

The Exact Order of Mixing in a Problem of Ward: A Seven-Term Polynomial with No Sparser Multiple

Alper Ferudun*

October 3, 2026

Abstract

Problem A of T. Ward's list "Six problems in algebraic dynamics" concerns the polynomial $f = 1 + u_1 u_2 + u_1^2 u_2 + u_1^3 u_2 + u_1^4 + u_2^2 + u_1^4 u_2^2$ and the algebraic $\mathbb{Z}^2$ -action attached to the module $\mathbb{Z}[u_1^{\pm 1}, u_2^{\pm 1}]/\langle p, f \rangle$, where p is a prime for which f is irreducible modulo p . Its order of mixing M was known to satisfy $3 \leq M < 7$, and the problem asks for the exact value. We show that $M = 6$ for every prime p : the action is mixing of order 6 and not of order 7. The polynomial f is absolutely irreducible modulo every prime, so no prime has to be excluded. By results of Schmidt, Masser, and Derksen and Masser, and because the radical of the group generated by u_1, u_2 in the function field of the curve $f = 0$ consists only of constant multiples of monomials, $M + 1$ is the least number of terms of a non-zero multiple of f in $\mathbb{F}_p[u_1^{\pm 1}, u_2^{\pm 1}]$. Our main result is that every non-zero multiple of f with coefficients in an algebraic closure of $\mathbb{F}_p$ has at least seven terms. The proof is elementary: it sorts a multiple by powers of u_2 , reduces to two possible shapes of a six-term multiple by a Chebyshev-type recurrence, and excludes these by a local analysis at the points of the curve with $u_1^4 = -1$, a Frobenius splitting and an explicit computation. The prime 2 is treated separately. For $p = 2$ an exhaustive computer search confirms the main result for coefficients in $\mathbb{F}_2$, independently of this proof, and shows that the support of f is, up to equivalence, the only non-mixing set with seven points. This is an unrefereed note.

1 Introduction

Let $R_d = \mathbb{Z}[u_1^{\pm 1}, \dots, u_d^{\pm 1}]$. Every countable R_d -module N defines an action $\alpha = \alpha_N$ of $\mathbb{Z}^d$ by continuous automorphisms of the compact abelian group $X = \widehat{N}$, dual to multiplication by the monomials $u^n = u_1^{n_1} \cdots u_d^{n_d}$; see Schmidt's book [Sch95]. Let μ be the Haar measure of X . For $r \geq 2$ the action is called *mixing of order r* , or *r -mixing*, if

$$\mu(\alpha^{-n_1} B_1 \cap \cdots \cap \alpha^{-n_r} B_r) \longrightarrow \mu(B_1) \cdots \mu(B_r)$$

for all Borel sets $B_1, \dots, B_r \subset X$, as $n_1, \dots, n_r \in \mathbb{Z}^d$ vary so that $n_i - n_j \rightarrow \infty$ for all $i \neq j$. The *order of mixing* $M(\alpha)$ is the largest r for which α is r -mixing. Ledrappier's example [Led78], given by the module $R_2/\langle 2, 1 + u_1 + u_2 \rangle$, is mixing but not 3-mixing.

Problem A of Ward's list [War06] concerns the polynomial

$$f(u_1, u_2) = 1 + u_1 u_2 + u_1^2 u_2 + u_1^3 u_2 + u_1^4 + u_2^2 + u_1^4 u_2^2 \quad (1)$$

and the $\mathbb{Z}^2$ -action α defined by the module $R_2/\langle p, f \rangle$, where p is a prime for which f is irreducible modulo p . Concretely, α is the shift action on the group of all $x \in \mathbb{F}_p^{\mathbb{Z}^2}$ with $\sum_{n \in S(f)} x_{m+n} = 0$ for

*Mercury Software GmbH. Email: alper@mercurycodelab.com. GitHub: <https://github.com/AlperTheKing>.
Unrefereed preprint. CC BY 4.0.

every $m \in \mathbb{Z}^2$, where $S(f) = \{(0,0), (4,0), (0,2), (4,2), (1,1), (2,1), (3,1)\}$ is the support of f . Ward states that the methods of Einsiedler and Ward [EW03] and of Kitchens and Schmidt [KS93] give $3 \leq M < 7$ for the order of mixing $M = M(\alpha)$, and asks for its exact value. The problem is also record AMR-042-0001 of the UnsolvedMath dataset [ula26].

Theorem 1.1. *For every prime p the polynomial f is absolutely irreducible modulo p , and the action α defined by $R_2/\langle p, f \rangle$ is mixing of order 6 but not of order 7. So $M = 6$ for every prime p .*

Theorem 1.1 follows from known results on mixing in positive characteristic and the next statement. A *term* of a Laurent polynomial is a monomial that occurs in it with a non-zero coefficient.

Theorem 1.2. *Let p be a prime and let $\overline{\mathbb{F}}_p$ be an algebraic closure of $\mathbb{F}_p$. Every non-zero element of the ideal generated by f in $\overline{\mathbb{F}}_p[u_1^{\pm 1}, u_2^{\pm 1}]$ has at least seven terms.*

The polynomial f itself has seven terms. Only multiples with coefficients in $\mathbb{F}_p$ are needed for Theorem 1.1. We work over $\overline{\mathbb{F}}_p$ because the proof uses the roots of $1 + u_1^4$ and p -power roots of the coefficients, and because Remark 2.3 needs the statement over $\overline{\mathbb{F}}_p$.

What was known. Theorem 3.1 of [EW03] states that $R-1 \leq M(\alpha) < |S(f)|$ when f is irreducible and its Newton polygon is an R -gon. For (1) the Newton polygon is the rectangle $[0,4] \times [0,2]$ and $|S(f)| = 7$, which gives Ward's bounds. If all points of $S(f)$ are vertices of the polygon, this determines $M(\alpha)$ [EW03, Example 4.1]. Here the points $(1,1), (2,1), (3,1)$ lie in the interior, and [EW03, Remark 5.1(3)] says that this situation seems to require results on S -unit equations in positive characteristic. Such results were obtained later. Masser [Mas04] proved a conjecture of Schmidt: the order of mixing is detected by non-mixing sets (see Section 2). The 2006 update of Ward's list records this and adds that computing the order of mixing of non-trivial examples remains a considerable problem. Derksen and Masser then showed, in the third part [DM18] of a series of papers [DM12, DM15, DM18] and with the main result of the first part, that for every prime ideal of R_d containing a prime p the smallest order of non-mixing can be determined effectively, and also the non-mixing sets of that order, up to equivalence, if this order is at least 3. They work out two examples, both with order of mixing 3. They also remark that this is related to the problem of finding the shortest polynomial in an ideal, which is probably effectively solvable in positive characteristic, and that one may have to pass to Puiseux polynomials [DM18, Section 1]. So for each fixed p the number M is computable in principle. We found no determination of M for Ward's example in the literature. This note adds the explicit value, for all primes at once, with an elementary proof of Theorem 1.2.

Outline. Write $x = u_1$, $y = u_2$, $a = 1 + x^4$ and $b = x + x^2 + x^3$, so that $f = a(1 + y^2) + by$. Section 2 reduces Theorem 1.1 to Theorem 1.2. After preliminaries on the curve $f = 0$ (Section 3) we sort a multiple of f by powers of y into *rows* and show that a multiple with at most six terms has rows with 2, 2, 2 or 2, 1, 1, 2 terms (Section 4). These two shapes are excluded in Sections 5 and 6 by an analysis at the places of the curve over the roots of a , where $a \approx -by$. Section 8 reports machine checks.

2 From mixing to sparse multiples

Let P be a prime ideal of R_d with $P \cap \mathbb{Z} = p\mathbb{Z}$ for a prime p , and let $\alpha = \alpha_{R_d/P}$. Let K_0 be the quotient field of R_d/P , let F be its field of constants (the algebraic closure of $\mathbb{F}_p$ in K_0), and let $G \subset K_0^*$ be the group generated by the images of $u_1, \dots, u_d$. The *radical* of G is the group $\sqrt{G}$ of all $w \in K_0$ with $w^s \in G$ for some integer $s \geq 1$; it is taken inside K_0 , not in an algebraic closure. A set $\{m_1, \dots, m_r\} \subset \mathbb{Z}^d$ of cardinality r is called *mixing* for α if $\lim_{k \rightarrow \infty} \mu(\bigcap_i \alpha^{-km_i} B_i) = \prod_i \mu(B_i)$ for

all Borel sets $B_1, \dots, B_r$, and *non-mixing* otherwise. If α is r -mixing, then every set of cardinality r is mixing. We use three results from the literature.

- (S) *Schmidt's criterion* ([DM18, (1.1)]; compare [Mas04, p. 191], which refers to equations (27.4) and (27.5) of [Sch95]). A set $\{m_0, \dots, m_n\} \subset \mathbb{Z}^d$ of cardinality $n+1$ is non-mixing for α if and only if there are $a_0, \dots, a_n \in K_0$, not all zero, such that $a_0 u^{km_0} + \dots + a_n u^{km_n} = 0$ in K_0 for infinitely many positive integers k .
- (M) *Masser's theorem* ([Mas04, Corollary, p. 190]). Let $r \geq 2$. If every subset of $\mathbb{Z}^d$ of cardinality r is mixing for α , then α is r -mixing.
- (DM) *Derksen and Masser* ([DM18, Lemma 5]). Let $n \geq 2$, and suppose that α is n -mixing but not $(n+1)$ -mixing. Then there are $\xi_0, \dots, \xi_n \in \sqrt{G}$ with $\xi_0 + \dots + \xi_n = 0$ and $\xi_i/\xi_j \notin F^*$ for all $i \neq j$.

In [DM18] a projective point $(\xi_0 : \dots : \xi_n)$ as in (DM) is called *pre-broad*. Lemma 5 of [DM18] says more: the non-mixing sets of cardinality $n+1$ form finitely many equivalence classes, which correspond to the pre-broad points of a finite set (see Section 8). The hypothesis $n \geq 2$ is needed. The proof of (DM) combines (S), (M) and [DM18, Lemma 4]: if $n \geq 2$, α is n -mixing, $a_1, \dots, a_n \in K_0$, and the equation $a_1 x_1 + \dots + a_n x_n = 1$ has a broad set of solutions in $(\sqrt{G})^n$ (a notion from [Mas04]), then $a_1, \dots, a_n \in \sqrt{G}$.

Now let $d = 2$, $x = u_1$, $y = u_2$ and $P = \langle p, f \rangle$ with f as in (1). By Lemma 3.1 below, P is a prime ideal for every prime p , and $K_0 = \mathbb{F}_p(x)[y]/(f)$ is a subfield of $K = \overline{\mathbb{F}}_p(x)[y]/(f)$.

Lemma 2.1. *Let p be a prime, and let G be the subgroup of K_0^* generated by x and y .*

- (a) *If $w \in K^*$, $c \in \overline{\mathbb{F}}_p^*$, $s \geq 1$ and $(d, e) \in \mathbb{Z}^2$ satisfy $w^s = c x^d y^e$, then $s \mid d$, $s \mid e$ and $w = c' x^{d/s} y^{e/s}$ with $c' \in \overline{\mathbb{F}}_p^*$.*
- (b) *$\overline{\mathbb{F}}_p \cap K_0 = \mathbb{F}_p$ and $\sqrt{G} = \mathbb{F}_p^* \cdot G$. Moreover $x^d y^e \in \overline{\mathbb{F}}_p$ only if $d = e = 0$.*

The proof is given in Section 3.

Proposition 2.2. *Let p be a prime and let α be the action defined by $R_2/\langle p, f \rangle$. Then $M(\alpha) + 1$ is the least number of terms of a non-zero element of the ideal generated by f in $\mathbb{F}_p[x^{\pm 1}, y^{\pm 1}]$.*

Proof. A Laurent polynomial over $\mathbb{F}_p$ lies in this ideal if and only if it vanishes in K_0 (Section 3). Let L be the least number of terms; $L \leq 7$.

Let $g = \sum_{i=1}^L c_i u^{n_i}$ be an element of the ideal with $c_i \in \mathbb{F}_p^*$ and pairwise distinct n_i . For every $e \geq 0$ we have $\sum_i c_i u^{p^e n_i} = g^{p^e} = 0$ in K_0 . By (S), with $k = p^e$, the set $\{n_1, \dots, n_L\}$ is non-mixing. So α is not L -mixing, and $M(\alpha) \leq L - 1$.

The action α is 2-mixing. Otherwise, by (M) and (S), there are $m_0 \neq m_1$ and $(a_0, a_1) \neq (0, 0)$ with $a_0 u^{km_0} + a_1 u^{km_1} = 0$ for infinitely many k . Then $a_0 a_1 \neq 0$, and $u^{k(m_1 - m_0)} = -a_0/a_1$ for two different values of k , so some monomial u^m with $m \neq 0$ equals 1 in K_0 . This contradicts Lemma 2.1(b). Hence $n = M(\alpha)$ satisfies $2 \leq n \leq L - 1$, and α is n -mixing but not $(n+1)$ -mixing. By (DM) and Lemma 2.1(b) there are $c_i \in \mathbb{F}_p^*$ and $e_i \in \mathbb{Z}^2$ ($0 \leq i \leq n$) with $\sum_i c_i u^{e_i} = 0$ in K_0 and $c_i u^{e_i}/(c_j u^{e_j}) \notin \mathbb{F}_p^*$ for $i \neq j$, that is, with pairwise distinct e_i . So $\sum_i c_i u^{e_i}$ is a non-zero element of the ideal with exactly $n+1$ terms, and $L \leq M(\alpha) + 1$. $\square$

Remark 2.3 (Another form of the criterion). The reduction can also be based on an older theorem, and then Lemmas 4 and 5 of [DM18] are not needed. According to [Mas04, Section 5] and [LS19, Section 3], a theorem of Masser, published as [KS93, Theorem 2.1] and as Theorem 28.1 of [Sch95],

contains the following implication, which is all we use. Let $x_0, \dots, x_n \in K_0^*$, and suppose that there are $a_i \in K_0$, not all zero, with $\sum_i a_i x_i^k = 0$ for infinitely many $k \geq 1$. Then there are integers $s, t \geq 1$ and s -th roots y_i of the x_i in an algebraic closure such that $y_0^t, \dots, y_n^t$ are linearly dependent over $\overline{\mathbb{F}}_p$.

Let $s \geq 1$, and let X, Y be s -th roots of x, y in an algebraic closure of K . *Claim: every relation $\sum_i c_i X^{d_i} Y^{e_i} = 0$ with $c_i \in \overline{\mathbb{F}}_p^*$ and pairwise distinct $(d_i, e_i) \in \mathbb{Z}^2$ has at least seven terms.* Raising the relation to a power of p , we may assume $p \nmid s$. By Lemma 2.1(a), $x^d y^e$ is an s -th power in K only if $s \mid d$ and $s \mid e$. So $[K(X, Y) : K] = s^2$ by Kummer theory, and the s^2 elements $X^i Y^j$ with $0 \leq i, j < s$, which span $K(X, Y)$ over K , are linearly independent over K . Sort the terms of the relation by the residues of (d_i, e_i) modulo s . The terms in a residue class (i, j) , divided by $X^i Y^j$, form a Laurent polynomial in x, y that vanishes in K , that is, a multiple of f . By Theorem 1.2 every non-empty class has at least seven terms.

If now $\{m_0, \dots, m_n\}$ is a non-mixing set, apply (S) and the implication quoted above to $x_i = u^{m_i}$. Each y_i^t is a root of unity times a monomial in X, Y with exponent vector tm_i . We get a relation as in the claim with at most $n + 1$ terms, so $n + 1 \geq 7$, and α is 6-mixing by (M). This argument uses Theorem 1.2 and Lemma 2.1(a) over $\overline{\mathbb{F}}_p$.

Remark 2.4. Both properties in Lemma 2.1(b) matter. For $\langle 2, (1 + u_1 + u_1^2)^3 + u_2 \rangle$ the radical contains $1 + u_1 + u_1^2$, and the order of mixing is 3, although the shortest polynomial in the ideal probably has six terms [DM18]. For $\langle 2, 1 + u + u^2 + uv + v^2 \rangle$ the constant field is $\mathbb{F}_4$, and there is a non-mixing set with three points, although no element of the ideal has three terms [LS19, Section 3].

3 The curve $f = 0$

From now on p is a prime, $x = u_1$, $y = u_2$, and

$$a = 1 + x^4, \quad b = x + x^2 + x^3 = x(1 + x + x^2), \quad f = ay^2 + by + a.$$

Lemma 3.1. (a) *The polynomials a and b are coprime in $\overline{\mathbb{F}}_p[x]$.* (b) *The polynomial f is irreducible in $\overline{\mathbb{F}}_p(x)[y]$ and in $\overline{\mathbb{F}}_p[x^{\pm 1}, y^{\pm 1}]$.*

Proof. (a) We have $a(0) = 1$, and if $w^2 + w + 1 = 0$, then $w^3 = 1$ and $a(w) = 1 + w = -w^2 \neq 0$.

(b) By (a), f is primitive as a polynomial in y over $\overline{\mathbb{F}}_p[x]$, and it is not divisible by x or y . By Gauss's lemma it suffices to show that f has no root in $\overline{\mathbb{F}}_p(x)$.

Let p be odd. If f has a root in $\overline{\mathbb{F}}_p(x)$, then $D = b^2 - 4a^2 = (b - 2a)(b + 2a)$ is a square in $\overline{\mathbb{F}}_p[x]$. We exhibit a simple root of D . Over $\mathbb{Z}$ we have $b + 2a = (2x^2 + 3x + 2)(x^2 - x + 1)$. Let ω be a root of $x^2 - x + 1$. Then $\omega^2 = \omega - 1$ and $\omega^3 = -1$, and one computes $2\omega^2 + 3\omega + 2 = 5\omega$ and $(b - 2a)(\omega) = 4\omega^2$. If $p \neq 3, 5$, then $x^2 - x + 1$ has two simple roots, and its cofactor in D takes the value $20\omega^3 = -20 \neq 0$ at ω . So ω is a simple root of D . If $p \in \{3, 5\}$, let $i^2 = -1$. Then $a(i) = 2$ and $b(i) = -1$, so $D(i) = -15 = 0$, while $D'(i) = 2b(i)b'(i) - 8a(i)a'(i) = 4 + 60i \neq 0$. So i is a simple root of D .

Let $p = 2$. The substitution $y = (b/a)(Y + a/b)$ turns $f = 0$ into $Y^2 + Y = a/b$, and $a/b = (1 + x)^4/(x(1 + x + x^2))$ has a simple pole at $x = 0$. If $h \in \overline{\mathbb{F}}_p(x)$ has a pole of order $m \geq 1$ at 0, then $h^2 + h$ has a pole of order $2m$ there, and otherwise $h^2 + h$ has no pole at 0. So $Y^2 + Y = a/b$ has no solution in $\overline{\mathbb{F}}_p(x)$. $\square$

By Lemma 3.1, $K = \overline{\mathbb{F}}_p(x)[y]/(f)$ is a field, quadratic over $\overline{\mathbb{F}}_p(x)$ with basis $\{1, y\}$. By Gauss's lemma, a Laurent polynomial in $\overline{\mathbb{F}}_p[x^{\pm 1}, y^{\pm 1}]$ lies in the ideal (f) if and only if its image in K is zero. The same holds over $\overline{\mathbb{F}}_p$ with $K_0 = \overline{\mathbb{F}}_p(x)[y]/(f) \subset K$, and comparing coordinates in the basis $\{1, y\}$ shows $\overline{\mathbb{F}}_p \cap K_0 = \overline{\mathbb{F}}_p$. In K we have $y \neq 0$ and

$$y + \frac{1}{y} = t := -\frac{b}{a}, \quad 1 + y^2 = ty. \quad (2)$$

Since $y^2 f(x, 1/y) = f(x, y)$ and $x^4 f(1/x, y) = f(x, y)$, the field K has the automorphisms $\sigma: y \mapsto 1/y$ (fixing x) and $\tau: x \mapsto 1/x$ (fixing y); and if $g \in (f)$, then $g(x, 1/y) \in (f)$. Define $S_n \in \mathbb{Z}[T]$ by $S_{-2} = -1$, $S_{-1} = 0$ and $S_n = TS_{n-1} - S_{n-2}$ for $n \geq 0$. So $S_0 = 1$, $S_1 = T$, and S_n is monic of degree n for $n \geq 0$. From $y^2 = ty - 1$ we get by induction

$$y^n = S_{n-1}(t)y - S_{n-2}(t) \quad (n \geq 0). \quad (3)$$

Since t is not constant, $S_n(t) \neq 0$ in $\overline{\mathbb{F}}_p(x)$ for every $n \geq 0$.

Places. We use places of the function field $K/\overline{\mathbb{F}}_p$; see [Sti09]. For a place $\mathfrak{q}$ let $v_{\mathfrak{q}}$ be its valuation. A non-singular point (x_0, y_0) of the affine curve $f = 0$ is the centre of exactly one place of K , and $x - x_0$ is a local parameter there if $\partial f / \partial y \neq 0$ at the point; likewise $y - y_0$ if $\partial f / \partial x \neq 0$ [Ful89, Section 3.2, Theorem 1]. If π is a local parameter at $\mathfrak{q}$ and $h \in K^*$ has $v_{\mathfrak{q}}(h) = v$, the residue of h/π^v in $\overline{\mathbb{F}}_p^*$ is called the *leading coefficient* of h at $\mathfrak{q}$. If $h_1 + \dots + h_n = 0$ in K with all $h_i \neq 0$, then

the minimal $v_{\mathfrak{q}}(h_i)$ occurs at least twice, and the leading coefficients of these h_i add up to 0. (4)

Lemma 3.2. (a) *Let p be odd and let x_0 be one of the four roots of a , the primitive 8th roots of unity. Then $(x_0, 0)$ is a non-singular point of the curve. Let $\mathfrak{q}$ be the place centred there. Both $x - x_0$ and y are local parameters at $\mathfrak{q}$, and with respect to y the leading coefficient of a at $\mathfrak{q}$ is $-b(x_0) = -x_0(1 + x_0 + x_0^2) \neq 0$. The place $\mathfrak{q}' = \sigma(\mathfrak{q})$ is different from $\mathfrak{q}$, and $v_{\mathfrak{q}'}(x - x_0) = 1$, $v_{\mathfrak{q}'}(y) = -1$.*

(b) *Let $p = 2$ and $\pi = 1 + x$, so that $a = \pi^4$ and $b = 1 + \pi^3$. There are exactly two places of K over $x = 1$, say $\mathfrak{q}$ and $\mathfrak{q}' = \sigma(\mathfrak{q})$. At both, π is a local parameter. Moreover $v_{\mathfrak{q}}(y) = 4$, $v_{\mathfrak{q}'}(y) = -4$, and $y = \pi^4 Y$ with $v_{\mathfrak{q}}(Y - 1 - \pi^3) \geq 4$.*

Proof. (a) We have $f(x_0, 0) = a(x_0) = 0$, and at this point $\partial f / \partial x = a'(x_0) = 4x_0^3 \neq 0$ and $\partial f / \partial y = b(x_0) \neq 0$ by Lemma 3.1(a). By (2), $a = -by/(1 + y^2)$, and b and $1 + y^2$ are units at $\mathfrak{q}$ with residues $b(x_0)$ and 1. Since $1/y$ vanishes at $\sigma(\mathfrak{q})$, this place is different from $\mathfrak{q}$.

(b) We have $b = (1 + \pi)(1 + \pi + \pi^2) = 1 + \pi^3$. Let $\mathfrak{q}$ be a place over $x = 1$ with ramification index e . Then $v_{\mathfrak{q}}(y + 1/y) = v_{\mathfrak{q}}(b/a) = -4e$, so $v_{\mathfrak{q}}(y) = \pm 4e$. As σ interchanges y and $1/y$, there are two places over $x = 1$, one with $v(y) > 0$, which we call $\mathfrak{q}$, and $\sigma(\mathfrak{q})$. Since $[K : \overline{\mathbb{F}}_p(x)] = 2$, both are unramified. So π is a local parameter and $v_{\mathfrak{q}}(y) = 4$. Finally, $\pi^4(1 + y^2) = (1 + \pi^3)y$ gives $Y = y/\pi^4 = (1 + y^2)/(1 + \pi^3) \equiv 1 + \pi^3 \pmod{\pi^4}$, because $v_{\mathfrak{q}}(y^2) = 8$ and $(1 + \pi^3)^2 \equiv 1 \pmod{\pi^4}$. $\square$

To treat both cases together, put $\kappa = 1$ if p is odd and $\kappa = 4$ if $p = 2$, and let $x_0 = 1$ if $p = 2$. For each root x_0 of a we then have two places $\mathfrak{q}, \mathfrak{q}'$ over $x = x_0$ with

$$v_{\mathfrak{q}}(y) = \kappa, \quad v_{\mathfrak{q}'}(y) = -\kappa, \quad v_{\mathfrak{q}}(r) = v_{\mathfrak{q}'}(r) = \text{ord}_{x_0}(r) \quad \text{for } 0 \neq r \in \overline{\mathbb{F}}_p[x^{\pm 1}]. \quad (5)$$

Lemma 3.3. *Let Q be a power of p .*

(a) $K = \bigoplus_{c=0}^{Q-1} x^c K^Q$.

(b) *If $p = 2$, then $x^d y^e \in K^2$ if and only if d and e are even.*

Proof. (a) The extension $K/\overline{\mathbb{F}}_p(x)$ is separable, because $\partial f / \partial y = 2ay + b$ is not zero in K . The map $h \mapsto h^Q$ is an isomorphism of K onto K^Q which maps $\overline{\mathbb{F}}_p(x)$ onto $\overline{\mathbb{F}}_p(x^Q)$. So $[K^Q : \overline{\mathbb{F}}_p(x^Q)] = 2$ and $[K : K^Q] = [K : \overline{\mathbb{F}}_p(x^Q)]/2 = Q$. The extension $K/K^Q(x)$ is purely inseparable and, since $K^Q(x) \supset \overline{\mathbb{F}}_p(x)$, also separable. Hence $K = K^Q(x)$. As $x^Q \in K^Q$, the powers $1, x, \dots, x^{Q-1}$ span K over K^Q , and they are a basis because $[K : K^Q] = Q$.

(b) Let ∂ be the derivation of K that extends d/dx ; it exists because $K/\overline{\mathbb{F}}_p(x)$ is separable. Its kernel is a field that contains K^2 and does not contain x . Since $[K : K^2] = 2$, the kernel is K^2 . Applying ∂ to $f(x, y) = 0$ gives $b\partial y = b'y$, because $a' = 0$ and $2ay = 0$ in characteristic 2; here $b' = 1 + x^2$. Hence

$$\partial(x^d y^e) = x^{d-1} y^e (d + e\rho), \quad \rho = \frac{xb'}{b} = \frac{(1+x)^2}{1+x+x^2}.$$

As $\rho \notin \mathbb{F}_2$, this is zero if and only if d and e are even. $\square$

Proof of Lemma 2.1. Zeros and poles of y . Let $v_{\mathfrak{q}}(y) > 0$. If x had a pole at $\mathfrak{q}$, then dividing $a(1+y^2) = -by$ by x^4 would give the residue 1 on the left and 0 on the right. So x is finite at $\mathfrak{q}$, and its value x_0 satisfies $a(x_0) = 0$. Thus $\mathfrak{q}$ is one of the places of Lemma 3.2, and y has a zero of order κ there by (5). By σ , the poles of y are the places $\mathfrak{q}'$.

Zeros and poles of x . At a zero of x the function a is a unit, so y is finite, and its value y_0 satisfies $1 + y_0^2 = 0$. If p is odd, these are the points $(0, \pm i)$, where $\partial f/\partial y = \pm 2i \neq 0$; so x has two simple zeros. If $p = 2$, the only such point is $(0, 1)$, where $\partial f/\partial y = b(0) = 0$ and $\partial f/\partial x = b'(0) = 1$. So this point is non-singular, there is exactly one place over $x = 0$, and x has a zero of order 2 there. By τ , the poles of x have the same orders. At the zeros and poles of x the function y is a unit, and at the zeros and poles of y the function x is a unit.

(a) Let $w^s = c x^d y^e$. Comparing orders at a zero of x and at a zero of y gives $s \mid d$ and $s \mid e$ if p is odd, and $s \mid 2d$ and $s \mid 4e$ if $p = 2$. If p is odd, $w/(x^{d/s} y^{e/s})$ has s -th power c , so it is a constant. Let $p = 2$ and $s = 2^r s'$ with s' odd. Then $s' \mid d$ and $s' \mid e$, and as before $w^{2^r} = c_1 x^{d_1} y^{e_1}$ with $c_1 \in \overline{\mathbb{F}}_p^*$, $d_1 = d/s'$ and $e_1 = e/s'$. If $r \geq 1$, then $x^{d_1} y^{e_1}$ is a square in K , so d_1 and e_1 are even by Lemma 3.3(b), and taking square roots gives $w^{2^{r-1}} = c_1^{1/2} x^{d_1/2} y^{e_1/2}$. Induction on r proves (a).

(b) We saw $\overline{\mathbb{F}}_p \cap K_0 = \mathbb{F}_p$ after Lemma 3.1. If $w \in K_0$ and $w^s \in G$, then by (a) $w = c' x^{d/s} y^{e/s}$ with $c' \in \overline{\mathbb{F}}_p \cap K_0 = \mathbb{F}_p$. If $x^d y^e = c \in \overline{\mathbb{F}}_p$, then (a), applied with $w = 1$ and an arbitrary s , gives $s \mid d$ and $s \mid e$ for all s . $\square$

For $k \geq 1$ let $\psi_k(Z) = 1 + Z + \cdots + Z^{k-1}$.

Lemma 3.4 (Binomials). *Let $\beta = x^m(\gamma + \gamma' x^n) \in \overline{\mathbb{F}}_p[x^{\pm 1}]$ with $n \geq 1$ and $\gamma\gamma' \neq 0$, and let x_0 be a root of a .*

- (a) *The order of vanishing of β at x_0 is 0 or the p -part $p^{v_p(n)}$ of n .*
- (b) *Let p be odd and $a \mid \beta$. Then $n = 4k$ and $\beta = \gamma x^m(1 - z^k)$ with $z = -x^4$. If $k = qk'$ with $q = p^{v_p(k)}$, then $\beta = \gamma x^m(a\psi_{k'}(z))^q$. So β vanishes to order exactly q at each root of a , and with respect to the local parameter y its leading coefficient at the place $\mathfrak{q}$ of Lemma 3.2(a) is $\gamma k' x_0^m (-b(x_0))^q$.*
- (c) *Let $p = 2$ and $a \mid \beta$. Then $\gamma' = \gamma$, the order of vanishing of β at $x = 1$ is $q = 2^{v_2(n)} \geq 4$, and $\beta/\pi^q \equiv \gamma x^m \pmod{\pi^q}$.*

Proof. (a) If $\beta(x_0) = 0$, then $\gamma + \gamma' x^n = \gamma(1 - (x/x_0)^n) = \gamma(1 - (x/x_0)^{n'})^q$ with $n = qn'$, $q = p^{v_p(n)}$, and $1 - X^{n'}$ has a simple root at $X = 1$.

(b) The four roots x_0 of a give the same value $x_0^n = -\gamma/\gamma'$. If ζ is one of them, ζ^3 is another, so $\zeta^{2n} = 1$ and $4 \mid n$. With $n = 4k$ we get $x_0^n = (-1)^k$, so $\gamma' = -(-1)^k \gamma$ and $\gamma + \gamma' x^n = \gamma(1 - z^k)$. Further $1 - z^k = (1 - z^{k'})^q = ((1 - z)\psi_{k'}(z))^q$ and $1 - z = a$. At a root of a we have $z = 1$ and $\psi_{k'}(1) = k' \neq 0$ in $\overline{\mathbb{F}}_p$. The leading coefficient follows from Lemma 3.2(a) and $k'^q = k'$.

(c) Evaluating at $x = 1$ gives $\gamma' = \gamma$. With $n = qn'$, n' odd, $1 + x^n = (1 + x^{n'})^q = \pi^q \psi_{n'}(x)^q$, and $\psi_{n'}(x) \equiv \psi_{n'}(1) = 1 \pmod{\pi}$. So $\psi_{n'}(x)^q \equiv 1 \pmod{\pi^q}$, and $q \geq 4$ because $\pi^4 \mid \beta$. $\square$

Lemma 3.5 (Frobenius splitting). *Let Q be a power of p , let $G_1, \dots, G_n \in \overline{\mathbb{F}}_p[x^{\pm 1}, y^{\pm 1}]$ and $m_1, \dots, m_n \in \mathbb{Z}$, and suppose that $\sum_i x^{m_i} G_i^Q \in (f)$. Then for every $c \in \{0, \dots, Q-1\}$ the sum of the $x^{(m_i-c)/Q} G_i$ over all i with $m_i \equiv c \pmod{Q}$ lies in (f) .*

Proof. In K we have $0 = \sum_c x^c (\sum_{m_i \equiv c} x^{(m_i-c)/Q} G_i)^Q$. By Lemma 3.3(a) each of the Q -th powers is zero. $\square$

Lemma 3.6. *Let $U = 1/(1+y^2) \in K$ and $\varepsilon = a^2/b^2$. Then*

$$a = -byU, \quad y^2U = 1 - U, \quad U(1 - U) = \varepsilon, \quad (6)$$

and $\{1, U\}$ is a basis of K over $\overline{\mathbb{F}}_p(x)$. Let q be a power of p .

(a) *Let p be odd, $D = b^2 - 4a^2$, $\delta = \frac{1}{4} - \varepsilon = D/(4b^2)$ and $W = U - \frac{1}{2}$. Then $W^2 = \delta$ and*

$$\varepsilon U^{q-1} = \left(\frac{1}{4} - \delta^{(q+1)/2}\right) + \frac{1}{2}(\delta^{(q-1)/2} - 1)W, \quad \varepsilon(1-U)^{q-1} = \left(\frac{1}{4} - \delta^{(q+1)/2}\right) - \frac{1}{2}(\delta^{(q-1)/2} - 1)W.$$

(b) *Let $p = 2$, $q = 2^d$ and $E_d = \varepsilon + \varepsilon^2 + \varepsilon^4 + \dots + \varepsilon^{q/2}$ (so $E_0 = 0$). Then*

$$U^q = U + E_d, \quad \varepsilon U^{q-1} = \varepsilon + E_d(1+U), \quad \varepsilon(1+U)^{q-1} = \varepsilon + E_dU.$$

Proof. By (2), $1 + y^2 = ty = -by/a$; this gives the first formula in (6), and the others follow. If $U \in \overline{\mathbb{F}}_p(x)$, then $y^2 \in \overline{\mathbb{F}}_p(x)$ and $y = (1 + y^2)/t \in \overline{\mathbb{F}}_p(x)$, which is false.

(a) $W^2 = U^2 - U + \frac{1}{4} = \delta$. So $U^q = (\frac{1}{2} + W)^q = \frac{1}{2} + \delta^{(q-1)/2}W$, and $\varepsilon U^{q-1} = (1 - U)U^q = (\frac{1}{2} - W)(\frac{1}{2} + \delta^{(q-1)/2}W)$. Multiplying out gives the first formula. The second follows by the symmetry $U \leftrightarrow 1 - U$, $W \leftrightarrow -W$.

(b) $U^2 = U + \varepsilon$, and by induction $U^{2^d} = U + E_d$, since $E_{d+1} = \varepsilon + E_d^2$. Then $\varepsilon U^{q-1} = (1 + U)U^q = (1 + U)(U + E_d) = \varepsilon + E_d(1 + U)$, and $\varepsilon(1 + U)^{q-1} = U(1 + U)^q = U(1 + U + E_d) = \varepsilon + E_dU$. $\square$

4 Rows: at least six terms

Let $0 \neq g \in (f) \subset \overline{\mathbb{F}}_p[x^{\pm 1}, y^{\pm 1}]$. After multiplication by a monomial we can write $g = \sum_{j=0}^B g_j(x) y^j$ with $g_j \in \overline{\mathbb{F}}_p[x^{\pm 1}]$ and $g_0 \neq 0 \neq g_B$. We call the g_j the *rows* of g and B its *height*. The number of terms of g is the sum of the numbers of terms of its rows.

Lemma 4.1. *We have $B \geq 2$, and a divides g_0 and g_B in $\overline{\mathbb{F}}_p[x^{\pm 1}]$. In particular g_0 and g_B have at least two terms each.*

Proof. Write $g = fh$. The lowest and the highest row of fh are a times the lowest and the highest row of h , and the heights differ by 2. A monomial is not divisible by a . $\square$

Lemma 4.2. (i) *Some row g_j with $0 < j < B$ is non-zero.*

(ii) *If exactly one row g_j with $0 < j < B$ is non-zero, then it is not a monomial.*

Proof. (i) If $g = g_0 + g_B y^B$, then by (3) the coefficient of y in the basis $\{1, y\}$ is $g_B S_{B-1}(t)$, which is not zero.

(ii) The coefficient of y gives $g_j S_{j-1}(t) + g_B S_{B-1}(t) = 0$. Suppose $g_j = \gamma x^m$. Then $g_B = -\gamma x^m R(t)$ with $R = S_{j-1}/S_{B-1} \in \overline{\mathbb{F}}_p(T)$. Since $0 \leq \deg S_{j-1} < \deg S_{B-1}$, R is not a polynomial, so it has a pole at some $T_0 \in \overline{\mathbb{F}}_p$. We claim that there is $x_1 \in \overline{\mathbb{F}}_p$ with $x_1 \neq 0$, $a(x_1) \neq 0$ and $t(x_1) = T_0$, that is, $b(x_1) + T_0 a(x_1) = 0$. If $T_0 \neq 0$, the polynomial $T_0 x^4 + x^3 + x^2 + x + T_0$ has a root $x_1 \neq 0$, and $a(x_1) = 0$ would imply $b(x_1) = 0$, against Lemma 3.1(a). If $T_0 = 0$, take a root of $1 + x + x^2$. Now $R(t(x))$ has a pole at x_1 , so g_B is not a Laurent polynomial. This is a contradiction. $\square$

Proposition 4.3. *A non-zero element of (f) has at least six terms. If it has exactly six terms, its non-zero rows are, from bottom to top,*

- *three binomials g_0, g_j, g_B with $0 < j < B$ (shape $(2, 2, 2)$), or*
- *a binomial g_0 , two monomials g_{h_1}, g_{h_2} with $0 < h_1 < h_2 < B$, and a binomial g_B (shape $(2, 1, 1, 2)$).*

Proof. By Lemmas 4.1 and 4.2(i), g_0 and g_B have at least two terms each and there is a non-zero row in between. With at most five terms, there is exactly one row in between, and it is a monomial. This contradicts Lemma 4.2(ii). With six terms, the rows in between have one term or two terms in total; one term is excluded in the same way. $\square$

5 The shape $(2, 1, 1, 2)$

Proposition 5.1. *No element of (f) has shape $(2, 1, 1, 2)$.*

Proof. Suppose that $g = g_0 + \gamma_1 x^{m_1} y^{h_1} + \gamma_2 x^{m_2} y^{h_2} + g_B y^B \in (f)$ with $0 < h_1 < h_2 < B$, $\gamma_1 \gamma_2 \neq 0$ and binomials g_0, g_B . By Lemma 4.1, a divides g_0 and g_B .

Let p be odd. By Lemma 3.4(b), $g_0 = \gamma_0 x^{m_0} (1 - z^{k_0})$, and g_0 vanishes to order $q_0 = p^{v_p(k_0)}$ at every root x_0 of a . Write $k_0 = q_0 k'_0$. At the place $\mathfrak{q}$ over x_0 the four summands of g have the valuations q_0, h_1, h_2 and at least B . By (4) and $h_1 < h_2 < B$ we get $q_0 = h_1$, and the leading coefficients of the first two summands cancel:

$$\gamma_0 k'_0 x_0^{m_0} (-b(x_0))^{q_0} + \gamma_1 x_0^{m_1} = 0 \quad \text{for all four roots } x_0 \text{ of } a.$$

Since $b = x(1 + x + x^2)$ and q_0 is odd, this reads $(1 + x_0 + x_0^2)^{q_0} = c x_0^e$ with $c = \gamma_1 / (\gamma_0 k'_0) \neq 0$ and $e = m_1 - m_0 - q_0$. Put $\xi = x_0^{q_0}$. As q_0 is a power of p , the left side is $1 + \xi + \xi^2$. As q_0 is odd, $x_0 \mapsto x_0^{q_0}$ permutes the four primitive 8th roots of unity, and $x_0 = \xi^r$ for an integer r with $r q_0 \equiv 1 \pmod{8}$. Hence

$$1 + \xi + \xi^2 = c \xi^{e'} \quad \text{for all four primitive 8th roots of unity } \xi,$$

where $0 \leq e' \leq 7$ and $e' \equiv r e \pmod{8}$. So $x^4 + 1$ divides $1 + x + x^2 - c x^{e'}$. Modulo $x^4 + 1$ the monomial $c x^{e'}$ is $\pm c x^{e' \bmod 4}$. So the remainder of $1 + x + x^2 - c x^{e'}$ has at least two non-zero coefficients. This is a contradiction.

Let $p = 2$. We use induction on B and assume that no element of (f) of shape $(2, 1, 1, 2)$ has smaller height. By Lemma 3.4(c), $g_0 = \gamma_0 x^{m_0} (1 + x^{n_0})$ with $q_0 = 2^{v_2(n_0)} \geq 4$. At the place $\mathfrak{q}$ of Lemma 3.2(b) the four summands of g have the valuations $q_0, 4h_1, 4h_2$ and at least $4B$. By (4), $q_0 = 4h_1$. We divide $g = 0$ by π^{q_0} and reduce modulo π^4 . The last two summands disappear, because $4h_2 - q_0 \geq 4$. By Lemma 3.4(c) and Lemma 3.2(b),

$$\gamma_0 (1 + \pi)^{m_0} + \gamma_1 (1 + \pi)^{m_1} (1 + h_1 \pi^3) \equiv 0 \pmod{\pi^4}.$$

So $\gamma_0 = \gamma_1$ and $(1 + \pi)^{m_0 - m_1} \equiv 1 + h_1 \pi^3 \pmod{\pi^4}$. Now $(1 + \pi)^n \bmod \pi^4$ depends only on n modulo 4, and it equals $1, 1 + \pi, 1 + \pi^2, 1 + \pi + \pi^2 + \pi^3$ for $n \equiv 0, 1, 2, 3$. Hence $m_1 \equiv m_0 \pmod{4}$ and h_1 is even. The same argument for $y^B g(x, 1/y) \in (f)$ gives $m_2 \equiv m_B \pmod{4}$ and $B - h_2$ even, where $g_B = \gamma_B x^{m_B} (1 + x^{n_B})$. Since n_0 and n_B are even, the exponent vectors of the three terms of $g_0 + \gamma_1 x^{m_1} y^{h_1}$ are congruent to $c_I = (m_0, 0)$ modulo 2, and those of the three terms of $\gamma_2 x^{m_2} y^{h_2} + g_B y^B$ are congruent to $c_{II} = (m_B, B)$. As $\overline{\mathbb{F}}_p$ is perfect, we can write

$$g = u^{c_I} W_I^2 + u^{c_{II}} W_{II}^2$$

with Laurent polynomials W_I, W_{II} that have three terms each. By Proposition 4.3 they are not in (f) , so they are non-zero in K , and $u^{c_I - c_{II}} = (W_{II}/W_I)^2 \in K^2$. By Lemma 3.3(b), $c_I \equiv c_{II} \pmod{2}$. Then $g = u^{c_I} H^2$ with $H = W_I + u^{(c_{II} - c_I)/2} W_{II}$. The Laurent polynomial H vanishes in K , so $H \in (f)$. Its terms are the square roots of the terms of $u^{-c_I} g$. So H has shape $(2, 1, 1, 2)$ and height $B/2 < B$. This contradicts the induction hypothesis. $\square$

6 The shape $(2, 2, 2)$

Proposition 6.1. *No element of (f) has shape $(2, 2, 2)$.*

Proof. Suppose that $g = g_0 + g_j y^j + g_B y^B \in (f)$ with $0 < j < B$ and three binomial rows $g_\bullet = x^{m_\bullet}(\gamma_\bullet + \gamma'_\bullet x^{n_\bullet})$, $\bullet \in \{0, j, B\}$. By Lemma 4.1, a divides g_0 and g_B . Let x_0 be a root of a , and let q_0, q_B and ν be the orders of vanishing of g_0, g_B and g_j at x_0 . By Lemma 3.4, q_0 and q_B are powers of p that do not depend on x_0 , and $q_0, q_B \geq 4$ if $p = 2$.

Step 1: valuations. By (5) the three summands of g have the valuations $q_0, \kappa j + \nu, \kappa B + q_B$ at $\mathfrak{q}$, and $q_0, \nu - \kappa j, q_B - \kappa B$ at $\mathfrak{q}'$. Put

$$v_1 = q_0, \quad v_2 = \kappa j + \nu, \quad v_3 = \kappa B + q_B, \quad v'_1 = q_B, \quad v'_2 = \kappa(B - j) + \nu, \quad v'_3 = \kappa B + q_0.$$

By (4) the minimum is attained at least twice in (v_1, v_2, v_3) and in (v'_1, v'_2, v'_3) . Suppose v_3 is minimal in the first triple. Then $v_1 \geq v_3 > q_B = v'_1$, so $v'_3 > v'_1$, and therefore $v'_1 = v'_2$. Now $v_3 \leq v_2$ reads $\kappa B + \kappa(B - j) + \nu \leq \kappa j + \nu$, which is false. By symmetry v'_3 is not minimal in the second triple. Hence $v_1 = v_2$ and $v'_1 = v'_2$:

$$q_0 = \kappa j + \nu, \quad q_B = \kappa(B - j) + \nu. \quad (7)$$

In particular ν does not depend on x_0 .

Step 2: $\nu \geq 1$. Suppose $\nu = 0$.

Let p be odd. Then $j = q_0$. By (4) the leading coefficients of g_0 and $g_j y^j$ at $\mathfrak{q}$ cancel. By Lemma 3.4(b), with $g_0 = \gamma_0 x^{m_0}(1 - z^{k_0})$ and $k_0 = q_0 k'_0$, this gives

$$\gamma_0 k'_0 x_0^{m_0 + q_0} (1 + x_0 + x_0^2)^{q_0} = x_0^{m_j} (\gamma_j + \gamma'_j x_0^{n_j}) \quad \text{for all four roots } x_0 \text{ of } a.$$

Substituting $\xi = x_0^{q_0}$ and $x_0 = \xi^r$ as in the proof of Proposition 5.1, we get $1 + \xi + \xi^2 = c_1 \xi^{e_1} + c_2 \xi^{e_2}$ for all four primitive 8th roots of unity ξ , with $c_1 c_2 \neq 0$ and $0 \leq e_1, e_2 \leq 7$. So $x^4 + 1$ divides $1 + x + x^2 - c_1 x^{e_1} - c_2 x^{e_2}$. Modulo $x^4 + 1$ the two monomials change at most two of the three coefficients of $1 + x + x^2$. This is a contradiction.

Let $p = 2$. Then $q_0 = 4j$ and $q_B = 4(B - j)$, so $j = 2^{e_0}$ and $B - j = 2^{e_B}$ are powers of 2. Write $n_0 = q_0 n'_0$ with n'_0 odd. Then $1 + x^{n_0} = (1 + x^{4n'_0})^j = (a\Psi_0)^j$ with $\Psi_0 = \psi_{n'_0}(x^4)$, and likewise $1 + x^{n_B} = (a\Psi_B)^{B-j}$. With $a = byU$ and $y^2 U = 1 + U$ from (6) we get

$$g = y^j [C_0 U^j + g_j + C_B (1 + U)^{B-j}], \quad C_0 = \gamma_0 x^{m_0} (\Psi_0 b)^j, \quad C_B = \gamma_B x^{m_B} (\Psi_B b)^{B-j}.$$

By Lemma 3.6(b), $U^j = U + E_{e_0}$ and $(1 + U)^{B-j} = 1 + U + E_{e_B}$. Comparing coordinates in the basis $\{1, U\}$ gives $C_0 = C_B$ and $g_j = C_0(1 + E_{e_0} + E_{e_B})$. Let ω be a root of $1 + x + x^2$. The order of $\Psi_0 b$ at ω is 1 or 5, because b has a simple zero there and $1 + x^{4n'_0} = (1 + x^{n'_0})^4$ has a zero of order 4 or 0, while $a(\omega) \neq 0$. So the order of C_0 at ω is j or $5j$, and that of C_B is $B - j$ or $5(B - j)$. As j and $B - j$ are powers of 2, $C_0 = C_B$ forces $j = B - j$. Then $E_{e_0} = E_{e_B}$ and $g_j = C_0 = \gamma_0 x^{m_0} (\Psi_0 b)^j$. This has as many terms as $\Psi_0 b = \sum_{i < n'_0} (x^{4i+1} + x^{4i+2} + x^{4i+3})$, namely $3n'_0 \geq 3$. But g_j is a binomial. This is a contradiction.

So $\nu \geq 1$. By Lemma 3.4(a), $\nu = q_j := p^{v_p(n_j)}$, and (7) becomes

$$q_0 = \kappa j + q_j, \quad q_B = \kappa(B - j) + q_j, \quad q_0, q_j, q_B \text{ powers of } p \text{ with } q_j < q_0 \text{ and } q_j < q_B. \quad (8)$$

Step 3: reduction to the smallest q_j . Let p be odd. Then g_j vanishes at the four roots of a , so $a \mid g_j$. By Lemma 3.4(b), $g_\bullet = \gamma_\bullet x^{m_\bullet} (1 - z^{k_\bullet})$ and $q_\bullet = p^{v_p(k_\bullet)}$ for $\bullet \in \{0, j, B\}$. Put $Q = q_j$. By (8), Q divides q_0 and q_B , hence also j and B , and Q divides k_0, k_j, k_B . So $g = \sum_\bullet x^{m_\bullet} G_\bullet^Q$ with $G_\bullet = \gamma_\bullet^{1/Q} (1 - z^{k_\bullet/Q}) y^{h_\bullet/Q}$, where $h_\bullet = 0, j, B$. By Lemma 3.5, for each residue class of the $m_\bullet$ modulo Q the corresponding partial sum is in (f) . A non-zero element of (f) has at least three non-zero rows (Lemmas 4.1 and 4.2(i)). So the three $m_\bullet$ are congruent modulo Q , and $g = x^c H^Q$ with $H \in (f)$ of shape $(2, 2, 2)$ and with $q_j(H) = 1$. We replace g by H and assume $q_j = 1$. Then $j = q_0 - 1$ and $B - j = q_B - 1$ by (8), and $q_0, q_B \geq p$.

Let $p = 2$. Write $q_0 = 2^{d_0} q_j$ and $q_B = 2^{d_B} q_j$ with $d_0, d_B \geq 1$. Then (8) reads $4j = q_j(2^{d_0} - 1)$ and $4(B - j) = q_j(2^{d_B} - 1)$. So $4 \mid q_j$, and now $a = (1 + x)^4$ divides g_j . With $Q = q_j/4$, which divides n_0, n_j, n_B, j and B , the same argument as for odd p lets us assume $q_j = 4$. Then all $n_\bullet$ are divisible by 4, and with $z = x^4$ we have $g_\bullet = \gamma_\bullet x^{m_\bullet} (1 + z^{k_\bullet})$, where k_j is odd, and $j = 2^{d_0} - 1$, $B - j = 2^{d_B} - 1$.

In both cases put $q = q_0/q_j$ and $q' = q_B/q_j$, and write $k_0 = qk'_0$, $k_B = q'k'_B$ with $p \nmid k'_0 k'_B$; here $z = -x^4$ if p is odd and $z = x^4$ if $p = 2$. Then $q, q' \geq p$, $p \nmid k_j$,

$$j = q - 1, \quad B - j = q' - 1, \quad g_0 = \gamma_0 x^{m_0} (a\psi_0)^q, \quad g_j = \gamma_j x^{m_j} a\psi_j, \quad g_B = \gamma_B x^{m_B} (a\psi_B)^{q'}, \quad (9)$$

where $\psi_0 = \psi_{k'_0}(z)$, $\psi_j = \psi_{k_j}(z)$ and $\psi_B = \psi_{k'_B}(z)$. As a polynomial in x , $\psi_k(z)$ has degree $4(k - 1)$ and constant term 1.

Step 4: a computation in the basis $\{1, U\}$. We substitute $a = -byU$ and $y^2U = 1 - U$ from (6) into (9). Since $B = (q - 1) + (q' - 1)$, we have $y^{q'} U^{q'} y^B = y^q U (y^2 U)^{q'-1} = y^q U (1 - U)^{q'-1}$, and so

$$g = -bU y^q [A_0 U^{q-1} + A_j + A_B (1 - U)^{q'-1}]$$

with $A_0 = \gamma_0 x^{m_0} \psi_0^q b^{q-1}$, $A_j = \gamma_j x^{m_j} \psi_j$ and $A_B = \gamma_B x^{m_B} \psi_B^{q'} b^{q'-1}$ (if p is odd, $(-b)^{q-1} = b^{q-1}$). So the bracket vanishes in K .

Let p be odd. Multiply the bracket by ε and use Lemma 3.6(a). The coordinates in the basis $\{1, W\}$ give

$$A_0 (\delta^{(q-1)/2} - 1) = A_B (\delta^{(q'-1)/2} - 1), \quad (10)$$

$$A_0 \left(\frac{1}{4} - \delta^{(q+1)/2}\right) + A_B \left(\frac{1}{4} - \delta^{(q'+1)/2}\right) + \varepsilon A_j = 0. \quad (11)$$

Recall $\delta = D/(4b^2)$, and note $2^{q-1} = 1$ in $\mathbb{F}_p$. Put

$$\Pi_q = b^{q-1} (\delta^{(q-1)/2} - 1) = D^{(q-1)/2} - b^{q-1}, \quad \Omega_q = 4b^{q+1} \left(\frac{1}{4} - \delta^{(q+1)/2}\right) = b^{q+1} - D^{(q+1)/2}.$$

Since $D = -4x^8 + \dots - 4$ has degree 8 and b has degree 3 and $b(0) = 0$, the polynomial Π_q has degree $4(q - 1)$ and $\Pi_q(0) \neq 0$, and Ω_q has degree $4(q + 1)$ and $\Omega_q(0) \neq 0$. Now (10) says $\gamma_0 x^{m_0} \psi_0^q \Pi_q = \gamma_B x^{m_B} \psi_B^{q'} \Pi_{q'}$. Comparing the orders at $x = 0$ and the degrees gives $m_0 = m_B$ and $4(k'_0 - 1)q + 4(q - 1) = 4(k'_B - 1)q' + 4(q' - 1)$, that is, $k_0 = k_B$. Hence $q = q'$, $k'_0 = k'_B$, $\gamma_0 = \gamma_B$ and $A_0 = A_B$. Then (11), multiplied by $4b^2$, becomes

$$\gamma_0 x^{m_0} \psi_0^q \Omega_q + 2a^2 \gamma_j x^{m_j} \psi_j = 0.$$

Comparing degrees (after division by the powers of x) gives $4(k'_0 - 1)q + 4(q + 1) = 8 + 4(k_j - 1)$, that is, $k_j = k'_0 q$. But $p \nmid k_j$ and $p \mid q$. This is a contradiction.

Let $p = 2$, $q = 2^d$ and $q' = 2^{d'}$ with $d, d' \geq 1$. Multiply the bracket by ε and use Lemma 3.6(b). The coordinates in the basis $\{1, U\}$ give

$$A_0 E_d = A_B E_{d'}, \quad A_0 (\varepsilon + E_d) + \varepsilon A_j + \varepsilon A_B = 0. \quad (12)$$

Put $N_d = b^q E_d = \sum_{i < d} a^{2^{i+1}} b^{q-2^{i+1}}$. The summand with $i = d-1$ is a^q ; the others have smaller degree and vanish at $x = 0$. So N_d is a polynomial of degree $4q$ with $N_d(0) = 1$. It is divisible by a^2 , and N_d/a^2 has degree $4q - 8$ and constant term 1. The first equation in (12) says $\gamma_0 x^{m_0} \psi_0^q N_d = \gamma_B x^{m_B} \psi_B^{q'} N_{d'}$. Orders at 0 and degrees give $m_0 = m_B$ and $k'_0 q = k'_B q'$. Hence $q = q'$, $k'_0 = k'_B$, $\gamma_0 = \gamma_B$ and $A_0 = A_B$. The second equation then gives $A_j = A_0 E_d / \varepsilon$, that is,

$$\gamma_j x^{m_j} \psi_j = \gamma_0 x^{m_0} \psi_0^q \cdot b \cdot N_d / a^2.$$

Since $b = x(1 + x + x^2)$, comparing degrees after division by the powers of x gives $4(k_j - 1) = 4(k'_0 - 1)q + 2 + 4q - 8$, that is, $4k_j = 4k'_0 q - 2$. This is impossible. $\square$

7 Proofs of the theorems

Proof of Theorem 1.2. By Proposition 4.3 a non-zero element of (f) with fewer than seven terms has exactly six terms and shape $(2, 2, 2)$ or $(2, 1, 1, 2)$. These shapes are excluded by Propositions 5.1 and 6.1. $\square$

Proof of Theorem 1.1. The polynomial f is absolutely irreducible modulo every prime p by Lemma 3.1. By Theorem 1.2 the least number of terms of a non-zero multiple of f in $\mathbb{F}_p[x^{\pm 1}, y^{\pm 1}]$ is 7, attained by f . By Proposition 2.2, $M(\alpha) = 6$. $\square$

8 Machine checks

The proofs above do not depend on computation. The following computations, all in exact arithmetic, test them.

An exhaustive search for $p = 2$. For $p = 2$ we searched for all relations $X_1 + \dots + X_r = 0$ in $K_0 = \mathbb{F}_2(x)[y]/(f)$ between r pairwise distinct monomials $X_i = u^{n_i}$, $n_i \in \mathbb{Z}^2$, with no vanishing proper subsum. A multiple of f over $\mathbb{F}_2$ with r terms is such a relation or splits into shorter ones. The method is a descent along the binary digits of the exponents. As in Lemma 3.3(a), every $h \in K_0$ is uniquely of the form $h = \Lambda_0(h)^2 + x\Lambda_1(h)^2$. Normalise $X_r = 1$. After t digits, the vector with entries $u^{\lfloor n_i/2^t \rfloor}$ lies in the K_0 -subspace $V_t = \{Z \in K_0^r : \sum_i u^{n_i \bmod 2^t} Z_i^{2^t} = 0\}$, and equations for V_{t+1} are obtained from those for V_t and the next digits by applying Λ_0 and Λ_1 . A state consists of V_t and of the partition of the indices by the digits read so far. A state is dropped if it forces $Z_r = 0$ or a vanishing proper subsum. Otherwise the finitely many possible final vectors are tested, and every relation found is verified in K_0 . States are stored and not expanded twice. We have no a priori bound for the number of states: that the set of reachable states is finite is established by the terminating runs, not in advance. A run that terminates is exhaustive.

The runs for $r = 3, 4, 5, 6$ expanded 1, 11, 23 and 8004 states and found no relation between distinct monomials. (For $r = 6$ the search meets 48 solutions with coinciding monomials, which are discarded.) This confirms Theorem 1.2 for $p = 2$ and coefficients in $\mathbb{F}_2$, the case needed for Theorem 1.1, independently of Sections 4 to 6. Three variants of the search, without the symmetry reduction or without the pruning by subsums, gave the same result for $r = 5$ and $r = 6$. The search engine was validated on seven other curves $ay^2 + by + c$ over $\mathbb{F}_2$: in 17 comparisons its list of relations with 4, 5 or 6 terms in a box agreed with a brute-force enumeration.

The run for $r = 7$ expanded 38031 states. A second pass without stored states then listed all relations: every relation between seven distinct monomials is a monomial times f^{2^e} , $e \geq 0$. Following [DM18], call two finite sets $S, S' \subset \mathbb{Q}^2$ *equivalent* if $\lambda S = S' + v$ for some positive rational number

λ and some $v \in \mathbb{Q}^2$. By [DM18, Lemma 5] the equivalence classes of non-mixing sets of cardinality 7 correspond to classes of pre-broad points on $X_0 + \cdots + X_6 = 0$, and by Lemma 2.1 these points are the relations between seven distinct monomials. So, for $p = 2$, the computation shows that every non-mixing set of cardinality 7 is equivalent to $S(f)$. This statement rests on the computer search. We did not study the corresponding question for odd p .

Further checks. (i) The formulas of Lemma 3.6 and the statements on Π_q , Ω_q , N_d and N_d/a^2 were checked by machine for several p and q , and D is not a constant times a square modulo any odd prime below 5000. For $p = 3, 5, 7$ there are multiples of f with four rows whose two middle rows are monomials. On six of them the two predictions of the first part of the proof of Proposition 5.1 were confirmed.

(ii) A program enumerated all $\mathbb{F}_p$ -linear relations with at most six terms between the monomials $x^i y^j$, $0 \leq i \leq A$, $0 \leq j \leq B$, modulo f . It found none for $(p; A, B) = (2; 48, 24)$, $(3; 32, 16)$, $(5; 26, 13)$, $(7; 20, 10)$, $(11; 16, 8)$, $(13; 14, 7)$, in smaller boxes for $p = 17, 19, 23, 41, 43$, and in flat or tall boxes up to $(2; 200, 6)$ and $(2; 12, 100)$. A second program, written separately and based on a different method, found none for $(2; 128, 64)$, $(3; 80, 40)$, $(5; 64, 31)$, $(7; 52, 25)$, $(11; 30, 14)$, $(13; 26, 12)$ and in flat or tall boxes up to $(2; 400, 8)$ and $(2; 16, 200)$. A third program, also written separately, maps the monomials to $(\mathbb{F}_p[y]/(m))[x]/(f)$ for a random polynomial m and tests every hit by exact division. It found none for $(2; 88, 84)$, $(3; 60, 56)$, $(5; 42, 40)$ and $(7; 34, 32)$. In 20 small boxes, for f and four other polynomials, its list of the multiples with at most six terms (seven for f) agreed with a complete enumeration of all multiples in the box. With seven terms it found only monomial multiples of the powers f^{p^e} in the boxes $(2; 48, 24)$, $(3; 16, 12)$, $(5; 20, 10)$ and $(7; 12, 8)$.

(iii) By (3), the lowest and the highest row of a multiple with three or four non-zero rows are determined by the rows in between, through the polynomials $a^n S_n(t) \in \mathbb{F}_p[x]$. With arithmetic in $\mathbb{F}_p[x]$ only, a program found no multiple of shape $(2, 2, 2)$ with $B \leq 60$ and exponents up to 6000, and none of shape $(2, 1, 1, 2)$ with $B \leq 30$ and $|m_2 - m_1| \leq 200$, for the 15 primes $p \leq 47$. A second program tested shape $(2, 2, 2)$ for all $B \leq 200$ ($p \leq 13$) and $B \leq 120$ ($17 \leq p \leq 47$) with no bound on the exponents, and found nothing. As a positive control, b was replaced by $x + x^3$. Then $(1 + x^4)(1 + y^2) + (x + x^3)y$ has shape $(2, 2, 2)$, and the programs report it and its Frobenius twists. For $p = 2$ this control polynomial is reducible, being divisible by $(1 + x)^2$; there the control only tests that multiples with the prescribed row shapes are detected.

Verification. Theorems 1.1 and 1.2 are proved in Sections 2 to 7; the proof of Theorem 1.1 depends on the published results (S), (M) and (DM). The computations of Section 8 are checks, except that the statement on non-mixing sets of cardinality 7 for $p = 2$ rests on them. The programs and their outputs are in the source files.

1. The author's programs (the first ones in each item of Section 8) were all run again for this version, with the same results. The longest runs, the search for $r = 7$ and one variant of the search for $r = 6$ (696 151 states), take 40 to 45 minutes each.
2. A first independent verification run, which was AI-assisted, re-derived every step of the proofs in the form in which they were first written down, and repeated the runs of the author's programs, except the longest variant of the search for $r = 6$, with identical results. It also produced the second programs of Section 8, written without using the author's code, and Remark 2.3.
3. A second independent verification run, also AI-assisted, checked the text of this paper line by line, including three arguments that were rewritten after the first run (the proofs of Lemmas 3.1 and 3.3 and Step 1 in the proof of Proposition 6.1), and compared (S), (M), (DM) and [DM18,

Lemma 4] with the sources. It wrote the third program of Section 8 (ii) and a program that checks the displayed identities exactly (424 checks). It also ran the author’s programs and the searches of the second programs again from the source archive, with the same results.

Scope and priority. Theorem 1.1 answers Problem A of [War06] as stated there and in the dataset record [ula26]: the exact order of mixing is 6, and the condition on p holds for all primes. We do not determine the non-mixing sets of cardinality 7 for odd p . We read [Mas04] and [DM18] in the published versions (the latter in the form first published online, from the first author’s web page), and [War06], [EW03] and [LS19] in the freely available versions. Of [DM12] we read only the introduction, in the arXiv version, and we did not see [DM15]. The statements (S), (M), (DM) and the numbering of the lemmas are taken from the published versions. We did not consult Schmidt’s book [Sch95], [KS93] or [Led78]; what we say about them is taken from [Mas04, DM18, LS19, War06]. The textbooks [Sti09, Ful89] are cited for standard facts and were not consulted again for this note. The reduction in Proposition 2.2 is an application of [Mas04, DM18]. The procedure of Derksen and Masser is effective, and they point out the relation to shortest polynomials; we claim only what seems to be the first explicit determination of M for Ward’s example, uniform in p . We searched arXiv, Crossref, OpenAlex and zbMATH, the lists of works citing [EW03, KS93, Mas04, DM12, DM15, DM18], and the web, and found no determination of M for this example. This negative search is not a proof of priority.

AI-assisted tools were used to search the problem corpus and the literature, to find the arguments, to write and run the verification programs, and to help prepare this manuscript. The author has reviewed the proofs, the computations and the sources and is responsible for all claims. The note has not been peer reviewed.

References

- [DM12] H. Derksen and D. Masser. Linear equations over multiplicative groups, recurrences, and mixing I. *Proceedings of the London Mathematical Society*, 104(5):1045–1083, 2012. Preprint: arXiv:1010.4519. doi:10.1112/plms/pdr040.
- [DM15] H. Derksen and D. Masser. Linear equations over multiplicative groups, recurrences, and mixing II. *Indagationes Mathematicae*, 26(1):113–136, 2015. doi:10.1016/j.indag.2014.08.002.
- [DM18] H. Derksen and D. Masser. Linear equations over multiplicative groups, recurrences, and mixing III. *Ergodic Theory and Dynamical Systems*, 38(7):2625–2643, 2018. doi:10.1017/etds.2016.137.
- [EW03] M. Einsiedler and T. Ward. Asymptotic geometry of non-mixing sequences. *Ergodic Theory and Dynamical Systems*, 23(1):75–85, 2003. Preprint: arXiv:math/0204174. doi:10.1017/S0143385702000950.
- [Ful89] W. Fulton. *Algebraic Curves: An Introduction to Algebraic Geometry*. Advanced Book Classics. Addison-Wesley, Redwood City, CA, 1989. Reprint of the 1969 original.
- [KS93] B. Kitchens and K. Schmidt. Mixing sets and relative entropies for higher-dimensional Markov shifts. *Ergodic Theory and Dynamical Systems*, 13(4):705–735, 1993. doi:10.1017/S014338570000763X.
- [Led78] F. Ledrappier. Un champ markovien peut être d’entropie nulle et mélangeant. *Comptes Rendus de l’Académie des Sciences de Paris, Séries A et B*, 287(7):A561–A563, 1978.

- [LS19] D. Lind and K. Schmidt. Mahler’s work and algebraic dynamical systems. In *The Legacy of Kurt Mahler: A Mathematical Selecta*, volume 8 of *Documenta Mathematica Series*, pages 57–77. EMS Press, 2019. Preprint: arXiv:1803.05862. doi:10.4171/dms/8/6.
- [Mas04] D. W. Masser. Mixing and linear equations over groups in positive characteristic. *Israel Journal of Mathematics*, 142:189–204, 2004. doi:10.1007/BF02771532.
- [Sch95] K. Schmidt. *Dynamical Systems of Algebraic Origin*, volume 128 of *Progress in Mathematics*. Birkhäuser, Basel, 1995. doi:10.1007/978-3-0348-0277-2.
- [Sti09] H. Stichtenoth. *Algebraic Function Fields and Codes*, volume 254 of *Graduate Texts in Mathematics*. Springer, Berlin, second edition, 2009. doi:10.1007/978-3-540-76878-4.
- [ula26] ulamai. UnsolvedMath. Hugging Face dataset, <https://huggingface.co/datasets/ulamai/UnsolvedMath>, 2026. Version 1.6.0; record AMR-042-0001.
- [War06] T. Ward. Six problems in algebraic dynamics (updated December 2006). Problem list, <https://www.imath.kiev.ua/~skolyada/kevin.pdf>, 2006. Accessed October 2026.