Joachim von zur Gathen (2010). Shift-invariant polynomials and Ritt's Second Theorem. In *Finite Fields : Theory and Applications*, Gary McGuire, Gary L. Mullen, Daniel Panario & Igor E. Shparlinski, editors, volume 518 of *Contemporary Mathematics*, 161–184. ISBN 0-8218-4786-4 (ISBN-10), 978-0-8218-4786-2 (ISBN-13). The *Local PDF* is a corrected version.

Shift-invariant polynomials and Ritt's Second Theorem

Joachim von zur Gathen

Abstract. Ritt's Second Theorem deals with compositions $g \circ h = g^* \circ h^*$ of univariate polynomials over a field, where $\deg g = \deg h^*$. Joseph Fels Ritt (1922) presented two types of such decompositions. His main result here is that these comprise all possibilities, up to some linear transformations. A recently established normal form describes Ritt's compositions concisely. This form is unique unless the characteristic divides the larger of the two component degrees. The present paper studies this case, which is best understood with methods from invariant theory. Examples of nonuniqueness are presented, as well as a method for determining all of them and estimating their number. Some of the results are conjectural.

Section 1. Introduction

For several problems concerning the composition of polynomials (univariate over a field), one has to understand the "collisions" $g \circ h = g^* \circ h^*$. There are two obvious types of such collisions, called *exponential* and *trigonometric*; details are given below. Ritt's Second Theorem, from Ritt (1922), is a celebrated result in this area. It says that, under certain mild conditions, the above two types are essentially the only possibilities. The history and the sharpest previous versions of this result are given in Zannier (1993) and Schinzel (2000).

In the usual versions of Ritt's Second Theorem, a total of four unspecified linear functions appear. This makes the Theorem somewhat difficult to use, since a uniqueness property in Ritt's Second Theorem is not obvious. Indeed, Beardon & Ng (2000) are puzzled by its absence. On their page 128, they write, translated to the notation used below, "Now these rules are a little less transparent, and a little less independent, than may appear at first sight. First, we note that [the First Case], which is stated in its conventional form, is rather loosely defined, for the k and w are not uniquely determined by the form $x^k w(x^\ell)$; for instance, if $w(0) = 0$, we can equally well write this expression in the form $x^{k+\ell}\tilde{w}(x^\ell)$, where $\tilde{w} = w/x$. Next, $T_2(x, 1) = x^2 - 2$ differs by a linear component from x^2, so that in some circumstances it is possible to apply [the Second Case] to $T_2(x, 1)$, then [a linear composition], and then (on what is essentially the same factor) [the Second Case]. These observations perhaps show why it is difficult to use Ritt's result."

2000 *Mathematics Subject Classification.* 68W30, 11T06, 12E05, 13A50, 05A16.

These well-motivated concerns are essentially settled in von zur Gathen (2008b), where a normal form for such collisions is provided. The exponential collisions ("First Case") are parametrized by a polynomial w and a field element a; see Fact 2.1 below. This normal form is uniquely determined by the composition and the larger value m of the two component degrees involved, provided that the characteristic p does not divide m. The present paper studies the ambiguities in this parametrization when p does divide m.

For perspective, we note that the *tame case*, where p does not divide the degree of the composition, is now reasonably well understood. The present contribution deals with the more difficult *wild case*, with divisibility by p.

There is a sequence of papers whose overall goal it is to approximate, with small relative error, the number of decomposable polynomials of degree n over $\mathbb{F}_q$. One readily finds that the major contribution to this number comes from components of degrees ℓ and n/ℓ, where ℓ is the smallest prime factor of the composite integer n. An essential step is to estimate the number of collisions with these degrees. The present paper throws more light on these collisions, in a special case. See von zur Gathen (2008a) for the counting result, and also von zur Gathen (2008b) for the multivariate case. Bodin *et al.* (2009) state estimates for these problems.

Section 2. Distinct-degree collisions of decompositions

A nonzero polynomial $f \in F[x]$ over a field F is *monic* if its leading coefficient equals 1. We call f *original* if its graph contains the origin, that is, $f(0) = 0$. For $g, h \in F[x]$,

$$f = g \circ h = g(h) \in F[x]$$

is their *composition*. If $\deg g, \deg h \geq 2$, then (g, h) is a *decomposition* of f. One can normalize any decomposition so that h is monic and original. By a harmless (and unique) linear transformation, one may also assume f and g to be monic and original. See von zur Gathen (2008a) for more details.

The following is an example of a collision, called *exponential*:

$$x^k w^\ell \circ x^\ell = x^{k\ell} w^\ell(x^\ell) = x^\ell \circ x^k w(x^\ell),$$

for any polynomial $w \in F[x, y]$, where F is a field (or even a ring). We define the (bivariate) *Dickson polynomials of the first kind* $T_m \in F[x, y]$ by $T_0 = 2$, $T_1 = x$, and

$$T_m = xT_{m-1} - yT_{m-2} \text{ for } m \geq 2.$$

The monograph of Lidl, Mullen & Turnwald (1993) provides extensive information about these polynomials. We have $T_m(x, 0) = x^m$, and $T_m(x, 1)$ is closely related to the *Chebyshev polynomial* $C_n = \cos(n \arccos x)$, as $T_n(2x, 1) = 2C_n(x)$. T_m is monic (for $m \geq 1$) of degree m, and

$$T_m = \sum_{0 \leq i \leq m/2} \frac{m}{m-i} \binom{m-i}{i} (-y)^i x^{m-2i} \in F[x, y].$$

Furthermore,

$$T_m(x, y^\ell) \circ T_\ell(x, y) = T_{\ell m}(x, y) = T_\ell(x, y^m) \circ T_m(x, y),$$

and if $\ell \neq m$, then substituting any $z \in F$ for y yields a collision, called *trigonometric*.

Ritt's Second Theorem is the central tool for understanding distinct-degree collisions. It says that, under certain conditions, the examples above are essentially the only distinct-degree collisions. It was first proved by Ritt (1922). He worked with $F = \mathbb{C}$ and used analytic methods. Subsequently, his approach was replaced by algebraic methods, in the work of Levi (1942) and Dorey & Whaples (1974), and Schinzel (1982) presented an elementary but long and involved argument. Thus Ritt's Second Theorem was also shown to hold in positive characteristic p. The original versions of this required $p > \deg(g \circ h)$. Zannier (1993) reduced this to the milder and more natural requirement $g'(g^*)' \neq 0$. His proof works over an algebraically closed field, and Schinzel's 2000 monograph adapts it to finite fields. The following normal form is proved in von zur Gathen (2008b).

Fact 2.1. *Let F be a field of characteristic p, let $m > \ell \geq 2$ be integers, and $n = \ell m$. Furthermore, we have monic original $f, g, h, g^*, h^* \in F[x]$ satisfying*

$$\gcd(\ell, m) = 1, \deg g = \deg h^* = m, \deg h = \deg g^* = \ell, \tag{2.2}$$

$$f = g \circ h = g^* \circ h^*, \tag{2.3}$$

$$g'(g^*)' \neq 0, \tag{2.4}$$

where $g' = \partial g/\partial x$ is the derivative of g. Then either (i) or (ii) hold, and (iii) is also valid.

(i) *(First Case) There exist a monic polynomial $w \in F[x]$ of degree s and $a \in F$ so that*

$$f = (x - a^{k\ell} w^\ell(a^\ell)) \circ x^{k\ell} w^\ell(x^\ell) \circ (x + a), \tag{2.5}$$

where $m = s\ell + k$ is the division with remainder of m by ℓ, with $1 \leq k < \ell$. Furthermore

$$kw + \ell x w' \neq 0 \text{ and } p \nmid \ell. \tag{2.6}$$

Conversely, any (w, a) for which (2.6) holds yields a collision satisfying (2.2) through (2.4) via the above formulas. If $p \nmid m$, then (w, a) is uniquely determined by f and ℓ.

(ii) *(Second Case) There exist $z, a \in F$ with $z \neq 0$ so that*

$$f = (x - T_n(a, z)) \circ T_n(x, z) \circ (x + a).$$

Now (z, a) is uniquely determined by f. Furthermore we have $p \nmid n$. Conversely, if $p \nmid n$, then any (z, a) as above yields a collision satisfying (2.2) through (2.4) via the above formulas.

(iii) *When $\ell \geq 3$, the First and Second Cases are mutually exclusive. For $\ell = 2$, the Second Case is included in the First Case.*

In each case, there are also explicit formulas for the four components, which we omit. Based on this normal form, one can determine the number of distinct-degree collisions exactly in the tame case. One generalization covers the case where $g'(g^*)'$ is allowed to vanish. A second generalization allows ℓ and m to have a nontrivial gcd, but assumes that $p \nmid \ell m$; this is based on a result by Tortrat (1988).

The goal in the present paper is to investigate the (lack of) uniqueness in (2.5). As a simplification, we leave out the left hand linear component. (5.2) and Conjecture 8.3 justify this. Furthermore, $x^{k\ell} w^\ell(x^\ell) \circ (x+a) = x^\ell \circ x^k w(x^\ell) \circ (x+a)$,

and if two such expressions are equal, then so are their ℓth roots, since w is monic. Thus we set for monic $w \in F[x]$ and $a \in F$

$$\rho_{w,a} = x^k w(x^\ell) \circ (x+a) \tag{2.7}$$

and ask:

$$\text{when is } \rho_{w,a} = \rho_{\tilde{w},\tilde{a}} \text{ for monic } w, \tilde{w} \in F[x] \text{ and } a, \tilde{a} \in F? \tag{2.8}$$

Section 3. The ring of invariants under additive shifts

We embed the additive group of a field F of positive characteristic p into $\mathrm{GL}(2,F)$ by mapping $a \in F$ to

$$\begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}(2,F).$$

The minus sign makes the notation compatible with the nomenclature in the motivating example (2.7), since the resulting action τ on $F[x,y]$ is given by

$$\tau_a(x) = x + ay, \quad \tau_a(y) = y.$$

This is the 2-dimensional special case of transvections; see e.g., Neusel & Smith (2002), Section 6.2. For an additive subgroup $G \subseteq F$, this induces an action of G on $F[x,y]$. A standard task of invariant theory is, in this special case, to determine the ring of invariants

$$F[x,y]^G = \{f \in F[x,y] \colon \forall a \in G \quad \tau_a(f) = f\}.$$

Our application is somewhat nonstandard, in that we start, a priori, with arbitrary subsets $G \subseteq F$ and $P \subseteq F[x,y]$. We first note that $F[x,y]^G$ and the stabilizer (or isotropy group)

$$\mathrm{stab}^P = \{a \in F \colon \forall f \in P \quad \tau_a(f) = f\} \subseteq F$$

are subrings and subgroups, respectively.

Lemma 3.1. *Let $P \subseteq F[x,y]$ and $G \subseteq F$ be nonempty.*

(i) *$\mathrm{stab}^P \subseteq F$ is an additive subgroup.*
(ii) *$F[x,y]^G \subseteq F[x,y]$ is a subring containing F.*
(iii) *If $G = \{0\}$, then $F[x,y]^G = F[x,y]$.*

Proof. (i) For $a, b \in \mathrm{stab}^P$, we have for all $f \in P$

$$f(x + (a-b)y) = f((x-by) + ay) = f(x-by) = f((x-by)+by) = f(x).$$

(ii) and (iii) are clear. □

This action provides an (antitone) Galois correspondence between the subgroups of F and the subrings of $F[x,y]$.

For an integer r, we denote as $r^{\mathbb{N}_{\geq 1}} = \{r^i \colon i \geq 1\}$ the set of powers of r. When $r \in p^{\mathbb{N}_{\geq 1}}$ and $\mathbb{F}_r \subseteq F$, a polynomial $\lambda \in F[x]$ is an *r-polynomial* (*or linearized*) if and only if $\lambda(a+b) = \lambda(a) + \lambda(b)$ and $\lambda(ua) = u\lambda(a)$ for all $a, b \in F$ and $u \in \mathbb{F}_r$. Equivalently, x^i has a nonzero coefficient in λ only if $i \in r^{\mathbb{N}_{\geq 1}}$. For a finite additive subgroup $G \subseteq F$, we let

$$\lambda^G = \prod_{a \in G} (x-a) = \prod_{a \in G} (x+a).$$

Then λ^G is a p-polynomial of degree $\#G$. Since $0 \in G$, the coefficient λ_0^G of x in $\lambda^G = \prod_{a\in G}(x-a)$ equals

$$\lambda_0^G = \prod_{\substack{a\in G\\ a\neq 0}} a \neq 0. \tag{3.2}$$

(3.2) implies that $(x^{-1}\lambda^G)(0) \neq 0$. We let $\tilde\lambda^G = y^{\#G}\cdot\lambda^G(x/y) = \prod_{a\in G}(x-ay) \in F[x,y]$ be the homogenization of λ^G.

A result of Kemper (1996) determines $F[x,y]^G$ as follows. This is also shown in Smith (1995), Theorem 8.2.13.

Theorem 3.3. *Let $G \subseteq F$ be a finite subgroup. Then*

$$F[x,y]^G = F[\tilde\lambda^G, y].$$

Proof. For any $b \in G$, we have

$$\tau_b(\tilde\lambda^G) = \prod_{a\in G}\big((x-by)-ay\big) = \tilde\lambda^G,$$

so that $\tilde\lambda^G$ is invariant. So is y, and the two are algebraically independent. We have $\deg\lambda^G\cdot\deg y = \#G$, and Kemper (1996), Proposition 16, implies the claim. □

Theorem 3.3 generalizes a result of Landweber & Stong (1987), who show that for $G = F = \mathbb{F}_q$, we have

$$\mathbb{F}_q[x,y]^G = \mathbb{F}_q[x^q - xy^{q-1}, y];$$

see also Smith (1995), Proposition 8.2.5. Here, $x^q - xy^{q-1} = \tilde\lambda^G$. Almkvist (1983) considers the situation $G = \mathbb{F}_p \subseteq F = \mathbb{F}_q$.

Invariant theory usually works with homogeneous polynomials. But our motivating question is inhomogeneous, and for just two variables, the inhomogeneous version becomes typographically somewhat simpler. In content, the two versions are equivalent, and the transitions both ways are standard. Namely, a homogeneous $f \in F[x,y]$ becomes $f(x,1) \in F[x]$, and $g \in F[x]$ of degree n becomes $y^n g(x/y) \in F[x,y]$. In this language, corresponding to τ we have the *shift* (or *translation*, or *transvection) action* σ of the additive group F on $F[x]$, given by $\sigma_a(f) = f\circ(x+a)$ for $a \in F$ and $f \in F[x]$. For $P \subseteq F[x]$,

$$\operatorname{stab}^P = \{a \in F\colon \forall f \in P\ \sigma_a(f) = f\} \subseteq F \tag{3.4}$$

is the *stabilizer* of P, an additive subgroup of F. For $G \subseteq F$, we let

$$F[x]^G = \{f \in F[x]\colon \forall a \in G\ \sigma_a(f) = f\}$$

be the ring of polynomials that are invariant under shifts from G. In the case of singletons, we write stab^f and $F[x]^a$ for $\operatorname{stab}^{\{f\}}$ and $F[x]^{\{a\}}$, respectively. In characteristic 0, there are no nonconstant shift-invariant polynomials. σ corresponds to the homogeneous transvection action ρ of G on $F[x,y]$. The inhomogeneous version of Theorem 3.3 reads as follows.

Corollary 3.5. *Let $G \subseteq F$ be a finite additive subgroup. Then*

$$F[x]^G = F[\lambda^G].$$

For all $a \in G$ and $f \in F[x]^G$, we have $f(a) = f(0)$. If $G \subset H \subseteq F$ are subgroups, then $\deg \lambda^G < \deg \lambda^H$. It follows that $\mathrm{stab}^{F[x]^G} = G$.

As an aside, we note that $F(x) \supseteq F(\lambda^G)$ is the splitting field of the irreducible polynomial $\lambda^G(t) - \lambda^G(x) \in F(\lambda^G)[t]$, where t is a new indeterminate. Its Galois group is G.

Section 4. Shift-invariant polynomials in x^ℓ

We now consider the following variation on shift-invariant polynomials. We are given a positive integer ℓ coprime to p and a finite additive subgroup $G \subseteq F$, and set

$$F[x]_\ell^G = \{w \in F[x] \colon w(x^\ell) \in F[x]^G\}.$$

Theorem 4.1. *Let $G \subseteq F$ be a finite additive subgroup and $\ell \geq 1$. The following hold.*

(i) *If $x^{-1}\lambda^G \notin F[x^\ell]$, then $F[x]_\ell^G = F$.*
(ii) *If $x^{-1}\lambda^G \in F[x^\ell]$, say $x^{-1}\lambda^G = u(x^\ell)$ with $u \in F[x]$, then $p \nmid \ell$ and*

$$F[x]_\ell^G = F[xu^\ell].$$

Proof. Both claims are clear when $G = \{0\}$ or $\ell = 1$, and we now assume $G \neq \{0\}$ and $\ell \geq 2$.

(i) We always have $F \subseteq F[x]_\ell^G$, and so suppose that $w \in F[x]_\ell^G \smallsetminus F$. Thus $w(x^\ell) \in F[x]^G = F[\lambda^G]$, so that $w(x^\ell) = v(\lambda^G)$ for some $v \in F[x] \setminus F$. We let $\#G = p^d$ with $d \geq 1$ and write

$$\lambda^G = \sum_{0 \leq i \leq d} \lambda_i x^{p^i}, \quad v = \sum_j v_j x^j \tag{4.2}$$

with all $\lambda_i, v_j \in F$ and $\lambda_0 \neq 0$, by (3.2). We let

$$I = \{i \leq d \colon \lambda_i \neq 0, \ell \nmid p^i - 1\},$$
$$b = \min\{j \colon j \geq 1 \text{ and } v_j \neq 0\}.$$

We first assume $I \neq \varnothing$, and let $h = \min I$. The unique term of smallest positive degree in

$$v(\lambda^G) = v_0 + v_b(\lambda^G)^b + \cdots \tag{4.3}$$

is $v_b\lambda_0^b x^b$, so that $\ell \mid b$, since $v(\lambda^G) \in F[x^\ell]$. Now the term in (4.3) of the smallest degree not divisible by ℓ is

$$v_b \cdot (\lambda_0 x)^{b-1}\lambda_h x^{p^h} = v_b\lambda_0^{b-1}\lambda_h x^{b+p^h-1}.$$

Now it follows that $\ell \mid b + p^h - 1$ and hence $\ell \mid p^h - 1$, contradicting $h \in I$. We conclude that $I = \varnothing$ and hence $x^{-1}\lambda^G \in F[x^\ell]$.

(ii) We first note that $\ell \mid \deg(x^{-1}\lambda^G) = p^d - 1$, so that $p \nmid \ell$. Furthermore, we have

$$x^\ell \circ \lambda^G = x^\ell \circ (x \cdot u(x^\ell)) = xu^\ell \circ x^\ell \in F[x^\ell], \tag{4.4}$$

and now show that

$$F[x]_\ell^G \circ x^\ell = F[x^\ell \circ \lambda^G]. \tag{4.5}$$

For one inclusion, let $w \in F[x]_\ell^G$. Then $w \circ x^\ell \in F[x]^G = F[\lambda^G]$, say $w \circ x^\ell = v \circ \lambda^G$ for some $v \in F[x]$. We claim that $v \in F[x^\ell]$, and use induction on $\deg w \geq 0$. The claim is clear for $\deg w = 0$, since then $v \in F$. For $\deg w > 0$, we write $t = \deg v$ and have $\ell \cdot \deg w = t \cdot \#G$ and $\gcd(\ell, \#G) = 1$, so that $\ell \mid t$ and $x^t \in F[x^\ell]$. If $v = x^t$, the claim is proven. Otherwise, we have from (4.4)

$$x^t \circ \lambda^G = x^{t/\ell} \circ xu^\ell \circ x^\ell \in F[x^\ell].$$

We let $w^* = w - x^{t/\ell} \circ xu^\ell$. Then

$$\begin{aligned} w^* \circ x^\ell &= w \circ x^\ell - x^{t/\ell} \circ xu^\ell \circ x^\ell \\ &= v \circ \lambda^G - x^{t/\ell} \circ x^\ell \circ \lambda^G \\ &= (v - x^t) \circ \lambda^G \in F[\lambda^G] = F[x]^G, \end{aligned}$$

so that $w^* \in F[x]_\ell^G$ and $\deg w^* < \deg w$. By induction, we have $v - x^t \in F[x^\ell]$ and hence $v \in F[x^\ell]$, as claimed. Writing $v = v^* \circ x^\ell$ with $v^* \in F[x]$, we have

$$\begin{aligned} w \circ x^\ell = v^* \circ x^\ell \circ \lambda^G \in F[x^\ell \circ \lambda^G], \\ F[x]_\ell^G \circ x^\ell \subseteq F[x^\ell \circ \lambda^G]. \end{aligned}$$

For the reverse inclusion in (4.5), we take $v \in F[x]$ and $w = v \circ xu^\ell$. Then

$$v \circ x^\ell \circ \lambda^G = v \circ xu^\ell \circ x^\ell = w \circ x^\ell,$$

so that $w \in F[x]_\ell^G$, $v \circ x^\ell \circ \lambda^G \in F[x]_\ell^G \circ x^\ell$, and hence $F[x^\ell \circ \lambda^G] \subseteq F[x]_\ell^G \circ x^\ell$. This proves (4.5).

For any $w \in F[x]$, we have

$$\begin{aligned} w \in F[x]_\ell^G &\iff w \circ x^\ell \in F[x]_\ell^G \circ x^\ell = F[x^\ell \circ \lambda^G] = F[xu^\ell \circ x^\ell] = F[xu^\ell] \circ x^\ell \\ &\iff w \in F[xu^\ell]. \end{aligned}$$

□

In terms of the coefficients (4.2) of λ^G, we set

$$\mu_\ell^G = x\Big(\sum_{0 \leq i \leq d} \lambda_i x^{(p^i-1)/\ell}\Big)^\ell = xu^\ell = x \cdot (x^{-1}\lambda^G(x^{1/\ell}))^\ell,$$

where the assumption $x^{-1}\lambda^G \in F[x^\ell]$ justifies the notation $x^{1/\ell}$. Then we have shown that

$$F[x]_\ell^G = \begin{cases} F[\mu_\ell^G] & \text{if } x^{-1}\lambda^G \in F[x^\ell], \\ F & \text{otherwise.} \end{cases}$$

For all $a \in G$ and $w \in F[x]_\ell^G$, we have $\mu_\ell^G(a^\ell) = 0$ and $w(a^\ell) = w(0)$.

$\mathbb{F}_q$ contains a primitive ℓth root of unity ζ if and only if $\ell \mid q - 1$. Kemper (2009) has pointed out that if F contains such a ζ, then $F[x^\ell] = F[x]^{\langle\zeta\rangle}$ and $F[x^\ell] \cap F[x]^G = F[x]^H$, where $\langle\zeta\rangle$ is the multiplicative group generated by ζ acting on $F[x]$ via $x \mapsto \zeta x$, and $H = \langle G, \zeta\rangle$ consists of all transformations of the form $x \mapsto \zeta^i x + a$ with a a sum of terms $\zeta^j b$ with $b \in G$. In Theorem 4.1(ii), one can conclude that $F[x]^H = F[(\lambda^G)^\ell]$ and $\#H = \ell\#G$, so that H is a semidirect product of G and $\langle\zeta\rangle$. It is not clear whether this observation may lead to a simpler proof of Theorem 4.1.

We can characterize the assumption in Theorem 4.1(ii) as follows.

Lemma 4.6. *Let F be a field of characteristic p, let $\ell \geq 1$ with $p \nmid \ell$, $c = \operatorname{ord}_\ell p$, $r = p^c$, and $G \subseteq F$ a nonzero finite additive subgroup. The following hold.*

(i)

$$\begin{aligned} \mathbb{F}_r \subseteq F \text{ and } G \text{ is } \mathbb{F}_r\text{-linear} &\iff \lambda^G \text{ is an } r\text{-polynomial} \\ &\iff x^{-1}\lambda^G \in F[x^\ell]. \end{aligned}$$

(ii) *If $x^{-1}\lambda^G \in F[x^\ell]$, then F contains a primitive ℓth root of unity.*

Proof. (i) Let λ^G be an r-polynomial $\sum \lambda_i x^{r^i}$, and embed F and $\mathbb{F}_r$ in a common superfield. Then for all $a \in G$, $z \in \mathbb{F}_r$, and $i \in \mathbb{N}$ we have $z^{r^i} = z$ and $\lambda^G(za) = z\lambda^G(a) = 0$, so that $za \in G \subseteq F$ and G is $\mathbb{F}_r$-linear. Taking a nonzero $a \in G \subseteq F$, we find $z = za/a \in F$, and thus $\mathbb{F}_r \subseteq F$. The other direction in the first equivalence follows from Theorem 3.52 of Lidl & Niederreiter (1983). (Their context and statement assume F to be finite, but their explicit calculations in the proof do not make use of this assumption.) For the second equivalence, we write $\#G = p^d$ and $\lambda^G = \sum_{0 \leq i \leq d} \lambda_i x^{p^i}$ with all $\lambda_i \in F$. Then

$$\begin{aligned} &\lambda^G \text{ is an } r\text{-polynomial} \\ &\iff \forall i \leq d \quad \lambda_i = 0 \text{ or } c \mid i \\ &\iff \forall i \leq d \quad \lambda_i = 0 \text{ or } p^i \equiv 1 \bmod \ell \\ &\iff \forall i \leq d \quad \lambda_i = 0 \text{ or } \ell \mid p^i - 1 \\ &\iff x^{-1}\lambda^G \in F[x^\ell]. \end{aligned}$$

(ii) Let s be the largest power of p so that λ^G is an s-polynomial, and write $\lambda^G = \sum_i \lambda_i x^{s^i}$ and $I = \{i \geq 0 \colon \lambda_i \neq 0\}$. Then $\gcd(I) = 1$, by the maximality of s. We take $t_i \in \mathbb{Z}$ for $i \in I$ with $\sum_{i \in I} i t_i = 1$. For all $i \in I$, we have $x^{s^i - 1} \in F[x^\ell]$ and hence $s^i \equiv 1 \bmod \ell$. It follows that

$$s - 1 = \prod_{i \in I} (s^i)^{t_i} - 1 \equiv 0 \bmod \ell.$$

(We can also conclude that $s \in r^{\mathbb{N}_{\geq 1}}$.) Using (i), we have $\mathbb{F}_s \subseteq F$ and $\ell \mid s - 1$, so that $\mathbb{F}_s$ contains a primitive ℓth root of unity. □

Lidl & Niederreiter (1983) also give an explicit description of λ^G in terms of an $\mathbb{F}_r$-basis, via "q-Vandermonde" determinants.

Section 5. Shift-invariant polynomials of the form $x^k w(x^\ell)$

Next we take positive integers k and ℓ with $p \nmid \ell$, a finite additive subgroup $G \subseteq F$, and, with a view to (2.7), we set

$$F[x]^G_{k,\ell} = \{w \in F[x] \colon x^k w(x^\ell) \in F[x]^G\}.$$

Theorem 5.1. *Let k and ℓ be positive integers with $p \nmid \ell$, and let $G \subseteq F$ be a finite additive subgroup. If $x^{-1}\lambda^G \in F[x^\ell]$, say $x^{-1}\lambda^G = u \circ x^\ell$ with $u \in F[x]$, then $F[x]^G_{k,\ell} = u^k \cdot F[x]^G_\ell = u^k \cdot F[xu^\ell]$, and for $a \in G$ and $w \in F[x]^G_{k,\ell}$ we have $a^k w(a^\ell) = 0$.*

Proof. The claim is trivial for $G = \{0\}$, and we now assume $G \neq \{0\}$.

For the inclusion "$\subseteq$", we take some $w \in F[x]^G_{k,\ell}$, so that $x^k w(x^\ell) \in F[x]^G = F[\lambda^G]$, and there is some $h_1 \in F[x]$ with $x^k w(x^\ell) = h_1 \circ \lambda^G$. By (3.2) and since $x^k \mid h_1 \circ \lambda^G$, it follows that $x^k \mid h_1$. We set $h = x^{-k} h_1 \in F[x]$. Then

$$\begin{aligned} x^k w(x^\ell) &= (x^k h) \circ \lambda^G = (x \cdot x^{-1}\lambda^G)^k \cdot (h \circ \lambda^G) \\ &= x^k \cdot (u \circ x^\ell)^k \cdot (h \circ \lambda^G), \\ w \circ x^\ell &= (u^k \circ x^\ell) \cdot (h \circ \lambda^G), \\ u^k \circ x^\ell &\mid w \circ x^\ell. \end{aligned}$$

We take the division with remainder $w = su^k + r$ with $s, r \in F[x]$ and $\deg r < \deg u^k$. Then

$$\begin{aligned} u^k \circ x^\ell \mid w \circ x^\ell - (s \circ x^\ell) \cdot (u^k \circ x^\ell) \\ = (w - su^k) \circ x^\ell = r \circ x^\ell, \end{aligned}$$

so that $r = 0$ and $u^k \mid w$. It follows that

$$\begin{aligned} (u^{-k} w) \circ x^\ell = h \circ \lambda^G \in F[\lambda^G] = F[x]^G, \\ u^{-k} w \in F[x]^G_\ell = F[xu^\ell], \\ w \in u^k \cdot F[xu^\ell], \end{aligned}$$

where we have used Corollary 3.5 and Theorem 4.1.

We have shown one inclusion. For the reverse, we take some $w \in u^k \cdot F[xu^\ell]$, so that $w = u^k \cdot v(xu^\ell)$ for some $v \in F[x]$. Then, using (4.4) we find

$$\begin{aligned} x^k w(x^\ell) &= x^k \cdot \big((u^k \cdot v(xu^\ell)) \circ x^\ell\big) \\ &= x^k \cdot (u \circ x^\ell)^k \cdot (v \circ xu^\ell \circ x^\ell) \\ &= x^k \cdot (x^{-1}\lambda^G)^k \cdot (v \circ x^\ell \circ \lambda^G) \\ &= (x^k \cdot (v \circ x^\ell)) \circ \lambda^G \in F[\lambda^G] = F[x]^G, \end{aligned}$$

and hence $w \in F[x]^G_{k,\ell}$. We conclude that $u^k \cdot F[x]^G_\ell = F[x]^G_{k,\ell}$. The last claim follows from

$$a^k w(a^\ell) = (x^k w(x^\ell))(a) = (h_1 \circ \lambda^G)(a) = h_1(0) = 0. \tag{5.2}$$

□

When the assumption of Theorem 5.1 is not satisfied, we offer the following conjecture.

Conjecture 5.3. *Let $\ell < p$ be a prime, $1 \le k < \ell$, and let $G \subseteq F$ be a nonzero finite additive subgroup. If $x^{-1}\lambda^G \notin F[x^\ell]$, then $F[x]^G_{k,\ell} = \{0\}$.*[1]

More generally we might allow any positive integers k and ℓ with $p \nmid \ell$. But the above is sufficient for our final goal of counting decomposable polynomials.

We obtain a partial answer to the question (2.8) by splitting the map ρ into two components:

$$\begin{array}{ccccc} P_s \times F & \stackrel{\tau}{\to} & P_s \times P_m & \stackrel{\pi_2}{\longrightarrow} & P_m. \\ \varphi \downarrow & & & & \\ F & & & & \end{array} \tag{5.4}$$

Here P_d consists of the monic polynomials in $F[x]$ of degree d for any $d \ge 0$, $\tau(w,a) = (w, \rho_{w,a})$, and φ and π_2 are the second projections. Thus $\pi_2 \circ \tau = \rho$. If Conjecture 8.3 below holds, then π_2 is injective, and to understand the fibers of ρ, it is sufficient to know those of τ.

For any $w \in F[x]$, $\operatorname{stab}^w_{k,\ell} = \operatorname{stab}^{x^k w(x^\ell)}$, as in (3.4), is a subgroup of F.

Lemma 5.5. *Let $w \in F[x]$ be monic of degree s, and $G = \operatorname{stab}^w_{k,\ell}$. Then the sets $\varphi(\tau^{-1}((w, \rho_{w,a})))$ for $a \in F$ are precisely the cosets of G. When $F = \mathbb{F}_q$, then $\#\{\rho_{w,a} \colon a \in \mathbb{F}_q\} = q/\#G$.*

Proof. Let $a, b \in F$. Then

$$\begin{aligned} \rho_{w,a} = \rho_{w,b} &\iff x^k w(x^\ell) \circ (x+a) = x^k w(x^\ell) \circ (x+b) \\ &\iff x^k w(x^\ell) = x^k w(x^\ell) \circ (x+b-a) \\ &\iff b - a \in G. \end{aligned}$$

Thus each nonempty fiber of the map $\mathbb{F}_q \to P_m$ (over $\mathbb{F}_q$) with $a \mapsto \rho_{w,a}$ is a coset of G and has $\#G$ elements. □

This implies that over $\mathbb{F}_q$, we have

$$\#\operatorname{im}\tau = \sum_{w \in P_s} \frac{q}{\#\operatorname{stab}^w_{k,\ell}}. \tag{5.6}$$

Section 6. The regularity condition

In the normal form of Ritt's Second Theorem, we have the regularity condition (2.6):

$$kw + \ell x w' \ne 0.$$

We evaluate this for $w \in F[x]^G_{k,\ell}$, using the explicit description of Theorem 5.1. So let G and u be as in Theorem 4.1(ii), $\lambda^G = \sum_{0 \le i} \lambda_i x^{p^i}$ with all $\lambda_i \in F$ and $\lambda_0 \ne 0$

[1] The published journal version has the typographical error ∅ instead of {0}.

by (3.2), $v \in F[x]$ monic and $w = u^k v(xu^\ell)$. We have

$$\begin{aligned}
-x^{-1}u(x^\ell) + x^{-1}\lambda_0 &= -x^{-2}\lambda^G + x^{-1}(\lambda^G)' = (x^{-1}\lambda^G)' \\
&= (u(x^\ell))' = u'(x^\ell)\cdot \ell x^{\ell-1}, \\
(-u+\lambda_0)\circ x^\ell &= -u\circ x^\ell + \lambda_0 = \ell x^\ell u'(x^\ell) = \ell x u'\circ x^\ell, \\
-u+\lambda_0 &= \ell x u', \\
kw + \ell x w' &= kw + \ell x\big(ku^{k-1}u'v(ux^\ell) + u^k v'(xu^\ell)(u^\ell + \ell x u^{\ell-1}u')\big) \\
&= ku^k v(xu^\ell) + (-ku^k + k\lambda_0 u^{k-1})v(xu^\ell) \\
&\quad + \ell x u^k v'(xu^\ell)(u^\ell - u^\ell + \lambda_0 u^{\ell-1}) \\
&= k\lambda_0 u^{k-1} v(xu^\ell) + \ell\lambda_0 x u^{k+\ell-1} v'(xu^\ell) \\
&= \lambda_0 u^{k-1}\cdot((kv + \ell x v')\circ xu^\ell).
\end{aligned}$$

Since $\lambda_0 u^{k-1} \neq 0$, we have

$$kw + \ell x w' \neq 0 \iff (kv + \ell x v')\circ xu^\ell \neq 0 \iff kv + \ell x v' \neq 0. \tag{6.1}$$

We write $t = \deg v$ and $v = \sum_{0\le i\le t} v_i x^i$ with all $v_i \in F$ and $v_t = 1$. Then

$$\begin{aligned}
kv + \ell x v' = 0 &\iff \forall i \le t \quad (k+i\ell)v_i = 0 \\
&\iff \forall i \le t \quad p \mid k + i\ell \text{ or } v_i = 0.
\end{aligned}$$

Since $\ell \mid r-1$, ℓ is invertible modulo p, and the latter condition fixes all v_i to be 0 except when $p \mid k + i\ell$. In particular, we have

$$p \nmid k + t\ell \Longrightarrow (2.6) \text{ holds.} \tag{6.2}$$

If $p \mid k + t\ell$, then there are exactly $q^{\lfloor t/p\rfloor}$ values of v of degree t that violate the condition in (6.1).

Section 7. Generating and counting invariant polynomials

This section describes a way of generating the elements of some fixed degree in $F[x]^G_{k,\ell}$. We start with an arbitrary field F of characteristic p, and then find a more precise description when F is finite.

For fields $E \subseteq F$ and $d \ge 0$, we denote as $\mathbb{G}_E(d, F)$ the set of d-dimensional vector spaces $G \subseteq F$ over E. For $\mathbb{F}_r \subseteq \mathbb{F}_q$, with $q = r^b$, the size of this Grassmannian is

$$\#\mathbb{G}_{\mathbb{F}_r}(d, \mathbb{F}_q) = \frac{\prod_{0\le i<d}(r^b - r^i)}{\prod_{0\le i<d}(r^d - r^i)}.$$

Given q, ℓ, and m, the following algorithm constructs ambiguities as in (2.8).

Algorithm 7.1. Generating shift-invariant polynomials of the form $x^k w(x^\ell)$.
Input: A finite field $\mathbb{F}_q$ of characteristic p, and integers $m > \ell \ge 2$ with $\gcd(\ell, m) = 1$ and $p \mid m$.
Output: A set S of (w, G), where $w \in \mathbb{F}_q[x]$ is monic of degree s and $G = \mathrm{stab}^w_{k,\ell}$. Here $m = s\ell + k$ is the division with remainder, with $1 \le k < \ell$.

1. Set $S \leftarrow \varnothing$, write $q = p^e$, $c \leftarrow \mathrm{ord}_\ell\, p$, $r \leftarrow p^c$, and let μ be the multiplicity of p in m. If $c \nmid e$ or $\mu < c$, then return $\varnothing$.
2. For $d = \lfloor \mu/c \rfloor$ down to 1 do Steps 3 through 5.
3. $t \longleftarrow (mr^{-d} - k)/\ell$.

4. For all $G \in \mathbb{G}_{\mathbb{F}_r}(d, \mathbb{F}_q)$ and for all monic $v \in \mathbb{F}_q[x]$ of degree t do Steps 4 and 5.
5.

$$\lambda^G = \sum_{0 \le i \le d} \lambda_i x^{r^i} \longleftarrow \prod_{a \in G} (x - a),$$
$$u \longleftarrow \sum_{0 \le i \le d} \lambda_i x^{(r^i - 1)/\ell},$$
$$w \longleftarrow u^k \cdot (v \circ x u^\ell).$$

6. If $(w, H) \in S$ for some $H \subseteq F$ then $S \longleftarrow S$ else $S \longleftarrow S \cup \{(w, G)\}$.
7. Return S.

Theorem 7.2. *Let S be the output of Algorithm 7.1. For any $(w, G) \in S$, $w \in F[x]$ is monic of degree s and $G = \operatorname{stab}^w_{k,\ell}$. If Conjecture 5.3 holds, then S consists of all such (w, G). If $(w, G) \in S$ and v is chosen in Step 3, then (2.6) is equivalent to*

$$kv + \ell x v' \neq 0. \tag{7.3}$$

When $d = \mu/c$, this is satisfied for any v.

Proof. We have $r \equiv 1 \bmod \ell$ and $m = k \bmod \ell$, so that $\ell \mid mr^{-d} - k$ and in Step 3, t is an integer. Furthermore, $mr^{-d} - k \ge 1 - k > -\ell$, so that $t \ge 0$. The assignment to u in Step 5 is well-defined by Lemma 4.6, and $x^{-1}\lambda^G = u \circ x^\ell \in \mathbb{F}_q[x^\ell]$.

We first verify that any $(w, G) \in S$ as computed in the algorithm satisfies the output conditions. We have $\#G = r^d = \deg \lambda^G \ge r$ and $\deg u = (r^d - 1)/\ell$. Since u and v are monic, so is w, and

$$\begin{aligned} \deg w &= \frac{k(r^d - 1)}{\ell} + t(1 + (r^d - 1)) \\ &= \frac{k(r^d - 1)}{\ell} + \frac{(mr^{-d} - k)r^d}{\ell} = \frac{m - k}{\ell} = s. \end{aligned}$$

From Theorem 5.1, we have $w \in u^k \cdot \mathbb{F}_q[xu^\ell] = \mathbb{F}_q[x]^G_{k,\ell}$. If $w \in \mathbb{F}_q[x]^H_{k,\ell}$ and $G \subset H$, then also $w \in \mathbb{F}_q[x]^G_{k,\ell}$. The condition in Step 6 guarantees that only the maximal such H, namely $(w, \operatorname{stab}^w_{k,\ell})$, is included in S. Thus the output specifications are satisfied. The theorem's last claim follows from (6.1).

For the reverse inclusion, we take some monic $w \in \mathbb{F}_q[x]^G_{k,\ell}$ of degree s with $G = \operatorname{stab}^w_{k,\ell} \neq \{0\}$. By Conjecture 5.3 and Theorem 5.1, there exist monic $\tilde{u}, \tilde{v} \in \mathbb{F}_q[x]$ so that

$$x^{-1}\lambda^G = \tilde{u}(x^\ell)$$

and $w = \tilde{u}^k \cdot \tilde{v}(x\tilde{u}^\ell)$. Lemma 4.6 implies that $\mathbb{F}_r \subseteq \mathbb{F}_q$ and thus $c \mid e$, so that the return statement in Step 1 is correct. Furthermore, G is a vector space over $\mathbb{F}_r$, and we claim that for $d = \dim_{\mathbb{F}_r} G$ and $t = \deg \tilde{v}$, the value (w, G) is included in S at Step 5. We note that $G \in \mathbb{G}_{\mathbb{F}_r}(d, \mathbb{F}_q)$, so that G is one of the choices in Step 4.

Since $\deg \tilde{u} = ((\deg \lambda^G) - 1)/\ell = (r^d - 1)/\ell$, we have

$$\begin{aligned} s = \deg w &= \frac{k(r^d-1)}{\ell} + tr^d, \\ m &= \ell s + k = (\ell t + k) r^d, \\ t &= \frac{mr^{-d} - k}{\ell} \geq 0, \\ r^d &= p^{cd} \mid m, \\ cd &\leq \mu. \end{aligned}$$

The conditions in the algorithm are satisfied, and for the choice $v = \tilde{v}$ in Step 4, (w, G) is included in S. □

In the algorithm, we have $k + t\ell = mr^{-d}$. According to (6.2), for $d = \mu/c$ all v in Step 4 satisfy (7.3). For smaller values of d and any G, the number of v violating (7.3) equals $mr^{-d}/p = mp^{-cd-1} \geq 1$.

Due to the inclusion-exclusion of Step 6, the size of S is somewhat complicated to determine. We obtain an upper bound by ignoring the exclusion, and a lower bound by just taking $d = 1$. We have $S = \varnothing$ if $c \nmid e$ or $\mu < c$, and otherwise

$$\frac{r^{e/c} - 1}{r - 1} r^{e(m/r-k)/\ell c} \leq \#S \leq \sum_{1 \leq d \leq \mu/c} \frac{\prod_{0 \leq i < d} (r^{e/c} - r^i)}{\prod_{0 \leq i < d} (r^d - r^i)} r^{e(mr^{-d}-k)/\ell c}.$$

Our next goal is to compute $\#\operatorname{im} \tau$. We let

$$X^0 = \{(w, G) \colon w \in \mathbb{F}_q[x] \text{ monic of degree } s,\ G = \operatorname{stab}^w_{k,\ell}\},$$

so that $\#X^0 = q^s$. According to Lemma 5.5, each fiber of $\tau \restriction \{w\} \times F$ has size $q \cdot (\#\operatorname{stab}^w_{k,\ell})^{-1}$, and hence

$$\#\operatorname{im} \tau = \sum_{(w,G) \in X^0} \frac{q}{\#G},$$

as in (5.6). Setting

$$X = \{(w, G) \in X^0 \colon G \neq \{0\}\}$$

isolates the "interesting" cases, and

$$\begin{aligned} \#\operatorname{im} \tau &= q \cdot \#X^0 - q \cdot \#X + \sum_{(w,G) \in X} \frac{q}{\#G} \\ &= q^{s+1} - q \sum_{(w,G) \in X} \frac{\#G - 1}{\#G}. \end{aligned}$$

Example 7.4. As a concrete example, we take $F = \mathbb{F}_9$, $m = 9 > 2 = \ell$, so that $s = 4$, $k = 1$, $c = 1 \mid 2 = e$, $r = 3$, and $\mu = 2$. We start with $d = 2 = \mu/c$. We have $t = 0$ and the only choices $\mathbb{F}_9 \in \mathbb{G}_{\mathbb{F}_3}(2, \mathbb{F}_9)$ and $v = 1$ in Step 4. This yields $\lambda^{\mathbb{F}_9} = x^9 - x$, $u = x^4 - 1 = w$, and $(x^4 - 1, \mathbb{F}_9)$ is added to S.

Next comes $d = 1$. We have $t = 1$ and $v = x + v_0$ with $v_0 \in \mathbb{F}_9 = \mathbb{F}_3[\alpha]$ with $\alpha^2 = -1$. Each of the four $G \in \mathbb{G}_{\mathbb{F}_3}(1, \mathbb{F}_9)$ can be written as $G = a \cdot \mathbb{F}_3$, with some nonzero $a \in \mathbb{F}_9$. In Step 5, we have $w = u \cdot (xu^2 + v_0) = xu^3 + v_0 u$. Table 7.1 gives the relevant quantities for the four G.

a	λ^G	u	xu^ℓ	u^3	w
1	x^3-x	$x-1$	x^3+x^2+x	x^3-1	$x^4-x+v_0(x-1)$
α	x^3+x	$x+1$	x^3-x^2+x	x^3+1	$x^4+x+v_0(x+1)$
$\alpha+1$	$x^3+\alpha x$	$x+\alpha$	$x^3+\alpha x^2-x$	$x^3-\alpha$	$x^4-\alpha x+v_0(x+\alpha)$
$\alpha-1$	$x^3-\alpha x$	$x-\alpha$	$x^3+\alpha x-x$	$x^3+\alpha$	$x^4+\alpha x+v_0(x-\alpha)$

TABLE 7.1. The four $\mathbb{F}_3$-lines in $\mathbb{F}_9$ and their polynomials.

We have $4 \cdot 9 = 36$ pairs (a, v_0), but since $G \subseteq \mathbb{F}_9$, four of them yield the same $w = x^4 - 1$, namely $(1,1)$, $(\alpha,-1)$, $(\alpha+1,\alpha)$, and $(\alpha-1,-\alpha)$, which are already taken care of by $(x^4-1, \mathbb{F}_9)$. We thus have one $(w,G) \in S$ with $\#G = 9$, and 32 with $\#G = 3$.

Thus X consists of $(x^4-1, \mathbb{F}_q)$ and the $4 \cdot 8 = 32$ pairs (w,G) with $\#G = 3$ in Table 7.1, with x^4-1 excluded in each of the four lines. It follows that

$$\#\operatorname{im}\tau = 9^5 - 9(1 \cdot \frac{8}{9} + 32 \cdot \frac{2}{3}) = 9^5 - 200 = 58\,849 > 0.9966 \cdot 9^5.$$

In the first line of Table 7.1, we have $G = \langle 1 \rangle = \mathbb{F}_3$ and $w = x^4 + x + 1$ for $v_0 = -1$. The entry means that $\rho_{w,0} = \rho_{w,1} = \rho_{w,-1}$. We can compose on the right with any $a \in \mathbb{F}_9$, and thus have $\rho_{w,\alpha} = \rho_{w,\alpha+1} = \rho_{w,\alpha-1}$ and $\rho_{w,-\alpha} = \rho_{w,-\alpha+1} = \rho_{w,-\alpha-1}$. Thus the nine values (w,a) with $a \in \mathbb{F}_9$ yield three polynomials $xw(x^2) \circ (x+a) = \rho_{w,a}$, corresponding to the reduction factor $q/\#G = 3$ in Lemma 5.5. The projection π_2 in (5.4) is injective, and we conclude that, the total number of distinct $\rho_{w,a} = x^k w(x^\ell) \circ (x+a)$, as in (2.7), is 58 849.

The condition $kv + \ell x v' \neq 0$ in (6.1) is satisfied for $v = 1$, and for $v = x + v_0$ it reads $0 \neq x + v_0 + 2x = v_0$, which for each a as above holds for 7 out of 8 admissible values of v_0. Thus of the 200 values of $\rho_{w,a}$ counted above, $9(8/9 + 32 \cdot 7/8 \cdot 2/3) = 176$ satisfy (6.1). In terms of the normal form (2.5), the first line in Table 7.1 with $v_0 = 1$ provides the example

$$\begin{aligned} f &= x^{18} + x^{10} + x^2 \\ &= (x^9 + x^5 + a^2(-x^4 + x^3 + x^2 - x) + x) \circ (x^2 - ax) \\ &= x^2 \circ (x^9 - x) \end{aligned}$$

for all $a \in \mathbb{F}_3$. No example of such a "collision of collisions" seems to be in the literature. ◊

Example 7.5. We now take $q = 9$, $m = 15$, $\ell = 2$ so that $k = 1$, $s = 7$, $c = 1$, $r = 3$, $\mu = 1$. We have $c \mid e = 2$ and $\mu/c = 1$, so that only $d = 1$ is considered in the algorithm.

In Step 3, we have $t = 2$ and consider the four $G \in \mathbb{G}_{\mathbb{F}_3}(1, \mathbb{F}_9)$, $v = x^2 + v_1 x + v_0$, with $v_0, v_1 \in \mathbb{F}_9$, and $w = u \cdot (v \circ xu^2)$. Table 7.2 is arranged as Table 7.1 and shows the current value of w.

This gives $\#X = 4 \cdot 9^2 = 324$ different values of w, while $\#X^0 = 9^4$ and $9^5 = 59\,049$. For each of the four nonzero G in Table 7.2, we have $\#G = 3$ and $\#\mathbb{F}_9[x]^G_{1,2} = 81$. Thus $\#X = 4 \cdot 81 = 324$, and

$$\#\operatorname{im}\tau = 9^5 - 9 \cdot 4 \cdot 81 \cdot \frac{3-1}{3} = 9^5 - 1\,944 = 57\,105 = \frac{235}{243} \cdot 9^5 > 0.9670 \cdot 9^5.$$

a	u	w
1	$x-1$	$x^7+x^6+x^5-x^4-x^3-x^2+v_1(x^4-x)+v_0(x-1)$
α	$x+1$	$x^7-x^6+x^5+x^4-x^3+x^2+v_1(x^4+x)+v_0(x+1)$
$\alpha+1$	$x+\alpha$	$x^7-\alpha x^6-x^5-\alpha x^4-x^3+\alpha x^2+v_1(x^4-\alpha x)+v_0(x+\alpha)$
$\alpha-1$	$x-\alpha$	$x^7+\alpha x^6-x^5+\alpha x^4-x^3-\alpha x^2+v_1(x^4+\alpha x)+v_0(x-\alpha)$

TABLE 7.2. The current value of w.

The regularity condition (7.3) becomes

$$0 \neq kv + \ell x v' = 5x^2 + v_0.$$

It is alway satisfied, as in Theorem 7.2. ◊

Section 8. Nonuniqueness of (w, a) if $p \mid m$

We now turn to the question (2.8) that motivated this work. We provide two answers. In this section, we present a conjecture under which the approach presented above would solve the problem. In the next section, we prove an unconditional but weak estimate.

We use the following notation. For positive integers k and ℓ, $w \in F[x]$ monic, and $a \in F$, we let

$$\psi_{w,a} = (x - a^{k\ell} w^\ell(a^\ell)) \circ x^{k\ell} w^\ell(x^\ell) \circ (x + a), \tag{8.1}$$

as in (2.5). Furthermore, for monic $w, \tilde{w} \in F[x]$ we set

$$\mathrm{eq}_{k,\ell}^{w,\tilde{w}} = \{a \in F \colon \psi_{\tilde{w},0} = \psi_{w,a}\}.$$

(The word *equalizer* is formed in analogy with *stabilizer*). We always have $0 \in \mathrm{eq}_{k,\ell}^{w,w}$. The lower index will usually have the value k, ℓ and we drop it at times without further notice.

A *t-way ambiguity* is a set of t pairs (w, a) so that $\psi_{w,a}$ is the same for all pairs. An equalizer of size t yields a t-way ambiguity in (2.5). The connection to Section 5 is that when $a^k w(a^\ell) = 0$, we have

$$\begin{aligned} \psi_{w,a} &= x^\ell \circ \rho_{w,a}, \\ \psi_{w,0} = \psi_{w,a} &\iff \rho_{w,0} = \rho_{w,a} \\ &\iff a \in \mathrm{stab}_{k,\ell}^w \iff w \in F[x]_{k,\ell}^a. \end{aligned}$$

Thus $\mathrm{eq}_{k,\ell}^{w,w} = \mathrm{stab}_{k,\ell}^w$ in this case. We first note that the choice of 0 as an argument in the definition actually covers the general case.

Lemma 8.2. *Let k and ℓ be positive integers, $w, \tilde{w} \in F[x]$ monic, and $a, \tilde{a} \in F$. Then the following hold.*

(i) $\psi_{w,a} = \psi_{\tilde{w},\tilde{a}} \iff a - \tilde{a} \in \mathrm{eq}^{w,\tilde{w}}$.
(ii) *If $0 \in \mathrm{eq}_{k,\ell}^{w,\tilde{w}}$, then $w = \tilde{w}$.*

PROOF. (i) Let

$$\begin{aligned} \tilde{u} &= x + \tilde{a}^{k\ell} \tilde{w}^\ell(\tilde{a}^\ell), \\ u &= \tilde{u} - a^{k\ell} w^\ell(a^\ell). \end{aligned}$$

Then

$$\begin{aligned}\tilde{u} \circ \psi_{\tilde{w},\tilde{a}} \circ (x - \tilde{a}) &= \tilde{u} \circ (\tilde{u}^{-1} \circ x^{k\ell} \tilde{w}^{\ell}(x^{\ell}) \circ (x + \tilde{a})) \circ (x - \tilde{a}) \\ &= x^{k\ell} \tilde{w}^{\ell}(x^{\ell}) = \psi_{\tilde{w},0}, \\ \tilde{u} \circ \psi_{w,a} \circ (x - \tilde{a}) &= \tilde{u} \circ (x - a^{k\ell} w^{\ell}(a^{\ell})) \circ x^{k\ell} w^{\ell}(x^{\ell}) \circ (x + a) \circ (x - \tilde{a}) \\ &= u \circ x^{k\ell} w^{\ell}(x^{\ell}) \circ (x + a - \tilde{a}).\end{aligned}$$

This polynomial is monic and original, so that it equals $\psi_{w,a-\tilde{a}}$. Since the linear components are invertible, we have

$$\psi_{w,a} = \psi_{\tilde{w},\tilde{a}} \iff \psi_{\tilde{w},0} = \psi_{w,a-\tilde{a}} \iff a - \tilde{a} \in \mathrm{eq}^{w,\tilde{w}}.$$

(ii) We have

$$0 \in \mathrm{eq}_{k,\ell}^{w,\tilde{w}} \Longrightarrow x^{k\ell} \tilde{w}^{\ell}(x^{\ell}) = x^{k\ell} w^{\ell}(x^{\ell}) \Longrightarrow w = \tilde{w}. \qquad \square$$

Conjecture 8.3. *Let F be a field of characteristic $p \geq 2$, let $\ell < p$ be a prime, $1 \leq k < \ell$, let $w, \tilde{w} \in F[x]$ be monic, and assume that $a \in \mathrm{eq}_{k,\ell}^{w,\tilde{w}}$. Then $w = \tilde{w}$ and $a^k w(a^{\ell}) = 0$.*

The validity of this conjecture, and also of Conjecture 5.3, has been verified experimentally for $q^{1+\deg w} \leq 10^4$. Its truth would imply that π_2 in (5.4) is injective, $\#\operatorname{im}\pi = \#\operatorname{im}\rho$, and the counting results of Section 7 would apply to ρ.

As noted after Conjecture 5.3, we might allow, more generally, any $k \geq 1$ and $\ell \geq 2$ with $p \nmid \ell$.

Lemma 8.4. *Assume that $p > \ell > k \geq 1$, $p \mid m = \ell s + k$, and $s = 1$. Then Conjecture 8.3 holds. Furthermore, if $\mathrm{eq}_{k,\ell}^{w,w} \neq \varnothing$, then $k = 1$, $p = \ell + 1$, and $w = x - b^{p-1}$ for some $b \in F$.*

Proof. We write $w = x + w_0$ and $\tilde{w} = x + \tilde{w}_0$ with $w_0, \tilde{w}_0 \in F$, and take some $a \in \mathrm{eq}^{\tilde{w},w}$ (having interchanged w and $\tilde{w}$). We note that $p \leq m = \ell s + k = \ell + k < 2\ell < 2p$, so that $m = p = k + \ell$. Then

$$\begin{aligned}\psi_{\tilde{w},0} &= x^{k\ell} \tilde{w}^{\ell}(x^{\ell}) = (x^k (x^{\ell} + \tilde{w}_0))^{\ell}, \\ \psi_{w,a} &= (x - a^{k\ell}(a^{\ell} + w_0)^{\ell}) \circ x^{k\ell}(x^{\ell} + w_0)^{\ell} \circ (x + a) \\ &= (x + a)^{k\ell}((x + a)^{\ell} + w_0)^{\ell} - (a^k (a^{\ell} + w_0))^{\ell} \\ &= ((x + a)^{k+\ell} + w_0 (x + a)^k)^{\ell} - (a^{k+\ell} + w_0 a^k)^{\ell} \\ &= (x^p + a^p + w_0 (x + a)^k)^{\ell} - (a^p + w_0 a^k)^{\ell}.\end{aligned}$$

We first assume that $a^k w(a^{\ell}) \neq 0$. Let E be an extension field of F containing a primitive ℓth root of unity. Then

$$\psi_{w,a} = \prod_{\zeta^{\ell}=1} (x^p + a^p + w_0 (x + a)^k - \zeta(a^p + w_0 a^k)).$$

We consider some $\zeta \in E$ with $\zeta^{\ell} = 1$ and evaluate the factor given above at $x = 0$:

$$a^p + w_0 a^k - \zeta(a^p + w_0 a^k) = (1 - \zeta) a^k (a^{\ell} + w_0).$$

This vanishes only if $\zeta = 1$, and therefore only this factor is divisible by x. From $\psi_{\tilde{w},0} = \psi_{w,a}$, we find that

$$x^{k\ell} \mid x^p + a^p + w_0(x+a)^k - (a^p + w_0 a^k) = x^p + w_0((x+a)^k - a^k).$$

The coefficient of x on the right hand side equals $w_0 k a^{k-1}$. Since $k\ell \geq \ell \geq 2$, this coefficient vanishes, and hence $w_0 = 0$. It follows that

$$(x^p + \tilde{w}_0 x^k)^\ell = (x^p + a^p)^\ell - a^{p\ell}.$$

The coefficient of x^p on the right hand side is $\ell a^{p(\ell-1)} \neq 0$, while on the left hand side it vanishes. Thus the assumption $a^k w(a^\ell) \neq 0$ leads to a contradiction.

We have shown that $a^k w(a^\ell) = 0$. If $a = 0$, then $x^k w(x^\ell) = x^k \tilde{w}(x^\ell)$, which implies $w = \tilde{w}$. Thus we may assume that $a \neq 0$. Then $a^\ell + w_0 = 0$ and $w_0 \neq 0$, and

$$x^p + \tilde{w}_0 x^k = x^p + a^p + w_0(x+a)^k.$$

The coefficient of x on the right hand side is $w_0 k a^{k-1} \neq 0$, so that $k = 1$, $\ell = p-1$, $\tilde{w}_0 = w_0$, and $\tilde{w} = w = x - a^{p-1}$. □

Section 9. An unconditional estimate

The assumption of Conjectures 5.3 and 8.3 leads to a satisfactory answer to the question (2.8). We now present a result without assumptions. The resulting bound in Theorem 9.8 is weaker than what we expect to be true.

We start with a result which shows that in a special situation a factor of degree k automatically implies one of degree $k\ell$.

Lemma 9.1. *Let F be a field of characteristic p, $w \in F[x]$, $a \in F$ nonzero, and $k, \ell \geq 1$ with $p \nmid \ell$. The following are equivalent.*

(i) $(x - a^\ell)^k \mid w$,

(ii) $(x-a)^k \mid w(x^\ell)$,

(iii) $(x^\ell - a^\ell)^k \mid w(x^\ell)$.

Proof. (i) $\Longrightarrow$ (iii) follows by substituting x^ℓ for x, and (iii) $\Longrightarrow$ (ii) from the fact that $x - a \mid x^\ell - a^\ell$. It remains to show (ii) $\Longrightarrow$ (i). This is clear for $k = 1$. For an induction on k, we let

$$w = u_0 + u_1(x - a^\ell) + u_2(x-a^\ell)^2 + \cdots + u_{k-1}(x-a^\ell)^{k-1} + u_k(x-a^\ell)^k$$

be the Taylor expansion of w around a^ℓ, with $u_0, \ldots, u_{k-1} \in F$ and $u_k \in F[x]$. Then $u_0 = w(a^\ell) = 0$ by the conclusion for $k = 1$, and $w_1 = w/(x - a^\ell) \in F[x]$. We observe that

$$\frac{x^\ell - 1}{x - 1}(1) = (x^{\ell-1} + \cdots + 1)(1) = \ell \neq 0,$$

since $p \nmid \ell$. It follows that $\gcd(x-1, \frac{x^\ell - 1}{x-1}) = 1$. Substituting x by x/a, we find

$$\gcd(x-a, \frac{x^\ell - a^\ell}{x-a}) = 1,$$
$$\gcd((x-a)^k, x^\ell - a^\ell) = x - a,$$
$$(x-a)^{k-1} \mid \frac{w(x^\ell)}{x^\ell - a^\ell} = w_1(x^\ell).$$

Applying the induction hypothesis to w_1, we find that

$$u_1 = \cdots = u_{k-1} = 0,$$
$$(x - a^\ell)^k \mid w. \qquad \square$$

The case $p \mid \ell$, say $\ell = \ell^* p^d$ with $p \nmid \ell^*$ and an integer $d \geq 1$, is not covered by the lemma. But we can conclude from (ii) that

$$(x-a)^{p^d \cdot \lceil k/p^d \rceil} \mid w(x^{\ell^*})^{p^d},$$
$$(x - a^{\ell^*})^{\lceil k/p^d \rceil} \mid w,$$

where we take $w \in \mathbb{F}_p[x]$ for simplicity.

Lemma 9.2. *Let $a \in \mathrm{eq}^{w,\tilde{w}}$ be nonzero and $\lambda = \gcd(\ell, p-1)$. Then the following hold.*

(i) *If $w(a^\ell) \neq 0$, then*
$$(x - a^\ell)^{k\ell - 1} \mid kw + \ell x w',$$
$$(x - (-a)^\ell)^{k\ell - 1} \mid k\tilde{w} + \ell x \tilde{w}'.$$

(ii) *If $w(a^\ell) = 0$, then*
$$(x - a^\ell)^k \mid w,$$
$$(x - (-a)^\ell)^k \mid \tilde{w}.$$

(iii) *If $w = \tilde{w}$ and $w(a^\ell) \neq 0$, then*
$$(x^{(p-1)/\lambda} - a^{\ell(p-1)/\lambda})^{k\ell - 1} \mid kw + \ell x w'.$$

(iv) *If $w = \tilde{w}$ and $w(a^\ell) = 0$, then*
$$(x^{(p-1)/\lambda} - a^{\ell(p-1)/\lambda})^k \mid w.$$

Proof. We set $b = a^{k\ell} w^\ell(a^\ell)$ and

$$f = (x-b) \circ x^k w^\ell \circ x^\ell \circ (x+a) = \psi_{w,a} = \psi_{\tilde{w},0} = x^k \tilde{w}^\ell \circ x^\ell. \tag{9.3}$$

(i) (9.3) implies that

$$f' = \big((x^{k-1} w^{\ell-1}(kw + \ell x w')) \circ ((x+a)^\ell)\big) \cdot \ell (x+a)^{\ell-1}$$
$$= \big((x^{k-1} \tilde{w}^{\ell-1}(k\tilde{w} + \ell x \tilde{w}')) \circ x^\ell\big) \cdot \ell x^{\ell-1}.$$

Now $x \nmid w((x+a)^\ell)$, so that

$$x^{k\ell-1} = x^{(k-1)\ell} \cdot x^{\ell-1} \mid (kw + \ell x w') \circ (x+a)^\ell. \tag{9.4}$$

Composing on the right with $x - a$, it follows from Lemma 9.1 that

$$(x - a^\ell)^{k\ell - 1} \mid kw + \ell x w'.$$

The second claim in (i) follows similarly.

(ii) We have $u = x$ and

$$f = (x+a)^{k\ell} w^\ell((x+a)^\ell) = x^{k\ell}\tilde{w}^\ell(x^\ell),$$
$$(x+a)^k w((x+a)^\ell) = x^k \tilde{w}(x^\ell);$$

the latter follows since both sides are monic polynomials whose ℓth powers are equal. Thus $x^k \mid w((x+a)^\ell)$ and $(x-a)^k \mid w(x^\ell)$. Lemma 9.1 implies that

$$(x-a^\ell)^k \mid w.$$

Similarly, we find that

$$(x-(-a)^\ell)^k \mid \tilde{w}.$$

(iii) and (iv). We claim that for $i \geq 0$ we have

$$f = (x-ib) \circ x^k w^\ell \circ x^\ell \circ (x+ia). \tag{9.5}$$

(We identify an integer i with $i \bmod p$ in $\mathbb{F}_p \subseteq F$.) When i is 0 or 1, this follows from (9.3). For $i \geq 1$, we have inductively

$$\begin{aligned}&(x-(i+1)b) \circ x^k w^\ell \circ x^\ell \circ (x+(i+1)a)\\ &= (x-b) \circ (x-ib) \circ x^k w^\ell \circ x^\ell \circ (x+ia) \circ (x+a)\\ &= (x-b) \circ x^k w^\ell \circ x^\ell \circ (x+a) = f.\end{aligned}$$

We let $S = (\mathbb{F}_p^\times)^\lambda$ be the set of λth powers in $\mathbb{F}_p^\times \subseteq F^\times$. Then $\#S = (p-1)/\lambda$, and

$$\operatorname*{lcm}_{1\leq i<p}(x-i^\ell) = \prod_{j\in S}(x-j) = x^{(p-1)/\lambda} - 1,$$
$$\operatorname*{lcm}_{1\leq i<p}(x-(ia)^\ell) = \prod_{j\in S}(x-ja^\ell) = x^{(p-1)/\lambda} - a^{\ell(p-1)/\lambda}.$$

From (9.5) we find for $1 \leq i < p$ that

$$\begin{aligned}0 = f(0) &= (x - ia^{k\ell} w^\ell(a^\ell)) \circ (ia)^{k\ell} w^\ell((ia)^\ell)\\ &= (ia)^{k\ell} w^\ell((ia)^\ell) - ia^{k\ell} w^\ell(a^\ell).\end{aligned}$$

Thus if $w(a^\ell) \neq 0$, then also $w((ia)^\ell) \neq 0$. As in (9.4), it follows that

$$x^{k\ell-1} \mid (kw + \ell x w') \circ (x+ia)^\ell,$$
$$(x-(ia)^\ell)^{k\ell-1} \mid kw + \ell x w'$$

for all i with $1 \leq i < p$, so that

$$(x^{(p-1)/\lambda} - a^{\ell(p-1)/\lambda})^{k\ell-1} \mid kw + \ell x w'.$$

If $w(a^\ell) = 0$, then $b = 0$ and $x^{k\ell}$ divides $w^\ell((x+ia)^\ell)$ for $1 \leq i < p$, so that

$$(x-ia)^k \mid w(x^\ell) \text{ for } 1 \leq i < p,$$
$$(x^{(p-1)/\lambda} - a^{\ell(p-1)/\lambda})^k \mid w. \qquad \square$$

We can also deal with (iv) in the language of Section 5. We set $G = a \cdot \mathbb{F}_p \subseteq \mathbb{F}_q$. Then $\lambda^G = x^p - a^{p-1}x$ and $w \in \mathbb{F}_q[x]^G_{k,\ell}$, and Conjecture 5.3, if true, implies that $\ell \mid p-1$, so that $\lambda = \ell$, and in Theorem 5.1 we have $u = x^{(p-1)/\ell} - a^{p-1}$ and $w = u^k \cdot v(xu^\ell)$ for some $v \in \mathbb{F}_q[x]$. In particular $u^k \mid w$, which is the claim of (iv).

For a bound on the number of ambiguities, we have to quantify the effect of the divisibilities in (i) and (iii). To this end, we write $P_d^* = \{f \in F[x]\colon \deg f \le d\}$, so that P_d, as defined after (5.4), consists of the monic $f \in P_d^*$ of degree d. Now let $d \ge 1, g \in P_d$, and consider the vector space

$$W_g = \{w \in P_s^*\colon g \mid kw + \ell x w'\}.$$

Lemma 9.6. *Let F be a field of characteristic $p \ge 2$, let $m > \ell \ge 2$, $s = \lfloor m/\ell \rfloor$ and $d \ge 1$, with $p \mid m$ and $p \nmid \ell$ be as above, and $g \in P_d$. If $d = 1$, we assume that $p \nmid s$. The following hold.*

(i) $\dim W_g \le s$.
(ii) *$W_g \cap P_s \subseteq P_s$ is an affine linear subset of dimension* $(\dim W_g) - 1$.
(iii) $\#(W_g \cap P_s) \le q^{s-1}$.

Proof. (i) We write $g = \sum_{0 \le i \le d} g_i x^i$ with all $g_i \in F$ and $g_d = 1$. Furthermore, we take $g_i = 0$ if $i > d$ or $i < 0$. We consider the following $(\lfloor s/p \rfloor + 1) \times (s - d + 1)$ matrix R_g. Its rows are indexed by b, with $0 \le b \le s/p$, and its columns by j, with $0 \le j \le s - d$, and the entries are

$$(R_g)_{b,j} = g_{d-bp+j}.$$

Thus the product

$$\sum_{0 \le j \le s-d} (R_g)_{b,j} v_{s-d-j} = \sum_{0 \le j \le s-d} g_{d-bp+j} v_{s-d-j}$$

of the bth row with the transposed coefficient vector $(v_{s-d}, \dots, v_0)^{\mathsf{T}}$ of $v = \sum_{0 \le j \le s-d} v_j x^j \in P_{s-d}^*$ equals the coefficient of x^{s-bp} in $g \cdot v$. In other words, R_g is the matrix of multiplication by g in the standard basis, where only every pth row is taken.

As an example, we take $p = 2, m = 20, \ell = 3, d = 2$, so that $s = 6$, $s/p = 3$, $s - d = 4$, and

$$R_g = \begin{pmatrix} g_2 & 0 & 0 & 0 & 0 \\ g_0 & g_1 & g_2 & 0 & 0 \\ 0 & 0 & g_0 & g_1 & g_2 \\ 0 & 0 & 0 & 0 & g_0 \end{pmatrix}.$$

We note that $g_2 = 1$. The rank of R_g is at least 3, and it equals 4 if and only if $g_0 g_1 \ne 0$. Furthermore, $R_g \cdot (v_4, v_3, v_2, v_1, v_0)^{\mathsf{T}}$ consists of the coefficients of $g \cdot v$ at x^6, x^4, x^2, x^0.

We consider the two F-linear maps

$$\begin{array}{ccccc} P_s^* & & & & P_{s-d}^* \\ & \searrow^{\delta} & & \swarrow_{\mu} & \\ & & P_s^* & & \end{array}$$

with $\delta(w) = kw + \ell x w'$ and $\mu(v) = gv$. Then $W_g = \delta^{-1}(\operatorname{im} \mu)$. Let $w = \sum_{0 \le i \le s} w_i x^i \in P_s^*$, with all $w_i \in F$. Then

$$kw + \ell x w' = \sum_{0 \le i \le s} (k + i\ell) w_i x^i. \tag{9.7}$$

We have $p \mid m = k + \ell s$, so that $\deg \delta(w) < s$. Since $p \nmid \ell$, we have for $0 \le i \le s$ that

$$k + i\ell \equiv 0 \bmod p \Longleftrightarrow i \equiv s \bmod p.$$

Thus the coefficient of x^i in $\delta(w)$ is zero if $p \mid s - i$. There are $\lfloor s/p \rfloor + 1$ such $i \le s$. This imposes $\lfloor s/p \rfloor + 1$ linear conditions on $\operatorname{im}\delta$ which are linearly independent, so that $\dim \operatorname{im}\delta \le s - \lfloor s/p \rfloor$ and $\dim \ker \delta = \dim P_s^* - \dim \operatorname{im}\delta \ge s + 1 - (s - \lfloor s/p \rfloor) = \lfloor s/p \rfloor + 1$. On the other hand, if $w^* \in P_s^*$ satisfies these linear conditions, then (9.7) can be solved for $w \in P_s^*$ with $w^* = \delta(w)$. It follows that equality holds in the dimension estimates above.

Furthermore, the multiplication map μ is injective. For $v \in P_{s-d}^*$, $R_v \cdot v$ consists of the coefficients of gv at the x^i with $p \mid s - i$. It follows that

$$\begin{aligned} \mu(v) = gv \in \operatorname{im}\delta &\Longleftrightarrow R_g \cdot v = 0, \\ \dim(\operatorname{im}\delta \cap \operatorname{im}\mu) &= s - d + 1 - \operatorname{rank} R_g, \\ \dim W_g = \dim(\delta^{-1}(\operatorname{im}\mu)) &= \dim(\operatorname{im}\delta \cap \operatorname{im}\mu) + \dim\ker\delta \\ &= s - d - \operatorname{rank} R_g + \lfloor s/p \rfloor + 2. \end{aligned}$$

We let $r = \lfloor (s-d)/p \rfloor + 1$ and consider the $r \times r$-submatrix U of R_g consisting of the top r rows with $0 \le b < r$ and columns $j = bp$ for $0 \le b < r$. Now $(R_g)_{b,bp} = g_d$ and $(R_g)_{b,j} = g_{d-bp+j} = 0$ for $j > bp$. Thus U is a lower triangular matrix with $g_d \ne 0$ on the diagonal. U is indeed a submatrix of R_g, since for row b we have $b \le \lfloor (s-d)/p \rfloor \le \lfloor s/p \rfloor$, and for the maximal value of j we have $(r-1)p \le \lfloor (s-d)/p \rfloor \cdot p \le s - d$. Setting $t = \lfloor (s-d)/p \rfloor - \lfloor s/p \rfloor + d - 1$, it follows that

$$\begin{aligned} &\operatorname{rank} R_g \ge \operatorname{rank} U = r, \\ &\dim W_g \le s - d - \left\lfloor \frac{s-d}{p} \right\rfloor + \left\lfloor \frac{s}{p} \right\rfloor + 1 = s - t. \end{aligned}$$

If $d \ge 3$, then

$$t \ge \frac{s-d}{p} - \frac{p-1}{p} - \frac{s}{p} + d - 1 = \frac{1}{p}((d-2)(p-1) - 1) \ge 0.$$

If $d = 2$, then

$$t \ge \left\lfloor \frac{s}{p} \right\rfloor - 1 - \left\lfloor \frac{s}{p} \right\rfloor + 2 - 1 = 0,$$

and if $d = 1$, then $p \nmid s$ by assumption, and

$$t = \left\lfloor \frac{s}{p} \right\rfloor - \left\lfloor \frac{s}{p} \right\rfloor + 1 - 1 = 0.$$

In all cases, we have shown $\dim W_g \le s$.

(ii) $P_s \subseteq P_s^*$ is an affine hyperplane, and $0 \in W_g \setminus P_s$. Therefore $W_g \cap P_s$ is an affine hyperplane in W_g, of dimension $\dim W_g - 1 \le s - 1$. (iii) follows from this. □

We let $P_n^{(0)} \subseteq P_n^*$ be the set of original polynomials of degree n, and now determine a lower bound, admittedly weak, on the number of non-ambiguities.

Theorem 9.8. *Let $\mathbb{F}_q$ be a finite field, $m > \ell \geq 2$ integers with $\gcd(\ell, m) = 1$ and $p \mid m$, $m = s\ell + k$ the division with remainder, with $1 \leq k < \ell$, $n = \ell m$, and*

$$\begin{aligned}\psi\colon P_s \times F &\longrightarrow P_n^{(0)}\\ (w, a) &\longmapsto (x - a^{k\ell} w^\ell(a^\ell)) \circ x^{k\ell} w^\ell(x^\ell) \circ (x + a),\end{aligned}$$

as in (2.5) and (8.1). Then

$$q^{s+1}(1 - 4q^{-1}) \leq \#\operatorname{im}\psi \leq q^{s+1}.$$

Proof. Clearly, $\#\operatorname{im}\psi \leq \#(P_s \times F) = q^{s+1}$. We denote as

$$M = \{f \in P_n^{(0)}\colon \#\psi^{-1}(f) \geq 2\}$$

the set of "ambiguous" polynomials, and consider the action φ of F on $P_n^{(0)}$, given for $a \in F$ by

$$\varphi_a\colon f \longmapsto (x - f(a)) \circ f \circ (x + a).$$

We take some $f = \psi_{\tilde{w},\tilde{a}}$ and $(w, a) \in P_s \times F$. Using Lemma 8.2 we find

$$\psi_{\tilde{w},\tilde{a}} = \psi_{w,a} \iff a - \tilde{a} \in \operatorname{eq}^{w,\tilde{w}}.$$

Similarly for $b \in F$, we have

$$\psi_{\tilde{w},\tilde{a}+b} = \varphi_a(f) = \psi_{w,a} \iff a - (\tilde{a} + b) \in \operatorname{eq}^{w,\tilde{w}}.$$

It follows that

$$\begin{aligned}\#\psi^{-1}(f) = \#\{(w, a)\colon f = \psi_{w,a}\} &= \sum_{w \in P_s} \#\operatorname{eq}^{w,\tilde{w}}\\ &= \#\psi^{-1}(\varphi_a(f)),\end{aligned}$$

so that $\#\psi^{-1}$ is constant on the orbits of φ. Thus M is a union of φ-orbits. Furthermore, each such orbit contains $\psi_{\tilde{w},0}$ for some $\tilde{w} \in P_s$, and also $\psi_{w,a}$ for some $(w, a) \in P_s \times F$ with $a \neq 0$.

We now take some nonzero $a \in F$ and bound $V_a = \{w \in P_s\colon \psi_{w,a} \in M\}$. Following the parts of Lemma 9.2, we distinguish four cases.

$$\begin{aligned}V_a^i &= \{w \in V_a\colon a \in \operatorname{eq}^{w,\tilde{w}} \text{ for some } \tilde{w} \neq w, \text{ and } w(a^\ell) \neq 0\},\\ V_a^{ii} &= \{w \in V_a\colon a \in \operatorname{eq}^{w,\tilde{w}} \text{ for some } \tilde{w} \neq w, \text{ and } w(a^\ell) = 0\},\\ V_a^{iii} &= \{w \in V_a\colon a \in \operatorname{eq}^{w,w} \text{ and } w(a^\ell) \neq 0\},\\ V_a^{iv} &= \{w \in V_a\colon a \in \operatorname{eq}^{w,w} \text{ and } w(a^\ell) = 0\}.\end{aligned}$$

For V_a^i, we set $g = (x - a^\ell)^{k\ell - 1}$, so that $V_a^i \subseteq W_g \cap P_s$ by Lemma 9.2(i), and by Lemma 9.6(iii), $\#V_a^i \leq q^{s-1}$, where $d = k\ell - 1 \geq 1$. If $d = 1$, then $k = 1$ and $p \nmid s$, since otherwise $p \mid m - \ell s = k$. For $w \in V_a^{ii}$, we have $(x - a^\ell)^k \mid w$, and $\#V_a^{ii} \leq q^{s-k} \leq q^{s-1}$. For $w \in V_a^{iii}$, we have, with λ from Lemma 9.2 and $g = (x^{(p-1)/\lambda} - a^{\ell(p-1)/\lambda})^k$, $V_a^{iii} \subseteq W_g \cap P_s$ and hence $\#V_a^{iii} \leq q^{s-1}$, where again $p \nmid s$ if $d = 1$. For $w \in V_a^{iv}$, we have $(x - a^\ell)^k \mid w$ and $\#V_a^{iv} \leq q^{s-k} \leq q^{s-1}$. Overall, we find

$$\begin{aligned}\#V_a &\leq \#V_a^i + \#V_a^{ii} + \#V_a^{iii} + \#V_a^{iv} \leq 4q^{s-1},\\ \#M &\leq \sum_{a \in F} \#\psi(V_a \times \{a\}) \leq 4q^s,\\ \#\operatorname{im}\psi &\geq q^{s+1} - \#M \geq q^{s+1}(1 - 4q^{-1}).\end{aligned}$$

□

From Conjectures 5.3 and 8.3, if true, would follow a bound on $\#\operatorname{im}\psi$ much closer to q^{s+1} than the lower bound proven here, which we have made no attempt to optimize.

Section 10. Acknowledgments

Konstantin Ziegler performed extensive experiments on $F[x]^G_{k,\ell}$ whose analysis led to some of the present results. Thanks go to Gregor Kemper for discussions on invariant theory.

This work was supported by the B-IT Foundation and the Land Nordrhein-Westfalen.

References

Gert Almkvist (1983). Invariants of Z/pZ in characteristic p. In *Invariant Theory*, F. Gherardelli, editor, volume 996 of *Lecture Notes in Mathematics*, 109–117. Springer-Verlag, Berlin, Heidelberg. ISBN 3-540-12319-9.

A. F. Beardon & T. W. Ng (2000). On Ritt's Factorization of Polynomials. *Journal of the London Mathematical Society* **62**, 127–138. URL `http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=58787`.

Arnaud Bodin, Pierre Dèbes & Salah Najib (2009). Indecomposable polynomials and their spectrum. *Acta Arithmetica* **139(1)**, 79–100.

F. Dorey & G. Whaples (1974). Prime and Composite Polynomials. *Journal of Algebra* **28**, 88–101. URL `http://dx.doi.org/10.1016/0021-8693(74)90023-4`.

Joachim von zur Gathen (2008a). Counting decomposable multivariate polynomials. *Preprint,* 21 pages. URL `http://arxiv.org/abs/0811.4726`.

Joachim von zur Gathen (2008b). Counting decomposable univariate polynomials. *Preprint,* 92 pages. URL `http://arxiv.org/abs/0901.0054`.

Gregor Kemper (1996). Calculating Invariant Rings of Finite Groups over Arbitrary Fields. *Journal of Symbolic Computation* **21**, 351–366.

Gregor Kemper (2009). Personal communication.

Peter S. Landweber & Robert E. Stong (1987). The depth of rings of invariants over finite fields. In *Number Theory*, D. V. Chudnovsky, G. V. Chudnovsky, H. Cohn & M. B. Nathanson, editors, volume 1240 of *Lecture Notes in Mathematics*, 259–274. Springer-Verlag, New York.

H. Levi (1942). Composite Polynomials with coefficients in an arbitrary Field of characteristic zero. *American Journal of Mathematics* **64**, 389–400.

R. Lidl, G. L. Mullen & G. Turnwald (1993). *Dickson polynomials.* Number 65 in Pitman Monographs and Surveys in Pure and Applied Mathematics. Longman Scientific & Technical. ISBN 0-582-09119-5.

Rudolf Lidl & Harald Niederreiter (1983). *Finite Fields.* Number 20 in Encyclopedia of Mathematics and its Applications. Addison-Wesley, Reading MA.

Mara D. Neusel & Larry Smith (2002). *Invariant Theory of Finite Groups*, volume 94 of *Mathematical Surveys and Monographs.* American Mathematical Society, USA.

J. F. Ritt (1922). Prime and Composite Polynomials. *Transactions of the American Mathematical Society* **23**, 51–66. URL `http://www.jstor.org/stable/1988911`.

Andrzej Schinzel (1982). *Selected Topics on Polynomials.* Ann Arbor; The University of Michigan Press. ISBN 0-472-08026-1.

Andrzej Schinzel (2000). *Polynomials with special regard to reducibility.* Cambridge University Press, Cambridge, UK. ISBN 0521662257.

Larry Smith (1995). *Polynomial Invariants of Finite Groups*, volume 6 of *Research Notes in Mathematics.* A K peters, Ltd., Wellesley, MA. ISBN 978-1-56881-053-9.

Pierre Tortrat (1988). Sur la composition des polynômes. *Colloquium Mathematicum* **55**(2), 329–353.
U. Zannier (1993). Ritt's Second Theorem in arbitrary characteristic. *Journal für die reine und angewandte Mathematik* **445**, 175–203. URL http://www.digizeitschriften.de/index.php?id=loader&tx_jkDigiTools_pi1[IDDOC]=503382.

B-IT, Universität Bonn, D-53113 Bonn
E-mail address: gathen@bit.uni-bonn.de