Ledrappier's system is almost mixing of all orders

L. ARENAS-CARMONA†, D. BEREND‡ and V. BERGELSON§

† *Department of Mathematics, University of Chile, Casilla 653, Santiago, Chile*
(e-mail: learenas@uchile.cl)

‡ *Departments of Mathematics and Computer Science, Ben-Gurion University, Beer Sheva 84105, Israel*
(e-mail: berend@math.bgu.ac.il)

§ *Department of Mathematics, Ohio State University, Columbus, OH 43210, USA*
(e-mail: vitaly@math.ohio-state.edu)

Abstract. We consider Ledrappier's dynamical system, which was the first example of a $\mathbb{Z}^2$-action which is 2-mixing but not 3-mixing. Our main result is that, excluding certain small 'constructible' sets, the system is mixing of every order.

1. *Introduction*

Let k be a positive integer and let $T_1, \dots, T_k$ be invertible commuting measure-preserving transformations of a probability space $(X, \mathcal{B}, \mu)$. The corresponding $\mathbb{Z}^k$-action $(T^{\mathbf{n}})_{\mathbf{n}\in\mathbb{Z}^k}$ (where, for $\mathbf{n} = (n_1, \dots, n_k)$, we denote $T^{n_1} \dots T^{n_k}$ by $T^{\mathbf{n}}$) is r-*mixing*, for $r \geq 2$, if

$$\mu\left(\bigcap_{i=1}^{r} T^{\mathbf{n}_i} A_i\right) \longrightarrow \prod_{i=1}^{r} \mu(A_i), \qquad A_i \in \mathcal{B},\ 1 \leq i \leq r,$$

as $\mathbf{n}_i - \mathbf{n}_j \to \infty$ for all $i \neq j$†. The question whether, for $\mathbb{Z}$-actions, 2-mixing necessarily implies r-mixing for every r is an old open problem in ergodic theory. One of the first non-trivial results related to this problem was established by Rohlin [**6**], who proved that mixing (i.e. 2-mixing) endomorphisms of compact groups are mixing of all orders. Rohlin's result was generalized subsequently to more general classes of transformations, and conventional wisdom started leaning towards the belief that the answer to the above question is positive. Since there is, on the face of it, nothing special about $\mathbb{Z}$-actions, there was also a tendency to think that the situation for $\mathbb{Z}^k$-actions with $k > 1$ is similar. To the surprise of many, Ledrappier [**3**] proved that this is not the case. He provided an example of a pair σ, τ

† Note that the conventional notion of mixing, namely the condition that $\mu(A \cap T^{\mathbf{n}} B) \underset{\mathbf{n}\to\infty}{\longrightarrow}$ for all A, $B \in \mathcal{B}$, corresponds to 2-mixing.

of commuting mixing automorphisms of a compact abelian group G, such that for some measurable set $A \subseteq G$ one has

$$\mu(A \cap \sigma^{2^n} A \cap \tau^{2^n} A) \underset{n\to\infty}{\not\longrightarrow} \mu(A)^3.$$

Ledrappier's work has served as an impetus for new and interesting developments and, indeed, has led to the creation of a new branch of ergodic theory, which studies $\mathbb{Z}^d$-actions by automorphisms of compact abelian groups, and has strong connections to abstract algebra and number theory (see [**8**]).

Our goal in this paper is to undertake a deeper study of the higher-order mixing properties of Ledrappier's example, in the hope that this will shed new light on other similar (and more general) examples. One of the natural questions addressed in this paper concerns the nature of the obstacles to higher-order mixing. We will show that, in fact, Ledrappier's example is 'almost mixing of all orders'.

To formulate this result formally (and to prove it), we have to review first Ledrappier's construction and introduce some notation and definitions. To give the reader a feeling of the kind of results to be proved subsequently, we will formulate now a special case of our main result, which describes completely the obstacle to 3-mixing in Ledrappier's example.

Put

$$\mathcal{L} = \{\{(a, b), (a + 2^k, b), (a, b + 2^k)\} : a, b \in \mathbb{Z},\ k \in \mathbb{Z}_+\}$$

(where $\mathbb{Z}_+ = \{0, 1, 2, \ldots\}$). We view $\mathcal{L}$ as a set of triangles in the two-dimensional integer lattice, obtained from the single triangle $\{(0, 0), (1, 0), (0, 1)\}$ by dilations by powers of 2 and translations.

Denote by $\rho(C, D)$ the Hausdorff distance between subsets C and D of $\mathbb{Z}^2$,

$$\rho(C, D) = \max\left\{\sup_{c\in C}\inf_{d\in D} \|c - d\|, \sup_{d\in D}\inf_{c\in C} \|c - d\|\right\}, \quad C, D \subseteq \mathbb{Z}^2,$$

where $\|\cdot\|$ is the maximum norm on $\mathbb{Z}^2$†. As usual in metric spaces, we shall also denote by $\rho(C, \mathcal{D})$ the minimal distance between a set $C \subseteq \mathbb{Z}^2$ and a collection $\mathcal{D}$ of subsets of $\mathbb{Z}^2$.

Two sequences $(\mathbf{v}_n)$ and $(\mathbf{w}_n)$ of pairs of integers *grow apart as* $n \to \infty$ if $\|\mathbf{v}_n - \mathbf{w}_n\| \underset{n\to\infty}{\longrightarrow} \infty$.

THEOREM 1.1. *Let* σ, τ *be the automorphisms from Ledrappier's example, and*

$$((a_1^{(n)}, b_1^{(n)}), (a_2^{(n)}, b_2^{(n)}), (a_3^{(n)}, b_3^{(n)}))_{n=1}^{\infty}$$

be a sequence of triples of integer pairs. Then

$$\mu(\sigma^{a_1^{(n)}} \tau^{b_1^{(n)}} A \cap \sigma^{a_2^{(n)}} \tau^{b_2^{(n)}} B \cap \sigma^{a_3^{(n)}} \tau^{b_3^{(n)}} C) \underset{n\to\infty}{\longrightarrow} \mu(A)\mu(B)\mu(C), \qquad A, B, C \in \mathcal{B},$$

if and only if the following conditions hold.

† We shall denote by ρ the Hausdorff distance between subsets of $(\mathbb{Z}^2)^r$ as well.

(1) *The sequences* $(a_i^{(n)}, b_i^{(n)})$ *and* $(a_j^{(n)}, b_j^{(n)})$ *grow apart for* $i \neq j$.

(2) $\rho(D_n, \mathcal{L}) \underset{n\to\infty}{\longrightarrow} \infty$, *where* $D_n = \{(a_i^{(n)}, b_i^{(n)}) : 1 \le i \le 3\}$†.

The theorem indicates that the only obstruction to 3-mixing in Ledrappier's system is what happens along powers of 2. Our main interest in this paper is to understand (the generalized form of) this phenomenon for mixing of any order. The main result of the paper is that, roughly speaking, the only obstacle to mixing of any order is connected to exceptional behavior of Ledrappier's system along certain explicitly described rarified sets. For example, we obtain mixing of all orders along 'most' systems of polynomial sequences (cf. Proposition 8.2 and Theorem 8.18).

Let us mention in passing a few relevant results regarding high-order mixing of algebraic dynamical systems, and the related algebraic tools required to tackle such systems. First, note that the situation is simpler for connected groups. In fact, Schmidt and Ward [**9**] showed that, for such systems, mixing implies mixing of all orders. When passing from this ergodic-theoretical result, by duality, to the equivalent algebraic claim, one obtains certain equations over fields of characteristic 0, which have to be shown to admit only finitely many solutions. This step is accomplished using [**7**]. Ledrappier's example shows that, in the totally disconnected case, mixing does not imply mixing of higher orders. When studying the degree of mixing of such systems, one needs again to consider various equations arising from considering the dual action. This time, the equations are over fields of finite characteristic. The relevant algebraic tools are now provided by results such as those obtained by Masser [**5**], who was able to show that the degree of mixing of such systems is completely determined by the non-mixing shapes. (For another result in this realm, which deals only with the special case of S-unit equations with two indeterminates, see Voloch [**13**].)

In §2 we present Ledrappier's example in detail, and provide some more background. The main result of the paper is presented in §3. In §4 we digress to study in detail the case of 4-mixing, where we manage to draw a complete picture of the situation. Section 5 contains an algebraic result, which is crucial for the proof of the main theorem. In §6 we prove the main theorem. We would like to note that some of our arguments are reminiscent of those in [**1**]. Section 7 provides more details on the sets which are the obstacles to high-order mixing. Finally, in §8 we treat some general examples which confirm the claim contained in the title of the paper.

2. *Ledrappier's example*

Let $\mathbb{F}_2$ denote the field of two elements. We start with the set $\mathbb{F}_2^{\mathbb{Z}^2}$, considered as the set of all double sequences over $\mathbb{F}_2$. Equipped with the product topology and coordinate-wise addition, $\mathbb{F}_2^{\mathbb{Z}^2}$ forms a compact abelian group. The Haar measure on $\mathbb{F}_2^{\mathbb{Z}^2}$ is the product measure obtained by taking the normalized counting measure of $\mathbb{F}_2$. On $\mathbb{F}_2^{\mathbb{Z}^2}$ we have a

† Thus, for example, if $a_n, b_n, |a_n - b_n| \underset{n\to\infty}{\longrightarrow} \infty$, then

$$\mu(A_1 \cap \sigma^{a_n} A_2 \cap \tau^{b_n} A_3) \underset{n\to\infty}{\longrightarrow} \mu(A_1)\mu(A_2)\mu(A_3)$$

for any measurable sets A_1, A_2, A_3.

leftward shift σ and a downward shift τ. The former is defined by

$$\sigma((v_{mn})_{m,n=-\infty}^{\infty}) = (v_{m+1,n})_{m,n=-\infty}^{\infty}$$

and the latter by

$$\tau((v_{mn})_{m,n=-\infty}^{\infty}) = (v_{m,n+1})_{m,n=-\infty}^{\infty}.$$

Obviously, the set

$$G = \{(v_{mn})_{m,n=-\infty}^{\infty} : v_{mn} + v_{m+1,n} + v_{m,n+1} = 0,\ (m,\ n) \in \mathbb{Z}^2\}$$

is a compact subgroup of $\mathbb{F}_2^{\mathbb{Z}^2}$, invariant under both σ and τ. Our object of study is the measure-preserving $\mathbb{Z}^2$-action $(\sigma^m\tau^n)_{m,n=-\infty}^{\infty}$ on the probability space $(G,\ \mathcal{B},\ \mu)$, where $\mathcal{B}$ is the Borel field of G and μ the normalized Haar measure on G.

It will be convenient to identify points in $\mathbb{F}_2^{\mathbb{Z}^2}$ with formal power series. Namely, a point $(v_{mn})_{m,n=-\infty}^{\infty} \in \mathbb{F}_2^{\mathbb{Z}^2}$ is identified with the power series

$$\sum_{m,n=-\infty}^{\infty} v_{mn}x^{-m}y^{-n}.$$

Thus the actions of σ and τ correspond to multiplication by x and y, respectively. (Note that the set of all power series $\sum_{m,n=-\infty}^{\infty} v_{mn}x^{-m}y^{-n}$ does not admit a 'natural' multiplication operation, but the product of a power series and a polynomial in $\mathbb{F}_2[x^{\pm 1},\ y^{\pm 1}]$ is well defined.) The dual group of $\mathbb{F}_2^{\mathbb{Z}^2}$ may be identified with $\mathbb{F}_2[x^{\pm 1},\ y^{\pm 1}]$ as follows. The value of a character (corresponding to a polynomial $\sum_{(m,n)\in S} x^m y^n$) at the point $(v_{mn})_{m,n=-\infty}^{\infty} \in \mathbb{F}_2^{\mathbb{Z}^2}$ is $(-1)^{\sum_{(m,n)\in S} v_{mn}}$, where we note that the exponent $\sum_{(m,n)\in S} v_{mn}$ is the free term of the product $\sum_{(m,n)\in S} x^m y^n \cdot \sum_{m,n=-\infty}^{\infty} v_{mn}x^{-m}y^{-n}$. The duals $\hat{\sigma}$ and $\hat{\tau}$ correspond to multiplication by x and y on $\mathbb{F}_2[x^{\pm 1},\ y^{\pm 1}]$. The subgroup G corresponds to the set of those power series $\sum_{m,n=-\infty}^{\infty} v_{mn}x^{-m}y^{-n}$ for which $(1 + x + y) \cdot \sum_{m,n=-\infty}^{\infty} v_{mn}x^{-m}y^{-n} = 0$. The annihilator of G corresponds to the set of all polynomials divisible by $1 + x + y$. Hence $\hat{G}$ is the quotient $\mathbb{F}_2[x^{\pm 1},\ y^{\pm 1}]/\langle 1 + x + y\rangle$, which may be identified with the ring of all rational functions over $\mathbb{F}_2$ whose denominator is of the form $x^k(1 + x)^l$.

Let us briefly mention that there is another way of viewing the system, which may be more convenient for certain purposes. Take $\mathbb{F}_2^{\mathbb{Z}_+}$ as the underlying group. Let σ' be the (one-sided) shift, and let $\tau' = I + \sigma'$, where I is the identity map. The dual group now is $\mathbb{F}_2[x]$, and the dual actions are again multiplications by x and by $1 + x$. This system is basically equivalent to Ledrappier's system. (More accurately, the transformations σ' and τ' are non-invertible, and one should pass to the natural extension to obtain exactly the same system.)

3. *The main result*

As mentioned above, our main result in this paper says, roughly speaking, that the only obstacle to mixing of all orders in Ledrappier's system is what happens along powers of 2. We shall proceed to state this result in a precise form in the general case.

In view of Theorem 1.1, the following definition is natural.

Definition 3.1. Let $(X, \mathcal{B}, \mu, (T^{\mathbf{n}})_{\mathbf{n}\in\mathbb{Z}^2})$ be a measure-preserving system, $r \geq 2$ an integer and $\mathcal{M} \subseteq (\mathbb{Z}^2)^r$. The system is *$r$-mixing modulo $\mathcal{M}$* if for any $A_1, A_2, \ldots, A_r \in \mathcal{B}$ one has

$$\lim \mu\biggl(\bigcap_{i=1}^{r} T^{\mathbf{n}_i} A_i\biggr) = \prod_{i=1}^{r} \mu(A_i)$$

as $\rho((\mathbf{n}_1, \ldots, \mathbf{n}_r), \mathcal{M}) \to \infty$ and $\mathbf{n}_i - \mathbf{n}_j \to \infty$ for $i \neq j$.

Obviously, the smaller the set $\mathcal{M}$ is, the stronger is the assertion that a system is mixing modulo $\mathcal{M}$. Also, if $\mathcal{M}_1$ and $\mathcal{M}_2$ are two sets of r-tuples, such that the distance from every point of $\mathcal{M}_1$ to the set $\mathcal{M}_2$ is at most C for some constant C, then mixing modulo $\mathcal{M}_2$ implies mixing modulo $\mathcal{M}_1$. In particular, if the Hausdorff distance between $\mathcal{M}_1$ and $\mathcal{M}_2$ is finite (which means that there exists a constant C such that every point of $\mathcal{M}_1$ is at a distance at most C from $\mathcal{M}_2$ and *vice versa*), then mixing modulo $\mathcal{M}_2$ coincides with mixing modulo $\mathcal{M}_1$.

With this terminology, Theorem 1.1 is equivalent to the assertion that Ledrappier's system is 3-mixing modulo $\mathcal{L}$†.

Definition 3.1 may be easily modified for $\mathbb{Z}^d$-actions for any d. Note that a $\mathbb{Z}^d$-action $(T^{\mathbf{n}})_{\mathbf{n}\in\mathbb{Z}^d}$ on a probability space $(X, \mathcal{B}, \mu)$ is strongly mixing, i.e. it satisfies the condition

$$\mu(A \cap T^{\mathbf{n}}B) \underset{\mathbf{n}\to\infty}{\longrightarrow} \mu(A)\mu(B), \qquad A, B \in \mathcal{B},$$

if and only if it is 2-mixing modulo the empty set. Similarly, one can check that a $\mathbb{Z}^d$-action $(T^{\mathbf{n}})_{\mathbf{n}\in\mathbb{Z}^d}$ is weakly mixing, i.e. satisfies the condition

$$\mu(A \cap T^{\mathbf{n}}B) \underset{\substack{\mathbf{n}\to\infty \\ \mathbf{n}\notin\mathcal{M}_0}}{\longrightarrow} \mu(A)\mu(B), \qquad A, B \in \mathcal{B},$$

for some $\mathcal{M}_0 \subseteq \mathbb{Z}^d$ of density 0, if and only if there exists a set $\mathcal{M} \subseteq (\mathbb{Z}^d)^2$ of density 0 such that the action is 2-mixing modulo $\mathcal{M}$. It is worth mentioning that, in general, the exceptional set $\mathcal{M}$ distinguishing weak mixing from strong mixing may be not too small, in the following sense. Given any positive sequence (c_n) satisfying $c_n = o(n^d)$, one can show [**14**] that there exists a weakly mixing system $(X, \mathcal{B}, \mu, (T^{\mathbf{n}})_{\mathbf{n}\in\mathbb{Z}^d})$, sets $A, B \in \mathcal{B}$ and $\varepsilon > 0$ such that

$$|\{1 \leq \|\mathbf{n}\| \leq N : |\mu(A \cap T^{\mathbf{n}}B) - \mu(A)\mu(B)| > \varepsilon\}| > c_N$$

for all sufficiently large N.

Our objective in this paper is to find, for any r, necessary and sufficient conditions on a set of r-tuples $\mathcal{M}_r$, so that Ledrappier's system will be r-mixing modulo $\mathcal{M}_r$. Moreover, we would like these conditions to be as explicit as possible. As we shall see, Ledrappier's system is r-mixing modulo $\mathcal{M}_r$ for rather small sets $\mathcal{M}_r$.

† Note that we could avoid the condition that the distance between $\mathbf{n}_i$ and $\mathbf{n}_j$ tends to infinity for $i \neq j$ by adjoining to $\mathcal{M}$ the set

$$\bigcup_{1\leq i<j\leq r} \{(\mathbf{n}_1, \mathbf{n}_2, \ldots, \mathbf{n}_r) \in (\mathbb{Z}^2)^r : \mathbf{n}_i = \mathbf{n}_j\}.$$

However, since in any non-trivial system the convergence in question (for all r-tuples of measurable sets) implies that the sequences $\mathbf{n}_i$ and $\mathbf{n}_j$ grow apart, it seems more natural to put the condition as part of the definition.

Definition 3.2. A finite r-element set $\{(a_1, b_1), (a_2, b_2), \ldots, (a_r, b_r)\}$ in $\mathbb{Z}^2$ is a *special r-gon* if

$$x^{a_1}(1+x)^{b_1} + x^{a_2}(1+x)^{b_2} + \cdots + x^{a_r}(1+x)^{b_r} = 0.$$

The set of all special r-gons will be denoted by $\mathcal{L}_r$.

Denote by Λ_r the set of all r-element sets in $\mathbb{Z}^2$, containing a special s-gon for some $s \leq r$.

The following theorem is the main result of this paper.

THEOREM 3.3. *For every $r \geq 3$, Ledrappier's system is r-mixing modulo Λ_r.*

As we shall see in the following, the set Λ_r is rather small, which justifies the title of this paper. (See §§4, 7 and 8.)

4. *An explicit form of Theorem 3.3 for $r = 4$*

Theorem 1.1 is in principle a special case of Theorem 3.3, but uses the additional knowledge as to how special triangles look like (see Lemma 5.6). To obtain an 'explicit' form of Theorem 3.3 for a specific r, one needs to know explicitly the family $\mathcal{L}_r$, as well as all families $\mathcal{L}_s$ for $s < r$. Unfortunately, these families tend to become quite cumbersome as r increases. We shall now give an explicit description of $\mathcal{L}_4$. Given two polynomials

$$P_1(x, y) = \sum_{i=1}^{r} x^{a_{i1}} y^{b_{i1}}, \quad P_2(x, y) = \sum_{i=1}^{r} x^{a_{i2}} y^{b_{i2}},$$

we will say that the polynomial $P_2(x, y)$ is *obtained from $P_1(x, y)$ by an (a, b)-translation and 2^k-dilation* if

$$P_2(x, y) = x^a y^b P_1(x, y)^{2^k},$$

namely if

$$a_{i2} = a + 2^k a_{i1}, \quad b_{i2} = b + 2^k b_{i1}, \quad 1 \leq i \leq r,$$

for some integers a, b and $k \geq 0$. Analogous terminology will be used for r-gons. Now consider the following families of quadrangles (of which the first consists of a single quadrangle):

$$\begin{aligned}
\mathcal{Q}_1 &= \{\{(0,0), (0,3), (3,0), (1,1)\}\},\\
\mathcal{Q}_2 &= \{\{(0,0), (0,2^k+1), (2^k,0), (1,2^k)\} : k \geq 0\},\\
\mathcal{Q}_3 &= \{\{(0,0), (0,2^k), (2^k+1,0), (2^k,1)\} : k \geq 0\},\\
\mathcal{Q}_4 &= \{\{(0,0), (0,2^k), (2^k-1,0), (2^k-1,1)\} : k \geq 1\},\\
\mathcal{Q}_5 &= \{\{(0,0), (0,2^k-1), (2^k,0), (1,2^k-1)\} : k \geq 1\},\\
\mathcal{Q}_6 &= \{\{(0,0), (0,2^k+1), (1,0), (2^k,1)\} : k \geq 0\},\\
\mathcal{Q}_7 &= \{\{(0,0), (0,1), (2^k+1,0), (1,2^k)\} : k \geq 0\},\\
\mathcal{Q}_8 &= \{\{(0,1), (0,2^k), (1,0), (2^k,0)\} : k \geq 1\},\\
\mathcal{Q}_9 &= \{\{(0,2^k+1), (0,2^k), (1,0), (2^k+1,0)\} : k \geq 0\},\\
\mathcal{Q}_{10} &= \{\{(2^k+1,0), (2^k,0), (0,1), (0,2^k+1)\} : k \geq 0\}.
\end{aligned}$$

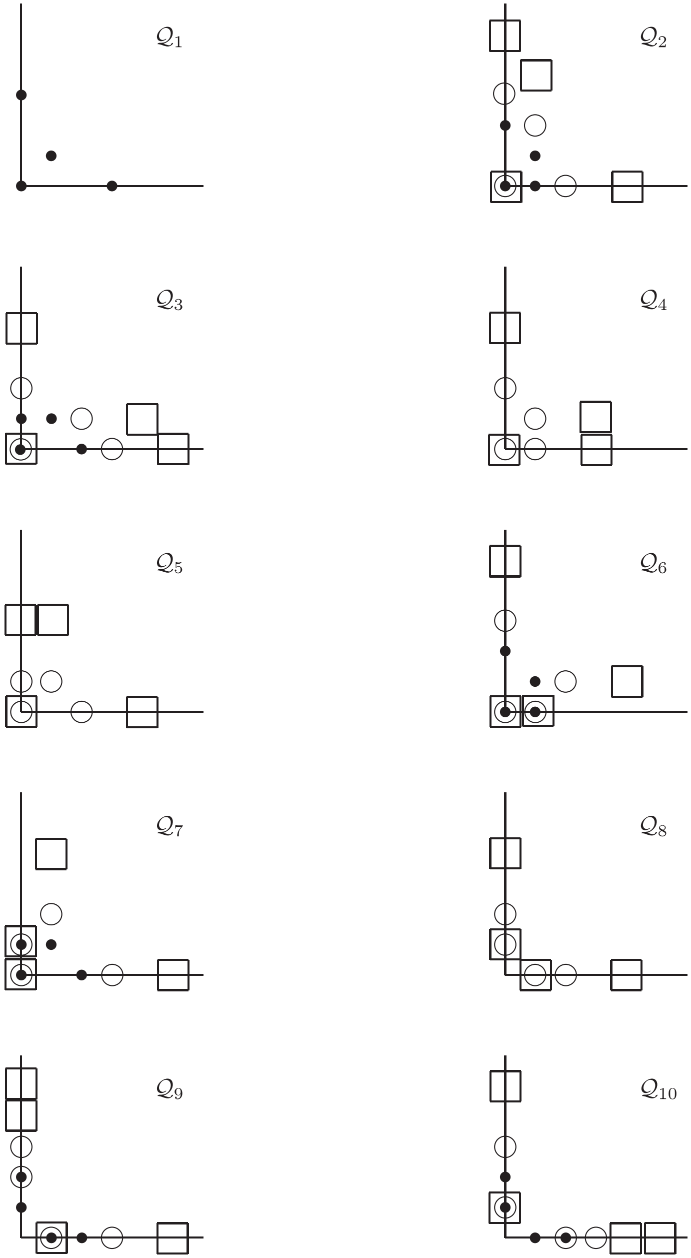

FIGURE 1. Some quadrangles in $\mathcal{L}_4$.

See Figure 1, where the quadrangles for some selected values of k are shown. A bold dot corresponds to $k = 0$ (where relevant), a circle to $k = 1$, and a square to $k = 2$.

PROPOSITION 4.1. *$\mathcal{L}_4$ is the family of all quadrangles obtained from some quadrangle in $\bigcup_{i=1}^{10} \mathcal{Q}_i$ by an (a, b)-translation and a 2^k-dilation for some (a, b) and k.*

The proof will be provided in the next section.

5. *Auxiliary algebraic results*

PROPOSITION 5.1. *Suppose*

$$\sum_{i=1}^{r} P_i(x)x^{a_i}(1+x)^{b_i} = 0 \tag{1}$$

for some $P_1(x), P_2(x), \dots, P_r(x) \in \mathbb{F}_2[x]$, *not all* 0, *of degrees not exceeding* 2^s, *and integers* $a_1, b_1, a_2, b_2, \dots, a_r, b_r$. *Then there exist constants* $c_1, c_2, \dots, c_r \in \mathbb{F}_2$, *not all* 0, *and* $\alpha_i, \beta_i \in \{0, 1\}$ *such that*

$$\sum_{i=1}^{r} c_i x^{[a_i/2^s]+\alpha_i}(1+x)^{[b_i/2^s]+\beta_i} = 0. \tag{2}$$

Proof. The proof is by induction on s. It will be convenient to prove slightly more than required. Namely, we shall prove in addition that, if $\deg P_i(x) < 2^s$ and $a_i \equiv b_i \equiv 0 \pmod{2^s}$ for some i, then $\alpha_i = \beta_i = 0$.

For $s = 0$, all the P_i are either constant or linear. Those which are constant are already as required. If $P_i(x) = x$, take $c_i = 1$ and increase a_i by 1 (that is, $\alpha_i = 1$). Similarly, if $P_i(x) = 1 + x$, take $c_i = 1$ and increase b_i by 1. Obviously, these changes bring the equation to the required form. Moreover, if $\deg P_i(x) < 2^0 = 1$ for some i, then the corresponding term is already of the required form, so that $\alpha_i = \beta_i = 0$.

Suppose the proposition has been established when all polynomials $P_i(x)$ are of degree not exceeding 2^{s-1}, and suppose now that (1) holds, with all degrees not exceeding 2^s. Let $P_{i1}(x)$ be the polynomial obtained from $P_i(x)$ upon multiplying it by x if a_i is odd and by $1 + x$ if b_i is odd. (Note that $P_{i1}(x) = x(1 + x)P_i(x)$ in case both a_i and b_i are odd.) Write (1) in the form

$$\sum_{i=1}^{r} P_{i1}(x)(x^{[a_i/2]}(1+x)^{[b_i/2]})^2 = 0. \tag{3}$$

If not all polynomials $P_{i1}(x)$, $1 \le i \le r$, are squares, then differentiate both sides to obtain

$$\sum_{i=1}^{r} P'_{i1}(x)(x^{[a_i/2]}(1+x)^{[b_i/2]})^2 = 0, \tag{4}$$

where not all derivatives $P'_{i1}(x)$ vanish. Since the derivative of any polynomial over $\mathbb{F}_2$ is a square, we may rewrite (4) as

$$\sum_{i=1}^{r} P^2_{i2}(x)(x^{[a_i/2]}(1+x)^{[b_i/2]})^2 = 0, \tag{5}$$

where $P^2_{i2}(x) = P'_{i1}(x)$, and thus

$$\sum_{i=1}^{r} P_{i2}(x)x^{[a_i/2]}(1+x)^{[b_i/2]} = 0. \tag{6}$$

We have

$$\deg P_{i2}(x) = \frac{\deg P'_{i1}(x)}{2} \le \frac{\deg P_{i1}(x) - 1}{2} \le \frac{\deg P_i(x) + 1}{2} \le \frac{2^s + 1}{2},$$

and therefore deg $P_{i2}(x) \le 2^{s-1}$. Employing the induction hypothesis on (6) we arrive at an equality of the form

$$\sum_{i=1}^{r} c_i x^{[[a_i/2]/2^{s-1}]+\alpha_i}(1+x)^{[[b_i/2]/2^{s-1}]+\beta_i} = 0. \tag{7}$$

Since $[[m/2]/2^{s-1}] = [m/2^s]$ for any integer m, this is in fact an equality as required. Note that, if deg $P_i(x) < 2^s$ for some i, and $a_i \equiv b_i \equiv 0 \pmod{2^s}$, then $P_{i1}(x) = P_i(x)$ so that deg $P_{i2}(x) \le 2^{s-1} - 1$ and $[a_i/2] = a_i/2 \equiv 0 \pmod{2^{s-1}}$, $[b_i/2] = b_i/2 \equiv 0 \pmod{2^{s-1}}$. The induction hypothesis guarantees that in this case we shall have $\alpha_i = \beta_i = 0$.

Now assume all polynomials $P_{i1}(x)$, $1 \le i \le r$, are squares. Pass from (3) directly to (5), and then to (6), where this time $P_{i2}^2(x) = P_{i1}(x)$. Now

$$\deg P_{i2}(x) = \frac{\deg P_{i1}(x)}{2} \le \frac{\deg P_i(x) + 2}{2} \le \frac{2^s + 2}{2} = 2^{s-1} + 1. \tag{8}$$

If we could bound the degrees of the $P_{i2}(x)$ from above by 2^{s-1}, the induction hypothesis could be applied as before. Going over the chain of inequalities in (8), we see that the left-hand side equals the right-hand side if and only if deg $P_i(x) = 2^s$ and both a_i and b_i are odd. Denote by I_0 the set of those indices i for which all three conditions are satisfied. According to our assumption, if $i \in I_0$ then $P_{i2}^2(x) = P_{i1}(x) = x(1+x)P_i(x)$, and therefore $P_{i2}(x)$ is divisible by $x(1+x)$. Put

$$P_{i3}(x) = \begin{cases} \dfrac{P_{i2}(x)}{x}, & i \in I_0,\ a_i \not\equiv 2^s - 1 \pmod{2^s}, \\ \dfrac{P_{i2}(x)}{1+x}, & i \in I_0,\ a_i \equiv 2^s - 1 \pmod{2^s},\ b_i \not\equiv 2^s - 1 \pmod{2^s}, \\ \dfrac{P_{i2}(x)}{x(1+x)}, & i \in I_0,\ a_i \equiv b_i \equiv 2^s - 1 \pmod{2^s}, \\ P_{i2}(x), & i \notin I_0. \end{cases} \tag{9}$$

Reordering the terms in (2), we may assume that the first condition in (9) is satisfied for $1 \le i \le r_1$, the second for $r_1 + 1 \le i \le r_2$, and so forth. Rewrite (6) as

$$\sum_{i=1}^{r_1} P_{i3}(x)x^{[a_i/2]+1}(1+x)^{[b_i/2]} + \sum_{i=r_1+1}^{r_2} P_{i3}(x)x^{[a_i/2]}(1+x)^{[b_i/2]+1}$$
$$+ \sum_{i=r_2+1}^{r_3} P_{i3}(x)x^{[a_i/2]+1}(1+x)^{[b_i/2]+1} + \sum_{i=r_3+1}^{r} P_{i3}(x)x^{[a_i/2]}(1+x)^{[b_i/2]} = 0. \tag{10}$$

Apply the induction hypothesis to (10). We obtain an equality of the required form, except that the resulting α_i and β_i may seem to be possibly 2 instead of either 0 or 1. To this end, we note the following.

(1) For $1 \le i \le r_1$ we have $[a_i/2] \not\equiv 2^{s-1} - 1 \pmod{2^{s-1}}$, and therefore $[([a_i/2] + 1)/2^{s-1}] = [a_i/2^s]$.

(2) For $r_1 + 1 \le i \le r_2$ we analogously have $[([b_i/2] + 1)/2^{s-1}] = [b_i/2^s]$.

(3) For $r_2 + 1 \le i \le r_3$ we are in the special situation where $\deg P_{i3}(x) = 2^{s-1}$ and $[a_i/2] + 1 \equiv [b_i/2] + 1 \equiv 0 \pmod{2^{s-1}}$, for which the induction hypothesis ensures that the exponents of x and $1 + x$ in the resulting equality will be $[([a_i/2] + 1)/2^{s-1}] = [a_i/2^s] + 1$ and $[([b_i/2] + 1)/2^{s-1}] = [b_i/2^s] + 1$, respectively.

(4) For $r_3 + 1 \le i \le r$ we clearly obtain in the reduced equality terms as required. Moreover, those terms in (10) which arose from terms in (1) with $\deg P_i(x) < 2^s$ and $a_i \equiv b_i \equiv 0 \pmod{2^s}$ give rise to terms of the form $P_{i3}(x)x^{a_i/2}(1 + x)^{b_i/2}$, where $\deg P_{i3}(x) < 2^{s-1}$, whence the induction hypothesis ensures that $\alpha_i = \beta_i = 0$.

This completes the proof. □

Definition 5.2. A polynomial $P(x, y) \in \mathbb{F}_2[x^{\pm 1}, y^{\pm 1}]$ is an *L-polynomial* (L for Ledrappier) if $P(x, 1 + x) = 0$†.

In other words, L-polynomials are those polynomials belonging to the ideal $\langle 1 + x + y\rangle$. (In fact, if $P(x, y) \in \langle 1 + x + y\rangle$, then $P(x, y) = (1 + x + y)Q(x, y)$ for some polynomial $Q(x, y)$, so that $P(x, 1 + x) = (1 + x + 1 + x)Q(x, 1 + x) = 0$. On the other hand, if $P(x, 1 + x) = 0$, then $1 + x$ is a root of the polynomial $R(y) = P(x, y) \in \mathbb{F}_2(x)[y]$, and therefore $P(x, y)$ is divisible by $y - (1 + x) = 1 + x + y$.)

Definition 5.3. A polynomial $P(x, y) = \sum_{i=1}^{r} x^{a_i} y^{b_i} \in \mathbb{F}_2[x^{\pm 1}, y^{\pm 1}]$ is in *reduced form* if $(a_i, b_i) \ne (a_j, b_j)$ for $i \ne j$. The *length* of a polynomial is the number of monomials r in its reduced form.

In view of Proposition 5.1, the main thing we need to do is characterize L-polynomials. This is relatively easy for ‘short’ polynomials.

LEMMA 5.4. *There are no L-polynomials of length* 1 *or* 2.

Definition 5.5. A polynomial $T(x, y) \in \mathbb{F}_2[x^{\pm 1}, y^{\pm 1}]$ is *triangular* if

$$T(x, y) = x^a y^b + x^{a+2^k} y^b + x^a y^{b+2^k}$$

for some integers a, b and $k \ge 0$.

LEMMA 5.6. *A polynomial of length* 3 *is an L-polynomial if and only if it is triangular.*

Proof. The ‘if’ direction is immediate. For the inverse direction, let $P(x, y) = \sum_{i=1}^{3} x^{a_i} y^{b_i}$ be an L-polynomial. Consider the three pairs (a_i, b_i) modulo $(2, 2)$. We distinguish between three cases according to the distribution of the pairs among the residue classes. It will be convenient to deal with the cases in the following order.

Case 1. Two of the pairs belong to the same residue class, while the third pair belongs to another.

Multiplying by some $x^\alpha y^\beta$ we may assume the class containing the two pairs to be $(0, 0)$. Differentiating both sides of the identity

† Thus, a finite set $((a_1, b_1), (a_2, b_2), \ldots, (a_r, b_r))$ is a special r-gon if and only if the polynomial $x^{a_1} y^{b_1} + x^{a_2} y^{b_2} + \cdots + x^{a_r} y^{b_r}$ is an L-polynomial.

$$\sum_{i=1}^{3} x^{a_i}(1+x)^{b_i} = 0,$$

we are left with a single monomial on the left-hand side, which is a contradiction.

Case 2. The pairs belong to three distinct classes.

Without loss of generality, we may assume that $(a_1, b_1) \equiv (0, 0) \pmod{(2, 2)}$, $(a_2, b_2) \equiv (1, 0) \pmod{(2, 2)}$ and $(a_3, b_3) \equiv (0, 1) \pmod{(2, 2)}$. Differentiate to obtain

$$x^{a_2-1}(1+x)^{b_2} + x^{a_3}(1+x)^{b_3-1} = 0.$$

This implies $a_2 = a_3 + 1$ and $b_2 = b_3 - 1$. Consequently

$$x^{a_2}(1+x)^{b_2} + x^{a_3}(1+x)^{b_3} = x^{a_3}(1+x)^{b_2}(x + (1+x)) = x^{a_3}(1+x)^{b_2},$$

so that $a_1 = a_3$ and $b_1 = b_2$. Hence $P(x, y)$ is triangular.

Case 3. All pairs lie in the same residue class.

It suffices to prove our claim for the case where $a_i, b_i \geq 0$ for each i. We do it by induction on the total degree of $P(x, y)$. For degree 0 the claim is trivial, since there exist no such polynomials of length 3. Suppose the proposition holds for polynomials of degree not exceeding $d - 1$, and let $P(x, y)$ be of degree d. If all the a_i are odd, then the polynomial $P(x, y)/x$ is still an L-polynomial with all pairs of exponents in the same residue class. By the induction hypothesis, $P(x, y)/x$ is triangular, and therefore so is $P(x, y)$ itself. Similarly, we may assume all the b_i to be even. Since

$$\sum_{i=1}^{3} x^{a_i}(1+x)^{b_i} = \left(\sum_{i=1}^{3} x^{a_i/2}(1+x)^{b_i/2}\right)^2,$$

the polynomial $\sum_{i=1}^{3} x^{a_i/2} y^{b_i/2}$ is also an L-polynomial, and it is of degree smaller than that of $P(x, y)$. If all its coefficients lie in the same residue class modulo $(2, 2)$, then by the induction hypothesis the polynomial is triangular, and hence so is $P(x, y)$. In the other case, it satisfies either the conditions of Case 1 or those of Case 2, and again we are done. □

Remark 5.7. Actually, we could have given a simpler proof, as follows. Multiplying by some $x^\alpha y^\beta$ we may assume all the a_i and b_i to be non-negative with at least one of the a_i and at least one of the b_i being 0. The two substitutions $x = 0$ (and $y = 1$) and $x = 1$ (and $y = 0$) show that at least two of the a_i and at least two of the b_i are 0. Thus $P(x, y) = 1 + x^a + y^b$. The equality $(1 + x)^b = 1 + x^a$ now gives $b = a = 2^s$. However, the proof we have given is more in line with the techniques we employ in the paper, and it is instructive to have prior to the characterization of L-polynomials of length 4.

The characterization of L-polynomials of length 4 is essentially the contents of Proposition 4.1, which will now be proved.

Proof of Proposition 4.1. Let $P(x, y) = \sum_{i=1}^{4} x^{a_i} y^{b_i}$ be an L-polynomial. Similarly to the proof of the preceding lemma, two of the a_i may be assumed to vanish, and so may two of the b_i. Thus, there are two cases to consider.

Case 1. $1+(1+x)^{b_2}+x^{a_3}+x^{a_4}(1+x)^{b_4}=0$.

Consider the four integers b_2, a_3, a_4, b_4. We proceed by looking at how many of them are even and how many odd.

Subcase 1(i). All four numbers are odd.

Differentiating both sides of the equality we obtain

$$(1+x)^{b_2-1}+x^{a_3-1}+x^{a_4-1}(1+x)^{b_4-1}=0.$$

According to Lemmas 5.4 and 5.6, this means that we have here some rearrangement of the equality

$$1+x^{2^k}+(1+x)^{2^k}=0.$$

One possibility is having $b_2-1=a_3-1=2^k$ and $a_4-1=b_4-1=0$. Substituting in the original equality, we obtain

$$1+(1+x)^{2^k+1}+x^{2^k+1}+x(1+x)=0,$$

which is easily seen to yield $k=1$, so that $b_2=a_3=3$ and $a_4=b_4=1$. This yields the quadrangle $\mathcal{Q}_1$. If either $a_4-1=2^k$ or $b_4-1=2^k$, then the last summand on the left-hand side of our equality is a polynomial of degree 2^k+2, while the first three are of lower degree, and therefore we do not get a solution.

Subcase 1(ii). Exactly one of the numbers is odd.

Since three of the numbers are even, three of the terms on the right-hand side of our equality are squares, and so is their sum, whereas the fourth cannot be a square. Thus this case is impossible.

Subcase 1(iii). Exactly two of the numbers are odd.

The reasoning in the preceding case shows that not both b_2 and a_3 are even. If b_2 and a_3 are both odd, then a differentiation takes us to the case $r=2$, which shows that $b_2=a_3=1$, which does not lead to a solution. Suppose b_2 and exactly one of a_4 and b_4 are odd. If b_4 is the odd one, then a differentiation again leads to an L-polynomial of length 2, implying that $a_4=0$. This case is the intersection of Case 1 with Case 2, and will be considered within the framework of Case 2 later. If a_4 is odd, then differentiate

$$(1+x)^{b_2-1}+x^{a_4-1}(1+x)^{b_4}=0,$$

to obtain $a_4=1$ and $b_2-1=b_4$. The original equality reduces then to

$$1+(1+x)^{b_4}+x^{a_3}=0,$$

which by Lemma 5.6 gives $a_3=b_4=2^k$ for some k. Altogether we have $b_2=2^k+1$, $a_3=b_4=2^k$, $a_4=1$, which yields the quadrangle $\mathcal{Q}_2$. The case where a_3 and exactly one of a_4 and b_4 are odd may be transformed to the preceding case by replacing x by $1+x$, and this gives therefore the solution $b_2=a_4=2^k$, $a_3=2^k+1$, $b_4=1$, namely $\mathcal{Q}_3$.

Subcase 1(iv). Exactly three of the numbers are odd.

If the even number is b_2, then differentiation shows that $a_3 = a_4, b_4 = 1$, and substituting in the original equality we obtain $b_2 = 2^k$, $a_3 = 2^k - 1$. Thus $b_2 = 2^k$, $a_3 = a_4 = 2^k - 1$, $b_4 = 1$, that is the quadrangle $\mathcal{Q}_4$. If the even number is a_3, then by symmetry $b_2 = b_4 = 2^k - 1$, $a_3 = 2^k$, $a_4 = 1$, namely $\mathcal{Q}_5$. If the even number is a_4, then by differentiation

$$(1+x)^{b_2-1} + x^{a_3-1} + x^{a_4}(1+x)^{b_4-1} = 0.$$

According to Lemma 5.6, this yields three possible solutions. One of these is $b_2 = 2^k + 1$, $a_3 = b_4 = 1$, $a_4 = 2^k$, leading to $\mathcal{Q}_6$. The other two are $b_2 = 1$, $a_3 = 2^k + 1$, $a_4 = 0$, $b_4 = 2^k + 1$, and $b_2 = a_3 = 2^k + 1$, $a_4 = 0$, $b_4 = 1$. Both of these solutions correspond to Case 2 also, and will be treated there. If the even number is b_4, then by symmetry we arrive at the solution $b_2 = a_4 = 1$, $a_3 = 2^k + 1$, $b_4 = 2^k$, leading to $\mathcal{Q}_7$.

Subcase 1(v). None of the numbers are odd.

Divide all four numbers by the largest possible power of 2, thus reverting to one of the former cases.

Case 2. $(1+x)^{b_1} + (1+x)^{b_2} + x^{a_3} + x^{a_4} = 0$.

Similarly to Case 1, we separate to subcases according to the number of even and odd numbers among the integers b_1, b_2, a_3, a_4.

Subcase 2(i). All four numbers are odd.

Multiply both sides by x and differentiate to obtain

$$(1+x)^{b_1-1} + (1+x)^{b_2-1} = 0,$$

which we know from Lemma 5.4 to be impossible.

Subcase 2(ii). Exactly one of the numbers is odd.

Exactly as in Subcase 1(ii), this yields no solutions.

Subcase 2(iii). Exactly two of the numbers are odd.

A differentiation gives an equality involving only two terms. By Lemma 5.4 the derivative must then vanish trivially, which means that one of the a_i and one of the b_i are 1, say $b_1 = a_3 = 1$. This yields

$$1 + (1+x)^{b_2} + x^{a_4} = 0,$$

and Lemma 5.6 implies $b_2 = a_4 = 2^k$ for some k, which gives $\mathcal{Q}_8$.

Subcase 2(iv). Exactly three of the numbers are odd.

After differentiation we revert to the case of Lemma 5.6. If the even exponent is one of the b_i, say b_2, then this implies $a_3 - 1 = 0$ and $b_1 - 1 = a_4 - 1 = 2^k$ for some k, namely $a_3 = 1$, $b_1 = a_4 = 2^k + 1$. Then

$$(1+x)^{b_2} = (1+x)^{2^k+1} + x + x^{2^k+1} = (1+x)^{2^k},$$

which gives $b_2 = 2^k$. This yields the quadrangle $\mathcal{Q}_9$. If the even exponent is one of the a_i then we obtain symmetrically $\mathcal{Q}_{10}$.

Subcase 2(v). None of the numbers are odd.

Similarly to Case 1 we may divide them all by the largest possible power of 2, which brings us back to one of the former subcases. □

6. *Proof of Theorem 3.3*

The proof of Theorem 3.3, to be presented in this section, hinges on Lemma 6.2 below. We will find it convenient to view in this section the condition of mixing modulo $\mathcal{M}$ somewhat differently. Namely, this condition may be rephrased in terms of the sequences of r-tuples of pairs of integers for which the intersections we consider converge to the correct limit. Thus, we start by defining the following notion.

Definition 6.1. A sequence $((a_1^{(n)}, b_1^{(n)}), (a_2^{(n)}, b_2^{(n)}), \dots, (a_r^{(n)}, b_r^{(n)}))_{n=1}^{\infty}$ of r-tuples of pairs of integers, with $(a_i^{(n)}, b_i^{(n)})$ and $(a_j^{(n)}, b_j^{(n)})$ growing apart for $i \neq j$, is *mixing* if

$$\int \prod_{i=1}^{r} \sigma^{a_i^{(n)}} \tau^{b_i^{(n)}} f_i \, d\mu \underset{n\to\infty}{\longrightarrow} \prod_{i=1}^{r} \int f_i \, d\mu$$

for every $f_1, f_2, \dots, f_r \in L^{\infty}$.

Note that a sequence $((a_1^{(n)}, b_1^{(n)}), (a_2^{(n)}, b_2^{(n)}), \dots, (a_r^{(n)}, b_r^{(n)}))_{n=1}^{\infty}$ is mixing if and only if any bounded perturbation thereof is such. Clearly, if the system is mixing modulo $\mathcal{M}$, then every sequence $((a_1^{(n)}, b_1^{(n)}), (a_2^{(n)}, b_2^{(n)}), \dots, (a_r^{(n)}, b_r^{(n)}))_{n=1}^{\infty}$, with $(a_i^{(n)}, b_i^{(n)})$ and $(a_j^{(n)}, b_j^{(n)})$ growing apart for $i \neq j$, and the distance of whose elements from $\mathcal{M}$ tends to infinity, is mixing according to this definition. On the other hand, once we characterize those sequences which are mixing according to this definition, we have actually characterized those sets $\mathcal{M}$ for which the system is mixing modulo $\mathcal{M}$.

For systems consisting of compact abelian groups and endomorphisms thereof, it is usually most convenient to test mixing properties by studying corresponding properties of the dual action. For Ledrappier's system, the condition is given by the following lemma, which we present without proof. (See [**8**, p. 263] for full details.)

LEMMA 6.2. *A sequence* $((a_1^{(n)}, b_1^{(n)}), (a_2^{(n)}, b_2^{(n)}), \dots, (a_r^{(n)}, b_r^{(n)}))_{n=1}^{\infty}$ *is mixing if and only if for any polynomials* $P_1(x), P_2(x), \dots, P_r(x)$, *not all* 0, *the equation*

$$P_1(x)x^{a_1^{(n)}}(1+x)^{b_1^{(n)}} + P_2(x)x^{a_2^{(n)}}(1+x)^{b_2^{(n)}} + \cdots + P_r(x)x^{a_r^{(n)}}(1+x)^{b_r^{(n)}} = 0 \quad (11)$$

has only a finite number of solutions n.

Example 6.3. The sequence $((0, 0), (2^n, 0), (0, 2^n))_{n=1}^{\infty}$ is not mixing since

$$1 \cdot x^0(1+x)^0 + 1 \cdot x^{2^n}(1+x)^0 + 1 \cdot x^0(1+x)^{2^n} = 0$$

for each n.

Proof of Theorem 3.3. Let $((a_1^{(n)}, b_1^{(n)}), (a_2^{(n)}, b_2^{(n)}), \dots, (a_r^{(n)}, b_r^{(n)}))_{n=1}^{\infty}$ be a non-mixing sequence, such that $(a_i^{(n)}, b_i^{(n)})$ and $(a_j^{(n)}, b_j^{(n)})$ grow apart for each $i \neq j$. We have to show that it contains a subsequence, consisting of elements which are at a bounded distance from Λ_r. Indeed, by Lemma 6.2 there exist polynomials $P_1(x), P_2(x), \dots, P_r(x)$, not all 0, such that the equation

$$P_1(x)x^{a_1^{(n)}}(1+x)^{b_1^{(n)}} + P_2(x)x^{a_2^{(n)}}(1+x)^{b_2^{(n)}} + \cdots + P_r(x)x^{a_r^{(n)}}(1+x)^{b_r^{(n)}} = 0$$

has infinitely many solutions n. Apply Proposition 5.1 to each of these n. Ignoring the other n, and passing to a subsequence, we obtain constants $c_1, c_2, \ldots, c_r \in \mathbb{F}_2$, not all 0, and $\alpha_{in}, \beta_{in} \in \{0, 1\}$ such that

$$\sum_{i=1}^{r} c_i x^{[a_i^{(n)}/2^s]+\alpha_{in}}(1+x)^{[b_i^{(n)}/2^s]+\beta_{in}} = 0. \tag{12}$$

Passing again to a subsequence, we may assume that each of the sequences α_{in} and β_{in} is constant and that each of the sequences $(a_i^{(n)})$ and $(b_i^{(n)})$ is constant modulo 2^s, say $a_i^{(n)} \equiv a_i \pmod{2^s}$ and $b_i^{(n)} \equiv b_i \pmod{2^s}$. Reordering the pairs $(a_i^{(n)}, b_i^{(n)})$, we may finally rewrite (12) in the form

$$\sum_{i=1}^{r'} x^{[a_i^{(n)}/2^s]+\alpha_i}(1+x)^{[b_i^{(n)}/2^s]+\beta_i} = 0$$

for some $1 \le r' \le r$. Raising this equality to the power 2^s, we find that

$$\sum_{i=1}^{r'} x^{a_i^{(n)}-a_i+2^s\alpha_i}(1+x)^{b_i^{(n)}-b_i+2^s\beta_i} = 0.$$

Thus, the r'-gon

$$((a_1^{(n)}, b_1^{(n)}), (a_2^{(n)}, b_2^{(n)}), \ldots, (a_{r'}^{(n)}, b_{r'}^{(n)}))$$

is obtained from a special r'-gon by a bounded perturbation, so that the r-gon

$$((a_1^{(n)}, b_1^{(n)}), (a_2^{(n)}, b_2^{(n)}), \ldots, (a_r^{(n)}, b_r^{(n)}))_{n=1}^{\infty}$$

stays close to Λ_r. This completes the proof. □

7. *The structure of $\mathcal{L}_r$*

We have seen in §4 that the only reason for a sequence of r-tuples of pairs to be non-mixing is its proximity to some special r-gons (or the proximity of some of its components to some special s-gons for a certain $3 \le s < r$). For $r = 3$, the only special triangles are those corresponding to triangular polynomials. For $r = 4$ we have seen that there are several families of special quadrangles. Most of these, in fact all those obtained from one of the r-gons of the families $\mathcal{Q}_2, \ldots, \mathcal{Q}_{10}$ by translation and dilation by a power of 2, correspond to sums of two triangular polynomials. For example, for a typical quadrangle in $\mathcal{Q}_2$,

$$1 + y^{2^k+1} + x^{2^k} + xy^{2^k} = (1 + x^{2^k} + y^{2^k}) + y^{2^k}(1 + x + y).$$

The quadrangle in $\mathcal{Q}_1$ corresponds in two ways to sums of four triangular polynomials,

$$\begin{aligned} 1 + x^3 + y^3 + xy &= (1 + x + y) + x(1 + x^2 + y^2) + y(1 + x^2 + y^2) + xy(1 + x + y) \\ &= (1 + x^4 + y^4) + x^3(1 + x + y) + y^3(1 + x + y) + xy(1 + x^2 + y^2), \end{aligned}$$

but it may be shown to correspond to no shorter sum of triangular polynomials. Our main result in this section asserts that any special r-gon corresponds to a sum of triangular polynomials, and the number of addends is bounded above by some constant depending only on r.

THEOREM 7.1. *Every special r-gon corresponds to a sum of at most r^3 triangular polynomials.*

It will be convenient to denote by $h(r)$ the minimal number m such that every special r-gon corresponds to a sum of at most m triangular polynomials. Thus, the theorem asserts that $h(r)$ is finite for every r and, moreover, $h(r) \le r^3$.

Proof. We proceed by induction on r. For $r = 3$ the theorem is a weak version of Lemma 5.6. Assume that the theorem holds for polynomials of length up to $r - 1$, and let $P(x, y) \in \mathbb{F}_2[x^{\pm 1}, y^{\pm 1}]$ be an L-polynomial of length r, say $P(x, y) = \sum_{i=1}^{r} x^{a_i} y^{b_i}$. Assume without loss of generality that $a_i, b_i \ge 0$ for $1 \le i \le r$. For $\alpha, \beta \in \{0, 1\}$ put

$$R_{\alpha\beta} = \{1 \le i \le r : (a_i, b_i) \equiv (\alpha, \beta) \pmod 2\}$$

and

$$P_{\alpha\beta}(x, y) = \sum_{i \in R_{\alpha\beta}} x^{a_i} y^{b_i},$$

so that

$$P(x, y) = \sum_{\alpha,\beta=0}^{1} P_{\alpha\beta}(x, y).$$

Let

$$r_{\alpha\beta} = |R_{\alpha\beta}|, \qquad \alpha, \beta \in \{0, 1\}.$$

We may assume that $r_{\alpha\beta} < r$ for each (α, β). Indeed, if $r_{\alpha\beta} = r$ for some (α, β), then dividing $P(x, y)$ by $x^{\alpha} y^{\beta}$ and taking a square root, we obtain an L-polynomial of length r. Continuing the process, we eventually obtain an L-polynomial of length r satisfying the extra condition. Expressing this polynomial as a sum of several triangular polynomials, we easily decompose $P(x, y)$ into a sum of as many triangular polynomials.

Replacing $P(x, y)$ by one of the polynomials $xP(x, y)$, $yP(x, y)$ or $xyP(x, y)$, if needed, we may assume that $r_{00} \ge r_{10}, r_{01}, r_{11}$. Next, we note that the polynomial $P(y^{-1}, xy^{-1})$ is also an L-polynomial, and by replacing $P(x, y)$ by it we leave the set R_{00} intact and permute the other three $R_{\alpha\beta}$ cyclically. Hence, using this transformation or its inverse, we may assume that $r_{01} \ge r_{10}, r_{11}$.

Now we introduce a few more polynomials, as follows:

$$\begin{aligned} V_1(x, y) &= x^{-1} y P_{10}(x, y) + P_{01}(x, y) + x^{-1} P_{11}(x, y), \\ V_2(x, y) &= P_{00}(x, y) + x^{-1} P_{10}(x, y) + x^{-1} y P_{11}(x, y), \\ V_3(x, y) &= (1 + x^{-1} + x^{-1} y) P_{10}(x, y) + (1 + x^{-1} + x^{-1} y) P_{11}(x, y). \end{aligned}$$

All three are L-polynomials. In fact

$$V_1(x, 1 + x) = (1 + x)\frac{d}{dx} P(x, 1 + x) = 0,$$

$$V_2(x, 1 + x) = \frac{d}{dx}((1 + x) P(x, 1 + x)) = 0,$$

and

$$V_3(x, y) = x^{-1}(1 + x + y)(P_{10}(x, y) + P_{11}(x, y)), \tag{13}$$

so that

$$V_3(x, 1 + x) = 0.$$

Note that the length of $V_1(x, y)$ is $r - r_{00}$, the length of $V_2(x, y)$ is $r - r_{01}$, and (13) shows that $V_3(x, y)$ is a sum of $r_{10} + r_{11}$ triangular polynomials. As $P(x, y) = \sum_{i=1}^{3} V_i(x, y)$, this shows that $P(x, y)$ is a sum of triangular polynomials.

The construction above yields

$$h(r) \leq \max(h(r - r_{00}) + h(r - r_{01}) + r_{10} + r_{11}),$$

where the maximum is taken over all polynomials of length r satisfying our assumptions. In view of the assumptions on the $r_{\alpha\beta}$ it follows that

$$\begin{aligned} h(r) \leq \max\{h(r - d) + h(r - c) + a + b : a \leq b \leq c \leq d \in \mathbb{Z}_+, \\ a + b + c + d = r, d < r\}. \end{aligned}$$

It remains to prove that, for non-negative integers $a \leq b \leq c \leq d < r$ with $a + b + c + d = r$, we have

$$(r - d)^3 + (r - c)^3 + a + b \leq r^3. \tag{14}$$

In fact, the constraints guarantee that $c \geq (r - d)/3$, and therefore

$$(r - d)^3 + (r - c)^3 + a + b \leq (r - d)^3 + \left(\frac{2r + d}{3}\right)^3 + r.$$

We have to show that, for $d \in [r/4, r - 1]$, the right-hand side is bounded above by r^3. Routine calculations show that, considered as a function of d in that interval, the right-hand side is decreasing from $r/4$ up to some point and increases after that. Thus we need to check only the values at the endpoints $r/4$ and $r - 1$, and it is easily seen that both values are bounded above by r^3. This completes the proof. □

Remark 7.2. The upper bound of r^3 in the theorem can be easily replaced by a somewhat smaller power of r, but our method does not yield a bound of r^2. It would be interesting to know how fast $h(r)$ grows as a function of r. The 'worst' example (in the sense that it seems not to be representable as a sum of a few triangular polynomials) we have so far is the family of polynomials

$$1 + xy + x^3 + y^3,$$
$$1 + xy + x^3y^3 + x^7 + y^7,$$

and in general

$$1 + xy + x^3y^3 + x^7y^7 + \cdots + x^{2^{r-1}-1}y^{2^{r-1}-1} + x^{2^r-1} + y^{2^r-1}.$$

Denote the last polynomial by U_r. One can easily verify that

$$U_{r+1} = U_r + x^{2^r-1}(1 + x^{2^r} + y^{2^r}) + y^{2^r-1}(1 + x^{2^r} + y^{2^r}) + x^{2^r-1}y^{2^r-1}(1 + x + y).$$

As $U_1 = 1 + x + y$ is triangular, this shows that U_r is a sum of $3r - 2$ triangular polynomials. Note that, as U_r is of length $r + 2$, this example yields an infinite family of polynomials such that the polynomial of length r is a sum of $3r - 8$ triangular polynomials.

QUESTION 7.3. *Can any U_r be expressed as a sum of less than $3r-2$ triangular polynomials?*

QUESTION 7.4. *Assuming that the answer to the preceding question is negative, are there even sharper examples?*

While we have not found an exact expression for $h(r)$, we are able to find, for each special r-gon, the minimal number m such that the r-gon corresponds to a sum of m triangular polynomials. For an r-gon R (and corresponding L-polynomial $P(x, y)$), we denote this number by $h(R)$ (or $h(P)$). The correspondence between special r-gons and L-polynomials enables us to use geometrical terminology for the latter. In particular, it will be convenient to use the *diameter* of an L-polynomial, meaning the diameter of the corresponding subset of $\mathbb{Z}^2$. Distances in $\mathbb{Z}^2$ will be calculated by the maximum metric. Denote $\Delta = 1 + x + y$.

PROPOSITION 7.5. *Given a special r-gon R, it is possible to calculate $h(R)$ effectively. Moreover, if the diameter of R is D, then it is effectively possible to represent the corresponding special polynomial P as a sum of $h(R)$ triangular polynomials*

$$P(x, y) = \sum_{t\in\mathcal{T}} x^{a_t} y^{b_t} \Delta^{2^{k_t}}, \tag{15}$$

where each triangle is of diameter at most $5(5r^3)^{r^3}D$, and its Hausdorff distance from R is at most $(5r^3)^{r^3+1}D$.

We first need a lemma.

LEMMA 7.6. *Let R and P be as in Proposition 7.5, and suppose P has a representation as in (15), with $|\mathcal{T}| = h(P)$. Suppose there exist integers k and l, satisfying $D < 2^l/5$ and $h(P) < 2^{l-k}/5$, such that for every $t \in \mathcal{T}$ we have either $k_t \le k$ or $k_t \ge l$. Then $P(x, y)$ has an alternative representation as a sum of triangular polynomials, in which the terms $x^{a_t} y^{b_t} \Delta^{2^{k_t}}$ with $k_t < k$ are replaced by terms of the form $x^{a_t'} y^{b_t'} \Delta^{2^{k_t}}$, while the terms $x^{a_t} y^{b_t} \Delta^{2^{k_t}}$ with $k_t > l$ are replaced by terms of the form $x^{a_t'} y^{b_t'} \Delta^{2^{k_t-l}}$.*

Proof. If $\mathcal{T}$ is of the smallest possible size, then $|\mathcal{T}| = h(P)$. We can define a graph on $\mathcal{T}$, where two triangles in $\mathcal{T}$ are adjacent if they share a common vertex. This graph will be denoted by $\mathcal{T}$ as well. Since the partial sum corresponding to any connected component of $\mathcal{T}$ is itself an L-polynomial, we shall assume throughout that $\mathcal{T}$ is connected.

Split the set $\mathcal{T}$ into two parts, depending on the size of the triangles. Namely, write $\mathcal{T} = \mathcal{T}_{\rm b} \cup \mathcal{T}_{\rm s}$, where $\mathcal{T}_{\rm b} = \{t : k_t \ge l\}$ and $\mathcal{T}_{\rm s} = \{t : k_t \le k\}$. Let $\mathcal{T}_{{\rm b}1}, \dots, \mathcal{T}_{{\rm b}m}$ be the connected components of $\mathcal{T}_{\rm b}$ and $\mathcal{T}_{{\rm s}1}, \dots, \mathcal{T}_{{\rm s}u}$ the connected components of $\mathcal{T}_{\rm s}$. Note that, shrinking each $\mathcal{T}_{{\rm b}i}$ and $T_{{\rm s}i}$ into a single vertex, we obtain a bipartite connected graph $\mathbf{T}$.

Put

$$P_{di}(x, y) = \sum_{t\in\mathcal{T}_{di}} x^{a_t} y^{b_t} \Delta^{2^{k_t}}, \quad d \in \{{\rm b}, {\rm s}\}.$$

Let $R_{d,i}$ be the union of the triangles in $\mathcal{T}_{d,i}$, where d is either b or s. Note that all the first coordinates of elements in $R_{{\rm b},i}$ are congruent modulo 2^l. In terms of polynomials, this means that

$$P_{\mathrm{b}i} = x^{\tilde{a}_i} y^{\tilde{b}_i} g_i(x^{2^l}, y^{2^l}), \quad i = 1, 2, \dots, m,$$

for appropriate polynomials

$$g_i(x, y) = \sum_{j \in P_{b,i}} x^{(a_j - \tilde{a}_i)/2^l} y^{(b_j - \tilde{b}_i)/2^l} \Delta^{2^{k_j - l}}, \quad i = 1, 2, \dots, m,$$

where $|\tilde{a}_i| \le 2^{l-1}$, and $|\tilde{b}_i| \le 2^{l-1}$. Multiplying P by an appropriate monomial, we may assume $\tilde{a}_1 = \tilde{b}_1 = 0$.

For each $\mathcal{T}_{\mathrm{b},i}$, consider the shortest path from $\mathcal{T}_{\mathrm{b},1}$ to $\mathcal{T}_{\mathrm{b},i}$ in the bipartite graph $\mathbf{T}$. The change from the initial values $\tilde{a}_1 = 0$ and $\tilde{b}_1 = 0$ to the final values $\tilde{a}_i$ and $\tilde{b}_i$ is due to the sets $\mathcal{T}_{\mathrm{s},j}$ in the path. In fact, if $\mathcal{T}_{\mathrm{b},i_0}$, $\mathcal{T}_{\mathrm{s},j}$ and $\mathcal{T}_{\mathrm{b},i_1}$ are consecutive vertices of this path, then $|\tilde{a}_{i_1}| \le |\tilde{a}_{i_0}| + 2^k |\mathcal{T}_{\mathrm{s},j}|$. It follows that $|\tilde{a}_i| \le 2^k |\mathcal{T}_{\mathrm{s}}| \le 2^l/5$ and similarly $|\tilde{b}_i| \le 2^l/5$ for $i = 1, \dots, m$.

Since each $R_{\mathrm{s},j}$ must intersect some $R_{\mathrm{b},i}$, there must be at least one point in $R_{\mathrm{s},j}$ congruent to $(\tilde{a}_i, \tilde{b}_i)$. As the diameter of $R_{\mathrm{s},j}$ is at most $2^k |\mathcal{T}_{\mathrm{s},j}| < 2^l/5$, the nearest lattice point to all points $2^{-l}\mathbf{w}$ with $\mathbf{w} \in \mathcal{T}_{\mathrm{s},j}$ is the same point $(A_j, B_j) \in \mathbb{Z}^2$, Furthermore, the distance between $(2^l A_j, 2^l B_j)$ and any point of $R_{\mathrm{s},j}$ is at most $2^{l+1}/5$.

Let I be the set of all integer pairs (α, β) for which the monomial $x^\alpha y^\beta$ appears in one of the polynomials $g_i(x, y)$. Write $g_i(x, y) = \sum_{(\alpha,\beta) \in I} \epsilon_{i,\alpha,\beta} x^\alpha y^\beta$ for suitable $\epsilon_{i,\alpha,\beta}$. Then

$$\begin{aligned} P(x, y) &= \sum_{(\alpha,\beta) \in I} x^{2^l \alpha} y^{2^l \beta} \left(\sum_{i=1}^{m} \varepsilon_{i,\alpha,\beta} x^{\tilde{a}_i} y^{\tilde{b}_i} \right) + \sum_{j=1}^{u} \sum_{t \in \mathcal{T}_{\mathrm{s},j}} x^{a_t} y^{b_t} \Delta^{2^{k_t}} \\ &= \sum_{(\alpha,\beta) \in I} x^{2^l \alpha} y^{2^l \beta} \left(\sum_{i=1}^{m} \varepsilon_{i,\alpha,\beta} x^{\tilde{a}_i} y^{\tilde{b}_i} + \sum_{\{j | (A_j, B_j) = (\alpha,\beta)\}} \sum_{t \in \mathcal{T}_{\mathrm{s},v}} x^{a_t - 2^l \alpha} y^{b_t - 2^l \beta} \Delta^{2^{k_t}} \right). \end{aligned}$$

Notice that every term in the above sum is contained in a neighborhood of radius $2^{l+1}/5$ of the corresponding point $(2^l \alpha, 2^l \beta)$. Since the diameter of P is smaller than $2^l/5$, one of the terms in this sum yields P and the others vanish. By the assumption that $(0, 0) \in R$, it follows that P must equal the term corresponding to $(0, 0)$, namely

$$\sum_{i=1}^{m} \varepsilon_{i,0,0} x^{\tilde{a}_i} y^{\tilde{b}_i} + \sum_{\{j | (A_j, B_j) = (0,0)\}} \sum_{t \in \mathcal{T}_{\mathrm{s},j}} x^{a_t} y^{b_t} \Delta^{2^{k_t}} = P(x, y),$$

whereas for $(\alpha, \beta) \ne (0, 0)$ we have

$$\sum_{i=1}^{m} \varepsilon_{i,\alpha,\beta} x^{\tilde{a}_i} y^{\tilde{b}_i} + \sum_{\{j | (A_j, B_j) = (\alpha,\beta)\}} \sum_{t \in \mathcal{T}_{\mathrm{s},j}} x^{a_t - 2^l \alpha} y^{b_t - 2^l \beta} \Delta^{2^{k_t}} = 0.$$

It follows that in the above expansion for $P(x, y)$ we can replace $2^l \alpha$ and $2^l \beta$ by α and β, respectively, and write

$$
\begin{aligned}
P(x,y) &= \sum_{(\alpha,\beta)\in I} x^{\alpha} y^{\beta} \Bigg(\sum_{i=1}^{m} \varepsilon_{i,\alpha,\beta} x^{\tilde{a}_i} y^{\tilde{b}_i} + \sum_{\{j \mid (A_j,B_j)=(\alpha,\beta)\}} \sum_{t\in\mathcal{T}_{\mathrm{s},j}} x^{a_t-2^l\alpha} y^{b_t-2^l\beta} \Delta^{2^{k_t}} \Bigg) \\
&= \sum_{i=1}^{m} x^{\tilde{a}_i} y^{\tilde{b}_i} g_i(x,y) + \sum_{(\alpha,\beta)\in I} x^{\alpha} y^{\beta} \Bigg(\sum_{\{v \mid (A_v,B_v)=(\alpha,\beta)\}} \sum_{t\in\mathcal{T}_{\mathrm{s},v}} x^{a_t-2^l\alpha} y^{b_t-2^l\beta} \Delta^{2^{k_t}} \Bigg) \\
&= \sum_{i=1}^{m} \sum_{t\in P_{b,i}} x^{(\tilde{a}_i+(a_t-\tilde{a}_i)/2^l)} y^{(\tilde{b}_i+(b_t-\tilde{b}_i)/2^l)} \Delta^{2^{k_t-l}} \\
&\quad + \sum_{j=1}^{u} \sum_{t\in\mathcal{T}_{\mathrm{s},j}} x^{a_t-2^l A_v+A_v} y^{b_t-2^l B_v+B_v} \Delta^{2^{k_t}}. \qquad \square
\end{aligned}
$$

Proof of Proposition 7.5. Take a decomposition as in (15) with $\max_{t\in\mathcal{T}}$ as small as possible. Suppose, say, that $2^{k_1} \leq \cdots \leq 2^{k_{h(P)}}$. Set $k_0 = 0$. Let m be the maximal integer such that $2^{k_m} \leq 5D$. If $s \geq m$, and if $2^{k_{s+1}-k_s} > 5h(R)$, the assumptions of Lemma 7.6 are satisfied for $k = k_s$ and $l = k_{s+1}$, so that the conclusion of that lemma contradicts our minimality assumption. It follows that $2^{k_{s+1}} \leq 5h(P)\cdot 2^{k_s}$ for $s \geq m$, and therefore by Theorem 7.1 we have $2^{k_{h(P)}} \leq (5h(P))^{h(P)}(5D) \leq (5r^3)^{r^3}(5D)$. Hence every triangle in the decomposition must be within a radius of $\sum_{t\in\mathcal{T}} 2^{k_t} \leq (r^3)(5h(P))^{h(P)}(5D) \leq (5r^3)^{r^3}(5D)$ from some point of R.

Since the number of decompositions satisfying these properties is finite, this yields an algorithm for finding a decomposition as required. □

8. *Higher-order mixing along special sequences*

The main theme of this paper is that a sequence $((a_1^{(n)}, b_1^{(n)}), (a_2^{(n)}, b_2^{(n)}), \ldots, (a_r^{(n)}, b_r^{(n)}))_{n=1}^{\infty}$ is mixing unless it satisfies certain quite restrictive conditions. In this section we apply our previous results to study the conditions under which some 'natural' sequences, arising from polynomials or multiplicative semigroups, are mixing. The examples below give further validation to our theme. We start with the following straightforward consequence of Lemma 6.2.

PROPOSITION 8.1. *Each of the following three conditions implies that the sequence*

$$((a_1^{(n)}, b_1^{(n)}), (a_2^{(n)}, b_2^{(n)}), \ldots, (a_r^{(n)}, b_r^{(n)}))_{n=1}^{\infty}$$

is mixing:

(1) $|a_i^{(n)} - a_j^{(n)}| \underset{n\to\infty}{\longrightarrow} \infty$ *for* $1 \leq i < j \leq r$.

(2) $|b_i^{(n)} - b_j^{(n)}| \underset{n\to\infty}{\longrightarrow} \infty$ *for* $1 \leq i < j \leq r$.

(3) $|(a_i^{(n)} + b_i^{(n)}) - (a_j^{(n)} + b_j^{(n)})| \underset{n\to\infty}{\longrightarrow} \infty$ *for* $1 \leq i < j \leq r$.

In fact, if the first condition in the proposition is satisfied, then we can bound from above the power of x dividing the left-hand side of (11), if the second condition is satisfied then we do the same using $1 + x$, and if the third condition holds then we prove the proposition by considering the degree of the left-hand side of (11).

The following proposition is an immediate consequence of Proposition 8.1.

PROPOSITION 8.2. *Let $p_i(x), q_i(x) \in \mathbb{Q}[x]$ be polynomials without constant term such that $p_i(\mathbb{Z}), q_i(\mathbb{Z}) \subseteq \mathbb{Z}$, $1 \le i \le r$. Suppose at least one of the following conditions holds:*
(1) *the polynomials $p_i(x)$ are mutually distinct;*
(2) *the polynomials $q_i(x)$ are mutually distinct;*
(3) *the polynomials $p_i(x) + q_i(x)$ are mutually distinct.*
Then the sequence $((p_1(n), q_1(n)), (p_2(n), q_2(n)), \ldots, (p_r(n), q_r(n)))_{n=1}^{\infty}$ is mixing.

Denote by $\mathbb{Z}^*$ the multiplicative semigroup of non-zero integers. Let $\pi_j : \mathbb{Z}^m \to \mathbb{Z}$ be the jth coordinate map, $j = 1, \ldots, m$.

LEMMA 8.3. *Assume that G is a finitely generated subsemigroup of the m-fold cartesian product $(\mathbb{Z}^*)^m$ with a set of generators $h_1, \ldots, h_t$, such that for all $j = 1, \ldots, m$ and $i = 1, \ldots, t$ we have $\pi_j(h_i) \neq \pm 2^s$ for $s \ge 1$. Let $l : \mathbb{Z}^m \to \mathbb{Z}$ be a linear function, and let (g_n) be a sequence in G such that $|l(g_n)| \underset{n\to\infty}{\longrightarrow} \infty$. Then, for any fixed integer d and any fixed $a_0, \ldots, a_d$, there exist at most finitely many n for which $l(g_n)$ is of the form $\sum_{i=0}^{d} a_i 2^{r_i}$.*

Proof. Let $R = \{2, 4, 8, \ldots\}$ be the subsemigroup of $\mathbb{Z}^*$ generated by 2. The conditions of the lemma imply that, for every $g \in G$ and $j = 1, \ldots, m$, the set $R\pi_j(g) \cap \pi_j(G)$ is finite.

Let (g_n) be a sequence in G such that $|l(g_n)| \underset{n\to\infty}{\longrightarrow} \infty$. Let $l(x_1, \ldots, x_m) = \sum_{i=1}^{m} \alpha_i x_i$. Replacing $\{1, \ldots, m\}$ by a subset thereof and passing to subsequences if needed, we may assume that no subsum of $l(g_n) = \sum_{i=1}^{m} \alpha_i g_{ni}$ vanishes. Thus we need to prove that the equation

$$l(g_n) - \sum_{j=0}^{d} a_j 2^{r_j} = 0$$

has at most finitely many solutions $(n, r_0, \ldots, r_d)$ with no vanishing subsums of $\sum_{j=0}^{d} a_j 2^{r_j}$. (If there is a vanishing subsum, replace it by a shorter sum.) In particular, it suffices to prove that

$$l(2^{-r_0} g) - \sum_{j=1}^{d} b_j 2^{r_j - r_0} = 1, \quad b_j = a_0^{-1} a_j$$

has finitely many solutions $(g, r_0, \ldots, r_d)$ with $g \in G$ and r_i as above. By [**2**, Theorem 1.1], there exist at most finitely many solutions of

$$l(g') - \sum_{j=1}^{d} b_j 2^{s_j} = 1, \quad g' \in G',$$

without vanishing subsums, where G' is the subgroup of $\mathbb{R}^*$ generated by G and

$$\{(2, 1, \ldots, 1), (1, 2, \ldots, 1), \ldots, (1, 1, \ldots, 2)\}.$$

Since $R\pi_j(g) \cap \pi_j(G)$ is finite for each $g \in G$, it follows that for every such g' there are at most finitely many possible values for r_0 such that $g = 2^{r_0} g' \in G$. In particular, $l(g_n) - \sum_{j=0}^{d} a_j 2^{r_j} = 0$ has finitely many solutions without vanishing subsums. Now we take a general solution of $l(g_n) - \sum_{j=0}^{d} a_j 2^{r_j} = 0$ and rewrite it as a sum of minimal vanishing subsums. We decompose a linear function into subsums in the form

$$l(x_1, \dots, x_m) = \alpha_1 x_1 + \cdots + \alpha_m x_m = \sum_j l_j(x_1, \dots, x_m),$$

where $\mathbf{T} = \{T_1, \dots, T_s\}$ is a partition of $\{1, \dots, m\}$ and $l_j(x_1, \dots, x_m) = \sum_{i \in T_j} \alpha_i x_i$. The partition $\mathbf{T}$ can be chosen so that the minimal vanishing subsums are of the form $l_j(g_n) - \sum_{i \in U_j} a_i 2^{r_i}$, where $\mathbf{U} = \{U_1, \dots, U_s\}$ is a similar partition for $\{0, \dots, c\}$. For any such pair of partitions $(\mathbf{T}, \mathbf{U})$ there are a finite number of solutions so that the minimal vanishing subsums are actually $l_j(g_n) - \sum_{i \in U_j} a_i 2^{r_i}$, so that there are finitely many solutions of $l(g_n) - \sum_{j=0}^{d} a_j 2^{r_j} = 0$. □

LEMMA 8.4. *Let G be as in the preceding lemma and let $l_1, \dots, l_r, l'_1, \dots, l'_r : \mathbb{Z}^m \to \mathbb{Z}$ be linear. If $(\gamma_n)_{n=1}^{\infty}$ is a sequence in G satisfying*

$$\rho((l_i(\gamma_n), l'_i(\gamma_n)), (l_j(\gamma_n), l'_j(\gamma_n))) \underset{n\to\infty}{\longrightarrow} \infty, \quad 1 \le i < j \le r,$$

then the sequence $((l_1(\gamma_n), l'_1(\gamma_n)), \dots, (l_r(\gamma_n), l'_r(\gamma_n)))$ is mixing.

Proof. If the sequence in question is not mixing, then by passing to a subsequence we may assume that, say, the sequence $(l_1(\gamma_n) - l_2(\gamma_n))$ is unbounded and is composed of numbers each of which is of the form $\sum_{i=1}^{d} \varepsilon_i 2^{r_i}$, where $d \le h(r)$ and $\varepsilon_i = \pm 1$ for each i. This contradicts Lemma 8.3, and thus proves our lemma. □

Let $D \subseteq \mathbb{Z}^2$. A $\mathbb{Z}^2$-action (T, S) on $(X, \mathcal{B}, \mu)$ is *r-mixing along D* if

$$\mu\left(\bigcap_{i=1}^{r} T^{m_i(t)} S^{n_i(t)} A_i\right) \longrightarrow \prod_{i=1}^{r} \mu(A_i), \quad A_1, \dots, A_r \in \mathcal{B},$$

as $(m_i(t), n_i(t)) \in D$ for each i and $\rho((m_i(t), n_i(t)), (m_j(t), n_j(t))) \underset{t\to\infty}{\longrightarrow} \infty$ for $i \ne j$.

THEOREM 8.5. *Let Γ be a finitely generated multiplicative subsemigroup of $\mathbb{Z}^*$. If $\Gamma \cap \{2^n : n = 1, 2, \dots\} = \emptyset$, then Ledrappier's system is mixing of all orders along $\Gamma \times \Gamma$.*

Proof. If Ledrappier's system is not mixing of order r along $\Gamma \times \Gamma$, then there exist sequences $(m_i(t))_{t=1}^{\infty}$ and $(n_i(t))_{t=1}^{\infty}$ in Γ, for $i = 1, \dots, r$, such that

$$\mu\left(\bigcap_{i=1}^{r} \sigma^{m_i(t)} \tau^{n_i(t)} A_i\right) \underset{t\to\infty}{\longrightarrow} K \ne \prod_{i=1}^{r} \mu(A_i),$$

and $\rho((m_i(t), n_i(t)), (m_j(t), n_j(t))) \underset{t\to\infty}{\longrightarrow} \infty$. Let $\gamma_t = (m_1(t), \dots, m_r(t), n_1(t), \dots, n_r(t)) \in G = \Gamma^{2n}$ and apply Lemma 8.4. □

THEOREM 8.6. *Let Γ be as in Theorem 8.5, and $p_i(x), q_i(x)$ be integer polynomials without constant term, $1 \le i \le r$, such that the pairs $(p_i(x), q_i(x))$ are distinct. Then*

$$\mu\left(\bigcap_{i=1}^{r} \sigma^{p_i(m)} \tau^{q_i(n)} A_i\right) \underset{\substack{(m,n)\to\infty \\ (m,n)\in\Gamma\times\Gamma}}{\longrightarrow} \prod_{i=1}^{r} \mu(A_i).$$

Proof. We proceed similarly to the proof of the preceding theorem. If the conclusion does not hold, then there exist sequences $(m_t)_{t=1}^{\infty}$ and $(n_t)_{t=1}^{\infty}$ in Γ such that

$$\mu\biggl(\bigcap_{i=1}^{r} \sigma^{p_i(m_t)}\tau^{q_i(n_t)}A_i\biggr) \underset{t\to\infty}{\longrightarrow} K \neq \prod_{i=1}^{r} \mu(A_i).$$

Put $\gamma_t = (m_t, \dots, m_t^M, n_t, \dots, n_t^M) \in G = \Gamma^{2M}$, where

$$M = \max\{\deg p_1, \dots, \deg p_r, \deg q_1, \dots, \deg q_r\}.$$

Now apply Lemma 8.4. □

To present our next result, we first need the following.

Definition 8.7. A polynomial $p(x) \in \mathbb{Q}[x]$ with $p(\mathbb{Z}) \subseteq \mathbb{Z}$ is *2-exceptional* if the diophantine equation $p(m) = 2^n$ has infinitely many positive solutions (m, n).

LEMMA 8.8. *A non-constant polynomial $p(x) \in \mathbb{Q}[x]$ with $p(\mathbb{Z}) \subseteq \mathbb{Z}$ is 2-exceptional if and only if it satisfies the following conditions:*
(a) $p(x) = 2^j(ax + b)^k$ *for certain integers a, b, j, k with $b, k \geq 1$ and $j \geq 0$;*
(b) *a is odd and $(a, b) = 1$;*
(c) *b is congruent to some power of 2 modulo a.*

Proof. Suppose first that $p(x)$ is 2-exceptional. By [**11**, Theorem 10.2], the polynomial $p(x)$ has a single root, so that $p(x) = d(ax + b)^k$ for some d, a, b, k with $a, k \geq 1$. Clearly, d must be a power of 2, and we may assume that $(a, b) = 1$. Also, the arithmetic progression $(am + b)_{m=1}^{\infty}$ must be a power of 2 infinitely often. For $2^{n'}$ to be of the form $am + b$ we need to have $2^{n'} \equiv b \pmod{a}$.

The converse direction is trivial. □

Going over the proof, we see that, moreover, given such a 2-exceptional polynomial, it is easy to characterize the argument values for which it assumes a power of 2 value.

To formulate our next theorem, it will be convenient to use the following notion. Let $p_i(x)$ and $q_i(x)$ be integer polynomials (or, more generally, rational polynomials, assuming integer values at all rational points) for $1 \leq i \leq 2$. The quadruple $(p_1(x), q_1(x), p_2(x), q_2(x))$ is *exceptional* if up to some additive constants it is of one of the forms

$$\begin{gathered}(p, 0, 0, p), (0, p, p, 0), (0, -p, -p, -p), (-p, -p, 0, -p),\\ (-p, p, -p, 0), (-p, 0, -p, p),\end{gathered} \tag{16}$$

where $p(x) = 2^t(ax + b)^r$, for some integers t, a, b, r with $t \geq 0$, $r \geq 1$, a odd and b in the orbit of 2 modulo a.

THEOREM 8.9. *Let $p_i(x), q_i(x) \in \mathbb{Q}[x]$ with $p_i(\mathbb{Z}), q_i(\mathbb{Z}) \subseteq \mathbb{Z}$ for $i = 1, 2$. Then*

$$\mu(A_0 \cap \sigma^{p_1(m)}\tau^{q_1(n)}A_1 \cap \sigma^{p_2(m)}\tau^{q_2(n)}A_2) \underset{(m,n)\to\infty}{\longrightarrow} \mu(A_0)\mu(A_1)\mu(A_2), \quad A_0, A_1, A_2 \in \mathcal{B},$$

if and only if the quadruple $(p_1(x), q_1(x), p_2(x), q_2(x))$ is not exceptional.

Proof. The convergence condition is satisfied if and only if the triangle

$$\{(0,0),\ (p_1(n),\ q_1(n)),\ (p_2(n),\ q_2(n))\}$$

is not a special triangle, up to some additive constants, infinitely often. Then it must be of one of the forms in (16), where $p(n)$ must be a power of 2 infinitely often. It follows that p must be special. □

The analogue of Theorem 8.9 for higher-order mixing seems unreachable with the current knowledge on the emerging diophantine equations. We start with the following.

Example 8.10. Consider the polynomials

$$\begin{aligned} p_1(x) &= 0, & q_1(x) &= 0,\\ p_2(x) &= 0, & q_2(x) &= x^7,\\ p_3(x) &= x^7 - x^3, & q_3(x) &= 0,\\ p_4(x) &= x^7 - x^3, & q_4(x) &= x^3. \end{aligned}$$

According to Theorem 3.3 and Proposition 4.1 (using the quadrangle Q_3), the sequence

$$((p_1(n),\ q_1(n)),\ (p_2(n),\ q_2(n)),\ (p_3(n),\ q_3(n)),\ (p_4(n),\ q_4(n)))_{n=1}^{\infty}$$

is not mixing.

Thus, a 4-tuple of pairs of polynomial sequences may fail to mix even if some of the differences are not (up to a constant) polynomials with a single root. The reason is that these polynomials are allowed to assume values which are sums or differences of two powers of 2, so that [**11**, Theorem 10.2] is not applicable any more. To characterize mixing polynomial sequences of length 4, one would need to find which polynomials are guaranteed not to assume infinitely often values of the form $2^k \pm 2^l$. In view of Theorem 7.1, to characterize r-tuples of pairs of polynomials for which we have mixing, one needs first to find which polynomials may assume infinitely often values which are sums and/or differences of up to some fixed number of powers of 2. As even the resolution of the very special case, of finding which sums/differences of up to three powers of 2 are squares, is quite recent [**12**] (see also [**4**] and [**10**]), it seems that a lot of work still needs to be done to that end. In this connection, we raise the following.

QUESTION 8.11. *Given any positive integer C, characterize those polynomials $p(x)$ for which the diophantine equation*

$$p(m) = 2^{n_1} \pm 2^{n_2} \pm \cdots \pm 2^{n_C}$$

may have infinitely many solutions $m, n_1, \ldots, n_C$.

THEOREM 8.12. *If $\{(a_1, b_1), (a_2, b_2), \ldots, (a_r, b_r)\}$ is a special r-gon, then:*

(1) *the set $\{1 \le k \le r : a_k = \min_{1 \le i \le r} a_i\}$ is of even size, and in particular consists of at least two numbers;*

(2) *the set $\{1 \le k \le r : b_k = \min_{1 \le i \le r} b_i\}$ is of even size, and in particular consists of at least two numbers;*

(3) *the set $\{1 \le k \le r : a_k + b_k = \max_{1 \le i \le r}(a_i + b_i)\}$ is of even size, and in particular consists of at least two numbers.*

Proof.

(1) Suppose $\min_{1\le i\le r} a_i = a$, where the minimum is obtained an odd number of times. Then the coefficient of x^a in the polynomial $\sum_{1\le i\le r} x^{a_i}(1+x)^{b_i}$ is 1, and in particular $\sum_{1\le i\le r} x^{a_i} y^{b_i}$ is not an L-polynomial.

(2) This part follows from the preceding part by interchanging x and $1+x$.

(3) Suppose $\max_{1\le i\le r}(a_i+b_i) = c$, where the maximum is obtained an odd number of times. Then the coefficient of x^c in the polynomial $\sum_{1\le i\le r} x^{a_i}(1+x)^{b_i}$ is 1, and in particular $\sum_{1\le i\le r} x^{a_i} y^{b_i}$ is not an L-polynomial. □

Example 8.13.

$$\mu\Bigl(\bigcap_{i=1}^{r} \sigma^{n^i}\tau^n A_i\Bigr) \underset{n\to\infty}{\longrightarrow} \prod_{i=1}^{r} \mu(A_i), \qquad A_1, A_2, \ldots, A_r \in \mathcal{B}.$$

Example 8.14. If $p_1(x), q_1(x), p_2(x), q_2(x), \ldots, p_r(x), q_r(x)$ are polynomials of mutually distinct degrees, then

$$\mu\Bigl(\bigcap_{i=1}^{r} \sigma^{p_i(n)}\tau^{q_i(n)} A_i\Bigr) \underset{n\to\infty}{\longrightarrow} \prod_{i=1}^{r} \mu(A_i), \qquad A_1, A_2, \ldots, A_r \in \mathcal{B}.$$

Definition 8.15. A set $E \subseteq \mathbb{N}$ is *rarified* if

$$|\{1 \le n \le N : n \in E\}| = O((\log N)^C)$$

for some constant C.

We shall sometimes say *rarified with exponent C* when we wish to specify the constant C in the definition.

LEMMA 8.16. *If $E_1, \ldots, E_n$ are rarified with exponent C, then so is $E_1 \cup \cdots \cup E_n$.*

Proof. Trivial. □

LEMMA 8.17. *Let $f : \mathbb{Z} \to \mathbb{Z}$ be a function that is at most M-to-1 and that satisfies $f(n) = O(n^R)$ for positive constants M and R. If E is rarified of exponent C then so is $f^{-1}(E)$. In particular, the inverse image in $\mathbb{N}$ of a rarified set under a polynomial map, which maps $\mathbb{N}$ into itself, is rarified as well.*

Proof. Let K be a constant such that $f(n) \le Kn^R$ and $F = f^{-1}(E)$. Put $E(N) = \{1 \le n \le N : n \in E\}$ and $F(N) = \{1 \le n \le N : n \in F\}$. Then $y \in F(N)$ implies $f(y) \in E(KN^R)$. Therefore,

$$\begin{aligned} |F(N)| \le M|E(KN^R)| &= O([\log(KN^R)]^C) \\ &= O([\log K + R\log N]^C) \\ &= O((\log N)^C). \end{aligned}$$ □

Note that this lemma applies in particular to polynomials.

THEOREM 8.18. *Let $p_i(x), q_i(x) \in \mathbb{Q}[x]$, $1 \le i \le r$, be polynomials without constant term, assuming integer values on integers, and such that $(p_i(x), q_i(x)) \ne (p_j(x), q_j(x))$ for $i \ne j$. Then there exists a rarified set $E \subseteq \mathbb{Z}$ such that*

$$\mu\left(\bigcap_{i=1}^{r} \sigma^{p_i(n)}\tau^{q_i(n)}A_i\right) \longrightarrow \prod_{i=1}^{r} \mu(A_i), \qquad A_1, A_2, \ldots, A_r \in \mathcal{B},$$

as $n \to \infty$ along values outside E.

For the proof, we need the following fact.

PROPOSITION 8.19. *Let $((a_1^{(n)}, b_1^{(n)}), (a_2^{(n)}, b_2^{(n)}), \ldots, (a_r^{(n)}, b_r^{(n)}))_{n=1}^{\infty}$ be a non-mixing sequence of r-tuples of pairs of integers. Denote by $D_h \subset \mathbb{Z}$ the set of sums and differences of at most h powers of 2, where $h = h(r)$ is as defined in §7. Then, for some pair $(i, j) = (i(n), j(n))$, the difference $a_i^{(n)} - a_j^{(n)}$ tends to ∞ with n, and is at a bounded distance from D_h infinitely often.*

Note that, by passing to a subsequence, we may assume (i, j) to be constant.

In the proof we shall use the notion of a *minimal special r-gon*, which is a special r-gon, containing properly no other special r-gon.

Proof of Proposition 8.19. Obviously, any special r-gon is a union of minimal ones. Let f be a minimal special r-gon. Write $f = \sum_{i=1}^{h} \Delta_i$, where each Δ_i is triangular. Consider the graph G, whose vertices are $v_1, \ldots, v_h$ and containing the edge $\overline{v_i v_j}$ if the triangular polynomials Δ_i and Δ_j have a common term. In other words, a term of Δ_i may cancel a term of Δ_j only if $\overline{v_i v_j}$ is an edge of G. In particular, there can be no cancellation between triangular polynomials corresponding to vertices in different connected components of G. It follows that, if $v_{i_1}, \ldots, v_{i_r}$ are all vertices in one connected component, then $\sum_{u=1}^{r} \Delta_{i_u}$ is a subsum of f and a special r'-gon. Thus, G must be connected. Note that we can transform a triple (i, a, b), where $x^a y^b$ is a term of Δ_i, to any other by a sequence of alternating steps of the following types:
(1) replace (i, a, b) by (j, a, b) if $x^a y^b$ is also a term of Δ_j;
(2) replace (i, a, b) by (i, c, d) if $x^c y^d$ is also a term of Δ_i.
Only the steps of the second type change the coordinates, and they do so only by a power of 2. By choosing a simple path in the graph, we see that we need at most h steps of type 1, and therefore also at most h steps of type 2. Since a special r-gon contains at least 3 non-collinear points, it follows that every special r-gon contains terms of the form $x^a y^b$ and $x^c y^d$, where $a - c \in D_h$. Now the result follows from Theorem 3.3. □

Proof of Theorem 8.18. By Proposition 8.19, if every sequence $a_i(n) - a_j(n)$ or $b_i(n) - b_j(n)$ that tends to infinity gets away from D_h, then the sequence

$$((a_1(n), b_1(n)), (a_2(n), b_2(n)), \ldots, (a_r(n), b_r(n)))$$

is mixing. Note that there exists a rarified set E, such that the sequence $x(n)$ satisfies $\rho(x(n), D) \underset{n\to\infty}{\longrightarrow} \infty$ provided that $x(n) \underset{n\to\infty}{\longrightarrow} \infty$ and $x(n) \notin E$. Now the result follows from Lemmas 8.16 and 8.17. □

Acknowledgements. The first author's research was partially supported by Fondecyt, Grant #1040227, and the Chilean Presidential Cathedra in Number Theory. The second author's research was supported in part by the Israel Research Foundation (Grant #186/01). The third author's research was partially supported by NSF Grant DMS-0600042.

The authors would like to express their gratitude to the referee, who pointed out several relevant references, which helped to put the paper in the right context.

References

[1] M. Einsiedler and T. Ward. Asymptotic geometry of non-mixing sequences. *Ergod. Th. & Dynam. Sys.* **23** (2003), 75–85.

[2] J.-H. Evertse, H. P. Schlickewei and W. M. Schmidt. Linear equations in variables which lie in a multiplicative group. *Ann. of Math.* (2) **155** (2002), 807–836.

[3] F. Ledrappier. Un champ markovien peut être d'entropie nulle et mélangeant. *C. R. Acad. Sci. Paris* **287** (1978), 561–563.

[4] F. Luca. The Diophantine equation $x^2 = p^a \pm p^b + 1$. *Acta Arith.* **112** (2004), 87–101.

[5] D. W. Masser. Mixing and linear equations over groups in positive characteristic. *Israel J. Math.* **142** (2004), 189–204.

[6] V. A. Rohlin. On endomorphisms of compact commutative groups (Russian). *Izv. Akad. Nauk SSSR, Ser. Mat.* **13** (1949), 329–340.

[7] H. P. Schlickewei. S-unit equations over number fields. *Invent. Math.* **102** (1990), 95–107.

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

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

[10] R. Scott. Elementary treatment of $p^a \pm p^b + 1 = x^2$. http://arxiv.org/pdf/math/0608796 (also available at http://www.homepage.villanova.edu/robert.styer/ReeseScott/index.htm).

[11] T. N. Shorey and R. Tijdeman. *Exponential Diophantine Equations*. Cambridge University Press, Cambridge, 1986.

[12] L. Szalay. The equations $2^N \pm 2^M \pm 2^L = z^2$. *Indag. Math. (N.S.)* **13** (2002), 131–142.

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

[14] B. Weiss. Personal communication.