{d}(\mu, \nu) = \limsup_{D} \bar{d}_D(\mu, \nu). \tag{1.4}$$

We amend the definition of ABI in [21] as follows. Let σ be the shift on $S^{\mathbb{Z}^2}$. Say that a stationary $\mathbb{Z}^2$ process x has *almost block independence* if for any $\epsilon > 0$ there is an N_ϵ such that if $n \geq N_\epsilon$, $R = [0, n-1) \times [0, n-1) \cap \mathbb{Z}^2$, and y is another process with

(1) $\bar{d}_R(\sigma_{n(a,b)}(y), x) = 0$ for all $(a,b) \in \mathbb{Z}^2$ and
(2) y restricted to $n(a,b) + R$ is independent of y restricted to $n(a',b') + R$ if $(a,b) \neq (a',b')$,

then $\bar{d}(x,y) \leq \epsilon$.

Notice that it is sufficient to produce a process y with properties (1), (2) and having $\bar{d}(x,y) < \epsilon$ since the properties together determine y in the sense that if y, $\bar{y}$ are processes sharing (1) and (2), then $\bar{d}(y, \bar{y}) = 0$.

For a $\mathbb{Z}$ process with finite state space Theorem 2 in [21] shows that almost block independence implies finitely determined and hence a process with almost block independence is a stationary coding of a Bernoulli process. We explain in Appendix B how the ideas of [21] may be applied to $\mathbb{Z}^2$ to show the process is finitely determined. The equivalence of finitely determined to Bernoulli is shown in [2], §1.

§2. Almost block independence

For each pair $n, m \in \mathbb{Z}$ with $n \leq m$, define

$$S(n,m) = \{(a,b) \in \mathbb{Z}^2 \mid n \leq a \leq m\}.$$

For a trigonometric polynomial $f : \mathbb{T} \to \mathbb{C}$, let $H(f)$ denote the highest frequency in f, so that if $f(t) = \sum c_k \mathbf{e}(kt)$ then $H(f) = \max\{|k| \mid c_k \neq 0\}$. For a character χ on $\mathbb{T}^{\mathbb{Z}}$, let $H(\chi)$ denote the largest frequency appearing in χ, so $H(\chi_{\mathbf{a}}) = \max\{|a_k|\}$. For a polynomial $a_n u^{-n} + \cdots + a_m u^m$ in $\mathbb{Z}[u, u^{-1}]$ call $n + m$ the degree of the polynomial.

Lemma 2.1. *Let $m(N,D) = \min\{r \mid (r-2)! \geq N^2(D+1)\}$. If*

$$f(\mathbf{x}) = \prod_{S(-p,0)} f_{ij}(x_{ij}) \times \prod_{S(m,m+q)} f_{ij}(x_{ij})$$

where each f_{ij} is a trigonometric polynomial with $H(f_{ij}) \leq N$, $f_{ij} \equiv 1$ if $|j| > D$, and $m > m(N,D)$, then $\int f d\mu_\infty = \int f d\nu_\infty$, where μ_∞ is Haar measure on $\pi^{(S(-p,0)\cup S(m,m+q))}(X)$, and ν_∞ is the independent concatenation of Haar measure on $\pi^{S(-p,0)}(X)$ and Haar measure on $\pi^{S(m,m+q)}(X)$.

Proof. Notice that the values of a point $\mathbf{x} \in X$ on the coordinates in the vertical line $L(n) = \{(a,b) \in \mathbb{Z}^2 \mid a = n\}$ determine (by (1.1)) the coordinates in $S(n,m)$ for any $m \geq n$. Thus we can identify a function h on $\pi^{S(n,m)}(X)$ with a function on $\pi^{L(n)}(X)$.

If χ is any character on $\pi^{S(-n,0)}(X) \times \pi^{S(m,m+q)}(X)$ then, by the usual theory of characters,

$$\int \chi \mathrm{d}\nu_\infty = \begin{cases} 1, & \text{if } \chi \equiv 1; \\ 0, & \text{if not.} \end{cases}$$

Similarly,

$$\int \chi \mathrm{d}\mu_\infty = \begin{cases} 1, & \text{if } \chi \equiv 1 \text{ on } \pi^{(S(-p,0)\cup S(m,m+q))}(X); \\ 0, & \text{if not.} \end{cases}$$

As above, identify the dual group of $\pi^{L(a)}(X)$ with $\mathbb{Z}[u,u^{-1}]$, and let

$$E_N = \{f \in \mathbb{Z}[u,u^{-1}] \mid H(f) \leq N\}.$$

The way in which a character $\chi_{\mathbf{a}}$ on $\pi^{L(a)}(X)$ determines (or induces) a character on the adjacent lines $\pi^{L(a-1)}(X)$ and $\pi^{L(a+1)}(X)$ is as follows.

On $\pi^{L(a-1)}(X)$, $\chi_{\mathbf{a}}$ induces the character $\chi_{S\mathbf{a}}$ where S is the injective homomorphism of $\mathbb{Z}[u,u^{-1}]$ dual to the surjective homomorphism that transforms $\mathbf{x} = (x_{a-1,k})_{k\in\mathbb{Z}}$ to $\mathbf{y} = (x_{a,k})_{k\in\mathbb{Z}}$ according to the rule (1.1). Thus S may be expressed explicitly in two different ways. If the character $\chi_{\mathbf{a}}$ is identified with the finitely supported infinite integer vector $\mathbf{a}$, then S has the upper triangular matrix form

$$S = \left|\begin{array}{cccccc} \ddots & \ddots & & & & \\ & -1 & -1 & & & \\ & & -1 & -1 & & \\ & & & -1 & -1 & \\ & & & & \ddots & \ddots \\ & & & & & \ddots \end{array}\right| \tag{2.1}$$

(here the diagonal is $(\dots,-1,-1,-1,\dots)$). If the character $\chi_{\mathbf{a}}$ is identified with the polynomial $f_{\mathbf{a}}(u)$ then the action of S is multiplication by the polynomial $-(1+u^{-1})$. By duality, the kernel of the map $\mathbf{x} \mapsto \mathbf{y}$ (a circle) is dual to the cokernel of S.

On $\pi^{L(a+1)}(X)$, $\chi_{\mathbf{a}}$ induces the character $\chi_{T\mathbf{a}}$ say. Here T is not a homomorphism since S is not invertible (equivalently: the values $\mathbf{x} = (x_{a+1,k})_{k\in\mathbb{Z}}$ do not determine the values $\mathbf{z} = (x_{a,k})_{k\in\mathbb{Z}}$). To describe the map T, let $\mathbf{a} = (\dots,0,a_{-n},\dots,a_m,0,\dots)$ have $a_{-n} \neq 0$, and notice that

$$\chi_{T\mathbf{a}}(\mathbf{x}) = \int_{\{\widehat{S}\mathbf{y}=\mathbf{x}\}} \chi_{\mathbf{a}}(\mathbf{y})d\mathbf{y}$$

where the integration is with respect to Haar measure on $\pi^{L(a+1)}(X) = \mathbb{T}^{\mathbb{Z}}$, which is the infinite product of Lebesgue measure on $\mathbb{T}$. Label the coordinates $(x_{a+1,k})_{k\in\mathbb{Z}}$ by $(x_n)_{n\in\mathbb{Z}}$, where $x_{a+1,k} = x_{-k}$. Applying the rule (1.1), this simplifies to give

$$\begin{aligned}
\chi_{T\mathbf{a}}(\mathbf{x}) &= \int \mathbf{e}(a_{-n}y_{-n})\mathbf{e}(a_{-n+1}(x_{-n+1}-y_{-n}))\mathbf{e}(a_{-n+2}(x_{-n+2}-y_{-n+1}))\dots \\
&\qquad\qquad \mathbf{e}(a_m(x_m - y_{m-1}))dy_{-n}\dots dy_{m-1} \\
&= \int \mathbf{e}(a_{-n}y_{-n})\mathbf{e}(a_{-n+1}(-x_{-n+1}-y_{-n})) \\
&\qquad\qquad \mathbf{e}(a_{-n+2}(-x_{-n+2}+x_{-n+1}+y_{-n}))\dots \\
&\qquad\qquad \mathbf{e}(a_m(-x_m + x_{m-1} - x_{m-2} + \dots - (-1)^{m+n}y_{-n}))dy_{-n}
\end{aligned}$$

which vanishes unless $\sum(-1)^k a_k = 0$. We conclude that

$$\chi_{T\mathbf{a}}(\mathbf{x}) = \begin{cases} \chi_{\mathbf{b}}(\mathbf{x}), & \text{if } \sum(-1)^k a_k = 0; \\ 0, & \text{if } \sum(-1)^k a_k \neq 0 \end{cases} \tag{2.2}$$

where $\mathbf{b}$ (written in polynomial form) is given by

$$f_{\mathbf{b}}(u) = u^{-n+1}(-a_{-n+1}+a_{-n+2}-\cdots\pm a_m)+u^{-n+2}(-a_{-n+2}+\cdots\mp a_m)+\cdots+u^m(-a_m). \tag{2.3}$$

The support of the Fourier transform of the function induced on $L(m)$ by the function $\prod_{S(m,m+q)} f_{ij}$ lies in the set

$$E_N + SE_N + S^2E_N + \dots S^qE_N; \tag{2.4}$$

the support of the Fourier transform of the function induced on $L(0)$ by the function $\prod_{S(m,m+q)} f_{ij}$ lies in

$$S^m(E_N + SE_N + S^2E_N + \dots S^qE_N), \tag{2.5}$$

while that of the function induced by $\prod_{S(-p,0)} f_{ij}$ on $L(0)$ lies in

$$T^pE_N + T^{p-1}E_N + \dots TE_N + E_N. \tag{2.6}$$

The integrals will therefore agree if, for $m > m(N,D)$,

$$(E_N + SE_N + S^2E_N + \dots S^qE_N) \cap T^m(T^pE_N + T^{p-1}E_N + \dots TE_N + E_N) = \{0\}. \tag{2.7}$$

We claim that if $m > m(N,D) = \min\{r \mid (r-2)! \geq N^2(D+1)\}$, then

$$T^m(T^pE_N + T^{p-1}E_N + \dots TE_N + E_N) \cap (\mathbb{Z}[u,u^{-1}]) = \{0\}. \tag{2.8}$$

Consider the expression (2.6). By iterating the map T, we see from (2.2) that $T^k(f_{\mathbf{a}}) = 0$ unless the coefficients of $f_{\mathbf{a}}$, written as $\mathbf{a}' = (a_{-n},\dots,a_m)^t$ with $a_{-n} \neq 0$, satisfy the equations $M\mathbf{a}' = 0$ where $M = M(n+m-1,k)$ is the $(n+m+1)\times k$ matrix

$$M = \begin{bmatrix} 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & \dots & (-1)^{m+n+1} \\ 0 & 1 & -2 & 3 & -4 & 5 & -6 & 7 & -8 & \dots & \\ 0 & 0 & 1 & -3 & 5 & -7 & 9 & -11 & 13 & \dots & \\ \vdots & & & & & & & & & & \vdots \\ 0 & 0 & \dots & \dots & 1 & -(1+k-1) & \dots & & & & \end{bmatrix}$$

whose j^{th} row comprises $(j-1)$ zeros followed by the alternating arithmetic progression

$$1, -(1+(j-1)), (1+2(j-1)), -(1+3(j-1)), \dots, (-1)^{m+n-j}(1+(m+n+2-j)(j-1)).$$

Notice that since $a_{-n} \neq 0$, $a_l \neq 0$ for some $l > k$; if not the maximal rank of the matrix shows that the only solution has $\mathbf{a}' = 0$. By the appendix, at least one element of $\mathbf{a}$ has modulus no less than $(k-2)!$; since $f_{\mathbf{a}} \in E_N$, this can only happen if $(k-2)! \leq N$. Assuming that $N \geq 4$, this certainly requires $k \leq N$ (it requires much more, but for our purposes any estimate will do here). We deduce that the only terms that contribute to the left hand side of (2.8) are those of the form $T^k E_N$ with $k \leq N$.

In order to estimate the size of the coefficients of a polynomial in (2.6), let $f(u)$ be $a_{-n}u^{-n} + \cdots + a_m u^m$. Then the condition (2.2) becomes $Tf = 0$ unless $f = (1+u^{-1})g$ for some $g \in \mathbb{Z}[u, u^{-1}]$. The coefficients of Tf are given by (2.3); writing $[u^k]$ for the coefficient of u^k in Tf gives the following estimates. Firstly, $[u^{-n+1}] = -a_{-n}$ by the condition (2.2) so $|[u^{-n+1}]| \leq N$. Similarly, $[u^{-n+2}] = -a_{-n+1} - [u^{-n+1}]$ so $|[u^{-n+2}]| \leq 2N$. Continuing in this way gives $|[u^{-n+j}]| \leq jN$ for $j = 1, \ldots, n+m$. Starting at the other end of the polynomial gives $|[u^{m-j}]| \leq jN$ for $j = 1, \ldots, n+m$. This gives the estimate

$$|[u^j]| \leq N \times \min\{|j+n|, |j-m|\}$$

for each of the coefficients, and hence (recall that $n + m \leq 2D + 1$)

$$H(Tf) \leq N \times (D+1). \tag{2.9}$$

Now the map T sends a polynomial like f to a polynomial with the same positive (or leading) degree (m), and a lower negative degree ($n-1$) (or larger positive trailing degree if $n < 0$); in either case the quantity $m+n$ is diminished by one while m is preserved. Thus, if g is a polynomial that appears in (2.6), with $g(u) = a_l u^l + \cdots + a_k u^k$ say, then a_k is made up of contributions from T acting on a degree 2 polynomial, T^2 acting on a degree 3 polynomial, and so on, up to degree N only. The coefficient a_{k-1} is made up of contributions from T acting on a degree 3 polynomial, T^2 acting on a degree 4 polynomial, and so on, up to degree N only. By (2.9),

$$H(g) \leq N + N(D+1) + \cdots + N(D+1) \leq N^2(D+1). \tag{2.10}$$

Apply Appendix A again: $T^m g$ is a trivial character unless $N^2(D+1) \leq (m-2)!$. Choose $m(N,D) = \min\{r \mid (r-2)! \geq N^2(D+1)\}$; then for $m > m(N,D)$, (2.8) holds. □

Following [3], Definition 3, say that a partition $P = \{P_1, \ldots, P_r\}$ of $\mathbb{T}$ is nice if, for each $N \in \mathbb{N}$, there is a set $E(N)$, with $\mu(E(N)) < N^{-2}$, and there exist trigonometric polynomials $\{f_1, \ldots, f_r\}$ on $\mathbb{T}$ of the form

$$f_k(t) = \sum_{j=-N^{10}}^{N^{10}} c_j^{(k)} \mathbf{e}(jt) \tag{2.11}$$

with

$$f_k(t) \geq 1 \text{ on } P_k \backslash E(N), \quad \sup \sum_k f_k < 1 + N^{-2}, \tag{2.12}$$

and

$$f_k(t) \leq N^{-2} \text{ on } X \backslash (P_k \cup E(N)). \tag{2.13}$$

Such partitions exist: consider a partition P each of whose elements is an interval in $\mathbb{T}$. Choose, for each P, the Fejér sum of order N^{10} of the corresponding interval to get the required trigonometric polynomials approximating the atoms of the partition.

A partition P of $\mathbb{T}$ determines a partition of X, also denoted P, the "time zero" partition whose atoms are sets of the form $\{\mathbf{x} \in X \mid x_{(0,0)} \in P_i\}$.

We need the following Lemma, due to Katznelson (Lemma 1 of [3]). Let $(Y, \mathcal{B}, \nu)$ be a probability space. Two finite partitions $P = \{P_1, \dots, P_s\}$, $Q = \{Q_1, \dots, Q_t\}$, of Y are ϵ–independent if

$$\sum_{i=1}^{s} \sum_{j=1}^{t} |\nu(P_i \cap Q_j) - \nu(P_i)\nu(Q_j)| < \epsilon.$$

Lemma 2.2. *Let $P = \{P_1, \dots, P_s\}$ and $Q = \{Q_1, \dots, Q_t\}$ be finite partitions of Y. Assume that there is a set $E \subset Y$, $\nu(E) < \epsilon^2$, and for each $i = 1, \dots, s$ and $j = 1, \dots, t$ there are nonnegative measurable functions f_i and g_j on Y such that*

$$f_i \geq 1 \text{ on } P_i \backslash E, \;\; g_j \geq 1 \text{ on } Q_j \backslash E,$$

$$\sum_{i=1}^{s} \int_Y f_i d\nu < 1 + \epsilon^2, \;\; \sum_{j=1}^{t} \int_Y g_j d\nu < 1 + \epsilon^2$$

and

$$\int_Y f_i g_j d\nu = \int_Y f_i d\nu \int_Y g_j d\nu.$$

Then P and Q are 11ϵ–independent.

Lemma 2.3. *For any partition P of X arising as the time zero partition of a nice partition of $\mathbb{T}$, the finite state process (X, P) is almost block independent.*

Proof. Let $R \subset \mathbb{Z}^2$ be a finite subset. A function of the form $f(\mathbf{x}) = \prod_{(i,j) \in R} f_{ij}(x_{ij})$ where each f_{ij} has the form (2.11) will be called an (N, R)–function. Say that a partition Q of X is approximated to within δ by an (N, R) function if there is a collection of (N, R)–functions $\mathfrak{F}$ with the property that there is a set E, $\mu(E) < \delta^2$, and for each atom $Q_i \in Q$ there is an $f_i \in \mathfrak{F}$ with $f_i \geq 1$ on $Q_i \backslash E$, $f_i < \delta$ on $X \backslash (Q_i \cup E)$ and $\sum_i \int f_i d\mu < 1 + \delta^2$. Let

$$P^{(R)} = \bigvee_{(n,m) \in R} \alpha_{(n,m)}(P).$$

Claim that if $\mathfrak{F}$ is an $(N, \{0,0\})$ collection of functions that approximates P to within δ, then $\mathfrak{F}^{(R)}$, the set of all functions of the form $f(\mathbf{x}) = \prod_{(i,j) \in R} f_{ij}(x_{ij})$ where each $f_{ij} \in \mathfrak{F}$, is an (N, R) collection of functions that approximates $P^{(R)}$ to within $\delta^{(R)}$ where $\delta^{(R)} = \max\{\delta\sqrt{\#R}, \sqrt{(1+\delta^2)^{(\#R)} - 1}\}$. Firstly, the error set $E^{(R)}$ is at most the union of the fibres above the error sets on each coordinate, hence has measure no more than $\#R\mu(E)$. Now claim by induction on $\#R$ that $\sum_{\text{atoms of } P^R} f < (1+\delta^2)^{(\#R)}$: assume this for $\#R = k$, and then $\sum_n \sum_i \int f_n f_i = \sum_n \int \sum_i f_n f_i = \sum \int f_n \sum_i f_i$. Writing $\sum_i f_i = 1 + e$, where e is a function on $\mathbb{T}$ with $\int |e| < \delta^2$, this is bounded by $\sum_n \int f_n(1+e) \leq (1+\delta^2)^k + \delta^2(1+\delta^2)^k$.

We will say that two collections of functions $\mathfrak{F}$ and $\mathfrak{G}$ are independent if, for any $f \in \mathfrak{F}$, $g \in \mathfrak{G}$, $\int fg d\mu = \int f d\mu \int g d\mu$.

For any $\delta > 0$, define

$$m_\delta(S) = \min\{r \mid (r-2)! \geq (S-1) \times (N_0(S,\delta))^{20}\}$$

where

$$N_0(S,\delta) = \min\{N \mid \max\{N^{-1}\sqrt{S}, \sqrt{(1+N^{-2})^{(S^2)} - 1} < \delta\}\}.$$

By Lemma 2.1 and the above remarks, if $\mathfrak{T} = \{T_i\}$ is a collection of $S \times S$ tiles (squares of coordinates with side S in $\mathbb{Z}^2$) that are placed with gaps of size at least $m_\delta(S)$ between adjacent tiles, then for each i there is a $(N_0(S,\delta), T_i)$ collection of functions $\mathfrak{F}_i$ that approximates $P^{(T_i)}$ to within δ and has $\mathfrak{F}_i$ independent of $\mathfrak{F}_j$ for $i \neq j$.

Now let $\epsilon > 0$ be given, and put $\epsilon_k = \frac{6}{k^2\pi^2}\epsilon$ so that $\sum \epsilon_k = \epsilon$. Let

$$n_0(\epsilon) = \min\{n \mid 4m_{\epsilon_1}(k(n + m_{\epsilon_1}(n)) < \epsilon_1 \times n \text{ for some } k \in \mathbb{N}\},$$

and notice that this always exists because, for fixed δ, $m_\delta(n)$ grows very slowly in n.

Given $n > n_0(\epsilon)$, let

$$S_1 = \max\{s \mid 4m_{\epsilon_1}(s) < n\epsilon_1 \text{ and } (n + m_{\epsilon_1}(s)) | s\},$$

where we amend $m_{\epsilon_1}(s)$ by adding some number no larger than n to allow divisibility if needed. Now, by Lemma 2.1, we may tile the square $\mathfrak{S}_1$ of side S_1 symmetrically with tiles of size $n \times n$ spaced a distance $m_{\epsilon_1}(S_1)$ apart and on each tile T_i, $i = 1, \dots, p(1) = S_1^2/(n+m_{\epsilon_1})^2$ there is an $(N_0(S_1, \epsilon), T_i)$ collection of functions $\mathfrak{F}$ which has $\mathfrak{F}$ restricted to T_i independent of $\mathfrak{F}$ restricted to T_j for $i \neq j$, and the functions of the form $\prod_{l=1}^{p(1)} f_l((x_{ij})_{ij \in T_l})$, $f_l \in \mathfrak{F}$, approximate $P^{(T_1 \cup \dots T_{p(1)})}$ to within ϵ. Moreover, if λ_1 is the proportion of $\mathfrak{S}_1$ that is not covered by some $n \times n$ tile, then

$$\lambda_1 \leq \big(\frac{S_1}{n}\big)^2 \times 4nm_{\epsilon_1}\frac{S_1}{S_1^2} \leq 4m_{\epsilon_1}\frac{S_1}{n} < \epsilon_1.$$

Now define an inductive procedure for extending the tiling. For $k \geq 1$, let

$$S_{k+1} = \max\{s \mid 4m_{\epsilon_{(k+1)}}(s) < S_k \epsilon_{(k+1)}\}.$$

By construction, if $\mathfrak{S}_k$ is a square of side S_k, then we may tile all but ϵ_k (in proportion) of $\mathfrak{S}_k$ with $p(k)$ square tiles of side S_{k-1} and on each smaller tile T_i, $i = 1, \dots, p(k)$ we have a $(N_0(S_k, \epsilon), T_i)$ collection of functions $\mathfrak{F}$ such that $\mathfrak{F}^{(p(k)}$ (i.e. products of the form $\prod_{l=1}^{p(k)} f_l((x_{ij})_{ij \in T_l})$, $f_l \in \mathfrak{F}$), approximate $P^{(T_1 \cup \dots T_{p(k)})}$ to within ϵ. Moreover, $\mathfrak{F}$ restricted to T_i is independent of $\mathfrak{F}$ restricted to T_j if $i \neq j$.

Each smaller tile is an $\mathfrak{S}_{k-1}$ square, which may be covered to within ϵ_{k-1} with tiles of side S_{k-2}, and there is a corresponding collection of functions approximating the join of P over all these yet smaller tiles to within ϵ. The proportion of $\mathfrak{S}_k$ that is not covered by these smaller tiles of side $S_{(k-2)}$ is no greater than $\epsilon_k + \epsilon_{k-1}$.

Continuing, we obtain a tiling, by $n \times n$ squares, of all but ϵ (in proportion) of $\mathbb{Z}^2$; on each tile T_i there is a collection of functions $\mathfrak{F}_i$ with $\mathfrak{F}_i$ independent of $\mathfrak{F}_j$ for $i \neq j$, and $\mathfrak{F}_{i(1)} \times \dots \times \mathfrak{F}_{i(l)}$ approximating $P^{(T_{i(1)} \cup \dots T_{i(l)})}$ to within ϵ.

We claim that the process (X, P) is n–block 11ϵ–independent. To see this, let $\{i(1), \dots, i(p)\}$ be a finite collection of $n \times n$ tiles, and let $i \notin \{i(1), \dots, i(p)\}$. We show that $P^{(T_i)}$ is 11ϵ–independent of $P^{(T_{i(1)} \cup \dots \cup T_{i(p)})}$. Choose k large enough to ensure that there is a square $\mathfrak{S}_k$ of side S_k with $T_i \cup (T_{i(1)} \cup \dots \cup T_{i(p)}) \subset \mathfrak{S}_k$. Then, by construction, there are families of functions $\mathfrak{F}$, $\mathfrak{F}'$ (depending only on the tiled coordinates in $\mathfrak{S}_k$ and made up as a product of pairwise-independent functions), with the following properties:

(1) $\mathfrak{F}$ approximates $P^{(T_{i(1)} \cup \dots \cup T_{i(p)})}$ to within ϵ,
(2) $\mathfrak{F}'$ approximates $P^{(T_i)}$ to within ϵ, and
(3) $\mathfrak{F}$ and $\mathfrak{F}'$ are independent.

By Lemma 2.2, we deduce that the partitions $P^{(T_{i(1)} \cup \dots \cup T_{i(p)})}$ and $P^{(T_i)}$ are 11ϵ–independent.

Thus, (X, P) restricted to the tiles is an n–block 11ϵ–independent process; call this process Z. Now $\bar{d}((X, P), Z) \leq \epsilon$ because Z can be exactly copied into (X, P)

all but ϵ (in proportion) of the time. Let Y denote the process obtained by independently concatenating (X, P) restricted to each tile; Y is then n–block independent.

Now Lemma 6.3 of [20] extends easily to $\mathbb{Z}^2$, showing that a δ–independent process is within 4δ ($\bar{d}$) of an independent one (the n–blocks do not affect this because we can simply think of the processes as having a larger finite state space). We deduce that $\bar{d}(Z, Y) \leq 44\epsilon$ and hence $\bar{d}((X, P), Y) \leq 45\epsilon$, showing almost block independence. □

Theorem 2.4. *The system (X, α) is measurably isomorphic to a Bernoulli shift.*

Proof. For each partition P of X arising from a nice partition of $\mathbb{T}$ at time zero, the finite state process (X, P) is almost block independent by Lemma 2.3, and hence is finitely determined by Appendix B. It follows from [2], Theorem 1.1 that $(X, P^{(\mathbb{Z}^2)})$ is measurably isomorphic to a Bernoulli shift.

Let P_k be the partition of $\mathbb{T}$ into k intervals of the form $[\frac{j}{k!}, \frac{j+1}{k!})$. If $\mathbf{x} \neq \mathbf{y}$ are distinct points in X then they differ in some position, so for some k they lie in different atoms of $P_k^{(\mathbb{Z}^2)}$. Thus the algebra generated by $P_k^{(\mathbb{Z}^2)}$ increases to the whole σ–algebra $\mathcal{B}$ modulo null sets. By [15], §III, Theorem 5 (the Monotone Theorem for amenable group actions), we conclude that $(X, \mathcal{B}, \mu)$ is measurably isomorphic to a Bernoulli shift. □

§3. An example

In this section we describe four $\mathbb{Z}^2$ actions which are all measurably isomorphic but not pair–wise topologically conjugate. The $\mathbb{Z}^2$ actions by measure preserving transformations (X_1, α^1) and (X_2, α^2) (each equipped with an invariant probability measure) are *measurably isomorphic* if there are sets of full measure $Y_1 \subset X_1$, $Y_2 \subset X_2$ and an isomorphism of measure spaces $\psi : Y_1 \to Y_2$ with the property that $\psi\alpha^1_{(n,m)} = \alpha^2_{(n,m)}\psi$ for all $(n, m) \in \mathbb{Z}^2$. If X_1, X_2 are compact topological spaces and the actions are by homeomorphisms, then the systems are *topologically conjugate* if there is a homeomorphism $\psi : X_1 \to X_2$ with $\psi\alpha^1_{(n,m)} = \alpha^2_{(n,m)}\psi$ for all $(n, m) \in \mathbb{Z}^2$. If X_1 and X_2 are compact groups then the systems are *algebraically conjugate* if a topological conjugacy ψ may be chosen to be an isomorphism of the compact groups.

Let $R = \mathbb{Z}[x^{\pm1}, y^{\pm1}]$. If L is an R–module, then the module structure defines a $\mathbb{Z}^2$ action α^L on the compact dual group $X_L = \widehat{L}$ (see [7] or [11] for the details). Consider the following ideals of R: $\mathfrak{p} =< 1 + x + y >$, $\mathfrak{q} =< 1 + x^{-1} + y >$, $\mathfrak{s} =< 1 + x^{-1} + y^{-1} >$, and $\mathfrak{t} =< 1 + x + y^{-1} >$. For each ideal $\mathfrak{f}$ there is a $\mathbb{Z}^2$ system $(X_{R/\mathfrak{f}}, \alpha^{R/\mathfrak{f}})$; call the four systems corresponding to the above ideals P, Q, S, and T respectively. Notice that P is the system considered above defined by (1.1) and (1.2). The compact groups in the systems Q, S and T are given by

$$X_{R/\mathfrak{q}} = \{\mathbf{x} \in \mathbb{T}^{\mathbb{Z}^2} \mid x_{(n,m)} + x_{(n-1,m)} + x_{(n,m+1)} = 1 \text{ for all } n, m \in \mathbb{Z}\}, \quad (3.1)$$

$$X_{R/\mathfrak{s}} = \{\mathbf{x} \in \mathbb{T}^{\mathbb{Z}^2} \mid x_{(n,m)} + x_{(n-1,m)} + x_{(n,m-1)} = 1 \text{ for all } n, m \in \mathbb{Z}\}, \tag{3.2}$$

and

$$X_{R/\mathfrak{t}} = \{\mathbf{x} \in \mathbb{T}^{\mathbb{Z}^2} \mid x_{(n,m)} + x_{(n+1,m)} + x_{(n,m-1)} = 1 \text{ for all } n, m \in \mathbb{Z}\}. \tag{3.3}$$

Since no orientation was used in the arguments above, §2 shows that P, Q, S and T are each measurably isomorphic to a Bernoulli $\mathbb{Z}^2$ action. The entropies all coincide with that of $\alpha^{R/\mathfrak{p}}$ by [11], so we conclude that the systems are all measurably isomorphic by [5], §5. A more detailed description of Ornstein's isomorphism theorem for Bernoulli $\mathbb{Z}^d$ actions is given in [6], Theorem 2.

We claim that they are not pair–wise topologically conjugate. This may be seen from the relationship between the algebraic structure of the modules and the dynamical properties of the systems described in [7] and [19].

If the systems are topologically conjugate, then [19], Theorem 4.2, shows that they must be algebraically conjugate, and Corollary 4.3 then shows that $R/\mathfrak{p}$, $R/\mathfrak{q}$, $R/\mathfrak{s}$, and $R/\mathfrak{t}$ must be isomorphic as R–modules, which is not the case: the set of associated primes of each module is $\{\mathfrak{p}\}$, $\{\mathfrak{q}\}$, $\{\mathfrak{s}\}$, and $\{\mathfrak{t}\}$ respectively. This is the higher dimensional analogue of noting that two toral automorphisms cannot be topologically conjugate unless their corresponding integer matrices have the same characteristic equation.

A cruder subdivision can be made by considering periodic points: this shows that neither one of P and S is topologically conjugate to Q or T.

Let Fix_Γ^L denote the subgroup of points in X_L that are invariant under the action of the subgroup $\Gamma \subset \mathbb{Z}^2$. If $\Gamma = (a,b)\mathbb{Z} + (c,d)\mathbb{Z}$ then the dual of Fix_Γ^L is given by

$$\widehat{Fix_\Gamma^L} \cong \frac{L}{< 1 - x^a y^b, 1 - x^c y^d > L}.$$

For the lattice $\Gamma = (1,1)\mathbb{Z} + (-2,2)\mathbb{Z}$, we compute directly that $\widehat{Fix_\Gamma^{R/\mathfrak{p}}} \cong \mathbb{Z}/3\mathbb{Z}$ and $\widehat{Fix_\Gamma^{R/\mathfrak{q}}} \cong \mathbb{Z}/15\mathbb{Z}$, so $|Fix_\Gamma^{R/\mathfrak{p}}| \neq |Fix_\Gamma^{R/\mathfrak{q}}|$ and the systems P and Q are therefore not topologically conjugate.

Notice that periodic points will not distinguish P from S or Q from T. Because any lattice Γ is invariant under the action of $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ it is clear that

$$Fix_\Gamma^{R/\mathfrak{p}} \cong Fix_\Gamma^{R/\mathfrak{s}} \text{ and } Fix_\Gamma^{R/\mathfrak{q}} \cong Fix_\Gamma^{R/\mathfrak{t}}$$

for any Γ. Thus the four systems P, Q, R, and S between them provide examples of measurable isomorphism without topological conjugacy, isomorphic periodic point groups of all periods without topological conjugacy, and equal dynamical zeta functions without topological conjugacy.

Appendix A

In this appendix we prove the assertion used in the proof of Lemma 2.1 concerning the size of integer solutions of the equation $M(n,k)\mathbf{a} = 0$. Recall that $M(n,k)$ is the matrix (where we label the variables so that $\mathbf{a}$ is the integer vector $(a_1,\dots,a_n)$ with $a_1 \neq 0$):

$$M(n,k) = \begin{bmatrix} 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & \dots \\ 0 & 1 & -2 & 3 & -4 & 5 & -6 & 7 & -8 & 9 & \dots \\ 0 & 0 & 1 & -3 & 5 & -7 & 9 & -11 & 13 & -15 & \dots \\ \vdots & & & & & & & & & & \\ 0 & 0 & \dots & \dots & 1 & -(1+k-1) & \dots & & & & \end{bmatrix}$$

whose j^{th} row comprises $(j-1)$ zeros followed by the alternating arithmetic progression

$$1, -(1+(j-1)), (1+2(j-1)), -(1+3(j-1)), (1+4(j-1)), \dots,$$
$$(1+(n+2-j)(j-1)(-1)^{n-j}).$$

The k equations may be written in the form $\mathbf{r}_j \cdot \mathbf{a}$ for $j = 1,\dots,k$ where $\mathbf{r}_j$ is the j^{th} row of the matrix. The matrix can be row–reduced in such a way that the top row becomes $(1,0,\dots,0,c_1,\dots,c_{n-k})$ in which there are $(k-1)$ zeros. The row reduction is determined by this property because the matrix has maximal rank. Let $n! = 1$ for $n \leq 1$.

Lemma A.1. *If $\lambda_1,\dots,\lambda_k$ have $\mathbf{r}_1+\lambda_1\mathbf{r}_2+\dots+\lambda_{k-1}\mathbf{r}_k = (1,0,\dots,0,c_1,\dots,c_{n-k})$ as above then $\lambda_s = (s-2)!$.*

Proof. For $s \leq 4$ this can be seen from the matrix. Let m_{ij} be the $(i,j)^{\text{th}}$ entry in M; this is given by

$$m_{ij} = (-1)^{i+j}(1+(j-i)(i-1))$$

if $i \leq j$ and is 0 otherwise. The value of λ_{s+1} is obtained from the values of $\lambda_1,\dots\lambda_s$ by performing the row reduction to simplify the first s entries in the first row and then seeing what appears as the $(s+1)^{\text{th}}$ entry in the first row:

$$\lambda_{s+1} = -\sum_{p=1}^{s} \lambda_p m_{p,s+1} = -\sum_{p=1}^{s} \lambda_p (1+(p-1)(s-p+1))(-1)^{s+p+1}.$$

Assume that $\lambda_t = (t-2)!$ for $t \leq s$. Then

$$\sum_{p=1}^{s} \lambda_p (1+(p-1)(s-p+1))(-1)^{s+p+1} - \sum_{p=1}^{s-1} \lambda_p (1+(p-1)(s-p+1))(-1)^{s+p}$$
$$= -(1+(s-1))(s-2)! = -(s-2)!(s)$$

so $-\lambda_{s+1} = -s(s-2)! + \lambda_s = (s-1)(s-2)! = (s-1)!$, which shows the lemma.□

This forces the coefficients c_j to be large when they are non–zero:

Lemma A.2. *Each c_j is divisible by $(k-2)!$.*

Proof. Consider c_1:

$$|c_1| = |\sum_{p=1}^{k} \lambda_p (1+(p-1)(k-p+1))(-1)^{k+p+1}| = |\lambda_k| = (k-2)!$$

Now $\lambda_{k+1} = \sum_{p=1}^{k+1} \lambda_p m_{p,k+2}$ so

$$c_2 = \sum_{p=1}^{k} \lambda_p m_{p,k+2} = \lambda_{k+1} - \lambda_k m_{p,k+2} \in (k-2)!\mathbb{Z},$$

with similar formulæ for c_3, c_4 and so on. □

Now return to the equation $M(n,k)\mathbf{a} = 0$; since $a_1 \neq 0$ by assumption, the row–reduced equation is

$$a_1 + \sum_{j=1}^{n-k} c_j a_{k+j} = 0$$

and Lemma 2 then shows that $|a_1| \geq (k-2)!$ as required.

Appendix B

In this appendix we show how the method of [21] applies in our situation to show that the three dot dynamical system is finitely determined. The equivalence of finitely determined with Bernoullicity for $\mathbb{Z}^d$ actions is shown in §1 of [2], where five characterizations of Bernoullicity for $\mathbb{Z}^d$ actions are shown to be equivalent. Recall from §1 the definition of the $\bar{d}_R$ metric for a subset $R \subset \mathbb{Z}^2$; for a finite state $\mathbb{Z}^2$ process X let σ denote the shift action of $\mathbb{Z}^2$. Let $R(n) = [0, n-1] \times [0, n-1] \cap \mathbb{Z}^2$. Given points $\mathbf{n}, \mathbf{m} \in \mathbb{Z}^2$ let $Rect(\mathbf{n}, \mathbf{m})$ denote the rectangle

$$Rect(\mathbf{n}, \mathbf{m}) = \{(a,b) \in \mathbb{Z}^2 \mid (\mathbf{n})_1 \leq a \leq (\mathbf{m})_1, (\mathbf{n})_2 \leq b \leq (\mathbf{m})_2\}.$$

In this section processes are stationary finite state $\mathbb{Z}^2$ processes.

Definition B1. A stationary process X has almost block independence if for any $\epsilon > 0$ there exists N_ϵ such that if $n > N_\epsilon$ and Y is another process with

(1) $\bar{d}_R(n)(\sigma_{n(a,b)}(Y), X) = 0$ for all $(a,b) \in \mathbb{Z}^2$, and $(B1)$

(2) Y restricted to $n(a,b)+R(n)$ is independent of Y restricted to $n(a',b')+R(n)$ if $(a,b) \neq (a',b')$, $(B2)$

then

$$\bar{d}(x,y) \leq \epsilon. \qquad (B3)$$

Definition B2. A stationary process X is finitely determined if given $\epsilon > 0$ there is a $\delta > 0$ and N such that if Y is a stationary ergodic process with the same size state space as X and

$$\bar{d}_{R(N)}(X,Y) < \delta \qquad (B4)$$

and

$$|h(X) - h(Y)| < \delta \qquad (B5)$$

then $\bar{d}(X,Y) < \epsilon$.

Notice that §1 of [2] shows that the finitely determined processes are exactly those arising as codings of $\mathbb{Z}^2$ i.i.d. processes.

Theorem B3. *If X has almost block independence then X is finitely determined.*

Consider an array of binary digits in $\{0,1\}^{\mathbb{Z}^2}$. An $R(n)$–*cell* is a square block of side $(n-1)$ consisting of 1's, surrounded by 0's. A binary array $r \in \{0,1\}^{R(m)}$ is a δ–n–*array* if there is a disjoint collection of $R(n)$–cells covering at least $(1-\delta)$ of $R(m)$ (that is, containing at least $(1-\delta)m^2$ coordinates).

Lemma B4. *Assume X is almost block independent and $\epsilon > 0$. There is an $N \in \mathbb{N}$ and a $\delta > 0$ such that if $n \geq N$ we can find M so that $m \geq M$ implies that if $r \in \{0,1\}^{R(m)}$ is a δ–n array and $\bar{Y}$ is a process with*

(1) $\bar{d}_{R(n)}(\sigma_{(a,b)}(\bar{Y}), X) = 0$ *if $(a,b) + R(n+1)$ is an $R(n)$–cell and* $(B6)$

(2) $\bar{Y}$ *restricted to $(a,b) + R(n)$ is independent of $\bar{Y}$ restricted to $(a,b) + n(a',b') + R(n), (a',b') \neq (0,0)$* $(B7)$

then

$$\bar{d}_{R(m)}(X, \bar{Y}) \leq \epsilon. \qquad (B8)$$

Proof. Choose $N_1 = N_{\epsilon/3}$ from Definition B1. Define a process Y so that Y satisfies (B1), (B2) with $n = N_1$. By (B3) we may choose $N_2 > N_1$ such that

$$\bar{d}_{R(n)}(X,Y) \leq \epsilon/3 \text{ if } n \geq N_2. \qquad (B9)$$

Fix $n \geq N_2 + 2N_1$. Let $r \in\in \{0,1\}^{R(m)}$ be a $\delta - n$–array, and let Y be a process satisfying (B6) and (B7).

Let $\mathbf{k}_1, \mathbf{k}_2, \dots \in \mathbb{Z}^2$ be defined as follows.

$\mathbf{k}_1 + R(n+1)$ is an $R(n)$–cell in r,

$\mathbf{k}_2 + R(n+1)$ is an $R(n)$–cell in r restricted to $R(m)\backslash\{\mathbf{k}_1 + R(n+1)\}$,

$\mathbf{k}_3 + R(n+1)$ is an $R(n)$–cell in r restricted to $R(m)\backslash(\{\mathbf{k}_1 + R(n+1)\} \cup \{\mathbf{k}_2 + R(n+1)\}$,

and so on.

Consider the $R(n)$–cell $\mathbf{k}_j + R(n+1)$; define points $\mathbf{p}_j$, $\mathbf{m}_j \in \mathbb{Z}^2$ as follows. Let a be the least multiple of N_1 exceeding $(\mathbf{k}_j)_1$, b the least multiple of N_1 exceeding $(\mathbf{k}_j)_2$, and put $\mathbf{p}_j = (a, b)$. To define $\mathbf{m}_j$, let c be the least multiple of N_1 exceeding $(\mathbf{k}_j)_1 + n$, d the least multiple of N_1 exceeding $(\mathbf{k}_j)_2 + n$, and set $\mathbf{m}_j = (c, d)$.

Notice that $Rect(\mathbf{p}_j, \mathbf{m}_j)$ sits inside an $R(n)$–cell. Also, $\mathbf{p}_j$, $\mathbf{m}_j$ are separated by at least N_2 in each coordinate by choice of n. Thus the distribution of Y restricted to $Rect(\mathbf{p}_j, \mathbf{m}_j)$ is identical to the distribution of X restricted to $Rect(\mathbf{p}_j, \mathbf{m}_j)$ (Y satisfies (B6)).

Hence

$$\bar{d}_{Rect(\mathbf{p}_j,\mathbf{m}_j)}(\sigma_{\mathbf{p}_j}\bar{Y}, \sigma_{\mathbf{p}_j}Y) \leq \epsilon/3 \qquad (B10)$$

by choice of N_1.

We now join the Y process to the $\bar{Y}$ process. On coordinates $\mathbf{i} \in Rect(\mathbf{p}_j, \mathbf{m}_j)$ for some j use (B10) to join Y to $\bar{Y}$ $\bar{d}$ closely. On the remaining coordinates, join arbitrarily. We have a $\bar{d} \leq \epsilon/3$ joining on all but a proportion δ of the coordinates, so for sufficiently small δ, $\bar{d}_{R(m)}(\bar{Y}, Y) < 2\epsilon/3$ say. Hence $\bar{d}_{R(m)}(X, \bar{Y}) \leq \bar{d}_{R(m)}(X, Y) + \bar{d}_{R(m)}(Y, \bar{Y}) \leq \epsilon$. □

Proof of Theorem B3. Let X have state space $S = \{s_1, \dots, s_k\}$. Let N, δ be the numbers corresponding to a given $\epsilon > 0$ according to Lemma B4; put $n = N$. Define a $S^{R(N)}$–valued i.i.d. process Z with measure on $S^{R(N)}$ given by the $R(n)$ block measure $\mu_{R(n)}$ of the S–valued process X. Let R be a $\{0,1\}$–valued $\mathbb{Z}^2$ coding of an i.i.d. process with the property that $r \in R$, when restricted to $R(m)$, is a δ–N–array with probability at least $(1 - \delta)$. Call this property (B11). Such a process may be obtained from a Rohlin $R(N+1)$ tower built in a $\mathbb{Z}^2$ Bernoulli shift B (see [14] or [5] for the Rohlin lemma for $\mathbb{Z}^d$ actions); on this tower label the edge with a 0 and the rest with a 1. Let W be the $S^{R(N)} \times \{0,1\}$–valued $\mathbb{Z}^2$ process (Z, R) where the two processes are independent. Since W is a coding of the i.i.d. process (Z, B), it is finitely determined. Fix a point $x \in S$ and define $\bar{Y}$ as follows. If $r \in R$ is a δ–N–array, with N–cells $\{\mathbf{q}_i + R(N+1)\}$, then on each cell let $\bar{Y}_{\mathbf{q}_i + R(N)} = Z_{\mathbf{q}_i}$ (recall that Z is $S^{R(N)}$–valued). On the remainder of R, let $\bar{Y} = x$. Then $\bar{Y}$ is a coding of an i.i.d. process and is therefore finitely determined. Moreover, (B8) and (B3) together show that $\bar{d}(X, \bar{Y}) < \epsilon$.

Thus X is a $\bar{d}$ limit of finitely determined processes; hence X is finitely determined (Theorem 4, §III of [15]). □

Corollary B5. *A stationary finite state $\mathbb{Z}^2$ process that is almost block independent is measurably isomorphic to a Bernoulli shift.*

Proof. A special case of §III.1 of [15] shows that entropy classifies finitely determined $\mathbb{Z}^d$ processes up to measurable isomorphism. Since Bernoulli processes are clearly finitely determined, this shows that a finitely determined process is measurably isomorphic to a Bernoulli process. □

ACKNOWLEDGEMENT

The author would like to thank Profs. Dan Rudolph and Janet Kammeyer for instruction on the $\bar{d}$ metric and almost block independence, and for frequent helpful discussions; also the referee for substantial correction and clarification.

References

[1] H. Chu, *Some results on affine transformations of compact groups*, Inventiones mathematicæ **28** (1975), 161–183.

[2] J. Kammeyer, *A complete classification of the two–point extensions of a multidimensional Bernoulli shift*, Journal d'Analyse Mathématique **54** (1990), 113–163.

[3] Y. Katznelson, *Ergodic automorphisms of* $\mathbb{T}^n$ *are Bernoulli shifts*, Israel Journal of Mathematics **10** (1971), 186–95.

[4] Y. Katznelson, *An introduction to Harmonic Analysis*, Dover, New York, 1976.

[5] Y. Katznelson and B. Weiss, *Commuting measure–preserving transformations*, Israel Journal of Mathematics **12** (1972), 161–173.

[6] J. Kieffer, *The isomorphism theorem for generalized Bernoulli schemes*, Studies in Probability and Ergodic Theory, Adv. in Math. Supp. Studies **2** (1978), 251–267.

[7] B. Kitchens and K. Schmidt, *Automorphisms of compact groups*, Ergodic Theory & Dynamical Systems **9** (1989), 691–735.

[8] F. Ledrappier, *Un champ markovien peut être d'entropie nulle et melangeant*, Comptes Rendus Acad. Sci. Paris **Ser. A, 287** (1978), 561–562.

[9] D. Lind, *Ergodic automorphisms of the infinite torus are Bernoulli*, Israel Journal of Mathematics **17** (1974), 162–168.

[10] D. Lind, *The structure of skew products with ergodic group automorphisms*, Israel Journal of Mathematics **28** (1977), 205–248.

[11] D. Lind, K. Schmidt and T. Ward, *Mahler measure and entropy for commuting automorphisms of compact groups*, Inventiones mathematicæ **101** (1990), 593–629.

[12] G. Miles and R. K. Thomas, *Generalised torus automorphisms are Bernoullian*, Studies in Probability and Ergodic Theory, Adv. in Math. Supp. Studies **2** (1978), 231–249.

[13] D. S. Ornstein, *Mixing Markov shifts of kernel type are Bernoulli*, Advances in Mathematics **10** (1973), 143–6.

[14] D. S. Ornstein & B. Weiss, *Ergodic theory of amenable group actions I: the Rohlin lemma*, Bull. Amer. Math. Soc. **2** (1980), 161–164.

[15] D. S. Ornstein & B. Weiss, *Entropy and isomorphism theorems for actions of amenable groups*, Journal d'Analyse Mathématique **48** (1987), 1–141.
[16] V. A. Rokhlin, *Metric properties of endomorphisms of compact commutative groups*, Amer. Math. Soc. Transl.(3) **64** (1967), 244–252.
[17] A. Rosenthal, *Weak Pinsker property and Markov processes*, Ann. Inst. Henri Poincaré Prob. Stat. **22** (1986), 347–369.
[18] K. Schmidt, *Mixing automorphisms of compact groups and a theorem by Kurt Mahler*, Pacific Journal of Mathematics **137** (1989), 371–385.
[19] K. Schmidt, *Automorphisms of compact abelian groups and affine varieties*, Proc. London Math. Soc. **61** (1990), 480–496.
[20] P. Shields, *The theory of Bernoulli shifts*, Univ. of Chicago Press, 1973.
[21] P. Shields, *Almost block independence*, Z. Wahrsch. verw. Geb. **49** (1979), 119–123.

Department of Mathematics, Ohio State University, Columbus, OH 43210.