Ergod. Th. & Dynam. Sys. (First published online 2017), page 1 of 19*
doi:10.1017/etds.2016.137
*Provisional—final page numbers to be inserted when paper edition is published

Linear equations over multiplicative groups, recurrences, and mixing III

H. DERKSEN† and D. MASSER‡

† *Department of Mathematics, University of Michigan, East Hall, 530 Church Street, Ann Arbor, Michigan 48104, USA*
(e-mail: hderksen@umich.edu)

‡ *Departement Mathematik und Informatik, Universität Basel, Spiegelgasse 1, 4051 Basel, Switzerland*
(e-mail: David.Masser@unibas.ch)

(*Received* 21 *October* 2016 *and accepted in revised form* 14 *November* 2016)

Abstract. Given an algebraic $\mathbf{Z}^d$-action corresponding to a prime ideal of a Laurent ring of polynomials in several variables, we show how to find the smallest order $n+1$ of non-mixing. It is known that this is determined by the non-mixing sets of size $n+1$, and we show how to find these in an effective way. When the underlying characteristic is positive and $n \geq 2$, we prove that there are at most finitely many classes under a natural equivalence relation. We work out two examples, the first with five classes and the second with 134 classes.

1. *Introduction*

Not long ago the second author published a paper [**M**] about linear equations over multiplicative groups in positive characteristic. This was specifically aimed at an application to a problem about mixing for dynamical systems of algebraic origin, and as a result about linear equations it lacked some of the simplicity of the classical results in zero characteristic. A new feature was the appearance of $n-1$ independently operating Frobenius maps; here n is the number of variables.

Soon afterwards the first author published a paper [**D**] about recurrences in positive characteristic. He proved an analogue of the Skolem–Lech–Mahler theorem famous in zero characteristic. A new feature was the appearance of integer sequences involving combinations of $d-2$ powers of the characteristic; here d is the order of the recurrence.

It turns out that these two new features are identical. In positive characteristic the vanishing of a recurrence with d terms can be regarded as a linear equation in $d-1$ variables to be solved in a multiplicative group (so in particular $n-1=d-2$). This observation can be developed in three directions.

In Part I of this series [**DM1**] we gave an improved version of the result of [**M**] in a form more closely related to that in zero characteristic. In Part II [**DM2**] we applied this to recover the result of [**D**], and indeed we generalized it to sums of recurrences. Here in Part III we present some new applications to mixing problems for dynamical systems of algebraic origin. In an earlier version we gave an effective algorithm to determine the smallest order $n+1$ of non-mixing of any basic action associated with a given prime ideal in a Laurent polynomial ring. This solved the problem (3) mentioned by Schmidt in [**S**, p. 283].

Thanks to the work of [**M**], we know that this non-mixing comes from sets of cardinality $n+1$ which are themselves non-mixing for α (see later for definitions). After receiving our solution mentioned above, Schmidt in a message dated 12th July 2006 asked us if it is possible to determine all these non-mixing sets (or 'shapes') effectively. This we do in the present paper, which also includes a different method of determining n.

For a positive integer d, let α be a $\mathbf{Z}^d$-action on a compact abelian group. We have three possibilities:

(I) there is n-mixing but not $(n+1)$-mixing for some unique $n = n(\alpha) \geq 2$;
(II) there is no 2-mixing;
(III) there is n-mixing for all $n \geq 2$.

In case (II) we may write $n(\alpha) = 1$, and in case (III) we may write $n(\alpha) = \infty$.

Write $\mathcal{R} = \mathcal{R}_d$ for the Laurent polynomial ring $\mathbf{Z}[u_1, u_1^{-1}, \ldots, u_d, u_d^{-1}]$. As in [**S**, Lemma 5.1, p. 36], for any countable $\mathcal{R}$-module $\mathcal{M}$ there is a corresponding $\mathbf{Z}^d$-action $\alpha = \alpha^{\mathcal{M}}$ by automorphisms of the compact metric group $\widehat{\mathcal{M}}$. We may therefore write $n(\alpha) = n(\mathcal{M})$. By [**S**, Theorem 27.2(1), p. 264], the mixing properties of α are determined by the mixing properties of the actions $\alpha^{\mathcal{R}/\mathcal{P}}$ corresponding to the prime ideals $\mathcal{P}$ of $\mathcal{R}$ associated with $\mathcal{M}$. In particular,

$$n(\mathcal{M}) = \min_{\mathcal{P}} n(\mathcal{R}/\mathcal{P}).$$

So, in some sense it suffices to consider just these $\alpha = \alpha^{\mathcal{R}/\mathcal{P}}$. Certainly if $\mathcal{M}$ is Noetherian there are only finitely many $\mathcal{P}$ to consider, and it is well known that these can often be effectively found (for example if $\mathcal{M}$ is an ideal of $\mathcal{R}$).

Then for $\alpha = \alpha^{\mathcal{R}/\mathcal{P}}$ a set $\{\mathbf{m}_0, \ldots, \mathbf{m}_n\}$ in $\mathbf{Z}^d$ of cardinality $n+1$ is non-mixing if and only if there are $a_0, \ldots, a_n$ in the quotient field K of the integral domain $\mathcal{R}/\mathcal{P}$, not all zero, such that

$$a_0 \mathbf{u}^{\mathbf{m}_0 k} + \cdots + a_n \mathbf{u}^{\mathbf{m}_n k} = 0 \tag{1.1}$$

(in K) for infinitely many positive integers k, where $\mathbf{u}^{\mathbf{m}} = u_1^{m_1} \cdots u_d^{m_d}$ for $\mathbf{m} = (m_1, \ldots, m_d)$.

If the characteristic of $\mathcal{R}/\mathcal{P}$ is zero (so that $\mathcal{P} \cap \mathbf{Z}$ is zero) and α is mixing, then it is known that α is n-mixing for every $n \geq 3$. See [**S**, Theorem 27.3(2), p. 265] for the proof, due to Schmidt and Ward [**SW**], which amounts to showing that it is equivalent to the classical results of Evertse, Schlickewei, and van der Poorten about linear equations over multiplicative groups in zero characteristic. Thus, in this case $n(\mathcal{R}/\mathcal{P})$ must be either 1 or ∞.

The dichotomy here can be resolved in several ways; here is one possibility.

It is known that $\alpha^{\mathcal{R}/\mathcal{P}}$ is 2-mixing (that is, just mixing) if and only if $u_1, \ldots, u_d$ stay multiplicatively independent in K. See for example [**S**, Theorem 6.5(2), p. 47]. Now it is not difficult to determine whether $u_1, \ldots, u_d$ become multiplicatively dependent modulo constants of K; a good estimate in terms of the variety in $\mathbf{C}^d$ associated with $\mathcal{P}$ is given as [**BMZ**, Lemma 3.2, p. 14], for example. If there is such a dependence, then using a simple induction we can even determine all relations

$$\mathbf{u}^{\mathbf{b}} = \beta \tag{1.2}$$

in the form of a basis $(\mathbf{b}_1, \ldots, \mathbf{b}_r)$ for the group B of all $\mathbf{b} = (b_1, \ldots, b_d)$ in $\mathbf{Z}^d$ for which there exists a constant β in (1.2). These β must be algebraic over $\mathbf{Q}$; call them $\beta_1, \ldots, \beta_r$ corresponding to the basis elements. Now it is clear that $u_1, \ldots, u_d$ become multiplicatively dependent in K if and only if $\beta_1, \ldots, \beta_r$ are themselves multiplicatively dependent. This latter can be determined in a standard way using heights; for a good estimate in a typical situation, see [**LM**, Corollary 3.2, p. 281], for example.

Let us assume that $\beta_1, \ldots, \beta_r$ are indeed multiplicatively dependent, so that the smallest order of non-mixing is 2.

Then in a similar way one can determine the group of all $\mathbf{c} = (c_1, \ldots, c_r)$ in $\mathbf{Z}^r$ such that $\beta_1^{c_1} \cdots \beta_r^{c_r} = 1$. Via $\mathbf{b} = c_1\mathbf{b}_1 + \cdots + c_r\mathbf{b}_r$, this leads easily to the subgroup B_1 of B for which $\mathbf{u}^{\mathbf{b}} = 1$. And, B_1 is of finite index in a unique primitive subgroup $\sqrt{B_1}$ in $\mathbf{Z}^d$; this is the set of $\mathbf{b}$ for which there exists a root of unity ζ with

$$\mathbf{u}^{\mathbf{b}} = \zeta. \tag{1.3}$$

Now it is an easy exercise using (1.1) to show that the set $\{\mathbf{m}_0, \mathbf{m}_1\}$ of cardinality 2 in $\mathbf{Z}^d$ is non-mixing for α if and only if the non-zero $\mathbf{m}_0 - \mathbf{m}_1$ lies in $\sqrt{B_1}$.

Thus, the only real problems arise when the characteristic p of $R/\mathcal{P}$ is positive, and from now on we assume that this is the case. Then it is known that $n = n(\mathcal{R}/\mathcal{P}) < \infty$ (see below).

When $n = 1$, we can reason as in zero characteristic. Namely, the arguments of [**BMZ**, Lemma 3.2] stay valid in positive characteristic; the essential fact is that a field of rational functions over $\mathbf{F}_p$ in several variables is still a 'field with a proper set of absolute values satisfying a product formula'. Indeed, this fact was used throughout [**DM1**] to define all the heights there. But then β in (1.2) is algebraic over $\mathbf{F}_p$ and so a root of unity; thus, we are automatically in (1.3).

Now, by [**S**, Theorem 28.7, p. 275], the non-mixing property of a set is invariant under $\mathbf{Z}^d$-translation and also under multiplication by a positive integer. Also, from (1.1), it is trivially invariant under dividing by a positive integer as long as the set stays in $\mathbf{Z}^d$. Thus, in particular, it seems reasonable to think of the non-mixing sets as being in $\mathbf{Q}^d$ rather than $\mathbf{Z}^d$; further, they fall into natural equivalence classes as follows.

Define two finite sets M, $M^{\#}$ in $\mathbf{Q}^d$ to be equivalent if there is a positive rational x, and $\mathbf{f}$ in $\mathbf{Q}^d$, such that $xM = M^{\#} + \mathbf{f}$ (this is not quite the same definition as in Ward [**W**, p. 2]). We might without much confusion describe the equivalence classes also as 'shapes'. Clearly, every non-empty equivalence class contains an M in $\mathbf{Z}^d$. We can even take all the coordinates non-negative, and moreover make sure that the convex hull touches every coordinate hyperplane (for example when $d = 2$ we just push the set as far as it will go

south and west). This is the same as saying that the Laurent polynomial $\sum_{\mathbf{m}\in M} u^{\mathbf{m}}$ is a genuine polynomial and not divisible by any of $u_1, \ldots, u_d$. We could call such a set semi-reduced.

We can further assume that no $s^{-1}M$ $(s = 2, 3, \ldots)$ is in $\mathbf{Z}^d$; and this we call reduced.

It is not difficult to see that the reduced set in each class is unique (we will not need this until the examples). At first it is rather clear that if M, $M^{\#}$ are both semi-reduced and translates of each other, then they are equal (when $d = 2$, this is obvious from pushing). And, if M, $M^{\#}$ are both reduced and equivalent, then xM, $M^{\#}$ are semi-reduced and translates, so $xM = M^{\#}$. Writing $x = r/s$ for positive coprime r, s, we see that $(r/s)M$ is in $\mathbf{Z}^d$; but, as $(s/s)M$ is too, so is $(1/s)M$ in $\mathbf{Z}^d$. Thus, $s = 1$. Similarly, $r = 1$, so $M = M^{\#}$, proving the uniqueness.

Here is the main result of this paper.

THEOREM *Given a prime ideal $\mathcal{P}$ of $\mathcal{R}$ with $\mathcal{P} \cap \mathbf{Z} = p\mathbf{Z}$ $(p > 0)$, the smallest order $n + 1 = n(\mathcal{R}/\mathcal{P}) + 1 < \infty$ of non-mixing can be effectively determined. Further, if $n \geq 2$, then there are only finitely many equivalence classes of non-$(n+1)$-mixing sets, and these can be effectively determined.*

The discussion above shows that the condition $n \geq 2$ is important for the finiteness. In fact the arguments above make it clear that a non-mixing action in positive characteristic can have infinitely many equivalence classes of non-mixing sets, but that this happens if and only if the rank r of the analogue of the group $\sqrt{B_1}$ satisfies $r \geq 2$. For example, this happens when $\mathcal{P}$ in $\mathcal{R}_2$ contains both $u_1 - 1$ and $u_2 - 1$, but not when $\mathcal{P}$ in $\mathcal{R}_1$ contains $u_1 - 1$.

We note that there is usually no trouble to find an effective upper bound for the smallest order of non-mixing, for example if $\mathcal{P}$ is explicitly given in terms of generators. Just pick any P in $\mathcal{P}$ not in $p\mathcal{R}$, so that $P(u_1, \ldots, u_d) = 0$ in K, and take the $k = p^e$ powers $(e = 0, 1, 2, \ldots)$; the resulting equations then show by (1.1) that α is not $(N+1)$-mixing, where $N + 1 \geq 2$ is the number of non-zero terms in P reduced modulo p. So, $n(\mathcal{R}/\mathcal{P}) \leq N$.

Thus, it would seem that our work has something to do with the problem of finding the 'shortest' polynomial in a given ideal; see also [**S**, p. 282]. In zero characteristic this problem is surprisingly difficult and probably there is in general no effective algorithm. In one variable it is related to a conjecture of Posner and Rumsey; see for example the article [**SV**] of Schlickewei and Viola, which makes use of the subspace theorem in the form of an S-unit equation. However, the latter was one of the key objects in [**D**, **M**, **DM1**], and the lesson there is that things are much easier in positive characteristic. In this case it is quite likely that the work in [**DM1**] leads to an effective solution of the shortest polynomial problem, although we do not investigate this in the present paper. But actually there is an extra twist here, which arises from the main result of [**M**]. Namely, we may have to extend the Laurent ring to a Puiseux ring.

A nice example of this is given in [**S**, p. 278]. Here $\mathcal{P}$ is generated by $p = 2$ and

$$P = 1 + u_1 + u_1^3 + u_1^5 + u_1^6 + u_2 = (1 + u_1 + u_1^2)^3 + u_2, \tag{1.4}$$

where the shortest polynomial is probably P; at any rate α is not 6-mixing. But P, although irreducible in $\mathbf{F}_2[u_1, u_1^{-1}, u_2, u_2^{-1}]$, is clearly divisible by

$$Q = 1 + u_1 + u_1^2 + u_2^{1/3}$$

in $\mathbf{F}_2[u_1, u_2^{1/3}]$. Now the Q^{p^e} show equally well that α is not 4-mixing (see §4). The general situation for principal ideals $\mathcal{P}$ (when considered mod p) is clarified in terms of non-mixing sets by [**S**, Proposition 28.9, p. 276]. This shows how to find all non-mixing sets that are minimal in a certain sense. But it does not show how to find the ones of smallest cardinality. Here we illustrate our techniques by proving that α is 3-mixing with exactly five classes of non-mixing sets of cardinality 4.

The other examples in [**S**] all concern principal ideals. Here we consider also a non-principal ideal. It is generated by 2 and

$$P_1 = 1 + u_1 + u_1^2 + u_2, \quad P_2 = 1 + u_1 + u_1^3 + u_3. \tag{1.5}$$

Again there is certainly no 4-mixing, and again we will prove that there is 3-mixing. But this time there are exactly 134 classes of non-mixing sets of cardinality 4. The most complicated one comes from the fact that our ideal happens to contain

$$u_1^{25} + u_1^{20}u_2u_3 + u_2^{12} + u_3^4.$$

These examples should make it clear that the determination of the smallest order of non-mixing and the equivalence classes of corresponding non-mixing sets is not only effective but also fairly practical. By using the estimates in [**DM1**], it should also be possible to give explicit bounds for the sets in terms of $\mathcal{P}$ or more precisely its generators.

Our proof uses observations from [**M**] as well as one of the main results of [**DM1**].

More precisely, let V be a variety in projective n-space defined by linear equations in $X_0, \ldots, X_n$ over positive characteristic. The work of [**DM1**] shows how to find all points of V whose coordinates are in a given finitely generated group. The precise description can be complicated, involving (as we mentioned) as many as $n-1$ independently operating Frobenius maps, as well as cosets defined by equations $X_i = aX_j$ (see for example [**DM1**, Theorem 1, p. 1049]).

But for an action $\alpha = \alpha^{\mathcal{R}/\mathcal{P}}$ as above with $n = n(\alpha)$, the hyperplane V_n defined by the single equation $X_0 + \cdots + X_n = 0$, and the group as the radical inside K of the group generated by $u_1, \ldots, u_d$ in K, the description is much simpler. In particular, only a single Frobenius turns up, and apart from cosets we see only points $(\xi_0^{p^e}, \ldots, \xi_n^{p^e})$ ($e = 0, 1, \ldots$) for a finite set Π of $(\xi_0, \ldots, \xi_n)$. This is proved in Lemma 5. It also shows that Π is closely related to the desired equivalence classes; for that we need the concept of 'broad set' used in [**M**], which is crucial to control the coefficients $a_0, \ldots, a_n$ in (1.1).

Our paper is arranged as follows. In §2 we prove four lemmas as preparation for the fifth. Our theorem follows quickly in §3. Then §4 treats the example (1.4) and §5 the much more difficult example (1.5).

2. *Preliminaries*

Let k be any field (even of zero characteristic) and let $\mathcal{L}$ be a vector space of linear forms L in variables $X_0, \ldots, X_n$. We say that $L = \sum_{i \in I} c_i X_i$ in $\mathcal{L}$ is minimal if there is no non-empty subset $I' \neq I$ of I such that some non-zero $L' = \sum_{i \in I'} c_i' X_i$ is in $\mathcal{L}$.

LEMMA 1. *The space $\mathcal{L}$ is generated by its minimal forms.*

Proof. Compare [**BM**, Lemma 4, p. 431]. It suffices to prove that every non-zero form L in $\mathcal{L}$ can be written as a linear combination of minimal forms. This will be by induction on the length l of L, that is, the number of non-zero coefficients. The case $l = 1$ is trivial. So, assume for some $l \geq 2$ that this holds for all forms of $\mathcal{L}$ of length strictly less than l. Take L in $\mathcal{L}$ of length exactly l. After a permutation, we can suppose that $L = c_0X_0 + \cdots + c_{l-1}X_{l-1}$. If L is already minimal, we are done. Otherwise we can assume after another permutation that some $L' = c'_0X_0 + \cdots + c'_{m-1}X_{m-1}$ lies in $\mathcal{L}$ with some $m < l$ and $c'_0 \neq 0$. Then L' and $L'' = c'_0L - c_0L' \neq 0$ are both of length strictly less than l, and so the induction hypothesis can be applied to $L = (c_0/c'_0)L' + (1/c'_0)L''$, giving the required assertion for L. This proves the lemma. □

With $\mathcal{R}$ and $\mathcal{P}$ as the theorem, we work in the quotient field K of $\mathcal{R}/\mathcal{P}$; then $\mathbf{F} = \overline{\mathbf{F}_p} \cap K$ is a finite field. We also work with the group G generated in K^* by the images of $u_1 \neq 0, \ldots, u_d \neq 0$. We write $\sqrt{G}$ for the radical of G inside K. This is well known to be finitely generated (see for example [**M**, p. 195]). It clearly also has rank d; let $v_1, \ldots, v_d$ be basis elements modulo torsion (that is, modulo $\mathbf{F}^*$). For $\mathbf{m} = (m_1, \ldots, m_d)$ in $\mathbf{Z}^d$, we abbreviate $v_1^{m_1} \cdots v_d^{m_d}$ as above to $\mathbf{v}^{\mathbf{m}}$.

We write $\mathbf{P}_n(\sqrt{G})$ for the points of projective n-space $\mathbf{P}_n$ whose coordinates can be taken in $\sqrt{G}$. For a variety V in $\mathbf{P}_n$ defined by linear equations, we write $V(\sqrt{G})$ for the intersection $V \cap \mathbf{P}_n(\sqrt{G})$. We are going to use some results of [**DM1**], in which we say that ψ from $\mathbf{P}_n$ to $\mathbf{P}_n$ is a $\sqrt{G}$-isomorphism if it is defined by $\psi(X_0, \ldots, X_n) = (g_0X_0, \ldots, g_nX_n)$ for $g_0, \ldots, g_n$ in $\sqrt{G}$; and that V is $\sqrt{G}$-isotrivial if there is such a ψ with $\psi(V)$ defined over $\mathbf{F}$. We say that a variety is transversal if each one of the projective variables $X_0, \ldots, X_n$ occurs in the defining equations with non-zero coefficient. We say that a variety is a torsion coset if it is defined by equations of the form $X_i = \zeta X_j$ $(i \neq j)$ with ζ in $\mathbf{F}^*$. A transversal torsion coset Z leads to a partition $I_1 \cup \cdots \cup I_h$ of $\{0, 1, \ldots, n\}$ into parts of size at least two together with $\zeta_0, \ldots, \zeta_n$ in $\mathbf{F}^*$, such that for each $j = 1, \ldots, h$ the equality of the quotients X_i/ζ_i $(i \in I_j)$ defines Z. We define the variety V_n by $X_0 + \cdots + X_n = 0$.

LEMMA 2. *Suppose for some $n \geq 1$ that α is $(n+1)$-mixing. Then there exists a finite collection $\mathcal{Z}$ of transversal torsion cosets Z in V_n such that*

$$V_n(\sqrt{G}) = \bigcup_{Z \in \mathcal{Z}} Z(\sqrt{G}).$$

Proof. This bears some resemblance to the Descent Step (a) over $\sqrt{G}$ of [**DM1**, p. 1047]. However, we cannot apply it here because V_n is not only $\sqrt{G}$-isotrivial but even defined over $\mathbf{F}_p$. The proof that follows is self-contained. Take any point $(\xi_0, \ldots, \xi_n)$ of $V_n(\sqrt{G})$, and write $\xi_i = \zeta_i \mathbf{v}^{\mathbf{r}_i}$ for torsion ζ_i $(i = 0, 1, \ldots, n)$. Frobenius leads to

$$\zeta_0 \mathbf{v}^{\mathbf{r}_0 q} + \cdots + \zeta_n \mathbf{v}^{\mathbf{r}_n q} = 0$$

for infinitely many prime powers $q = p^e$. We convert this into powers $\mathbf{u}^{\mathbf{m}}$ by a standard argument. Let $s = [\sqrt{G} : G]$. There is some r such that q is congruent to r modulo s for

infinitely many q, and we get

$$a_0\mathbf{u}^{\mathbf{m}_0 k} + \cdots + a_n\mathbf{u}^{\mathbf{m}_n k} = 0 \quad (k = (q-r)/s) \tag{2.1}$$

(in K) for these q, where

$$a_0 = \zeta_0\mathbf{v}^{\mathbf{r}_0 r}, \ldots, a_n = \zeta_n\mathbf{v}^{\mathbf{r}_n r}$$

and $\mathbf{m}_0, \ldots, \mathbf{m}_n$ are defined by

$$\mathbf{v}^{\mathbf{r}_0 s} = \mathbf{u}^{\mathbf{m}_0}, \ldots, \mathbf{v}^{\mathbf{r}_n s} = \mathbf{u}^{\mathbf{m}_n}.$$

As $k \to \infty$ in (2.1), this looks suspiciously like non-$(n+1)$-mixing (1.1), even with a non-$(n+1)$-mixing set $M = \{\mathbf{m}_0, \ldots, \mathbf{m}_n\}$. The only way out is that M has cardinality $h < n+1$. Writing $M = \{\mathbf{m}'_1, \ldots, \mathbf{m}'_h\}$ and I_j for the set of i with $\mathbf{m}_i = \mathbf{m}'_j$, we get a partition $I_1 \cup \cdots \cup I_h$ of $\{0, 1, \ldots, n\}$. Then

$$0 = \sum_{i=0}^{n} a_i\mathbf{u}^{\mathbf{m}_i k} = \sum_{j=1}^{h} b_j\mathbf{u}^{\mathbf{m}'_j k}$$

with $b_j = \sum_{i\in I_j} a_i$. Even then this looks like non-h-mixing; but this time the only way out is $b_1 = \cdots = b_h = 0$.

In particular, each I_j has cardinality at least two, and the quantities $\mathbf{v}^{\mathbf{r}_i s} = \mathbf{u}^{\mathbf{m}_i} = \mathbf{u}^{\mathbf{m}'_j}$ $(i \in I_j)$ are equal, so also the $\mathbf{r}_i$ $(i \in I_j)$, say to $\mathbf{m}''_j$. So, also the

$$\frac{\xi_i}{\zeta_i} = \mathbf{v}^{\mathbf{r}_i} = \mathbf{v}^{\mathbf{m}''_j} \quad (i \in I_j).$$

Thus, our point $(\xi_0, \ldots, \xi_n)$ lies in the corresponding transversal torsion coset Z. Also,

$$\sum_{i\in I_j} \zeta_i = \sum_{i\in I_j} a_i\mathbf{v}^{-\mathbf{r}_i r} = \mathbf{v}^{-\mathbf{m}''_j r} \sum_{i\in I_j} a_i = 0 \quad (j = 1, \ldots, h),$$

and this implies that Z lies in V_n. That completes the proof. □

If α is only n-mixing, then we cannot expect a conclusion as strong as that of Lemma 2. But the following is not too much weaker, where for $\pi = (\xi_0, \ldots, \xi_n)$ we write $\pi^l = (\xi_0^l, \ldots, \xi_n^l)$.

LEMMA 3. *Suppose for some $n \geq 2$ that α is n-mixing. Then there exist a finite collection $\mathcal{Z}$ of transversal torsion cosets Z in V_n and a finite set Π in $\mathbf{P}_n(\sqrt{G})$ such that*

$$V_n(\sqrt{G}) = \bigcup_{Z\in\mathcal{Z}} Z(\sqrt{G}) \cup \bigcup_{\pi\in\Pi} \bigcup_{e=0}^{\infty} \pi^{p^e}.$$

Proof. We apply the Descent Step (b) over $\sqrt{G}$ of [**DM1**, p. 1047] with ψ there as the identity and q there as p. Because V_n is not a coset, we obtain a finite collection $\mathcal{W}$ of proper $\sqrt{G}$-isotrivial linear subvarieties $W \neq V_n$ of V_n, also defined over K, such that

$$V_n(\sqrt{G}) = \bigcup_{W\in\mathcal{W}} \bigcup_{e=0}^{\infty} W(\sqrt{G})^{p^e}. \tag{2.2}$$

It will turn out that all the W here which are positive-dimensional can be taken as transversal torsion cosets.

Consider any W in (2.2), and pick a $\sqrt{G}$-isomorphism ψ with $\psi(W)$ defined over $\mathbf{F}$. We call W minimal if

$$W(\sqrt{G}) \neq \bigcup_{W' \in \mathcal{W}'} W'(\sqrt{G})$$

for a finite collection $\mathcal{W}'$ of $\sqrt{G}$-isotrivial linear subvarieties $W' \neq W$ of W with $\psi(W')$ defined over $\mathbf{F}$. If some W is not minimal, then we can replace it in (2.2) by lower-dimensional varieties. So, we can assume here that all W are minimal and of course that all $W(\sqrt{G})$ are non-empty.

Consider such a W with $\tilde{W} = \psi(W)$ defined over $\mathbf{F}$. Say

$$\psi(X_0, \ldots, X_n) = (g_0 X_0, \ldots, g_n X_n) = (\tilde{X}_0, \ldots, \tilde{X}_n)$$

for $g_0, \ldots, g_n$ in $\sqrt{G}$. We know from Lemma 1 that the ideal of $\tilde{W}$ is generated by the minimal forms. Let $\sum_{i \in I} \zeta_i \tilde{X}_i$ be one of these, with of course $\zeta_i \neq 0$ in $\mathbf{F}$. As $\dim \tilde{W} < \dim V_n = n - 1$, any n from $\tilde{X}_0, \ldots, \tilde{X}_n$ are dependent over $\mathbf{F}$ on $\tilde{W}$, and so I has cardinality at most n. Pick any $(\xi_0, \ldots, \xi_n)$ in $W(\sqrt{G})$. We get $\sum_{i \in I} \zeta_i g_i \xi_i = 0$. So, by Lemma 2 (in lower dimension), some $(\zeta_i g_i \xi_i)/(\zeta_j g_j \xi_j)$ $(i, j \in I, i \neq j)$ lies in $\mathbf{F}^*$, say ζ. It follows that

$$W(\sqrt{G}) = \bigcup_{ij\zeta} W_{ij\zeta}(\sqrt{G}),$$

where $W_{ij\zeta}$ is defined by $\zeta_i g_i X_i = \zeta \zeta_j g_j X_j$ in W. If all $W_{ij\zeta}$ here are $\neq W$, we contradict the minimality of W. So, some $W_{ij\zeta} = W$. This means that $\zeta_i \tilde{X}_i - \zeta \zeta_j \tilde{X}_j$ vanishes on $\tilde{W}$. Thus, that must have been the minimal form from the start.

Therefore, $\tilde{W}$ is a torsion coset. As it lies in the transversal $\psi(V_n)$ it must be itself transversal. Thus, it comes from a partition $I_1 \cup \cdots \cup I_h$ of $\{0, 1, \ldots, n\}$ into parts of size at least two together with $\tilde{\zeta}_i$ in $\mathbf{F}^*$, such that for each $j = 1, \ldots, h$ the equality of the quotients $\tilde{X}_i/\tilde{\zeta}_i$ $(i \in I_j)$ defines $\tilde{W}$. So, W is defined by the corresponding equality of the $X_i/\tilde{g}_i$ $(i \in I_j)$, where the $\tilde{g}_i = \tilde{\zeta}_i/g_i$ are still in $\sqrt{G}$.

Now the fact that W lies in V_n is easily seen to imply the equations

$$\sum_{i \in I_j} \tilde{g}_i = 0 \quad (j = 1, \ldots, h). \tag{2.3}$$

If $h = 1$, then of course W is the point $\pi = (\tilde{g}_0, \ldots, \tilde{g}_n)$, and we define Π as the finite set of points arising in this way.

So, we assume that $h \geq 2$ from now on. Now every sum in (2.3) involves at most n terms, so Lemma 2 is applicable as above. It yields a further partition $I_j = \bigcup I_{jk}$ into parts of size at least two together with more ζ_i' in $\mathbf{F}^*$ such that the $\tilde{g}_i/\zeta_i'$ $(i \in I_{jk})$ are equal. Further, just as in (2.3), we get $\sum_{i \in I_{jk}} \zeta_i' = 0$.

Consider now the linear variety Z_W defined by the equality of the X_i/ζ_i' $(i \in I_{jk})$ for each choice of j, k. It is a transversal torsion coset contained in V_n. It is not difficult to check that W lies in Z_W (but there might be more equations defining W, for example those connecting X_i in I_{jk} for different k). At any rate we have for each π in Π

$$W(\sqrt{G}) \cup \{\pi\} \subseteq Z_W(\sqrt{G}) \cup \{\pi\} \subseteq V_n(\sqrt{G}).$$

Raising to the power p^e and taking the union over all positive-dimensional W in $\mathcal{W}$, all π in Π, and $e = 0, 1, 2, \ldots$, we get $V_n(\sqrt{G})$ not only on the right but also on the left, by (2.2). It follows that this is also the middle term

$$\bigcup_W \bigcup_{e=0}^{\infty} Z_W(\sqrt{G})^{p^e} \cup \bigcup_{\pi\in\Pi} \bigcup_{e=0}^{\infty} \pi^{p^e}.$$

Now this completes the proof since each Z_W has only a finite set $\mathcal{Z}_W$ of conjugates over $\mathbf{F}_p$, and

$$\bigcup_{e=0}^{\infty} Z_W(\sqrt{G})^{p^e} = \bigcup_{Z\in\mathcal{Z}_W} Z(\sqrt{G}). \qquad \square$$

The following observation is crucial to get information about the coefficients arising from non-mixing. It seems convenient to work affinely for a bit. Recall from [**M**, p. 189] that a set Σ in $\sqrt{G}^n$ is called broad if:

(i) Σ is infinite;

(ii) for each g in $\sqrt{G}$ and each $i = 1, \ldots, n$, there are at most finitely many $(x_1, \ldots, x_n)$ in Σ with $x_i = g$;

(iii) if $n \geq 2$ then for each g in $\sqrt{G}$ and each i, $j = 1, \ldots, n$ with $i \neq j$ there are at most finitely many $(x_1, \ldots, x_n)$ in Σ with $x_i/x_j = g$.

LEMMA 4. *Suppose for some $n \geq 2$ that α is n-mixing, and that there exist $a_1, \ldots, a_n$ in K such that the equation $a_1x_1 + \cdots + a_nx_n = 1$ has a broad set of solutions in $(\sqrt{G})^n$. Then $a_1, \ldots, a_n$ lie in $\sqrt{G}$.*

Proof. The $a_1, \ldots, a_n$ lie in K^*, otherwise we would have non-n-mixing. This can be seen by writing the solutions as $x_i = \zeta_i \mathbf{v}^{\mathbf{m}_i}$ $(i = 1, \ldots, n)$ for ζ_i in $\mathbf{F}^*$ and then getting into G by reducing the exponents modulo s according to $\mathbf{m}_i = \mathbf{m}_{i0} + s\mathbf{q}_i$; then $\mathbf{v}^{s\mathbf{q}_i} = \mathbf{u}^{\mathbf{p}_i}$. If say $a_n = 0$ we get equations $a'_1\mathbf{u}^{\mathbf{p}_1} + \cdots + a'_{n-1}\mathbf{u}^{\mathbf{p}_{n-1}} = 1$ (in K) and now these give n-mixing (see [**S**, p. 263]) unless some $\mathbf{p}_i$ or some $\mathbf{p}_i - \mathbf{p}_j$ $(i \neq j)$ does not tend to infinity. But this would contradict the broadness.

Therefore, $a_1, \ldots, a_n$ satisfy the hypotheses of [**M**, Lemma 5, p. 197] with $\sqrt{G}$ in place of G.

Thus, either (aa) or (bb) of Lemma 5 holds. But (aa) would also lead to non-n-mixing. So, (bb) holds.

Now apply Lemma 5 to the new equation (16) of [**M**, p. 198]. Again (bb) must hold.

And so on for ever. By [**M**, Lemma 2, p. 193], this means that $a_1, \ldots, a_n$ must lie in $\sqrt{G}$. This proves the present lemma. $\square$

We call the projective $(\xi_0, \ldots, \xi_n)$ with non-zero coordinates pre-broad if no ξ_i/ξ_j $(i \neq j)$ lies in $\mathbf{F}^*$. We call two such points π, $\pi^{\#}$ proportional if there are positive integers l, $l^{\#}$ with $\pi^{l^{\#}} = \pi^{\#l}$.

We note that the equation defining V_n is invariant under the symmetric group S_{n+1} on $n + 1$ elements, so that this also acts on points of V_n. It also acts on proportionality classes.

LEMMA 5. *Suppose for some $n \geq 2$ that α is n-mixing but not $(n+1)$-mixing. Then there exist a finite collection $\mathcal{Z}$ of transversal torsion cosets Z in V_n, and a finite set Π in $\mathbf{P}_n(\sqrt{G})$, containing at least one pre-broad element, such that*

$$V_n(\sqrt{G}) = \bigcup_{Z \in \mathcal{Z}} Z(\sqrt{G}) \cup \bigcup_{\pi \in \Pi} \bigcup_{e=0}^{\infty} \pi^{p^e}.$$

Further, α has only finitely many equivalence classes of non-$(n+1)$-mixing sets, and these are in one-to-one correspondence with the S_{n+1}-orbits of the proportionality classes of the pre-broad π in Π.

Proof. Because α is non-$(n+1)$-mixing, there certainly exist non-$(n+1)$-mixing sets. Pick any such set. It is equivalent to some $\{0, \mathbf{m}_1, \ldots, \mathbf{m}_n\}$ in $\mathbf{Z}^d$ and then there are $a_1, \ldots, a_n$ in K and an infinite set of positive integers k such that

$$a_1 \mathbf{u}^{\mathbf{m}_1 k} + \cdots + a_n \mathbf{u}^{\mathbf{m}_n k} = 1$$

(in K). As there is 2-mixing, the $u_1, \ldots, u_d$ are multiplicatively independent in K, and in particular the hypotheses of Lemma 4 are satisfied. It follows that $a_1, \ldots, a_n$ lie in $\sqrt{G}$.

With the basis elements $v_1, \ldots, v_d$ of $\sqrt{G}$ modulo torsion, we can write $a_i = \zeta_i \mathbf{v}^{\mathbf{p}_i}$ $(i = 1, \ldots, n)$ and

$$u_h = \zeta_h' \mathbf{v}^{\mathbf{q}_h} \quad (h = 1, \ldots, d) \tag{2.4}$$

for ζ_i, ζ_h' in $\mathbf{F}^*$. Putting the rows $\mathbf{q}_1, \ldots, \mathbf{q}_d$ together to make an invertible integral matrix $\mathbf{Q}$, so that $\mathbf{u}^{\mathbf{m}} = \zeta_{\mathbf{m}} \mathbf{v}^{\mathbf{mQ}}$ for $\zeta_{\mathbf{m}}$ in $\mathbf{F}^*$, we obtain the points

$$(\xi_0^{(k)}, \xi_1^{(k)}, \ldots, \xi_n^{(k)}) = \pi^{(k)} = (-1, \zeta_1^{(k)} \mathbf{v}^{\mathbf{p}_1 + \mathbf{m}_1 \mathbf{Q} k}, \ldots, \zeta_n^{(k)} \mathbf{v}^{\mathbf{p}_n + \mathbf{m}_n \mathbf{Q} k})$$

on V_n as in Lemma 3, with $\zeta_i^{(k)}$ in $\mathbf{F}^*$. We are going to prove that there cannot exist two different k, k' such that $\pi^{(k)}$, $\pi^{(k')}$ lie in the same Z.

If $\pi^{(k)}$ lies in Z, then some $\xi_i^{(k)}/\xi_j^{(k)}$ $(i \neq j)$ would be in $\mathbf{F}^*$. If for example $i \neq 0$, $j \neq 0$, then this implies that $(\zeta_i^{(k)} \mathbf{v}^{\mathbf{p}_i + \mathbf{m}_i \mathbf{Q} k})/(\zeta_j^{(k)} \mathbf{v}^{\mathbf{p}_j + \mathbf{m}_j \mathbf{Q} k})$ lies in $\mathbf{F}^*$ and so $\mathbf{p}_i + \mathbf{m}_i \mathbf{Q} k = \mathbf{p}_j + \mathbf{m}_j \mathbf{Q} k$, that is, $\mathbf{p}_i - \mathbf{p}_j = -(\mathbf{m}_i - \mathbf{m}_j)\mathbf{Q} k$. Writing the same equation for $k' \neq k$ and subtracting gives a contradiction as $\mathbf{m}_i \neq \mathbf{m}_j$. A similar argument works if $i = 0$ or $j = 0$ (with $\mathbf{m}_0 = 0$).

Thus, for all sufficiently large k, the points $\pi^{(k)}$ must be the π^q for π in Π and $q = p^e$ $(e = 0, 1, 2, \ldots,)$. So, we can find two different k, k' and two q, q' with $\pi^{(k)} = \pi^q$, $\pi^{(k')} = \pi^{q'}$ for the same $\pi = (-1, \xi_1, \ldots, \xi_n)$ in Π. Writing $\xi_i = \tilde{\zeta}_i \mathbf{v}^{\mathbf{r}_i}$ for $\tilde{\zeta}_i$ in $\mathbf{F}^*$, we get as above

$$\mathbf{p}_i + \mathbf{m}_i \mathbf{Q} k = \mathbf{r}_i q, \quad \mathbf{p}_i + \mathbf{m}_i \mathbf{Q} k' = \mathbf{r}_i q' \quad (i = 1, \ldots, n).$$

Thus,

$$\mathbf{m}_i \mathbf{Q}(k - k') = \mathbf{r}_i (q - q') \quad (i = 1, \ldots, n). \tag{2.5}$$

In particular, $q \neq q'$, and so our set $\{0, \mathbf{m}_1, \ldots, \mathbf{m}_n\}$ is equivalent to

$$\pm\{0, \mathbf{r}_1, \ldots, \mathbf{r}_n\}\mathbf{Q}^{-1}. \tag{2.6}$$

As $\mathbf{Q}$ is fixed and there are only finitely many possibilities for $\{0, \mathbf{r}_1, \ldots, \mathbf{r}_n\}$ corresponding to the finite set Π, the finiteness assertion in the present lemma for non-mixing sets follows.

The existence of some pre-broad π in Π also follows, because the point

$$\pi = (-1, \xi_1, \ldots, \xi_n) = (-1, \tilde{\zeta}_1 \mathbf{v}^{\mathbf{r}_1}, \ldots, \tilde{\zeta}_n \mathbf{v}^{\mathbf{r}_n})$$

is pre-broad if and only if the set $\{0, \mathbf{r}_1, \ldots, \mathbf{r}_n\}$ has cardinality $n+1$.

But to prove the one-to-one assertion we must tighten things up a bit.

We first show how to eliminate the minus possibility in (2.6). We can suppose that the k, k' above are just two elements of an infinite set. We fix k and then make k' tend to infinity. Using heights as in [**M**, **DM1**], we can easily see (from $\mathbf{m}_n \neq 0$, for example) that the height of $\pi^{(k')}$ tends to infinity. Thus, the corresponding q' tends to infinity. Therefore, we can assume that $k < k'$ and $q < q'$. So, indeed, we can improve (2.6) to $\{0, \mathbf{r}_1, \ldots, \mathbf{r}_n\}\mathbf{Q}^{-1}$. □

We described above how a non-mixing set M gives rise to π. Suppose that two such sets M, $M^{\#}$ give rise to π, $\pi^{\#}$ in the same S_{n+1}-orbit of proportionality classes. We show that M, $M^{\#}$ are equivalent.

Say $\pi^{\#l} = ((-1)^l, \tilde{\zeta}_1^{\#l} \mathbf{v}^{\mathbf{r}_1^{\#} l}, \ldots, \tilde{\zeta}_n^{\#l} \mathbf{v}^{\mathbf{r}_n^{\#} l})$ comes, for example, from permuting the first two coordinates of $\pi^{l^{\#}} = ((-1)^{l^{\#}}, \tilde{\zeta}_1^{l^{\#}} \mathbf{v}^{\mathbf{r}_1 l^{\#}}, \ldots, \tilde{\zeta}_n^{l^{\#}} \mathbf{v}^{\mathbf{r}_n l^{\#}})$. Then

$$\tilde{\zeta}_1^{\#l} \mathbf{v}^{\mathbf{r}_1^{\#} l} = (-1)^{l^{\#} - l} \tilde{\zeta}_1^{-l^{\#}} \mathbf{v}^{-\mathbf{r}_1 l^{\#}},$$

$$\tilde{\zeta}_i^{\#l} \mathbf{v}^{\mathbf{r}_i^{\#} l} = (-1)^{-l} \tilde{\zeta}_1^{-l^{\#}} \mathbf{v}^{-\mathbf{r}_1 l^{\#}} \tilde{\zeta}_i^{l^{\#}} \mathbf{v}^{\mathbf{r}_i l^{\#}} \quad (i = 2, \ldots, n).$$

In particular,

$$\mathbf{r}_1^{\#} l = -\mathbf{r}_1 l^{\#},$$

$$\mathbf{r}_i^{\#} l = -\mathbf{r}_1 l^{\#} + \mathbf{r}_i l^{\#} \quad (i = 2, \ldots, n).$$

Thus, looking at the improved (2.6), we see that $M^{\#}$ is equivalent to

$$\{0, -\mathbf{r}_1, -\mathbf{r}_1 + \mathbf{r}_2, \ldots, -\mathbf{r}_1 + \mathbf{r}_n\}\mathbf{Q}^{-1}$$

in turn equivalent to $\{\mathbf{r}_1, 0, \mathbf{r}_2, \ldots, \mathbf{r}_n\}\mathbf{Q}^{-1}$ and so to M.

A similar argument works for any permutation, and thus the number of classes of non-mixing sets is at most the number of orbits.

To prove the opposite inequality, we note as in the proof of Lemma 2 that a pre-broad $\pi = (-1, \tilde{\zeta}_1 \mathbf{v}^{\mathbf{r}_1}, \ldots, \tilde{\zeta}_n \mathbf{v}^{\mathbf{r}_n})$ gives rise to a potential non-mixing set via

$$\tilde{\zeta}_1 \mathbf{v}^{\mathbf{r}_1 q} + \cdots + \tilde{\zeta}_n \mathbf{v}^{\mathbf{r}_n q} = 1$$

and $k = (q - r)/s$ to get

$$a_1 \mathbf{u}^{\mathbf{m}_1 k} + \cdots + a_n \mathbf{u}^{\mathbf{m}_n k} = 1$$

(in K) for

$$a_i = \tilde{\zeta}_i \mathbf{v}^{\mathbf{r}_i r}, \quad \mathbf{u}^{\mathbf{m}_i} = \mathbf{v}^{\mathbf{r}_i s} \quad (i = 1, \ldots, n).$$

Using (2.4), we see that $\mathbf{u}^{\mathbf{m}} = \zeta_{\mathbf{m}} \mathbf{v}^{\mathbf{mQ}}$ for some torsion $\zeta_{\mathbf{m}}$, and it follows that

$$\mathbf{m}_i \mathbf{Q} = \mathbf{r}_i s \quad (i = 1, \ldots, n) \tag{2.7}$$

consistent with (2.5). As π is pre-broad, the set $M = \{0, \mathbf{m}_1, \ldots, \mathbf{m}_n\}$ has cardinality $n+1$ and is therefore indeed non-mixing.

As above, it is now rather easy to see that if π, $\pi^\#$ give rise to equivalent M, $M^\#$ then they are in the same orbit, so we get the desired opposite inequality. Here it is convenient to note that any π is in the same orbit as some power $\pi^l = (1, \mathbf{v}^{\mathbf{s}_1}, \ldots, \mathbf{v}^{\mathbf{s}_n})$ and so the roots of unity play no role.

3. *Proof of theorem*

Because $n \geq 2$, our $\alpha = \alpha^{\mathcal{R}/\mathcal{P}}$ is 2-mixing. In the notation of the previous section, we look first at $V_2(\sqrt{G})$. From Lemma 3, we get $\mathcal{Z}$ and Π.

If some $\pi = (\xi_0, \xi_1, \xi_2)$ in Π is pre-broad, then α is non-3-mixing, else Lemma 2 would show that this π lies in some transversal torsion coset, forcing some ξ_i/ξ_j $(i \neq j)$ in $\mathbf{F}^*$. So, 3 is the smallest order of non-mixing, and by Lemma 5 there are only finitely many classes.

Otherwise no π in Π is pre-broad, and now Lemma 5 shows that α must be 3-mixing. We then jump to $V_3(\sqrt{G})$ and repeat the process. Eventually we must find some pre-broad point in some Π corresponding to some $V_n(\sqrt{G})$, and this leads to non-$(n+1)$-mixing. This $n+1$ is the required smallest order of non-mixing. And, as explained in §1, an *a priori* upper bound can be found in the usual way simply by taking any non-zero polynomial in $\mathcal{P}$. This completes the proof of the theorem.

The effectivity follows at once from the effectivity of [**DM1**].

4. *An example*

Before starting with this, we consider briefly the original Ledrappier example [**Led**], which is 2-mixing but not 3-mixing. It corresponds to $\mathcal{P}$ generated by 2 and $1 + u_1 + u_2$ in $\mathcal{R} = \mathbf{Z}[u_1, u_1^{-1}, u_2, u_2^{-1}]$. The group G has generators the images of u_1, u_2 in the quotient field K of $\mathcal{R}/\mathcal{P}$. We may identify K with $\mathbf{F}_2(t)$ and the generators with t, $1+t$, respectively. As these are clearly multiplicatively independent, we see already that α is 2-mixing. Equally clearly, $\sqrt{G}$ has generators t, $1+t$. To go further, we need the field $C = \mathbf{F}_2(t^2)$ of differential constants.

Now Leitner in [**Lei**, Theorem 1, p. 327] showed that

$$V_2(\sqrt{G}) = \bigcup_{\pi \in S_3(\pi_0)} \bigcup_{e=0}^{\infty} \pi^{2^e}$$

for $\pi_0 = (1, t, 1+t)$. Thus, we see at once from Lemma 5 and (2.7) that there is exactly one class of non-mixing sets of order 3, with representative $\{(0,0), (1,0), (0,1)\}$. See also Lemma 5.6 (p. 348) of the paper [**ABB**] of Arenas-Carmona *et al* (which is however more concerned with higher order mixing for Ledrappier away from this shape).

It might be fun to try $\mathcal{P}$ generated by 2 and $1 + u_1 + u_2 + u_3$ in $\mathbf{Z}[u_1, u_1^{-1}, u_2, u_2^{-1}, u_3, u_3^{-1}]$. But perhaps one should glance at §5 before starting. And one would have to work with two variables t, t'.

We return to the example of [**S**, p. 278]. Here $\mathcal{P}$ is generated by 2 and (1.4) in $\mathcal{R} = \mathbf{Z}[u_1, u_1^{-1}, u_2, u_2^{-1}]$. As already remarked, the factor $1 + u_1 + u_1^2 + u_2^{1/3}$ in $\mathbf{F}_2[u_1, u_2^{1/3}]$

shows that α is not 4-mixing, because in the quotient we have

$$0 = 1 + u_1^{2^e} + u_1^{2.2^e} + u_2^{2^e/3} = 1 + a_1 u_1^{3k} + a_2 u_1^{6k} + a_3 u_2^k$$

for all even e, with $a_1 = u_1$, $a_2 = u_1^2$, $a_3 = u_2^{1/3} = 1 + u_1 + u_1^2$ in K and $k = (2^e - 1)/3$; and so the non-mixing set is

$$\{(0, 0), (3, 0), (6, 0), (0, 1)\}. \tag{4.1}$$

We will prove here that α is 3-mixing and that there are exactly five equivalence classes for non-mixing sets of size 4.

The group G has generators the images of u_1, u_2 in the quotient field K of $\mathcal{R}/\mathcal{P}$. We may again identify K with $\mathbf{F}_2(t)$ and the generators with t, $(1 + t + t^2)^3$, respectively. As these are clearly multiplicatively independent, we see already that α is 2-mixing. It is easy to see that $\sqrt{G}$ has generators t, $1 + t + t^2$.

Already by Lemma 5 above with $n = 2$, the next lemma shows that α is 3-mixing.

LEMMA 6. *The set $V_2(\sqrt{G})$ is empty.*

Proof. It suffices to deduce a contradiction from the existence of x and y in $\sqrt{G}$ with $x + y = 1$. We follow the methods of [**Lei**].

Assume first that the C-vector space $Cx + Cy$ has dimension 2. Using a dot to indicate the derivative with respect to t, we deduce that $\dot{y}/y \neq \dot{x}/x$. We get in the usual way the identities

$$x = \frac{\dot{y}/y}{\dot{y}/y - \dot{x}/x}, \quad y = \frac{-\dot{x}/x}{\dot{y}/y - \dot{x}/x}. \tag{4.2}$$

Now, if $z = t^a(1 + t + t^2)^b$ is a typical element of $\sqrt{G}$, then

$$\frac{\dot{z}}{z} = \frac{a}{t} + \frac{b}{1 + t + t^2}$$

takes just four values

$$0, \quad \frac{1}{t}, \quad \frac{1}{1 + t + t^2}, \quad \frac{(1 + t)^2}{t(1 + t + t^2)}. \tag{4.3}$$

Since $\dot{y}/y - \dot{x}/x$ in (4.2) is $\dot{z}/z$ for $z = y/x$, it follows that x and y are non-zero quotients of these. But the presence of the 'stranger' $(1 + t)^2$ means that the only possibilities for $x \neq 1$ and $y \neq 1$ are

$$\frac{1 + t + t^2}{t}, \quad \frac{t}{1 + t + t^2}.$$

However, then $x + y \neq 1$.

If $Cx + Cy$ has dimension 1, then as $x + y = 1$ we see that x and y lie in C. There is a biggest power q of 2 with $x = x'^q$ and $y = y'^q$ for x' and y' not both in C. Now $x' + y' = 1$ with x' and y' still in $\sqrt{G}$, and $Cx' + Cy'$ has dimension 2; but we have just seen this to be impossible. Thus, the present lemma is proved. □

To go further, we welcome the above stranger into the bigger group H with generators t, $1 + t + t^2$, $(1 + t)^2$. Here $\sqrt{H}$ has generators t, $1 + t + t^2$, $1 + t$.

LEMMA 7. *We have*

$$V_2(\sqrt{H}) = \bigcup_{\pi \in S_3(\Pi_2)} \bigcup_{e=0}^{\infty} \pi^{2^e}$$

for the set Π_2 *consisting of*

$$(1, t, 1+t), \quad (1, t(1+t), 1+t+t^2),$$
$$(1, t^3, (1+t)(1+t+t^2)), \quad (1, (1+t)^3, t(1+t+t^2)),$$
$$(t, (1+t)^2, 1+t+t^2), \quad (t^2, 1+t, 1+t+t^2), \quad (t^3, (1+t)^3, 1+t+t^2).$$

Proof. Again it suffices to consider x and y in $\sqrt{H}$ with $x + y = 1$.

Assume first that $Cx + Cy$ has dimension 2. Now, if $z = t^a(1+t+t^2)^b(1+t)^c$ is a typical element of $\sqrt{H}$, then

$$\frac{\dot{z}}{z} = \frac{a}{t} + \frac{b}{1+t+t^2} + \frac{c}{1+t}.$$

These are the elements in (4.3) together with their sums with $1/(1+t)$; that is,

$$\frac{1}{1+t}, \quad \frac{1}{t(1+t)}, \quad \frac{t^2}{(1+t)(1+t+t^2)}, \quad \frac{1}{t(1+t)(1+t+t^2)}. \tag{4.4}$$

Therefore, x and y are non-zero quotients of elements of (4.3) and (4.4). This time we find no strangers; and in fact each of the possible 42 values for x in $\sqrt{H}$ leads also to $y = 1 - x$ in $\sqrt{H}$. We verify without difficulty that the resulting 42 solutions $(x, y, 1)$ fall in seven orbits under S_3, as stated in the present lemma. As in the proof of Lemma 6, the case of dimension 1 supplies the exponents 2^e, and this completes the proof. □

Next we move to V_3. We define the torsion coset Z_{01} in V_3 by $x_0 = x_1$, $x_2 = x_3$.

LEMMA 8. *We have*

$$V_3(\sqrt{G}) = \bigcup_{Z \in S_4(Z_{01})} Z(\sqrt{G}) \cup \bigcup_{\pi \in S_4(\Pi_3)} \bigcup_{e=0}^{\infty} \pi^{2^e} \tag{4.5}$$

for the set Π_3 *consisting of*

$$(1, t, t^2, 1+t+t^2), \quad (1, t^3, t^2(1+t+t^2), (1+t+t^2)^2),$$
$$(1, t^3, 1+t+t^2, t(1+t+t^2)),$$
$$(t, t^4, 1+t+t^2, (1+t+t^2)^2), \quad (1, t^6, t(1+t+t^2)^2, (1+t+t^2)^3).$$

Proof. Take a point (x_0, x_1, x_2, x_3) of $V_3(\sqrt{G})$ not in any $Z(\sqrt{G})$ in (4.5). Write d for the dimension of $Cx_0 + Cx_1 + Cx_2 + Cx_3$ over C. Since $[K : C] = 2$, we have $d = 1$ or $d = 2$.

Assume first that $d = 2$. We will prove that our point lies in $S_4(\Pi_3)$.

To this end, define i, j to be equivalent if x_i/x_j lies in C. We show that at least one of the classes is a singleton. This is clear if the number h of classes is 4 or even 3. As $d = 2$, we cannot have $h = 1$. So, assume that $h = 2$ and there is no singleton.

Now both classes must have two elements. But if say x_0/x_1 and x_2/x_3 lie in C, then neither quotient can be 1 and the identity

$$\frac{x_0}{x_3} = \frac{1 + x_2/x_3}{1 + x_1/x_0}$$

shows that x_0/x_3 lies in C. Thus, $h = 1$, a contradiction which shows that there must indeed be a singleton.

We can assume that this singleton consists of x_3. That means that

$$y_i = \frac{\dot{x}_i}{x_i} - \frac{\dot{x}_3}{x_3} \neq 0 \quad (i = 0, 1, 2).$$

Further, each y_i is itself a logarithmic derivative of something in $\sqrt{G}$, and so it lies in a finite subset of $\sqrt{H}$ by (4.3). Also, the equation

$$y_0 x_0 + y_1 x_1 + y_2 x_2 = 0$$

follows from $x_0 + x_1 + x_2 + x_3 = 0$ and its derivative. This remark is in fact a condensed version of the arguments of [**M**, pp. 198 and 199]. Therefore, we have a point of $V_2(\sqrt{H})$.

Thus, there are $q = 2^e$ and $\pi = (\xi_0, \xi_1, \xi_2)$ as in Lemma 7 such that

$$\frac{y_1 x_1}{y_0 x_0} = \left(\frac{\xi_1}{\xi_0}\right)^q, \quad \frac{y_2 x_2}{y_0 x_0} = \left(\frac{\xi_2}{\xi_0}\right)^q.$$

Already this leads to an algorithm for finding our point of $V_3(\sqrt{G})$. Namely, for each π, we know from Lemma 6 that not both of ξ_1/ξ_0, ξ_2/ξ_0 lie in $\sqrt{G}$. So, there is at most one q such that

$$\frac{x_1}{x_0} = \frac{y_0}{y_1}\left(\frac{\xi_1}{\xi_0}\right)^q, \quad \frac{x_2}{x_0} = \frac{y_0}{y_2}\left(\frac{\xi_2}{\xi_0}\right)^q \tag{4.6}$$

both lie in $\sqrt{G}$. Further, we can easily see by considering powers of $1 + t$ that $q = 1, 2$. For example, by Lemma 6, at least one of ξ_1/ξ_0, ξ_2/ξ_0 must involve $1 + t$; but then by (4.3) a resulting $(1 + t)^4$ could not be cancelled in (4.6). If there is such a q, then we need only check whether

$$\frac{x_3}{x_0} = 1 + \frac{x_1}{x_0} + \frac{x_2}{x_0}$$

also lies in $\sqrt{G}$. Here we are still allowed to permute x_0, x_1, x_2 and so we can use the symmetry to reduce the work by a factor of six.

The case $d = 1$ is dealt with as in the proof of Lemma 6 by reducing to $x + y + z = 1$ and using $x = x'^q$, $y = y'^q$ and now $z = z'^q$.

All this means that the left-hand side of (4.5) is contained in the right-hand side. As the converse assertion is quickly checked, this completes the proof. □

Now, thanks to the one-to-one assertion in Lemma 5, we can find all the non-mixing sets of size 4. As Π_3 has five pre-broad elements, all in different proportionality classes, there are five equivalence classes. We can identify generators v_1, v_2 of $\sqrt{G}$ with t, $1 + t + t^2$

respectively (so that $\mathbf{Q} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 3 \end{smallmatrix}\right)$ in the proof). We find the integral representatives

$$\{(0,0),(3,0),(6,0),(0,1)\},$$
$$\{(0,0),(9,0),(6,1),(0,2)\},$$
$$\{(0,0),(9,0),(0,1),(3,1)\},$$
$$\{(3,0),(12,0),(0,1),(0,2)\},$$
$$\{(0,0),(18,0),(3,2),(0,3)\},$$

the first of which appears in [**S**, p. 278] and (4.1) above.

5. *Another example*

Here we deal with a non-principal ideal, of which there are no examples in the mixing [**S**, Ch. 28]. It is the $\mathcal{P}$ generated by 2 and (1.5) in $\mathcal{R} = \mathbf{Z}[u_1, u_1^{-1}, u_2, u_2^{-1}, u_3, u_3^{-1}]$. Each of the displayed generators shows that the corresponding α is not 4-mixing by providing the non-mixing sets

$$\{(0,0,0),(1,0,0),(2,0,0),(0,1,0)\}, \quad \{(0,0,0),(1,0,0),(3,0,0),(0,0,1)\}. \tag{5.1}$$

We will prove here that α is 3-mixing and that there are exactly 134 equivalence classes for non-mixing sets of size 4.

The group G has generators the images of u_1, u_2, u_3 in the quotient field K of $\mathcal{R}/\mathcal{P}$. We may identify K with $\mathbf{F}_2(t)$ and the generators with t, $1+t+t^2$, $1+t+t^3$, respectively. As these irreducible polynomials are clearly multiplicatively independent, we see already that α is 2-mixing. It is easy to see that $G = \sqrt{G}$. To go further, we need again the field $C = \mathbf{F}_2(t^2)$ of differential constants.

Already by Lemma 5 above with $n = 2$, the next lemma shows that α is 3-mixing.

LEMMA 9. *The set $V_2(\sqrt{G})$ is empty.*

Proof. It suffices to deduce a contradiction from the existence of x and y in $\sqrt{G}$ with $x + y = 1$.

Assume first that the C-vector space $Cx + Cy$ has dimension 2. We get again (4.2). Now, if $z = t^a(1+t+t^2)^b(1+t+t^3)^c$ is a typical element of $\sqrt{G}$, then

$$\frac{\dot{z}}{z} = \frac{a}{t} + \frac{b}{1+t+t^2} + \frac{c(1+t^2)}{1+t+t^3}$$

takes eight values, which are

$$0, \quad \frac{1}{t}, \quad \frac{1}{1+t+t^2}, \quad \frac{(1+t)^2}{t(1+t+t^2)} \tag{5.2}$$

as in (4.3) together with

$$\frac{(1+t)^2}{1+t+t^3}, \quad \frac{1}{t(1+t+t^3)}, \quad \frac{t^4}{(1+t+t^2)(1+t+t^3)}, \quad \frac{(1+t)^4}{t(1+t+t^2)(1+t+t^3)}. \tag{5.3}$$

The presence of strangers leads now to 14 possibilities for quotients $x \neq 0, 1$ in $\sqrt{G}$. However, it is quickly checked that then $y = 1 + x$ is not among them. □

The case of dimension 1 follows just as in the proof of Lemma 6.

To go further, we need the bigger group H with generators t, $1+t+t^2$, $1+t+t^3$, $(1+t)^2$. Here $\sqrt{H}$ has generators t, $1+t+t^2$, $1+t+t^3$, $1+t$. For an element $\mathbf{e}=(a,b,c,d)$ of $\mathbf{F}_2^4$, we write

$$P(\mathbf{e})=t(1+t+t^2)(1+t+t^3)(1+t)\left(\frac{a}{t}+\frac{b}{1+t+t^2}+\frac{c(1+t^2)}{1+t+t^3}+\frac{d}{1+t}\right) \quad (5.4)$$

in $\mathbf{F}_2[t]$. Write $\mathbf{0}=(0,0,0,0)$ and $\mathbf{1}=(1,1,1,1)$.

LEMMA 10. *We have*

$$V_2(\sqrt{H})=\bigcup_{\pi\in\Pi}\bigcup_{e=0}^{\infty}\pi^{2^e}$$

for the set Π *consisting of the 168 elements* $(P(\mathbf{e}_0),\ P(\mathbf{e}_1),\ P(\mathbf{e}_2))$ *with*

$$\mathbf{e}_0+\mathbf{e}_1+\mathbf{e}_2=\mathbf{0},\quad \mathbf{e}_0\neq\mathbf{0},\mathbf{1},\quad \mathbf{e}_1\neq\mathbf{0},\mathbf{1},\quad \mathbf{e}_2\neq\mathbf{0},\mathbf{1}.$$

Proof. Now, if $z=t^a(1+t+t^2)^b(1+t+t^3)^c(1+t)^d$ is a typical element of $\sqrt{H}$, then

$$\frac{\dot z}{z}=\frac{a}{t}+\frac{b}{1+t+t^2}+\frac{c(1+t^2)}{1+t+t^3}+\frac{d}{1+t}$$

as in (5.4). And (4.2) shows that projectively $(x,\ y,\ 1)$ is $(\dot y/y,\ -\dot x/x,\ \dot y/y-\dot x/x)$, so after multiplication of all coordinates by $t(1+t)(1+t+t^2)(1+t+t^3)$ we get $\pi=(P(\mathbf{e}_0),\ P(\mathbf{e}_1),\ P(\mathbf{e}_2))$ with $\mathbf{e}_0+\mathbf{e}_1+\mathbf{e}_2=\mathbf{0}$. We certainly have to avoid $P(\mathbf{0})=0$, but all other $P(\mathbf{e})$ turn out to be in $\sqrt{H}$ with the single exception of $P(\mathbf{1})=(1+t^2+t^3)^2$. Then we check that the resulting π are all different (however that is not crucial to the rest of the argument). This completes the proof. □

For the move to V_3, we use as above the torsion coset Z_{01} defined by $x_0=x_1,\ x_2=x_3$. But if we had known the outcome we might have never started on this example.

LEMMA 11. *We have*

$$V_3(\sqrt{G})=\bigcup_{Z\in S_4(Z_{01})}Z(\sqrt{G})\cup\bigcup_{\pi\in S_4(\Pi_3)}\bigcup_{e=0}^{\infty}\pi^{2^e}$$

for a set Π_3 *consisting of 134 elements containing*

$$(1,t,t^2,1+t+t^2),\quad (1,t,t^3,1+t+t^3),$$
$$(t^2,t^3,1+t+t^2,1+t+t^3),\quad (t,1+t+t^2,t(1+t+t^2),1+t+t^3)$$

through to the Baby Gremlin

$$(t^{21}(1+t+t^3),t^{20}(1+t+t^2),(1+t+t^2)^{12},(1+t+t^3)^4)$$

and the Gremlin

$$(t^{25},t^{20}(1+t+t^2)(1+t+t^3),(1+t+t^2)^{12},(1+t+t^3)^4).$$

Proof. As in the proof of Lemma 8, we take a point (x_0, x_1, x_2, x_3) of $V_3(\sqrt{G})$ not in any $Z(\sqrt{G})$ with $d = 1$ or $d = 2$ for the dimension of $Cx_0 + Cx_1 + Cx_2 + Cx_3$ over C.

Assume first that $d = 2$. Just as in the proof of Lemma 8, we can assume that

$$y_i = \frac{\dot{x}_i}{x_i} - \frac{\dot{x}_3}{x_3} \neq 0 \quad (i = 0, 1, 2).$$

Further, each y_i is itself a logarithmic derivative of something in $\sqrt{G}$, and so it lies in a finite subset of $\sqrt{H}$ by (5.2) and (5.3). We also get the equation

$$y_0x_0 + y_1x_1 + y_2x_2 = 0$$

and so a point of $V_2(\sqrt{H})$.

Thus, there are $q = 2^e$ and $\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2$ satisfying the conditions of Lemma 10 such that

$$\frac{y_1x_1}{y_0x_0} = \left(\frac{P(\mathbf{e}_1)}{P(\mathbf{e}_0)}\right)^q, \quad \frac{y_2x_2}{y_0x_0} = \left(\frac{P(\mathbf{e}_2)}{P(\mathbf{e}_0)}\right)^q.$$

Already this leads to an algorithm for finding our point of $V_3(\sqrt{G})$. Namely, for each $\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2$, we know from Lemma 9 that not both of $P(\mathbf{e}_1)/P(\mathbf{e}_0)$, $P(\mathbf{e}_2)/P(\mathbf{e}_0)$ lie in $\sqrt{G}$. So, there is at most one q such that

$$\frac{x_1}{x_0} = \frac{y_0}{y_1}\left(\frac{P(\mathbf{e}_1)}{P(\mathbf{e}_0)}\right)^q, \quad \frac{x_2}{x_0} = \frac{y_0}{y_2}\left(\frac{P(\mathbf{e}_2)}{P(\mathbf{e}_0)}\right)^q \tag{5.5}$$

both lie in $\sqrt{G}$. Further, we can easily see that $q = 1, 2, 4$. For example, by Lemma 9, at least one of $P(\mathbf{e}_1)/P(\mathbf{e}_0)$, $P(\mathbf{e}_2)/P(\mathbf{e}_0)$ must involve $1 + t$; but then by (5.2) and (5.3) a resulting $(1 + t)^8$ could not be cancelled in (5.5). If there is such a q, then we need only check whether

$$\frac{x_3}{x_0} = 1 + \frac{x_1}{x_0} + \frac{x_2}{x_0}$$

also lies in $\sqrt{G}$. This was originally carried out in 2010 by means of an interactive procedure on Maple. With mounting horror we realized that the many solutions were not obligingly organizing themselves into a few classes. After 20 hours we drew up a list of representative solutions numbered from 1 to the Eddingtonian 137. But Alexandre Warin in his 2012 Master Thesis observed that seven solutions appeared twice, then found four more solutions, and showed that there are no others involving exponents at most 21. Finally in 2016 the interactive procedure was repeated more carefully over 60 hours to show that Warin's list is indeed complete. The resulting 134 is just about Beethovenian.

The case $d = 1$ is dealt with as in the proof of Lemma 8 using $x = x'^q$, $y = y'^q$, $z = z'^q$. This completes the present proof. □

Again thanks to the one-to-one assertion in Lemma 5, we can find all the non-mixing sets of size 4. It is quickly checked that every π in Π_3 is pre-broad. Also, the corresponding non-mixing set $M = \{\mathbf{m}_0, \mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3\}$ turns up naturally in semi-reduced form. Furthermore, the coordinates of $\mathbf{m}_0, \mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3$ are all coprime and so M is reduced. So, we have the unique representative, and we only have to check that they are all different (itself not entirely painless).

Two π come immediately from the generators, namely the polynomials

$$1 + u_1 + u_1^2 + u_2, \quad 1 + u_1 + u_1^3 + u_3$$

with respective non-mixing sets

$$\{(0, 0, 0), (1, 0, 0), (2, 0, 0), (0, 1, 0)\}, \quad \{(0, 0, 0), (1, 0, 0), (3, 0, 0), (0, 0, 1)\}$$

as in (5.1).

The next two simplest polynomials are perhaps

$$u_1^2 + u_1^3 + u_2 + u_3, \quad u_1 + u_2 + u_3 + u_1u_2$$

with

$$\{(2, 0, 0), (3, 0, 0), (0, 1, 0), (0, 0, 1)\}, \quad \{(1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 0)\}.$$

The two most complicated are

$$u_1^{21}u_3 + u_1^{20}u_2 + u_2^{12} + u_3^4, \quad u_1^{25} + u_1^{20}u_2u_3 + u_2^{12} + u_3^4$$

with

$$\{(21, 0, 3), (20, 1, 0), (0, 12, 0), (0, 0, 4)\},$$
$$\{(25, 0, 0), (20, 1, 1), (0, 12, 0), (0, 0, 4)\}.$$

Acknowledgement. We wish warmly to thank Klaus Schmidt for his interest in our work on orders of non-mixing and his encouragement to go further with the non-mixing sets themselves. The first author was partially supported by NSF grant DMS-1601229.

REFERENCES

[**ABB**] L. Arenas-Carmona, D. Berend and V. Bergelson. Ledrappier's system is almost mixing of all orders. *Ergod. Th. & Dynam. Sys.* **28** (2008), 339–365.

[**BM**] W. D. Brownawell and D. Masser. Vanishing sums in function fields. *Math. Proc. Cambridge Philos. Soc.* **100** (1986), 427–434.

[**BMZ**] E. Bombieri, D. Masser and U. Zannier. Anomalous subvarieties—structure theorems and applications. *Int. Math. Res. Not. IMRN* (2007), Article ID rnm057 (33 pages), doi:10.1093/imrn/rnm057.

[**D**] H. Derksen. A Skolem–Mahler–Lech theorem in positive characteristic and finite automata. *Invent. Math.* **168** (2007), 175–224.

[**DM1**] H. Derksen and D. Masser. Linear equations over multiplicative groups, recurrences, and mixing I. *Proc. Lond. Math. Soc.* **104** (2012), 1045–1083.

[**DM2**] H. Derksen and D. Masser. Linear equations over multiplicative groups, recurrences, and mixing II. *Indag. Math.* **26** (2015), 113–136.

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

[**Lei**] D. Leitner. Linear equations over multiplicative groups in positive characteristic. *Acta Arith.* **153** (2012), 325–347.

[**LM**] T. Loher and D. Masser. Uniformly counting points of bounded height. *Acta Arith.* **111** (2004), 277–297.

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

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

[**SV**] H. P. Schlickewei and C. Viola. Polynomials that divide many k-nomials. *Number Theory in Progress*. Eds. K. Győry, H. Iwaniec and J. Urbanowicz. Walter de Gruyter, Berlin, 1999, pp. 445–450.

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

[**W**] T. Ward. Three results on mixing shapes. *New York J. Math.* **3A** (1997), 1–10.