Ergod. Th. & Dynam. Sys. (2003), **23**, 75–85
DOI: 10.1017/S0143385702000950

Asymptotic geometry of non-mixing sequences

MANFRED EINSIEDLER† and THOMAS WARD‡

† *Mathematisches Institut, Universität Wien, Strudlhofgasse 4, A-1090, Vienna, Austria*
and Mathematics Department, State College, PA 16802, USA
(e-mail: manfred@mat.univie.ac.at)
‡ *School of Mathematics, University of East Anglia, Norwich NR4 7TJ, UK*
(e-mail: t.ward@uea.ac.uk)

(*Received* 28 *June* 2000 *and accepted in revised form* 27 *November* 2001)

Abstract. The exact order of mixing for zero-dimensional algebraic dynamical systems is not entirely understood. Here non-Archimedean norms in function fields of positive characteristic are used to exhibit an asymptotic shape in non-mixing sequences for algebraic $\mathbb{Z}^2$-actions. This gives a relationship between the order of mixing and the convex hull of the defining polynomial. Using these methods, we show that an algebraic dynamical system for which any shape of cardinality three is mixing, is mixing of order three, and for any $k \geq 1$ exhibit examples that are k-fold mixing but not $(k+1)$-fold mixing.

1. *Introduction and definitions*

Let α be a $\mathbb{Z}^d$-action by μ-preserving transformations of a non-trivial probability space $(X, \mathcal{S}, \mu)$. A sequence $(\mathbf{n}_1^{(j)}, \mathbf{n}_2^{(j)}, \ldots, \mathbf{n}_r^{(j)})$ of r-tuples of elements of $\mathbb{Z}^d$ with

$$\mathbf{n}_s^{(j)} - \mathbf{n}_t^{(j)} \to \infty \quad \text{as} \quad j \to \infty \quad \text{for every } s \neq t \tag{1.1}$$

is mixing for α, if for any sets $A_1, \ldots, A_r \in \mathcal{S}$,

$$\lim_{j\to\infty} \mu(\alpha^{-\mathbf{n}_1^{(j)}}(A_1) \cap \cdots \cap \alpha^{-\mathbf{n}_r^{(j)}}(A_r)) = \mu(A_1) \cdots \mu(A_r). \tag{1.2}$$

If any sequence satisfying (1.1) has (1.2), then α is mixing of order r. As usual, an action that is mixing of order 2 is simply called mixing. The maximum value of r for which (1.1) implies (1.2)—if this is finite—is the order of mixing of α. If (1.1) implies (1.2) for all r, then α is mixing of all orders. A finite set $\{\mathbf{n}_1, \ldots, \mathbf{n}_r\}$ of integer vectors is a mixing shape for α if

$$\lim_{k\to\infty} \mu(\alpha^{-k\mathbf{n}_1}(A_1) \cap \cdots \cap \alpha^{-k\mathbf{n}_r}(A_r)) = \mu(A_1) \cdots \mu(A_r). \tag{1.3}$$

In general, mixing properties of shapes have no bearing on the order of mixing. For example, there are non-mixing actions for which all shapes are mixing (see [**10**, Theorem 1.2]). An algebraic dynamical system—one in which X is assumed to be a compact metrizable abelian group, μ is a Haar measure, and α acts by automorphisms—with all shapes mixing is, in contrast, mixing of all orders. However, it is not clear whether non-mixing shapes detect the exact order of mixing for algebraic systems that are not mixing of all orders—see [**7**, §27, §28] for an overview of this problem, and [**6**] for Ledrappier's seminal example which showed that mixing algebraic $\mathbb{Z}^2$-actions need not be mixing of all orders.

CONJECTURE 1.1. *An algebraic dynamical system for which all shapes of cardinality r are mixing is mixing of order r.*

As remarked above, the (suitably interpreted) conjecture holds for '$r = \infty$' in the sense that all shapes mixing implies mixing of all orders for algebraic dynamical systems. It also holds when $r = 2$—that is, if each element $\alpha^{\mathbf{n}}$, $\mathbf{n} \neq 0$ is mixing, then the whole action α is mixing (see [**7**, Theorem 1.6]). In Theorem 3.2 below we show that the conjecture holds when $r = 3$ and $d = 2$. Moreover, a weaker lower bound for the order of mixing for $d = 2$ is shown in Theorem 3.1, and this is used to give examples with any given order of mixing in §4.

2. *Algebraic $\mathbb{Z}^2$-actions*

Let $R_2 = \mathbb{Z}[u_1^{\pm 1}, u_2^{\pm 1}]$, the ring of Laurent polynomials with integer coefficients in the commuting variables u_1, u_2. Following [**3**], associate a given algebraic $\mathbb{Z}^2$-action α on X to an R_2-module as follows. Let M be the countable Pontryagin (character) group of X, and define $u_j m = \hat{\alpha}^{\mathbf{e}_j}(m)$ for all $m \in M$, where $\mathbf{e}_j \in \mathbb{Z}^2$ is the jth unit vector and $\hat{\alpha}^{\mathbf{e}_j}$ is the automorphism of M dual to $\alpha^{\mathbf{e}_j}$ for $j = 1, 2$. A polynomial $F \in R_2$ has the form $F(u) = \sum_{\mathbf{n} \in \mathbb{Z}^2} c_F(\mathbf{n}) u^{\mathbf{n}}$, where $c_F(\mathbf{n}) \in \mathbb{Z}$ and $c_F(\mathbf{n}) = 0$ for all but finitely many $\mathbf{n} \in \mathbb{Z}^2$, and $u^{\mathbf{n}} = u_1^{n_1} u_2^{n_2}$. Then $Fm = \sum_{\mathbf{n} \in \mathbb{Z}^2} c_F(\mathbf{n}) \hat{\alpha}^{\mathbf{n}}(m)$ for every $m \in M$. Conversely, suppose that M is a countable R_2-module, and let $X_M = \widehat{M}$ be its compact abelian character group. Each u_j is a unit in R_2, so the map γ_j defined by $\gamma_j(m) = u_j m$ is an automorphism of M. Define an algebraic $\mathbb{Z}^2$-action α_M on X_M by $\alpha_M^{\mathbf{e}_j} = \widehat{\gamma_j}$. Thus $(X_M, \alpha_M) \leftrightarrow M$ gives a one-to-one correspondence between algebraic $\mathbb{Z}^2$-actions and countable R_2 modules. See [**7**, Ch. II] for a further explanation and many examples.

Approximating indicator functions of the sets $A_1, \ldots, A_r$ appearing in (1.2) with trigonometric polynomials shows that for the algebraic action α corresponding to the R_2-module M, (1.2) is equivalent to the statement that for any $m_1, \ldots, m_r \in M$, not all zero, the relation

$$m_1 u^{\mathbf{n}_1^{(j)}} + m_2 u^{\mathbf{n}_2^{(j)}} + \cdots + m_r u^{\mathbf{n}_r^{(j)}} = 0 \tag{2.1}$$

holds for only finitely many j.

In order to use this to make progress with Conjecture 1.1, the following well-known results are needed. A consequence of the algebraic formulation of mixing in (2.1) is that a given sequence $(\mathbf{n}_1^{(j)}, \mathbf{n}_2^{(j)}, \ldots, \mathbf{n}_r^{(j)})$ is mixing for α_M on X_M if and only if it is mixing for $\alpha_{R_2/\mathfrak{P}}$ on $X_{R_2/\mathfrak{P}}$ for every prime ideal $\mathfrak{P} \subset R_2$ associated to the module M (see [**4**, Theorem 3.3] or [**8**, Theorem 2.2] for a proof of this). Moreover, if X_M is connected (equivalently, if M is torsion-free as an additive group), then by [**8**], if α_M is mixing, it is also mixing of all orders. If $\mathfrak{P}$ is a prime ideal generated by a single prime number then $\alpha_{R_2/\mathfrak{P}}$ is the $\mathbb{Z}^2$-shift on $\{0, \dots, p-1\}^{\mathbb{Z}^2}$ and is, therefore, mixing of all orders. Finally, if $\mathfrak{P}$ contains a prime number and is not of the form $pR_2 + FR_2$ for some non-zero polynomial F, then $R_2/\mathfrak{P}$ is finite and $\alpha_{R_2/\mathfrak{P}}$ is not mixing. It follows that in order to show Conjecture 1.1 for $d = 2$, it is sufficient to study algebraic dynamical systems corresponding to cyclic modules of the form $R_2/\mathfrak{P}$ for some prime ideal $\mathfrak{P} \subset R_2$ of the form $\langle p, F\rangle = pR_2 + FR_2$, where p is a prime number and $F \in R_2$.

3. *Asymptotic geometry*

From now on we will only be working in rings of characteristic p for a fixed prime p, so we replace the coefficient ring $\mathbb{Z}$ with the finite field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. Notationally this replaces the polynomial $F \in R_2$ with the reduced mod p polynomial $f = \bar{F} \in R_2^{(p)} = \mathbb{F}_p[u_1^{\pm 1}, u_2^{\pm 2}]$, and the prime ideal

$$\mathfrak{P} = \langle p, F\rangle = pR_2 + FR_2$$

with the prime ideal

$$\mathfrak{p} = \langle f\rangle = fR_2^{(p)}.$$

Notice that $R_2/\mathfrak{P}$ is then the same ring as $R_2^{(p)}/\mathfrak{p}$. For simplicity of notation we will write $\alpha = \alpha_{R_2^{(p)}/\mathfrak{p}}$ for the $\mathbb{Z}^2$-action on the dual of $R_2^{(p)}/\mathfrak{p}$ where $\mathfrak{p} = \langle f\rangle$ for a fixed irreducible polynomial $f \in R_2^{(p)}$.

Let $\mathcal{M}(\alpha)$ denote the order of mixing—the largest value of r for which (1.1) implies (1.2)—of α. Finding $\mathcal{M}(\alpha)$ is difficult (see [**7**, §28]) even for this special class of systems. Our purpose here is to find a new inequality that relates $\mathcal{M}(\alpha)$ to the shape of the polynomial f.

The polynomial f can be written $f(u) = \sum_{\mathbf{n}\in\mathbb{Z}^2} c_f(\mathbf{n})u^{\mathbf{n}}$, where $c_f(\mathbf{n}) \in \mathbb{F}_p$ for all $\mathbf{n} \in \mathbb{Z}^2$ and $c_f(\mathbf{n}) = 0$ for all but finitely many $\mathbf{n} \in \mathbb{Z}^2$. Let

$$S(f) = \{\mathbf{n} \in \mathbb{Z}^2 \mid c_f(\mathbf{n}) \neq 0\}$$

denote the support of f, and $\mathcal{N}(f)$ the convex hull of $S(f)$. Thus $S(f)$ is some finite set of points in $\mathbb{Z}^2$, and $\mathcal{N}(f)$ is a convex polygon in $\mathbb{Z}^2$.

The relationships found below between the faces of $\mathcal{N}(f)$ and measurable properties of the action are essentially equivalent to the geometry of half-spaces and relative entropies in [**4**]. The approach taken here using non-Archimedean norms seems to be better adapted to attacking Conjecture 1.1.

THEOREM 3.1. *Assume that $\mathcal{N}(f)$ is an R-gon and f is irreducible. Let α be the $\mathbb{Z}^2$-action on the dual of $R_2^{(p)}/\langle f\rangle$. Then*

$$R - 1 \leq \mathcal{M}(\alpha) < |S(f)|.$$

THEOREM 3.2. *Conjecture 1.1 holds when $r = 3$ for algebraic $\mathbb{Z}^2$-actions.*

The method of proof of Theorem 3.1 is to show that an arbitrary non-mixing sequence for α must asymptotically reflect part of the structure of $\mathcal{N}(f)$ (the slopes of the faces). Theorem 3.2 then holds because the slopes of a triangle determine its shape.

The key step in the proof is to construct norms that reflect the geometry of $\mathcal{N}(f)$. To clarify this, an example is described in detail (see Example 3.3). Before doing this, some background on non-Archimedean norms and Newton polygons is needed (see [**2**, Ch. 2] or [**5**, Ch. IV.3] for more details). A non-Archimedean norm $|\cdot|$ on an integral domain S is a function

$$|\cdot| : S \to \mathbb{R}$$

with the properties:

(i) $|a| \geq 0$;

(ii) $|a| = 0$ only for $a = 0$;

(iii) $|a+b| \leq \max(|a|, |b|)$; and

(iv) $|ab| = |a| \cdot |b|$;

for every $a, b \in S$. A norm always extends uniquely to the field K of quotients of S (or to any intermediate ring between S and K). An immediate consequence of property (iii) is that given any equation in S of the form

$$\sum_{k=1}^{n} a_k = 0 \quad \text{with } a_k \neq 0, \tag{3.1}$$

there must be at least two indices $i \neq j$ for which

$$\max_{1 \leq k \leq n} |a_k| = |a_i| = |a_j|. \tag{3.2}$$

The Newton polygon of a polynomial $g(x) = g_0 + g_1 x + \cdots + g_n x^n \in S[x]$ with respect to the non-Archimedean norm $|\cdot|$ is defined to be the highest convex polygonal line joining $(0, -\log|g_0|)$ to $(n, -\log|g_n|)$ which passes on or below all the points $(j, -\log|g_j|)$ for $j = 0, \ldots, n$. The vertices of the Newton polygon are the points $(j, -\log|g_j|)$ where the slope changes, and the slope of a line in the Newton polygon joining the vertices $(i, -\log|g_i|)$ and $(j, -\log|g_j|)$ is $(\log|g_j| - \log|g_i|)/(i - j)$. The basic property of the Newton polygon is the following: for each slope λ, there is a non-Archimedean norm $|\cdot|^{(\lambda)}$ on the extension ring $S[x]/\langle g\rangle$, coinciding with $|\cdot|$ on the constant polynomials (identified with S), and with $|x|^{(\lambda)} = \exp(\lambda)$.

Up to a natural equivalence the (non-trivial) non-Archimedean norms on $S = \mathbb{F}_p[u^{\pm 1}]$ are those of the form

$$|f|_g = p^{-\mathrm{ord}_g f}, \tag{3.3}$$

(where $g \in \mathbb{F}_p[u]$ is an irreducible polynomial, and $\mathrm{ord}_g f$ denotes the multiplicity of g in the prime decomposition of f), together with the exceptional norm

$$|f|_\infty = p^{\deg f}.$$

Example 3.3. Let $f(u_1, u_2) = u_2 + u_1 + u_1^3 u_2$, and view f as an element of $\mathbb{F}_p[u_2^{\pm 1}][u_1^{\pm 1}]$ (cf. Figure 1, showing the support of f as dots). Choose the norm $|\cdot| = |\cdot|_{u_2}$ on $\mathbb{F}_p[u_2^{\pm 1}]$, so that $|u_2| = 1/p$.

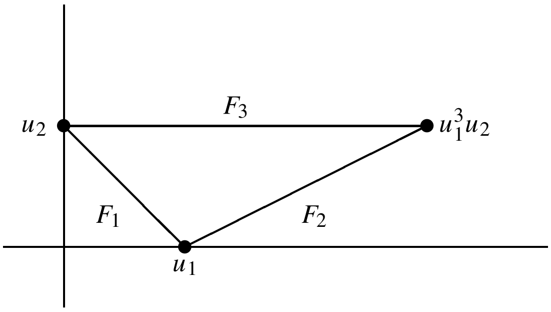

FIGURE 1. The faces of $\mathcal{N}(u_2 + u_1 + u_1^3u_2)$.

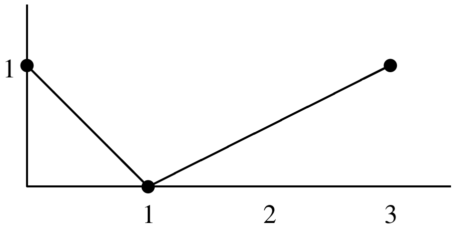

FIGURE 2. The Newton polygon of $f \in \mathbb{F}_p[u_2^{\pm 1}][u_1^{\pm 1}]$ with respect to $|\cdot|$.

How this norm extends to the ring extension $R = \mathbb{F}_p[u_2^{\pm 1}][u_1^{\pm 1}]/\langle f\rangle$ is determined by the Newton polygon of f viewed as a polynomial for u_1 with coefficients in $\mathbb{F}_p[u_2^{\pm 1}]$:

$$f(u_1) = u_2 \cdot u_1^0 + 1 \cdot u_1^1 + 0 \cdot u_1^2 + u_2 \cdot u_1^3. \tag{3.4}$$

The four points that define the Newton polygon (Figure 2) are therefore

$$\begin{aligned}(0, -\log_p|u_2|) &= (0, 1),\\ (1, -\log_p|1|) &= (1, 0),\\ (2, -\log_p|0|) &= (2, \infty)\end{aligned}$$

and

$$(3, -\log_p|u_2|) = (3, 1)$$

(logarithms base p are used for convenience, and $|0| = 0$). Notice that the Newton polygon of f shown in Figure 2 does not coincide with the convex hull of the support in Figure 1, but they have the same faces pointing towards negative powers of u_2 in Figure 1 (equivalently, towards monomials for which $|\cdot|$ is big).

From the Newton polygon in Figure 2, it follows that there are two norms $|\cdot|_1$, $|\cdot|_2$ extending $|\cdot|$ to R; the first has $|u_1|_1 = 1/p$ (from the line segment with slope -1) and the second has $|u_1|_2 = \sqrt{p}$ (from the line segment with slope $1/2$).

Thus the vector $\begin{pmatrix}\log_p|u_1|_1\\ \log_p|u_2|_1\end{pmatrix} = \begin{pmatrix}-1\\-1\end{pmatrix}$ is an outward normal to the face F_1 of $\mathcal{N}(f)$. The same expression using $|\cdot|_2$ gives $\begin{pmatrix}\log_p|u_1|_2\\ \log_p|u_2|_2\end{pmatrix} = \begin{pmatrix}1/2\\-1\end{pmatrix}$, an outward normal to the face F_2.

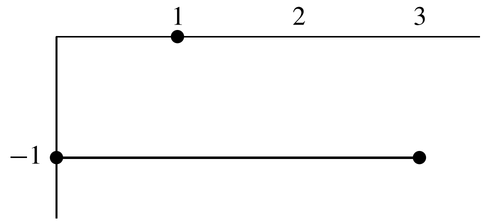

FIGURE 3. The Newton polygon of $f \in \mathbb{F}_p[u_2^{\pm 1}][u_1^{\pm 1}]$ with respect to $|\cdot|'$.

If the norm $|\cdot|' = |\cdot|_\infty$ on $\mathbb{F}_p[u_2]$ is chosen initially, then $|u_2|' = p$ (so that the monomials with big norm are in the upper-half plane of Figure 1) and the corresponding Newton polygon is determined from (3.4) by the points

$$\begin{aligned}(0, -\log_p|u_2|) &= (0, -1),\\ (1, -\log_p|1|) &= (1, 0),\\ (2, -\log_p|0|) &= (2, \infty)\end{aligned}$$

and

$$(3, -\log_p|u_2|) = (3, -1).$$

The resulting Newton polygon is shown in Figure 3; it shows there is only one extension to R, and the resulting norm $|\cdot|_1'$ has the property that $\left(\begin{smallmatrix}\log_p |u_1|_1'\\ \log_p |u_2|_1'\end{smallmatrix}\right) = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ is an outward normal to the face F_3 of $\mathcal{N}(f)$.

Example 3.3 generalizes to the next lemma.

LEMMA 3.4. *For each face F of the convex hull $\mathcal{N}(f)$, there is a norm $|\cdot|^{(F)}$ on the ring $R_2^{(p)}/\langle f\rangle$ with the property that the vector $\left(\begin{smallmatrix}\log_p |u_1|^{(F)}\\ \log_p |u_2|^{(F)}\end{smallmatrix}\right)$ is an outward normal to the face F of $\mathcal{N}(f)$.*

Proof. Choose a face F of $\mathcal{N}(f)$. If necessary, exchange u_1 and u_2 so that F is not a vertical line. By replacing u_2 with u_2^{-1} if necessary, assume further that F is one of the lower faces of $\mathcal{N}(f)$ when drawn in the plane (that is, like F_1 or F_2 in Figure 1). Let $|\cdot| = |\cdot|_{u_2}$ be the norm corresponding to the irreducible polynomial $g = u_2$ in $\mathbb{F}_p[u_2]$ as in (3.3). We may also assume (after multiplying f by a suitable monomial) that f is a polynomial in u_1 with coefficients in $\mathbb{F}_p[u_2]$,

$$f = \sum_{k=0}^{n} q_i(u_2)u_1^k \quad \text{with } q_0q_n \neq 0. \tag{3.5}$$

For each non-zero coefficient q_i, let m_i be the largest m for which u_2^m divides $q_i(u_2)$ in $\mathbb{F}_p[u_2]$. By the definition of the norm $|\cdot|$,

$$-\log_p|q_i(u_2)| = m_i.$$

On the other hand, the coefficient of $u_2^{m_i}$ in q_i is non-zero, so $(i, m_i) \in S(f)$. This shows that the points (i, m_i) appear both in the *Newton polygon* of f (considered as a polynomial in u_1 as in (3.5)) and in the *support* of f. Comparing $\mathcal{N}(f)$ and the Newton polygon shows that the lower faces of $\mathcal{N}(f)$ comprise exactly the Newton polygon of f. Thus, there is a norm $|\cdot|^{(F)}$ on $R_2^{(p)}/\langle f\rangle$ extending $|\cdot|$ for which

$$|u_1|^{(F)} = p^{\lambda},$$

where λ is the slope of F. The vector

$$\begin{pmatrix}\log_p |u_1|^{(F)}\\ \log_p |u_2|^{(F)}\end{pmatrix} = \begin{pmatrix}\lambda\\ -1\end{pmatrix}$$

is therefore an outward normal to the face F. □

These norms will now be used to show that a non-mixing sequence must asymptotically approximate some of the structure of $\mathcal{N}(f)$, by applying the simple observation that (3.1) implies (3.2) to the algebraic characterization of mixing, using a norm adapted to the shape $\mathcal{N}(f)$.

Proposition 3.5. *Assume that* $(A^{(j)}) = (\mathbf{n}_1^{(j)}, \mathbf{n}_2^{(j)}, \ldots, \mathbf{n}_r^{(j)})$ *is a sequence in* $(\mathbb{Z}^2)^r$ *with the property that*

$$m_1 u^{\mathbf{n}_1^{(j)}} + m_2 u^{\mathbf{n}_2^{(j)}} + \cdots + m_r u^{\mathbf{n}_r^{(j)}} = 0 \qquad (3.6)$$

for all j, *where* $m_1, \ldots, m_r \in R_2^{(p)}/\mathfrak{p}$ *are non-zero. Write* $\mathcal{N}(A^{(j)})$ *for the convex hull of the set* $\{\mathbf{n}_1^{(j)}, \mathbf{n}_2^{(j)}, \ldots, \mathbf{n}_r^{(j)}\}$. *Fix a face* F *of* $\mathcal{N}(f)$. *Then there is a constant* $K > 0$ *such that there is a face of* $\mathcal{N}(A^{(j)})$ *spanned (without loss of generality) by* $\mathbf{n}_1^{(j)}, \mathbf{n}_2^{(j)}$, *and a vector* $\mathbf{m}^{(j)}$ *with the property that the line through* $\mathbf{n}_1^{(j)}, \mathbf{m}^{(j)}$ *is parallel to* F *and* $\|\mathbf{m}^{(j)} - \mathbf{n}_2^{(j)}\| \le K$.

Proof. Pick a face F of $\mathcal{N}(f)$ and use Lemma 3.4 to find a norm $|\cdot|$ on $R_2^{(p)}/\langle f\rangle$ so that $\begin{pmatrix}\log_p |u_1|\\ \log_p |u_2|\end{pmatrix}$ is an outward normal to $\mathcal{N}(f)$ through F. Let $L = 2\max_{i=1,\ldots,r}\{|\log_p |m_i||\}$.

Fix j for now, and choose t to maximize $|u^{\mathbf{n}_t^{(j)}}|$. This corresponds to $\mathbf{n}_t^{(j)}$ being extremal in $\mathcal{N}(A^{(j)})$ in the direction of $\begin{pmatrix}\log_p |u_1|\\ \log_p |u_2|\end{pmatrix}$. Let ℓ be the line through $\mathbf{n}_t^{(j)}$ parallel to F. Assume that no other point of $A^{(j)}$ lies within L of the line ℓ. Then for $i \neq t$ we get

$$\mathbf{n}_i^{(j)}\begin{pmatrix}\log_p |u_1|\\ \log_p |u_2|\end{pmatrix} < \mathbf{n}_t^{(j)}\begin{pmatrix}\log_p |u_1|\\ \log_p |u_2|\end{pmatrix} - L$$

and

$$|m_i u^{\mathbf{n}_i^{(j)}}| = |m_i||u^{\mathbf{n}_i^{(j)}}| < p^{-L}|m_i||u^{\mathbf{n}_t^{(j)}}| \le |m_t u^{\mathbf{n}_t^{(j)}}|.$$

This shows that in (3.6) one term is bigger than all the others, which is impossible for a non-Archimedean norm $|\cdot|$. Therefore, for every j there must be a second point $\mathbf{n}_s^{(j)}$ within distance L of the line ℓ. If necessary, pass to a subsequence so that t and s are independent of j. By renaming the indices assume $t = 1, s = 2$.

Let V be the rational subspace normal to $\binom{\log_p |u_1|}{\log_p |u_2|}$, and choose a basis $\mathbf{b}_1, \mathbf{b}_2$ of $\mathbb{Z}^2$ with $\mathbf{b}_1 \in V$. For every $\mathbf{n} \in \mathbb{Z}^2$ the component $a_2\mathbf{b}_2$ of $\mathbf{n} = a_1\mathbf{b}_1 + a_2\mathbf{b}_2$ can be found by projection along V. Since $\mathbf{n}_2^{(j)}$ is within L of the line $\ell = V + \mathbf{n}_1^{(j)}$, the projections of $\mathbf{n}_1^{(j)}$ and $\mathbf{n}_2^{(j)}$ onto $\mathbf{b}_2\mathbb{Z}$ are close together, say within distance K. This shows that $\mathbf{n}_2^{(j)} = \mathbf{m}^{(j)} + \mathbf{c}^{(j)}$ with $\mathbf{m}^{(j)} \in V + \mathbf{n}_1^{(j)}$, $\|\mathbf{c}^{(j)}\| < K$. □

Proof of Theorem 3.1. First recall that $S(f)$ is automatically a non-mixing shape for α (see [**7**, Examples 27.1]), so

$$\mathcal{M}(\alpha) < |S(f)|.$$

On the other hand, the convex hull $\mathcal{N}(A^{(j)})$ is a convex polygon. By Proposition 3.5, each face F of $\mathcal{N}(f)$ must appear with a uniformly bounded error as one of the faces of $\mathcal{N}(A^{(j)})$. Since the differences in (1.1) go to infinity, the slope of this matching face approaches the slope of F. It follows that there must be at least R faces on the convex set $\mathcal{N}(A^{(j)})$, so $A^{(j)}$ must have at least R points. Thus $R - 1 \leq \mathcal{M}(\alpha)$. □

Proof of Theorem 3.2. If $\mathcal{N}(f)$ lies on a line, then α cannot be mixing. If $\mathcal{N}(f)$ is an R-gon with $R > 3$ then Theorem 3.1 shows that $\mathcal{M}(\alpha) \geq 3$. So assume that $\mathcal{N}(f)$ is a triangle, with vertices $\mathbf{d}_i$ for $i = 1, 2, 3$. Let F_i be the face of $\mathcal{N}(f)$ spanned by $\mathbf{d}_i, \mathbf{d}_{i+1}$ (reduce subscripts mod 3). Assume additionally that α is not mixing of order 3. We will deduce that there is a shape of cardinality 3 which is not mixing.

There exist non-zero elements $a, b, c \in R_2^{(p)}/\langle f\rangle$, and three sequences $\mathbf{n}_i^{(j)}$ of lattice points with (1.1) and

$$au^{\mathbf{n}_1^{(j)}} + bu^{\mathbf{n}_2^{(j)}} + cu^{\mathbf{n}_r^{(j)}} = 0 \tag{3.7}$$

for all j. Without loss of generality, assume that $\mathbf{n}_i^{(j)}, \mathbf{n}_{i+1}^{(j)}$ (again, reduce subscripts mod 3) are the two points in $A^{(j)}$ satisfying Proposition 3.5 for the face F_i. Let $i = 1$. Then there exists $\mathbf{m}^{(j)} \in \mathbb{Z}^2$ bounded away from $\mathbf{n}_2^{(j)}$ such that $\mathbf{m}^{(j)}$ is actually on the line through $\mathbf{n}_1^{(j)}$ parallel to F_1. Passing to a subsequence, we may assume that

$$\mathbf{n}_2^{(j)} - \mathbf{m}^{(j)} = \mathbf{k}$$

is independent of j. Transform equation (3.7) to

$$au^{\mathbf{n}_1^{(j)}} + (bu^{\mathbf{k}})u^{\mathbf{m}^{(j)}} + cu^{\mathbf{n}_r^{(j)}} = 0.$$

In other words, by changing b slightly we can assume that $\mathbf{n}_1^{(j)}$ and $\mathbf{n}_2^{(j)}$ are actually on a line parallel to F. By changing $\mathbf{n}_2^{(j)}$ if necessary by another bounded amount, we can assume that there exists $s_j \in \mathbb{N}$ with

$$\mathbf{n}_2^{(j)} - \mathbf{n}_1^{(j)} = s_j(\mathbf{d}_2 - \mathbf{d}_1).$$

Repeat this for the face F_2, and change the vector $\mathbf{n}_3^{(j)}$ accordingly. Now we know that there exists $t_j \in \mathbb{N}$ such that

$$\mathbf{n}_3^{(j)} - \mathbf{n}_2^{(j)} = t_j(\mathbf{d}_3 - \mathbf{d}_2).$$

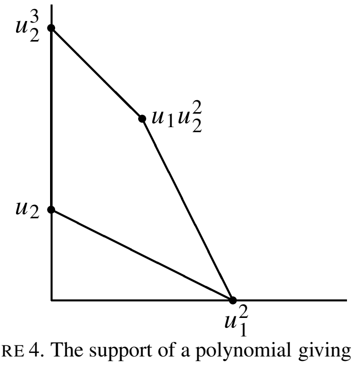

FIGURE 4. The support of a polynomial giving three-fold mixing.

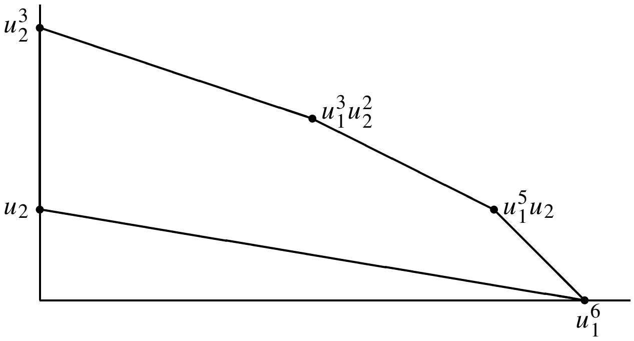

FIGURE 5. The support of a polynomial giving four-fold mixing.

Using Proposition 3.5 again for the face F_3, we see that $\mathbf{n}_1^{(j)}$ and $\mathbf{n}_3^{(j)}$ are almost on the same line parallel to F_3. This almost similarity between $\mathcal{N}(f)$ and $\mathcal{N}(A^{(j)})$ shows that $|s_j - t_j|$ is bounded. Changing $\mathbf{n}_3^{(j)}$, again by a uniformly bounded amount, we can achieve $s_j = t_j$ so the triangles are similar. This shows that $\{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ is a non-mixing shape. □

4. *Examples*

Example 4.1. Theorem 3.1 shows that if f is an irreducible polynomial for which the support $S(f)$ coincides with the extreme points of the Newton polygon $\mathcal{N}(f)$, then $\mathcal{M}(\alpha) = |S(f)| - 1$. In order to produce an example α with prescribed order of mixing $\mathcal{M}(\alpha) = k$, it is therefore sufficient to exhibit such an irreducible polynomial with $|S(f)| = k + 1$. This may be done using Eisenstein's irreducibility criterion (see [**1**] for a general norm-theoretic treatment of the Eisenstein criterion). Two simple examples will illustrate the method; it is clear from these how to build an example for any order of mixing. We are grateful to Klaus Schmidt for pointing out that an explicit construction of such a family of examples was not known previously.

(1) To find an example with order of mixing 3, consider $f(u_1, u_2) = u_1^2 + u_1u_2^2 + u_2^3 + u_2 \in \mathbb{F}[u_2][u_1]$; the prime $u_2 \in \mathbb{F}[u_2]$ divides the coefficients u_2^2 and $u_2^3 + u_2$, but u_2^2 does not divide the coefficient $u_2^3 + u_2$. By Eisenstein's criterion f is irreducible. The support of the polynomial is shown in Figure 4.

(2) To find an example with mixing of order 4, let $f(u_1, u_2) = u_1^6 + u_1^5 u_2 + u_1^3 u_2^2 + u_2 + u_2^3$. As before, this is seen to be irreducible by viewing it as a polynomial in u_1 with coefficients in $\mathbb{F}[u_2]$. The support of the polynomial is shown in Figure 5.

Note that in these examples we are choosing the shape of the support freely; it is also possible to find examples for which any prescribed shape is the minimal non-mixing shape by [**10**], though not in a constructive fashion.

Example 4.2. Theorem 3.2 shows that the system corresponding to the ideal $\mathfrak{p} = \langle 2, 1 + u_1 + u_2 + u_2^2\rangle$ is 3-mixing, answering a question in [**7**, p. 283].

Example 4.3. In the previous example, we used the fact from [**7**] that no shape with cardinality 3 is non-mixing. An alternative method to show this is to use a result of Voloch on solutions to $ax + by = 1$ in function fields. Consider again $\mathfrak{p} = \langle 2, 1 + u_1 + u_2 + u_2^2\rangle$; then Theorem 3.1 states that

$$2 \le \mathcal{M}(\alpha) < 4,$$

and we wish to show that $\mathcal{M}(\alpha) = 3$. To see this, assume that

$$(\mathbf{n}_1^{(j)}, \mathbf{n}_2^{(j)}, \mathbf{n}_3^{(j)} = 0)$$

is a non-mixing sequence for α with $\mathbf{n}_s^{(j)} - \mathbf{n}_t^{(j)} \to \infty$ as $j \to \infty$ for $s \neq t$. Then there are elements m_1, m_2, m_3 of $R_2/\mathfrak{p}$, not all zero, with

$$m_1 u^{\mathbf{n}_1^{(j)}} + m_2 u^{\mathbf{n}_2^{(j)}} = -m_3 \tag{4.1}$$

for infinitely many j. The field of fractions of $R_2/\mathfrak{p}$ may be identified with $\mathbb{F}_2(t)$ by the map $u_1 \mapsto t$, $u_2 \mapsto 1 + t + t^2$, and in this field (4.1) becomes

$$ax + by = 1 \tag{4.2}$$

with infinitely many solutions for x, y in the finitely generated multiplicative subgroup $G = \langle\langle t, 1 + t + t^2\rangle\rangle$ of $\mathbb{F}_2(t)^*$. By [**9**], it follows that (4.2) is a *G-trivial* equation: there is an $n \ge 1$ for which $a^n, b^n \in G$. Since G is generated by irreducible polynomials, this can only be true if $a, b \in G$. So there is an infinite family of equations

$$u^{\mathbf{m}_1^{(j)}} + u^{\mathbf{m}_2^{(j)}} = 1 \tag{4.3}$$

with $\mathbf{m}_1^{(j)}$, $\mathbf{m}_2^{(j)}$, and $\mathbf{m}_1^{(j)} - \mathbf{m}_2^{(j)} \to \infty$ as $j \to \infty$. By considering the shape of $\mathcal{N}(1 + u_1 + u_1^2 + u_2)$, this shows that the polynomial in (4.3) has the same shape as $\mathcal{N}(1 + u_1 + u_1^2 + u_2)$, so (without loss of generality), $\mathbf{m}_1^{(j)} = (0, m(j))$ and $\mathbf{m}_2^{(j)} = (2m(j), 0)$ for some $m(j) \to \infty$. Thus the equation reduces to

$$(1 + t + t^2)^{m(j)} = 1 + t^{2m(j)}. \tag{4.4}$$

Write $m(j) = 2^e \ell$, ℓ odd, for some $e \ge 0$. Then the left-hand side of (4.4) is

$$\begin{aligned}(1 + t + t^2)^{2^e \ell} &= (1 + t + O(t^2))^{2^e}\\ &= 1 + t^{2^e} + O(t^2)^{2^e}\\ &= 1 + t^{2^{e+1}\ell},\end{aligned}$$

which is impossible. It follows that $\mathcal{M}(\alpha) = 3$.

5. *Further results*

Remark 5.1. (1) An extension of the arguments above using the product formula for norms in function fields can be used to show that the set of ratios of lengths of faces of $\mathcal{N}(A^{(j)})$ in a sequence that witnesses the failure of R-fold mixing is bounded.

(2) If $S(f)$ is a rectangle, then the methods used above show that the only non-mixing sequences of cardinality 4 asymptotically arise from the non-mixing shape $S(f)$. However, if $|S(f)| = 4$ and $S(f)$ has a pair of non-parallel sides, then this approach does not give anything stronger than (1) above. Since this approach does not give Conjecture 1.1 even for $r = 4$ it has not been pursued further.

(3) Further progress on these problems, particularly when $S(f) \cap \mathcal{N}(f)^\circ$ is non-empty, seems to require Diophantine results on S-unit equations in finite characteristic.

The method used to prove Proposition 3.5 may also be applied to prove the following theorem, the first part of which relates to $\mathbb{Z}^d$-actions for any $d \geq 2$.

THEOREM 5.2. (1) *If a sequence* $(\mathbf{n}_1^{(j)}, \mathbf{n}_2^{(j)}, \ldots, \mathbf{n}_r^{(j)})$ *in* $(\mathbb{Z}^d)^r$ *has the property that, for every* $s \neq t$,

$$\frac{\mathbf{n}_s^{(j)} - \mathbf{n}_t^{(j)}}{\|\mathbf{n}_s^{(j)} - \mathbf{n}_t^{(j)}\|} \longrightarrow \mathbf{v}(s, t),$$

for some vector $\mathbf{v}(s, t)$ *whose entries are linearly independent over* $\mathbb{Q}$, *then the sequence is a mixing sequence for any mixing algebraic* $\mathbb{Z}^d$*-action.*

(2) *Now fix* $d = 2$ *and let* α *be determined by an irreducible polynomial* $f \in R_2^{(p)}$ *as above. Then any shape that does not contain all the faces of* $\mathcal{N}(f)$ *as directions of differences is a mixing shape for* α.

Note added in proof. Professor David Masser has now proved Conjecture 1.1.

Acknowledgement. M. Einsiedler was supported by EPSRC Grant GR/M 49588.

REFERENCES

[1] J. W. S. Cassels. *Local Fields*. Cambridge University Press, Cambridge, 1986.

[2] D. Goss. *Basic Structures of Function Field Arithmetic*. Springer, Berlin, 1998.

[3] B. Kitchens and K. Schmidt. Automorphisms of compact groups. *Ergod. Th. & Dynam. Sys.* **9**(4) (1989), 691–735.

[4] B. Kitchens and K. Schmidt. Mixing sets and relative entropies for higher-dimensional Markov shifts. *Ergod. Th. & Dynam. Sys.* **13**(4) (1993), 705–735.

[5] N. Koblitz. *p-adic Numbers, p-adic Analysis, and Zeta-functions*. Springer, New York, 1977.

[6] F. Ledrappier. Un champ markovien peut être d'entropie nulle et mélangeant. *C. R. Acad. Sci. Paris Sér. A–B* **287** (1978), A561–A563.

[7] K. Schmidt. *Dynamical Systems of Algebraic Origin*. Birkhäuser, Basel, 1995.

[8] K. Schmidt and T. Ward. Mixing automorphisms of compact groups and a theorem of Schlickewei. *Invent. Math.* **111**(1) (1993), 69–76.

[9] J. F. Voloch. The equation $ax + by = 1$ in characteristic p. *J. Number Theory* **73**(2) (1998), 195–200.

[10] T. Ward. Three results on mixing shapes. *Proceedings of the New York Journal of Mathematics Conference*, June 9–13, 1997. *New York J. Math.* **3A** (1997/1998), 1–10 (electronic).