DE GRUYTER

FINITE FIELDS AND THEIR APPLICATIONS

PROCEEDINGS OF THE 14TH INTERNATIONAL CONFERENCE ON FINITE FIELDS AND THEIR APPLICATIONS, VANCOUVER, JUNE 3-7, 2019

Edited by James A. Davis

PROCEEDINGS IN MATHEMATICS

James A. Davis (Ed.)
Finite Fields and their Applications

De Gruyter Proceedings in Mathematics

Finite Fields and their Applications

Proceedings of the 14th International Conference on Finite Fields and their Applications, Vancouver, June 3-7, 2019

Edited by
James A. Davis

DE GRUYTER

Editor
Dr. James A. Davis
Department of Mathematics
and Computer Science
University of Richmond
410 Westhampton Way
Richmond 23173
USA
jdavis@richmond.edu

ISBN 978-3-11-062123-5
e-ISBN (PDF) 978-3-11-062173-0
e-ISBN (EPUB) 978-3-11-062217-1

Library of Congress Control Number: 2020944447

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Typesetting: VTeX UAB, Lithuania
Printing and binding: CPI books GmbH, Leck

www.degruyter.com

Contents

John Sheekey and David Thomson
A note on depth-*b* normal elements —— 1

Yoshinori Hamahata
Values of Dirichlet–Goss series with periodic coefficients —— 11

Gove Effinger
On elements of normal depth-2 in quartic extensions of F_p —— 25

James Davis, John Polhill, and Ken Smith
Relative linking systems of difference sets and linking systems of relative difference sets —— 43

Christian Kaspers and Alexander Pott
On solving isomorphism problems about 2-designs using block intersection numbers —— 51

Yuta Kodera, Sylvain Duquesne, and Yasuyuki Nogami
Multiplication and squaring in cubic and quartic extensions for pairing based cryptography —— 71

Pascale Charpin
***Crooked* functions —— 87**

Michael D. Fried
Diophantine statements over residue fields: Galois stratification and uniformity —— 103

Dorian Goldfeld and Giacomo Micheli
The algebraic theory of fractional jumps —— 133

Gary McGuire and Daniela Mueller
Some results on linearized trinomials that split completely —— 149

Tahseen Rabbani and Ken W. Smith
Non-Abelian orthogonal building sets —— 165

Satoru Fukasawa and Katsushi Waki
Examples of plane rational curves with two Galois points in positive characteristic —— 181

Charles J. Colbourn and Violet R. Syrotiuk
Covering strong separating hash families —— 189

Eduardo Camps, Hiram H. López, Gretchen L. Matthews, and Eliseo Sarmiento
Monomial-Cartesian codes closed under divisibility —— 199

John Sheekey and David Thomson

A note on depth-*b* normal elements

Abstract: In this paper, we study elements $\beta \in \mathbb{F}_{q^n}$ having normal α-depth b; that is, elements for which $\beta, \beta-\alpha, \ldots, \beta-(b-1)\alpha$ are simultaneously normal elements of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$. In [1], the authors present the definition of normal 1-depth but mistakenly present results for normal α-depth for some fixed normal element $\alpha \in \mathbb{F}_{q^n}$. We explain this discrepancy and generalize the given definition of normal (1–)depth from [1] as well as answer some open questions presented in [1].

Keywords: Finite fields, normal bases, primary decomposition

MSC 2010: 11T30, 11T71, 12Y05

1 Introduction and notation

Throughout this document, we use the following standard notation. Let p be a prime and let q be a power of p, the finite field of q elements is denoted $\mathbb{F}_q$, and the finite degree n extension of $\mathbb{F}_q$ is denoted $\mathbb{F}_{q^n}$. The (relative) trace function is denoted $\mathrm{Tr}_{\mathbb{F}_{q^n}:\mathbb{F}_q}: \mathbb{F}_{q^n} \to \mathbb{F}_q$. We remark that the trace function is onto, and for any $k \not\equiv 0 \pmod{p}$, the element $k\alpha$ is also normal. For any positive integer n, denote by $e = \nu_p(n)$, the *p-ary valuation* of n; that is, the largest integer e such that p^e divides n but p^{e+1} does not divide n. We also denote by $\tau = p^e$; specifically, $\tau = 1$ ($e = 0$) if $\gcd(p, n) = 1$.

In Section 2, we derive conditions for elements to be normal that we will use later in the paper. In Section 3, we correct and generalize the notion of normal elements of depth b from [1]. Also motivated by [1], in Section 4 we observe that depth is not necessarily invariant under conjugation, and further analyze the depth of the conjugates of normal elements.

2 Finite fields as Frobenius modules

In this section, we follow [2, 3] and introduce finite fields as Frobenius modules. Let $\sigma_q: \overline{\mathbb{F}_q} \to \overline{\mathbb{F}_q}$ denote the Frobenius q-automorphism. Clearly, σ_q fixes $\mathbb{F}_q$ and for any $n > 0$ and $\alpha \in \overline{\mathbb{F}_q}$, $\sigma_q^n(\alpha) = \alpha$ if and only $\alpha \in \mathbb{F}_{q^n}$. Moreover, the Galois group of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ is cyclic of order n and generated by σ_q.

John Sheekey, University College Dublin, Belfield, Dublin 4, Ireland, e-mail: john.sheeekey@ucd.ie
David Thomson, Tutte Institute for Mathematics and Carleton University, 1125 Colonel By Dr., Ottawa, Ontario K1S 5B6, Canada, e-mail: dthomson@math.carleton.ca

https://doi.org/10.1515/9783110621730-001

Let $\alpha \in \mathbb{F}_{q^n}$ and let $\mathcal{B}$ consist of the Galois orbit of α; that is, $\mathcal{B} = \{\alpha, \alpha^q, \dots, \alpha^{q^{n-1}}\}$. If $\mathcal{B}$ is a linearly independent set, then α is a *normal element* of $\mathbb{F}_{q^n}$ and $\mathcal{B}$ is a *normal basis* of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$. We also call α a *cyclic vector* for $\mathbb{F}_{q^n}$ as a vector space over $\mathbb{F}_q$.

For $f(x) = \sum_{i=0}^{m} a_i x^i$, denote the action of f on $\overline{\mathbb{F}_q}$ by

$$f \circ \alpha = f(\sigma_q)(\alpha) = \sum_{i=0}^{m} a_i \alpha^{q^i}.$$

Clearly, $(f+g) \circ \alpha = f \circ \alpha + g \circ \alpha$ for any $f, g \in \mathbb{F}_q[x]$, and $(x^n - 1) \circ \alpha = 0$ if and only if $\alpha \in \mathbb{F}_{q^n}$. Moreover, $(fg) \circ \alpha = f \circ (g \circ \alpha)$, so that if $f \circ \alpha = 0$ for any $\alpha \in \mathbb{F}_{q^n}$, then f divides $x^n - 1$.

Definition 1.

1. For any $\alpha \in \mathbb{F}_{q^n}$, define the *annihilator* of α as the polynomial $\mathrm{ann}_\alpha \in \mathbb{F}_q[x]$ of smallest degree such that $\mathrm{ann}_\alpha \circ \alpha = 0$.
2. For any $f \in \mathbb{F}_q[x]$, define $\ker(f) = \{\alpha \in \mathbb{F}_{q^n} : \mathrm{ann}_\alpha = f\}$, the set of elements of $\mathbb{F}_{q^n}$ annihilated by f under $\circ$.

Observe that ann_α annihilates any linear combination of Galois conjugates of α. We have $\ker(x^n - 1) = \mathbb{F}_{q^n}$ and $\mathrm{ann}_\alpha(x)$ divides $x^n - 1$ for any α. Moreover, α is a normal element of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ if and only if $\mathrm{ann}_\alpha(x) = x^n - 1$ by linear independence of the conjugates of α. We summarize these observations in Proposition 1.

Proposition 1. *For any prime power q, the number of normal elements of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ is given by $\Phi_q(x^n - 1)$, where Φ_q is Euler's totient function over $\mathbb{F}_q$; that is, $\Phi_q(x^n - 1)$ is the number of polynomials in $\mathbb{F}_q[x]$ of degree less than n that are relatively prime with $x^n - 1$.*

Existence of normal elements can be gleaned directly from Proposition 1, since $\Phi_q(x^n - 1)$ is nonzero for all $n \geq 1$.

We now introduce a map central to the remainder of this work. Suppose $\alpha \in \mathbb{F}_{q^n}$ is normal and define the map $\phi_\alpha : \mathbb{F}_q[x] \to \mathbb{F}_{q^n}$ by $\phi(f) = f \circ \alpha$. Then $\ker(\phi_\alpha) = (x^n - 1)$, since α is normal; similarly ϕ_α is onto since the set $\mathcal{B}_\alpha = \{\alpha, \alpha^q, \dots, \alpha^{q^{n-1}}\}$ is a basis. Hence $\mathbb{F}_{q^n} \cong \mathbb{F}_q[x]/(x^n - 1)$ as Frobenius modules. We will abuse notation and refer to this isomorphism also as ϕ_α.

Let $g(x) = \sum_{i=0}^{n-1} g_i x^i \in \mathbb{F}_q[x]$, and $\beta = \phi_\alpha(g) = \sum_{i=0}^{n-1} g_i \alpha^{q^i}$. Then $\beta^q = \sum_{i=0}^{n-1} g_{i-1} \alpha^{q^i}$. Thus $\phi_\alpha^{-1}(\beta^q) = x\phi_\alpha^{-1}(\beta) \bmod (x^n - 1)$. Thus the Frobenius action on $\mathbb{F}_q[x]/(x^n - 1)$ is induced by $\overline{\sigma_q}(g) := xg(x)$, with $\overline{\sigma_q} = \phi_\alpha \sigma_q \phi_\alpha^{-1}$.

We exploit the decomposition of $\mathbb{F}_q[x]/(x^n - 1)$ as a Frobenius module. We follow the treatment in [4]. Let $e = \nu_p(n)$ be the valuation of n at p and let $x^n - 1 = f_1^{e_1} \cdots f_r^{e_r}$ be the primary factorization of $x^n - 1$; then $e_i = p^e = \tau$ for all $i = 1, \dots, r$. In particular, $\tau = 1$ if $\gcd(p, n) = 1$. Denote by $\overline{V_i} = \mathbb{F}_q[x]/(f_i^\tau)$, then

$$\mathbb{F}_q[x]/(x^n - 1) \cong \bigoplus_{i=1}^{r} \overline{V_i}. \tag{2.1}$$

Explicitly, we write the image of g in $\bigoplus_{i=1}^{r} \overline{V_i}$ as $(g \bmod f_1^\tau, \ldots, g \bmod f_r^\tau)$. We abuse notation slightly and write $V_i = \phi_\alpha(\overline{V_i})$.

$$\bigoplus_{i=1}^{r} V_i \cong \mathbb{F}_{q^n} \cong \mathbb{F}_q[x]/(x^n - 1) \cong \bigoplus_{i=1}^{r} \overline{V_i}. \tag{2.2}$$

Equation (2.2) is the *primary decomposition* of $\mathbb{F}_{q^n}$ as a Frobenius module. Moreover, we observe that each V_i is stable under σ_q.

Proposition 2. *Let α be a normal element of $\mathbb{F}_{q^n}$, and suppose $\beta = \phi_\alpha(g(x))$. Then* $\operatorname{ann}_\beta = \frac{x^n-1}{\gcd(x^n-1,g(x))}$, *and β is normal if and only if* $\gcd(x^n - 1, g(x)) = 1$. *Furthermore,* $V_i = \ker(f_i^\tau)$.

Proof. Let $f(x) = \sum_{i=0}^{m} a_i x^i$. Then $f \circ \beta = 0$ if and only if $f(x)g(x) \in (x^n - 1)$. The smallest degree polynomial satisfying $f(x)g(x) \in (x^n - 1)$ is clearly $\frac{x^n-1}{\gcd(x^n-1,g(x))}$, as claimed. Since β is normal if and only if $\operatorname{ann}_\beta(x) = x^n - 1$, β is normal if and only if $\gcd(x^n - 1, g(x)) = 1$.

Now $\beta \in \phi_\alpha(\overline{V_i})$ if and only if f_j^τ divides $g(x)$ for all $j \neq i$, which occurs if and only if $f_i(x)^\tau g(x) \in (x^n - 1)$, if and only if $f_i^\tau \circ \beta = 0$, if and only if $\beta \in \ker(f_i^\tau)$. □

We summarize the characterizations of normal elements here.

Proposition 3. *Let α be a normal element of $\mathbb{F}_{q^n}$, and suppose $\beta = \phi_\alpha(g(x))$. Let $x^n - 1 = f_1^\tau \cdots f_r^\tau$, with the f_i being distinct irreducible polynomials in $\mathbb{F}_q[x]$. Let $g_i = g \bmod f_i^\tau$, and $\beta = \sum_{i=1}^{r} \beta_i$ for $\beta_i \in V_i$. Then the following are equivalent:*

1. *β is normal;*
2. *$\gcd(x^n - 1, g(x)) = 1$;*
3. *$\gcd(f_i, g_i) = 1$ for each i;*
4. *$\operatorname{ann}_{\beta_i} = f_i^\tau$ for each i;*
5. *$\beta_i \in \ker(f_i^\tau) \setminus \ker(f_i^{\tau-1})$ for each i.*

Proof. (1. $\iff$ 2.) This is Proposition 2.

(2. $\iff$ 3.) Let $g_i = g \bmod f_i^\tau$, then $g = hf_i^\tau + g_i$ for some $h \in \mathbb{F}_q[x]$. Then (the irreducible) f_i divides g_i for some $1 \leq i \leq r$ if and only if f_i divides g, contradicting $\gcd(x^n - 1, g(x)) = 1$.

(3. $\iff$ 4.) Let $g_i = g \bmod f_i^\tau$ with $\beta_i = \phi_\alpha(g_i) \in V_i$. Clearly, $\beta_i \in V_i$ if and only if $\beta_i \in \ker(f_i^\tau)$, so $\operatorname{ann}_{\beta_i} = f_i^k$ for $1 \leq k \leq \tau$. Now, $f_i^k \circ \beta_i = 0$ if and only if $f_i^k g_i \circ \alpha = 0$ if and only if $f_i^k g_i \in (x^n - 1)$. Now $\gcd(f_i, g_i) = 1$ if and only if $k = \tau$ for all i.

(4. $\iff$ 5.) By the minimality of $\operatorname{ann}_{\beta_i}$, we have $\operatorname{ann}_{\beta_i} = f_i^\tau$ if and only if $\beta_i \in \ker(f_i^\tau)$ and $\beta_i \notin \ker(f_i^{\tau-1})$. □

If $\gcd(p, n) = 1$, then $\tau = 1$, and thus we get the following.

Corollary 1. *Let* $\gcd(p, n) = 1$ *and let $\beta = \beta_1 + \beta_2 + \cdots + \beta_r$ with $\beta_i \in V_i$; then β is a normal element if and only if $\prod_{i=1}^{r} \beta_i \neq 0$.*

3 Depth-b normal elements

Definition 2. Let $b \in \mathbb{N}$ with $b \leq p$. If $\beta \in \mathbb{F}_{q^n}$ is such that $\beta, \beta-\alpha, \ldots, \beta-(b-1)\alpha$ are normal elements of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ for some $\alpha \in \mathbb{F}_{q^n}$; then we say that β has *normal α-depth b*.

In [1], the authors introduced normal depth, where the definition was for $\theta = 1$. However, the results in [1] are in fact referring to normal α-depth, for some fixed normal element α. We will explain the discrepancy below, and consider the more general problem.

We remark that Definition 2 can be extended for $b \geq p$ when q is a power of p by imposing an ordering on the elements of $\mathbb{F}_q$ (or even further still, on $\mathbb{F}_{q^n}$). Since [1] and Section 4 are mostly concerned with depth 2, we will not treat these sorts of extensions in this work.

We recap (and generalize) the main question from [1].

Question 1. To what extent do the conjugates of an element β having normal α-depth b also have normal α-depth b?

In particular in [1], they focus on normal depth 2 and search for *lonely elements*: that is, normal elements of depth 2 having a conjugate that fails to have normal depth 2.

Lemma 1. *Without loss of generality, fix a normal element α of $\mathbb{F}_{q^n}$ satisfying $\mathrm{Tr}_{\mathbb{F}_{q^n}:\mathbb{F}_q}(\alpha) = n/\tau$, since if α' is any normal element with $\mathrm{Tr}_{\mathbb{F}_{q^n}:\mathbb{F}_q}(\alpha') = k \neq 0$, the element $\alpha = \alpha' \frac{\tau}{nk}$ is normal (since $\tau/n \not\equiv 0 \pmod p$). Then:*

1. *$\phi_\alpha^{-1}(\alpha^{q^i}) = x^i$ and $\phi_\alpha^{-1}(1) = (\tau/n)\frac{x^n-1}{x-1}$;*
2. *the image of α in $\bigoplus_{i=1}^r \overline{V_i}$ is $(1,1,\ldots,1)$, and the image of 1 is $((x-1)^{\tau-1}, 0, \ldots, 0)$.*

Proof.

1. For $f(x) = \sum_{i=0}^m a_i x^i$, we have $\phi_\alpha(f) = f \circ \alpha = \sum_{i=0}^m f_i \alpha^{q^i}$. Hence, $x^i \circ \alpha = \alpha^{q^i}$ or $\phi_\alpha^{-1}(\alpha^{q^i}) = x^i$. Similarly,

$$\mathrm{Tr}_{\mathbb{F}_{q^n}:\mathbb{F}_q}(\alpha) = \left(\sum_{i=0}^{n-1} x^i\right) \circ \alpha = n/\tau,$$

   so by linearity, $\phi_\alpha^{-1}(1) = (\tau/n)\sum_{i=0}^{n-1} x^i = (\tau/n)\frac{x^n-1}{x-1}$.

2. Since $\alpha = \phi_\alpha(1)$, we have $g_i = g \bmod f_i^\tau = 1$ for all $1 \leq i \leq r$. Similarly, $1 = (\tau/n)\phi_\alpha(\frac{x^n-1}{x-1})$, and with $\tau = p^{\nu_p(n)}$ and by linearity of Frobenius,

$$\begin{aligned}\frac{x^n-1}{x-1} &= \sum_{i=0}^{n-1} x^i \equiv (n/\tau)\sum_{i=0}^{\tau-1} x^i \bmod (x^\tau - 1)\\ &= (n/\tau)\frac{x^\tau-1}{x-1} = (n/\tau)(x-1)^{\tau-1}.\end{aligned}$$

□

Proposition 4. *Let $\alpha \in \mathbb{F}_{q^n}$ be normal. An element $\beta = \phi_\alpha(g(x))$ has normal α-depth b if and only if $\gcd(x^n - 1, g(x) - c) = 1$ for all $c \in \{0, \dots, b-1\}$.*

Proof. The proof is immediate from the linearity of ϕ and from Proposition 3, Remark 2. □

In [1], the number $\#\{g : \gcd(x^n - 1, g(x) - c) = 1 \; \forall \; c \in \{0, \dots, b-1\}\}$ was defined as $\Phi_b(x^n - 1)$.

Theorem 1. *Let $\alpha \in \mathbb{F}_{q^n}$ be normal with $\mathrm{Tr}_{\mathbb{F}_{q^n}:\mathbb{F}_q}(\alpha) = \tau/n$, let $e = v_p(n)$, and let $\beta = \phi_\alpha(g(x))$ also be normal. Then:*

1. *if $e > 0$, then β has normal 1-depth p; moreover, $\beta - c$ is normal for all $c \in \mathbb{F}_q$;*
2. *if $e = 0$, then β has normal 1-depth b if and only if $g(1) \geq b$ (under a suitable implicit ordering of the elements of $\mathbb{F}_q$). In particular, $\beta - c$ is normal if and only if $g(1) \neq c$.*

Proof. Let $g_i = g \bmod f_i^\tau$. Then the image of $\beta - c$ in $\bigoplus_{i=1}^r \overline{V_i}$ is $(g_1 - c(x-1)^{\tau-1}, g_2, \dots, g_r)$.

If $e > 0$ and β is normal, then $\gcd(g_1, (x-1)^\tau) = 1$. If $\beta - c$ is not normal, then $(x-1)$ divides $g_1 - c(x-1)^{\tau-1}$, implying $(x-1)$ divides g_1, a contradiction. Thus $\beta - c$ is normal for all $c \in \mathbb{F}_q$.

If $e = 0$, then $g_1 = g(1)$, and the image of g is $(g(1) - c, g_2, \dots, g_r)$. By Corollary 1, β is normal if and only if $g_i \neq 0$ for each i. Hence, $\beta - c$ is *not* normal if and only if $g(1) = c$. □

In [1], the authors mistakenly state that the number of elements having normal 1-depth b is equal to $\Phi_b(x^n - 1)$. This assumably arose by the erroneous assumption that $\phi_\alpha(1) = 1$. Instead, since $\phi_\alpha(1) = \alpha$, $\Phi_b(x^n - 1)$ refers to the number of elements having normal α-depth b, and so for the remainder of this paper we focus on this case as well.

4 Conjugates: lonely and sociable elements

Throughout this section, we use the notation from Section 3; in particular, $x^n - 1 = (f_1 \cdots f_r)^\tau$ where $n = \tau m$ with $\gcd(m, \tau) = 1$, and f_i is irreducible for $1 \leq i \leq r$. Suppose $\beta = \phi_\alpha(g(x))$ has normal α-depth b. We consider the normal α-depth of its conjugates. Recall that $\beta^{q^i} = \phi_\alpha(x^i g(x))$. Thus we need to consider the common divisors of $x^i g(x) - c$ with $x^n - 1$, or equivalently $g(x) - cx^i$ with $x^n - 1$.

Definition 3. An element $\beta \in \mathbb{F}_{q^n}$ is (α, b)-*lonely* if β has normal α-depth b, but β^{q^i} does not have normal α-depth b for some i. If β^{q^i} has normal α-depth b for all i, we say that β is (α, b)-*sociable*.

Similar to Proposition 3, we have a number of equivalent characterizations of sociable elements.

Theorem 2. *Let $x^n - 1 = f_1^\tau f_2^\tau \cdots f_r^\tau$ with f_i irreducible, $1 \le i \le r$. Let $\beta \in \mathbb{F}_{q^n}$ with $g(x) = \phi_\alpha^{-1}(\beta)$ and let $g_i = g \bmod f_i^\tau$. Then the following are equivalent:*

1. *β is (α, b)-sociable;*
2. *$\gcd(x^n - 1, g(x) - cx^j) = 1$ for all $j \in \{0, \dots, n-1\}$, $c \in \{0, \dots, b-1\}$;*
3. *$\gcd(f_i, g_i - cx^j) = 1$ for all $i \in \{1, \dots, r\}$, $j \in \{0, \dots, n-1\}$, $c \in \{0, \dots, b-1\}$;*
4. *$g(\theta) \notin \{c\theta^j : c \in \{0, \dots, b-1\}, j \in \{0, \dots, n-1\}\}$ and θ a root of $x^n - 1$.*

Proof. The equivalence of items 1, 2, 3 come directly from applying Proposition 3 to Definition 3. Here, we prove only $3 \iff 4$.

Suppose $\gcd(f_i, g_i - cx^j) \neq 1$ for some $1 \le i \le r$, $0 \le j \le n-1$, which occurs if and only if $f_i(\theta_i) = g_i(\theta_i) - c\theta_i^j = 0$ for some $\theta_i \in \mathbb{F}_{q^{\deg(f_i)}}$; that is, $g_i(\theta_i) = c\theta_i^j$. The fourth equivalence follows, since $g_i = g \bmod f_i^\tau$, so $g(\theta_i) = g_i(\theta_i)$. □

The number of β that are (α, b)-sociable is the number of g satisfying the conditions on their roots given in the fourth equivalence of Theorem 2.

Lemma 2. *Let $x^n - 1 = f_1^\tau \cdots f_r^\tau$ and let θ_i be a root of f_i, $1 \le i \le r$. Then there are exactly $q^{\deg(f_i)}$ possible values for $g(\theta_i)$ for $g \in \mathbb{F}_q[x]$. Furthermore, let $\theta_{ij} = \theta_i^{q^j}$ for $j = 0, 1, \dots, \deg(f_i)$ be the roots of f_i in $\mathbb{F}_q(\theta_i)$, and fix $\gamma_i \in \mathbb{F}_q(\theta_i)$, $1 \le i \le r$. Then there exist precisely $q^{\frac{n(\tau-1)}{\tau}}$ polynomials g of degree at most n with $g(\theta_{ij}) = \gamma_i^{q^j}$ for all $1 \le i \le r$, $0 \le j \le \deg(f_i) - 1$.*

Proof. Clearly, $g(\theta_i) \in \mathbb{F}_q(\theta_i) = \mathbb{F}_{q^{\deg(f_i)}}$, and so there are at most $q^{\deg(f_i)}$ possible values for $g(\theta_i)$. As g has coefficients in $\mathbb{F}_q$, we have that $g(\theta_i^{q^j}) = g(\theta_i)^{q^j}$ for any j.

With $n = n_0\tau$, two polynomials g and h in $\mathbb{F}_q[x]$ agree on all n_0-th roots of unity if and only if $f_1 f_2 \cdots f_r$ divides $g - h$. As $\deg(f_1 f_2 \cdots f_r) = n_0$, there are $q^{n-n_0} = q^{n_0(\tau-1)}$ such polynomials h of degree at most n. □

For β that are (α, b)-sociable, Theorem 2 provides a number of forbidden values for $g(\theta_i)$. The precise number of forbidden values that ensure that β is (α, b)-sociable is complicated in general, but we can solve it completely in some cases left open in [1].

Proposition 5. *The number of elements in $\mathbb{F}_{q^n}$ that are (α, b)-sociable is at most*

$$q^{\frac{n(\tau-1)}{\tau}} \prod_{i=1}^{r} \left(q^{\deg(f_i)} - n(b-1) - 1\right)$$

Proof. By Lemma 2, there are at most $q^{\deg(f_i)}$ choices for $g(\theta_i)$ for each $i = 1, \dots, r$. By the final assertion of Theorem 2, an upper bound on the number of forbidden choices of $g(\theta_i)$ occurs when all of $c\theta_i^j$ are distinct for all $c \in \{1, \dots, b-1\}$ and $j \in \{0, \dots, n-1\}$. This gives $n(b-1)$ forbidden values for $g(\theta_i)$, and the further restriction $g(\theta_i) \neq 0$ together with Lemma 2 completes the proof. □

Proposition 6. *Suppose* $(n, q-1) = 1$. *Then the number of elements that are* (α, b)*-sociable is*

$$q^{\frac{n(r-1)}{r}} \prod_{i=1}^{r} (q^{\deg(f_i)} - (b-1)\mathrm{ord}(\theta_i) - 1),$$

where θ_i *is a root of* f_i, $1 \le i \le r$.

Proof. Since $(n, q-1) = 1$, $x^n - 1$ has only one root in $\mathbb{F}_q$, namely 1. Thus as each θ_i^j is an nth root of 1 (in some extension field), we have that $\theta_i^j \notin \mathbb{F}_q$ for all i and all $1 < j < \mathrm{ord}(\theta_i)$. Therefore, $\#\{c\theta_i^j : c \in \{1, \dots, b-1\}, j \in \{0, \dots, n-1\}\} = (b-1)\mathrm{ord}(\theta_i)$. As $g(\theta_i) \neq 0$, there are $q^{\deg(f_i)} - (b-1)\mathrm{ord}(\theta_i) - 1$ choices for $g(\theta_i)$ for each i for which $\phi_\alpha(g)$ is (α, b)-sociable. The factor $q^{\frac{n(r-1)}{r}}$ follows from Lemma 2. □

Corollary 2. *Suppose* $n = q^s$. *Then the number of elements that are* (α, b)*-sociable is*

$$q^{q^s - q^{s-1}}(q-b).$$

For a specific example of Corollary 2, taking $q = n$, $b = 2$, we get that there are $q^{q-1}(q-2)$ elements which are $(\alpha, 2)$-sociable in $\mathbb{F}_{q^q}$.

Corollary 3. *Suppose* n *is prime,* $n \notin \{p, q-1\}$, *and let* $x^n - 1 = (x-1)f_2 \cdots f_r$. *Then the number of elements that are* (α, b)*-sociable is*

$$(q-b)\prod_{i=2}^{r} (q^{\deg(f_i)} - (b-1)n - 1).$$

In [1], focus is applied to the case $b = 2$, the case of $(\alpha, 2)$-lonely/sociable elements. We now apply Theorem 2 to this situation.

Proposition 7. *Suppose* $n|(q-1)$. *Then the number of elements that are* $(\alpha, 2)$*-sociable is*

$$\prod_{i=1}^{n} \left(q - \frac{n}{(i,n)} - 1 \right). \tag{4.1}$$

Proof. As $n|(q-1)$, $x^n - 1$ factorizes in to a product of distinct linear factors over $\mathbb{F}_q$. Let $f_i = x - \theta_i$. Then β is $(\alpha, 2)$-sociable if and only if $g(\theta_i) \neq 0, \theta_i^j$ for any j. Thus the number of forbidden choices for $g(\theta_i)$ is $\mathrm{ord}(\theta_i) + 1$. Letting θ be a primitive nth root of unity in $\mathbb{F}_q$, and letting $\theta_i = \theta^i$, then $\mathrm{ord}(\theta_i) = \frac{n}{(i,n)}$ and the result follows. □

Remark 1. Note that Formula (4.1) is not true in general. Issues arise when there exist $c_1, c_2 \in \{0, \dots, b-1\}$ such that $c_1 = c_2\theta_i^j$, in which case $\#\{c\theta_i^j : c \in \{0, \dots, b-1\}, j \in \{0, \dots, n-1\}\}$ is more difficult to calculate. The conditions of the previous two theorems were chosen to avoid this possibility.

The following example of Proposition 7 provides an answer to the first open question left in [1].

Example 1. Suppose $n = 3$, and suppose $x^3 - 1$ factors into distinct linear factors over $\mathbb{F}_q$, say $x^3 - 1 = (x-1)(x-\lambda)(x-\mu)$; equivalently, if $q \equiv 1 \pmod 3$. Then $\phi_\alpha(g)$ has normal α-depth 2 if and only if $\{0,1\} \cap \{g(1), g(\lambda), g(\mu)\} = \emptyset$. Similarly, $\phi_\alpha(g)^{q^i}$ has normal α-depth 2 if and only if $\{0,1\} \cap \{g(1), \lambda^i g(\lambda), \mu^i g(\mu)\} = \emptyset$. Since a polynomial of degree at most three is uniquely determined by its evaluation at three different elements of $\mathbb{F}_q$, then there are $(q-2)^3$ elements of α-depth 2, of which $(q-2)(q-4)^2$ are not lonely. Thus there are $4(q-2)(q-3)$ lonely elements.

We can also apply Proposition 7 to provide a partial answer to the second open question in [1].

Example 2. Suppose $n = 4$, $q = 5$, then $x^4 - 1 = (x-1)(x-2)(x-3)(x-4)$, with $\mathrm{ord}(1) = 1$, $\mathrm{ord}(4) = 2$ and $\mathrm{ord}(2) = \mathrm{ord}(3) = 4$. A direct application of Proposition 7 shows that there are no $(\alpha, 2)$-sociable elements.

Example 2 generalizes in an obvious way.

Proposition 8. *Let $q = n+1$, then there are no $(\alpha, 2)$-sociable elements of $\mathbb{F}_{q^n}$.*

Proof. Let θ be a primitive element in $\mathbb{F}_q$. Then $x^n - 1 = x^{q-1} - 1 = \prod_{\lambda \in \mathbb{F}_q^*}(x - \lambda) = \prod_{i=0}^{q-2}(x - \theta^i)$. Hence for $\beta = \phi_\alpha(g)$ to be $(\alpha, 2)$-sociable it would require that $g(\theta) \neq 0$ and $g(\theta) \neq \theta^i$ for any $0 \leq i \leq q-2$, which is impossible as $g(\theta) \in \mathbb{F}_q$. □

The following was proved in [1, Proposition 4.3]. We include an alternative proof here.

Proposition 9. *Suppose $\frac{x^n-1}{x-1}$ is irreducible over $\mathbb{F}_q$. Then the number of elements that are $(\alpha, 2)$-sociable is*

$$(q-2)(q^{n-1} - n - 1),$$

and the number of elements that are $(\alpha, 2)$-lonely is

$$(q-2)(n-1).$$

Proof. Recall that $\frac{x^n-1}{x-1}$ is irreducible over $\mathbb{F}_q$ if and only if q is primitive modulo n. Then $\{\theta^{q^i} : i = 0, \ldots, n-2\} = \{\theta^i : i = 1, \ldots, n-1\}$ is the set of distinct roots of $\frac{x^n-1}{x-1}$. Thus, an element $\phi_\alpha(g)$ is $(\alpha, 2)$-sociable if and only if $g(1) \neq 0, 1$ and $g(\theta) \neq \theta^i$ for $i = 1, 2, \ldots, n-1$. Hence, there are $(q-2)(q^{n-1} - n + 1)$ elements that are $(\alpha, 2)$-sociable and $(q-2)(n-1)$ lonely elements in $\mathbb{F}_{q^n}$. □

The following example examines two cases of $(\alpha, 3)$-sociable elements, giving the first directions toward the third open problem in [1].

Example 3.

1. Consider the case $q = 7$, $n = 3$, $b = 3$. Then $x^n - 1 = (x-1)(x-2)(x-4)$, and $2^3 = 1$. Now the set $\{c\theta^j : c \in \{0,1,2\}, j \in \{0,1,2\}\}$ is equal to $\{0,1,2\}$ for $\theta = 1$ and $\{0,1,2,4\}$ for $\theta = 2,4$. Thus the number of $(\alpha, 3)$-sociable elements is $(7-3)(7-4)^2 = 36$.
2. In the case $q = 13$, $n = 3$, $b = 3$, we have $x^n - 1 = (x-1)(x-3)(x-9)$. Now the set $\{c\theta^j : c \in \{0,1,2\}, j \in \{0,1,2\}\}$ is equal to $\{0,1,2\}$ for $\theta = 1$ and $\{0,1,2,3,5,6,9\}$ for $\theta = 3,9$. Thus the number of $(\alpha, 3)$-sociable elements is $(13-3)(13-7)^2 = 36$.

These two examples illustrates how extra care must be taken when an element of $\{0,\dots,b-1\}$ is a nontrivial nth root of unity.

5 Conclusions and future directions

In this paper, we study a generalization of normal elements of depth b, as presented in [1]. Since depth is not invariant under conjugation, we further analyze the depth of the conjugates of normal elements.

The notion of depth readily lends itself to further generalization. One such "natural" generalization is as follows. Given some total ordering $\mathcal{O}$ of the elements of $\mathbb{F}_{q^n}$, say $\mathcal{O} = \{o_0, o_1, \dots, o_{q^n-1}\}$, an element $\beta \in \mathbb{F}_{q^n}$ has normal $(\mathcal{O}, \alpha)$-depth b if $\beta - o_0\alpha, \beta - o_1\alpha, \dots, \beta - o_{b-1}\alpha$ are simultaneously normal. Here, β has α-depth b if $o_i = i$ for $i = 0,\dots,b-1$. Some interesting questions here occur when α is a normal element of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ and $\mathcal{O}_\zeta = (0, \zeta, \zeta^2, \dots, \zeta^{q^n-2})$, for a primitive element $\zeta \in \mathbb{F}_{q^n}$. Determining conditions for which β has $(\mathcal{O}_\zeta, \alpha)$-depth 2, or statistics on the possible values of b for which β has $(\mathcal{O}_\zeta, \alpha)$-depth b is the subject of future work.

Bibliography

[1] G. Effinger and G. L. Mullen, Two extended Euler functions with applications to latin squares and bases of finite field extensions, *Bull. Inst. Comb. Appl.*, **85** (2019), 92–111.
[2] H. W. Lenstra and R. Schoof, Primitive normal bases for finite fields, *Math. Comput.*, **48** (1987), 217–231.
[3] L. Reis and D. Thomson, Existence of primitive 1-normal elements in finite fields, *Finite Fields Appl.*, **51** (2018), 238–269.
[4] A. Steel, A new algorithm for the computation of canonical forms of matrices over fields, *J. Symb. Comput.*, **24** (1997), 409–432.

Yoshinori Hamahata

Values of Dirichlet–Goss series with periodic coefficients

Abstract: Let $\overline{\mathbb{F}_q(T)}$ be an algebraic closure of $\mathbb{F}_q(T)$. For a function $f : \mathbb{F}_q[T] \to \overline{\mathbb{F}_q(T)}$, the series $L(s,f) = \sum_{a\in\mathbb{F}_q[T]:\mathrm{monic}} f(a)a^{-s}$ is called the Dirichlet–Goss series. We examine the values $L(p^n,f)$ $(n \geq 0)$ for a nonzero periodic function f with an irreducible period and prove their transcendence results.

Keywords: Periodic function, Dirichlet series, L-function, function field

MSC 2010: Primary 11J72, Secondary 11M38, 11M41, 11R60

1 Introduction

For a function $f : \mathbb{N} \to \mathbb{C}$, the series $L(s,f) = \sum_{n=1}^{\infty} f(n)n^{-s}$ is called the *Dirichlet series*. In the early 1960s, Sarvadaman Chowla [6] made the following conjecture: *Let f be a nonzero rational-valued periodic function defined on the integers with prime period such that $f(p) = 0$. Then $L(1,f) \neq 0$ if this converges.*

Chowla [6] proved this in the case where f is an odd function. Subsequently, he [7] asked if there exists a nonzero rational valued periodic function $f : \mathbb{Z} \to \mathbb{Q}$ with prime period p such that $L(1,f)$ converges and is zero without the condition $f(p) = 0$. In 1973, Baker, Birch, and Wirsing [4] answered the above conjecture using Baker's theory of linear forms in logarithms. Their theorem is as follows.

Theorem 1.1 (Baker, Birch, and Wirsing [4]). *Let m be a positive integer and f a nonzero function defined on the integers with algebraic values and period m such that* (i) *$f(r) = 0$ if $1 < \gcd(r,m) < m$*; (ii) *The mth cyclotomic polynomial Ψ_m is irreducible over $\mathbb{Q}(f(1),\dots,f(m))$. Then $L(1,f) \neq 0$ if this converges.*

Under the same assumption of Theorem 1.1, Adhikari, Saradha, Shorey, and Tijdeman [1] proved that $L(1,f)$ is a transcendental number if this converges. Concerning the values $L(k,f)$ $(k \geq 1)$, Okada [15] obtained the following theorem.

Theorem 1.2 (Okada [15]). *Let k be a positive integer, and let f be a nonzero function defined on the integers with algebraic values and period $m > 2$ such that* (i) *f is even or odd according to k is even or odd*; (ii) *$f(n) = 0$ if $\gcd(n,m) > 1$*; (iii) *The mth cyclotomic polynomial Ψ_m is irreducible over $\mathbb{Q}(f(1),\dots,f(m))$. Then $L(k,f) \neq 0$ if this converges.*

Yoshinori Hamahata, Department of Applied Mathematics, Okayama University of Science, Ridai-cho 1-1, Okayama 700-0005, Japan, e-mail: hamahata@xmath.ous.ac.jp

https://doi.org/10.1515/9783110621730-002

Let $A = \mathbb{F}_q[T]$ and let K be the quotient field of A. We write $\overline{K}$ for an algebraic closure of K. For a function $f : A \to \overline{K}$, the series $L(s,f) = \sum_{a\in A:\text{monic}} f(a)a^{-s}$ is called the *Dirichlet–Goss series*. When f is the inclusion map $\iota : A \to \overline{K}$, the corresponding series is the Carlitz zeta function $\zeta_C(s) = \sum_{a\in A:\text{monic}} a^{-s}$. It is well known that for any positive integer n, $\zeta_C(n)$ is transcendental over $\overline{K}$. For a monic irreducible polynomial $P \in A$, any character $\chi : (A/P)^* \to \overline{K}$ can be extended to a Dirichlet character $\chi : A \to \overline{K}$ by $\chi(a) = 0$ whenever $P|a$. This is a periodic function with period P. The corresponding series $L(s,\chi)$ is called the *Goss L-function*. Lutes and Papanikolas [12] investigated transcendence properties of this value $L(1,\chi)$. In [11], we investigated $L(s,f)$ for a nonzero periodic function f and established an analog of Theorem 1.2. In this paper, using the results from [2, 12], we examine the values $L(p^n,f)$ $(n \geq 0)$ for a nonzero periodic function f with an irreducible period and prove their transcendence results.

The remainder of the paper is organized as follows. In Section 2, we recall some results needed for our study. In Section 3, we prove the results for periodic functions on A. In Section 4, we state our results for the Dirichlet–Goss series. In Section 5, we provide the proofs of our results stated in Section 4. In Section 6, we apply our results to polygamma functions, the Hurwitz zeta function, and the Euler–Lehmer constants in function fields.

2 Preliminaries

Let $\mathbb{F}_q$ be the finite field with q elements, where q is a power of the prime number p. Let $A = \mathbb{F}_q[T]$ and $K = \mathbb{F}_q(T)$. Let $K_\infty = \mathbb{F}_q((T^{-1}))$ be the completion of K at $\infty = (T^{-1})$, and let $\mathbb{C}_\infty$ be the completion of an algebraic closure $\overline{K_\infty}$ of K_∞. Let $P \in A_+$ be an irreducible element of degree $d > 0$. For a ring R, let R^* be the unit group of R.

2.1 Carlitz exponential

We write $A\{\tau\}$ for the twisted polynomial ring whose multiplication is defined by $\tau a = a^q\tau$ $(a \in A)$. The $\mathbb{F}_q$-linear ring homomorphism $\rho : A \to A\{\tau\}$, defined by $1 \mapsto \tau^0$ and $T \mapsto \rho_T = T\tau^0 + \tau$, is called the *Carlitz A-module*. Using each $M \in A \setminus \{0\}$, ρ associates an additive polynomial $\rho_M(x)$ given by $\rho_M(x) := \rho_M(\tau)(x) \in A[x]$. This is called the *Carlitz M-polynomial*. For $M \in A \setminus \{0\}$, let $\rho[M] = \{\alpha \in \mathbb{C}_\infty \mid \rho_M(\alpha) = 0\}$ be the set of Carlitz M-torsion points. The set $\rho[M]$ is a cyclic A-module and its generator (as a Carlitz A-module) is called the *primitive Carlitz M-torsion point*. The minimal polynomial $\Phi_M(x)$ of any primitive M-torsion point over K is called the *Carlitz M-th cyclotomic polynomial*. The polynomials $\rho_M(x)$ and $\Phi_M(x)$ have degrees $q^{\deg M}$ and $\varphi(M)$, respectively, where $\varphi(M) := \#(A/MA)^*$. For details on these polynomials, we refer the reader to [3].

There exists a unique entire function $e(z)$ over $\mathbb{C}_\infty$ such that for each $a \in A$, we have $\rho_a(e(z)) = e(az)$ (see [9, Chapter 3]). The function $e(z)$ is called the *Carlitz exponential*. Let L be the set of all zeros of $e(z)$. Then L is a rank one free A-module (see [9, Corollary 3.2.9]). It is well known that $L = \overline{\pi}A$ is analogous to $\pi\mathbb{Z}$.

2.2 Anderson's log-algebraicity formulas

We recall log-algebraicity formulas from Anderson's work [2].

We set $\mathbf{e}(z) = e(\overline{\pi}z)$. According to Anderson [2], for $a \in A$ with $P \nmid a$ and $m \geq 1$, there exist unique $\mathbf{e}_m^*(a) \in \overline{K}$ such that for $a, b \in A$ relatively prime to P,

$$\sum_{m=1}^{q^d-1} \mathbf{e}_m^*(a)\mathbf{e}(b/P)^m = P \cdot \delta_{ab} = \begin{cases} P & \text{if } a \equiv b \pmod{P}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

Furthermore, for $m, n \in \{1, \dots, q^d - 1\}$, it holds that

$$\sum_{\substack{a \in A_+ \\ \deg a < d}} \mathbf{e}_m^*(a)\mathbf{e}(a/P)^n = P \cdot \delta_{mn} = \begin{cases} P & \text{if } m = n, \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

Theorem 2.1 (Anderson [2]). *Let m be a nonnegative integer. Then the power series*

$$S_m(x, z) := e\Big(\sum_{a \in A_+} \frac{\rho_a(x)^m}{a} z^{q^{\deg a}} \Big) \in K[x][[z]] \tag{2.3}$$

belongs to $A[x, z]$.

For $m, n \in \mathbb{N}$, let

$$\ell_m(n, z) = \sum_{a \in A_+} \frac{\mathbf{e}(az)^m}{a^n}. \tag{2.4}$$

Then, using (2.3), we obtain

$$e(\ell_m(1, 1/P)) = S_m(\mathbf{e}(1/P), 1) \in A[\mathbf{e}(1/P)] \subset \overline{K}. \tag{2.5}$$

Using (2.1), we obtain

$$\sum_{\substack{a \in A_+ \\ a \equiv b \pmod{P}}} \frac{1}{a} = \frac{1}{P} \sum_{m=1}^{q^d-1} \mathbf{e}_m^*(b)\ell_m(1, 1/P) \tag{2.6}$$

for any $b \in A$ relatively prime to P.

2.3 Lutes and Papanikolas' transcendence results

We recall the results of Lutes and Papanikolas [12, 16].

Theorem 2.2 (Papanikolas [16]). *Let $\lambda_1, \dots, \lambda_r \in \mathbb{C}_\infty$ satisfy $e(\lambda_i) \in \overline{K}$ for each $i = 1, \dots, r$. If $\lambda_1, \dots, \lambda_r$ are linearly independent over K, then they are algebraically independent over $\overline{K}$.*

Let

$$\mathcal{N} = \{1\} \cup \{m \in \mathbb{N} \mid 1 \le m \le q^d - 1, m \not\equiv 1 \pmod{q-1}\}.$$

Then $\ell_m(1, 1/P)$ $(m \in \mathcal{N})$ are linearly independent over K. Using (2.5) and Theorem 2.2, $\ell_m(1, 1/P)$ $(m \in \mathcal{N})$ are algebraically independent over $\overline{K}$.

Theorem 2.3 (Lutes and Papanikolas [12]). *Let Θ_P be the group of Dirichlet characters modulo P on A. Then, for each $\chi \in \Theta_P$, $L(1, \chi)$ is transcendental over $\overline{K}$. Furthermore, it holds that*

$$\mathrm{tr.deg}_{\overline{K}} \overline{K}(L(1,\chi) \mid \chi \in \Theta_P) = \frac{(q^d - 1)(q-2)}{q-1} + 1.$$

3 Periodic functions

Let $P \in A_+$ be an irreducible polynomial with $\deg P = d > 0$.

3.1 Parity conditions

The following theorem demonstrates that each function on A can be decomposed into some functions satisfying parity conditions.

Lemma 3.1. *For any function $f : A \to \mathbb{C}_\infty$, there exist unique functions $f_i : A \to \mathbb{C}_\infty$ $(i = 1, \dots, q-1)$ such that:*
(i) $f_i(\epsilon z) = \epsilon^i f(z) \quad (\epsilon \in \mathbb{F}_q^*)$;
(ii) $f = f_1 + \cdots + f_{q-1}$.

Proof. Let ζ be a generator of $\mathbb{F}_q^*$. We first prove the existence of $f_1(z), \dots, f_{q-1}(z)$. For $i = 1, \dots, q-1$, let

$$g_i(z) = -\sum_{n=1}^{q-1} \zeta^{ni} f(\zeta^n z).$$

It holds that

$$f(z) = g_1(z) + \cdots + g_{q-1}(z),$$
$$g_i(\zeta z) = \zeta^{q-1-i} g_i(z) \quad (i = 1, \dots, q-1).$$

Hence, letting $f_i(z) = g_{q-1-i}(z)$, we obtain $f_1(z), \dots, f_{q-1}(z)$ satisfying (i) and (ii).

We next prove the uniqueness of $f_1(z), \dots, f_{q-1}(z)$. To this end, it is sufficient to demonstrate that $f_1 = \cdots = f_{q-1} = 0$ if $f = 0$. We substitute $\zeta^i z$ $(i = 1, \dots, q-1)$ into $f_1(z) + \cdots + f_{q-1}(z) = 0$, successively. Thus, we obtain

$$\begin{pmatrix} \zeta & \zeta^2 & \cdots & \zeta^{q-1} \\ \zeta^2 & \zeta^4 & \cdots & \zeta^{2(q-1)} \\ \vdots & \vdots & & \vdots \\ \zeta^{q-1} & \zeta^{2(q-1)} & \cdots & \zeta^{(q-1)(q-1)} \end{pmatrix} \begin{pmatrix} f_1(z) \\ f_2(z) \\ \vdots \\ f_{q-1}(z) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Because the matrix in the left-hand side is invertible, $f_1 = \cdots = f_{q-1} = 0$. □

3.2 Fourier expansions

For a subfield E of $\mathbb{C}_\infty$, we write $\mathcal{F}(P; E)$ for the set of E-valued periodic functions on A with period P. We call $f \in \mathcal{F}(P; E)$ *of Dirichlet type* if $f(P) = 0$. Set $\mathcal{FD}(P; E) = \{f \in \mathcal{F}(P; E) \mid f(P) = 0\}$. These sets $\mathcal{F}(P; E)$ and $\mathcal{FD}(P; E)$ become vector spaces over E. Let $\nu_P : A \to \overline{K}$ be a periodic function with period P defined by $\nu_P(a) = 1$ if $P|a$, and 0 otherwise. Then it holds that

$$\mathcal{F}(P; E) = \mathcal{FD}(P; E) \oplus E \cdot \nu_P.$$

When $f \in \mathcal{FD}(P; \mathbb{C}_\infty)$, for $a \in A$ with $P \nmid a$ and $m \in \{1, \dots, q^d - 1\}$,

$$\widehat{f}_m(a) := \frac{1}{P} \sum_{\substack{b \in A \\ \deg b < d}} f(b) \mathbf{e}_m^*(ab) \tag{3.1}$$

is called the *Fourier transform* of f. The following theorem is the *Fourier inversion formula* for periodic functions.

Proposition 3.2. *For any $f \in \mathcal{FD}(P; \mathbb{C}_\infty)$, there exist unique $c_m \in \mathbb{C}_\infty$ $(m = 1, \dots, q^d - 1)$ such that*

$$f(a) = \sum_{m=1}^{q^d-1} c_m \mathbf{e}(a/P)^m. \tag{3.2}$$

Proof. We first prove the existence of c_m. Let $c_m = \widehat{f}_m(1)$. Then, using (2.1) and (3.1), it holds that for $a \in A$ with $P \nmid a$,

$$\sum_{m=1}^{q^d-1} c_m \mathbf{e}(a/P)^m = \frac{1}{P} \sum_{\substack{b \in A \\ \deg b < d}} f(b) \sum_{m=1}^{q^d-1} \mathbf{e}_m^*(b) \mathbf{e}(a/P)^m = f(a).$$

We next prove the uniqueness of c_m. Assume that $f(a)$ can be written as (3.2). Then, using (2.2) for $n \in \{1, \dots, q^d - 1\}$,

$$\widehat{f}_n(1) = \frac{1}{P} \sum_{\substack{a \in A \\ \deg a < d}} f(a)\mathbf{e}_n^*(a) = \frac{1}{P} \sum_{m=1}^{q^d-1} c_m \sum_{\substack{a \in A \\ \deg a < d}} \mathbf{e}_n^*(a)\mathbf{e}(a/P)^m = c_n. \qquad \square$$

For a subfield E of $\mathbb{C}_\infty$ including $\mathbb{F}_q$, set

$$\mathcal{FD}_i(P; E) = \{f \in \mathcal{FD}(P, E) \mid f(\epsilon z) = \epsilon^i f(z) \ (\epsilon \in \mathbb{F}_q^*)\}$$

for $i = 1, \dots, q-1$. These become vector spaces over E. Using Lemma 3.1, we obtain

$$\mathcal{FD}(P; E) = \bigoplus_{i=1}^{q-1} \mathcal{FD}_i(P; E).$$

For a positive integer m, we define a periodic function $h_m : A \to \overline{K}$ by $h_m(z) = e(z/P)^m$.

Proposition 3.3. *Let $i \in \{1, \dots, q-1\}$.*

(i) *The set $\{h_m \mid 1 \le m \le q^d - 1, m \equiv i \pmod{q-1}\}$ is a basis of $\mathcal{FD}_i(P; \mathbb{C}_\infty)$ over $\mathbb{C}_\infty$. Hence,*

$$\dim_{\mathbb{C}_\infty} \mathcal{FD}_i(P; \mathbb{C}_\infty) = \frac{q^d - 1}{q - 1}.$$

(ii) *The set $\{h_m \mid 1 \le m \le q^d - 1, m \equiv i \pmod{q-1}\}$ is a basis of $\mathcal{FD}_i(P; \overline{K})$ over $\overline{K}$. Hence,*

$$\dim_{\overline{K}} \mathcal{FD}_i(P; \overline{K}) = \frac{q^d - 1}{q - 1}.$$

Proof. (i) It is obvious that this set is contained in $\mathcal{F}_i(P; \mathbb{C}_\infty)$. Using Proposition 3.2, this set is a basis of $\mathcal{F}_i(P; \mathbb{C}_\infty)$ over $\mathbb{C}_\infty$.

(ii) It is obvious that this set is contained in $\mathcal{F}_i(P; \overline{K})$. If $f \in \mathcal{F}_i(P; \overline{K})$, by definition, $\widehat{f}_m(1) \in \overline{K}$ for any m. Hence, using Proposition 3.2, f can be written uniquely as a linear combination of $\{h_m \mid 1 \le m \le q^d - 1, m \equiv i \pmod{q-1}\}$ over $\overline{K}$. $\square$

4 Dirichlet–Goss series

4.1 Results under parity conditions

Let $P \in A_+$ be an irreducible polynomial of $\deg P = d > 0$. In [11], we proved the following theorem, which is an analog of Theorem 1.2.

Theorem 4.1 ([11]). *Let n be a positive integer. We take a nonzero $g \in \mathcal{FD}(P;\overline{K})$ such that:*
(i) $g(\epsilon z) = \epsilon^n g(z) \quad (\epsilon \in \mathbb{F}_q^*)$;
(ii) *The Carlitz Pth cyclotomic polynomial Φ_P is irreducible over $K(g(a) \mid a \in A/PA)$.*

Then $L(n,g) \neq 0$.

Example 4.2. Let $P = T$. We define $g : A \to \overline{K}$ by $g(z) = \mathbf{e}(z/T)^{q-1}$. Clearly, $g(a) = -T$ if $T \nmid a$, and 0 if $T|a$. This implies that $g \in \mathcal{FD}_{q-1}(T;\overline{K})$. Using Theorem 4.1, $L(q-1,g) \neq 0$. This value can be written as $L(q-1,g) = (T^{2-q} - T)\zeta_C(q-1)$. Using BC_{q-1}, which is the Bernoulli–Carlitz number of order $q-1$ (Goss [9]), we obtain $\zeta_C(q-1) = BC_{q-1}\overline{\pi}^{q-1}$, which yields

$$L(q-1,g) = \left(\frac{1}{T^{q-2}} - T\right) BC_{q-1}\overline{\pi}^{q-1}.$$

For a positive integer n, let $P_n(z) = \sum_{l \in L}(z+l)^{-n}$. According to Goss [8], there exists a monic polynomial $G_n(X) \in K[X]$ of degree n such that $P_n(z) = G_n(e(z)^{-1})$. This is called the *Goss polynomial*. For example, $G_n(X) = X^n$ if $n \leq q$. To prove Theorem 4.1, we used the following lemma.

Lemma 4.3 ([11]). *If $g \in \mathcal{FD}(P;\overline{K})$ satisfies the condition (i) of Theorem 4.1, then*

$$\begin{aligned} L(n,g) &= \left(\frac{\overline{\pi}}{P}\right)^n \sum_{\substack{b \in A_+ \\ \deg b < d}} g(b) P_n(\overline{\pi} b/P) \\ &= \left(\frac{\overline{\pi}}{P}\right)^n \sum_{\substack{b \in A_+ \\ \deg b < d}} g(b) G_n\big(\mathbf{e}(b/P)^{-1}\big) \end{aligned} \tag{4.1}$$

Here, the sum on the right-hand side of (4.1) belongs to $\overline{K}$. Because $\overline{\pi}^n$ is transcendental over $\overline{K}$, we obtain the following theorem.

Theorem 4.4.
(i) *If $g \in \mathcal{FD}(P;\overline{K})$ satisfies $g(\epsilon z) = \epsilon^n g(z)$ $(\epsilon \in \mathbb{F}_q^*)$ and $L(n,g) \neq 0$, then $L(n,g)$ is transcendental over $\overline{K}$, and can be written as $\overline{\pi}^n \alpha$ for a nonzero $\alpha \in \overline{K}$.*
(ii) *Under the assumption of Theorem 4.1, $L(n,g)$ is transcendental over $\overline{K}$, and can be written as $\overline{\pi}^n \alpha$ for a nonzero $\alpha \in \overline{K}$.*

Remark 4.5. If $g \in \mathcal{FD}(P;\overline{K})$ does not satisfy (ii) of Theorem 4.1, it is possible that $L(n,g) = 0$. For example, let $q = 3$ and $P = T^2 + 1$. Then P is irreducible in A. If

$$g(\epsilon) = \mathbf{e}(\epsilon/P), \quad g(\epsilon T) = -\mathbf{e}(\epsilon T/P) \qquad (\epsilon \in \mathbb{F}_q^*),$$

then the nonzero periodic function $g : A \to \overline{K}$ with period P can be obtained. Using (4.1) and $G_1(X) = X$, $L(1,g) = 0$.

4.2 Main theorems

For a nonnegative integer n, there exists $i \in \{1,\dots,q-1\}$ such that $p^n \equiv i \pmod{q-1}$. The following three theorems are our main results.

Theorem 4.6. *Let $f \in \mathcal{FD}(P;\overline{K})$ be a nonzero periodic function with the decomposition $f = \sum_{j=1}^{q-1} f_j$ as in Lemma 3.1.*

(i) *If the cyclotomic polynomial Φ_P is irreducible over $K(f_i(b) \mid b \in A/PA)$, then $L(p^n,f)$ is transcendental over $\overline{K}$.*

(ii) *If $f \neq f_i$, then $L(p^n,f)$ is transcendental over $\overline{K}$.*

Theorem 4.7. *We have*

$$\operatorname{tr.deg}_{\overline{K}}\overline{K}(L(p^n,f) \mid f \in \mathcal{FD}_j(P;\overline{K})) = \begin{cases} (q^d-1)/(q-1) & \text{if } j \neq i, \\ 1 & \text{if } j = i. \end{cases} \tag{4.2}$$

$$\begin{aligned} &\operatorname{tr.deg}_{\overline{K}}\overline{K}(L(p^n,f) \mid f \in \mathcal{F}(P;\overline{K})) \\ &\quad = \operatorname{tr.deg}_{\overline{K}}\overline{K}(L(p^n,f) \mid f \in \mathcal{FD}(P;\overline{K})) = \operatorname{tr.deg}_{\overline{K}}\overline{K}(\ell_{mp^n}(p^n,1/P) \mid m \in \mathcal{N}) \\ &\quad = \frac{(q^d-1)(q-2)}{q-1} + 1. \end{aligned} \tag{4.3}$$

Theorem 2.3 can be generalized as follows.

Theorem 4.8. *Let Θ_P be the group of Dirichlet characters modulo P on A. Then, for each $\chi \in \Theta_P$, $L(p^n,\chi)$ is transcendental over $\overline{K}$. Furthermore, it holds that*

$$\operatorname{tr.deg}_{\overline{K}}\overline{K}(L(p^n,\chi) \mid \chi \in \Theta_P) = \frac{(q^d-1)(q-2)}{q-1} + 1.$$

5 Proofs of Theorems 4.6, 4.7, and 4.8

5.1

For $j \in \{1,\dots,q-1\}$, set

$$\mathcal{L}_j(p^n) = \{L(p^n,f) \mid f \in \mathcal{FD}_j(P;\overline{K})\},$$

which becomes a vector space over $\overline{K}$ under the operations

$$\begin{aligned} L(p^n,f) + L(p^n,g) &= L(p^n,f+g), \\ c \cdot L(p^n,f) &= L(p^n,c\cdot f) \quad (f,g \in \mathcal{FD}_j(P;\overline{K}), c \in \overline{K}). \end{aligned}$$

Using Proposition 3.3 and Lemma 4.3, we have

$$\dim_{\overline{K}} \mathcal{L}_j(p^n) \leq \begin{cases} (q^d-1)/(q-1) & \text{if } j \neq i, \\ 1 & \text{if } j = i. \end{cases}$$

Because $\ell_m(1,1/P)$ $(m \in \mathcal{N})$ are algebraically independent over $\overline{K}$, $\ell_{mp^n}(p^n,1/P) = \ell_m(1,1/P)^{p^n}$ $(m \in \mathcal{N})$ are also algebraically independent over $\overline{K}$. Using $p^n \equiv i \pmod{q-1}$, we have $\ell_{p^n}(p^n,1/P) \in \mathcal{L}_i(p^n)$, which yields that $\ell_{p^n}(p^n,1/P)$ is a basis of $\mathcal{L}_i(p^n)$ and $\dim_{\overline{K}} \mathcal{L}_i(p^n) = 1$. Observing $\ell_{mp^n}(p^n,1/P) \in \sum_{j\neq i} \mathcal{L}_j(p^n)$ $(m \in \mathcal{N} \setminus \{1\})$, it follows that $\dim_{\overline{K}} \sum_{j\neq i} \mathcal{L}_j(p^n) = (q^d-1)(q-2)/(q-1)$, which yields $\ell_{mp^n}(p^n,1/P)$ $(m \in \mathcal{N}\setminus\{1\}, mp^n \equiv j \pmod{q-1})$ form a basis of $\mathcal{L}_j(p^n)$ and $\dim_{\overline{K}} \mathcal{L}_j(p^n) = (q^d-1)/(q-1)$.

5.2 Proof of Theorem 4.6

For $f \in \mathcal{FD}(P;\overline{K})$, $L(p^n,f)$ can be written uniquely as a linear combination of $\ell_{mp^n}(p^n, 1/P)$ $(m \in \mathcal{N})$ over $\overline{K}$. Clearly, $L(p^n,f) = L(p^n,f_i) + L(p^n,f-f_i)$ and that $L(p^n,f_i) = \alpha \cdot \ell_{p^n}(p^n,1/P)$ and $L(p^n,f-f_i) = \sum_{m\in\mathcal{N}\setminus\{1\}} \beta_m \ell_m(p^n,1/P)$ for some $\alpha, \beta_m \in \overline{K}$. Hence, it holds that

$$L(p^n,f) = 0 \Leftrightarrow L(p^n,f_i) = L(p^n,f-f_i) = 0. \tag{5.1}$$

(i) When $f = f_i$, the claim is obvious from Theorem 4.4.

When $f \neq f_i$, $f - f_i \neq 0$. If $f(z) - f_i(z) = \sum_{m\in\mathcal{N}\setminus\{1\}} \beta_m \mathbf{e}(z/P)^m$, then there exists $m \in \mathcal{N}\setminus\{1\}$ such that $\beta_m \neq 0$. Then we obtain

$$L(p^n,f-f_i) = \sum_{m\in\mathcal{N}\setminus\{1\}} \beta_m \ell_m(1,1/P)^{p^n} \neq 0.$$

Using (5.1), $L(p^n,f) \neq 0$, which yields (i).

(ii) This follows from the proof of (i). □

5.3 Proof of Theorem 4.7

The equality (4.2) follows from the discussion in 5.1. We prove (4.3). Let $\gamma =: \zeta_C(1) = \sum_{a\in A_+} a^{-1}$ be the Euler constant, which is known to be transcendental over $\overline{K}$. If $\rho_P(z) = \sum_{i=0}^d \rho_i z^{q^i}$, then $\sum_{i=0}^d \rho_i \mathbf{e}(a/P)^{q^i} = \rho_P(\mathbf{e}(a/P)) = 0$ for $a \in A_+$ with $P \nmid a$. Observing $\rho_0 = P$, $1 = -P^{-1}\sum_{i=1}^d \rho_i \mathbf{e}(a/P)^{q^i-1}$ is obtained. Hence, we obtain

$$\sum_{\substack{a\in A_+\\ P\nmid a}} \frac{1}{a} = -\frac{1}{P}\sum_{i=1}^d \rho_i \ell_{q^i-1}(1,1/P).$$

Because $\gamma = \sum_{\substack{a\in A_+\\ P\nmid a}} a^{-1} + \gamma/P$,

$$\gamma = \frac{P}{P-1}\sum_{\substack{a\in A_+\\ P\nmid a}} \frac{1}{a} = \frac{1}{1-P}\sum_{i=1}^d \rho_i \ell_{q^i-1}(1,1/P),$$

which implies

$$L(p^n,\nu_P)=\sum_{\substack{a\in A_+\\ P|a}}\frac{1}{a^{p^n}}=\left(\frac{\gamma}{P}\right)^{p^n}=\frac{1}{P^{p^n}(1-P^{p^n})}\sum_{i=1}^{d}\rho_i^{p^n}\ell_{(q^i-1)p^n}(p^n,1/P).$$

Therefore, we obtain $\{L(p^n,f)\mid f\in\mathcal{F}(P;\overline{K})\}=\{L(p^n,f)\mid f\in\mathcal{FD}(P;\overline{K})\}$, which yields the first equality. From the discussion in 5.1, $\{\ell_{mp^n}(p^n,1/P)\mid m\in\mathcal{N}\}$ is a basis of $\sum_{j=1}^{q-1}\mathcal{L}_j(p^n)$. This yields the second and third equalities. □

5.4 Proof of Theorem 4.8

For any $\chi\in\Theta_P$, there exists $j\in\{1,\dots,q-1\}$ such that $\chi\in\mathcal{FD}_j(P;\overline{K})$. When $j\neq i$, using Theorem 4.4(ii), $L(p^n,\chi)$ is transcendental over $\overline{K}$. When $j=i$, using $L(p^n,\chi)\neq 0$ and Theorem 4.4(i), $L(p^n,\chi)$ is transcendental over $\overline{K}$.

We next prove the latter part. Raising both sides of (2.6) to the p^nth power, we obtain

$$\sum_{\substack{a\in A_+\\ a\equiv b\ (\mathrm{mod}\ P)}}\frac{1}{a^{p^n}}=\frac{1}{P^{p^n}}\sum_{m=1}^{q^d-1}\mathbf{e}_m^*(b)^{p^n}\ell_{mp^n}(p^n,1/P).\tag{5.2}$$

For $\chi\in\Theta_P$, multiplying (5.2) by $\chi(b)$, and taking the sum over b, we obtain

$$L(p^n,\chi)=\sum_{m=1}^{q^d-1}\frac{1}{P^{p^n}}\sum_{\substack{0\neq b\in A\\ \deg b<d}}\chi(b)\mathbf{e}_m^*(b)^{p^n}\ell_{mp^n}(p^n,1/P).\tag{5.3}$$

Using (2.1), we obtain

$$\ell_{mp^n}(p^n,1/P)=\sum_{\chi\in\Theta_P}\left(\sum_{\substack{0\neq b\in A\\ \deg b<d}}\chi^{-1}(b)\mathbf{e}(b/P)^{mp^n}\right)L(p^n,\chi).\tag{5.4}$$

Combining (5.3) with (5.4), we obtain

$$\begin{aligned}&\mathrm{tr.deg}_{\overline{K}}\overline{K}(L(p^n,\chi)\mid\chi\in\Theta_P)\\&\quad=\mathrm{tr.deg}_{\overline{K}}\overline{K}(\ell_{mp^n}(p^n,1/P)\mid m\in\mathcal{N})=\frac{(q^d-1)(q-2)}{q-1}+1.\end{aligned}$$

□

6 Applications: polygamma values, Hurwitz zeta values, and Euler–Lehmer constants

In this section, we apply the results in Section 4 to polygamma functions, the Hurwitz zeta function, and the Euler–Lehmer constants in function fields. Set $A_0 = A_+ \cup \{0\}$. Let $P \in A_+$ be an irreducible polynomial with $\deg P = d > 0$.

6.1 Polygamma values

For a nonnegative integer k, the function

$$\psi_k(z) = (-1)^{k+1} \sum_{a \in A_0} \frac{1}{(z+a)^{k+1}}$$

is called the *kth polygamma function*. In particular, when $k = 0$, $\psi(z) := \psi_0(z)$ is called the *digamma function*, which is the logarithmic derivative of the geometric gamma function $\Gamma(z) = z^{-1} \prod_{a \in A_+} (1 + z/a)^{-1}$. Note that $\psi_k(z)$ can be obtained from $\psi(z)$ by $\psi_k(z) = \mathcal{D}_k \psi(z)$, where $\mathcal{D}_k$ is the kth hyperdifferential operator in z, as discussed by Bosser and Pellarin [5].

Theorem 6.1. *Let n be a nonnegative integer. For any nonzero $b \in A$ with $\deg b < \deg P$, $\psi_{p^n-1}(b/P)$ is transcendental over $\overline{K}$. Furthermore, we obtain*

$$\operatorname{tr.deg}_{\overline{K}} \overline{K}(\psi_{p^n-1}(b/P) \mid 0 \neq b \in A, \deg b < \deg P) = \frac{(q^d-1)(q-2)}{q-1} + 1. \tag{6.1}$$

Proof. First, we prove the case $n = 0$. We define $f \in \mathcal{FD}(P; \overline{K})$ by

$$f(a) = \begin{cases} 1 & \text{if } a \equiv b \pmod P, \\ 0 & \text{otherwise.} \end{cases}$$

Then it holds that

$$\psi(b/P) = -P \sum_{a \in A_0} \frac{1}{b + Pa} = \begin{cases} -P \cdot L(1,f) & \text{if } b \in A_+, \\ -P/b - P \cdot L(1,f) & \text{if } b \notin A_+ \end{cases} \tag{6.2}$$

When $q = 2$, $f = f_1$. Using Theorem 4.4(ii), $L(1,f)$ is transcendental over $\overline{K}$. When $q \neq 2$, $f \neq f_1$. Using Theorem 4.6(ii), $L(1,f)$ is transcendental over $\overline{K}$. Thus, $\psi(b/P)$ is transcendental over $\overline{K}$.

Using (2.6) and (6.2),

$$\psi(b/P) = \begin{cases} -\sum_{m=1}^{q^d-1} \mathbf{e}_m^*(b) \ell_m(1, 1/P) & \text{if } b \in A_+, \\ -P/b - \sum_{m=1}^{q^d-1} \mathbf{e}_m^*(b) \ell_m(1, 1/P) & \text{if } b \notin A_+. \end{cases} \tag{6.3}$$

On the other hand, using (2.2), for $1 \le n \le q^d - 1$,

$$\ell_n(1,1/P) = \begin{cases} -P^{-1} \sum_{\substack{b \in A_+ \\ \deg b < d}} \mathbf{e}(b/P)^n \psi(b/P) & \text{if } b \in A_+, \\ -P^{-1} \sum_{\substack{b \in A_+ \\ \deg b < d}} \mathbf{e}(b/P)^n (\psi(b/P) + P/b) & \text{if } b \notin A_+. \end{cases} \tag{6.4}$$

Using Theorem 4.7, (6.3), and (6.4), the left-hand side of (6.1) becomes

$$\operatorname{tr.deg}_{\overline{K}} \overline{K}(\ell_m(1,1/P) \mid m \in \mathcal{N}) = \frac{(q^d-1)(q-2)}{q-1} + 1.$$

The general case $n \ge 0$ follows from the fact $\psi_{p^n-1}(z) = \psi(z)^{p^n}$ and Theorem 4.7. □

6.2 Hurwitz zeta values

For $-x \in K_\infty \setminus A_0$, we define the *Hurwitz zeta function* $\zeta(s,x)$ by

$$\zeta(s,x) = \sum_{a \in A_0} \frac{1}{(a+x)^s} \quad (s \in \mathbb{N}).$$

When $x = b/P$ $(0 \ne b \in A, \deg b < d)$, $\zeta(p^n, b/P) = (-1)^{p^n} \psi_{p^n-1}(b/P)$. Using Theorem 6.1, the following theorem holds.

Theorem 6.2. *Let n be a nonnegative integer. For any nonzero $b \in A$ with $\deg b < \deg P$, $\zeta(p^n, b/P)$ is transcendental over $\overline{K}$. Furthermore, we obtain*

$$\operatorname{tr.deg}_{\overline{K}} \overline{K}(\zeta(p^n, b/P) \mid 0 \ne b \in A, \deg b < \deg P) = \frac{(q^d-1)(q-2)}{q-1} + 1.$$

6.3 Euler–Lehmer constants

Let $\gamma = \sum_{a \in A_+} a^{-1}$ be the Euler constant used in 5.3. This number can be generalized as follows. We take $M \in A_+$ with $\deg M > 0$, and $b \in A$ with $\deg b < \deg M$. Then the infinite sum

$$\gamma(b,M) = \sum_{\substack{a \in A_+ \\ a \equiv b \pmod M}} \frac{1}{a}$$

is called a *Euler–Lehmer constant*. This number has the following properties:

$$\sum_{\substack{b \in A \\ \deg b < \deg M}} \gamma(b,M) = \gamma, \tag{6.5}$$

$$\gamma(0,M) = \gamma/M, \tag{6.6}$$

$$\gamma(b,M) = \begin{cases} -\psi(b/M)/M & \text{if } b \in A_+, \\ -\psi(b/M)/M - 1/b & \text{if } b \notin A_+. \end{cases} \tag{6.7}$$

Using Theorem 6.1, (6.5), (6.6), and (6.7), the following theorem holds.

Theorem 6.3. *For any $b \in A$ with $\deg b < \deg P$, $\gamma(b,P)$ is transcendental over $\overline{K}$. Furthermore, we obtain*

$$\begin{aligned} &\mathrm{tr.deg}_{\overline{K}}\overline{K}(\gamma(b,P) \mid b \in A, \deg b < \deg P) \\ &= \mathrm{tr.deg}_{\overline{K}}\overline{K}(\gamma(b,P) \mid 0 \neq b \in A, \deg b < \deg P) = \frac{(q^d-1)(q-2)}{q-1} + 1. \end{aligned}$$

Remark 6.4. For the classical case, we state related results and conjectures.

Let K be an algebraic number field over which the mth cyclotomic polynomial Ψ_m is irreducible. The classical kth polygamma function $\psi_k(z)$ is defined by $\psi_k(z) = (-1)^{k+1}k!\sum_{n=0}^{\infty}(z+n)^{-k-1}$. Then Murty and Saradha [13] conjectured that the $\varphi(m)$ numbers, $\psi_k(a/m)$ $(1 \le a \le m, \gcd(a,m) = 1)$ are linearly independent over K.

Let $n > 1$ and $m > 2$. According to [10], Chowla–Milnor conjecture is that the $\varphi(m)$ classical Hurwitz zeta values $\zeta(n,a/m)$ $(1 \le a \le m, \gcd(a,m) = 1)$ are linearly independent over $\mathbb{Q}$. Gun, Murty, and Rath [10] provided a nontrivial lower bound of the dimension of the $\mathbb{Q}$-span of these numbers over $\mathbb{Q}$.

The classical Euler constant γ is defined by $\gamma = \lim_{x\to\infty}(\sum_{n\le x} n^{-1} - \log x)$. For a positive integer m and a nonnegative integer a with $0 \le a < m$, the classical Euler–Lehmer constant $\gamma(a,m)$ is defined by $\gamma(a,m) = \lim_{x\to\infty}(\sum_{\substack{n\le x \\ n\equiv a \pmod m}} n^{-1} - (\log x)/m)$. Murty and Saradha [14] proved that at most one number in the infinite list of numbers $\gamma(a,m)$ $(m \ge 2, 1 \le a < m)$ is an algebraic number. Therefore, it appears that these are all transcendental numbers.

Bibliography

[1] S. D. Adhikari, N. Saradha, T. N. Shorey, and R. Tijdeman, Transcendental infinite sums, *Indag. Math.*, **12** (2001), 1–14.

[2] G. W. Anderson, Log-algebraicity of twisted A-harmonic series and special values of L-series in characteristic p, *J. Number Theory*, **60** (1996), 165–209.

[3] S. Bae, The arithmetic of Carlitz polynomials, *J. Korean Math. Soc.*, **35** (1998), 341–360.

[4] A. Baker, B. J. Birch, and E. A. Wirsing, On a theorem of Chowla, *J. Number Theory*, **5** (1973), 224–236.

[5] V. Bosser and F. Pellarin, Hyperdifferential properties of Drinfeld quasi-modular forms, *Int. Math. Res. Not.*, **2008** (2008), 56.

[6] S. Chowla, A special infinite series, *Norske Vid. Selsk. Forth. (Trondheim)*, **37** (1964), 85–87.

[7] S. Chowla, The nonexistence of nontrivial linear relations between the roots of a certain irreducible equation, *J. Number Theory*, **2** (1970), 120–123.

[8] D. Goss, The algebraist's upper half-plane, *Bull. Am. Math. Soc.*, **2** (1980), 391–415.
[9] D. Goss, Basic Structures of Function Field Arithmetic, Springer, 1996.
[10] S. Gun, M. R. Murty, and P. Rath, On a conjecture of Chowla and Milnor, *Can. J. Math.*, **63** (2011), 1328–1344.
[11] Y. Hamahata, Chowla's theorem over function fields, *Int. J. Number Theory*, **14** (2018), 1689–1698.
[12] B. A. Lutes and M. A. Papanikolas, Algebraic independence of values of Goss *L*-functions, *J. Number Theory*, **133** (2013), 1000–1011.
[13] M. R. Murty and N. Saradha, Special values of the polygamma functions, *Int. J. Number Theory*, **5** (2009), 257–270.
[14] M. R. Murty and N. Saradha, Euler–Lehmer constants and a conjecture of Erdös, *J. Number Theory*, **130** (2010), 2671–2682.
[15] T. Okada, On an extension of a theorem of S. Chowla, *Acta Arith.*, **38** (1981), 341–345.
[16] M. A. Papanikolas, Tannakian duality for Anderson–Drinfeld motives and algebraic independence of Carlitz logarithms, *Invent. Math.*, **171** (2008), 123–174.

Gove Effinger

On elements of normal depth-2 in quartic extensions of $\mathbf{F}_p$

Abstract: Let p be a prime number and let θ be a fixed normal element of the extension field $\mathbf{F}_{p^n}$ of $\mathbf{F}_p$. An element $\alpha \in \mathbf{F}_{p^n}$ is said to have normal θ-depth-2 (or simply normal depth-2) if α and $\alpha - \theta$ are both normal. A central question is: To what extent are normal bases preserved by this depth operation; that is, if α has normal depth-2, do all of α's conjugates also have normal depth-2? The answer, in general, is that some bases are preserved and some are not. In previous work, specific counts of preserved bases are computed (as functions of p) for all quadratic and cubic extensions. In this paper, we obtain analogous counts for all quartic extensions $\mathbf{F}_{p^4}$ of $\mathbf{F}_p$.

Keywords: Finite fields, normal bases, Euler Φ-function

MSC 2010: 11T30, 11T55, 11A25

1 Introduction

We begin with a definition originally made in Section 4 of [1].

Definition 1.1. Let p be a prime number and let θ be a fixed normal element of the extension field $\mathbf{F}_{p^n}$ of $\mathbf{F}_p$. An element $\alpha \in \mathbf{F}_{p^n}$ is said to have *normal θ-depth-2* (or simply *normal depth-2* once θ is fixed) if α and $\alpha - \theta$ are both normal. The set of elements of normal depth-2 in $\mathbf{F}_{p^n}$ is denoted N_2 throughout this paper.

We note at the outset that in [1] it is incorrectly stated that normal depth-2 involves the normality of α and $\alpha - 1$. The correct objects are α and $\alpha - \theta$. See [4], which includes a helpful characterization of normal elements.

The focus in [1] and in this paper is the question of to what extent normal bases of $\mathbf{F}_{p^n}$ over $\mathbf{F}_p$ are preserved by this "depth operation"; that is, if α has normal depth-2, do all of α's conjugates also have normal depth-2? The answer, in general, is that some bases are preserved in N_2 and some are not. To study this question, we make the following definition.

Definition 1.2. The element α of normal depth-2 is called *lonely* if not all of its conjugates also have normal depth-2.

In [1] and here, we endeavor to count, for given p and n, the number of lonely elements in $\mathbf{F}_{p^n}$, which will then yield the number of preserved bases in N_2 once we make

Acknowledgement: The author wishes to thank David Thomson and John Sheekey for their careful reading of and helpful observations on [1], without which this paper would not have been possible. He also thanks the two anonymous reviewers for their excellent ideas on improving the manuscript.

Gove Effinger, Skidmore College, Saratoga Springs, NY, USA, e-mail: effinger@skidmore.edu

https://doi.org/10.1515/9783110621730-003

use of the function defined below, which was first defined more generally in Section 3 of [1]. This is an extension of the polynomial Euler Φ-function which counts polynomials over a given finite field of degree less than the argument polynomial which are relatively prime to that polynomial. The crucial connection to our work here is contained in Theorem 2.39 of [3], namely that $\Phi(t^n - 1)$ counts the number of normal elements of $\mathbf{F}_{p^n}$.

Definition 1.3. Suppose the polynomial f of positive degree is in $\mathbf{F}_p[t]$. The *extended polynomial Euler function* $\Phi_2(f)$ is the number of polynomials A of degree less than the degree of f such that $\gcd(A, f) = 1$ and $\gcd(A - 1, f) = 1$.

It is shown in Section 4 of [1] that the number of elements in N_2 is counted by $\Phi_2(t^n - 1)$. This then tells us that the number of normal bases preserved in N_2 by the depth operation will always be $(\Phi_2(t^n - 1) - \text{number of lonely elements})/n$. We remark also that though the status of specific elements will depend upon our choice of the fixed element θ, the *counts* of these elements will not be affected.

In [1] and [2], specific counts of preserved bases in N_2 are computed (as functions of p) for all quadratic and cubic extensions. In this paper, we obtain analogous counts for all quartic extensions $\mathbf{F}_{p^4}$ of $\mathbf{F}_p$. Some data for this $n = 4$ case is as follows. We saw in previous work and will see here that the factorization of the polynomial $t^n - 1$ plays a crucial role.

n	p	Factorization of $t^n - 1$	Order(N_2) = $\Phi_2(t^n - 1)$	Normal bases in N_2	Lonely elements
4	3	$(t-1)(t+1)(t^2+1)$	7	0	7
	5	$(t-1)(t+1)(t+2)(t+3)$	81	0	81
	7	$(t-1)(t+1)(t^2+1)$	1175	220	295
	11	$(t-1)(t+1)(t^2+1)$	9639	2088	1287
	13	$(t-1)(t+1)(t+5)(t+8)$	14,641	1760	7601
	17	$(t-1)(t+1)(t+4)(t+13)$	50,625	7560	20,385
	19	$(t-1)(t+1)(t^2+1)$	103,751	24,208	6909

In [1], there is no analysis of quartic extensions. The main analysis is of combinations of n and p for which $t^n - 1$ factors into $t - 1$ and an irreducible factor of degree $n - 1$, which clearly implies that n is prime. In particular, cubic extensions with p of the form $3k + 2$ are of this form, but the case of $3k + 1$ is more complicated and is analyzed separately in [2] and [4]. Hence the situation is fully understood for all quadratic and cubic extensions of $\mathbf{F}_p$ for all primes p. Here, we fully analyze quartic extensions, where the analyses are significantly different depending on whether p is of the form $4k + 1$ or $4k + 3$. The latter case is somewhat easier (since $t^2 + 1$ is then irreducible over $\mathbf{F}_p$), so we deal with that case first.

In all that follows, θ is a fixed normal element of $\mathbf{F}_{p^n}$, and for all elements $\alpha \in \mathbf{F}_{p^n}$, we write α and $\alpha - \theta$ in terms of the normal basis generated by θ (i. e.,

$$\alpha = a_{n-1}\theta^{p^{n-1}} + a_{n-2}\theta^{p^{n-2}} + \cdots + a_1\theta^p + a_0\theta$$

and

$$\alpha - \theta = a_{n-1}\theta^{p^{n-1}} + a_{n-2}\theta^{p^{n-2}} + \cdots + a_1\theta^p + (a_0 - 1)\theta).$$

Then α and $\alpha - \theta$ are in one-to-one correspondence with the "conventional" polynomials $\overline{\alpha} = a_{n-1}t^{n-1} + a_{n-2}t^{n-2} + \cdots + a_1 t + a_0$ and $\overline{\alpha} - 1 = a_{n-1}t^{n-1} + a_{n-2}t^{n-2} + \cdots + a_1 t + (a_0 - 1)$. That is, α is an element of $\mathbf{F}_{p^n}$; $\overline{\alpha}$ is a polynomial of degree less than n in the variable t. All of the following analyses involve both field elements such as α, β and conj 1 and the corresponding polynomials $\overline{\alpha}$, $\overline{\beta}$, $\overline{\text{conj}\,1}$.

2 The case $p = 4k + 3$

Suppose $p = 4k + 3$ for some k. By quadratic residue theory, we know that $t^2 + 1$ is irreducible over $\mathbf{F}_p$, and so $t^4 - 1$ has as its prime factorization $(t-1)(t+1)(t^2+1)$. Using long division, it is easy to observe that the following criteria hold for the polynomial $\overline{\alpha} = at^3 + bt^2 + ct + d$ over $\mathbf{F}_p$:

1. $\overline{\alpha}$ is divisible by $t - 1$ if and only if $a + b + c + d = 0$.
2. $\overline{\alpha}$ is divisible by $t + 1$ if and only if $a - b + c - d = 0$.
3. $\overline{\alpha}$ is divisible by $t^2 + 1$ if and only if $a = c$ and $b = d$.

We start here with elements α of $\mathbf{F}_{p^4}$, represented as polynomials $\overline{\alpha}$ of degree less than 4 over $\mathbf{F}_p$, which are normal, i. e., $\overline{\alpha}$ are relatively prime to $t^4 - 1$. The set of all normal elements in $\mathbf{F}_{p^4}$ is denoted by N and has $\Phi(t^4 - 1)$ elements. Among these, elements α which have the property that *both* $\overline{\alpha}$ and $\overline{\alpha} - 1$ are relatively prime to $t^4 - 1$ form the set of elements of normal depth-2, which set is denoted N_2 and has order $\Phi_2(t^4 - 1)$. We know that by definition if β is a p-conjugate of $\alpha \in N$, then $\overline{\beta}$ is also relatively prime $t^4 - 1$, but it may well be that $\overline{\beta} - 1$ is *not* relatively prime to $t^4 - 1$, and hence α will be lonely. Our strategy here (a sort of "backdoor" approach) for identifying lonely elements will be to look at normal elements β which have the property that $\overline{\beta} - 1$ is divisible by $t + 1$, or by $t^2 + 1$, or by both $t + 1$ and $t^2 + 1$. For each set and each such β, we will seek and count p-conjugates of these elements whose corresponding polynomials are divisible by *none* of the three factors of $t^4 - 1$, i. e., which are in the set N_2. Such elements will then be lonely. In order to implement the approach just described, we need to count the various elements β which are in N but not in N_2 which could then possibly have p-conjugates which lie in N_2. For ease of notation, we shall denote $\overline{\beta}$ by (a, b, c, d).

We shall throughout make use of the key fact that the p-conjugates of a normal element β are the first conjugate "conj 1," whose corresponding polynomial $\overline{\text{conj}\,1}$ is (b, c, d, a), second conjugate "conj 2," with $\overline{\text{conj}\,2} = (c, d, a, b)$, and third conjugate "conj 3," with $\overline{\text{conj}\,3} = (d, a, b, c)$; that is, we obtain the p-conjugates of normal elements by simply rotating their coefficients. (Again, see [3], Chapter 2, Section 3.) We start by observing that if β is in N but $\overline{\beta} - 1$ is divisible by $t - 1$, i. e., if the sum of $\overline{\beta}$'s

coefficients is 1, then this will automatically be true as well for all of its conjugates, so no lonely elements can be generated from this set of elements. The point is that the sum of coefficients is constant under rotation.

We shall now focus on three subsets of $N - N_2$ which are disjoint from the set just mentioned (elements β for which $\overline{\beta} - 1$ is divisible by $t - 1$) and are disjoint from each other.

1. S_1 is the set of elements $\beta \in N$ which have the property that $\overline{\beta} - 1$ is divisible by $t + 1$ but not by $t - 1$ and $t^2 + 1$. The order of S_1 is $\Phi_2((t-1)(t^2+1)) = (p-2)(p^2-2)$.
2. S_2 is the set of elements $\beta \in N$ which have the property that $\overline{\beta} - 1$ is divisible by $t^2 + 1$ but not by $t - 1$ and $t + 1$. The order of S_2 is $\Phi_2((t-1)(t+1)) = (p-2)^2$.
3. S_3 is the set of elements $\beta \in N$ which have the property that $\overline{\beta} - 1$ is divisible by $t + 1$ and $t^2 + 1$ but not by $t - 1$. The order of S_3 is $\Phi_2(t-1) = p - 2$.

Just double-checking these orders, on the one hand the order of $N - N_2$ is

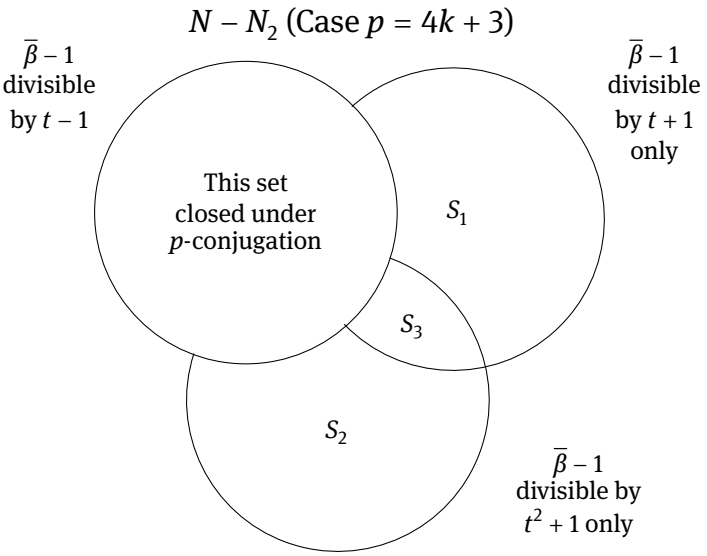
$$\Phi(t^4 - 1) - \Phi_2(t^4 - 1) = (p-1)^2(p^2-1) - (p-2)^2(p^2-2) = 2p^3 - 2p^2 - 6p + 7.$$

On the other hand, the set of $\beta \in N$ such that $t-1$ divides $\overline{\beta}-1$ (with no other conditions) has $(p-1)^2(p^2-1)/(p-1)$ elements, and so counting that set and the sets S_1, S_2, and S_3, which are all pairwise disjoint, we get

$$\begin{aligned}&(p-1)(p^2-1) + (p-2)(p^2-2) + (p-2)^2 + (p-2)\\ &\quad = (p^3 - p^2 - p + 1) + (p^3 - 2p^2 - 2p + 4) + (p^2 - 4p + 4) + (p-2) = 2p^3 - 2p^2 - 6p + 7,\end{aligned}$$

as desired.

Here then is a Venn diagram of the set $N - N_2$:



We observe now that if β is in the set S_1, then $\overline{\beta} - 1$ is divisible by $t + 1$, i. e., $a - b + c - d = 1$. It follows that $\overline{\text{conj } 2}$ (i. e., (c, d, a, b)) will also have an alternating sum of 1, *but* $\overline{\text{conj } 1}$ and $\overline{\text{conj } 3}$ will have an alternating sum of –1 and so their corresponding elements could lie in N_2.

A similar observation holds for the set S_2 of elements β for which $\overline{\beta} - 1$ is divisible by $t^2 + 1$, which means that $\overline{\beta}$ is of the form $(a, b, a, b + 1)$. Here, we see that none of the conjugate polynomials $(b, a, b + 1, a), (a, b + 1, a, b), (b + 1, a, b, a)$ satisfy that the first and third coefficients are equal *and* the fourth coefficient is 1 greater than the second coefficient, and so all three conjugates are potential lonely elements.

Finally, in the set S_3, we know that $\overline{\beta}$ must be of the form $(a, b, a, b + 1)$ and we must have $a - b + a - b - 1 = 1$, i. e., $2b = 2a - 2$, so $b = a - 1$ and so $\overline{\beta}$ is of the form $(a, a - 1, a, a)$. None of the conjugate polynomials are divisible by $t^2 + 1$, and $\overline{\text{conj } 1}$ and $\overline{\text{conj } 3}$ have an alternating sum of –1, and so could be lonely.

So now it is time to count our lonely elements. We must observe that even though S_1, S_2, and S_3 are pairwise disjoint, any element in them could have as many as three conjugates which are in N_2, so we must be on the lookout for overlaps among these as we move from set to set.

We start by counting the number of lonely elements which arise from the p-conjugates of elements of S_1. Recall that S_1 has $(p-2)(p^2-2)$ elements β. It may be helpful to look at the simplest case available to us, namely the case $p = 3$, for which S_1 has 7 elements. We shall hereafter refer to p-conjugates as simply conjugates.

$\overline{\beta}$	$\overline{\text{conj } 1}$	$\overline{\text{conj } 2}$	$\overline{\text{conj } 3}$
(0, 0, 0, 2)	(0, 0, 2, 0)	(0, 2, 0, 0)	**(2, 0, 0, 0)**
(1, 0, 2, 2)	(0, 2, 2, 1)	(2, 2, 1, 0)	(2, 1, 0, 2)
(1, 1, 2, 1)	(1, 2, 1, 1)	(2, 1, 1, 1)	(1, 1, 1, 2)
(1, 2, 2, 0)	(2, 2, 0, 1)	(2, 0, 1, 2)	(0, 1, 2, 2)
(2, 0, 1, 2)	(0, 1, 2, 2)	(1, 2, 2, 0)	(2, 2, 0, 1)
(2, 1, 1, 1)	**(1, 1, 1, 2)**	(1, 1, 2, 1)	(1, 2, 1, 1)
(2, 2, 1, 0)	(2, 1, 0, 2)	(1, 0, 2, 2)	(0, 2, 2, 1)

How many lonely elements are represented in this chart? As observed previously, since $\overline{\beta}$ has an alternating sum of 1, both $\overline{\text{conj } 1}$ and $\overline{\text{conj } 3}$ will have alternating sums of –1, and hence are potential lonely elements. Focusing on the set $\overline{\text{conj } 1}$, we note that only one of them (namely (1, 1, 1, 2)) is not lonely since its element lies in S_2 (i. e., its first and third coefficients are equal and its fourth coefficient is 1 greater than its second coefficient), and hence its element is not lonely (i. e., is not in N_2). Hence the set $\overline{\text{conj } 1}$ represents 6 lonely elements. Focusing now on the set $\overline{\text{conj } 3}$, we see that all of them except (2, 0, 0, 0) have already appeared as first conjugates, and hence should not be counted. However, we *should* count (2, 0, 0, 0). (The reason it does not appear as a first conjugate is that it is itself the first conjugate of (0, 2, 0, 0), whose element lies in S_2 and not in S_1.) We highlight these two polynomials in our chart, one

of which ((1, 1, 1, 2)) must be removed from our count and one ((2, 0, 0, 0)) which must be added. Hence we arrive at exactly 7 lonely elements generated from S_1 in this case. Note that with the addition and subtraction of one element each, we arrived at the count of lonely elements from S_1 being exactly the order of S_1.

Let us now generalize the above argument to the S_1 count of lonely elements for an arbitrary prime p satisfying $p = 4k+3$ for some k. Again, we have $(p-2)(p^2-2)$ elements β in S_1, and the alternating sum of each of their first and third conjugate polynomials is -1, so these are potential lonely elements. However, we must first throw out any first conjugates which lie in S_2, i. e., for which in $\overline{\text{conj}\,1}\ d = b$ and $c = a - 1$, so we are looking for polynomials of the form $(b, a-1, b, a)$. But we know that the alternating sum $2b - 2a + 1$ is -1, so $2b = 2a - 2$, i. e., $b = a - 1$, so our elements are of the form $(a-1, a-1, a-1, a)$ (just like (1, 1, 1, 2) when $p = 3$). Finally, we know that the sum of the coefficients cannot be 0 or 1 (since neither $\overline{\beta}$ nor $\overline{\beta} - 1$ is divisible by $t - 1$), so $4a - 3 \neq \{0, 1\}$, so $a \neq (4^{-1})\{4, 3\}$, i. e., a cannot be 1 or $(4^{-1})(3)$, and we see that we have $p - 2$ choices for a. Hence we must remove $p - 2$ elements from the set of first conjugates.

We now must ask how many third conjugates have the property that they are not the same as some first conjugate. This will occur when conj 2 lies in S_2, and an argument exactly parallel to the one just given show that there are $p - 2$ such second conjugates. Hence the corresponding third conjugates do not appear as first conjugates, and so they must be added to our count. We conclude then that the set S_1 generates

$$(p-2)(p^2-2) - (p-2) + (p-2) = (p-2)(p^2-2)$$

lonely elements.

We now turn to the set S_2, which has $(p-2)^2$ elements β for which $\overline{\beta}$ is of the form $(a, b, a, b+1)$. We note that all three conjugate polynomials $(b, a, b+1, a)$, $(a, b+1, a, b)$ and $(b+1, a, b, a)$ are not elements which lie in S_2, and hence are potential lonely elements. $\overline{\beta}$'s alternating sum cannot be 0 or 1, but it can be -1, in which case its first and third conjugate polynomials have an alternating sum of 1, and hence their elements cannot be lonely, but its second conjugate will be lonely since its in neither S_1 nor S_2. It is easy again to show that such a second conjugate polynomial is of the form $(a-1, a, a-1, a-1)$ for $p - 2$ choices of a. All other polynomials will have an alternating sum which is none of $\{0, 1, -1\}$, so all three of their conjugates are lonely. Now the $(p-2)^2$ elements of S_2 are evenly divided among the $p - 2$ possible alternating sum values (not 0 or 1), and we divide the polynomials into those for which the alternating sum is -1 (one lonely per each element) and those for which the alternating is none of $\{0, 1, -1\}$ (3 lonely per element), giving us a count of lonely elements generated to be $(p-2) + 3(p-3)(p-2) = (p-2)(3p-8)$.

However, it may not be the case that all of these elements are *new* lonely elements, i. e., they may have already been discovered as arising from S_1. This is indeed the case. The $p-2$ polynomials $(a-1, a, a-1, a-1)$ with alternating sum -1, pointed out above,

are the first conjugates of polynomials of the form $(a-1, a-1, a, a-1)$ whose alternating sum is 1, so the corresponding elements are in S_1 and the polynomials $(a-1, a, a-1, a-1)$ *have already been counted*. No other lonely elements arising from S_2 can have already been counted since their alternating sums are none of 0, 1 or –1. Hence we must remove this overlap of $p-2$ elements, and our final count for *new* lonely elements arising from S_2 is

$$(p-2)(3p-8) - (p-2) = (p-2)(3p-9) = 3(p-2)(p-3).$$

Finally, we need to seek lonely elements arising from S_3. There are only $p-2$ elements in S_3, all of whose polynomials are of the form $(a, a-1, a, a)$ (i. e., with an alternating sum of 1). Note that $\overline{\text{conj } 2}$ is $(a, a, a, a-1)$, whose element lies in S_1, so $\overline{\text{conj } 3}$ $(a, a, a-1, a)$ represents a lonely element but has already been counted. Now the first conjugate $(a-1, a, a, a)$ is also lonely, but it has also already been counted in our analysis of S_1, as these are precisely those elements there which occurred as third conjugates for elements in S_1 whose second conjugate (with polynomial $(a, a-1, a, a)$) lies in S_2. Hence there are no *new* lonely elements arising from S_3.

We have arrived at the following.

Theorem 2.1. *Suppose that p is of the form* $4k+3$ *for some* k. *Among the* $\Phi_2(t^4-1) = (p-2)^2(p^2-2)$ *elements of* $\mathbf{F}_{p^4}$ *which are of normal depth-2, there are* $(p-2)(p^2+3p-11)$ *lonely elements and hence* $(p-2)(p-3)(p^2-5)/4$ *normal bases.*

Proof. Adding up the lonely elements from S_1, S_2, and S_3, we get a total of

$$(p-2)(p^2-2) + 3(p-2)(p-3) + 0 = (p-2)(p^2+3p-11).$$

The number of normal bases then will be

$$\begin{aligned}&\big((p-2)^2(p^2-2) - (p-2)(p^2+3p-11)\big)/4\\ &\quad = (p-2)(p^3-3p^2-5p+15)/4 = (p-2)(p-3)(p^2-5)/4\end{aligned}$$

which completes the proof. □

These counts are confirmed by Mathematica© for the cases $p = 3, 7, 11$, and 19, as displayed at the opening of this paper.

3 The case $p = 4k+1$

Suppose now that $p = 4k+1$ for some k. Again by quadratic reciprocity, we know that t^4-1 factors into the four linear polynomials $t-1$, $t+1$, $t-\iota$, and $t+\iota$, where ι is a square root of –1 in $\mathbf{F}_p$. In this case, the set N of normal elements of $\mathbf{F}_{p^4}$ has $\Phi(t^4-1) = (p-1)^4$ elements and the set N_2 of normal elements of depth-2 has $\Phi_2(t^4-1) = (p-2)^4$ elements.

Just as in the previous section, we shall denote by $\overline{\alpha}$ the polynomial of degree less than 4 which corresponds to the expression of the element $\alpha \in \mathbf{F}_{p^4}$ in terms of the normal basis generated by the fixed element θ. In the arguments below, if the context is clear, we may abuse language a bit by saying, for example, "the polynomial $\overline{\alpha} = (a, b, c, d)$ is lonely," meaning, more precisely, "the element α is lonely," and so on.

As before, using long division, it is easy to observe that the following criteria hold for the polynomial $\overline{\alpha} = at^3 + bt^2 + ct + d$ over $\mathbf{F}_p$:

1. $\overline{\alpha}$ is divisible by $t - 1$ if and only if $a + b + c + d = 0$.
2. $\overline{\alpha}$ is divisible by $t + 1$ if and only if $a - b + c - d = 0$.
3. $\overline{\alpha}$ is divisible by $t - \iota$ if and only if $(d - b) + (c - a)\iota = 0$.
4. $\overline{\alpha}$ is divisible by $t + \iota$ if and only if $(d - b) - (c - a)\iota = 0$.

In the latter two conditions, we use the fact that $\iota^2 = -1$. We now observe that either or both of Conditions 3 and 4 are true if and only if

$$0 = \big((d - b) + (c - a)\iota\big)\big((d - b) - (c - a)\iota\big) = (d - b)^2 + (c - a)^2 \pmod{p}.$$

Hence we have a single criterion we can use to identify what polynomials are divisible by $t - \iota$ or $t + \iota$ (or both), and so we can analyze this case (i. e., $p = 4k + 1$) in a similar fashion as we did with the case $p = 4k + 3$, though the criterion is a bit more complicated in this case (and, by the way, will lead to many more lonely elements).

We shall again throughout make use of the key fact that the polynomials of the p-conjugates of a normal element β whose polynomial is $\overline{\beta} = (a, b, c, d)$ are $\overline{\mathrm{conj}\,1} = (b, c, d, a)$, $\overline{\mathrm{conj}\,2} = (c, d, a, b)$, and $\overline{\mathrm{conj}\,3} = (d, a, b, c)$. We start by observing that if β is in N but $\overline{\beta} - 1$ is divisible by $t - 1$, i. e., if the sum of $\overline{\beta}$'s coefficients 1, then this will automatically be true as well for all of its conjugates, so no lonely elements can be generated from this set of elements. The point, as before, is that the sum of coefficients is constant under rotation.

Continuing, we shall again focus on three subsets of $N - N_2$ which are disjoint from the set just mentioned (elements $\beta \in N$ for which $\overline{\beta} - 1$ is divisible by $t - 1$ with no other conditions) and are disjoint from each other.

1. T_1 is the set of elements $\beta \in N$ which have the property that $\overline{\beta} - 1$ is divisible by $t + 1$ but not by $t - 1$, $t - \iota$ and $t + \iota$. The order of T_1 is

$$\Phi_2\big((t - 1)(t - \iota)(t + \iota)\big) = (p - 2)^3.$$

2. T_2 is the set of elements β which have the property that $\overline{\beta} - 1$ is divisible by $t - \iota$ or $t + \iota$ (or both) but not by $t - 1$ and $t + 1$. The order of T_2 is

$$\begin{aligned}&\Phi_2\big((t - 1)(t + 1)(t + \iota)\big) + \Phi_2\big((t - 1)(t + 1)(t - \iota)\big) + \Phi_2\big((t - 1)(t + 1)\big)\\ &\quad = 2(p - 2)^3 + (p - 2)^2 = (p - 2)^2(2p - 3).\end{aligned}$$

3. T_3 is the set of elements $\beta \in N$ which have the property that $\overline{\beta} - 1$ is divisible by $t+1$ and $t-\iota$ or $t+\iota$ (or both), but not by $t-1$. The order of S_3 is

$$\begin{aligned}&\Phi_2\big((t-1)(t+\iota)\big) + \Phi_2\big((t-1)(t-\iota)\big) + \Phi_2(t-1)\\ &\quad = 2(p-2)^2 + (p-2) = (p-2)(2p-3).\end{aligned}$$

Double-checking these orders, on the one hand the order of $N - N_2$ is

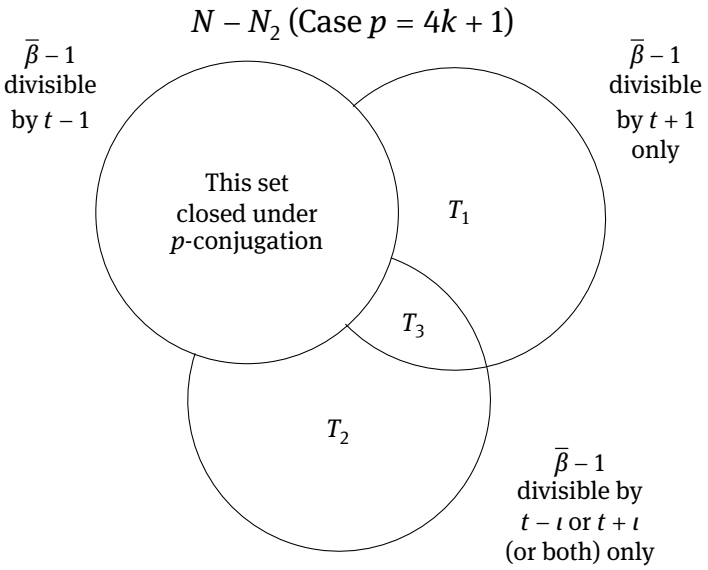
$$\Phi(t^4-1) - \Phi_2(t^4-1) = (p-1)^4 - (p-2)^4 = 4p^3 - 18p^2 + 28p - 15.$$

On the other hand, the set of $\beta \in N$ such that $t-1$ divides $\overline{\beta}-1$ with no other conditions has $(p-1)^4/(p-1) = (p-1)^3$ elements, and so counting that set and the sets T_1, T_2, and T_3, which are all pairwise disjoint, we get

$$\begin{aligned}&(p-1)^3 + (p-2)^3 + (p-2)^2(2p-3) + (p-2)(2p-3)\\ &\quad = (p^3 - 3p^2 + 3p - 1) + (p^3 - 6p^2 + 12p - 8) + (2p^3 - 11p^2 + 20p - 12) + (2p^2 - 7p + 6)\\ &\quad = 4p^3 - 18p^2 + 28p - 15,\end{aligned}$$

as desired.

The Venn diagram for this case is as follows:



We now embark on counting the number of lonely elements which arise among the conjugates of elements β lying in our three sets T_1, T_2, and T_3. We remark, as before, that since any element β in any one of these sets can give rise to as many as three lonely elements among its conjugates, we must be on the lookout for overlaps.

Concerning "overlaps," there are in fact three types of these in the analysis below:

a. For a given set T_1, T_2, or T_3 and for a given conjugate type conj 1, conj 2, or conj 3 associated with that set, there can be overlaps of elements satisfying the various necessary conditions;
b. for a given set T_1, T_2, or T_3, there can be overlaps across the three conjugate types, and
c. there can be overlaps of lonely conjugate elements across the three sets T_1, T_2, and T_3.

As before, we shall denote by ι a fixed square root of -1 on $\mathbf{F}_p$.

We shall employ the following notation throughout the analysis below: if $\overline{\beta} = (a, b, c, d)$ is the polynomial corresponding to an element of one of the three sets T_1, T_2, or T_3, we set

$$X = c - a \quad \text{and} \quad Y = d - b.$$

Using this notation, we have

1. $\overline{\beta} - 1$ is divisible by $t - \iota$ or $t + \iota$ (or both) if and only of $X^2 + (Y-1)^2 = 0$;
2. For $\overline{\text{conj } 1} = (b, c, d, a)$, $\overline{\text{conj } 1} - 1$ is so if and only if $Y^2 + (X+1)^2 = 0$;
3. For $\overline{\text{conj } 2} = (c, d, a, b)$, $\overline{\text{conj } 2} - 1$ is so if and only if $X^2 + (Y+1)^2 = 0$, and
4. For $\overline{\text{conj } 3} = (d, a, b, c)$, $\overline{\text{conj } 3} - 1$ is so if and only if $Y^2 + (X-1)^2 = 0$.

Moreover, as polynomials $\overline{\beta}$ range over those with a fixed sum $a + b + c + d = v$ and a fixed alternating sum $a - b + c - d = u$ (called the "alt sum" below), by adding we get $2a + 2c = v + u$, so $c - a = 2^{-1}(v + u) + 2a$, and we see that as a ranges over $\mathbf{F}_p$, we get p possible values of X. By subtracting, we get the same conclusion for Y, so for these fixed values v and u, we get p^2 ordered pairs (X, Y).

3.1 Counting lonely elements among the conjugates of elements β in T_1

So we first turn to the set T_1 of elements $\beta \in N$ for which $\overline{\beta} - 1$ is divisible by $t+1$ only. We fix the sum to a value $v \neq 0$ or 1. Since $\overline{\beta}$'s alt sum is 1, all second conjugates must also have an alt sum of 1, so no lonely elements can arise from this set. However, all first and third conjugates have an alt sum of -1, so lonely elements *can* arise from these two sets, and we need to count them and count overlaps. By definition of T_1, we know that for $\beta \in N_1$, $X^2 + Y^2 \neq 0$ and $X^2 + (Y-1)^2 \neq 0$. In order for conj 1 to be lonely, we must also have $Y^2 + (X+1)^2 \neq 0$, so we seek the intersection arising from three inequalities. It will be easier then to compute the union of three *equalities* and subtract from the total number p^2 of pairs (X, Y).

We first count the single sets. If $X^2 + Y^2 = 0$, then $Y = \pm \iota X$, so we get two pairs for each nonzero X and one more when $X = 0$ for a total of $2(p-1)+1 = 2p-1$ pairs. Similar

arguments obtain the same counts for the single sets satisfying $X^2 + (Y-1)^2 = 0$ and $Y^2 + (X+1)^2 = 0$.

Looking at double overlaps now, if $X^2+Y^2 = 0$ and $X^2+(Y-1)^2 = 0$, then subtracting we obtain $2Y - 1 = 0$, i. e., $Y = 2^{-1}$, and plugging this into the first condition gives $X^2 = -4^{-1}$, so $X = \pm\iota 2^{-1}$. Hence we get two pairs in this overlap, and a similar argument gives two pairs in the overlap of $X^2 + Y^2 = 0$ and $Y^2 + (X+1)^2 = 0$. The final double overlap is a bit different but with the same outcome. Combining $X^2 + (Y-1)^2 = 0$ and $Y^2 + (X+1)^2 = 0$, we get $2X + 1 - (-2Y + 1) = 0$, so $Y = -X$, which gives $2X^2 + 2X + 1 = 0$, and we arrive at $X = 4^{-1}(-2 \pm \sqrt{4-8}) = -2^{-1} \pm 2\iota$, and we see our two pairs. It is quick to check that the 6 ordered pairs found are distinct, so the triple overlap is empty. Hence we have that the complement of the set we seek has $3(2p-1) - 6 = 6p - 9$ pairs, and so the count of conj 1 lonely elements for the fixed sum of v is $p^2 - 6p + 9 = (p-3)^2$. Now freeing up the $p - 2$ allowable sum values, the total count of conj 1 lonely elements arising from elements of T_1 is $(p-2)(p-3)^2$.

A parallel argument will give the same count for conj 3 elements.

In order to finish our T_1 analysis, we need to count the overlap of lonely conj 1 and conj 3 elements. This is accomplished by doing inclusion/exclusion counting the complement, now including the fourth condition $(X-1)^2 + Y^2 = 0$, which again alone will yield $2p - 1$ pairs. Now we have $\binom{4}{2} = 6$ double overlaps, 3 of which are computed above and 3 more of which again yield two pairs each. Moreover, all 12 of those pairs are distinct, so the 4 triple overlaps and single quadruple overlap are all empty. We arrive then at a complement count of $4(2p-1) - 12 = 8p - 16$, so the lonely overlap count for a fixed sum is $p^2 - 8p + 16 = (p-4)^2$, and freeing the sum we get a total overlap count of $(p-2)(p-4)^2$. We conclude then that the total number of lonely elements arising from conjugates of element of T_1 is

$$(p-2)\big(2(p-3)^2 - (p-4)^2\big) = (p-2)(p^2 - 4p + 2).$$

This count is confirmed by Mathematica© for $p = 5$, 13, and 17.

3.2 Counting lonely elements among the conjugates of elements β in T_2

The analysis of the set T_2 is similar to that of T_1 but is more complicated, especially in counting overlaps, because now lonely elements can occur among all three of conj 1, conj 2, and conj 3. We again fix the coefficient sum of $\overline{\beta}$ for $\beta \in T_2$ to be $v \neq \{0, 1\}$. Since the alt sum of $\overline{\beta}$ also cannot be 0 or 1, conj 2 can as well have $p - 2$ alt sums (matching $\overline{\beta}$'s). However, if $\overline{\beta}$'s alt sum is –1, then the corresponding conj 1 and conj 3 polynomials will have an alt sum of 1, and hence are not lonely. Thus we can only allow $p - 3$ alt sums with those two sets.

We do a careful analysis of the set $\overline{\text{conj}\,1}$. Here, we assume a fixed sum $v \neq \{0,1\}$ and a fixed alt sum $w \neq \{0,1,-1\}$. The three conditions to be met in order for a polynomial (b,c,d,a) to be lonely are $X^2+Y^2 \neq 0$, $X^2+(Y-1)^2 = 0$, and $(X+1)^2+Y^2 \neq 0$. We choose again to use inclusion-exclusion on the complement in the set of p^2 ordered pairs (X,Y). Hence our conditions now are $X^2+Y^2 = 0$, $X^2+(Y-1)^2 \neq 0$, or $(X+1)^2+Y^2 = 0$.

Counting the single sets, the first and third are just as in the T_1 case, so their counts are each $2p-1$. The middle condition count is $p^2-2p+1 = (p-1)^2$.

Looking at double overlaps, for $X^2+Y^2 = 0$ and $(X+1)^2+Y^2 = 0$, again just as above, we get 2 pairs, which are in fact $(-2^{-1}, \iota 2^{-1})$ and $(-2^{-1}, -\iota 2^{-1})$. The other two double overlaps are different since they involve an inequality. We first consider $X^2+Y^2 = 0$ and $X^2+(Y-1)^2 = u \neq 0$. Combining and simplifying, we get $Y = 2^{-1}(1-u)$, and substituting back we get $X = \pm 2^{-1}(1-u)$. Hence if $u \neq 1$ we get $2(p-2)$ pairs, and when $u = 1$, we get just one, for a total of $2p-3$ pairs.

The final double overlap is for the conditions $(X+1)^2+Y^2 = 0$ and $X^2+(Y-1)^2 = u \neq 0$. Combining, we get $Y = -X-2^{-1}u$, and substituting into the first and using the quadratic formula, we get $X = 4^{-1}(-2-u \pm \iota(u-2))$. Hence if $u \neq 2$, we again get $2(p-2)$ pairs, and if $u = 2$, we get one, for a total of $2p-3$ pairs.

Finally, for the triple overlap, we put the two pairs in the first double overlap above into $X^2+(Y-1)$ to see if we get 0 or not. We do the first such pair here: $(-2^{-1})^2+(\iota 2^{-1}-1)^2 = 4^{-1}-4^{-1}-\iota+1 = 1-\iota \neq 0$. Hence this element (and similarly the other) satisfies all three conditions, so the triple overlap has a count of 2.

Our complement count in $\overline{\text{conj}\,1}$ is then $2(2p-1)+(p-1)^2-2-2(2p-3)+2 = p^2-2p+5$. We conclude that the count of lonely elements in conj 1 for fixed sum v and fixed alt sum w is $p^2-(p^2-2p+5) = 2p-5$. Now freeing up the sum and alt sum, we get a final count of

$$(p-2)(p-3)(2p-5).$$

Parallel arguments, which we will not repeat here, give the same count for $\overline{\text{conj}\,3}$; but for $\overline{\text{conj}\,2}$, as noted at the outset of this part, the total count will be

$$(p-2)^2(2p-5).$$

We must now identify any elements which are lonely but occur in two or all three of the conjugate types, since such elements must be counted only once. We begin with possible $\overline{\text{conj}\,1}$ and $\overline{\text{conj}\,3}$ overlaps. If $\overline{\text{conj}\,1} = (b,c,d,a)$ appears as a $\overline{\text{conj}\,3}$ as well, then it arises from $\overline{\beta} = (c,d,a,b)$, which is in N_2, so both of the conditions $X^2+(Y-1)^2 = 0$ and $X^2+(Y+1)^2 = 0$ must hold, and combining we get that $Y = 0$ and $X = \pm\iota$. Hence we must have $d = b$ and $c = a \pm \iota$.

We continue to assume that we are working with T_2 elements with a fixed sum of v and fixed alt sum of w. First, suppose $c = a+\iota$. Hence $a+b+(a+\iota)+b = v$ and $a-b+(a+\iota)-b = w$, and adding and subtracting we get $a = 4^{-1}(v+w-2\iota)$ and

$b = 4^{-1}(v - w)$. Similarly, if $c = a - \iota$, we get $a = 4^{-1}(v + w + 2\iota)$ and $b = 4^{-1}(v - w)$. Hence we get exactly two overlap elements, and we must finally confirm that they are lonely. But $(X+1)^2 + Y^2 = (\pm\iota + 1)^2 = \pm 2\iota$, and likewise $(X-1)^2 + Y^2 = (\pm\iota - 1)^2 = \pm 2\iota$, so neither expression equals 0, confirming that these two elements are indeed lonely.

We now move to possible overlaps among lonely conj 1 and conj 2 elements. This case is trickier since such an element must arise from elements of T_2 whose alt sums are negatives of each other. That is, if γ, with alt sum $-w$, is both a lonely conj 1 and conj 2 for T_2, then $\overline{\gamma}$ arises from a $\overline{\beta}_1$ whose alt sum is w and from a $\overline{\beta}_2$ whose alt sum is $-w$. So suppose $\overline{\beta}_1 = (a_1, b_1, c_1, d_1)$ and $\overline{\beta}_2 = (a_2, b_2, c_2, d_2)$, then $\overline{\gamma} = (b_1, c_1, d_1, a_1) = (c_2, d_2, a_2, b_2)$, and so we have $a_2 = d_1$, $b_2 = a_1$, $c_2 = b_1$ and $d_2 = c_1$, i. e., $\overline{\beta}_2 = (d_1, a_1, b_1, c_1)$. Since we can now express everything in terms of $\overline{\beta}_1$'s entries, we drop the subscripts, i. e., $\overline{\beta}_1 = (a, b, c, d)$ and $\overline{\beta}_2 = (d, a, b, c)$. Since both β_1 and β_2 are in N_2, we have $X^2 + (Y-1)^2 = 0$ and $(X-1)^2 + Y^2 = 0$, which gives us $Y = X$, and plugging back in, using the quadratic formula, and simplifying we arrive at $Y = X = 2^{-1}(1 \pm \iota)$. This then again will give us two overlap polynomials (which can be written down explicitly using $c = a + X$, etc.). We must finally check that the corresponding elements are indeed lonely, that is, $\overline{\gamma} = (b, c, d, a)$ must satisfy $(X+1)^2 + Y^2 \neq 0$. If $Y = X = 2^{-1}(1 + \iota)$, we have

$$\begin{aligned}(X+1)^2 + Y^2 &= \left(2^{-1}(1+\iota) + 1\right)^2 + \left(2^{-1}(1+\iota)\right)^2 \\ &= 4^{-1}(2\iota) + (1+\iota) + 1 + 4^{-1}(2\iota) = \iota + 1 + \iota + 1 = 2(\iota + 1) \neq 0,\end{aligned}$$

as desired, and the other pair (X, Y) evaluates to $2(-\iota + 1)$, which is also not 0.

We dispense with the parallel conj 2 and conj 3 argument, which again arrives at two elements in the overlap for each fixed allowable sum and alt sum.

Finally, we observe that if an element γ among the lonely elements arising from T_2 were in all three conjugate sets, then simultaneously $Y = X$, $Y = 0$, and $X = \pm\iota$, which is impossible. Hence no such γ exists.

We are now in a position to write down the count of lonely elements arising from T_2. That count is

$$\begin{aligned}&2(p-2)(p-3)(2p-5) + (p-2)^2(2p-5) - 6(p-2)(p-3) \\ &\quad = (p-2)\left(4p^2 - 22p + 30 + 2p^2 - 9p + 10 - 6p + 18\right) = (p-2)\left(6p^2 - 37p + 58\right).\end{aligned}$$

This count is confirmed by Mathematica© for $p = 5$, 13, and 17.

3.3 Counting lonely elements among the conjugates of elements β in T_3

This case is much simpler to work out than the two previous cases. Since T_3 is the intersection of T_1 and T_2, we can count its lonely conjugate elements by using the above T_2 analysis, but now the alt sum is fixed at 1, so

a. no second conjugates are possible since their alt sum is also 1, and
b. there will be no factor of $p-3$ in the counts.

Hence the total count is

$$2(p-2)(2p-5)-2(p-2)=2(p-2)(2p-6)=4(p-2)(p-3).$$

This count is confirmed by Mathematica© for $p=5$, 13, and 17.

3.4 Counting the overlaps of lonely elements across T_1, T_2, and T_3

We again do the analyses assuming a fixed sum $v\neq\{0,1\}$, and hence must multiply by $p-2$ to get total counts. We continue to use the notation $X=c-a$ and $Y=d-b$.

We start with possible overlaps of lonely conjugate elements arising from T_1 and T_2. If γ_1 is a conj 1 or conj 3 lonely element arising from T_2, then its alt sum w cannot be –1, but to be in T_1 its alt sum *must* be –1. Hence overlaps can only be lonely conj 2 elements $\overline{\gamma}_2=(c,d,a,b)$ from T_2 whose alt sum is –1, for then $\overline{\gamma}_2$ could occur as a $\overline{\text{conj 1}}$ of (b,c,d,a) or as a $\overline{\text{conj 2}}$ of (d,a,b,c). Both of these elements have an alt sum of 1 and so are candidates for being in T_1. We claim that at least one of them *is* in T_1. Suppose that neither (b,c,d,a) nor (d,a,b,c) is in T_1, then we have $(X+1)^2+Y^2=0$ and $(X-1)^2+Y^2=0$, and combining we get $X=0$ and $Y=\pm\iota$. Since $\overline{\gamma}_2=(c,d,a,b)$ is in T_2, we know that $X^2+(Y-1)^2=0$, but plugging in $X=0$ and $Y=\pm\iota$, we get $0+(\pm\iota-1)^2=\pm2\iota\neq0$. This contradiction shows that $\overline{\gamma}_2$ is in the overlap, and we established above that elements of its type are the only possibilities. We conclude then that the count of T_1 and T_2 overlaps is exactly the number of lonely conj 2 elements from T_2 whose alt sum is –1, and that count is

$$(p-2)(2p-5).$$

We move now to possible overlaps of lonely elements arising from T_1 and T_3. We note that since lonely elements arising either T_1 or T_3 must be of type conj 1 or conj 3 (since conj 2 element have an alt sum of 1), all overlap elements must have an alt sum of –1. We know that there are $2p-5$ such lonely conj 1 polynomials $\overline{\gamma}=(b,c,d,a)$ arising from $\overline{\beta}=(a,b,c,d)$ in T_3, and we know also that $X^2+(Y-1)^2=0$. The question is: how many of these elements arise as lonely conj 3 elements from a polynomial (c,a,b,d) in T_1. Suppose (c,d,a,b) is not in T_1, then we must have $X^2+(Y+1)^2=0$, and combining with the other condition above gives $Y=0$ and $X=\pm\iota$. Taking the case $X=\iota$, we must eliminate elements $\overline{\gamma}$ of the form $(b,a+\iota,b,a)$. But since $\overline{\gamma}$'s alt sum is –1, we have $2b-2a-\iota=-1$, i. e., $b=a+2^{-1}(\iota-1)$. Moreover, for a fixed sum v, we have then $a+2^{-1}(\iota-1)+a+\iota+a+2^{-1}(\iota-1)+a=v$, i. e., $4a+1-\iota+\iota=v$, so $a=4^{-1}(v-1)$. We conclude that we must eliminate a single element for the case $X=\iota$, and the same will hold for the case $X=-\iota$. We conclude that $2p-7$ lonely conj 1 elements arising from T_3 are in the T_1 overlap. A parallel argument will give $2p-7$ conj 3 elements, and

finally no γ can be simultaneously a conj 1 and conj 3 for the same β in T_3 since then we would have $d = b$ and $c = a$, i. e., $X = Y = 0$. Thus $X^2 + Y^2 = 0$, which cannot be true for normal elements (i. e., of elements of N), and we conclude that the overlap count in the T_1 and T_3 case is

$$2(p-2)(2p-7).$$

For the third double overlap possibility, namely lonely conjugate elements arising from T_2 and T_3, suppose $\overline{\gamma} = (b, c, d, a)$ is a lonely conj 1 polynomial arising from $\overline{\beta} = (a, b, c, d)$ in T_3, then again $\overline{\gamma}$'s alt sum is –1 and $X^2 + (Y-1)^2 = 0$. The element γ cannot be a conj 1 or a conj 3 arising from T_2 since elements of T_2 do not have an alt sum of 1, but γ could be a conj 2 element arising from (d, a, b, c), which will be in T_2 exactly when $(X-1)^2 + Y^2 = 0$. Combining this with the condition above, we get $-2X + 1 + 2Y - 1 = 0$, so $Y = X$ and plugging in we get $2X^2 - 2X + 1 = 0$, which will yield two distinct nonzero values, $u_1 = 2^{-1}(1+\iota)$ and $u_2 = 2^{-1}(1-\iota)$. Hence (d, a, b, c) is in T_2 only if it's of the form $(b+u_1, a, b, a+u_1)$ or $(b+u_2, a, b, a+u_2)$, so $\overline{\gamma}$ is of the form $(b, a+u_1, b+u_1, a)$ or $(b, a+u_2, b+u_2, a)$. But now applying the conditions that the sum is fixed at v and the alt sum at –1, we see that there are exactly two lonely elements γ which are simultaneously conj 1 elements out of T_3 and conj 2 elements out of T_2. A parallel argument shows the same result for elements γ which are simultaneously conj 3 elements out of T_3 and conj 2 elements out of T_2. The final count then is

$$4(p-2).$$

Finally, we must count the triple overlap. Here, we claim that for the fixed sum v, the two lonely conj 1 (from T_3) elements γ we identified just above in the T_2 and T_3 overlap are in fact conj 3 elements arising from T_1. Let $u_1 = 2^{-1}(1+\iota)$ be the first of two values there and suppose $\overline{\gamma}$ is of the form $(b, a+u_1, b+u_1, a)$, then $\overline{\gamma}$ is a conj 3 for the polynomial $(a+u_1, b+u_1, a, b)$ which will be in T_1 provided that $(a-a-u_1)^2+(b-b-u_1-1)^2$ is nonzero. But that expression is

$$\begin{aligned} u_1^2 + (u_1+1)^2 &= 2u_1^2 + 2u_1 + 1 = 2\big(4^{-1}(2\iota)\big) + 2\big(2^{-1}(1+\iota)\big) + 1 \\ &= \iota + 1 + \iota + 1 = 2(\iota+1) \neq 0. \end{aligned}$$

Similar calculations for the three other elements γ identified above will also yield a nonzero result, so all four of them are in the triple overlap, and there obviously can be no others. Hence our total triple overlap count is

$$4(p-2).$$

3.5 Putting all the pieces together

We are now finally in a position to count the total number of lonely elements arising in quartic extensions of $\mathbf{F}_p$ for p of the form $4k+1$. Since $p-2$ appears as a factor in

every count, we factor it out from the start:

$$\begin{aligned} &T_1 + T_2 + T_3 - T_1T_2 - T_1T_3 - T_2T_3 + T_1T_2T_3 \\ &\quad = (p-2)\left[(p^2 - 4p + 2) + (6p^2 - 37p + 58) + 4(p-3) - (2p-5) - 2(2p-7) - 4 + 4\right] \\ &\quad = (p-2)(7p^2 - 43p + 67). \end{aligned}$$

This count is confirmed by Mathematica© for the cases $p = 5$, 13, and 17, as displayed at the opening of this paper. We have then the following theorem.

Theorem 3.1. *Suppose that p is of the form $4k + 1$ for some k. Among the $\Phi_2(t^4 - 1) = (p-2)^4$ elements of $\mathbf{F}_{p^4}$ which are of normal depth-2, there are $(p-2)(7p^2 - 43p + 67)$ lonely elements, and hence $(p-2)(p-3)(p-5)^2/4$ normal bases.*

Proof. The lonely element count is given above. The number of normal bases of depth-2 is then

$$\begin{aligned} \left((p-2)^4 - (p-2)(7p^2 - 43p + 67)\right)/4 &= (p-2)(p^3 - 13p^2 + 55p - 75)/4 \\ &= (p-2)(p-3)(p-5)^2/4, \end{aligned}$$

which completes the proof. □

4 Conclusion

We observe that there seems to be little similarity in the formulas for the counts of the lonely elements for the two cases $p = 4k + 3$ and $p = 4k + 1$ (except of course for the ever-present factor of $p - 2$), but the two resulting formulas for the counts of normal bases of depth-2 are intriguingly similar, the only difference being the factor of $p^2 - 5$ in the former and $(p-5)^2$ in the latter. This would seem to point to a structure which is certainly not apparent from our somewhat complicated analysis, but may reveal itself by a some wholly different approach to the problem. As an example, in [4] the authors, using more advanced tools, produce a relatively short argument predicting the counts of depth-2 normal bases in $\mathbf{F}_{p^3}$ for p of the form $3k + 1$, which count matches exactly the count predicted by this author in [2] for that same case using a much more lengthy counting argument. For another example, in [4] the authors prove quickly that if $p = n + 1$ then there are *no* depth-2 normal bases. In this paper, we observe (with considerable effort) this fact for the case $p = 5$ since for *all* p of the form $4k + 1$, the number of depth-2 normal bases has a factor of $p - 5$ in it. It is of course reassuring that two very different approaches seem to be yielding identical results.

Bibliography

[1] G. Effinger and G. L. Mullen, Two extended Euler functions with applications to latin squares and bases of finite field extensions, *Bull. Inst. Comb. Appl.*, **85** (2019), 92–111.
[2] G. Effinger, On elements of normal depth-2 in cubic extensions of $\mathbf{F}_p$, available upon request.
[3] R. Lidl and H. Niederreiter, Finite Fields, Encyclopedia of Mathematics and its Applications, Vol. 20, Sec. ed., Cambridge University Press, Cambridge, 1997.
[4] J. Sheekey and D. Thomson, A note on depth-*b* normal elements, in this volume.

James Davis, John Polhill, and Ken Smith

Relative linking systems of difference sets and linking systems of relative difference sets

Abstract: In this paper, we consider the problem of linking systems of difference sets as well as variations. We review recent results on the subject, give new constructions of linked relative difference sets and relative linking systems of nonreversible difference sets, and conclude with several open questions.

Keywords: Difference set, relative difference set, linking system

MSC 2010: 05E30, 05B10

1 Introduction

Difference sets and their variations have been studied extensively due to their many combinatorial connections and applications, for example, with designs, graphs, error-correcting codes, and cryptographic schemes to name a few [1], [5]. We begin with the basic definitions before proceeding to the notion of linking.

Definition 1. Let G be a finite group of order v and let D be a subset of order k. Suppose further that the differences $d_1 d_2^{-1}$ for $d_1, d_2 \in D, d_1 \neq d_2$ represent each of the nonidentity elements in G exactly λ times. Then we call D a (v, k, λ)-*difference set (DS)* in G.

Definition 2. Let G be a finite group of order v and let D be a subset of order k. Suppose further that the differences $d_1 d_2^{-1}$ for $d_1, d_2 \in D, d_1 \neq d_2$ represent each of the nonidentity elements in D exactly λ times and each nonidentity element in $G \setminus D$ exactly μ times. Then we call D a (v, k, λ, μ)-*partial difference set (PDS)* in G.

Note that if we require that the identity element $e \notin D$ and that for all $d \in D$, $d^{-1} \in D$, then D is a *regular* partial difference set. All of the PDSs in this paper are regular. A partial difference set having parameters $(n^2, r(n-1), n + r^2 - 3r, r^2 - r)$ is called a *Latin square type PDS*. Similarly, a partial difference set having parameters $(n^2, r(n+1), -n + r^2 + 3r, r^2 + r)$ is called a *negative Latin square type PDS*.

Definition 3. Let G be a finite group of order v with a normal subgroup N of order n, and assume that $v = mn$. A subset R of cardinality k is called an (m, n, k, λ)-*relative*

James Davis, Department of Mathematics and Computer Science, University of Richmond, VA 23173, USA, e-mail: jdavis@richmond.edu
John Polhill, Department of Mathematics, Computer Science, and Statistics, Bloomsburg University, Bloomsburg, PA 17815, USA
Ken Smith, Department of Mathematics, Sam Houston State University, Huntsville, TX, USA

https://doi.org/10.1515/9783110621730-004

difference set (RDS) in G relative to N if the differences $d_1 d_2^{-1}$ for $d_1, d_2 \in R, d_1 \neq d_2$ represent each nonidentity element of $G \setminus N$ exactly λ times and each element of N zero times.

N is called the *forbidden subgroup*. If $G = H \times N$, where H is a subgroup of G, then R is called a *splitting RDS*. RDSs are said to be *semiregular* when $k - \lambda n = 0$. The RDSs constructed in this paper will be semiregular $(p^a, p^b, p^a, p^{(a-b)})$-RDSs, which have been studied extensively, for example, [8], [12], and [14].

For further study, the text of Beth, Jungnickel, and Lenz provides an excellent overview of difference sets [1]. Similarly, Ma's survey of partial difference sets [9] and Pott's text on relative difference sets [14] provide a thorough background on PDSs and RDSs, respectively.

Often difference sets and their variations are studied within the context of the group ring $\mathbb{Z}[G]$. For a subset D in G, we can abuse notation and write $D = \sum_{d \in D} d$ and $D^{(-1)} = \sum_{d \in D} d^{-1}$. Contextual clues tell us whether D will represent the difference set D or the element $\sum_{d \in D} d$ in the group ring $\mathbb{Z}[G]$. The following equations are the group ring equivalents for the various types of difference sets (for the RDS equation, the forbidden subgroup is N):

- $((v, k, \lambda)$-DS): $DD^{(-1)} = \lambda G + (k - \lambda)1_G$.
- $((v, k, \lambda, \mu)$-PDS): $DD^{(-1)} = D^2 = \mu(G - D) + \lambda D + (k - \mu)1_G$.
- $((m, n, k, \lambda)$-RDS): $RR^{(-1)} = \lambda(G - N) + k1_G$.

The remainder of the paper is organized as follows. Section 2 discusses variations of linked systems of difference sets. Section 3 gives a construction of linked systems where the linked objects are relative difference sets. The fourth and final section gives a general construction of a relative linking system of nonreversible difference sets and also suggests some possible directions for further research.

2 Linking difference sets

Until recently, the only infinite family of linking systems of symmetric designs was constructed by Cameron and Seidel in 1973 using bent functions arising from Kerdock codes [2]. The more recent work of Martin, Van Dam, and Muzychuk [10] showed that a system of linked symmetric designs is exactly equivalent to a 3-class Q-antipodal cometric association scheme and can be used to construct a 4-class Q-antipodal Q-bipartite cometric association scheme and real mutually unbiased bases. This generated renewed interest in discovering examples of linked systems of symmetric designs, and the notion of linked difference sets was introduced [3]. Indeed linking systems of difference sets will provide examples of linking systems of designs. The original definition of linking system of difference sets has now been simplified in [6] and is given below.

Definition 4. Let G be a finite group of order v and let $\ell \geq 2$. Suppose $\mathcal{L} = \{D_{i,j} | 1 \leq i, j \leq \ell, i \neq j\}$ is a set of (v, k, λ) difference sets in G. Then $\mathcal{L}$ is a $(v, k, \lambda; \ell)$-*linking system of difference sets in G of size* ℓ if there are integers $\mu, \nu \in \mathbb{Z}$ such that for all distinct i, j, k we have

$$D_{i,j}D_{j,k}^{(-1)} = (\mu - \nu)D_{i,k} + \nu G.$$

Variations on this idea have been explored in [4]. For example, if $D_{i,j}D_{j,k}^{(-1)} = \mu D_{i,k} + \nu(G - D_{i,k}) + cH$ for some subgroup H, then we will call this a *relative linking system*. In the next section, we will construct yet another variation where the sets D are relative difference sets. We will call this a linking system of relative difference sets.

Definition 5. Let G be a finite group of order v and let $\ell \geq 2$. A collection $\{R_{i,j} \mid 0 \leq i, j \leq \ell, i \neq j\}$ of (m, n, k, λ) relative difference sets in G all relative to the same subgroup N with $v = mn$ is an $(m, n, k, \lambda; \ell)$-*linking system of relative difference sets* for $\mu, \nu \in \mathbb{Z}$ if $R_{i,j}R_{j,k}^{(-1)} = (\mu - \nu)R_{i,k} + \nu G$, for some relative difference set $R_{i,k}$ relative to N.

3 Construction of linked systems of relative difference sets

In this section, we will make use of partial difference set partitions of abelian groups to form relative difference sets that form a linking system. What we require is the following.

Definition 6. Let G be an Abelian p-group of order $q^{2r}, q = p^t, p$ prime, whose non-identity elements can be partitioned into q partial difference sets of the Latin square type, $C_0, C_1, \ldots, C_{q-1}$ such that $|C_0| = (M + 1)(q^r - 1)$ and $|C_i| = M(q^r - 1)$ for $i \neq 0$ where $M = q^{r-1}$ (for convenience in later computations). We call this partition a q-quasi-hyperplane partition.

One such q-quasi-hyperplane partition can be constructed as follows. Let $\mathbb{F}_{q^r}$ be the finite field with $q^r = (p^t)^r$ elements for p prime, $t \in \mathcal{M}$. The 2-dimensional vector space $V = \mathbb{F}_{q^r}^2$ has $q^r + 1$ hyperplanes (1-dimensional subspaces), $H_i, 0 \leq i \leq q^r$. Then $C_0 = \bigcup_{i=0}^{q^{r-1}}(H_i \backslash \{0\})$ and $C_j = \bigcup_{i=jq^{r-1}+1}^{(j+1)q^{r-1}}(H_i \backslash \{0\}), 1 \leq j \leq (q - 1)$, is a q-quasi-hyperplane partition of V.

There are other Abelian groups which are not elementary Abelian that achieve this same partition. For example, [13] gives the following result (simplification of Corollary 6.1 in that paper).

Theorem 1. *Let* $G = (Z_p)^{2s_1} \times (Z_{p^2})^{2s_2} \times \cdots \times (Z_{p^{2k}})^{2s_k}$ *where the* s_i *are nonnegative integers, and let* $|G| = n^2$. *Then G has a partition into partial difference sets which is a q-quasi-hyperplane partition.*

Due to a result of Van Dam [15], we also have that the union of any of the PDSs in the partition is a PDS. We encapsulate the PDS information in the following group ring equations. The first two are immediate, while the latter pair require some calculation and using this fact that $C_i \cup C_j$ will be a partial difference set as well:

- $C_i{}^2 = M(q^r - 1) + (q^r + M^2 - 3M)C_i + (M^2 - M)(G^* - C_i)$ for $i \neq 0$
- $C_0{}^2 = (M+1)(q^r - 1) + (q^r + (M+1)^2 - 3(M+1))C_0 + ((M+1)^2 - (M+1))(G^* - C_0)$
- $C_iC_j = (M^2 + M)(C_i + C_j) + M^2(G^* - C_i - C_j)$
- $(1 + C_0)C_i = M^2(C_0 + C_i) + (M^2 + M)(G^* - C_0 - C_i)$

Now we are ready to define the relative difference sets in the group $\mathbb{F}_q \times G$, where G is a group with a q-quasi-hyperplane partition. Let $\mathbb{F}_q = \{0_q, x_1, x_2, \ldots, x_{q^r-1}\}$. Then for each $x \in \mathbb{F}_q, x \neq 0_q$ define R_x as follows:

$$R_x = (0_q, 1 + C_0) \cup (xx_1, C_1) \cup (xx_2, C_2) \cup \cdots \cup (xx_{q-1}, C_{q-1}).$$

For proof that R_x is a $(q^{2r}, q, q^{2r}, q^{2r-1})$ RDS in $\mathbb{F}_q \times G$ relative to $\mathbb{F}_q \times \{0_G\}$, see [8] and [12]. The following theorem shows that these RDSs form a linking system of RDSs.

Theorem 2. *The relative difference sets* $R_x = (0, 1 + C_0) \cup (xx_1, C_1) \cup (xx_2, C_2) \cup \cdots \cup (xx_{q-1}, C_{q-1})$ *for* $x \in \mathbb{F}_q^*$ *form a* $(q^{2r}, q, q^{2r}, q^{2r-1}; q-1)$*-linking system of relative difference sets in* $\mathbb{F}_q \times G$.

Proof. We will show that for $x, y \in \mathbb{F}_q^*$ and $x \neq y$, that the following is true:

$$R_xR_y{}^{(-1)} = (q^{2r-1} - q^{r-1})(\mathbb{F}_q \times G) + q^r R_{(x-y)^{-1}}.$$

In group ring notation, we have

$$\begin{aligned} R_xR_y{}^{(-1)} = {} & \big((0, (1 + C_0)) + (xx_1, C_1) + (xx_2, C_2) + \cdots + (xx_{q-1}, C_{q-1})\big) \\ & \times \big((0, (1 + C_0)) - (yx_1, C_1) - (yx_2, C_2) - \cdots - (yx_{q-1}, C_{q-1})\big). \end{aligned}$$

We begin computing the coefficient of $(0, 0_G)$: we only need consider the term $(1 + C_0)^2$, which gives $1 + (M+1)(q^r - 1) = q^{2r-1} - q^{r-1} + q^r$. Group elements in $(0, C_0)$ have $2 + q^r + (M+1)^2 - 3(M+1)$ from the $(1 + C_0)^2$ term and $q - 1$ terms with M^2. This gives $2 + q^r + (M+1)^2 - 3(M+1) + (q-1)M^2 = q^{2r-1} - q^{r-1} + q^r$. Coefficients for the terms $(0, C_i)$ are calculated from $(M+1)^2 + (M+1)$ from the $(1 + C_0)^2$ term, $M^2 - M$ from the two terms involving C_i and $q - 3$ remaining terms with M^2. Putting this together yields $q^{2r-1} - q^{r-1}$.

Now we consider terms $(z, g), z \in \mathbb{F}_q^*, g \in G$. Notice that $xa - ya = z$ has the unique solution $a = z(x - y)^{-1}$. So in our expansion, when we collect terms with z we get

$$\left(z, C_{x^{-1}z}(1 + C_0) + (1 + C_0)C_{-y^{-1}z} + \sum_{i,j:xx_i - yx_j = z} C_iC_j\right).$$

The coefficient of (z, C_a) will be derived from the first two terms yielding $M^2 + M$, the one term giving $q^r + qM^2 - 3M$ and the remaining $q - 3$ terms with M^2. Putting it together yields: $q^{2r-1} - q^{r-1} + q^r$. The $(z, 0_G)$ term comes solely from ${C_a}^2$, and is $q^{2r-1} - q^{r-1}$. There are several other cases to consider, but essentially each gets $M^2 + M - M\delta$ from the $(1 + C_0)$ terms, $M^2 - M$ from the C_a term, and $M^2 - M\delta$ from the other terms. The key is that $\delta = 1$ exactly twice in every case, and 0 otherwise. So we get

$$2(M^2 + M) + M^2 - M + (q - 3)M^2 - 2M = qM^2 - M = q^{2r-1} - q^{r-1}.$$

Putting all the pieces together, we have

$$R_x {R_y}^{(-1)} = (q^{2r-1} - q^{r-1})(\mathbb{F}_q \times G) + q^r R_{(x-y)^{-1}}. \qquad \square$$

4 Relative linking systems with McFarland difference sets and questions to ponder

The paper by Martin, Van Dam, and Muzychuk investigates the relationship of linked systems of designs with other association schemes [10]. Kodalen has now constructed linked systems of designs with Hadamard parameters that are not a power of two [7]. In the case of linked systems of difference sets, however, all known examples have Hadamard parameters and are in 2-groups. There are the original constructions of Cameron and Seidel [2], those given in the paper by Davis, Martin, and Polhill [3], and additional examples in the work of Jedwab, Li, and Simon [6]. In the latter article, the authors also show that neither the McFarland nor Spence families of difference sets can form a linking system.

Research problem 1. *Find linked systems of difference sets with parameters that are not Hadamard.*

Research problem 2. *Find linked systems of difference sets with parameters that are Hadamard but not in 2-groups.*

Research problem 3. *Rule out various difference set families as candidates for forming a linking system.*

In this paper and [4], examples of linked systems of relative and almost difference sets are given. Are there more constructions of these or other difference set variations or perhaps additional examples of relative linking systems other than the one given in [4]?

Research problem 4. *Find new examples of linked systems of relative difference sets, almost difference sets, etc.*

We can construct a family of relative linking systems generalizing the one on 4000 points from [4]. That example came from the remarkable reversible McFarland difference sets in $\mathbb{F}_{32} \times \mathbb{F}_{125}$. The new examples are not reversible, and hence are the first family of nonreversible examples for any type of linking system other than those constructed in the Hadamard family by Jedwab, Li, and Simon [6].

Theorem 3. *For p prime and $q = p^r$, suppose that $q+2$ is also a prime power. Then there is a $(q^2(q+2), q(q+1), q; q+1)$-relative linking system of difference sets in the group $G = \mathbb{F}_{q+2} \times \mathbb{F}_q \times \mathbb{F}_q$ relative to $\{0\} \times \mathbb{F}_q \times \mathbb{F}_q$.*

Proof. Let $\mathbb{F}_{q+2} = \{0, x_1, x_2, x_3, \ldots, x_{q+1}\}$. We construct McFarland difference sets using the hyperplane decomposition of $\{0\} \times \mathbb{F}_q \times \mathbb{F}_q$, denoting the $q+1$ hyperplanes $(0, H_1), (0, H_2), \ldots, (0, H_{q+1})$. Then for each nonzero element $x_a \in \mathbb{F}_{q+2}$ the associated McFarland difference set is

$$D_{x_a} = (x_a x_1, H_1) + (x_a x_2, H_2) + \cdots + (x_a x_{q+1}, H_{q+1}), 1 \le a \le q+1.$$

Note that $(0, H_i)(0, H_j) = \{0\} \times \mathbb{F}_q \times \mathbb{F}_q$ if $i \neq j$ while $((0, H_i))^2 = q(0, H_i)$. McFarland proved in [11] that D_{x_a} is a difference set, so we need only show that we have a closed relative linking system and specifically we show for $x_a \neq x_b$ that $D_{x_a} D_{x_b}^{(-1)} = (q-1)G + 2(\{0\} \times \mathbb{F}_q \times \mathbb{F}_q) + qD_{(x_a - x_b)}$.

$$\begin{aligned} D_{x_a} D_{x_b}^{(-1)} = {} & \big((x_a x_1, H_1) + (x_a x_2, H_2) + \cdots + (x_a x_{q+1}, H_{q+1})\big)\big(-(x_b x_1, H_1) - (x_b x_2, H_2) - \cdots \\ & - (x_b x_{q+1}, H_{q+1})\big). \end{aligned}$$

Since for all i, $x_a x_i \neq x_b x_i$, then all nonidentity elements of the form $(0, z_1, z_2), z_1, z_2 \in \mathbb{F}_q$ have coefficients determined as follows:

$$\left(0, \sum_{(i,j: x_a x_i = x_b x_j)} H_i H_j\right) = (q+1)\big(\{0\} \times \mathbb{F}_q \times \mathbb{F}_q\big).$$

On the other hand, for the terms $(y, z_1, z_2), y \in \mathbb{F}_{q+2}, z_1, z_2 \in \mathbb{F}_q$, we see that $x_a x_i - x_b x_i = y$ has the solution $c = y(x_a - x_b)^{-1}$. Thus,

$$\begin{aligned} & \left(y, \sum_{(i,j: x_a x_i - x_b x_j = y)} H_i H_j\right) \\ & \quad = \left(y, \sum_{(i \neq j): x_a x_i - x_b x_j = y} H_i H_j + {H_c}^2\right) = (q-1)(y, \mathbb{F}_q \times \mathbb{F}_q) + q(y, H_c). \end{aligned}$$

Putting this all together, we obtain

$$D_{x_a} D_{x_b}^{(-1)} = (q-1)G + 2\big(\{0\} \times \mathbb{F}_q \times \mathbb{F}_q\big) + qD_{(x_a - x_b)}. \qquad \square$$

This shows that some of the McFarland difference set family fits into a relative linking system, though they cannot form a linking system as proved in [6].

We conclude with one last problem. While [10] gives clear evidence of the importance of linking systems of difference sets, the variations have not been explored out of the context of difference sets.

Research problem 5. *Explore the variations of linked systems of difference sets in other contexts. In particular, what interesting properties have the generalized bent functions derived from the linking systems of relative difference sets given in this paper?*

Bibliography

[1] T. Beth, D. Jungnickel, and H. Lenz, Design Theory, Encyclopedia of Mathematics and Its Applications, Vol. 78, Second ed. Cambridge University Press, Cambridge, 1999.

[2] P. J. Cameron and J. J. Seidel, Quadratic forms over GF (2). *Indag. Math.*, **35** (1973), 1–8.

[3] J. A. Davis, W. Martin, and J. Polhill, Linking systems in nonelementary Abelian groups, *J. Comb. Theory, Ser. A*, **123** (2014), 92–103.

[4] J. A. Davis, J. Polhill, and K. Smith, Relative and almost linking systems, *J. Algebraic Comb.*, **50**(1) (2019), 113–118.

[5] C. Ding, Codes from Difference Sets, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2015.

[6] J. Jedwab, S. Li, S. Simon, Linking systems of difference sets, *J. Comb. Des.*, **27**(3) (2019), 161–187.

[7] B. Kodalen, Linked systems of symmetric designs, *Algebraic Combin.*, **2**(1) (2019), 119–147.

[8] K. H. Leung and S. L. Ma, Constructions of partial difference sets and relative difference sets on *p*-groups, *Bull. Lond. Math. Soc.*, **22** (1990), 533–539.

[9] S. L. Ma, A survey of partial difference sets, *Des. Codes Cryptogr.*, **4**(3) (1994), 221–261.

[10] E. R. van Dam, W. Martin, and M. Muzychuk, Uniformity in association schemes and coherent configurations: cometric Q-antipodal schemes and linked systems, *J. Comb. Theory, Ser. A*, **120**(7) (2013), 1401–1439.

[11] R. McFarland, A family of difference sets in non-cyclic groups, *J. Comb. Theory, Ser. A*, **15** (1973), 1–10.

[12] J. Polhill, A construction of layered relative difference sets using Galois rings, *Ars Comb.*, **78** (2006), 83–94.

[13] J. Polhill, New negative Latin square type partial difference sets in nonelementary abelian 2-groups and 3-groups, *Des. Codes Cryptogr.*, **46**(3) (2008), 365–377.

[14] A. Pott, Finite Geometry and Character Theory, Springer-Verlag, Berlin, 1995.

[15] E. R. van Dam, Strongly regular decompositions of the complete graph, *J. Algebraic Comb.*, **17** (2003), 181–201.

Christian Kaspers and Alexander Pott

On solving isomorphism problems about 2-designs using block intersection numbers

Abstract: In this paper, we give a partial solution to a new isomorphism problem about 2-$(v, k, k-1)$ designs from disjoint difference families in finite fields and Galois rings. Our results are obtained by carefully calculating and bounding some block intersection numbers, and we give insight on the limitations of this technique. Moreover, we present results on cyclotomic numbers, the multiplicities of block intersection numbers of certain designs and on the structure of Galois rings of characteristic p^2.

Keywords: Disjoint difference family, Galois ring, combinatorial design, isomorphism problem, intersection number, cyclotomic number

MSC 2010: 05B05, 05B10, 11T22

1 Introduction

In their previous work [15], the present authors studied two constructions of difference families in Galois rings by Davis, Huczynska, and Mullen [9] and by Momihara [17]. Both constructions were inspired by a classical construction of difference families in finite fields which was introduced by Wilson [21] in 1972. Various types of difference families have long been studied in combinatorial literature [1, 3, 5, 6, 7, 12, 14, 21]. They have applications in coding theory and communications and information security [18], and they are related to many other combinatorial objects. In particular, every difference family gives rise to a combinatorial design. Combinatorial designs themselves have been extensively studied since the first half of the nineteenth century, they have many applications in group theory, finite geometry and cryptography [3, 8].

Whenever a new construction of difference families is given, the natural question arises whether the associated designs are also new or whether they are isomorphic to known designs. By calculating and bounding some block intersection numbers, the present authors [15] solved this isomorphism problem for the difference families from Momihara [17] and Wilson [21] and for those from Davis, Huczynska, and Mullen [9] and Wilson [21]. In this paper, we obtain new difference families from the ones constructed by Davis, Huczynska, and Mullen [9]. These new difference families also have an analogue in finite fields from Wilson's [21] construction. Motivated by the present authors' previous results, we will use the same technique as in their paper [15] to study whether the associated designs are isomorphic or not. It will become clear that the ap-

Christian Kaspers, Alexander Pott, Insitute for Algebra and Geometry, Otto von Guericke University Magdeburg, 39106 Magdeburg, Germany, e-mails: christian.kaspers@ovgu.de, alexander.pott@ovgu.de

https://doi.org/10.1515/9783110621730-005

proach to use block intersection numbers as a tool to solve isomorphism problems is promising for certain types of designs but has its limitations in general.

We start by defining the objects we study in this paper. First, we need the following notation: Let G be an additively written Abelian group, $A, B \subseteq G$ and $g \in G$. We define multisets

$$\begin{aligned} \Delta A &= \{a - a' : a, a' \in A, a \neq a'\}, \\ A - B &= \{a - b : a \in A, b \in B, a \neq b\}, \\ A + g &= \{a + g : a \in A\}. \end{aligned}$$

We will sometimes use these notation to denote sets, not multisets. It will be clear from the context whether we mean the multiset or the respective set.

Definition 1. Let G be an Abelian group of order v, and let $D_1, D_2, \ldots, D_b$ be k-subsets of G. The collection $D = \{D_1, D_2, \ldots, D_b\}$ is called a (v, k, λ) *difference family in* G if each nonzero element of G occurs exactly λ times in the multiset union

$$\bigcup_{i=1}^{b} \Delta D_i.$$

If the subsets $D_1, D_2, \ldots, D_b$ are mutually disjoint, they form a *disjoint difference family*. If $b = 1$, one speaks of a (v, k, λ) *difference set*. We call D *near-complete* if the subsets $D_1, D_2, \ldots, D_b$ partition $G \setminus \{0\}$.

In this paper, we focus on near-complete $(v, k, k-1)$ disjoint difference families. For more background on this type of difference families, the reader is referred to the survey by Buratti [5] who summarizes many results and introduces a powerful new construction. His construction includes many known constructions, including the one by Davis, Huczynska, and Mullen [9]. However, it seems to be too general to use it for studying isomorphism problems, at least when using block intersection numbers. Eventually, we remark that every near-complete disjoint difference family is also an external difference family [7, 9, 15].

As mentioned above, every difference family gives rise to a combinatorial design.

Definition 2. Let P be a set with v elements that are called *points*. A t-(v, k, λ) *design*, or *t-design*, in brief, is a collection of k-subsets, called *blocks*, of P such that every t-subset of P is contained in exactly λ blocks.

The associated designs of difference families are 2-designs which are often referred to as *balanced incomplete block designs (BIBD)*. They are constructed as the development of a difference family.

Definition 3. Let G be an Abelian group, and let $D = \{D_1, D_2, \ldots, D_b\}$ be a collection of subsets of G. The *development of* D is the collection

$$\operatorname{dev}(D) = \{D_i + g : D_i \in D, g \in G\}$$

of all the translates of the sets $D_1, D_2, \ldots, D_b$. The sets $D_1, D_2, \ldots, D_b$ are called the *base blocks* of $\mathrm{dev}(D)$.

In other words, the development of D is the union of the orbits of the sets contained in D under the action of G. If all orbits have full length, $\mathrm{dev}(D)$ contains vb blocks. The following well-known proposition relates difference families to 2-designs.

Proposition 1.1. *Let D be a (v, k, λ) difference family in an Abelian group G. The development $\mathrm{dev}(D)$ of D forms a 2-(v, k, λ) design with point set G.*

2 Galois rings

In this section, we give a short introduction to Galois rings and present some of their well-known properties needed in this paper. We refer to the work by McDonald [16] and Wan [20] for extended general background on this topic. Let p be a prime, and let $f(x) \in \mathbb{Z}_{p^m}[x]$ be a monic basic irreducible polynomial of degree $r \geq 1$, which means that the image of f modulo p in $\mathbb{F}_p[x]$ is irreducible. The factor ring

$$\mathbb{Z}_{p^m}[x]/\langle f(x)\rangle$$

is called a *Galois ring* of characteristic p^m and extension degree r. It is denoted by $\mathrm{GR}(p^m, r)$, and its order is p^{mr}. Since any two Galois rings of the same characteristic and order are isomorphic, we will speak of *the* Galois ring $\mathrm{GR}(p^m, r)$.

Galois rings are local commutative rings. The unique maximal ideal of the ring $R = \mathrm{GR}(p^m, r)$ is

$$\mathcal{I} = pR = \{pa : a \in R\}.$$

The factor ring $R/\mathcal{I}$ is isomorphic to the finite field $\mathbb{F}_{p^r}$ with p^r elements. As a system of representatives of $R/\mathcal{I}$, we take the *Teichmüller set*

$$\mathcal{T} = \{0, 1, \xi, \ldots, \xi^{p^r-2}\},$$

where ξ denotes a root of order $p^r - 1$ of $f(x)$. It is convenient to choose the *generalized Conway polynomial*, that is the Hensel lift from $\mathbb{F}_p[x]$ to $\mathbb{Z}_{p^m}[x]$ of the Conway polynomial, as our polynomial $f(x)$. Then $x + \langle f\rangle$ is a generator of the Teichmüller group, and we set $\xi = x + \langle f\rangle$. Zwanzger [22, Section 1.3] provides more information on the generalized Conway polynomial and its construction. Every $a \in R$ has a unique *p-adic representation* $a = \alpha_0 + p\alpha_1 + \cdots + p^{m-1}\alpha_{m-1}$, where $\alpha_0, \alpha_1, \ldots, \alpha_{m-1} \in \mathcal{T}$.

The elements of $R \setminus \mathcal{I}$ are all the units of R. We denote this *unit group* by R^*. It has order $p^{mr} - p^{(m-1)r}$ and is the direct product of the cyclic *Teichmüller group*

$$\mathcal{T}^* = \mathcal{T} \setminus \{0\}$$

of order $p^r - 1$ and the *group of principal units* $\mathbb{P} = 1 + \mathcal{I}$ of order $p^{(m-1)r}$. If p is odd or if $p = 2$ and $m \leq 2$, then $\mathbb{P}$ is a direct product of r cyclic groups of order p^{m-1}. If $p = 2$ and $m \geq 3$, then $\mathbb{P}$ is a direct product of a cyclic group of order 2, a cyclic group of order 2^{m-2}, and $r - 1$ cyclic groups of order 2^{m-1}. In this paper, we will only consider Galois rings of characteristic p^2. In this case, $(1 + p\alpha)(1 + p\beta) = 1 + p(\alpha + \beta)$ for any $\alpha, \beta \in \mathcal{T}$, and every unit $u \in \mathrm{GR}(p^2, r)^*$ has a unique representation

$$u = \alpha_0(1 + p\alpha_1),$$

where $\alpha_0 \in \mathcal{T}^*$ and $\alpha_1 \in \mathcal{T}$. Moreover, the group of principal units $\mathbb{P}$ is a direct product of r cyclic groups of order p and thus has the structure of an elementary Abelian group of order p^r.

3 Construction of disjoint difference families

In this section, we describe three constructions of disjoint difference families. The constructions from Theorem 3.2 and Theorem 3.3 are well known. The third construction, in Theorem 3.4, follows from results by Furino [12]. As the first two constructions also fall into Furino's very general framework, we will present his result first. We only restate a special case of his construction.

Theorem 3.1 ([12, Theorem 3.3 and Corollary 3.5]). *Let R be a commutative ring with an identity. Denote the cardinality of R by v and the unit group of R by R^*. Let B be a subgroup of R^* of order k such that ΔB is a subset of R^*. Denote by S a system of representatives of the cosets of B in $R \setminus \{0\}$. The collection $\{sB : s \in S\}$ is a $(v, k, k-1)$ disjoint difference family in the additive group of R.*

Note that, by abuse of denotation, we also call sets sB where s is not a unit a coset of B. Because of the condition $\Delta B \subseteq R^*$, these cosets also have cardinality k, and all the cosets partition $R \setminus \{0\}$.

Next, we present the construction of disjoint difference families in finite fields by Wilson [21]. It makes use of the cyclotomy of the e-th powers in a finite field.

Theorem 3.2. *Let $\mathbb{F}_q$ be the finite field with q elements, and let α be a generator of the multiplicative group $\mathbb{F}_q^*$ of $\mathbb{F}_q$. Moreover, let e, f be integers satisfying $ef = q - 1$, where $e, f \geq 2$, and let*

$$C_i = \{\alpha^t : t \equiv i \pmod{e}\},$$

where $i = 0, 1, \ldots, e-1$, be the cosets of the unique subgroup C_0 of index e and order f that is formed by the e-th powers of α in $\mathbb{F}_q^*$. Then the collection $C = \{C_0, C_1, \ldots, C_{e-1}\}$ is a near-complete $(q, f, f-1)$ disjoint difference family in the additive group of $\mathbb{F}_q$.

We now present the construction of disjoint difference families by Davis, Huczynska, and Mullen [9]. We use the same notation as in section 2, and we remark that this theorem also follows from Theorem 3.1 and from a result by Buratti [5].

Theorem 3.3 ([9, Theorem 4.1]). *Let p be a prime, and let r be a positive integer such that $p^r \geq 3$. Denote by $\mathcal{T}$ the Teichmüller set of the Galois ring $GR(p^2, r)$ and by $\mathcal{T}^*$ the Teichmüller group $\mathcal{T}^* = \mathcal{T} \setminus \{0\}$. The collection*

$$E = \{(1 + p\alpha)\mathcal{T}^* : \alpha \in \mathcal{T}\} \cup p\mathcal{T}^*$$

forms a near-complete $(p^{2r}, p^r - 1, p^r - 2)$ disjoint difference family in the additive group of $GR(p^2, r)$.

Since $p^r - 1$ divides $p^{2r} - 1$, there exists a disjoint difference family in the additive group of $\mathbb{F}_{p^{2r}}$ that has the exact same parameters as the difference family from Theorem 3.3. It can be constructed using Theorem 3.2 by taking the (p^r+1)-th powers in $\mathbb{F}_{p^{2r}}$. Inspired by Theorem 3.3 and the work by Furino [12], we noticed that if p is odd, we obtain a new disjoint difference family by taking the cosets of the group of *Teichmüller squares*.

Theorem 3.4. *Let p be an odd prime and let r be a positive integer such that $p^r \geq 5$. Moreover, let*

$$\mathcal{T}^* = \{1, \xi, \xi^2, \ldots, \xi^{p^r-2}\}$$

be the Teichmüller group of the Galois ring $GR(p^2, r)$, and let $\mathcal{T} = \mathcal{T}^* \cup \{0\}$. By

$$\mathcal{T}_S^* = \{1, \xi^2, \ldots, \xi^{p^r-3}\}$$

we denote the set of squares and by

$$\mathcal{T}_N^* = \{\xi, \xi^3, \ldots, \xi^{p^r-2}\}$$

we denote the set of nonsquares in $\mathcal{T}^*$. The collection

$$E^H = \{(1 + p\alpha)\mathcal{T}_S^* : \alpha \in \mathcal{T}\} \cup \{p\mathcal{T}_S^*\} \cup \{(1 + p\alpha)\mathcal{T}_N^* : \alpha \in \mathcal{T}\} \cup \{p\mathcal{T}_N^*\}$$

forms a near complete $(p^{2r}, \frac{p^r-1}{2}, \frac{p^r-3}{2})$ disjoint difference family in the additive group of $GR(p^2, r)$.

Proof. Denote by $\mathcal{I} = p\mathrm{GR}(p^2, r)$ the maximal ideal of $\mathrm{GR}(p^2, r)$. The Teichmüller set $\mathcal{T}$ is a system of representatives of $\mathrm{GR}(p^2, r)/\mathcal{I}$. This factor ring is isomorphic to the finite field $\mathbb{F}_{p^r}$. Consequently, the difference of two distinct elements of the Teichmüller group $\mathcal{T}^*$ is a unit, hence $\Delta\mathcal{T}^* \subseteq \mathrm{GR}(p^2, r)^*$. As $\mathcal{T}_S^*$ is a subgroup of $\mathcal{T}^*$, it follows that $\Delta\mathcal{T}_S^*$ is a subset of the unit group $\mathrm{GR}(p^2, r)^*$. In this case, according to Theorem 3.1, the collection of the cosets of $\mathcal{T}_S^*$ in $\mathrm{GR}(p^2, r) \setminus \{0\}$ forms a disjoint difference family in the additive group of $\mathrm{GR}(p^2, r)$. □

Note that the difference family E^H from Theorem 3.4 can be obtained from the difference family E presented in Theorem 3.3 by cutting the base blocks of E into halves, hence the name E^H. Furthermore, note that there exists a difference family C^H in the finite field $\mathbb{F}_{p^{2r}}$ which has the same parameters as E^H. According to Theorem 3.2, the cosets of the subgroup C_0^H of the $2(p^r+1)$-th powers in $\mathbb{F}_{p^{2r}}^*$ form a $(p^{2r}, \frac{p^r-1}{2}, \frac{p^r-3}{2})$ disjoint difference family in the additive group of $\mathbb{F}_{p^{2r}}$. In the following section, we will study the isomorphism problem for the difference families C^H and E^H from finite fields and Galois rings. Moreover, we present additional isomorphism invariants of the designs in finite fields coming from Theorem 3.2.

4 A partial solution to the isomorphism problem

Denote by C the (p^{2r}, p^r-1, p^r-2) difference family and by C^H the $(p^{2r}, \frac{p^r-1}{2}, \frac{p^r-3}{2})$ difference family in the additive group of $\mathbb{F}_{p^{2r}}$ which are constructed using Theorem 3.2. Denote by E the (p^{2r}, p^r-1, p^r-2) difference family and by E^H the $(p^{2r}, \frac{p^r-1}{2}, \frac{p^r-3}{2})$ difference family in the additive group of $\mathrm{GR}(p^2, r)$ which are constructed using Theorem 3.3 and Theorem 3.4, respectively.

In their previous work, the present authors [15] solved the isomorphism problem for the 2-(p^{2r}, p^r-1, p^r-2) designs $\mathrm{dev}(C)$ and $\mathrm{dev}(E)$. They showed that the designs are nonisomorphic for all combinations of p and r except $p=3$ and $r=1$. In this section, we will give a partial solution to the isomorphism problem for the 2-$(p^{2r}, \frac{p^r-1}{2}, \frac{p^r-3}{2})$ designs $\mathrm{dev}(C^H)$ and $\mathrm{dev}(E^H)$. Note that these designs can be obtained from $\mathrm{dev}(C)$ and $\mathrm{dev}(E)$, respectively, by cutting every block into two halves.

Remark 1. The fact that two designs $\mathcal{D}_1, \mathcal{D}_2$ are nonisomorphic does not imply that two designs $\mathcal{D}_1^H, \mathcal{D}_2^H$ that are obtained by cutting the blocks of $\mathcal{D}_1$ and $\mathcal{D}_2$ into smaller blocks are nonisomorphic. This is shown in the following example which was given in the context of skew Hadamard difference sets by Feng and Xiang [11, Example 3.3]. Denote by C_0 the subgroup of the 14-th powers of the multiplicative group of the finite field $\mathbb{F}_{11^3}$ and by $C_0, C_1, \ldots, C_{13}$ the cosets of C_0. It follows from Theorem 3.2 that the collection $C = \{C_0, C_1, \ldots, C_{13}\}$ is a disjoint difference family in the additive group of $\mathbb{F}_{11^3}$. The collections

$$\begin{aligned}
D_1 &= \{\{C_0 \cup C_2 \cup C_4 \cup C_6 \cup C_8 \cup C_{10} \cup C_{12}\},\\
&\qquad \{C_1 \cup C_3 \cup C_5 \cup C_7 \cup C_9 \cup C_{11} \cup C_{13}\}\},\\
D_2 &= \{\{C_0 \cup C_1 \cup C_2 \cup C_3 \cup C_4 \cup C_5 \cup C_6\},\\
&\qquad \{C_7 \cup C_8 \cup C_9 \cup C_{10} \cup C_{11} \cup C_{12} \cup C_{13}\}\},\\
D_3 &= \{\{C_0 \cup C_1 \cup C_3 \cup C_4 \cup C_5 \cup C_6 \cup C_9\},\\
&\qquad \{C_2 \cup C_7 \cup C_8 \cup C_{10} \cup C_{11} \cup C_{12} \cup C_{13}\}\},
\end{aligned}$$

are also disjoint difference families in the additive group of $\mathbb{F}_{11^3}$. Consider their associated designs $\mathrm{dev}(D_1), \mathrm{dev}(D_2), \mathrm{dev}(D_3)$. Their full automorphism groups $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ have orders $|\mathcal{A}_1| = 5310690, |\mathcal{A}_2| = 252890$ and $|\mathcal{A}_3| = 758670$. Thus, the designs are pairwise nonisomorphic. However, it is clear that from all three difference families, we can obtain the difference family C by cutting their base blocks into the cyclotomic cosets $C_0, C_1, \ldots, C_{13}$. Hence, from the nonisomorphic designs $\mathrm{dev}(D_1), \mathrm{dev}(D_2)$, $\mathrm{dev}(D_3)$, we can obtain the exact same design $\mathrm{dev}(C)$.

The present authors [15] solved the isomorphism problem for $\mathrm{dev}(C)$ and $\mathrm{dev}(E)$ by comparing the block intersection numbers of these designs.

Definition 4. We call an integer N a *block intersection number* of a t-design $\mathcal{D}$, if $\mathcal{D}$ contains two distinct blocks B and B' that intersect in N elements.

Block intersection numbers are invariant under isomorphism. For a given design $\mathcal{D}$, they can be easily computed as the entries of the matrix $M^T M$, where M is the incidence matrix of $\mathcal{D}$ with the rows corresponding to the points and the columns corresponding to the blocks of $\mathcal{D}$. Note, however, that there also exist designs that have the exact same block intersection numbers but are nonisomorphic. One example are the designs given in Remark 1. These designs are pairwise nonisomorphic, but they all share the intersection numbers $0, 332, 333$. For the designs $\mathrm{dev}(C^H)$ and $\mathrm{dev}(E^H)$, however, block intersection numbers seem to distinguish the designs as the following example shows.

Example 1. The constructions from Theorem 3.2 and Theorem 3.4 yield $(625, 12, 11)$ disjoint difference families C^H and E^H in the additive groups of $\mathbb{F}_{5^4}$ and $\mathrm{GR}(25, 2)$, respectively. The associated 2-$(625, 12, 11)$ designs have the following block intersection numbers: for $\mathrm{dev}(C^H)$, they are $0, 1, 5, 6$, and for $\mathrm{dev}(E^H)$, they are $0, 1, 2, 5, 6$. Hence, the two designs are nonisomorphic.

Before we start with the actual calculation of our block intersection numbers, we focus on their multiplicities.

Remark 2. Not only the block intersection numbers themselves but also their multiplicities are isomorphism invariants of a combinatorial design. Hence, in the following, we will not only state the intersection numbers but also their multiplicities whenever it is possible. Although we will not use the multiplicities to solve an isomorphism problem in this paper, they might be useful for further research. To determine the multiplicity of an intersection number N, we will first count the number of pairs (i, j) such that two blocks B_i and B_j intersect in N elements without considering that $B_i \cap B_j = B_j \cap B_i$. In the end, we divide this number by 2.

Example 2. The multiplicities of the block intersection numbers of the designs from Example 1 are as follows: In $\mathrm{dev}(C^H)$, the intersection numbers $0, 1, 5, 6$ occur with multiplicities 410 328 750, 117 000 000, 195 000, and 585 000, respectively. In $\mathrm{dev}(D^H)$,

the intersection numbers $0, 1, 2, 5, 6$ have multiplicities 417 078 750, 100 687 500, 10 312 500, 7 500 and 22 500, respectively. Hence, the multiplicities distinguish the designs.

For all three designs given in Remark 1, the multiplicities of the block intersection numbers $0, 332, 333$ are 1331, 2 655 345, and 885 115, respectively. Hence, in this case, neither the intersection numbers nor their multiplicities distinguish the designs.

In the remainder of this section, we will first generally describe the block intersection numbers and their multiplicities of the designs coming from the disjoint difference families in $\mathbb{F}_q$ that we presented in Theorem 3.2. From this result, we will derive the block intersection numbers and their multiplicities of the designs $\operatorname{dev}(C)$ and $\operatorname{dev}(C^H)$. Note that the present authors, in their previous paper [15], already gave the intersection numbers of $\operatorname{dev}(C)$. We will now contribute the associated multiplicities. Finally, we will establish bounds on the intersection numbers of $\operatorname{dev}(E^H)$ which will lead to a partial solution of the isomorphism problem of the designs $\operatorname{dev}(C^H)$ and $\operatorname{dev}(E^H)$.

The block intersection numbers of the design from Theorem 3.2 are strongly related to the so-called cyclotomic numbers: As in Theorem 3.2, let $C_0, C_1, \dots, C_{e-1}$ be the cosets of the subgroup C_0 of the e-th powers in $\mathbb{F}_q^*$. For fixed nonnegative integers $i, j \le e-1$, the *cyclotomic number* $(i,j)_e$ *of order* e is defined as

$$(i,j)_e = \big|(C_i + 1) \cap C_j\big|.$$

Denote by $n_e(N)$ the number of pairs (i,j), where $i, j \le e-1$, such that the cyclotomic number $(i,j)_e = N$.

Proposition 4.1. *Let $e, f \ge 2$ be integers such that $ef = q-1$, and let C be a $(q, f, f-1)$ disjoint difference family in the additive group of $\mathbb{F}_q$ constructed with Theorem 3.2. The block intersection numbers of the 2-$(q, f, f-1)$ design* $\operatorname{dev}(C)$ *are* 0 *and the values of the cyclotomic numbers $(i,j)_e$ of order e. The intersection number* 0 *has multiplicity $\frac{1}{2}q(q-1)n_e(0) + \frac{1}{2}qe(e-1)$, each nonzero intersection number N has multiplicity $\frac{1}{2}q(q-1)n_e(N)$.*

Proof. Denote by α a primitive element of the finite field $\mathbb{F}_q$. Let $C = \{C_0, C_1, \dots, C_{e-1}\}$ be a disjoint difference family from Theorem 3.2 in the additive group of $\mathbb{F}_q$. Take two arbitrary distinct blocks $C_i + a$ and $C_j + b$ of $\operatorname{dev}(C)$. If we want to calculate the cardinality,

$$\big|(C_i + a) \cap (C_j + b)\big|$$

of their intersection, we need to determine the number of solutions (s,t) of the equation

$$\alpha^{se+i} + a = \alpha^{te+j} + b. \tag{4.1}$$

If $a = b$, then obviously only the case $i \neq j$ is relevant. As C_i and C_j are disjoint, there are no solutions in this case and

$$\left|(C_i + a) \cap (C_j + a)\right| = 0.$$

Since there are q choices for a and $e(e-1)$ choices for (i,j) such that $i \neq j$, the block intersection number 0 occurs $qe(e-1)$ times in this context. Removing repeated intersections, this multiplicity reduces to $\frac{qe(e-1)}{2}$.

If $a \neq b$, then $a - b = \alpha^r$ for some $r \in \{0, \ldots, q-1\}$. Write $r = me + r'$ such that $0 \leq r' \leq e-1$. Now, we can rewrite (4.1) as

$$\alpha^{(s-m)e+(i-r')} + 1 = \alpha^{(t-m)e+(j-r')}.$$

Consequently,

$$\left|(C_i + a) \cap (C_j + b)\right| = \left|(C_{i-r'} + 1) \cap C_{j-r'}\right|,$$

where the subscripts are calculated modulo e. The right-hand side of the above equation is exactly the cyclotomic number $(i - r', j - r')_e$. We have $q(q-1)$ choices for (a,b) such that $a \neq b$, and the difference $a - b$ covers all the elements of $\mathbb{F}_q^*$ the same number of times. Consequently, each cyclotomic number $(i,j)_e$ that equals N contributes with $q(q-1)$ to the multiplicity of the block intersection number N. Removing repeated intersections, this contribution reduces to $\frac{q(q-1)}{2}$. □

Using a result by Baumert, Mills, and Ward [2, Theorems 2 and 4], the present authors [15] showed that the cyclotomic numbers of order $p^r + 1$ in $\mathbb{F}_{p^{2r}}$ are given as

$$\begin{aligned} (0,0)_{p^r+1} &= p^r - 2, \\ (0,i)_{p^r+1} = (i,0)_{p^r+1} = (i,i)_{p^r+1} &= 0 \qquad \text{for } i \neq 0, \\ (i,j)_{p^r+1} &= 1 \qquad \text{for } i \neq j \text{ and } i,j \neq 0. \end{aligned} \tag{4.2}$$

With the help of Theorem 4.1, we can now determine the block intersection numbers of the 2-$(p^{2r}, p^r - 1, p^r - 2)$ design $\mathrm{dev}(C)$ and their multiplicities. While the present authors [15] presented these block intersection numbers before, the results about their multiplicities are new.

Corollary 4.2. *The 2-$(p^{2r}, p^r - 1, p^r - 2)$ design $\mathrm{dev}(C)$ has exactly the following block intersection numbers:*

block intersection number	*multiplicity*
0	$\frac{1}{2}(3p^{5r} + p^{4r} - 2p^{3r})$
1	$\frac{1}{2}(p^{6r} - p^{5r} - p^{4r} + p^{3r})$
$p^r - 2$	$\frac{1}{2}(p^{4r} - p^{2r})$

Proof. Let $e = p^r + 1$. Denote by $n_e(N)$ the number of cyclotomic numbers of order e that equal N. In $\mathbb{F}_{p^{2r}}$, according to (4.2), we have

$$n_e(0) = 3p^r, \quad n_e(1) = p^r(p^r - 1) \quad \text{and} \quad n_e(p^r - 2) = 1. \tag{4.3}$$

We multiply these numbers with the factor $\frac{1}{2}p^{2r}(p^{2r}-1)$ from Theorem 4.1. This gives us the multiplicities of the block intersection numbers 1 and $p^r - 2$. To obtain the multiplicity for 0, according to Theorem 4.1, we additionally need to add $\frac{1}{2}p^{2r}(p^r+1)p^r$. □

Next, we determine the cyclotomic numbers of order $2(p^r+1)$ in $\mathbb{F}_{p^{2r}}$ which are the intersection numbers of $\mathrm{dev}(C^H)$. Unfortunately, these parameters no longer match the conditions of the theorems by Baumert, Mills, and Ward [2] that were used to obtain the cyclotomic numbers of order p^r+1. Nevertheless, we can deduce these cyclotomic numbers from (4.2) with the help of the following well-known lemma.

Lemma 4.3 ([10, §67], [19, Theorem 2]). *Let p be an odd prime. Let S be the set of nonzero squares and N be the set of nonsquares in the finite field $\mathbb{F}_{p^r}$. Denote by QQ the number of squares $s \in S$ for which $s+1$ is a nonzero square and by QN the number of $s \in S$ for which $s+1$ is not a square. Moreover, let NN denote the number of nonsquares $n \in N$ for which $n+1$ is not a square and NQ the number of $n \in N$ for which $n+1$ is a nonzero square.*

– *If $p^r - 1 \equiv 0 \pmod 4$, then*

$$QQ = \frac{p^r-5}{4}, \quad QN = \frac{p^r-1}{4}, \quad NN = \frac{p^r-1}{4}, \quad NQ = \frac{p^r-1}{4}.$$

– *If $p^r - 1 \equiv 2 \pmod 4$, then*

$$QQ = \frac{p^r-3}{4}, \quad QN = \frac{p^r+1}{4}, \quad NN = \frac{p^r-3}{4}, \quad NQ = \frac{p^r-3}{4}.$$

Combining Theorem 4.3 with (4.2), we obtain the following result.

Proposition 4.4. *Let p be an odd prime, and let $e = p^r + 1$ for some positive integer r. In $\mathbb{F}_{p^{2r}}$, the cyclotomic numbers of order $2e$ are as follows:*

– *If $p^r - 1 \equiv 0 \pmod 4$, then*

$$(0,0)_{2e} = \frac{p^r-5}{4},$$
$$(0,e)_{2e} = (e,0)_{2e} = (e,e)_{2e} = \frac{p^r-1}{4}.$$

– *If $p^r - 1 \equiv 2 \pmod 4$, then*

$$(0,e)_{2e} = \frac{p^r+1}{4},$$
$$(0,0)_{2e} = (e,0)_{2e} = (e,e)_{2e} = \frac{p^r-3}{4}.$$

In both of the above cases,

$$\begin{aligned}(0,i)_{2e} &= (i,0)_{2e} = (i,i)_{2e} = (i,e)_{2e}\\ &= (e,i)_{2e} = (i,e+i)_{2e} = (e+i,i)_{2e} = 0 \quad \textit{for } i \notin \{0,e\}.\end{aligned}$$

Out of the remaining cyclotomic numbers,

$$(i,j)_{2e}, (i,j+e)_{2e}, (i+e,j)_{2e}, (i+e,j+e)_{2e}, \quad \textit{where } i,j \neq 0 \textit{ and } i \neq j,$$

for each choice of i and j, exactly one cyclotomic number is 1 *and the other three cyclotomic numbers are* 0*, but it is not known which one is* 1.

Proof. Let α be a generator of $\mathbb{F}_{p^{2r}}^*$, let C_0 be the unique subgroup of order p^r-1 of $\mathbb{F}_{p^{2r}}^*$ formed by the (p^r+1)-th powers, and let $C_0, C_1, \ldots, C_{p^r}$ be the cosets of C_0. The finite field $\mathbb{F}_{p^{2r}}$ contains a unique subfield $\mathbb{F}_{p^r}$ with p^r elements. Hence, the group C_0 is the multiplicative group $\mathbb{F}_{p^r}^*$ of the subfield $\mathbb{F}_{p^r}$. As p^r is odd, C_0 consists of $\frac{1}{2}(p^r-1)$ squares and nonsquares in $\mathbb{F}_{p^r}$ each. Consequently,

$$C_0 = C_0^H \cup C_e^H,$$

where

$$C_0^H = \{\alpha^t \mid t \equiv 0 \pmod{2(p^r+1)}\}$$

is the set of squares and

$$C_e^H = \{\alpha^t \mid t \equiv e \pmod{2(p^r+1)}\}$$

is the set of nonsquares in $\mathbb{F}_{p^r}^*$. The values of the cyclotomic numbers $(i,j)_{2e}$, where $i,j \in \{0,e\}$, now follow from Theorem 4.3. In the same way as before, we can divide each of the cosets $C_0, C_1, \ldots, C_{p^r}$, of C_0 into two cosets C_i^H and C_{e+i}^H of C_0^H. Since

$$C_i = C_i^H \cup C_{e+i}^H$$

for all $i = 0, 1, \ldots, p^r$, we obtain

$$(C_i+1) \cap C_j = \bigcup_{\substack{k \in \{i,e+i\}\\ \ell \in \{j,e+j\}}} (C_k^H+1) \cap C_\ell^H$$

for $0 \le i,j \le p^r$. In terms of cyclotomic numbers, this means

$$(i,j)_e = \sum_{\substack{k \in \{i,e+i\}\\ \ell \in \{j,e+j\}}} (k,\ell)_{2e} \tag{4.4}$$

for $0 \le i,j \le p^r$. The values of the cyclotomic numbers $(i,j)_{2e}$, where $i,j \notin \{0,e\}$, now follow from combining (4.4) with (4.2). □

Unfortunately, the exact values of the cyclotomic numbers $(i,j)_{2e}$, $(i,j+e)_{2e}$, $(i+e,j)_{2e}$, $(i+e,j+e)_{2e}$, where $i,j \neq 0$ and $i \neq j$, in $\mathbb{F}_{p^{2r}}$ are not known in general. It is an open problem to determine those. However, Theorem 4.4 immediately gives us the block intersection numbers of the 2-design $\mathrm{dev}(C^H)$ as well as their multiplicities.

Theorem 4.5. *Let C^H be a $(p^{2r}, \frac{p^r-1}{2}, \frac{p^r-3}{2})$ difference family in the additive group of $\mathbb{F}_{p^{2r}}$ constructed using Theorem 3.2. The associated 2-$(p^{2r}, \frac{p^r-1}{2}, \frac{p^r-3}{2})$ design $\mathrm{dev}(C^H)$ has exactly the following block intersection numbers.*
If $p^r - 1 \equiv 0 \pmod 4$, then

block intersection number	*multiplicity*
0	$\frac{1}{2}(3p^{6r} + 9p^{5r} + p^{4r} - 3p^{3r} + 2p^{2r})$
1	$\frac{1}{2}(p^{6r} - p^{5r} - p^{4r} + p^{3r})$
$\frac{1}{4}(p^r - 5)$	$\frac{1}{2}(p^{4r} - p^{2r})$
$\frac{1}{4}(p^r - 1)$	$\frac{1}{2}(3p^{4r} - 3p^{2r})$.

If $p^r - 1 \equiv 2 \pmod 4$, then

block intersection number	*multiplicity*
0	$\frac{1}{2}(3p^{6r} + 9p^{5r} + p^{4r} - 3p^{3r} + 2p^{2r})$
1	$\frac{1}{2}(p^{6r} - p^{5r} - p^{4r} + p^{3r})$
$\frac{1}{4}(p^r - 3)$	$\frac{1}{2}(3p^{4r} - 3p^{2r})$
$\frac{1}{4}(p^r + 1)$	$\frac{1}{2}(p^{4r} - p^{2r})$.

Proof. Let $e = p^r + 1$. It follows from Theorem 4.1 that the block intersection numbers of $\mathrm{dev}(C^H)$ are exactly 0 and the cyclotomic numbers from Theorem 4.4. We obtain their multiplicities using (4.4): Every cyclotomic number of order e that equals 0 splits into four cyclotomic numbers of order $2e$ that equal 0. Every cyclotomic number of order e that takes the value 1 splits into three cyclotomic numbers of order $2e$ that equal 0 and one cyclotomic number of order $2e$ that equals 1. If $p^r - 1 \equiv 0 \pmod 4$, the unique cyclotomic number of order e that equals $p^r - 2$ splits into one cyclotomic number of order $2e$ that equals $\frac{1}{4}(p^r - 5)$ and three cyclotomic numbers of order $2e$ that equal $\frac{1}{4}(p^r - 1)$. If $p^r - 1 \equiv 2 \pmod 4$, then we obtain $\frac{1}{4}(p^r - 3)$ three times and $\frac{1}{4}(p^r + 1)$ once.

Denote by $n_e(N)$ the number of cyclotomic numbers of order e that equal N. These numbers were given in (4.3). By the above argumentation, we obtain the following values for $n_{2e}(N)$. If $p^r - 1 \equiv 0 \pmod 4$, then

$$\begin{aligned} n_{2e}(0) &= 4n_e(0) + 3n_e(1), & n_{2e}(1) &= n_e(1) \\ n_{2e}((p^r-5)/4) &= n_e(p^r-2), & n_{2e}((p^r-1)/4) &= 3n_e(p^r-2). \end{aligned} \tag{4.5}$$

If $p^r - 1 \equiv 2 \pmod 4$, then

$$\begin{aligned} n_{2e}(0) &= 4n_e(0) + 3n_e(1), & n_{2e}(1) &= n_e(1) \\ n_{2e}((p^r-3)/4) &= 3n_e(p^r-2), & n_{2e}((p^r+1)/4) &= n_e(p^r-2). \end{aligned} \tag{4.6}$$

From Theorem 4.1, it follows that we need to multiply these numbers with $\frac{1}{2}p^{2r}(p^{2r}-1)$ to obtain the multiplicities of the respective block intersection numbers. For the block intersection number 0, we additionally need to add $\frac{1}{2}p^{2r}(2p^{2r}+2)(2p^{2r}+1)$. □

Next, we examine the intersection numbers of $\mathrm{dev}(E^H)$, the design associated to the disjoint difference family E^H in the Galois ring $\mathrm{GR}(p^2, r)$ from Theorem 3.4. Since the design $\mathrm{dev}(E^H)$ is constructed by letting the additive group of $\mathrm{GR}(p^2, r)$ act on the difference family E^H, there is a strong connection between differences and block intersection numbers: Let $E_i^H, E_j^H \in E^H$ be two distinct base blocks of $\mathrm{dev}(E^H)$, and let d be a difference occurring N_d times in the multiset $E_i^H - E_j^H$. Then N_d is the block intersection number $|E_i^H \cap (E_j^H + d)|$ of the blocks E_i^H and $E_j^H + d$ of $\mathrm{dev}(E^H)$. Hence, to calculate the block intersection number $|E_i^H \cap (E_j^H + d)|$ we need to calculate the multiplicity N_d of d in $E_i^H - E_j^H$. We will do exactly this for certain base blocks of $\mathrm{dev}(E^H)$.

Let ξ be a generator of the Teichmüller group $\mathcal{T}^*$ and let $\mathcal{T} = \mathcal{T}^* \cap \{0\}$. As in Theorem 3.4, we denote by $\mathcal{T}_S^*$ the subgroup of Teichmüller squares and by $\mathcal{T}_N^*$ the set of Teichmüller nonsquares. Furthermore, we call a coset of type

$$(1 + p\alpha)\mathcal{T}_S^*,$$

where $\alpha \in \mathcal{T}$, a *square coset of* $\mathcal{T}_S^*$, and a coset of type

$$(1 + p\alpha)\mathcal{T}_N^* = (1 + p\alpha)\xi\mathcal{T}_S^*,$$

where $\alpha \in \mathcal{T}$, a *nonsquare coset of* $\mathcal{T}_S^*$. In the remaining part of this section, we will establish bounds on block intersection numbers of $\mathrm{dev}(E^H)$ that come from the multisets $\Delta\mathcal{T}_S^*$ and $\mathcal{T}_S^* - \mathcal{T}_N^*$. We begin by analyzing the structure of these multisets.

Lemma 4.6. *Let p be an odd prime. Using the same notation as above, consider the multisets $\Delta\mathcal{T}_S^*$ and $\mathcal{T}_S^* - \mathcal{T}_N^*$ in the Galois ring $GR(p^2, r)$.*

- *If $p^r - 1 \equiv 0 \pmod 4$, then $\Delta\mathcal{T}_S^*$ contains $\frac{p^r-5}{4}$ square cosets and $\frac{p^r-1}{4}$ nonsquare cosets of $\mathcal{T}_S^*$, and $\mathcal{T}_S^* - \mathcal{T}_N^*$ contains $\frac{p^r-1}{4}$ square and nonsquare cosets of $\mathcal{T}_S^*$ each.*
- *If $p^r - 1 \equiv 2 \pmod 4$, then $\Delta\mathcal{T}_S^*$ contains $\frac{p^r-3}{4}$ square and nonsquare cosets of $\mathcal{T}_S^*$ each, and $\mathcal{T}_S^* - \mathcal{T}_N^*$ contains $\frac{p^r-3}{4}$ square cosets and $\frac{p^r+1}{4}$ nonsquare cosets of $\mathcal{T}_S^*$.*

Proof. Denote by $\mathcal{I}$ the maximal ideal of $\mathrm{GR}(p^2, r)$. The Teichmüller set $\mathcal{T}$ is a system of representatives of $\mathrm{GR}(p^2, r)/\mathcal{I}$ which is isomorphic to the finite field $\mathbb{F}_{p^r}$. Hence, the sets of Teichmüller squares $\mathcal{T}_S^*$ and Teichmüller nonsquares $\mathcal{T}_N^*$ act in the same way as

the respective sets of squares and nonsquares in $\mathbb{F}_{p^r}$. The result now follows from Theorem 4.3: Let $d = s - s'$ be the difference of two distinct nonzero squares s, s' in $\mathbb{F}_{p^r}$. Equivalently,

$$sd^{-1} - s'd^{-1} = 1.$$

Note that d^{-1} is a square if and only if d is a square. Using the notation from Theorem 4.3, the equation $s - s' = d$ has QQ solutions for s, s' if d is a square, and NN solutions if d is a nonsquare. Analogously, we obtain the number of solutions for s, n of $s - n = d$, where s is a nonzero square and n is a nonsquare in $\mathbb{F}_{p^r}$. □

Furthermore, we need the following properties of squares and nonsquares in the Galois ring $\mathrm{GR}(p^2, r)$.

Proposition 4.7. *Consider the Galois ring $GR(p^2, r)$, where p is odd. Denote by $\mathcal{T}_S^*$ the set of Teichmüller squares and by $\mathcal{T}_N^*$ the set of Teichmüller nonsquares.*

1. *If $p^r - 1 \equiv 0 \pmod 4$, then -1 is a Teichmüller square, and $\mathcal{T}_S^* = -\mathcal{T}_S^*$ and $2\mathcal{T}_S^* \subseteq \Delta\mathcal{T}_S^*$. If $p^r - 1 \equiv 2 \pmod 4$, then -1 is a Teichmüller nonsquare, and $\mathcal{T}_S^* = -\mathcal{T}_N^*$ and $2\mathcal{T}_S^* \subseteq \mathcal{T}_S^* - \mathcal{T}_N^*$.*
2. *If $p^r - 1 \equiv 0 \pmod{12}$, then $1 \in \Delta\mathcal{T}_S^*$, and $\mathcal{T}_S^* \subseteq \Delta\mathcal{T}_S^*$. If $p^r - 1 \equiv 6 \pmod{12}$, then $1 \in \mathcal{T}_N^* - \mathcal{T}_S^*$, and $\mathcal{T}_S^* \subseteq \mathcal{T}_N^* - \mathcal{T}_S^*$.*
3. *If $p^r - 1 \equiv 0$ or $6 \pmod 8$, then 2 is a square, and $2\mathcal{T}_S^*$ is a square coset of $\mathcal{T}_S^*$. If $p^r - 1 \equiv 2$ or $4 \pmod 8$, then 2 is a nonsquare, and $2\mathcal{T}_S^*$ is a nonsquare coset of $\mathcal{T}_S^*$.*

Proof. Let ξ be a generator of the Teichmüller group $\mathcal{T}^*$ in the Galois ring $\mathrm{GR}(p^2, r)$.

1. The present authors [15] proved that if p is odd, -1 is contained in $\mathcal{T}^*$, in particular $-1 = \xi^{\frac{1}{2}(p^r-1)}$. The exponent $\frac{1}{2}(p^r-1)$ is even if $p^r - 1 \equiv 0 \pmod 4$, then -1 is a square in $\mathcal{T}^*$. If $p^r - 1 \equiv 2 \pmod 4$, the exponent $\frac{1}{2}(p^r - 1)$ is odd, hence -1 is a nonsquare in $\mathcal{T}^*$.
2. If $p^r - 1 \equiv 0 \pmod 6$, the equation $x^6 = 1$ has exactly six solutions in the Teichmüller group $\mathcal{T}^*$, namely $\xi^{k(p^r-1)/6}$, where $k \in \{0, 1, \ldots, 5\}$. We show that the sum of these elements is 0. It is easy to see that

$$\xi^{(p^r-1)/6} \sum_{k=0}^{5} \xi^{k(p^r-1)/6} = \sum_{k=0}^{5} \xi^{k(p^r-1)/6}.$$

Hence,

$$(\xi^{(p^r-1)/6} - 1) \sum_{k=0}^{5} \xi^{k(p^r-1)/6} = 0.$$

As we have shown in the proof of Theorem 3.4, the element $\xi^{(p^r-1)/6} - 1$ is a unit. It follows that

$$\sum_{k=0}^{5} \xi^{k(p^r-1)/6} = 0. \tag{4.7}$$

By the same reasoning, $\sum_{k=0}^{2} \xi^{k(p^r-1)/3} = 0$. Consequently, we can rewrite (4.7) as

$$\xi^{5(p^r-1)/6} - \xi^{2(p^r-1)/3} = 1.$$

If $p^r - 1 \equiv 0 \pmod{12}$, the elements $\xi^{5(p^r-1)/6}$ and $\xi^{2(p^r-1)/3}$ are squares and, consequently, $1 \in \Delta\mathcal{T}_S^*$. If $p^r - 1 \equiv 6 \pmod{12}$, then $\xi^{5(p^r-1)/6}$ is a nonsquare and $\xi^{2(p^r-1)/3}$ is a square, hence $1 \in \mathcal{T}_N^* - \mathcal{T}_S^*$.

3. We first consider $r = 1$. Note that $\mathrm{GR}(p^2, 1) = \mathbb{Z}_{p^2}$. The following classical results about quadratic residues were first systematically given by Gauss [13]. An element a relatively prime to an odd prime p is a square in $\mathbb{Z}_{p^m}$ if and only if a is a square in $\mathbb{Z}_p$. In $\mathbb{Z}_p$, the element 2 is a square if $p - 1 \equiv 0$ or $6 \pmod 8$, and 2 is a nonsquare if $p - 1 \equiv 2$ or $4 \pmod 8$. This solves the problem for $r = 1$.
Now, let $r \geq 2$. Let

$$\mathcal{T}_1^* = \{1, \zeta, \zeta^2, \ldots, \zeta^{p-2}\}$$

denote the Teichmüller group of $\mathrm{GR}(p^2, 1)$, and let $\mathcal{T}_1 = \mathcal{T}_1^* \cup \{0\}$. For a fixed prime p, the Galois ring $\mathrm{GR}(p^2, 1)$ is a subring of $\mathrm{GR}(p^2, r)$ for all $r \geq 1$. If $\mathcal{T}^* = \langle \xi \rangle$ denotes the Teichmüller group of $\mathrm{GR}(p^2, r)$, then $\mathcal{T}_1^*$ is a subgroup of $\mathcal{T}^*$ and we write

$$\mathcal{T}_1^* = \{1, \xi^{(p^r-1)/(p-1)}, \xi^{2(p^r-1)/(p-1)}, \ldots, \xi^{(p-2)(p^r-1)/(p-1)}\},$$

where $\zeta^k = \xi^{k(p^r-1)/(p-1)}$. Since 2 is a unit in $\mathrm{GR}(p^2, 1)$, we can write $2 = (1 + p\alpha_0)\alpha_1$ for unique α_0, α_1, where $\alpha_0 \in \mathcal{T}_1$ and $\alpha_1 \in \mathcal{T}_1^*$. It follows that $\alpha_1 = \zeta^\ell$ for some $\ell \in \{0, 1, \ldots, p-2\}$. In $\mathrm{GR}(p^2, r)$, we consequently obtain

$$2 = (1 + p\alpha_0)\xi^{\ell(p^r-1)/(p-1)}.$$

Hence, 2 is a square, and thereby $2\mathcal{T}_S^*$ is a square coset of $\mathcal{T}_S^*$, if at least one of the two numbers ℓ and $(p^r - 1)/(p-1)$ is even. The second number is even if and only if r is even. In this case, $p^r - 1 \equiv 0 \pmod 8$. Hence, if r is odd, the number ℓ needs to be even. This is the case if and only if 2 is a square in $\mathrm{GR}(p^2, 1)$, which, according to the case $r = 1$, holds whenever $p - 1 \equiv 0$ or $6 \pmod 8$. If r is odd, $p^r \equiv p \pmod 8$. The result follows. □

By combining all three results from Theorem 4.7, we obtain the following corollary.

Corollary 4.8. *Consider the Galois ring $GR(p^2, r)$, where p is odd. Denote by $\mathcal{T}_S^*$ the set of Teichmüller squares and by $\mathcal{T}_N^*$ the set of Teichmüller nonsquares.*

- *If $p^r - 1 \equiv 0 \pmod{12}$, then the multiset $\Delta\mathcal{T}_S^*$ contains both $\mathcal{T}_S^*$ and $2\mathcal{T}_S^*$, and the set $2\mathcal{T}_S^*$ is a square coset of $\mathcal{T}_S^*$ if and only if $p^r - 1 \equiv 0 \pmod{24}$.*
- *If $p^r - 1 \equiv 6 \pmod{12}$, then the multiset $\mathcal{T}_S^* - \mathcal{T}_N^*$ contains both $\mathcal{T}_N^*$ and $2\mathcal{T}_S^*$, and the set $2\mathcal{T}_S^*$ is a nonsquare coset of $\mathcal{T}_S^*$ if and only if $p^r - 1 \equiv 18 \pmod{24}$.*

Note that $p^r - 1 \equiv 0 \pmod{24}$ holds whenever the prime $p \geq 5$ and r is even. To continue, we need the following result about when 2 is a Teichmüller square.

Lemma 4.9. *Consider the Galois ring $GR(p^2, r)$, where p is odd. Denote by $\mathcal{T}_S^*$ the set of Teichmüller squares and by $\mathcal{T}_N^*$ the set of Teichmüller nonsquares.*
- *If both $p^r - 1 \equiv 0$ or $6 \pmod 8$ and $2^{p-1} \equiv 1 \pmod{p^2}$, then $\mathcal{T}_S^* = 2\mathcal{T}_S^*$.*
- *If both $p^r - 1 \equiv 2$ or $4 \pmod 8$ and $2^{p-1} \equiv 1 \pmod{p^2}$, then $\mathcal{T}_N^* = 2\mathcal{T}_S^*$.*

Proof. The equation $\mathcal{T}_S^* = 2\mathcal{T}_S^*$ holds if and only if 2 is a square in the Teichmüller group $\mathcal{T}^*$. According to Theorem 4.7, the element 2 is a square in $\mathrm{GR}(p^2, r)$ if and only if $p^r - 1 \equiv 0$ or $6 \pmod 8$. Since $\mathcal{T}^*$ has order $p^r - 1$, the element 2 is contained in $\mathcal{T}^*$ if and only if $2^{p^r-1} \equiv 1 \pmod{p^2}$. Since 2 is also an element of $\mathbb{Z}_{p^2} = \mathrm{GR}(p^2, 1)$ which is a subring of $\mathrm{GR}(p^2, r)$, this condition can be reduced to $2^{p-1} \equiv 1 \pmod{p^2}$.

On the other hand, the equation $\mathcal{T}_N^* = 2\mathcal{T}_S^*$ holds if and only if 2 is a nonsquare and $2 \in \mathcal{T}^*$. The second statement now follows by analogous reasoning as above from Theorem 4.7. □

Primes that solve $2^{p-1} \equiv 1 \pmod{p^2}$ are called *Wieferich primes*. So far, the only known Wieferich primes are 1093 and 3511. Thus, the only known Galois rings of characteristic p^2 satisfying $\mathcal{T}_S^* = 2\mathcal{T}_S^*$ are $\mathrm{GR}(1093^2, r)$, where r is even, and $\mathrm{GR}(3511^2, r)$ for arbitrary r. The only known Galois ring of characteristic p^2 satisfying $\mathcal{T}_N^* = 2\mathcal{T}_S^*$ is $\mathrm{GR}(1093^2, r)$, where r is odd.

With the help of the first result given in Theorem 4.7, we now establish a lower bound on the multiplicities of certain differences of Teichmüller elements. This is an analogue of the present authors' previous result [15, Lemma 5.9] for the design $\mathrm{dev}(E)$ from Theorem 3.3:

Lemma 4.10. *Consider the Galois ring $GR(p^2, r)$, where p is odd. Denote by $\mathcal{T}_S^*$ the set of Teichmüller squares and by $\mathcal{T}_N^*$ the set of Teichmüller nonsquares.*
- *If $p^r - 1 \equiv 0 \pmod 4$, then all differences $d \in \Delta\mathcal{T}_S^*$ where $d \notin 2\mathcal{T}_S^*$ have multiplicity $N_d > 1$ in $\Delta\mathcal{T}_S^*$.*
- *If $p^r - 1 \equiv 2 \pmod 4$, then all differences $d \in \mathcal{T}_S^* - \mathcal{T}_N^*$ where $d \notin 2\mathcal{T}_S^*$ have mulitplicity $N_d > 1$ in $\mathcal{T}_S^* - \mathcal{T}_N^*$.*

Proof. We prove the first result. The proof of the second statement is analogous. Let p be a prime and r be a positive integer such that $p^r - 1 \equiv 0 \pmod 4$. Moreover, let $d \in \Delta\mathcal{T}_S^*$, which means that $d = s - s'$ is the difference of two distinct Teichmüller squares $s, s' \in \mathcal{T}_S^*$. According to Theorem 4.7, $\mathcal{T}_S^* = -\mathcal{T}_S^*$. Hence, if $s' \neq s$, then $(-s') - (-s) = d$ is a second representation of d in $\Delta\mathcal{T}_S^*$. Note that all these differences occur in pairs. If $s' = -s$, however, the two representations are the same, and $d = 2s$, thus $d \in 2\mathcal{T}_S^*$. The statement follows. □

In the following lemma, we will establish an upper bound on the multiplicity of certain differences in $\Delta\mathcal{T}_S^*$ and $\mathcal{T}_S^* - \mathcal{T}_N^*$. For our main theorem, only the first part of

the lemma is relevant. However, we also state the second part as it is easily obtained from the previous results.

Lemma 4.11. *Let p be an odd prime such that $2^{p-1} \not\equiv 1 \pmod{p^2}$. Consider the Galois ring $GR(p^2, r)$, and denote by $\mathcal{T}_S^*$ the set of Teichmüller squares and by $\mathcal{T}_N^*$ the set of Teichmüller nonsquares.*

- *If $p^r - 1 \equiv 0 \pmod{24}$, then all differences $d \in \Delta\mathcal{T}_S^*$ where d is a square have multiplicity $N_d < \frac{p^r-5}{4}$ in $\Delta\mathcal{T}_S^*$*
- *If $p^r - 1 \equiv 18 \pmod{24}$, then all differences $d \in \mathcal{T}_S^* - \mathcal{T}_N^*$ have multiplicity $N_d < \frac{p^r+1}{4}$ in $\mathcal{T}_S^* - \mathcal{T}_N^*$.*

Proof. The condition $2^{p-1} \not\equiv 1 \pmod{p^2}$ ensures that $2 \notin \mathcal{T}^*$, and consequently $\mathcal{T}_S^* \neq 2\mathcal{T}_S^* \neq \mathcal{T}_N^*$ as we showed in Theorem 4.9. Assume $p^r - 1 \equiv 0 \pmod{24}$, and let d be the difference of two Teichmüller squares, $d \in \Delta\mathcal{T}_S^*$, such that d is a square. Denote by N_d the multiplicity of d in $\Delta\mathcal{T}_S^*$. From Theorem 4.6, we know that $\Delta\mathcal{T}_S^*$ contains $\frac{1}{4}(p^r - 5)$ not necessarily distinct square cosets of $\mathcal{T}_S^*$. It follows that $N_d \leq \frac{1}{4}(p^r - 5)$, and $N_d = \frac{1}{4}(p^r - 5)$ if and only if $\Delta\mathcal{T}_S^*$ contains only exactly one square coset of $\mathcal{T}_S^*$ with multiplicity $\frac{1}{4}(p^r - 5)$. Assume $N_d = \frac{1}{4}(p^r - 5)$. If $p^r - 1 \equiv 0 \pmod{24}$, then according to Theorem 4.8, both $\mathcal{T}_S^*$ and $2\mathcal{T}_S^*$ are subsets of $\Delta\mathcal{T}_S^*$, and $2\mathcal{T}_S^*$ is a square coset of $\mathcal{T}_S^*$. This is a contradiction.

Now, assume $p^r - 1 \equiv 18 \pmod{24}$, and let d be the difference of a Teichmüller square and a Teichmüller nonsquare, $d \in \mathcal{T}_S^* - \mathcal{T}_N^*$, such that d is a nonsquare. Denote by N_d the multiplicity of d in $\mathcal{T}_S^* - \mathcal{T}_N^*$. Analogously to above, we conclude from Theorem 4.6 that $N_d \leq \frac{1}{4}(p^r + 1)$, and $N_d = \frac{1}{4}(p^r + 1)$ if and only if $\mathcal{T}_S^* - \mathcal{T}_N^*$ contains only exactly one nonsquare coset of $\mathcal{T}_S^*$. Assume $N_d = \frac{1}{4}(p^r + 1)$. If $p^r - 1 \equiv 18 \pmod{24}$, then both $\mathcal{T}_N^*$ and $2\mathcal{T}_S^*$ are contained in $\mathcal{T}_S^* - \mathcal{T}_N^*$, and $2\mathcal{T}_S^*$ is a nonsquare coset of $\mathcal{T}_S^*$. Again, we obtain a contradiction. □

As we have mentioned above, the multiplicity of a difference d in $\Delta\mathcal{T}_S^*$ corresponds directly to the block intersection number $|\mathcal{T}_S^* \cap (\mathcal{T}_S^* \cap d)|$. Hence, we obtain from the previous lemmas the following theorem which is our main theorem.

Theorem 4.12. *Let p be an odd prime such that $2^{p-1} \not\equiv 1 \pmod{p^2}$. Let C^H be a $(p^{2r}, \frac{p^r-1}{2}, \frac{p^r-3}{2})$ disjoint difference family in the additive group of the finite field $\mathbb{F}_{p^{2r}}$ constructed with Theorem 3.2, and let E^H be a disjoint difference family with the same parameters in the additive group of the Galois ring $GR(p^2, r)$ constructed with Theorem 3.4. If $p^r - 1 \equiv 0 \pmod{24}$, the 2-$(p^{2r}, \frac{p^r-1}{2}, \frac{p^r-3}{2})$ designs $\mathrm{dev}(E^H)$ and $\mathrm{dev}(C^H)$ are nonisomorphic.*

Proof. Let p be an odd prime and r be an integer such that $p^r - 1 \equiv 0 \pmod{24}$. Recall from Theorem 4.5 that in this case the block intersection numbers of our design $\mathrm{dev}(C^H)$ are given as $0, 1, \frac{1}{4}(p^r - 5), \frac{1}{4}(p^r - 1)$. Now, consider the Galois ring $\mathrm{GR}(p^2, r)$, and denote by $\mathcal{T}_S^*$ the set of Teichmüller squares and by $\mathcal{T}_N^*$ the set of Teichmüller

nonsquares. By combining Theorem 4.10 and Theorem 4.11, we obtain

$$1 < \left|\mathcal{T}_S^* \cap (\mathcal{T}_S^* + d)\right| < \frac{p^r - 5}{4}$$

for all squares $d \in \Delta\mathcal{T}_S^* \setminus 2\mathcal{T}_S^*$. Consequently, the design $\mathrm{dev}(E^H)$ has an intersection number different from the ones of $\mathrm{dev}(C^H)$, and the designs are nonisomorphic. □

We remark that $p^r - 1 \equiv 0 \pmod{24}$ holds for all p and r where $p \geq 5$ and r is even. Furthermore, we remark that for the Wieferich primes 1093 and 3511, that satisfy $2^{p-1} \equiv 1 \pmod{p^2}$, the designs $\mathrm{dev}(C^H)$ and $\mathrm{dev}(E^H)$ are nonisomorphic for all $r > 1$. With the help of `Magma` [4], we computed the multisets $\Delta\mathcal{T}_S^*$ and $\mathcal{T}_S^* - \mathcal{T}_N^*$ for $r = 1$ and checked that these multisets contain more than one square coset and more than two nonsquare coset of $\mathcal{T}_S^*$ each. Consequently, there will be at least as many square and nonsquare cosets of $\mathcal{T}_S^*$ in $\Delta\mathcal{T}_S^*$ and $\mathcal{T}_S^* - \mathcal{T}_N^*$, respectively, for $r > 1$. So, the bounds on the respective block intersection numbers established in the previous proof hold.

To conclude this section, we give an example that demonstrates why our block intersection number approach fails if $p^r - 1 \not\equiv 0 \pmod{24}$. We choose as an example the case $p^r - 1 \equiv 18 \pmod{24}$ since, in the previous lemmas, we have already obtained several results about this case that followed immediately from the results for $p^r - 1 \not\equiv 0 \pmod{24}$.

Example 3. Let p and r such that $p^r - 1 \equiv 18 \pmod{24}$. In this case, according to Theorem 4.5, the design $\mathrm{dev}(C^H)$, has block intersection numbers 0, 1, $\frac{1}{4}(p^r - 3)$, $\frac{1}{4}(p^r + 1)$. For the design $\mathrm{dev}(E^H)$, using Theorem 4.10 and Theorem 4.11, we obtain

$$1 < \left|\mathcal{T}_S^* \cap (\mathcal{T}_N^* + d)\right| < \frac{p^r + 1}{4}$$

for all $d \in (\mathcal{T}_S^* - \mathcal{T}_N^*) \setminus 2\mathcal{T}_S^*$. However, this result is of little use as it is still possible that there exists $d \in (\mathcal{T}_S^* - \mathcal{T}_N^*) \setminus 2\mathcal{T}_S^*$ such that $|\mathcal{T}_S^* \cap (\mathcal{T}_N^* + d)| = \frac{1}{4}(p^r - 3)$ or that two completely different blocks intersect in $\frac{1}{4}(p^r + 1)$ elements. In fact, the multiset $p\mathcal{T}_S^* - p\mathcal{T}_N^*$ contains $\frac{1}{4}(p^r + 1)$ times the set $p\mathcal{T}_N^*$ and $\frac{1}{4}(p^r - 3)$ times the set $p\mathcal{T}_S^*$. Hence, these two numbers actually occur as the block intersection numbers

$$\left|p\mathcal{T}_S^* \cap (p\mathcal{T}_N^* + d)\right|,$$

where $d \in \mathcal{I} \setminus \{0\}$. Hence, we cannot show the existence of an intersection number N such that $1 < N < \frac{1}{4}(p^r - 3)$.

5 Conclusion and open questions

Motivated by the present authors' [15] recent results, we tried to use the same technique to solve another isomorphism problem about 2-$(v, k, k-1)$ designs. Thanks to

the algebraic structure of our designs, we were able to solve the problem for many cases, and, in doing so, obtained some interesting results about cyclotomic numbers and the structure of Galois rings of characteristic p^2. But the isomorphism problem is still not solved for all cases.

Our results demonstrate that using block intersection numbers as a method to tackle isomorphism problems about combinatorial designs has its limitations. One needs designs that have a sufficiently strong algebraic structure to calculate or even bound these numbers. We still consider this approach promising, especially if the designs are constructed as the developments of some difference structures. During our studies, we discovered the following interesting open problems:

- Our computations hint that Theorem 4.12 holds for all p and r, where p is odd. However, our examination of intersection numbers did not lead to the results necessary to prove this conjecture. We leave this task to future work.
- The construction of a disjoint difference family in $\mathrm{GR}(p^2, r)$ presented in Theorem 3.4 does not only work for the subgroup of squares in the Teichmüller group but for all its subgroups. Moreover, there will always be an analogue in $\mathbb{F}_{p^{2r}}$. It would be interesting to study the isomorphism problem for the associated designs in all these cases. It might be possible to deduce more block intersection numbers from the ones given in this paper and in [15].
- As mentioned before, nonisomorphic designs can have the same block intersection numbers. It would be interesting to find more difference families as in Remark 1 for which their associated designs have the same intersection numbers but are still nonisomorphic.

Bibliography

[1] R. J. R. Abel and M. Buratti, Difference families, In: Handbook of Combinatorial Designs, Ed. by C. J. Colbourn and J. H. Dinitz. 2nd ed. Boca Raton: Chapman & Hall/CRC Press, 2007, pp. 392–410.

[2] L. D. Baumert, W. H. Mills, and R. L. Ward, Uniform cyclotomy, *J. Number Theory*, **14**(1) (1982), 67–82.

[3] T. Beth, D. Jungnickel, and H. Lenz, Design Theory, 2nd ed., Cambridge University Press, Cambridge, 1999.

[4] W. Bosma, J. Cannon, and C. Playoust, The Magma algebra system. I. The user language, *J. Symb. Comput.*, **24**(3–4) (1997), 235–265.

[5] M. Buratti, On disjoint $(v, k, k-1)$ difference families, *Des. Codes Cryptogr.*, Special Issue: Finite Geometries (2018).

[6] M. Buratti and D. Jungnickel, Partitioned difference families versus zero-difference balanced functions, *Des. Codes Cryptogr.*, (2019), 1–7.

[7] Y. Chang and C. Ding, Constructions of external difference families and disjoint difference families, *Des. Codes Cryptogr.*, **40**(2) (2006), 167–185.

[8] C. J. Colbourn and J. H. Dinitz, eds. Handbook of Combinatorial Designs, 2nd ed., Chapman & Hall/CRC Press, Boca Raton, 2007.

[9] J. A. Davis, S. Huczynska, and G. L. Mullen, Near-complete external difference families, *Des. Codes Cryptogr.*, **84** (2017), 415–424.
[10] L. E. Dickson, Linear Groups: With an Exposition of the Galois Field Theory, Dover Publications, New York, 1958.
[11] T. Feng and Q. Xiang, Cyclotomic constructions of skew Hadamard difference sets, *J. Comb. Theory, Ser. A*, **119**(1) (2012), 245–256.
[12] S. Furino, Difference families from rings, *Discrete Math.*, **97**(1–3) (1991), 177–190.
[13] C. F. Gauss, Untersuchungen über höhere Arithmetik, Trans. by H. Maser, Chelsea Publishing Company, New York, 1981.
[14] J. Jedwab and S. Li, Construction and nonexistence of strong external difference families, *J. Algebraic Comb.*, **49**(1) (2019), 21–48.
[15] C. Kaspers and A. Pott, Solving isomorphism problems about 2-designs from disjoint difference families, *J. Comb. Des.*, **27**(5) (2019), 277–294.
[16] B. R. McDonald, Finite Rings with Identity, Marcel Dekker, Inc., New York, 1974.
[17] K. Momihara, Disjoint difference families from Galois rings, *Electron. J. Comb.*, **24**(3) (2017), 3.23.
[18] S.-L. Ng and M. B. Paterson, Disjoint difference families and their applications, *Des. Codes Cryptogr.*, **78**(1) (2016), 103–127.
[19] T. Ralston, On the distribution of squares in a finite field, *Geom. Dedic.*, **8**(2) (1979), 207–212.
[20] Z.-X. Wan, Lectures on Finite Fields and Galois Rings, World Scientific, Singapore, 2003.
[21] R. M. Wilson, Cyclotomy and difference families in elementary Abelian groups, *J. Number Theory*, **4**(1) (1972), 17–47.
[22] J. Zwanzger, Computergestützte Suche nach optimalen linearen Codes über endlichen Kettenringen unter Verwendung heuristischer Methoden, PhD thesis, University of Bayreuth, Bayreuth, 2011.

Yuta Kodera, Sylvain Duquesne, and Yasuyuki Nogami

Multiplication and squaring in cubic and quartic extensions for pairing based cryptography

Abstract: This paper proposes multiplication and squaring algorithms over cubic and quartic extensions for implementing pairing-friendly fields as the tower of fields efficiently. The algorithms are designed by introducing the Gauss periods of type (h, m) for constructing the normal basis. The costs are compared with several well-known methods such as the Karatsuba method, and as a result, it is found that the proposed methods work with smaller or almost equal costs than the conventional methods.

Keywords: Gauss periods, normal bases, vector arithmetic, pairing-friendly fields, cubic and quartic extensions

MSC 2010: 11T71

1 Introduction

Many protocol researchers have begun to employ pairing based cryptography [20] for the constructions of functional cryptographic applications such as homomorphic encryption [22], identity based cryptography [4], broadcast encryption [6], and short signature schemes [5], for example. On the other hand, since the fundamental arithmetic of pairing based cryptography is handled in extension fields such as $\mathbb{F}_{p^{12}}$, $\mathbb{F}_{p^{16}}$, or $\mathbb{F}_{p^{18}}$ for example, the performance of the pairing-based cryptography depends on the efficiency of arithmetic in such extension fields. Typically, these extension fields can be constructed as towers of fields, and Koblitz et al. [18] defined a field $\mathbb{F}_{p^k}$ as being pairing-friendly if conditions $p \equiv 1 \pmod{12}$ and $k = 2^i 3^j (i, j \geq 0)$ are satisfied. It is known that the pairing-friendly fields are suitable for implementing cryptographic bilinear pairings.

Among the researches surrounding the pairing, reduction of the costs for the base field arithmetic is one of the nonnegligible factors of efficient implementations. For example, the Karatsuba method [15, 17] is the first multiplication algorithm which works with smaller costs than the schoolbook method (distributive law) by reusing the com-

Acknowledgement: This work was partly supported by the JSPS Research Fellowships for Young Scientists KAKENHI 19J1179411.

Yuta Kodera, Yasuyuki Nogami, Okayama University, Japan, e-mail: p7ye9j8w@s.okayama-u.ac.jp
Sylvain Duquesne, Université Rennes I, France

https://doi.org/10.1515/9783110621730-006

mon term of the equation. As the natural generalization of the Karatsuba method, the Toom–Cook method [25, 7, 17] is the well-known algorithm for large integer arithmetic. However, since these algorithms are mainly designed for polynomial bases, a modular polynomial for defining the extension field should be carefully chosen so that the algorithm can perform adequately and efficiently. Such parameter selection is one of the attractive topics for implementing the pairing-based cryptography, and the readers can obtain further knowledge in [9].

As well as the algorithms for polynomial bases, there are several types of research and algorithms for efficient field arithmetic based on normal bases. For example, Mullin et al. have proposed type I and II optimal normal bases in [21], and the concept has been extended by Kato et al. in [16]. Furthermore, Nekado has generalized the concept of the algorithms in his Ph. D. thesis [23] by focusing on the so-called Gauss periods of type (h, m) [10, 12].

In this paper, the authors focus on so-called cyclic vector multiplication algorithm (CVMA), which was originally proposed by Nogami et al. in [24] for efficient vector arithmetic, and generalized by Nekado in [23]. Since this paper aims to improve the algorithms for the pairing-based cryptography, multiplication, and squaring over cubic extension and quartic extensions are focused on, in particular. More precisely, the authors review the structure of the CVMA with the Gauss periods of type (h, m), and the vector arithmetic over cubic and quartic extensions are optimized by fixing the parameters h and m.

The proposed methods are compared with several well-known methods. As a consequence, it is found that the costs of the proposed method are slightly lower than that of the conventional methods. Though these methods are assumed to be used as the base of the tower field, nontrivial chances will arise for selecting a suitable basis for implementation from the following reasons:

1. In practice, pairings of higher security level use relatively higher degree extension fields such as $\mathbb{F}_{p^{18}}$ [14] and $\mathbb{F}_{p^{24}}$ [2], for example. In that case, these extension fields are constructed as a tower of fields so that rational points defined over the fields can be considered as points over $\mathbb{F}_{p^3}$ and $\mathbb{F}_{p^4}$ via sextic twist,[1] respectively. In other words, the base field arithmetic has a great influence on the performance of an implementation.
2. In addition, though the operations over the defining field such as $\mathbb{F}_{p^{18}}$ and $\mathbb{F}_{p^{24}}$ are heavy, the improvements of their base field arithmetics bring opportunities to be used as a new basis representation for these fields which can provide faster arithmetic than the classical methods. In fact, the cost of multiplication in $\mathbb{F}_{p^{18}}$ based on the CVMA is smaller than the Karatsuba method (see Section 5).

1 Let E and E' be elliptic curves over $\mathbb{F}_p$. E' is called a *twist* of degree d of E if there exists an isomorphism $\psi_d : E' \to E$ defined over $\mathbb{F}_{p^d}$ and d is minimal. The possible degree of twist is $d = 1, 2, 3, 4$, or 6, and sextic twist is the twist of degree $d = 6$. The readers can refer to [3, 13, 20] for the detail.

3. Since the condition required for both conventional and proposed methods do not overlap completely, the proposed methods provide novel prime numbers for curves.

Considering further applicability of the proposed methods, the authors would like to expand the algorithms for general prime powers, and advanced applications including such expansion are future works.

2 Preliminaries

This section briefly introduces the notation and reviews the fundamentals of finite fields [19].

2.1 Notation

For a prime number p and a positive integer m, let $\mathbb{F}_p$ and $\mathbb{F}_{p^m}$ be a prime field and its extension field, respectively. An element $\boldsymbol{a} = (a_0, \ldots, a_{m-1}) \in \mathbb{F}_{p^m}$ with a basis $\{\beta_0, \ldots, \beta_{m-1}\}$ is represented by the polynomial form as follows:

$$\boldsymbol{a} = \sum_{i=0}^{m-1} a_i \beta_i.$$

For simplicity of comparisons, the authors evaluate the costs of operations by focusing on addition (A), multiplication (M), squaring (S), multiplication with a constant integer (c), and division by 2 (D_2) in what follows. For examples, the costs of $(xy+2x+y^2)$ and $x^3 + xy^2 + \sigma y$ are given by M + S + 3A and 2M + 2S + c + 2A, respectively, where σ is a constant in $\mathbb{F}_p$. Another example is $\frac{a_0 b_0^2}{2} + a_1(a_1 - b_1)\alpha + 2\sigma a_3 \alpha^2$ with the basis $\{1, \alpha, \alpha^2\}$. The readers need to remember that an equation over $\mathbb{F}_{p^m} \supset \mathbb{F}_p$ is evaluated by the number of required operations in $\mathbb{F}_p$. Therefore, $\frac{a_0 b_0^2}{2} + a_1(a_1 - b_1)\alpha + 2\sigma a_3 \alpha^2$ is granted as a vector $(\frac{a_0 b_0^2}{2}, a_1(a_1 - b_1), 2\sigma a_3)$ and is evaluated as 2M + S + 2A + c + D_2.

2.2 Gauss periods and cyclotomic polynomial

The Gauss periods are often used to construct normal bases in finite fields. The mathematical structure and a systematic construction have been studied by Feisel et al. and Gao in their works [10, 12]. In this paper, since the authors concentrate to improve arithmetic over $\mathbb{F}_{p^3}$ and $\mathbb{F}_{p^4}$ for prime integers p, pairing-friendly fields $\mathbb{F}_{p^k}$ are assumed to be constructed as the extension of $\mathbb{F}_{p^m}$ with classical polynomial bases.

The Gauss periods of type (h, m) are obtained in the following way. For a prime number p, choose positive integers h, m satisfying the conditions [1, 11] below:

Condition 1.

1. $r = hm + 1$ is a prime number not equal to p;
2. $\gcd(hm/e, m) = 1$, where e is the multiplicative order of p in $\mathbb{F}_r$.

According to the Fermat's little theorem, we have $r \mid (p^{r-1} - 1)$. In addition, since $\mathbb{F}_r^\times$ forms a multiplicative group consisting of hm elements, there exists a primitive rth root of unity in $\mathbb{F}_{p^{hm}}$, which is denoted by $\beta \in \mathbb{F}_{p^{hm}}$. In other words, β is a zero of the cyclotomic polynomial $\Phi_r(x)$. Recall that since r is a prime and we have $\beta^r = 1$, the following equation holds:

$$\beta^r - 1 = (\beta - 1)\sum_{j=0}^{r-1} \beta^j = 0.$$

Therefore, we have $\sum_{j=0}^{r-1} \beta^j = 0$ ($\because \beta \neq 1$).

Let $\mathcal{K}$ be the unique subgroup of $\mathbb{F}_r^\times$ with $\#\mathcal{K} = h$. Then define $\alpha \in \mathbb{F}_r$ with a coset $p^i\mathcal{K}$ as follows:

$$\alpha^{p^i} = \sum_{j \in \mathcal{K}} \beta^{p^i j}. \tag{2.1}$$

This α^{p^i} is called a prime Gauss period, which becomes a normal element in $\mathbb{F}_{p^m}$ under the condition (Condition 1), and the set of conjugates $\{\alpha, \alpha^p, \ldots, \alpha^{p^{m-1}}\}$ forms a normal basis of $\mathbb{F}_{p^m}$.

2.3 Cyclic vector multiplication algorithm

The cyclic vector multiplication algorithm (CVMA) was proposed by Nogami et al. in [24] for efficient arithmetic over extension fields. Since the CVMA is designed to work with the normal basis constructed by the Gauss periods of type (h, m), the basis of $\mathbb{F}_{p^m}$ is given by $\{\alpha, \alpha^p, \ldots, \alpha^{p^{m-1}}\}$, where α^{p^i} is a normal element as defined in equation (2.1). For a primitive hth root $d \in \mathbb{F}_r$, let $\epsilon(\cdot)$ and $\eta(\cdot)$ be functions defined as follows:

$$\epsilon(p^s d^t \pmod r) = \begin{cases} m & \text{if } p^s d^t \pmod r = 0, \\ s & \text{otherwise,} \end{cases} \tag{2.2}$$

$$\eta(s, t, u) = \epsilon(p^{s'} d^{t'} (= p^s + p^t d^u \pmod r)). \tag{2.3}$$

It is noted that a basis element α^{p^i} $(0 \leq i < m)$ is represented by β such that $\Phi_r(\beta) = 0$ as equation (2.4), and multiplication $\boldsymbol{x} \times \boldsymbol{y}$, where $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}_{p^m}$, is derived by equation (2.5) as follows:

$$\alpha^{p^i} = \left(\sum_{k=0}^{h-1} \beta^{d^k}\right)^{p^i} \in \mathbb{F}_{p^m}, \tag{2.4}$$

$$\boldsymbol{x} \times \boldsymbol{y} = \left(\sum_{i=0}^{m-1} x_i \alpha^{p^i}\right)\left(\sum_{i=0}^{m-1} y_i \alpha^{p^i}\right) = \sum_{i=0}^{m-1}\sum_{j=0}^{m-1} x_i y_j \alpha^{p^i+p^j}, \tag{2.5}$$

where $x_i, y_i \in \mathbb{F}_p$. The functions $\epsilon(\cdot)$ and $\eta(\cdot)$ are used for determining the exponent part of the basis element truncated by $\Phi_r(\beta)$ from $\alpha^{p^i+p^j}$. The readers can know the derivation of the equation (2.6) and equation (2.7) by referring to [23].

Let $\boldsymbol{a} = (a_0, a_1, \ldots, a_{m-1})$ and $\boldsymbol{b} = (b_0, b_1, \ldots, b_{m-1})$ be vectors in $\mathbb{F}_{p^m}$ respectively. Let v_l be a variable defined in equation (2.6) by using equations (2.2), (2.3).

$$v_l = \sum_{0 \le s < t < m} (a_s - a_t)(b_s - b_t) \sum_{u=0}^{h-1} \delta_l(\eta(s,t,u)), \tag{2.6}$$

where $\delta_i(j)$ is the Kronecker delta, which is defined as follows:

$$\delta_i(j) = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{otherwise.} \end{cases}$$

Then the lth coefficient of the vector $\boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b}$ is derived by equation (2.7), where $\boldsymbol{c} = (c_0, c_1, \ldots, c_{m-1}) \in \mathbb{F}_{p^m}^{\times}$.

$$c_l = \begin{cases} hv_m - v_l - a_l b_l & \text{if } h \text{ is odd}, \\ -v_l - a_l b_l & \text{otherwise.} \end{cases} \tag{2.7}$$

Based on equation (2.7), the CVMA for the Gauss periods of type (h, m) is obtained as shown in Algorithms 1, 2.

Algorithm 1 Precomputation steps for the CVMA.

Require: A prime number p and an extension degree m.

Ensure: $\eta[s,t,u]$.

```
Find h, and prepare a primitive hth root d of unity in F_r^×.
ε[0] ← m.
for s = 0 to m − 1 do
    for t = 0 to h − 1 do
        ε[p^s d^t (mod r)] ← s.
    end for
end for
for s = 0 to m − 2 do
    for t = s + 1 to m − 1 do
        for u = 0 to h − 1 do
            η[s, t, u] ← ε[p^s + p^t d^u (mod r)].
        end for
    end for
end for
```

Algorithm 2 The CVMA for the Gauss periods of type (h, m).

Require: $\boldsymbol{a} = \sum_{s=0}^{m-1} a_s \alpha^{p^s}, \boldsymbol{b} = \sum_{s=0}^{m-1} b_s \alpha^{p^s} \in \mathbb{F}_{p^m}^{\times}$ and the precomputed values $\eta[s, t, u]$ for $0 \le s < t < m$ and $0 \le u < h$.

Ensure: $\boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b} = \sum_{s=0}^{m-1} c_s \alpha^{p^s}$.

```
for l = 0 to m − 2 do
    v_l ← a_l b_l.
end for
v_m ← 0.
for s = 0 to m − 2 do
    for t = s + 1 to m − 1 do
        w ← (a_s − a_t)(b_s − b_t).
        for u = 0 to h − 1 do
            v_η(s,t,u) ← v_η(s,t,u) + w.
        end for
    end for
end for
if h is an odd then
    w ← h v_m.
    for l = 0 to m − 1 do
        c_l ← w − v_l.
    end for
else
    for l = 0 to m − 1 do
        c_l ← −v_l.
    end for
end if
```

Recall that $r = hm + 1$ must be a prime according to the condition (Condition 1), and note that h decides the number of elements of $\mathcal{K}$, which is the unique subgroup for defining a normal element. Thus, this paper targets cubic and quartic extensions with the Gauss period of type $(2, 3)$ and $(1, 4)$ cases, in particular, so as to simplify the calculation and to minimize the additional costs. It is noted that though there are other choices of h and m for establishing the Gauss periods, larger h and m would cause the increase of the number of components of a basis element and as the result, the multiplication with such basis becomes more complex. For example, consider constructing a cubic extension with $(h, m) = (4, 3)$ ($\because r = hm + 1$ has to be a prime). In that case, the cyclotomic polynomial $\Phi_{13}(x)$ is used and a normal basis $\{\tau_1, \tau_2, \tau_3\}$ is given by

$$\{\tau_1, \tau_2, \tau_3\} = \{\beta + \beta^5 + \beta^8 + \beta^{12}, \beta^2 + \beta^3 + \beta^{10} + \beta^{11}, \beta^4 + \beta^6 + \beta^7 + \beta^9\},$$

where $\Phi_{13}(\beta) = 0$. Comparing to the case for using $(h, m) = (2, 3)$, whose a normal basis is given by $\{\tau_1, \tau_2, \tau_3\} = \{\beta + \beta^6, \beta^2 + \beta^5, \beta^3 + \beta^4\}$ for $\Phi_7(\beta) = 0$, one can confirm that the

basis becomes complex and as the result, it increases the cost of additions over $\mathbb{F}_p$. Therefore, h and m should be chosen smaller as possible, and the cyclotomic polynomials $\Phi_7(x)$ and $\Phi_5(x)$ are used for constructing the cubic extension and the quartic extension, respectively. The goal of this paper is to reduce the costs of Algorithm 2 in the particular case of degree 3 and 4.

3 Arithmetic over a cubic extension with the Gauss periods

This section reviews the algorithms and the costs of each operation over the cubic extension.

3.1 The normal basis of the cubic extension

Let $r = 7$ and β be such that $\Phi_7(\beta) = \frac{\beta^7-1}{\beta-1} = \sum_{i=0}^{6} \beta^i = 0$. The condition on a prime p for constructing the Gauss periods of type $(2, 3)$ is given by:

Condition 2. The Gauss periods of type $(2, 3)$ form a normal basis of $\mathbb{F}_{p^3}$ if $p \neq 7$ and $p \not\equiv 1, 6 \pmod 7$.

Since the unique subgroup of order 2 in $\mathbb{F}_7$ is $\{1, -1\}$, a normal element can be given by $\beta + \beta^6$ and the normal basis used for $\mathbb{F}_{p^3}$ in what follows is of the form

$$\{\tau_1, \tau_2, \tau_3\} = \{\beta + \beta^{-1}, \beta^2 + \beta^{-2}, \beta^3 + \beta^{-3}\}$$

be a basis. In addition, we have $-1 = \tau_1 + \tau_2 + \tau_3$ since $\Phi_7(\beta) = 0$.

3.2 Multiplication

Let $\boldsymbol{a} = (a_0, a_1, a_2), \boldsymbol{b} = (b_0, b_1, b_2)$ be vectors in $\mathbb{F}_{p^3}$, and $\{\tau_1, \tau_2, \tau_3\}$ the basis of $\mathbb{F}_{p^3}$. Since the basis is given by $\{\tau_1, \tau_2, \tau_3\} = \{\beta + \beta^{-1}, \beta^2 + \beta^{-2}, \beta^3 + \beta^{-3}\}$, τ_1, τ_2, and τ_3 hold the following relation:

$$\begin{aligned}
\tau_1^2 &= (\beta + \beta^{-1})^2 = \beta^2 + \beta^{-2} + 2 = \tau_2 + 2,\\
\tau_2^2 &= (\beta^2 + \beta^{-2})^2 = \beta^4 + \beta^{-4} + 2 = \tau_3 + 2,\\
\tau_3^2 &= (\beta^3 + \beta^{-3})^2 = \beta + \beta^{-1} + 2 = \tau_1 + 2,\\
\tau_1\tau_2 &= (\beta + \beta^{-1})(\beta^2 + \beta^{-2}) = \beta + \beta^{-1} + \beta^3 + \beta^{-3} = \tau_1 + \tau_3,\\
\tau_2\tau_3 &= (\beta^2 + \beta^{-2})(\beta^3 + \beta^{-3}) = \beta + \beta^{-1} + \beta^5 + \beta^{-5} = \tau_1 + \tau_2,\\
\tau_3\tau_1 &= (\beta^3 + \beta^{-3})(\beta + \beta^{-1}) = \beta^2 + \beta^{-2} + \beta^4 + \beta^{-4} = \tau_2 + \tau_3.
\end{aligned}$$

It is noted that β satisfies $\beta^7 = 1$ here. By using the above relationships, $\boldsymbol{a} \times \boldsymbol{b}$ can be expanded as follows:

$$\begin{aligned}
\boldsymbol{a} \times \boldsymbol{b} &= (a_0\tau_1 + a_1\tau_2 + a_2\tau_3)(b_0\tau_1 + b_1\tau_2 + b_2\tau_3) \\
&= (a_0b_0\tau_1^2 + a_0b_1\tau_1\tau_2 + a_0b_2\tau_1\tau_3) + (a_1b_0\tau_2\tau_1 + a_1b_1\tau_2^2 + a_1b_2\tau_2\tau_3) \\
&\quad + (a_2b_0\tau_3\tau_1 + a_2b_1\tau_3\tau_2 + a_2b_2\tau_3^2) \\
&= \{a_0b_1 + a_1b_0 + a_1b_2 + a_2b_1 + a_2b_2 - 2(a_0b_0 + a_1b_1 + a_2b_2)\}\tau_1 \\
&\quad + \{a_0b_2 + a_2b_0 + a_1b_2 + a_2b_1 + a_0b_0 - 2(a_0b_0 + a_1b_1 + a_2b_2)\}\tau_2 \\
&\quad + \{a_0b_1 + a_1b_0 + a_0b_2 + a_2b_0 + a_1b_1 - 2(a_0b_0 + a_1b_1 + a_2b_2)\}\tau_3
\end{aligned}$$

By focusing on the symmetricity of coefficients and using Karatsuba-like trick,[2] we have the below equivalent representation with respect for each coefficient of $\boldsymbol{a} \times \boldsymbol{b}$. As the result, additions decrease and the cost required for calculating all coefficients of $\boldsymbol{a} \times \boldsymbol{b}$ is then 6M + 12A in total.

$$\begin{aligned}
\text{Coefficient of } \tau_1: \quad & a_0b_1 + a_1b_0 + a_1b_2 + a_2b_1 + a_2b_2 - 2(a_0b_0 + a_1b_1 + a_2b_2) \\
& = (a_0 - a_1)(b_1 - b_0) + (a_1 - a_2)(b_2 - b_1) - a_0b_0, \qquad (3.1) \\
\text{Coefficient of } \tau_2: \quad & a_0b_2 + a_2b_0 + a_1b_2 + a_2b_1 + a_0b_0 - 2(a_0b_0 + a_1b_1 + a_2b_2) \\
& = (a_1 - a_2)(b_2 - b_1) + (a_0 - a_2)(b_2 - b_0) - a_1b_1, \qquad (3.2) \\
\text{Coefficient of } \tau_3: \quad & a_0b_1 + a_1b_0 + a_0b_2 + a_2b_0 + a_1b_1 - 2(a_0b_0 + a_1b_1 + a_2b_2) \\
& = (a_0 - a_2)(b_2 - b_0) + (a_0 - a_1)(b_1 - b_0) - a_2b_2. \qquad (3.3)
\end{aligned}$$

3.3 Squaring

According to equations (3.1), (3.2), (3.3), each coefficient of $\boldsymbol{a}^2 = \boldsymbol{a} \times \boldsymbol{a}$ is calculated by

$$\begin{aligned}
\text{Coefficient of } \tau_1: \quad & -(a_0 - a_1)^2 - (a_1 - a_2)^2 - a_0^2, \\
\text{Coefficient of } \tau_2: \quad & -(a_0 - a_2)^2 - (a_1 - a_2)^2 - a_1^2, \\
\text{Coefficient of } \tau_3: \quad & -(a_0 - a_2)^2 - (a_0 - a_1)^2 - a_2^2.
\end{aligned}$$

Since $\tau_1 + \tau_2 + \tau_3 = -1$, the bases $\{\tau_1, \tau_2, \tau_3\}$ and $\{1, \tau_1, \tau_2\}$ are flexibly convertible. Therefore, the authors consider the further transformations of each coefficient by intentionally applying the basis conversion for the sake of breaking the symmetrical form.

Let (t_0, t_1, t_2) be the coefficient of the vector which is represented by the basis $\{1, \tau_1, \tau_2\}$. Then, the coefficients of $\boldsymbol{a}^2$ are derived as follows:

$$t_0 = 2a_0^2 + a_1^2 + 2a_2^2 - 2a_0a_1 - 2a_0a_2,$$

2 It means that we attempt to reuse values as many as possible for computing the desired terms in the same manner as the Karatsuba method.

$$t_1 = a_2^2 - a_1^2 + 2a_1a_2 - 2a_0a_2,$$
$$t_2 = a_0^2 - a_1^2 + 2a_1a_2 - 2a_0a_1.$$

By focusing on the form of t_0, it is found that t_0 can be represented as the sum of three different squares, that is,

$$t_0 = (a_0 - a_1)^2 + (a_0 - a_2)^2 + a_2^2.$$

In addition, we can rewrite the formulas giving τ_1 and τ_2 so that $(a_0 - a_1)^2$ and $(a_0 - a_2)^2$ are also used for their computation so that their costs is lower. We get

$$t_1 = (a_0 - a_2)^2 - (a_0 - a_1)^2 - 2a_1(a_0 - a_2),$$
$$t_2 = (a_0 - a_1)^2 - 2a_1(a_1 - a_2).$$

Considering the fact that (t_0, t_1, t_2) is a coefficient vector of the basis $\{1, \tau_1, \tau_2\}$, the coefficient of $\boldsymbol{a}^2$ with the basis $\{\tau_1, \tau_2, \tau_3\}$ can be derived by $(t_1 - t_0, t_2 - t_0, -t_0)$. Therefore, each coefficient of $\boldsymbol{a}^2$ is calculated with pre-computed temporary values T_1 to T_6 as follows:

$$\begin{aligned}
&T_1 = a_2 - a_0, \quad T_2 = a_1T_1, \quad T_3 = a_2^2\\
&T_4 = (a_0 - a_1)^2, \quad T_5 = a_1(a_2 - a_1), \quad T_6 = T_1^2 + T_3\\
&\text{Coefficient of } \tau_1: \quad t_1 - t_0 = 2a_1(a_2 - a_0) - 2(a_0 - a_1)^2 - a_2^2\\
&\qquad\qquad = 2(T_2 - T_4) - T_3,\\
&\text{Coefficient of } \tau_2: \quad t_2 - t_0 = 2a_1(a_2 - a_1) - (a_2 - a_0)^2 - a_2^2\\
&\qquad\qquad = 2T_5 - T_6,\\
&\text{Coefficient of } \tau_3: \quad -t_0 = -(a_0 - a_1)^2 - (a_2 - a_0)^2 - a_2^2 = -T_4 - T_6.
\end{aligned}$$

Thus, required for calculating all coefficients of $\boldsymbol{a}^2$ is 2M + 3S + 11A.

4 Arithmetic over a quartic extension with the Gauss periods

This section discusses the algorithms and the costs of each operation over a quartic extension.

4.1 The normal basis of the direct quartic extension

The normal basis of a quartic extension used in this section is a type-I optimal normal basis. For simplicity, let $r = 5$ and β be such that $\Phi_5(\beta) = \frac{\beta^5 - 1}{\beta - 1} = \sum_{i=0}^{4} \beta^i = 0$. Then the condition on p is given by Condition 3 and the normal basis used in what follows is given by $\{\beta, \beta^2, \beta^3, \beta^4\}$.

Condition 3. The Gauss periods of type $(1, 4)$ form a normal basis of $\mathbb{F}_{p^4}$ if $p \neq r$ and $p \equiv 2, 3 \pmod 5$.

4.2 Multiplication

Let $\boldsymbol{a} = (a_0, a_1, a_2, a_3), \boldsymbol{b} = (b_0, b_1, b_2, b_3)$ be vectors in $\mathbb{F}_{p^4}^{\times}$, which are represented by the basis $\{\beta, \beta^2, \beta^3, \beta^4\}$ as linear combinations. According to the conventional CVMA algorithm, $\boldsymbol{a} \times \boldsymbol{b}$ is expanded as follows:

$$\begin{aligned}\boldsymbol{a} \times \boldsymbol{b} = &\{T_1 - (a_1 - a_3)(b_1 - b_3) - a_0 b_0\}\beta \\ &+ \{T_1 - (a_2 - a_3)(b_2 - b_3) - a_1 b_1\}\beta^2 \\ &+ \{T_1 - (a_0 - a_1)(b_0 - b_1) - a_2 b_2\}\beta^3 \\ &+ \{T_1 - (a_0 - a_2)(b_0 - b_2) - a_3 b_3\}\beta^4, \end{aligned} \tag{4.1}$$

where $T_1 = (a_0 - a_3)(b_0 - b_3) + (a_1 - a_2)(b_1 - b_2)$.

A multiplicative operation over $\mathbb{F}_p$ can be saved using again Karatsuba-like trick. T_1 can indeed be computed as

$$\begin{aligned}T_1 = &(a_0 + a_1 - a_2 - a_3)(b_0 + b_1 - b_2 - b_3) + (a_0 - a_1)(b_0 - b_1) \\ &- (a_0 - a_2)(b_0 - b_2) - (a_1 - a_3)(b_1 - b_3) + (a_2 - a_3)(b_2 - b_3).\end{aligned}$$

In addition, let $T_2, \ldots, T_9$ be temporary values which are defined as follows:

$$\begin{aligned}&T_2 = a_0 - a_2, \quad T_3 = a_1 - a_3, \quad T_4 = b_0 - b_2, \quad T_5 = b_1 - b_3, \quad T_6 = T_2 T_4, \\ &T_7 = T_3 T_5, \quad T_8 = (a_0 - a_1)(b_0 - b_1), \quad T_9 = (a_2 - a_3)(b_2 - b_3).\end{aligned}$$

Then it is found that T_1 and the coefficients $c_0, \ldots, c_3$ of $\boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b}$ can be represented by combinations of the temporary values as follows:

$$\begin{aligned}&T_1 = (T_2 + T_3)(T_4 + T_5) - T_6 - T_7 + T_8 + T_9, \\ &c_0 = T_1 - T_7 - a_0 b_0, \quad c_1 = T_1 - T_9 - a_1 b_1, \\ &c_2 = T_1 - T_8 - a_2 b_2, \quad c_3 = T_1 - T_6 - a_3 b_3.\end{aligned}$$

Therefore, the cost required for calculating all coefficients of $\boldsymbol{a} \times \boldsymbol{b}$ in $\mathbb{F}_{p^4}^{\times}$ is 9M + 22A.

4.3 Squaring

In equation (4.1), let $\boldsymbol{a}$ be equal to $\boldsymbol{b}$. Then, the square of $\boldsymbol{a}$ is obtained as follows:

$$\begin{aligned}\boldsymbol{a}^2 = &(2a_1a_3 - 2a_0a_3 - 2a_1a_2 + a_2^2)\beta + (2a_2a_3 - 2a_0a_3 - 2a_1a_2 + a_0^2)\beta^2 \\ &+ (2a_0a_1 - 2a_0a_3 - 2a_1a_2 + a_3^2)\beta^3 + (2a_0a_2 - 2a_0a_3 - 2a_1a_2 + a_1^2)\beta^4.\end{aligned} \tag{4.2}$$

Since the coefficients in equation (4.2) are symmetrical form, they have common factors such as $(a_0-a_1)(a_2-a_3)$ and $(a_0-a_2)(a_1-a_3)$. Therefore, it is possible to transform equation (4.2) as follows:

$$\begin{aligned}\boldsymbol{a}^2 = {} & (2(a_0-a_1)(a_2-a_3)-a_2(2a_0-a_2))\beta \\ & +(2(a_0-a_2)(a_1-a_3)-a_0(2a_1-a_0))\beta^2 \\ & +(2(a_0-a_2)(a_1-a_3)-a_3(2a_2-a_3))\beta^3 \\ & +(2(a_0-a_1)(a_2-a_3)-a_1(2a_3-a_1))\beta^4.\end{aligned}$$

Consequently, the coefficient vector (c_0,c_1,c_2,c_3) of $\boldsymbol{c}=\boldsymbol{a}^2$ is calculated as follows:

$$\begin{aligned}&T_1=a_0-a_2,\quad T_2=a_0-a_1,\quad T_3=a_2-a_3,\\ &T_4=a_1-a_3,\quad T_5=2T_2T_3,\quad T_6=2T_1T_4,\\ &c_0=T_5-a_2(a_0+T_1),\\ &c_1=T_6-a_0(a_1-T_2),\\ &c_2=T_6-a_3(a_2+T_3),\\ &c_3=T_5-a_1(a_3-T_4).\end{aligned}$$

Thus, the cost required for calculating all coefficients of $\boldsymbol{a}^2$ is 6M + 14A.

5 Comparisons and considerations

5.1 Comparisons

In this section, we compare the costs of multiplication and squaring over cubic and quartic extensions. Since the type-I, II optimal normal bases are involved in the CVMA based on the Gauss periods of type (h,m), we select algorithms for polynomial bases as the competitors by referring [8]. The comparisons are shown in Tables 1, 2, 3, and 4. It is noted that the costs of the Toom–Cook series have the lowest cost in terms of $\mathbb{F}_p$ multiplications in both multiplication and squaring in the cubic field but it is usually not used in practice because it involves much additions.

Table 1: Multiplication costs for cubic extensions.

Algorithm	Costs
Schoolbook	9M + 6A + 2c
Karatsuba	6M + 13A + 2c
Toom–Cook-3x	5M + 35A
CVMA	6M + 12A

Table 2: Squaring costs for cubic extensions.

Algorithm	Costs
Schoolbook	3M + 3S + 6A + 2c
Karatsuba	6S + 13A + 2c
Toom–Cook-3x	5S + 35A
Chung–Hasan-SQR2	2M + 3S + 10A + 2c
Chung–Hasan-SQR3	1M + 4S + 11A + 2c + D_2
CVMA	2M + 3S + 11A

Table 3: Multiplication costs for quartic extensions.

Algorithm	Costs
Schoolbook($\mathbb{F}_{p^4}$)/Schoolbook($\mathbb{F}_{p^2}$)	16M + 12A + 5c
Schoolbook($\mathbb{F}_{p^4}$)/Karatsuba($\mathbb{F}_{p^2}$)	12M + 16A + 4c
Karatsuba($\mathbb{F}_{p^4}$)/Schoolbook($\mathbb{F}_{p^2}$)	12M + 24A + 5c
Karatsuba($\mathbb{F}_{p^4}$)/Karatsuba($\mathbb{F}_{p^2}$)	9M + 25A + 4c
CVMA($\mathbb{F}_{p^4}$)	9M + 22A

Table 4: Squaring costs for quartic extensions.

Algorithm	Costs
Schoolbook($\mathbb{F}_{p^4}$)/Schoolbook($\mathbb{F}_{p^2}$)	6M + 4S + 10A + 4c
Schoolbook($\mathbb{F}_{p^4}$)/Karatsuba($\mathbb{F}_{p^2}$)	3M + 6S + 17A + 4c
Schoolbook($\mathbb{F}_{p^4}$)/Complex & Karatsuba($\mathbb{F}_{p^2}$)	7M + 19A + 6c
Karatsuba($\mathbb{F}_{p^4}$)/Schoolbook($\mathbb{F}_{p^2}$)	3M + 6S + 14A + 4c
Karatsuba($\mathbb{F}_{p^4}$)/Karatsuba($\mathbb{F}_{p^2}$)	9S + 20A + 4c
Karatsuba($\mathbb{F}_{p^4}$)/Complex & Karatsuba($\mathbb{F}_{p^2}$)	6M + 23A + 7c
Complex($\mathbb{F}_{p^4}$)/Schoolbook($\mathbb{F}_{p^2}$)	8M + 14A + 4c
Complex($\mathbb{F}_{p^4}$)/Karatsuba($\mathbb{F}_{p^2}$)	6M + 20A + 4c
CVMA($\mathbb{F}_{p^4}$)	6M + 14A

5.2 Considerations

According to the Tables 1, 2, 3, and 4, it is found that the proposed methods require less additions. It is noted that the costs of the conventional methods would increase depending on the choice of their modular polynomial. Since the conditions of primes for constructing a base extension field are different on the classical polynomial basis and the normal basis for the CVMA, the proposed method allows one for having more candidates for suitable representations of extension fields. For example, according to [9], conditions $u \not\equiv 7, 11 \pmod{12}$ for a BN prime $p = 36u^4 + 36u^3 + 24u^2 + 6u + 1$ [3] are a nice choice for reducing the cost of constant multiplications to implement a pairing on BN curves using the classical methods. In that case, p has to be represented by

$p = 12n + 7$ at least for a positive integer n. However, the CVMA gives an opportunity for having an efficient implementation with other primes such that $p \equiv 3 \pmod 5$, which is the case for using the quartic extension as a base field.

Moreover, the improvements of each proposed method are only in terms of additions in $\mathbb{F}_p$. However, it contributes to reducing the additions in higher degree extension fields as well as the classical methods. The cost of an operation would be the same level as the classical methods even if we treat a constant multiplication c as an addition A in the Karatsuba method. For example, consider constructing $\mathbb{F}_{p^{18}}$ as the tower of fields of the form $\mathbb{F}_{((p^3)^2)^3}$ using the Karatsuba method and the proposed method. In the case of conventional Karatsuba-based construction, the tower of fields is often built in the following strategy for minimizing the cost of calculation over $\mathbb{F}_{p^{18}}$:

$$\begin{aligned}\mathbb{F}_{p^3} &= \mathbb{F}_p[\alpha]/(\alpha^3 - z),\\ \mathbb{F}_{p^6} &= \mathbb{F}_{p^3}[\mu]/(\mu^2 - \alpha),\\ \mathbb{F}_{p^{18}} &= \mathbb{F}_{p^6}[\nu]/(\nu^3 - \mu),\end{aligned}$$

where z satisfies $z^{\frac{p-1}{2}} \neq 1$ and $z^{\frac{p-1}{3}} \neq 1$ in $\mathbb{F}_p$. On the other hand, consider replacing the first cubic extension field for applying the CVMA as follows:

$$\begin{aligned}\mathbb{F}_{p^3} &= \mathbb{F}_p[\beta]/(\Phi_7(\beta) = 0),\\ \mathbb{F}_{p^6} &= \mathbb{F}_{p^3}[\mu]/(\mu^2 - z),\\ \mathbb{F}_{p^{18}} &= \mathbb{F}_{p^6}[\nu]/(\nu^3 - \mu),\end{aligned}$$

where z is the same as the above.

To clarify the corresponding field for operations, let M_i, A_i, and c_i denote the multiplication, addition, and constant multiplication in $\mathbb{F}_{p^i}$ for $i \geq 1$, respectively. It is noted that M, A, and c are also used to describe as the general case. Then, the first one can have the cost for a single multiplication in $\mathbb{F}_{p^{18}}$ as $6\mathrm{M}_6 + 13\mathrm{A}_6 + 2\mathrm{c}_6 \approx 6\mathrm{M}_6 + 15\mathrm{A}_6$ by using Table 1. That is because an element in $\mathbb{F}_{p^{18}}$ is represented as a 3-dimensional vector of $\mathbb{F}_{p^6}$-elements. Second, remind that the cost of multiplication in a quadratic field with the Karatsuba method requires $3\mathrm{M} + 5\mathrm{A} + \mathrm{c} \approx 3\mathrm{M} + 6\mathrm{A}$ in general [8] and $\mathbb{F}_{p^6}$ is constructed as the 2-dimensional vector space whose coefficients are elements in $\mathbb{F}_{p^3}$. Thus, we can estimate M_6 and A_6 by $\mathrm{M}_6 = 3\mathrm{M}_3 + 6\mathrm{A}_3$ and $\mathrm{A}_6 = 2\mathrm{A}_3$, respectively. In the same manner, M_3 and A_3 are estimated by using Table 1 as follows:

$$\begin{aligned}M_3 &= \begin{cases}6\mathrm{M}_1 + 12\mathrm{A}_1 & \text{if the CVMA is used,}\\ 6\mathrm{M}_1 + 15\mathrm{A}_1 & \text{if the Karatsuba method is used,}\end{cases}\\ A_3 &= 3\mathrm{A}_1,\end{aligned}$$

respectively Therefore, the cost of multiplication in $\mathbb{F}_{p^{18}}$ using the CVMA is derived by

$$\begin{aligned}6\mathrm{M}_6 + 15\mathrm{A}_6 &= 6(3\mathrm{M}_3 + 6\mathrm{A}_3) + 15(2\mathrm{A}_3) = 18(6\mathrm{M}_1 + 12\mathrm{A}_1) + 66(3\mathrm{A}_1)\\ &= 108\mathrm{M}_1 + 414\mathrm{A}_1.\end{aligned}$$

On the other hand, the cost of multiplication in $\mathbb{F}_{p^{18}}$ using the Karatsuba method is derived by

$$\begin{aligned} 6\mathrm{M}_6 + 15\mathrm{A}_6 &= 6(3\mathrm{M}_3 + 6\mathrm{A}_3) + 15(2\mathrm{A}_3) = 18(6\mathrm{M}_1 + 15\mathrm{A}_1) + 66(3\mathrm{A}_1) \\ &= 108\mathrm{M}_1 + 468\mathrm{A}_1. \end{aligned}$$

The condition for constructing cubic fields as the base field and the required cost of multiplication with the CVMA or the Karatsuba method in $\mathbb{F}_{p^{18}}$ are summarized in Table 5. Moreover, considering the fact that the sextic twist can be applied for $\mathbb{F}_{((p^3)^2)^3}$, it is possible to handle rational points on $\mathbb{F}_{((p^3)^2)^3}$ as points on $\mathbb{F}_{p^3}$. In this sense, the improvement of base field arithmetic in this paper contributes to implementing pairing based cryptography efficiently as an alternative or a new class to classical methods.

Table 5: The summary of the conditions for constructing cubic fields as the base field and the comparison of the required cost for multiplication in $\mathbb{F}_{p^{18}}$.

	Condition for $\mathbb{F}_{p^3}$	**Modular polynomial for $\mathbb{F}_{p^3}$**	**The cost of multiplication in $\mathbb{F}_{p^{18}}$**
CVMA	$p \neq 7$ and $p \not\equiv 1, 6 \pmod 7$	$x^4 + x^3 + x^2 + x + 1$	$108\mathrm{M}_1 + 414\mathrm{A}_1$
Karatsuba	$3 \mid (p-1)$ and $z^{\frac{p-1}{3}} \not\equiv 1 \pmod p$	$x^2 - z$	$108\mathrm{M}_1 + 468\mathrm{A}_1$

6 Conclusion

This paper proposed algorithms which are designed based on the Gauss periods of type (h, m). The algorithms provide fundamental arithmetic over the cubic and the quartic extensions, which are useful for constructing pairing-friendly fields as towers of fields. To evaluate the efficiency, the costs were compared with the well-known algorithms for efficient extension fields arithmetic such as Karatsuba method.

As a consequence, it was found that the proposed methods require less additions. As shown in Tables 1, 2, 3, 4, and 5, the proposed methods contribute to construct the defining field for handling efficient arithmetic. It increases the options for selecting the suitable structure (basis representation) of the defining field for efficient implementations. Then, since the conditions for applicable primes are different, one can use both the proposed method and the classical methods depending on the features of the base field and that of target devices flexibly. Further improvement and practical results ought to be discussed for much more efficient and useful algorithms. For example, it should be theoretically possible to get an equivalent of the Toom–Cook or the Chung–Hasan-SQR3 method that maybe involves less additions and make it practically competitive. Such advanced researches are future works.

Bibliography

[1] D. W. Ash, I. F. Blake, and S. A. Vanstone, Low complexity normal bases, *Discrete Appl. Math.*, **25**(3) (1989), 191–210.

[2] P. S. L. M. Barreto, B. Lynn, and M. Scott, In: Constructing Elliptic Curves with Prescribed Embedding Degrees, SCN 2002, LNCS, Vol. 2576, Springer, 2002, pp. 257–267.

[3] P. S. L. M. Barreto and M. Naehrig, In: Pairing-Friendly Elliptic Curves of Prime Order, SAC 2005, LNCS, Vol. 3897, Springer, 2005, pp. 319–331.

[4] D. Boneh and M. Franklin, In: Identity-Based Encryption from the Weil Pairing, CRYPTO 2001, LNCS, Vol. 2139, Springer, 2001, pp. 213–229.

[5] D. Boneh, B. Lynn, and H. Shacham, Short signatures from the Weil pairing, *J. Cryptol.*, **17**(4) (2004), 297–319.

[6] D. Boneh, C. Gentry, and B. Waters, In: Collusion Resistant Broadcast Encryption with Short Ciphertexts and Private Keys, CRYPTO 2005, LNCS, Vol. 3621, Springer, 2005, pp. 258–275.

[7] S. A. Cook, On the Minimum Computation Time of Functions, PhD Thesis, Harvard University Department of Mathematics, 1966.

[8] A. J. Devegili, C. Ó hÉigeartaigh, M. Scott, and R. Dahab, Multiplication and Squaring on Pairing-Friendly Fields, IACR Eprint archive, 2006, http://eprint.iacr.org/2006/471 (visited 2019-4-22).

[9] S. Duquesne, N. E. Mrabet, S. Haloui, and F. Rondepierre, Choosing and generating parameters for low level pairing implementation on BN curves, *Appl. Algebra Eng. Commun. Comput.*, **29**(2) (2018), 113–147.

[10] S. Feisel, J. von zur Gathen, and M. A. Shokrollahi, Normal bases via general Gauss periods, *Math. Comput.*, **68**(255) (1999), 271–290.

[11] S. Gao, J. von zur Gathen, and D. Panario, In: Gauss Periods and Fast Exponentiation in Finite Fields Extended Abstract, LATIN '95, LNCS, Vol. 911, Springer, 1995, pp. 311–322.

[12] S. Gao, Abelian groups, Gauss periods, and normal bases, *Finite Fields Appl.*, **7**(1) (2001), 149–164.

[13] F. Hess, N. P. Smart, and F. Vercauteren, The eta pairing revisited, *IEEE Trans. Inf. Theory*, **52**(10) (2006), 4595–4602.

[14] E. J. Kachisa, E. F. Schaefer, and M. Scott, In: Constructing Brezing-Weng Pairing-Friendly Elliptic Curves Using Elements in the Cyclotomic Field, Pairing 2008, LNCS, Vol. 5209, Springer, 2008.

[15] A. Karatsuba and Yu. Ofman, Multiplication of multidigit numbers on automata, *Sov. Phys. Dokl.*, **7** (1963), 595–596.

[16] H. Kato, Y. Nogami, T. Yoshida, and Y. Morikawa, Cyclic vector multiplication algorithm based on a special class of Gauss period normal basis, *ETRI J.*, **29**(6) (2007), 769–778.

[17] D. E. Knuth, The Art of Computer Programming, volume 1 (3rd ed.): Fundamental Algorithms, Addison Wesley Longman Publishing Co., Inc., 1997.

[18] N. Koblitz and A. Menezes, In: Pairing-Based Cryptography at High Security Levels, Cryptography and Coding 2005, LNCS, Vol. 3796, Springer, 2005, pp. 13–36.

[19] R. Lidl and H. Niederreiter, Introduction to Finite Fields and Their Applications, Cambridge University Press, 1986.

[20] N. E. Mrabet and M. Joye, Guide to Pairing-Based Cryptography, Chapman and Hall/CRC, 2016.

[21] R. C. Mullin, I. M. Onyszchuk, S. A. Vanstone, and R. M. Wilson, Optimal normal bases in GF(p^n), *Discrete Appl. Math.*, **22**(2) (1989), 149–161.

[22] M. Naehrig, K. Lauter, and V. Vaikuntanathan, In: Can Homomorphic Encryption Be Practical? CCSW 2011, ACM, 2011, pp. 113–124.

[23] K. Nekado, Proposals of Multiplication and Inversion Methods in Extension Field for Scalable

Asymmetric-key and Fast Symmetric-key Cryptosystems, PhD Thesis, Graduate School of Natural Science and Technology, Okayama University, 2013.

[24] Y. Nogami, A. Saito, and Y. Morikawa, Finite extension field with modulus of all-one polynomial and representation of its elements for fast arithmetic operations, *IEICE Trans. Fundam.*, **E86-A**(9) (2003), 2376–2387.

[25] A. L. Toom, The complexity of a scheme of functional elements realizing the multiplication of integers, *Sov. Math.*, **4**(3) (1963), 714–716.

Pascale Charpin

Crooked functions

Abstract: Crooked permutations were introduced 20 years ago since they allow to construct interesting objects in graph theory. The field of applications was extended later. Crooked functions, bijective or not, correspond to APN functions and to some optimal codes. We adopt an unified presentation of crooked functions, explaining the connection with partially-bent functions. We then complete some known results and derive new properties. For instance, we observe that crooked functions allow to construct sets of bent functions and define some permutations.

Keywords: Vectorial function, Boolean function, derivative, differential set, plateaued function, partially-bent functions, bent functions, APN function, AB function, permutation

MSC 2010: 12Y05, 33B99

1 Introduction

The *crooked functions* have been introduced by Bending and Fon-Der-Flaass in 1998, as combinatorial objects of great interest [1]. Such a function has been defined from V to W, two n-dimensional vector spaces over $\mathbb{F}_2$, by the following property: the image set of any of its derivatives is the complement of a hyperplane. This characterization implies that a crooked function is bijective, and allows, in particular, to construct distance regular graphs. Later, several authors have developed this work, and have generalized the previous definition. They notably related the crooked functions with several optimal objects, which have applications both to cryptography and coding theory [12, 13].

This paper is a survey on crooked functions, including several new results. We recall what is known about crooked functions presently. We introduce another approach to study crooked functions by starting from the so-called *partially-bent* functions, and present some new results.

After preliminaries, we propose a brief survey on the (few) papers considering crooked functions. To our knowledge, the list of references includes all such papers. Section 4 is devoted to the structure of crooked functions. We begin by proving a link between partially-bent functions and crooked functions. We later differentiate the two cases: odd and even number of variables. The odd case was mainly treated in the first papers, since in this case, crooked functions could be permutations. In Sections 5

Acknowledgement: The author want to thank Gohar Kyureghyan for helpful comments and fruitful discussions.

Pascale Charpin, INRIA, 2 rue Simone Iff, Paris 75012, France, e-mail: Pascale.Charpin@inria.fr

https://doi.org/10.1515/9783110621730-007

and 6, we show how to construct, respectively, a set of bent functions and a set of permutations, using the nice structure of a crooked function. We conclude by the main conjecture about the existence of crooked functions.

2 Definitions, basic properties

Throughout this paper, $|E|$ denotes the cardinality of the set E, and $E^* = E \setminus \{0\}$. Let F be a mapping, from the finite field $\mathbb{F}_{2^n}$ to itself. Such a function is called *a vectorial function*, while a function f, from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$ is, as usually, a *Boolean function*. We denote by $\mathcal{I}m(\xi)$ the image set of any function ξ.

A vectorial function F, from $\mathbb{F}_{2^n}$ to itself, is said to be an *almost perfect nonlinear (APN)* function if and only if all the equations,

$$F(x) + F(x+a) = c, \quad a, c \in \mathbb{F}_{2^n}, \quad a \neq 0, \tag{2.1}$$

have zero or two solutions in $\mathbb{F}_{2^n}$, say x and $x+a$. Throughout this paper, we use U to denote a subfield of $\mathbb{F}_{2^n}$, usually $\mathbb{F}_{2^n}$ or $\mathbb{F}_2$. For $a \in \mathbb{F}_{2^n}^*$, the function from $\mathbb{F}_{2^n}$ to U, defined by

$$x \mapsto D_a F(x) = F(x) + F(x+a),$$

is called *derivative of F*, with respect to a. We call *differential set*, in point a, the image set of $D_a F$:

$$\mathcal{I}m(D_a F) = \{F(x) + F(x+a) \mid x \in \mathbb{F}_{2^n}\}. \tag{2.2}$$

Clearly, when F is APN, we have for any $a \in \mathbb{F}_{2^n}^*$:

$$D_a F(x) = D_a F(x+a) = c, \quad \text{for some } c,$$

for only one pair $(x, x+a)$. This means that $D_a F$ is a 2-to-1 function. Thus, one can formulate the APN property as follows.

Proposition 1. *A function $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ is APN if and only if all its differential sets have cardinality 2^{n-1}.*

The 2^n, so-called, *components of F* are the Boolean functions

$$f_\lambda : x \longmapsto \mathrm{Tr}(\lambda F(x)), \quad \lambda \in \mathbb{F}_{2^n},$$

where f_0 is the null function, by convention. They are linearly defined by means of the absolute trace on $\mathbb{F}_{2^n}$:

$$x \mapsto \mathrm{Tr}(x) = x + x^2 + \cdots + x^{2^{n-1}}.$$

The dual $V^\perp$, of any subspace V of $\mathbb{F}_{2^n}$, is the subspace of those y such that $\mathrm{Tr}(yx) = 0$, for all $x \in V$. The *Walsh transform* of a Boolean function f, is defined as

$$a \in \mathbb{F}_{2^n} \mapsto \mathcal{W}_f(a) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x) + \mathrm{Tr}(ax)}.$$

Recall the *Parseval's relation*:

$$\sum_{a\in\mathbb{F}_{2^n}} \left(\mathcal{W}_f(a)\right)^2 = 2^{2n}.$$

We will need the following result.

Lemma 1 ([6, Lemma V.2]). *Let f be a Boolean function over $\mathbb{F}_{2^n}$, and let V be a subspace of $\mathbb{F}_{2^n}$ of dimension k, $0 \le k \le n$. Then*

$$\sum_{v\in V} \left(\mathcal{W}_f(v)\right)^2 = 2^k \sum_{u\in V^\perp} \sum_{x\in\mathbb{F}_{2^n}} (-1)^{f(x)+f(x+u)}.$$

We now define particular APN functions, which exist for odd n only.

Definition 1. The function F is said to be an *almost bent* (AB) function if the numbers

$$\mu_F(a,\lambda) = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(\lambda F(x)+ax)}, \tag{2.3}$$

are equal to 0 or $\pm 2^{\frac{n+1}{2}}$ only, when $a \in \mathbb{F}_{2^n}$ and $\lambda \in \mathbb{F}_{2^n}^*$.

Note that $\mu_F(a,\lambda) = \mathcal{W}_{f_\lambda}(a)$ for any fixed λ.

A Boolean function f, over $\mathbb{F}_{2^n}$, is said to be *bent* when $\mathcal{W}_f$ takes two values $\{\pm 2^{n/2}\}$ only, in particular n must be even then. It is said to be *s-plateaued* when $\mathcal{W}_f$ takes three values,

$$\{0, \pm 2^{(n+s)/2}\}, \quad \text{with } 1 \le s \le n-2 \text{ and } n+s \text{ even}.$$

By convention, a bent function is 0-plateaued. The value $2^{(n+s)/2}$ is the *amplitude* of f. A *plateaued vectorial function* is a vectorial function whose components are plateaued Boolean functions. It is said that F is *plateaued with single amplitude*, when all components of F have the same amplitude.

The *sum-of-square* indicator of f is defined by

$$\nu(f) = \sum_{a\in\mathbb{F}_{2^n}} \mathcal{W}_{D_af}^2(0) = 2^{-n} \sum_{b\in\mathbb{F}_{2^n}} \mathcal{W}_f^4(b). \tag{2.4}$$

If f is s-plateaued, then $\nu(f) = 2^{2n+s}$. Moreover, the vectorial function F is APN if and only if

$$\sum_{\lambda\in\mathbb{F}_{2^n}^*} \nu(f) = 2^{2n+1}(2^n - 1) \tag{2.5}$$

(see [2, Corollary 1]). A Boolean function f is said to be *balanced* if it takes the values 0 and 1 the same number of times. Recall a well-known result:

Theorem 1. *A function $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ is a permutation if and only if all its components f_λ, $\lambda \in \mathbb{F}_{2^n}^*$, are balanced.*

3 Brief record

We use our terminology, of the previous section, rather than of the initial works on crooked functions. The next definition was proposed by Bending and Fon-Der-Flaass in [1], 20 years ago.

Definition 2. Let F be a function from $\mathbb{F}_{2^n}$ to itself. This function is called *crooked* if it satisfies the following three properties:
(i) $F(0) = 0$;
(ii) $F(x) + F(y) + F(z) + F(x + y + z) \neq 0$, for any three distinct x, y, z;
(iii) $D_aF(x) + D_aF(y) + D_aF(z) \neq 0$, for arbitrary x, y, z, and any $a \neq 0$.

This definition implies that such a function F is a bijection over $\mathbb{F}_{2^n}$, where n must be odd. Note that the condition (ii) means that F is APN (see (2.1)). Also, F is crooked if and only if any of its differential sets is a complement of a hyperplane. Further, other properties are studied, in [1], such as some relations of crooked permutations with bent functions of dimension $n - 1$, and with the so-called *Kerdock sets*.

Crooked permutations allow to construct some *distance regular graphs*. This was shown in [1], generalizing previous constructions. Later, van Dam and Fon-Der-Flaass proposed another construction, and then another generalization (see, in particular, Theorem 3 of [12]). Conversely, Godsil and Roy have shown that crooked permutations can be fully characterized by Preparata codes of minimum distance 5. Similarly, some distance-regular graphs provide crooked permutations [14, Theorems 3, 5].

Since the high interest for APN functions in cryptography and coding theory [10], the existence of crooked functions was later the core of the research on crooked functions. Kyureghyan proposed another definition of crooked functions, identifying all APN functions which are such that their differential sets are affine hyperplanes. She established the basic properties of such functions. She notably proved that the *monomial crooked* functions are quadratic [15, 16].

The APN quadratic functions are crooked. We do not know if crooked functions of higher algebraic degree do exist. This is a recurring question, about which only negative results have been obtained. An important result was obtained by Bierbrauer and Kyureghyan: *binomial crooked* functions are quadratic [4].

4 Structure of crooked functions

A hyperplane of $\mathbb{F}_{2^n}$ is an $(n-1)$-dimensional subspace of $\mathbb{F}_{2^n}$ over $\mathbb{F}_2$. For any hyperplane H there is a unique $\lambda \in \mathbb{F}_{2^n}^*$ such that

$$H = H_\lambda := \{y \in \mathbb{F}_{2^n} \mid \mathrm{Tr}(\lambda y) = 0\}. \tag{4.1}$$

We will denote by $\overline{H_\lambda}$ the complement of H_λ. The dual of H_λ is obviously $H_\lambda^\perp = \{0, \lambda\}$.

The next definition of *crooked* functions is due to Kyureghan [16]. The corpus of such functions includes the crooked permutations, but also a large variety of nonbijective crooked functions, especially in even dimension.

Definition 3. A function F, from $\mathbb{F}_{2^n}$ to itself, is called *crooked* when it is such that, for every $a \in \mathbb{F}_{2^n}^*$, its differential set

$$S_a = \{F(x) + F(x+a) \mid x \in \mathbb{F}_{2^n}\}$$

is an affine hyperplane.

We directly deduce from Proposition 1:

Claim 1. *Any crooked function is an APN function.*

Assuming that $S_a = H_\lambda$ or $\overline{H_\lambda}$, for some λ, we get

$$\mathrm{Tr}(\lambda(D_a F(x))) = c, \quad \text{for all } x, \text{ where } c \in \mathbb{F}_2 \text{ is fixed.}$$

This means that the derivative of the component f_λ of F, in point a, is a constant function. In this case, a is said to be a *linear structure* of f_λ. The *linear space* of f_λ is the subspace, including $a = 0$, of its linear structures. We will see that crooked functions have always components with nonzero linear structures.

4.1 Crooked and partially-bent functions

Let $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$ be a Boolean function of $\mathbb{F}_{2^n}$. Denote by N_d, the number of balanced derivatives of f, and by N_f, the size of the set $\{a \mid \mathcal{W}_f(a) = 0\}$. Partially-bent functions were introduced by Carlet in [9], as functions satisfying

$$(2^n - N_d)(2^n - N_f) = 2^n. \tag{4.2}$$

There is another characterization of partially-bent functions, which allows to determine their Walsh spectrum.

Theorem 2 ([9, Theorem]). *A Boolean function f is partially-bent if and only if its derivatives are either constant or balanced.*

This leads to the precise description of the Walsh spectrum of any partially-bent function. Note that bent functions are particular partially-bent functions, with $N_d = 2^n - 1$ and $N_f = 0$. Moreover, a partially-bent function is, in a certain sense, obtained by concatenating the same bent function, several times. The next result is partly given by [9, Proposition 2]. See also [16, Theorem 1], which concerns crooked functions but, actually, holds for any partially-bent function. For the Walsh spectrum, see a proof in [7, Proposition 4].

Corollary 1. *Let f be a partially-bent Boolean function. Assume that f has a linear space V of dimension $s > 0$.*

Then f is an s-plateaued function, such that $n+s$ is even, and $\mathcal{W}_f$ takes three values, 0 and $\pm 2^{(n+s)/2}$. The Walsh spectrum of f is given in Table 1.

Table 1: Walsh spectrum of the Boolean s-plateaued function f.

$\mathcal{W}_f(u)$	**Number of $u \in \mathbb{F}_{2^n}$**
0	$2^n - 2^{n-s}$
$2^{(n+s)/2}$	$2^{n-s-1} + (-1)^{f(0)} 2^{(n-s)/2-1}$
$-2^{(n+s)/2}$	$2^{n-s-1} - (-1)^{f(0)} 2^{(n-s)/2-1}$

The definition of a partially-bent Boolean function can be extended to the one of a vectorial partially-bent function as follows.

Definition 4. Let $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$. The function F is said to be a *partially-bent vectorial function* if every component of F is partially-bent. In particular, when n is even, some components can be bent.

Theorem 3. *Let F be a vectorial function over $\mathbb{F}_{2^n}$ with components f_λ. Set, for $a \in \mathbb{F}_{2^n}^*$,*

$$\Lambda_a = \{\lambda \in \mathbb{F}_{2^n}^* \mid D_a f_\lambda \text{ is constant}\} \cup \{0\},$$

and denote by $\ell(a)$ the dimension of Λ_a. Then we have:

- *Assume that F is partially-bent. Then the differential sets of F are affine subspaces. For any $a \in \mathbb{F}_{2^n}^*$, this subspace is of codimension $\ell(a)$, with $\ell(a) \geq 1$ and $D_a F$ is a $2^{\ell(a)}$-to-1 function.*
- *The function $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ is a crooked function if and only if F is partially-bent with $\ell(a) = 1$ for every nonzero a.*

Proof. Recall that $D_a f_\lambda(x) = \mathrm{Tr}(\lambda D_a F(x))$, for all x. Obviously, Λ_a is a subspace of $\mathbb{F}_{2^n}$. Now, fixing a and x, we compute

$$\begin{aligned} B(a,x) &= \sum_{\lambda \in \mathbb{F}_{2^n}} \sum_{y \in \mathbb{F}_{2^n}} (-1)^{D_a f_\lambda(x) + D_a f_\lambda(y)} \\ &= \sum_{y \in \mathbb{F}_{2^n}} \sum_{\lambda \in \mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(\lambda(D_a F(x) + D_a F(y)))} \\ &= 2^n \times \left|\{y \mid D_a F(x) = D_a F(y)\}\right|. \end{aligned}$$

On the other hand, we set for any $\lambda \in \mathbb{F}_{2^n}^*$:

$$B(\lambda) = \sum_{y \in \mathbb{F}_{2^n}} (-1)^{D_a f_\lambda(x) + D_a f_\lambda(y)} = (-1)^{D_a f_\lambda(x)} \sum_{y \in \mathbb{F}_{2^n}} (-1)^{D_a f_\lambda(y)}.$$

We first assume that F is partially bent. Note that $\ell(a) \geq 1$, since otherwise we would have that the Boolean function $D_a f_\lambda$ is balanced, for any $\lambda \in \mathbb{F}_{2^n}^*$, from Theorem 2. This is impossible since $D_a F$ cannot be a permutation (see Theorem 1).

Clearly, $B(\lambda) = 0$ if and only if the function $D_a f_\lambda$ is balanced. If it is not balanced, then $\lambda \in \Lambda_a$, and this function is constantly equal to either 0 or 1. In both cases, we get $B(\lambda) = 2^n$. Hence, we get

$$B(a, x) = \sum_{\lambda \in \Lambda_a} B(\lambda) = 2^n 2^{\ell(a)}, \quad \text{for all } x,$$

implying that the number of y such that $D_a F(x) = D_a F(y)$ equals $2^{\ell(a)}$, *i. e.*, $D_a F$ is $2^{\ell(a)}$-to-1. Since $\mathcal{I}m(D_a F)$ is contained in an affine subspace of codimension $\ell(a)$, according to the definition of Λ_a, $\mathcal{I}m(D_a F)$ is equal to this affine subspace.

Now, we suppose that F is a crooked function. Consider any component f_λ of F. Let $a \in \mathbb{F}_{2^n}^*$ such that $D_a f_\lambda$ is not constant. Set $V = \mathcal{I}m(D_a F)$. Then we have that any x satisfies:

$$\mathrm{Tr}\big(\lambda D_a F(x)\big) = 0 \quad \text{if and only if } x \in V \cap H_\lambda.$$

There are 2^{n-1} such x, since $D_a F$ is 2-to-1 and V is an affine hyperplane, which is neither H_λ nor its complement. Hence $D_a f_\lambda$ is balanced. We have proved that f_λ is partially-bent, completing the proof. □

Let F be any quadratic function:

$$F(x) = \sum_{0 \leq i < j < n} u_{i,j} x^{2^i + 2^j}, \quad u_{i,j} \in \mathbb{F}_{2^n}.$$

The derivatives of F are affine functions, say L_a for any $a \in \mathbb{F}_{2^n}^*$. Thus, F is partially-bent; it is crooked if and only if every L_a is an affine function with kernel of dimension 1.

Corollary 2. *Any quadratic vectorial function is partially-bent. It is crooked if and only if it is APN.*

4.2 Crooked functions on $\mathbb{F}_{2^n}$, n odd

In this section, we study crooked functions of odd dimension, bijective or not. Theorem 4 (below) is the main result, describing the exceptional properties of such functions. These results are quite known, but were partially presented in several papers [13, 15, 16]. First, it is easy to describe the set of crooked permutations.

Lemma 2. *Let F be a crooked function such that $F(0) = 0$, with differential sets S_a. Let a and λ be such that S_a equals either H_λ or $\overline{H_\lambda}$, where H_λ is defined by (4.1). Then we have*

$$S_a = \overline{H_\lambda} \iff \mathrm{Tr}\big(\lambda F(a)\big) = 1.$$

Besides, F is a permutation if and only if $S_a = \overline{H_\lambda}$, for any such pair (λ, a). In this case, n is odd.

Proof. By hypothesis, we have $\mathrm{Tr}(\lambda D_a F(x)) = c$ for all x, where $c \in \{0,1\}$. In particular, $\mathrm{Tr}(\lambda F(a)) = c$; further, $c = 0$ means that $\mathcal{I}m(D_a F) = H_\lambda$.

The function F is not bijective if and only if $F(x) = F(x+a)$, for some pair (x,a). Equivalently, there is (λ, a) such that $\mathcal{I}m(D_a F) = H_\lambda$, since 0 belongs to $\mathcal{I}m(D_a F)$.

When n is even, F cannot be a permutation, since it is an APN function, which is plateaued (see [2, Theorem 3]). □

When n is odd, a crooked function need not to be bijective, as we show by the next example.

Example 1. Assume that n is odd. It is well known that

$$F : x \mapsto x^{2^t+1}, \quad \text{with } \gcd(t,n) = 1,$$

is an AB permutation. So, it is a crooked permutation. Consider now the function $x \mapsto G(x) = x^{2^t+1} + x$, which is AB, too, and then crooked. Since $G(0) = G(1) = 0$, G is not a permutation.

Definition 5. A Boolean function is said to be *near-bent* if it is 1-plateaued, *i. e.*, its Walsh transform takes the values 0 and $\pm 2^{(n+1)/2}$ only.

Theorem 4. *Let n be odd, F be a crooked function over $\mathbb{F}_{2^n}$ with $F(0) = 0$. For every $a \in \mathbb{F}_{2^n}^*$, define $\lambda(a)$ as the unique element satisfying*

$$S_a = \beta + H_{\lambda(a)},$$

where $\beta \in \mathbb{F}_{2^n}$ is not unique. Then the following properties hold:

(i) *The differential sets S_a are pairwise distinct, which is equivalent to the statement that $a \mapsto \lambda(a)$ is bijective on $\mathbb{F}_{2^n}^*$.*

(ii) *Any component f_λ of F is a near-bent Boolean function with linear space of dimension 1, say $\{0, a\}$. When f_λ is balanced, its derivative in point a is equal to 1. This holds for any f_λ, when F is a permutation.*

(iii) *F is an AB function.*

Proof. Since F is crooked, it is partially-bent. Let f_λ, $\lambda \in \mathbb{F}_{2^n}^*$, be the components of F. Thus, for any λ and for any a, the derivative of f_λ, in point a, is either constant or balanced. Hence

$$\sum_{x \in \mathbb{F}_{2^n}} (-1)^{f_\lambda(x) + f_\lambda(x+a)} \in \{0, \pm 2^n\}.$$

For any fixed λ, there is at least one a, say $a(\lambda)$, such that $D_a f_\lambda$ is constant, since otherwise the function f_λ would be bent, which is impossible when n is odd. Thus, we get

the set of the $a(\lambda)$, whose size is at most $2^n - 1$. However, $a(\lambda) = a(\mu) = b$, for some $\mu \neq \lambda$, would mean the following: the Boolean functions

$$x \mapsto \mathrm{Tr}(\lambda D_b F(x)) \quad \text{and} \quad x \mapsto \mathrm{Tr}(\mu D_b F(x))$$

are both constant. This would imply that S_b is of codimension at least 2, a contradiction. To each λ corresponds one and only one a, completing the proof of (i).

We deduce that any component f_λ has only one nonzero linear structure, say a, *i. e.*, its linear space has dimension 1. Obviously, if f_λ is balanced, its derivative in point a is constantly equal to 1. When F is a permutation, all f_λ are balanced, completing the proof of (ii).

From Corollary 1, every f_λ is near-bent, providing

$$\mathcal{W}_{f_\lambda}(a) \in \{0, \pm 2^{(n+1)/2}\}, \quad \text{for all } a \in \mathbb{F}_{2^n}.$$

According to Definition 1, F is an AB function, completing the proof. □

From Table 1, we know that the support of the Walsh spectrum of any near-bent Boolean function on $\mathbb{F}_{2^n}$ has size 2^{n-1}. It could be an affine subspace of codimension 1, as it holds for components of some crooked functions.

Proposition 2. *Let $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, n odd, such that $F(0) = 0$. Assuming that F is crooked, the set*

$$W_\lambda = \{a \in \mathbb{F}_{2^n} \mid \mathcal{W}_{f_\lambda}(a) = 0\}$$

is an affine subspace of codimension 1, *for all $\lambda \in \mathbb{F}_{2^n}^*$.*

Conversely, assume that the sets W_λ are affine hyperplanes. In this case, if F is APN then F is crooked.

Proof. Assume that F is crooked and let $\lambda \in \mathbb{F}_{2^n}^*$. Since F is an AB function, f_λ is partially bent, with linear space $\{0, b\}$, for some nonzero b. Obviously, the function $g_a : x \mapsto f_\lambda(x) + \mathrm{Tr}(ax)$ is partially bent, too, with linear space $\{0, b\}$, for any $a \in \mathbb{F}_{2^n}$. We know that any partially-bent function is balanced if and only if it is not constant on its linear space (see [9, Proposition 2]). Thus, g_a is balanced if and only if

$$D_b g_a(x) = 1, \quad \textit{i. e.}, D_b f_\lambda(x) + \mathrm{Tr}(ab) = 1, \text{ for all } x.$$

Hence, we get $\mathrm{Tr}(ab) = 1 + c$, $c \in \mathbb{F}_2$, since $D_b f_\lambda$ is a constant function. We can suppose that $c = 0$. Thus, g_a is balanced if and only if $\mathrm{Tr}(ab) = 1$, that is $a \in \overline{H_b}$. Thus

$$W_\lambda = \{a \in \mathbb{F}_{2^n} \mid \mathrm{Tr}(ab) = 1\} = \overline{H_b}.$$

Similarly, with $c = 1$ we obtain $W_\lambda = H_b$.

Conversely, suppose that W_λ is an affine subspace of codimension 1, say $u + H_b$, $u \in \mathbb{F}_{2^n}$, for some b. So, $H_b^\perp = \{0, b\}$ and we have from Lemma 1:

$$\sum_{a\in H_b} \mathcal{W}_{f_\lambda}(a)^2 = 2^{n-1}\Big(\sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(\lambda D_0F(x))} + \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(\lambda D_bF(x))}\Big).$$

By Parseval's relation, the sum above on the left is equal either to 0 or to 2^{2n}. We deduce that D_bf_λ is a constant function. So, if D_bF is 2-to-1, then $\mathcal{I}m(D_bF)$ is equal to either H_λ or $\overline{H_\lambda}$. □

4.3 Crooked functions on $\mathbb{F}_{2^n}$, n even

When n is even, crooked functions are partially-bent functions which have bent components. We recall this property below in Theorem 5.

Theorem 5. *Let F be a plateaued function over $\mathbb{F}_{2^n}$, with n even and $n > 4$. Let $2^{(t_\lambda+n)/2}$ be the amplitude of any component of F, namely of any f_λ, $\lambda \in \mathbb{F}_{2^n}^*$. Denote by B the number of bent components of F. Then F is APN if and only if*

$$B = \sum_{\lambda\in\mathbb{F}_{2^n}^*, t_\lambda>0} (2^{t_\lambda} - 2). \tag{4.3}$$

This property holds, in particular, for crooked functions.

Consequently, B satisfies

$$\frac{2(2^n-1)}{3} \le B < 2^n - 2^{n/2}, \tag{4.4}$$

where the lower bound is reached if and only if $t_\lambda = 2$ for all nonzero t_λ.

Proof. Equality (4.3) has been proved in [11, Proposition 6], but may be computed by using (2.5). Indeed, since each f_λ is plateaued, we have $\nu(f_\lambda) = 2^{t_\lambda+2n}$, for any λ, and then:

$$A = \sum_{\lambda\in\mathbb{F}_{2^n}^*} \nu(f_\lambda) = 2^{2n} \sum_{\lambda\in\mathbb{F}_{2^n}^*} 2^{t_\lambda}.$$

According to (2.5), F is APN if and only if $A = 2^{2n+1}(2^n - 1)$, providing

$$B + \sum_{\lambda\in\mathbb{F}_{2^n}^*, t_\lambda>0} 2^{t_\lambda} = 2(2^n - 1) = 2(B + N), \tag{4.5}$$

since $2^n - 1 = B + N$ with $N = |\{\lambda \mid t_\lambda > 0\}|$. Hence, the equality above is equivalent to (4.3). Note that $t_\lambda > 0$ implies $t_\lambda \ge 2$, since $n + t_\lambda$ must be even.

The lower bound of B, in (4.4), is known (see [2, Corollary 3]). The upper bound has been proved by [18, Theorem 3]. It cannot be reached here, since we must have $B \equiv 2 \pmod 4$, from (4.5). □

By the study of bent components, one understand precisely the difference between odd and even cases. Let F be a crooked function on $\mathbb{F}_{2^n}$, with components f_λ. We have seen that, when n is odd, there is a one to one correspondence $\lambda \mapsto a(\lambda)$, where $a(\lambda)$ is the unique linear structure of f_λ. When n is even, the linear space of f_λ has dimension s, where s is even. So, either f_λ is bent ($s = 0$) or $s \geq 2$. The number of hyperplanes involved in the structure of F is at most $(2^n - 1)/3$. For any a, there is one and only one λ such that $\mathcal{I}m(D_aF)$ is equal to H_λ or to $\overline{H_\lambda}$. But, the function D_af_λ is constant for any a belonging to the linear space of f_λ. There are at least three such nonzero a, when f_λ is not bent.

Proposition 3. *Let F be a crooked function on $\mathbb{F}_{2^n}$, n even, with components f_λ. For any $a \in \mathbb{F}_{2^n}^*$, there is a unique λ such that the derivative of f_λ, in point a, is a constant function.*

Conversely, any function f_λ is either bent or with linear space V of dimension $k \geq 2$.

Proof. By hypothesis, $\mathcal{I}m(D_aF)$ is a hyperplane H_λ, or the complement of H_λ, for any $a \in \mathbb{F}_{2^n}^*$. Hence $\mathrm{Tr}(\lambda D_aF(x)) = c$, for all x, where $c \in \mathbb{F}_2$. Such a λ is unique, because otherwise $\mathcal{I}m(D_aF)$ would be an affine subspace of codimension 2.

Since F is a plateaued APN function, at least $2(2^n - 1)/3$ components of F are bent. Let f_λ be a nonbent component. Its linear set has an even dimension, so this dimension must be greater than or equal to 2. □

5 Bent functions from crooked functions

In this section, we consider crooked functions over $\mathbb{F}_{2^n}$, where n is odd. The components of such a function are *near-bent* Boolean functions. Moreover, every component has (only) one derivative which is a constant function. (see Theorem 4). We will show that a set of $2^n - 1$ bent functions of $n + 1$ variables can be derived, using this strong property. Our main reference is the construction of Leander and McGuire [17], on the near-bent Boolean functions, that we apply to vectorial functions which are crooked. We first recall some facts which are more or less known, maybe not in this form.

Lemma 3. *Let f be a near-bent Boolean function over $\mathbb{F}_{2^n}$, where n is odd. Assume that $f(0) = 0$. Then* (i) *and* (ii) *are equivalent:*

(i) *f has a constant derivative in point $a \in \mathbb{F}_{2^n}^*$.*

(ii) *There exists a such that*

$$\{u \in \mathbb{F}_{2^n} \mid \mathcal{W}_f(u) = 0\} = \{u \in \mathbb{F}_{2^n} \mid \mathrm{Tr}(ua) = 1 + f(a)\}.$$

Proof. Note that, f cannot have more than one linear structure, since it is near-bent. Assume that there is (a unique) a such that $D_af(x) = c$, for all x, where $c \in \mathbb{F}_2$. Thus,

$c = f(0) + f(a) = f(a)$. Now, we compute $\mathcal{W}_f(u)$:

$$\begin{aligned}\mathcal{W}_f(u) = \sum_{x\in\mathbb{F}_{2^n}} (-1)^{f(x)+\mathrm{Tr}(ux)} &= \sum_{x\in\mathbb{F}_{2^n}} (-1)^{f(x+a)+\mathrm{Tr}(u(x+a))}\\ &= \sum_{x\in\mathbb{F}_{2^n}} (-1)^{f(x)+f(a)+\mathrm{Tr}(u(x+a))}\\ &= (-1)^{f(a)+\mathrm{Tr}(ua)}\mathcal{W}_f(u),\end{aligned}$$

since $f(x) + f(x + a) = f(a)$. Clearly, if u is such that $\mathrm{Tr}(ua) + f(a) = 1$ then $\mathcal{W}_f(u) = 0$. But, there are 2^{n-1} such u, proving that (ii) holds.

Now suppose that (ii) holds, for some a. Recall that

$$H_a = \{u \in \mathbb{F}_{2^n} \mid \mathrm{Tr}(ua) = 0\} \quad \text{so that } H_a^{\perp} = \{0, a\}.$$

Thus, the set of those u such that $\mathcal{W}_f(u) = 0$ is either equal to the hyperplane H_a, or equal to its complement, according to either $f(a) = 1$ or $f(a) = 0$. So, we can apply Lemma 1:

$$\sum_{v\in H_a} \mathcal{W}_f(v)^2 = 2^{n-1}\Big(\sum_{x\in\mathbb{F}_{2^n}} (-1)^0 + \sum_{x\in\mathbb{F}_{2^n}} (-1)^{f(x)+f(x+a)}\Big).$$

By hypothesis, this sum equals 0 if $f(a) = 1$ and 2^{2n} if $f(a) = 0$, since the Parseval's relation. This is possible only if $D_a f$ is a constant derivative of f. It is the only one such derivative, since f is near-bent. □

Lemma 4. *Let f be a near-bent Boolean function on $\mathbb{F}_{2^n}$ (n odd) such that $f(0) = 0$. Assume that f has a linear structure a. Let g be the function from $G = \mathbb{F}_{2^n} \times \mathbb{F}_2$ to $\mathbb{F}_2$:*

$$g(x, y) = (y + 1)f(x) + y\big(f(x) + \mathrm{Tr}(a^{-1}x)\big) \tag{5.1}$$

Then g is a bent function of $n + 1$ variables.

Proof. For any u, we denote by f_u the Boolean function $x \mapsto f(x) + \mathrm{Tr}(ux)$. Let $b = a^{-1}$. The restriction of g to $\mathbb{F}_{2^n}$ ($y = 0$) and to its complement in G ($y = 1$) are respectively f and f_b which are both near-bent Boolean functions. Moreover,

$$D_a f(x) = f(a) \quad \text{and} \quad D_a f_b(x) = f(a) + \mathrm{Tr}(ba) = f(a) + 1 = f_b(a),$$

since $\mathrm{Tr}(ba) = \mathrm{Tr}(1) = 1$. Let $u \in \mathbb{F}_{2^n}^*$. Applying Lemma 3, f_u is balanced if and only if $\mathrm{Tr}(ua) = 1 + f(a)$ and f_{b+u} is balanced if and only if

$$\mathrm{Tr}\big(a(u + b)\big) = 1 + f(a), \quad \text{providing } \mathrm{Tr}(ua) = 1 + 1 + f(a) = f(a).$$

Thus f_u is balanced if and only if f_{b+u} is not balanced. This is equivalent to say that g is bent (see [6, Theorem V.3] or [17, Theorem 2]). □

Now, we consider a crooked function F, from $\mathbb{F}_{2^n}$ to itself. According to Theorem 4, we know that all components f_λ of F are near-bent with a (unique) constant derivative. This allows to derive a specific set of bent functions from any crooked function, in odd dimension.

Theorem 6. *Let F be a crooked function over $\mathbb{F}_{2^n}$ where n is odd, such that $F(0) = 0$. Denote by a_λ the linear structure of the component f_λ of F. Then we get a set $B(F)$ of $2^n - 1$ bent functions g_λ, each from $G = \mathbb{F}_{2^n} \times \mathbb{F}_2$ to $\mathbb{F}_2$:*

$$B(F) = \{g_\lambda(x,y) = (y+1)f_\lambda(x) + y(f_\lambda(x) + \mathrm{Tr}(a_\lambda^{-1}x)) \mid \lambda \in \mathbb{F}_{2^n}^*\}.$$

Proof. Since F is a crooked function, there is a bijective correspondence between the $\lambda \in \mathbb{F}_{2^n}^*$, and then the functions f_λ, and the linear structures a_λ of every f_λ. Thus, Lemma 4 applies for every λ, using a_λ^{-1}, providing $2^n - 1$ distinct bent functions. □

Remark 1. The construction, given by Theorem 6, could be of interest regarding the problem of the existence of crooked functions. It would be useful to exhibit some properties of $B(F)$, or to construct other sets of bent functions. For instance, g_λ could have an higher degree, by replacing f_λ by another Boolean function, affinely equivalent to f_λ. The question is to know if such a set of bent functions does exist, for some degree greater than 2.

Note that a set like $B(F)$ is not a subspace, but could generate a linear code with few weights.

To derive effectively bent functions from a given crooked function, it is necessary to have an expression of the linear structures a_λ. This is generally an open problem, even when F is a quadratic APN function.

Example 2. Let $F(x) = x^{2^k+1}$, $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, n odd. It is well known that F is crooked if and only if $\gcd(k,n) = 1$. For any $\lambda \in \mathbb{F}_{2^n}^*$, the unique linear structure of f_λ is $a_\lambda = \lambda^{-1/(2^k+1)}$. Thus, for any $\lambda \in \mathbb{F}_{2^n}^*$,

$$g_\lambda(x,y) = (y+1)\,\mathrm{Tr}(\lambda x^{2^k+1}) + y(\mathrm{Tr}(\lambda x^{2^k+1}) + \mathrm{Tr}(\lambda^{1/(2^k+1)}x)),$$

is a bent function, $(x,y) \mapsto g_\lambda(x,y)$, from $\mathbb{F}_{2^n} \times \mathbb{F}_2$ to $\mathbb{F}_2$.

6 Permutations from crooked functions

Any crooked function allows to construct a set of permutations, via their derivatives.

Theorem 7. *Let $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, such that $F(0) = 0$. Assume that F is a crooked function. Define, for any $a \in \mathbb{F}_{2^n}^*$:*

$$x \mapsto G_a(x) = F(x) + F(x+a) + F(a).$$

Let us define:

- $\lambda \in \mathbb{F}_{2^n}^*$, *defining the hyperplane* H_λ, *which is the image set of* G_a*;*
- β *be such that* $\mathrm{Tr}(\lambda\beta) = 1$, *i. e.*, $\beta \notin H_\lambda$*;*
- H_μ *be any hyperplane such that* $\mathrm{Tr}(\mu a) = 1$, *i. e.*, $a \notin H_\mu$.

Then the function

$$x \mapsto R_a(x) = G_a(x) + \beta\,\mathrm{Tr}(\mu x)$$

is a permutation, such that $R_a(H_\mu) = H_\lambda$.

Proof. Let $x \neq y$ such that $R_a(x) = R_a(y)$. We get

$$G_a(x) + G_a(y) = \beta(\mathrm{Tr}(\mu(x+y))).$$

If $\mathrm{Tr}(\mu(x+y)) = 0$, then $G_a(x) + G_a(y) = 0$, which implies $y = x + a$, since G_a is 2-to-1 and $G_a(x) = G_a(x+a)$. In this case, we get $\mathrm{Tr}(\mu(x+y)) = \mathrm{Tr}(\mu a) = 0$, a contradiction.

Now suppose that $\mathrm{Tr}(\mu(x+y)) = 1$. Hence $G_a(x) = G_a(y) + \beta$. But this is impossible, since $G_a(x)$ and $G_a(y)$ belong to H_λ while $\beta \notin H_\lambda$. Thus, we get again a contradiction and can conclude that R_a is a permutation.

By construction, $R_a(H_\mu) = H_\lambda$, since $\mathrm{Tr}(\mu x) = 0$ for all $x \in H_\mu$ and $G_a(H_\mu) = H_\lambda$. Indeed, G_a is 2-to-1 and its image set is H_λ. For any pair $(x, x+a)$, we have $x \in H_\mu$ if and only if $x + a \in \overline{H_\mu}$ (with $G_a(x) = G_a(x+a)$). □

7 Comments to conclude

As a prelude of this conclusion, we want to recall the main conjecture concerning crooked functions:

Conjecture. Let $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ be a crooked function, as defined by Definition 3. Then F is quadratic or, in other terms, has algebraic degree 2.

By the two previous sections, we aim to exhibit specific constructions which are possible only with crooked functions. Our purpose is to open new ways to study the existence of crooked functions. We are convinced that other properties have to be found, increasing the knowledge on crooked functions, especially on the existence of such functions.

On the other hand, the quadratic functions form a corpus of great interest about which many problems remain open. In particular, we are still far from understanding the different structures of quadratic APN functions of even dimension. The determination of the number of bent components of such a function would be of great interest. Negative answers, concerning the APN property of quadratic vectorial functions, have been obtained, for instance, in [2, 5]. In [3], a description of the corpus of binomial crooked functions is presented. As a conclusion, Bierbrauer has conjectured that, up to equivalence, no other binomial crooked functions exist.

Crooked functions of codimension k, $1 \le k \le n-1$, appeared during the cryptanalysis of a hash function, called MARACA [8]. Here, the differential sets are affine subspaces of same codimension. The authors of this cryptanalysis, Canteaut and Naya-Placienta, introduced the *crooked property*. They have shown that this property could make a cryptographic primitive very weak. We will explore, in a more general context, the *functions having the crooked property*, in a forthcoming work.

Bibliography

[1] T. Bending and D. Fon-Der-Flaass, Crooked functions, bent functions, and distance regular graphs, *Electron. J. Comb.*, **5**(1) (1998), R34.

[2] T. P. Berger, A. Canteaut, P. Charpin, and Y. Laigle-Chapuy, On almost perfect nonlinear functions, *IEEE Trans. Inf. Theory*, **52**(9) (2006), 4160–4170.

[3] J. Bierbrauer, A family of crooked functions, *Des. Codes Cryptogr.*, **50** (2009), 235–241.

[4] J. Bierbrauer and G. Kyureghyan, Crooked binomials, *Des. Codes Cryptogr.*, **46** (2008), 269–301.

[5] E. Byrne and G. McGuire, On the non-existence of quadratic APN and crooked functions on finite fields, In: Proc. of the Workshop on Coding and Cryptography, WCC 2004, pp. 316–324.

[6] A. Canteaut, C. Carlet, P. Charpin, and C. Fontaine, On cryptographic properties of the cosets of $R(1, m)$, *IEEE Trans. Inf. Theory*, **47**(4) (2001), 1494–1513.

[7] A. Canteaut and P. Charpin, Decomposing bent functions, *IEEE Trans. Inf. Theory*, **49**(8) (2003), 2004–2019.

[8] A. Canteaut and M. Naya-Plasencia, Structural weaknesses of permutations with a low differential uniformility and generalized crooked functions, In Finite Fields: Theory and Applications - FQ9 - Contemporary Mathematics, AMS, Vol. 518, 2010, pp. 55–71.

[9] C. Carlet, Partially-bent functions, *Des. Codes Cryptogr.*, **3** (1993), 135–145.

[10] C. Carlet, P. Charpin, and V. Zinoviev, Codes, bent functions and permutations suitable for DES-like cryptosystems, *Des. Codes Cryptogr.*, **15**(2) (1998), 125–156.

[11] P. Charpin and J. Peng, Differential uniformity and the associated codes of cryptographic functions, In: Advances in Mathematics of Communications, AIMS, Vol. 13(4), Special issue on Applications of discrete mathematics in secure communication, Ed. by S. Maitra, 2019, pp. 579–600.

[12] E. R. van Dam and D. Fon-Der-Flaass, Uniformly packed codes and more distance regular graphs from crooked functions, *J. Algebraic Comb.*, **12**(2) (2000), 115–121.

[13] E. R. van Dam and D. Fon-Der-Flaass, Codes, graphs, and schemes from nonlinear functions, *Eur. J. Comb.*, **24**(1), (2003) 85–98.

[14] C. Godsil and A. Roy. Two characterization of crooked functions, *IEEE Trans. Inf. Theory*, **54**(2) (2008), 864–866.

[15] G. Kyureghyan, The only crooked power functions are $x^{2^k+2^l}$, *Eur. J. Comb.*, **28** (2007), 1345–1350.

[16] G. Kyureghyan, Crooked maps in F_{2^n}, *Finite Fields Appl.*, **13**(3) (2007), 713–726.

[17] G. Leander and G. McGuire, Construction of bent functions from near-bent functions, *J. Comb. Theory, Ser. A*, **116** (2009), 960–970.

[18] A. Pott, E. Pasalic, A. Muratovic-Ribic, and S. Bajric, On the maximum number of bent component of vectorial functions, *IEEE Trans. Inf. Theory*, **64**(1) (2018), 403–411.

Michael D. Fried

Diophantine statements over residue fields: Galois stratification and uniformity

Abstract: We consider three general problems about Diophantine statements over finite fields that connect to the *Galois stratification* procedure for deciding such problems. To bring to earth the generality of these problems, we clarify each using the negative solution of U. Felgner's simply stated question relating pairs of finite fields $\mathbb{F}_p$ and $\mathbb{F}_{p^2}$ for large primes p.

For a diophantine problem, D, interpretable for almost all primes p, the paper plays on attaching a Poincaré series, $P_{D,p}$, to (D, p). The rationality and computability of $P_{D,p}$ as p varies gives some aspects of continuity. Most interesting, though, is how the coefficients of $P_{D,p}$ vary with p. The paper notes three ways to give counts for those coefficients:

- points mod p achieving particular conjugacy classes in a Galois stratification;
- points mod p on absolutely irreducible varieties; or
- traces of the p Frobenius on a Chow motive.

The gist of this paper, using one explicit well-known diophantine problem, is that Galois stratification naturally tethers these three abstract approaches. Among [19], [20], [10], and [22] is a history of the fundamental ideas and original motivations for extending the Galois stratification procedure. Others, [38], [28], and [36] referred to in Section 4 show a more extensive set of motivations for extending the topic today. So, it is time for an introduction relating the ideas to Deligne's proof of the Weil Conjectures and Langland's Program applied to two specific recognizable problems/examples.

Keywords: Galois stratification, field arithmetic, Dwork cohomology, ℓ-adic cohomology, Chow motives, Chebotarev density, Davenport's problem, Felgner's problem, Poincaré series, difference fields

MSC 2010: Primary 11G10, 12E30, 14F30, Secondary 11D72, 14G32

1 Three general problems on finite field equations

Traditional diophantine statements consider algebraic subsets of affine space with blocks of variables with some, say $\boldsymbol{x} = (x_1, \dots, x_m)$, unquantified, and others, say $\boldsymbol{y} = (y_1, \dots, y_n)$, quantified by the symbols $\exists$ and $\forall$. Questions over finite fields often assume the equation coefficients are in $\mathbb{Z}$. For example, they might consider $\boldsymbol{x}$ and

Michael D. Fried, Emeritus Full Professor of Mathematics, University of California at Irvine, USA, e-mail: michaeldavidfried@gmail.com

https://doi.org/10.1515/9783110621730-008

$\boldsymbol{y}$ taking values in the residue class fields $\mathbb{F}_p$, and ask if the statements are true for almost all p.

1.1 Introducing the three problems

Long time results on variants using *Galois stratification* (§2.1) have detailed literature as in [22, Chapters 30–31]. There are, however, three topics inadequately treated there.

(1.1a) Testing how a quantified D over distinct residue fields cohere – *uniformity in p* – using equivalent data from reducing an unquantified object over $\mathbb{Q}$ mod most primes p.

(1.1b) Coordinating (*) affine space arguments and (**) arithmetic homotopy with scheme or *projective geometry* language (the natural domain of arithmetic geometry).

(1.1c) Considering statements with variables taking values in the algebraic closure of $\mathbb{F}_p$, but fixed by respective powers of the Frobenius Fr_p:

$$\text{for } \boldsymbol{x} \text{ (resp. } \boldsymbol{y})(\mathrm{Fr}_p^{d_1}, \dots, \mathrm{Fr}_p^{d_m}) \stackrel{\text{def}}{=} \mathrm{Fr}_p^{\boldsymbol{d}} \text{ (resp. } (\mathrm{Fr}_p^{e_1}, \dots, \mathrm{Fr}_p^{e_n}) = \mathrm{Fr}_p^{\boldsymbol{e}}). \tag{1.2}$$

To help understand aspects of all three problems, we use Felgner's problem:

Can we define the fields $\mathbb{F}_p$ within the theory of the fields $\mathbb{F}_{p^2}$?

Comments on notation and background

We assume the reader knows the meaning of the words *cover of normal quasi-projective varieties*, where cover may, usually here, includes ramification, but is a *finite, flat* morphism (Grothendieck's definition). Although [27] is much more comprehensive, we still find it valuable to refer to [33] for proofs through special illuminating cases. Especially for Segre's proof that normalization in a function field of a quasi-projective variety is also quasiprojective [33, p. 400]. The constructions alluded to in this paper also depend on Chevalley's theorem that *the image of a constructible set is constructible* [33, p. 97].

Neither is illuminating on the phrase *nonregular Galois cover* for a very good reason; both do almost everything over algebraically closed fields. The Galois stratification procedure confronts situations in which $\varphi : Y \to X$ is a cover of absolutely irreducible varieties over a (perfect) field K (with algebraic closure $\bar{K}$), while its Galois closure $\hat{\varphi} : \hat{Y} \to X$, $\hat{Y}$ is not absolutely irreducible over K. In other words, $K(\hat{Y}) \cap \bar{K}$ is a nontrivial extension of K.

Comments on (1.1a)

For any prime p in (1.1a), a *Galois stratification* allows eliminating the quantifiers on the variables, producing a quantifier-free statement. The $\mathbb{Q}$ object in (1.1a) – a Ga-

lois stratification object, but over $\mathbb{Q}$ rather than over a finite field – has not been previously emphasized, though it has always been present. Though more general than a pure algebraic variety, it allows all the elimination theory and reductions mod primes.

At the heart of the Galois stratification procedure is its use of the *nonregular Chebotarev analog*. We use the acronym **NRC**. Both the nonregular adjective and the Chebotarev error term impact its use, as in the negative solution of Felgner's problem. It appears in two ways:

(1.3a) Eliminating quantifiers, the main reason we introduced the stratification procedure; and

(1.3b) estimating the unquantified variable values satisfying the diophantine conditions.

We can untangle the two steps of (1.3). So doing reveals two separate aspects of the stratification procedure. Its use in (1.3a) for collecting finite fields points. Then its generalization of quantifier elimination over many other collections of fields, as appears in [22, Chapter 30] as well as the variant given in (1.1c).

Starting with an elementary diophantine statement (problem) D, the procedure generalizes elementary statements to Galois stratifications $\mathcal{S}$. In each generalizing case, every step of the stratification procedure produces a new Galois stratification. The number of steps depends on the number of blocks of quantified variables. Section 2.1 gives the notation and explains the main theorem as given in several sources starting from [19].

That is, suppose there are k blocks of quantified variables. For each prime p, this produces $\bar{\mathcal{S}}_p \stackrel{\text{def}}{=} \{\mathcal{S}_{p,k}, \dots, \mathcal{S}_{p,0}\}$, a sequence of Galois stratifications (starting at k going to 0), and a finite Galois extension $\hat{K}/\mathbb{Q}$ (with group G_P). Excluding a finite computable set of primes p, these have the following property. For p a prime unramified in $\hat{K}$, denote an element of the conjugacy class of the Frobenius at p by Fr_p, and its fixed field in $\hat{K}$ by $\hat{K}_p$.

> For each conjugacy class C of values of the Frobenius, $\mathrm{Fr}_p \in G_P$, there is an object $\bar{\mathcal{S}}_{\mathrm{C}}$ over $\hat{K}_p$ whose reduction mod a prime above p gives $\bar{\mathcal{S}}_p$. (1.4)

That gives a start to many aspects of uniformity in p.

One form of the stratification procedure moves through field theory – after all, *Field Arithmetic* is the [22] title – based on *Frobenius fields*. While we still rely on details on the stratification procedure, Section 2.2 gives an overview hitting the most significant of those details. Especially it shows the role of the Chebotarev analog.

Section 2.3 emphasizes the word *nonregular* in that analog and its subtleties. For example, this gives the main distinctions, running over classes C in G, between the collections $\bar{\mathcal{S}}_{\mathrm{C}}$.

Comments on (1.1b)

From the beginning, there was an idea of attaching a more concise object, a Poincare series $P_p \stackrel{\text{def}}{=} P_{D,p}$, to $\bar{\mathcal{S}}_p$. The stratification procedure was the main tool in showing this object to be a computable rational function. The tension between (*) and (**) in (1.1b) appeared long before [19] as recounted in [18] where this author's applications to finite field questions used Riemann's work and projective geometry. The tension continued:

– on one hand, in applying Bombieri's use of Dwork's affine/singular theory of the zeta function to explicitly compute stratification Poincaré series *characteristics*; and;
– on the other, the Denef–Loeser (and Nicaise) production of Poincaré series coefficients as Chow motives. Deligne's proof of the Weil conjectures is in the background.

Continuing the comments on (1.1a), the *nonregular* adjective shows in Theorem 2.11. It gives how the stratification procedure varies with either the prime p (or in applying it to a specific finite field, the extension $\mathbb{F}_q$ for q a power of p).

Section 2.3.3 applies the main theorem to Felgner's problem. Here – I suspect a careful reader will agree – a mere cardinality estimate, rather than some precise understanding of the variance with either the prime p or the extension $\mathbb{F}_q$ for q a power of p, seems inadequate.

Yet, there is considerable flexibility in how those Poincaré series capture information. Section 3.3.1 illustrates with an a diophantine example the author has used with undergraduates. (Proofs are another matter.) This example helps picture how the reference to Chow motives, as pieces of étale cohomology groups, works.

Comments on (1.1c)

Section 1.2 explains Felgner's problem in very down-to-earth language. Felgner is a piece from the case $m = 1$, $d_1 = 1$, with all e_is equal to 2. It poses for Galois stratifications – expanded by *Frobenius vectors* (like $\mathrm{Fr}_q^{\boldsymbol{d}}$) – the nature of corresponding Poincaré series. For example, when can they be derived from standard Poincaré series?

This latter story passes through work of D. Wan, who introduced a Zeta function in the unquantified case, generalizing Artin's zeta where all d_is are 1, and Hrushovski and Tomasic who took a model theoretic approach, enhanced by Galois stratification.[1]

From constraints on space and time, we limit our responses to questions (1.1b) and (1.1c) to merely motivating seriously using these Poincaré series, as if their coefficients are Chow motives to which you apply the Frobenius Fr_p. Our goal is to ask questions

1 This starts with Ax and Kochen on the Artin conjecture on p-adic points on forms of degree d in projective d^2-space. To understand my motivation, consider the down-to-earth problems I connected to classical arithmetic geometry as recounted in the partially expository [17] and [18].

about their variance with p. Especially for what p certain especially simple results occur, as given explicitly in our example. We expect to expand on parts of Section 4.2 in a later paper.

1.2 Felgner clarifies what are statements over finite fields

A traditional question in the theory of finite fields starts with blocks of variables $\boldsymbol{y}_i \in \mathbb{A}^{n_i}$, $i = 1, \dots, k$. Below we use the notation $\sum_{j=1}^{k} n_j \stackrel{\text{def}}{=} N_k$. Assume these blocks are respectively quantified by $Q_1, \dots, Q_k = \boldsymbol{Q}$. Also we have a separate unquantified block of variables $\boldsymbol{x} \in \mathbb{A}^m$ together with an algebraic set X described by

$$\{(\boldsymbol{x}, \boldsymbol{y}_1, \dots, \boldsymbol{y}_k) \mid \varphi(\boldsymbol{x}, \boldsymbol{y}_1, \dots, \boldsymbol{y}_k)\}.$$

Here, φ is a collection of polynomial equalities and inequalities with coefficients in a ring R, defining a constructible subset of $\mathbb{A}^{m+N_k}$. A traditional question is, given φ and $Q_1, \dots, Q_k$, and no $\boldsymbol{x}$ variables, is there a useful procedure for deciding if the statement is true for *almost all* finite fields $\mathbb{F}_q$ in the following collections:

(1.5a) $R = \mathbb{F}_p$: q are powers of a given p.

(1.5b) $R = \mathcal{O}_K$, ring of integers of a number field K: $q = |\mathcal{O}_K/\boldsymbol{p}|$, orders of residue class fields.

(1.5c) $R = \mathbb{Z}$, the whole Universe case: All $\mathbb{F}_q$.

Section 2.1 explains the main theorem of Galois stratification. It is a stronger result than that above in that our starting statement usually includes unquantified variables $\boldsymbol{x}$. Then the result is that the starting statement – with, excluding $\boldsymbol{x}$, all variables quantified – is equivalent to one with no quantified variables. That is, it gives an *elimination of quantifiers*. We use Felgner's problem [14] to show its value and how it works.

In each case, in changing "almost all" to "all," our method (and generalizations) leaves checking a finite number of prime powers. For some problems those exceptional qs could be accidents. For others – motivated by classical cases – they could be significant. The Poincaré series approach (§3.1), seeking a significant rational function, has been the gold standard.

Felgner – in language reminiscent of the above (with $m = 1$) – asks if given $x' \in \mathbb{F}_{p^2}$,

$$\text{is } (x', \boldsymbol{y}_1, \dots, \boldsymbol{y}_k) \in X \text{ true for } Q_1\boldsymbol{y}_1 \dots Q_k\boldsymbol{y}_k \in (\mathbb{F}_{p^2})^{N_k}, \quad \text{if and only if } x' \in \mathbb{F}_p?$$

Except we are asking about a different set – $\{\mathbb{F}_{p^2} \mid p \text{ prime}\}$ – of fields than those included in any of the lists (1.5). Our method easily tolerates such flexibility. Further, the negative conclusion holds even if we asked for its truth replacing p by prime-powers q ($\mathbb{F}_p \mapsto \mathbb{F}_q$, $\mathbb{F}_{p^2} \mapsto \mathbb{F}_{q^2}$) and just for *infinitely* many q.

The conclusion about the set of such x' opens territory that is not hinted at by the list (1.5). Still, the nature of that set arising with actual problems in finite fields (as in Section 3.3), is the model for which the original result aimed. Here is an example of what we will exclude. Take $k = 2$, and $\boldsymbol{Q} = (\exists, \forall)$. Read this as follows:

(1.6a) Given $x' \in \mathbb{F}_q$, there exists $\boldsymbol{y}_1' \in (\mathbb{F}_{q^2})^{n_1}$ with $\boldsymbol{y}_2' \in (\mathbb{F}_{q^2})^{n_2}$, so that $(x', \boldsymbol{y}_1', \boldsymbol{y}_2') \in X$.

(1.6b) But if $x' \in \mathbb{F}_{q^2} \setminus \mathbb{F}_q$, then for any $\boldsymbol{y}_1' \in (\mathbb{F}_{q^2})^{n_1}$ there is $\boldsymbol{y}_2' \in (\mathbb{F}_{q^2})^{n_2}$ with $(x', \boldsymbol{y}_1', \boldsymbol{y}_2') \notin X$.

> It seems so simple: $x \in \mathbb{F}_{q^2}$ is in $\mathbb{F}_q$ if and only if $\mathrm{Fr}_q(x) \stackrel{\mathrm{def}}{=} x^q = x$.
> How could it fail that some algebraic statement generalizing (1.6) would encapsulate this?

2 The main theorem of Galois stratification

The main theorem inductively, Section 2.1, eliminates quantifiers for problems that

> start as elementary statements interpretable over *any* extension of a residue class field of the ring of integers of a number field K. (2.1)

First, we give the appropriate main theorem version that allows the elimination. This includes all possible statements that might have answered Felgner's question in the affirmative. To simplify discussing the procedure, take as fundamental the following problem.

Is a statement true for *almost all* residue class fields $\mathbb{Z}/p = \mathbb{F}_p$ (i. e., $R = \mathbb{Z}$ in (1.5)).

Given the elementary statement start of (2.1), we expediently consider the *exceptional stratification* primes (Definition 2.2) not included in the phrase "almost all."[2] That thereby allows the procedure to consider if a statement holds for *all* primes or variants of that question.

The philosophical differences between [19] and [22] are small, though there are more details in the latter. That is primed to handle – *Frobenius field* – results more general than the finite field case. It shows there is a *general Galois stratification* idea. Then you adjust by using an **NRC** replacement for each collection of (so-called) Frobenius fields.

This approach produces the finite field case by observing, if q is large, then the collection of finite fields behaves like a collection of Frobenius fields.[3] We do not redo that. Rather we show a path through [22, Chapter 30] in Section 2.2 – with an example – that works.

2 The word *exceptional* has two, incompatible uses in this paper, forced on us by the historical record. I have put the word *stratification* in this one to separate them.

3 That is no surprise since that was in the original motivations for [22].

2.1 Galois stratification and elimination of quantifiers

The starting field for the diophantine statement D here is either $\mathbb{Q}$ (or some fixed field K containing $\mathbb{Q}$) or a finite field $\mathbb{F}_q$ (usually $\mathbb{F}_p$). To simplify, we assume that we start either with $\mathbb{Q}$ (or with coefficients in $\mathbb{Z}$) or with $\mathbb{F}_p$. Here is the procedure abstractly, starting with the definition of Galois stratification.

In the rest of this subsection, our goal is consider residue class fields. Then Section 2.2 generalizes, as done in [22], but with different emphasis, how that gives objects over $\mathbb{Q}$.

(2.2a) There exists a stratification $\mathcal{S}_k$ (a disjoint union of normal algebraic sets)

$$\bigcup_{j=1}^{l_k} X_{j,k}, \quad \text{covering } \mathbb{A}^{m+N_k}.$$

(2.2b) For each (j,k), $1 \le j \le \ell_k$, there is a Galois (normal) cover $\hat{\varphi}_{j,k} : \hat{Y}_{j,k} \to X_{j,k}$[4] with group $G_{j,k}$ and a collection, $\mathbf{C}_{j,k}$, of conjugacy classes of $G_{j,k}$.[5]

(2.2c) Then, in place of $(x', \boldsymbol{y}'_1, \boldsymbol{y}'_2) \in$ (or $\notin$)$X(\mathbb{F}_q)$ we ask this. For $(\boldsymbol{x}', \boldsymbol{y}'_1, \dots, \boldsymbol{y}'_k) \in \mathcal{S}_k(\mathbb{F}_q)$,

> if $(\boldsymbol{x}', \boldsymbol{y}'_1, \dots, \boldsymbol{y}'_k) \in X_{j,k}$, then the Frobenius, $\mathrm{Fr}_{\boldsymbol{x}', \boldsymbol{y}'_1, \dots, \boldsymbol{y}'_k} \in G_{j,k}$ at the point, is in $\mathbf{C}_{j,k}$.

Ambiguity in the coefficients of equations

In [22], the covers are always taken to be unramified, but the essential is whether having the Frobenius in the conjugacy classes is unambiguous. We have purposely left ambiguous the coefficients of the equations defining the ingredients of (2.2a) and (2.2b). Indeed, the procedure works like this:

(2.3a) $\mathrm{G}_{\mathrm{aap}}$: The original statement is over $\mathbb{Q}$ (or even over $\mathbb{Z}$), but we do not care, for finitely many p, if it is not true or even for (2.2c) applicable.

(2.3b) $\mathrm{G}_{\mathrm{allp}}$: The original statement is over $\mathbb{Z}$, meaningful for each prime p and we have the option of deciding if, after finding it true for a. a. p, whether it is true for all p.

Definition 2.1. For each Galois cover $\hat{\varphi} : \hat{Y} \to X$ in a Galois stratification $\mathcal{S}$, we need an affine space description of the underlying spaces, and of the group of the cover. Running over all possible (j,k), we refer to this collection as the *characteristics* of $\mathcal{S}$.

It is easiest to describe the characteristics when the spaces are described as explicit open subsets of affine hypersurfaces, as they are in [22]. There are two possible

4 By putting a ˆ over φ we are reminding that this is a Galois cover with automorphisms defined over the field of definition of the cover.

5 For the problems in this paper, it suffices to take *Frobenius classes*. Each is a union of conjugacy classes of $G_{j,k}$ where if g is in one of these, then so is g^u if $(u, \mathrm{ord}(g)) = 1$.

situations where we might apply the stratification procedure assuming our initial goal is to decide if we can interpret it over $\mathbb{F}_p$ by reduction mod p, and then if it is true for almost all p.

These situations are compatible with the main theorem (2.4). It uses coefficients initially over $\mathbb{Q}$, producing a finite set, M_{k-1}, of primes for which (2.4) holds.

(2.4a) Either the resulting (over) $\mathbb{Q}$ equations do not make sense of (2.2c); or the fibers of the underlying stratification pieces are not flat over $\mathrm{Spec}(\mathbb{Z}_p)$; or

(2.4b) the Chebotarev estimates of (2.11) for some q divisible by p do not assure hitting all appropriate conjugacy classes.

The *flatness statement* means that reduction mod p preserves the characteristics of the stratification. In applying Denef–Löeser Section 3.3, what we call the characteristics is strengthened, but the idea is the same.

Definition 2.2. In the ith step of the induction procedure, refer to the respective primes, the *exceptional stratification primes* of (2.4), as $M_{i,a}$ and $M_{i,b}$, $i = k{-}1, \dots, 0$, and their union by M_i. We include in M_i the primes from M_{i+1}. So $M_0 \supset M_1 \supset \cdots \supset M_{k-1}$.

It is for the strong Poincaré series result of Theorem 3.1 that we have divided these considerations of primes. Following the stratification procedure format in Theorem 2.4, Section 2.3 gets into these significant issues: Reduction mod p; and the nonregular Chebotarev analog.

The extra point, not discussed in [22, Chapter 31], is how the diophantine illuminating properties of the Poincaré series in Theorem 3.1 vary as a function of $p \notin M_i$. Especially, how they depend on the nonregular analog as illustrated by the Section 3.3.1 example.

Inductive stratification notation

Below some additional notation replaces k by the index i. For example, we replace N_k by $N_i = \sum_{j=1}^{i} n_j$. Assume the problem has one quantifier for each block of variables. Our original example, illustrating a possible answer in Section 1.2 for Felgner, used two quantifiers.

Galois stratification allows eliminating one quantified block of variables, at a time, using a general, enhanced, *Nonregular Chebotarev Density Theorem*. Section 2.3 explains how that contributes to the main theorem (2.4) by which the elimination inductively forms a sequence of Galois stratifications: $\mathcal{S}_k, \dots \mathcal{S}_0$, with triples $(X_{\bullet,u}, G_{\bullet,u}, \mathbf{C}_{\bullet,u})$, $u = k, \dots, 0$.

For each u, the $\bullet$ indicates a sequence for $1 \le j \le \ell_u$ of underlying spaces in the stratification that satisfy the following properties.

As in (2.2b), saying $\mathrm{Fr}_{\boldsymbol{x}', \boldsymbol{y}'_1, \dots, \boldsymbol{y}'_i} \in \mathbf{C}_{\bullet,i}$ means that if the subscript point is in $X_{j,i}$, then $\mathrm{Fr}_{\boldsymbol{x}', \boldsymbol{y}'_1, \dots, \boldsymbol{y}'_i}$ refers to the conjugacy class value in the Galois cover $\hat{C}_{j,i} \to X_{j,i}$.

The variables $\boldsymbol{x}$ of the Galois stratification $(X_{\bullet,0}, G_{\bullet,0}, \mathbf{C}_{\bullet,0})$ of $\mathbb{A}^m$ are unquantified. For the ith stratification $\mathcal{S}_i$, the quantifiers are $Q_1, \dots, Q_i$; we suppress their notation.

Definition 2.3. For $\boldsymbol{x}' \in \mathbb{F}_q^m$, we say $\mathcal{S}_i(\boldsymbol{x}'; q)$ holds

$$\text{if } \mathrm{Fr}_{\boldsymbol{x}', \boldsymbol{y}_1', \dots, \boldsymbol{y}_i'} \in \mathbf{C}_{\bullet,i} \text{ for } Q_1 \boldsymbol{y}_1' \cdots Q_i \boldsymbol{y}_i' \in \mathbb{A}(\mathbb{F}_q)^{N_i}. \tag{2.5}$$

A relative version concentrates at the ith position. For $\boldsymbol{x}', \boldsymbol{y}_1', \dots, \boldsymbol{y}_{i-1}' \in \mathbb{F}_q^{m+N_{i-1}}$

$$\text{we say } \mathcal{S}_i(\boldsymbol{x}', \boldsymbol{y}_1', \dots, \boldsymbol{y}_{i-1}'; q) \text{ holds if } \mathrm{Fr}_{\boldsymbol{x}', \boldsymbol{y}_1', \dots, \boldsymbol{y}_i'} \in \mathbf{C}_{\bullet,i} \text{ for } Q_i \boldsymbol{y}_i' \in \mathbb{A}(\mathbb{F}_q)^{n_i}. \tag{2.6}$$

The notation $\bar{\mathcal{S}}_p \overset{\text{def}}{=} \{\mathcal{S}_{p,k}, \dots, \mathcal{S}_{p,0}\}$ from (1.4), for a sequence of Galois stratifications – with its attendant sequence of exceptional stratification primes $M_k \dots, M_0$ – appears at the end of the Theorem 2.4 statement. Also, Theorem 2.4 refers to the Frobenius in a finite Galois extension $\hat{K}/\mathbb{Q}$, and the Frobenius (conjugacy class), Fr_p attached to a prime p in this extension.

Theorem 2.4 (Main theorem [19]). *Assume quantifiers $Q_1, \dots, Q_k$ and, $\mathcal{S}_{k,p}$, an initial Galois stratification over $\mathbb{Q}$, in situation* (2.3a). *Then we may effectively form a stratification sequence $\bar{\mathcal{S}}_p$, as above, for each $p \notin M_0$ with the following properties:*

(2.7a) *Local property: Given $\boldsymbol{x}', \boldsymbol{y}_1', \dots, \boldsymbol{y}_{i-1}' \in \mathbb{F}_p^{m+N_{i-1}}$, $\mathrm{Fr}_{\boldsymbol{x}', \boldsymbol{y}_1', \dots, \boldsymbol{y}_{i-1}} \in \boldsymbol{C}_{\bullet,i-1}$ if and only if $\mathcal{S}_i(\boldsymbol{x}', \boldsymbol{y}_1', \dots, \boldsymbol{y}_{i-1}'; p)$ holds.*

(2.7b) *Inductive result: $\boldsymbol{x}' \in \mathbb{F}_p^m$, then $\mathrm{Fr}_{\boldsymbol{x}'} \in \boldsymbol{C}_0$ if and only if $S_k(\boldsymbol{x}', p)$ holds.*

Further, there is a finite Galois extension $\hat{K}/\mathbb{Q}$ with group $G \overset{\text{def}}{=} G_{\mathcal{S}_k}$ and a finite collection of stratification sequences ${}_j\bar{\mathcal{S}}$, $j = 1, \dots, \mu$, over $\hat{K}_p$ (see Section 2.2) with this property. For each $p \notin M_0$, $\bar{\mathcal{S}}_p$ is the reduction of ${}_j\bar{\mathcal{S}}$ mod p for some j depending only on Fr_p in G.

Remark 2.5. The last paragraph in Theorem 2.4 is there to compare the stratification procedure for a given p to a process that works uniformly in p. We are giving several approaches: the Frobenius field approach (Section 2.2); the finite field approach, and sometimes just deciding the truth of a diophantine statement for p. Still, all of them – including the ultimate count, as in the abstract, in the Poincaré series – depend on locating absolutely irreducible varieties related to the stratification procedure.

Remark 2.6 (Extend Theorem 2.4). For a given p, we may replace the quantification of the variables in $\mathbb{F}_p$ by quantification of its variables in $\mathbb{F}_q$ with q any power of p. Again, the correct ${}_j\bar{\mathcal{S}}$ depends only on $\mathrm{Fr}_q \in G$. See Remark 2.7 for the adjustment that applies to considering powers of q divisible by $p \in M_0$.

Remark 2.7. Notation of Definition 2.2 gives a procedure starting with the subscript $i = k$, ending with $i = 0$. Even, however, for primes in M_0 under hypothesis $\mathrm{G}_{\mathrm{allp}}$ (2.3b), the equation coefficients are in $\mathbb{F}_p$. We may then perform the stratification procedure. That allows deciding for each such p if the statement is true over $\mathbb{F}_q$ for *almost all*

powers q of p, so long as q is large enough for the Chebotarev conditions to hold. This extends Remark 2.6. We cannot, however, expect the stratification to come from an object over $\mathbb{Q}$ (by reduction mod p) or to be related in any obvious way to one of the stratifications ${}_j\bar{\mathcal{S}}$ given in Theorem 2.4.

2.2 The general stratification procedure

The qualitative point of the **NRC** is this.

> To know, for a Galois cover $\hat{\varphi} : \hat{Y} \to X$ over a finite field, what conjugacy classes (or cyclic subgroups) generate decomposition groups running over $\boldsymbol{x} \in X(\mathbb{F}_q)$. (2.8)

Section 2.3 shows the role of the Chebotarev theorem in Theorem 2.4. Especially that it gives a rough count of the $\boldsymbol{x}' \in \mathbb{F}_q^m$ for which $\mathcal{S}_k(\boldsymbol{x}', q)$ holds. This section outlines how [22, Chapters 30–31] gives Theorem 2.4. We trace the general form of the stratification procedure.

Definition 2.8. A profinite group G has the *embedding property* if given covers[6] of groups $\psi_{G,A} : G \to A$ and $\psi_{B,A} : B \to A$, for which *B is a quotient of G*, then

$$\text{there is a cover } \psi_{G,B} : G \to B \text{ with } \psi_{G,A} \circ \psi_{G,B} = \psi_{G,A}.$$

The name *superprojective* is thus apt for a group that is projective – in the category of profinite groups – with the embedding property.

That is, the embedding property is just projectivity with the extra condition that $\psi_{G,B}$ is a cover, given the necessary condition on B. This definition starts the analogy between finite fields and the very large class of Frobenius fields – again a reason for the existence of [22] – and whereby progress was achieved using Galois stratification [20].

[22, Section 24.1] develops the theory of *Frobenius fields M*:
(2.9a) M is a **PAC** field (all absolutely irreducible varieties over M have M points), and;
(2.9b) its absolute Galois group, G_M, has the embedding property.

> If M is Frobenius, then G_M is *superprojective* [22, Proposition 24.1.5].

The point of Frobenius fields

[22, Proposition 24.1.4] is the replacement for Frobenius fields of the Chebotarev analog. This is applied to the Galois covers $\hat{\varphi} : \hat{Y} \to X$ with group $G_{\hat{\varphi}}$ that appear in a

6 By a cover $\psi : H \to G$ of groups, we mean an onto homomorphism.

Galois stratification. This assumes $\hat{Y}$ is absolutely irreducible. Equivalently, $\hat{\varphi}$ is a *regular* cover, over M.

The conclusion of this Chebotarev analog is that we know precisely what *decomposition groups* – among the $H \le G_{\hat{\varphi}}$ – that $\hat{\varphi}$ has over fibers of $X(M)$. They are:

All subgroups H that are quotients of the absolute Galois group G_M.

The remainder of Chapter 30 through Section 30.5 is details of the Galois stratification procedure applied to the theory of *all* Frobenius fields containing a fixed given field [22, Theorem 30.6.1]. For simplicity, assume we start over $\mathbb{Q}$. Much of what looks technical is the pairing of decomposition groups of a cover $\hat{\varphi}_i$ of $\mathcal{S}_i$ with the decompositions groups that extend these to a given cover $\hat{\varphi}_{i+1}$ of $\mathcal{S}_{i+1}$ above $\hat{\varphi}_i$.

For example, [22, conditions (2) of Lemma 30.2.1] has this situation: $n_{i+1} = 1$ and $\boldsymbol{y}_{i+1} = y_{i+1}$. We continue Example 2.9 in Section 3.3.1. Again, M is notation for a Frobenius field.

Example 2.9 (Pairing two covers). Suppose X_2 is a hypersurface in $\mathbb{A}^{m+1}$ defined by the equation $\{(\boldsymbol{x}, y) \mid f(\boldsymbol{x}, y) = 0\}$ covered by $\hat{\varphi}_2 : \hat{C}_2 \to X_2$. Also, that $\mathrm{proj}_{\boldsymbol{x}} : \mathbb{A}^{m+1} \to \mathbb{A}^m$ restricted to X_2 is generically onto. So, it is a cover – not likely Galois in a practical case – over an open set in the image. Assume f is absolutely irreducible of degree u in y.

Take a Zariski open set X_1 in $\mathbb{A}^m$ and a Galois cover $\hat{\varphi}_1 : \hat{C}_1 \to X_1$ having an open map $\varphi : \hat{C}_1 \to \hat{C}_2 \to X_2$ factoring through X_1. To satisfy other constraints in [22], like being unramified, it may not be surjective to X_2. Note: $G_{\hat{\varphi}_1}$ has a natural degree u permutation representation $T : G_{\hat{\varphi}_1} \to S_u$ from its birational factorization through $\hat{\varphi}_2$.

The essence: Take for $\hat{\varphi}_1$ the minimal Galois closure of the cover $\mathrm{proj}_{\boldsymbol{x}} \circ \hat{\varphi}_2$. Here is what we want from the elimination of quantifiers, if say y was quantified by $\exists$.

Given a Frobenius field M, consider subgroups $H_1 \le G_{\hat{\varphi}_1}$ that are decomposition groups for $\hat{\varphi}_1$ over some $\boldsymbol{x}' \in X_1(M)$ that also fix a point $(\boldsymbol{x}', y') \in X_2(M)$. That is, H_1 fixes a letter in the representation T, so it restricts to a decomposition group H_2' of $\hat{\varphi}_2$. Then consider $\mathcal{H}$, the conjugacy classes of groups $H_2 \le G_{\hat{\varphi}_2}$ in Galois stratification data already attached to this cover in the Frobenius field case.

How to decide – for eliminating the quantifier $\exists$ – if H_1 will be in the conjugacy classes attached to the cover $\hat{\varphi}_1$? It is, if the H_2' described above corresponding to H_1, is in $\mathcal{H}$. □

Remark 2.10 (Finding H_1 in Example 2.9). The last line on finding those H_1s is what the Chebotarev analog quoted above does. In detail, you construct an absolutely irreducible variety for which an M point corresponds to such an H_1. Here, I allude to Chebotarev's *field crossing argument* – which is all over [22] starting on page 107 – and again in Section 2.3.1 and Section 3.3.1.

Continuing the stratification procedure

Reducing the general situation to Example 2.9 consists of many, conceptually easy, steps. A large collection of Frobenius fields satisfies the general theory of Frobenius fields containing any Hilbertian field K.

To compare with finite fields [22, Section 30.6] cuts things back from all Frobenius fields to much smaller collections of Frobenius fields. It does so by starting with $\mathcal{C}$ some *full formation* of finite groups. Then it limits G_M to be in the pro-$\mathcal{C}$ groups,

Cutting down further, [22, p. 720] considers a fixed superprojective group U. It restricts to those M containing K with G_M in the collection of quotients of U. Finally, it goes to the case totally compatible with finite fields: U is the profinite completion of $\mathbb{Z}$.

Restricting the full theory of Frobenius fields containing K to these special cases is quite conceptual. The final realization is that in this last case the stratification procedure is done entirely over $\mathbb{Q}$. Now reduce the whole stratification apparatus mod p for most primes p. This amounts to dealing with covers, over $\mathbb{Q}$ (or later, $\mathbb{F}_q$) especially those that are not regular.

Theorem 2.11 [15, Lemma 1] handles this. For a Galois cover $\hat{\varphi} : \hat{Y} \to X$ (again take both X and $\hat{Y}$ to be normal) with X absolutely irreducible having group $\hat{G}$ [15, Lemma 1].[7] §3.3.1 shows why this nonregularity, not easily eliminated, occurs. Take $\hat{\varphi}$ to be a K irreducible cover, K a number field, with ring of integers $\mathcal{O}_K$.

Assume $\hat{K}$ is the minimal field over which $\hat{\varphi}$ breaks into absolutely irreducible (Galois) covers of X. Take any component $\varphi' : Y' \to X$. It will be Galois with group identified with $G' \stackrel{\text{def}}{=} \{g \in \hat{G} \mid g \text{ maps } Y' \to Y'\}$. The Galois extension $\hat{K}/K$ has group $\hat{G}/G'$. For $\boldsymbol{p}$ a prime of $\mathcal{O}_K$ unramified in $\hat{K}$, denote an element of the conjugacy class of the Frobenius of $\boldsymbol{p}$ by $\mathrm{Fr}_{\boldsymbol{p}}$, and its fixed field (resp., residue class field) in $\hat{K}$ by $\hat{K}_{\boldsymbol{p}}$ (resp., $\mathbb{F}_{\boldsymbol{p}}$). Then let $\hat{\varphi}_{\boldsymbol{p}} : \hat{Y}_{\boldsymbol{p}} \to X$ be the union of the conjugates of $\varphi' : Y' \to X$ by $\mathrm{Fr}_{\boldsymbol{p}}$.

Theorem 2.11. *Excluding a finite set B of computable primes $\boldsymbol{p}$ of $\mathcal{O}_K$, reduction of $\hat{\varphi}$ mod $\boldsymbol{p}$ has an $\mathbb{F}_{\boldsymbol{p}}$ Galois cover component isomorphic to $\hat{\varphi}_{\boldsymbol{p}}$ by the reduction map. Take the conjugacy classes, $\boldsymbol{C}_{\boldsymbol{p}}$ associated to $\hat{\varphi}_{\boldsymbol{p}}$ to be those from $\boldsymbol{C}$ whose restriction to $\hat{K}_{\boldsymbol{p}}$ act trivially.*

Remark 2.12. The classes $\mathbf{C}_{\boldsymbol{p}}$ that appear in Theorem 2.11 are precisely those that can be realized as Frobenius elements, as in Comment on (2.12b) in Section 2.3.2.

Remark 2.13 (Proofsheet changes, Edition 2 vs. Edition 3 of [22]). We list the two that concern the topic of this paper. In both editions, a key stratification lemma repeats on consecutive pages as Lemma 30.2.6 and Lemma 30.3.1 (including the material in the 2nd edition that starts at the bottom of p. 711). Also, the references on p. 760 in Edition

7 This lemma does do the curve case, but the details necessary to generalize are the elimination results typical in commutative algebra.

2 between [W. D. Geyer and M. Jarden] and [W. Kimmerle, R. Lyons, R. Sanding and D. N. Teaque] are missing in Edition 2, but appear in Edition 3 on pages 766–771.

2.3 Choices and the nonregular Chebotarev

From the view of deciding diophantine statements, [19] is driven by actual diophantine applications, while [22] is stronger on logic's theories of fields. This paper returns to the first viewpoint: Galois stratification as enhancing understanding specific problems.

The variance of one diophantine problem with the prime p starts with using the Weil estimate systematically in Section 2.3.1, including finishing off Felgner in Section 2.3.3. Variance gets enhancement from attaching to the problems a Poincaré series, the topic of Section 3.

2.3.1 Choices in stratifying

Being courser – not stratifying excessively – on the formation of Galois stratifications allows staying closer to classical problems and explicit computation. At that, the original paper [19] and [22] are at opposite ends of the spectrum, despite the latter's details.

The former suggests restricting to flat covers when appropriate, involving blocks of variables of maximal length when possible. The latter takes blocks with just one variable and insists on unramified covers given by $\mathrm{Spec}(S) \to \mathrm{Spec}(R)$ with S generated, as a polynomial ring, over R, by a single element having no discriminant.[8] Section 2.2 outlined using the finer stratification procedure, as does this section which relies only on the original Weil estimate.

On the other hand, Section 3.3.1 gives our main example, a general situation that concludes with a Galois stratification containing just one cover, and by which we painlessly see infinitely many of the Poincaré series coefficients.

Now we show how the Chebotarev analog works, with a corollary to Main Theorem 2.4 that counts $\boldsymbol{x}' \in \mathbb{F}_q^m$ for which $S_k(\boldsymbol{x}', q)$ holds (for $q \notin M_0$). First, consider going from $\mathcal{S}_{i+1}$ to $\mathcal{S}_i$, dealing with the restriction of the stratification for $\mathcal{S}_{i+1}$ along the fibers of the

$$\text{natural projection } \mathbb{A}^{m+N_{i+1}} \xrightarrow{\mathrm{pr}_{i+1,i}} \mathbb{A}^{m+N_i}. \tag{2.10}$$

With ℓ_0 the number of elements in $\mathcal{S}_0$, and each $1 \le j \le \ell_0$, this gives the following [21, Theorem p. 104]:[9]

8 This guarantees that a Frobenius element is a well-defined conjugacy class, without demanding extra conditions on **C**. In, however, classical problem settings, using flexibility on **C** may be a good idea.

9 Qualitatively [15, Proposition 2] sufficed for [19], but there are more details in the unspecified constants, especially the dependency on the characteristics of the covers, in this reference.

(2.11a) an integer r_j between 0 and N_k and $\mu_j \in \mathbb{Q}^+$, a function of elements in $\mathbf{C}_{j,0}$;
(2.11b) giving the count, B_j, of $\boldsymbol{x}' \in \mathbb{A}^m$ with $\boldsymbol{x}' \in X_{j,0}$ and $\mathrm{Fr}_{\boldsymbol{x}'} \in \mathbf{C}_{j,0}$.

$$\text{Either: } (\dagger)\ B_j \text{ is } 0; \textit{ or } (\dagger\dagger)\ |B_j - \mu_j q^{r_j}|/q^{r_j-1/2} \text{ is bounded in } q.$$

These estimates [21, Section 3, especially Lemma 3.1] are based on [32]. They are explicit, much better than *primitive recursive*, as in the application concluding [19]. In [21, Theorem 6.1, 6.3 et al.], the main ingredients is this. Apply a Chebotarev density theorem version (as in [15, Proposition 2]) to a pair consisting of the Galois cover $\hat{\varphi} : \hat{C} \to X$ to estimate the number of points $\boldsymbol{x}' \in X(\mathbb{F}_q)$ for which $\mathrm{Fr}_{\boldsymbol{x}'}$ is in the conjugacy classes $\mathbf{C}$ attached to $\hat{\varphi}$. Comments on (2.12a) show why, in the inductive procedure, we are producing new covers with new Galois closures and corresponding conjugacy classes.

Tessellating X with hyperplane sections – akin to [32] – and using Chebotarev's own field crossing argument, reverts this to Weil's theorem on projective nonsingular curves over finite fields. [21, Section 4] explicitly traces these classical results. Yet, it still has an error estimate. So, it does not imply the rational function Poincaré series results in Section 3.1 or Section 3.3.

2.3.2 Other Chebotarev points

(2.12a) The actual quantifier elimination moving from $\mathcal{S}_{i+1}$ and $\mathcal{S}_i$ also uses (2.11). Indeed, that shows why there is no elimination of quantifiers through *elementary* statements.
(2.12b) Possibilities $(\dagger)$ and $(\dagger\dagger)$ in (2.11b) give very different error estimate contributions.
(2.12c) The distinctions in (2.12b) arise in a host of problems as illustrated in Section 3.3.1.

Comment on (2.12a)
Consider restriction of one of the terms of the stratification of $\mathcal{S}_{i+1}$ to the fibers $\mathrm{pr}_{i+1,i}$ of the projection (2.10). Much of the [19] proof assures the elimination theory allows applying (2.11), so as to pick out the conjugacy classes that will appear in each of the terms of the $\mathcal{S}_i$ stratification. Section 2.2 has that, though using Frobenius fields there avoided reference, at that point, to the Weil result.

So, for the quantifier $\exists$, the Chebotarev analog is essentially to assure that any conjugacy class that should occur in $\mathcal{S}_i$ actually does. Likewise, for $\forall$, that no conjugacy class that could change the result of the problem is excluded. Nevertheless, Chebotarev is giving error estimates, and not *precise* values that contribute to refined invariants as in Section 3.

Comment on (2.12b)

The distinction between (†) and (††) is consequent on the adjective *nonregular* analog of the Chebotarev (any dimensional base) version. Before [15, Section 2], it had been traditional to make an unwarranted assumption. That is, when considering a cover $\varphi : C \to X$ – say of normal varieties – defined and absolutely irreducible, say, over a field K, that some kind of manipulation would allow assuming that the functorially defined Galois closure $\hat{\varphi} : \hat{C} \to X$ of the cover could also be taken over K.

That will not work in considering the possibilities, based on one Galois stratification, in varying q, even in residue classes of a number field. Here is why (eschewing cautious details). Suppose a component of the Galois closure cover, $\hat{\varphi}$, has definition field $\hat{K} \neq K$ (with K a perfect field). Assume K is a number field with ring of integers $\mathcal{O}_K$.

Then, as we vary the residue class field $R_{\boldsymbol{p}} = \mathcal{O}_K/\boldsymbol{p}$, the corresponding Frobenius $\mathrm{Fr}_{\boldsymbol{x}'}$ for $\boldsymbol{x}' \in \mathbb{F}_q^m$ must restrict to the Frobenius $\mathrm{Fr}_{\boldsymbol{p}}$ on the residue class field $R_{\boldsymbol{p}}$. Suppose, for example, the order of $\mathrm{Fr}_{\boldsymbol{p}}$ does not divide the order of an element $g \in \mathbf{C}_{j,0}$.

Then that conjugacy class of g cannot possibly be $\mathrm{Fr}_{\boldsymbol{x}'}$.[10] That, however, is the only obstruction by the general Chebotarev result – the meaning in (2.4b) of hitting the correct classes – for realizing an element of $\mathbf{C}_{j,0}$ as a Frobenius, so long as q is sufficiently large.

Comment on (2.12c)

The comment on (2.12b) alludes, for $R = \mathcal{O}_K$, to the (†) and (††) conditions varying with the residue class field $R_{\boldsymbol{p}} = \mathcal{O}_K/\boldsymbol{p}$ in actual problems. Further, the value of the Frobenius in an extension of $\hat{K}/K$ attached to the characteristic of $R_{\boldsymbol{p}}$ measured this. The analog is true for problems over a given finite field: The Frobenius in an extension $\hat{\mathbb{F}}_q$ attached to the problem measures the variance with changing the extension of $\mathbb{F}_q$. Both justify the significance, and resistance to elimination, of the word *nonregular* in the Chebotarev analog. Section 3.3.1 is an example that makes this explicit.

2.3.3 Main theorem 2.4 implies "No" to Felgner

Back to Felgner's question with $m = 1$ and running over elements of $\mathbb{F}_{q^2}$ (not $\mathbb{F}_q$). Then the elimination of quantifiers has reduced Felgner's question to this.

(2.13a) Show the following is impossible for q large:

$$\text{with } M_w = \left|x' \in X_{w,0}(\mathbb{F}_{q^2}) \text{ and } \mathrm{Fr}_{x'} \in \mathbf{C}_{w,0}\right|, \quad M_{\mathcal{S}_0} \stackrel{\text{def}}{=} \sum_{w=1}^{\ell_0} M_w = |\mathbb{F}_q| = q.$$

10 That is, there is no error estimate for nonachievement of that class as a Frobenius, say by a bounding constant; it is just not achieved.

(2.13b) Main Theorem 2.4 and (2.11) implies for each w, M_w is either bounded (in q) or asymptotic to $t_w q^2$ for some nonzero t_w.

(2.13c) Neither is a bounded distance from q. Therefore, no elementary formula distinguishes $\mathbb{F}_q$ within $\mathbb{F}_{q^2}$, so long as q is large [21, Section 0].

Indeed, the same argument shows there is no need to take $m = 1$. No matter what formula, no matter the number of variables included in $\boldsymbol{x}$, you can't get q as the number of $\boldsymbol{x}'$ counted by (2.13a), so long as q is large.[11] So we used quantitative counting to *exclude* the elements of $\mathbb{F}_q$ as a result. This was rather than finding a qualitative device that eliminated existence of a formula that precisely nailed $\mathbb{F}_q$ among the elements of $\mathbb{F}_{q^2}$.[12]

Is that satisfactory? It gets to the heart of the Poincaré series/zeta function approach, which primarily aims to count points satisfying equations.

3 Diophantine invariants

[17, Section 7.3] discusses the history of attaching a Poincaré series (and zeta function) to diophantine problems attached to Galois stratifications. We refer to some of its highlights. This section pushes Chebotarev into the background; replacing it with expressions in precise point counts. Section 3.1 gives the Poincaré series definition.

Section 3.2 goes through the series' properties, especially the rationality theorem 3.1. These estimates use *Dwork cohomology* and specific results of Bombieri applied to it.

Section 3.3.1 ties together all threads of this paper with one specific problem – having a vast practical literature – that uses the Poincaré series. Then Section 3.3.2 briefly discusses the artistic extension – of Denef and Lóeser – from Galois stratification to *Chow Motive* coefficients. Thus, Section 3.3 is in the service of enhancing uniformity in p.

3.1 Poincaré series vs. coefficient estimates

We have already seen that the Galois stratification procedure is not canonical. So, it makes sense to address whether there is a *homotopy theoretic* approach based on the

11 [21] exists because the authors of [3] insisted in their first version that Galois stratification could not handle Felgner's question. This despite others at their conference, including one of the authors of [10] – I was not – telling them that was wrong, much akin to the end of [19].

12 That the count is bounded by a constant can't be excluded; see the *exceptional covers* of Section 3.3.

category of Galois stratifications that clarifies a natural equivalence on stratification. That is what these Poincaré series test.[13]

(3.1) is the definition of the Poincaré series attached to the Galois stratification $\mathcal{S}_k$ over $\mathbb{F}_q$ with triples $(X_{\bullet,k}, G_{\bullet,k}, \mathbf{C}_{\bullet,k})$, and its quantifiers $Q_1, \dots, Q_k$. The main theorem 2.4 replaces this by quantifying those points of the last stratification term whose Frobenius values are inside the requisite conjugacy classes. Though more complicated than counting points on a variety over a finite field, it is sufficiently akin to naturally extend classical methods.

The Poincaré series, in a variable t, for a given q attached to Galois Stratification $\mathcal{S}_k$ has this form with the coefficients $\mu(\mathcal{S}_k, q, m)$ explained below:

$$P(\mathcal{S}_k)_q(t) = \sum_{m=1}^{\infty} \mu(\mathcal{S}_k, q, m) t^m.^{14} \tag{3.1}$$

Those $\mu(\mathcal{S}_k, q, m)$s do depend on the quantifiers $\boldsymbol{Q}$. If the Galois stratification was an ordinary elementary statement, then to simplify we would abuse the notation by placing them outside reference to the variables. We mean each quantifier Q_i, respectively, applies to the variables $\boldsymbol{y}_i$, $i = 1, \dots k$. For example, $\boldsymbol{Q}\,\boldsymbol{x}\,\boldsymbol{y}_1\,\boldsymbol{y}_2$ is $\boldsymbol{x}\,Q_1\boldsymbol{y}_1 Q_2\boldsymbol{y}_2$.

So, similar to this, when quantifying the placement of the Frobenius elements, denote the quantified version by $\boldsymbol{Q}\mathrm{Fr}_{\boldsymbol{x}',\boldsymbol{y}'_1,\dots,\boldsymbol{y}'_k} \in \mathbf{C}_{j,k}$,

$$\text{running over } (\boldsymbol{x}', \boldsymbol{y}'_1, \dots, \boldsymbol{y}'_k) \in X_{j,k}(\mathbb{F}_{q^m})(\boldsymbol{x}', \boldsymbol{y}'_1, \dots, \boldsymbol{y}'_k), \quad 1 \le j \le \ell_k. \tag{3.2}$$

Then the coefficient $\mu(\mathcal{S}_k, q, m)$ in (3.1) is the point count of those $\boldsymbol{x}' \in X_{j,k}(\mathbb{F}_{q^m})$ for which $\boldsymbol{Q}\mathrm{Fr}_{\boldsymbol{x}',\boldsymbol{y}'_1,\dots,\boldsymbol{y}'_k} \in \mathbf{C}_{j,k}$ is constrained to over the quantified y-variables, $\boldsymbol{y}'_1, \dots, \boldsymbol{y}'_k$, with values in $\mathbb{F}_q$. Main Theorem 2.4 allows replacing $\mu(\mathcal{S}_k, q, m)$ by $\mu(\mathcal{S}_0, q, m)$: counting (rather than estimating as in (2.11b)) those

$$\boldsymbol{x}' \in X_{j,0}(\mathbb{F}_{q^m}) \text{ for which } \mathrm{Fr}_{\boldsymbol{x}'} \in \mathbf{C}_{j,0}, \quad 1 \le j \le \ell_0. \tag{3.3}$$

3.2 Poincaré properties

[22, Chapter 30 and 31] (Chapters 25 and 26 in the 1986 edition, pretty much the same) have complete details on the Poincaré and Zeta properties. A *Zeta function*, $Z(t)$, has an attached *Poincaré series* $\mathrm{P}(t)$. This is given by the logarithmic derivative:

$$t\frac{d}{dt}\log(Z(t)) = \mathrm{P}(t).$$

13 Take *homotopy theoretic* to mean that an outcome expressable in terms of some kind of cohomology. One that equivalences among structures related to Galois stratification or variants that result in the same cohomology results. In a sense that is the point of motivic cohomology.

14 The notation of Section 4.2 shows why we did not use the simpler notation $\mu(\mathcal{S}_k, q^m)$.

Add that $Z(0) = 1$, and each determines the other. The catch: $Z(t)$ rational (as a function of t) implies $\mathrm{P}(t)$ rational, but not always the converse.

Suppose D is an elementary diophantine statement, with quantifiers as we started this paper. Then, as earlier, take D in place of $\mathcal{S}_k$ in $P(\mathcal{S}_k)_q(t) = \sum_{m=1}^{\infty} \mu(\mathcal{S}_k, q, m)t^m$, and consider the coefficients referencing $(D, \boldsymbol{Q}, m)$ in place of $\mu(\mathcal{S}_k, q, m)$.

I do not know when Ax introduced considering such coefficients. He suggested to me meaningfully computing them at IAS in Spring 1968. Originally, I introduced the Galois stratification procedure to do just that, and to conclude the following result. Again, suppress notation $\boldsymbol{Q}$ for the quantifiers and give the result explicitly just when $R = \mathbb{Z}$ in (1.5).

The adjustments are clear for when R is a given finite field, or ring of integers of a number field. Use the notation around the main theorem 2.4.

Theorem 3.1 (Poincaré rationality). *For each prime $p \notin M_0$, $P(\mathcal{S}_k)_p(t)$ is a rational function $\frac{n_p(t)}{d_p(t)}$, with $n_p, d_p \in \mathbb{Q}[t]$ and computable. The corresponding $Z(\mathcal{S}_k)_q(t)$ has the form $\exp(m_p^*(t))(\frac{n_p^*(t)}{d_p^*(t)})^{\frac{1}{u_p}}$ with $m_p^*, n_p^*, d_p^* \in \mathbb{Q}[t]$ and $u_p \in \mathbb{Z}^+$ computable. Further, there are bounds, independent of p, for the degrees of all those functions of t. For any particular prime p all functions are computable (see §3.3.2).*

Comments on the proof of Theorem 3.1

We start with highlights from [22, Section 31.3] (or 1986 edition, Section 26.3; essentially identical) titled: Near rationality of the Zeta function of a Galois formula. A similar result bounds the degrees even if $p \in M_0$, assuming (2.3b).

As, however, Remark 2.7 notes, there will not be any expected relation with results for $p \notin M_0$. We cannot use the uniform estimate on the characteristics of the Galois stratification since they do not come from a reduction of a uniform stratification object as given in Theorem 2.11. Refer to the stratifications for $p \in M_0$ as *incidental*, with their incidental estimates. What we say here applies equally to uniform and incidental stratifications.

The conclusion of the Galois stratification procedure over the $\boldsymbol{x}$-space gives this computation for (3.3). Sum the number of $\boldsymbol{x}$ with values in $\mathbb{F}_{p^k}$ for which the Frobenius falls in conjugacy classes attached to the stratification piece going through $\boldsymbol{x}$.

The expression of that sum in *Dwork cohomology* is what makes the effectiveness statement in Theorem 3.1 possible. That also suggests its direct relation to Denef–Loeser. An ingredient for that is a formula of E. Artin. It computes any function on a group G that is constant on conjugacy classes as a $\mathbb{Q}$ linear combination of characters induced from the identity on cyclic subgroups of G.

Additional historical comments

A function on G that is 1 on a union of conjugacy classes, 0 off those conjugacy classes, is an example. [22, pp. 738–739] (1986 edition pp. 432–433) recognizes L-series

attached to that function as a sum of L-series attached to those special induced characters. I learned this from [2, p. 222] and had already used it in [15, Section 2]. Kiefe – working with Ax – learned it, as she used it in [31], from me as a student during my graduate course in Algebraic Number Theory at Stony Brook in 1971. The core of the course were notes from Brumer's Fall 1965 course at UM.[15]

Kiefe, however, applied it to a list-all-Gödel-numbered-proof procedure; not to Galois stratifications I showed her (see my Math Review of her paper, Nov. 1977, p. 1454). Consider the identity representation induced from a cyclic subgroup $\langle g\rangle, g \in G$. This L-Series is the zeta function for the quotient cover by $\langle g\rangle$ (exp. 7–9, p. 433, 1986 edition of [22]).

We do use a zeta function

Given a rational function in t, its *total degree* is the sum of the numerator and denominator degrees; assuming those two are relatively prime. The 1986 edition, Lemma 26.13 refers to combining [13] and [1] to do the zeta function of the affine hypersurface case for explicit bounds – dependent only on the degree of the hypersurface – on the total degree of the rational functions that give these zeta functions.

Then some devissage gets back to our case, given explicit computations dependent only on the degrees of the functions defining these algebraic sets. Lemma 26.14 assures the stated polynomials in t have coefficients in $\mathbb{Q}$, and it explicitly bounds their degrees. The trick is to take the logarithmic derivative of the rational function. Then the Poincaré series coefficients are power sums of the zeta-numerator zeros minus those of the zeta-denominator zeros. Using allowable normalizations, once you have gone up to the coefficients of the total degree, you have determined the appropriate numerator and denominator of $P(t)$.

One observation: We are left to uniformly bound in p the degrees of the zeta polynomials, etc. This follows from the comments above Theorem 2.11 giving the uniformity in p in the characteristics of the reduction of the stratification. Apply this to the degrees of polynomials describing the affine covers for Dwork–Bombieri.

3.3 Chow motive coefficients

As stated in Section 2.3.1, the error estimate that allowed the elimination of quantifiers is not appropriate for concluding either the estimate of the degrees of the polynomials in the Poincaré series, or that it is a rational function.

It is sensible to use as coefficients of the Poincaré series actual Galois stratifications. For $p \notin M_0$, those coefficients sum over the Galois covers $\hat{\varphi} : \hat{Y} \to X$ in the stratification $\mathcal{S}_0$ those $\boldsymbol{x}' \in X(\mathbb{F}_q)$ for which $\mathrm{Fr}_{\boldsymbol{x}'} \in \mathbf{C}$. It also works to replace the count on the

15 I submitted my paper in 1971. It had five different referees.

stratification pieces with absolutely irreducible algebraic varieties that give the same counts. This uses the field crossing argument mentioned several times previously, as in Remark 2.10.

The topic here is enhancing the Poincaré series coefficients, extending them to *Chow motive coefficients*; the last of the coefficient choices given in the abstract. After the example of Section 3.3.1, Section 3.3.2 briefly discusses the abstract setup of [11] and [34]. These have their own expositions on Galois stratification. Further discussion of these will occur in the extended version of this paper.

3.3.1 Distinguishing special primes

We produce an example where the associated Poincaré series over $\mathbb{F}_q$ has some coefficients that are *polynomials* in q. Yet, other coefficients are functions of the Frobenius Fr_q evaluated on a Chow motive: a linear expression including the (nontrivial) ℓ-adic cohomology of a nonsingular variety. [17, Section 3] takes a general diophantine property and casts it as an umbrella over two seemingly distinct diophantine properties about general covers.

These properties fit a *birational* rubric, or what we call *monodromy precision*. That is, the Galois closure of the covers alone, together with the corresponding permutation representations attached to their geometric and arithmetic monodromy, guaranties, precisely, their defining diophantine property. The **NRC** – as always – gives the count of achieved conjugacy classes *roughly*. Here, though, there will be no error term – unlike the Comment on (2.12b) in Section 2.3.2 – even though there is achievement of nontrivial conjugacy classes.

That allows stating, should we start with such a cover over a number field K, what are the primes for which the diophantine property has a precise formulation over a given residue class field $\mathcal{O}_K/\boldsymbol{p} = \mathbb{F}_{\boldsymbol{p}}$, as in [17, Definition 3.5 and Corollary 3.6]. Our example uses the simplest case: the *exceptional cover* property.

Example 3.2 (Example 2.9 continued). For this, in Example 2.9, take the following special case over K. The hypersurface in $\mathbb{A}^{m+1}$ is still defined by an equation $X_2 = \{(\boldsymbol{x}, y) \mid f(\boldsymbol{x}, y) = 0\}$, with f absolutely irreducible (over K) of degree $u > 1$ in y. Now, though, the cover $\hat{\varphi}_2$ is trivial (of degree 1). As before consider $\mathrm{proj}_{\boldsymbol{x}} : \mathbb{A}^{m+1} \to \mathbb{A}^m$ restricted to X_2 and these diophantine statements:

$$\begin{aligned} D_{\boldsymbol{p}}(\boldsymbol{x}): &\qquad \exists\, y \in \mathbb{F}_{\boldsymbol{p}} &&\mid f(\boldsymbol{x}, y) = 0; \\ D_{\boldsymbol{p}}: &\qquad \forall\, \boldsymbol{x} \in \mathbb{F}_{\boldsymbol{p}}^m\ \exists\, y \in \mathbb{F}_{\boldsymbol{p}} &&\mid f(\boldsymbol{x}, y) = 0; \\ D: &\quad D_{\boldsymbol{p}} \text{ is true for } \infty\text{-ly many } \boldsymbol{p}. \end{aligned} \tag{3.4}$$

Continue with the Example 2.9 notation, and the formation of the Galois cover $\hat{\varphi}_1 : \hat{C}_1 \to X_1$ with X_1 Zariski open in $\mathbb{A}^m$, and with group $G_{\hat{\varphi}_1}$ having its natural and

faithful, transitive, degree u permutation representation T. Consider the projective normalization, $\hat{C}_1^\dagger$, (resp., $X_2^\dagger$) of $\hat{\varphi}_1$ (resp., X_2) in the function field of $\hat{C}_1$ (resp., of X_2)

$$\hat{\varphi}_1^\dagger : \hat{C}_1^\dagger \to X_2^\dagger \to \mathbb{P}^m, \quad \mathbb{P}^m \supset \mathbb{A}^m.^{16}$$

Now extend the diophantine expressions of (3.4) to include $X_2^\dagger$. For example, $D_{\boldsymbol{p}}(\boldsymbol{x})$, for $\boldsymbol{x} \in \mathbb{P}^m(\mathbb{F}_q)$ means $\exists$ a point of $X_1^\dagger(\mathbb{F}_q)$ above $\boldsymbol{x}$. A similar meaning is given to $D_{\boldsymbol{p}}$. □

We make a simplifying assumption in Example 3.2 to match up with Section 3.3.2.

$$\text{Not only is } X_2 \text{ normal, but } X_2^\dagger \text{ is nonsingular.} \tag{3.5}$$

Though the spaces are given by projective, not affine, coordinates, we form a single cover that we may regard as a Galois stratification $\mathcal{S}_0$, with one stratification piece: $\hat{\varphi}_1^\dagger : \hat{C}_1^\dagger \to \mathbb{P}^m$ with attached conjugacy classes

$$\mathbf{C}_1 = \{g \in G_{\hat{\varphi}_1} \mid T(g) \text{ fixes a letter in the representation}\}.$$

Proposition 3.3. *Assume* (3.5). *There is a finite set, M_0, of primes $\boldsymbol{p}$ of $\mathcal{O}_K$ for which, given $\boldsymbol{p} \notin M_0$, then $D_{\boldsymbol{p}}$ is true if and only the following equivalent conditions hold:*
(3.6a) $\mathrm{Fr}_{\boldsymbol{x}'} \in \boldsymbol{C}_1$ *for each* $\boldsymbol{x}' \in \mathbb{P}^m(\mathbb{F}_{\boldsymbol{p}})$.
(3.6b) *There are infinitely many $\boldsymbol{p}$ for which the equivalence of* (3.6a) *holds.*

Assume (3.6). *Then for some finite set $M_0' \supset M_0$, for each $\boldsymbol{p} \notin M_0'$ for which* (3.6a) *holds,*

$$\begin{aligned} &\textit{if } \mathbb{F}_{\boldsymbol{p}} = \mathbb{F}_{q_0}, \text{ (3.6a) } \textit{holds with } \mathbb{F}_{q_0^m} \textit{ replacing } \mathbb{F}_{\boldsymbol{p}}. \textit{ Then, for } \infty\textit{-ly many } m, \\ &\quad \textit{the } m\textit{th coefficient of the Poincaré series } P(\mathcal{S}_0)_{\boldsymbol{p}}(t) \textit{ for } \mathcal{S}_0 \textit{ is } \tfrac{q_0^{m+1}-1}{q_0-1}. \end{aligned} \tag{3.7}$$

We also note:

$$\textit{There are also infinitely many } \boldsymbol{p} \textit{ for which } \text{(3.6a) } \textit{does not hold (see below).} \tag{3.8}$$

Definition 3.4 (Exceptional primes). The primes $\boldsymbol{p}$ for which (3.7) holds are called the *exceptional primes*, $E_{\hat{\varphi}_1^\dagger}$, of $\hat{\varphi}_1^\dagger$ (or of any other object natural attached to it). The main point is there are infinitely many of them, and their Poincaré series gives them away. Such a cover is called an *exceptional cover* over K.

Note, however, this is a case of a regular cover whose Galois closure is *not* regular. What this says in (3.6b) is that only the elements of $\mathbf{C}_1$ are achieved as Frobenius elements for the primes satisfying the conditions (3.6). For example, for those primes $\boldsymbol{p}$ (with $|\boldsymbol{p}|$ suitably large) for which reduction mod $\boldsymbol{p}$ gives a regular function field extension, the Chebotarev theorem says (3.6a) will *not* hold. It is very elementary – an easy case of Chebotarev for number fields – that there are infinitely many such $\boldsymbol{p}$.

16 $\hat{C}_1^\dagger$ is the projective normalization of $\mathbb{P}^m$ in the function field of $\hat{C}_1$; a canonical process.

[6] considered this hyperelliptic pencil with parameter λ: $y^2 - f(x) + \lambda$, $f \in \mathbb{F}_p(x)$. The difference between the expected number, p, of $\{(x,y) \in (\mathbb{Z}/p)^2\}$ and the actual value, $V_{p,\lambda}$, given by Weil's theorem caused them draw the following conclusion:

There is a constant $c_f > 0$ such that:
Running over $\lambda \in \mathbb{F}_p$, $\sum_{\lambda \in \mathbb{F}_p}(p - V_{p,\lambda})^2 > c_f p^2$, if and only if f is not exceptional.

[29] (recounted in [17, Section 7]) used a generalization of the Davenport–Lewis error term to count the summed squares of the multiplicity of (completely reducible) components of the monodromy (fundamental group) action on the 1st complex cohomology of a fiber of any family of nonsingular curves over a base S. His technique – also coming upon the exceptionality condition – used reduction mod p to get to the results of [9].

Remark 3.5. It has long been known that there are many exceptional covers. Between [16, Section 2], [25] and [17, Sections 6.1 and 6.2] those with $m = 1$ and $f(x,y) = f(y) - x$, f a rational function over a number field. The 1st and 3rd of these connect to Serre's open image theorem. That paper also introduced natural zeta function test cases of the Langland's program. One problem is a standout. For the wide class of these related to the GL_2 case (and only these) of Serre's open image theorem, the precise set $E_{\hat{\varphi}_1^\dagger}$ is a mystery appropriate for the non-Abelian class field theory of the Langland's program [17, Section 6.3].

3.3.2 Denef–Löser and Chow motives

The comments on Theorem 3.1 show we can express the coefficients in the Poincaré series from the trace of Frobenius iterates acting on the p-adic cohomology that underlies Dwork's zeta rationality result. Positive: The computation is effective. Negative: The cohomology underlying Dwork's construction varies with p. Nothing in 0 characteristic represents it.

Even, however, Dwork's cohomology [12] deals with stratifying your original variety. By "combining" the different pieces you conclude the rationality of the zeta function from information on the Frobenius action from the hypersurface case. The explicit estimates that Theorem 3.1 relies on for these hypersurface computations are from [1].

Denef and Loeser [10] applied Galois stratification (see the arXiv version of [26, App.]) to eliminate quantifiers in their p-adic problem goals. Their technique, as in [11], applies to consider – for almost all p – how to express those Poincaré series, as stated in the abstract to this paper, as elements in the Grothendieck group generated by completely reducible (by the action of the Frobenius) pieces of the weighted ℓ-adic cohomology – twisted by Tate Modules; a tensoring of the group by some power of the cyclotomic character – of *nonsingular projective* varieties. Thus, the coefficients – as in the previous cases – derive from the trace applied to restricting powers of the Frobenius for p to these coefficients. Section 3.4 has further clarification.

They use Galois stratification, with the field crossing argument (Remark 2.10) doing a lot of work in putting covers and conjugacy classes to the background. Yet, this artistic enhancement to the uniformity in p in the uniform stratification (Comments on the proof of Theorem 3.1) is still akin to the previous methods.

3.4 Deligne and motives

The word *motive* refers to the weighted pieces – rather than (pure) mth cohomology of a projective nonsingular variety – being a summand of this, tensored by a Tate twist. A correspondence – cohomologically idempotent – is attached to indicate the source of the projection that detaches a summand from the pure weighted cohomology. Such idempotents are compatible with their appearance in writing the Poincaré series attached to a conjugacy class count in terms of characters induced from cyclic subgroups (say, [22, Section 31.3]$_1$; as in Theorem 3.1) of the group G.

For one, from the main theorem of [9], the absolute values of the eigenvalues of the Frobenius on these pieces are known. That allows more carefully considering any cancellation of these eigenvalues. For example, [9, Theorem 8.1] gives a definitive result on the eigenvalues of a complete intersection of dimension n. The error term there is $O(q^{n/2})$ by contrast to the usual expectation that an error term is $O(q^{n-1/2})$ as in (2.11).

The Section 3.3.1 example has a naturally attached projective nonsingular variety, $X_2^\dagger$ to express the Poincaré series coefficients for a Galois stratification $\mathcal{S}_0$ for (almost) all p, not just for those in the exceptional set. Even this example shows an aspect of using Denef–Löser that the previous two approaches do not seem to have:

> Using (Chow) motive pieces to relate uniform stratification primes and incidental primes.

I have wanted to say this in print for a long time. It is common to think of ℓ-cohomology as if it could all be from the cohomology of Abelian varieties. For example, [7] expresses the Weil conjectures for the cohomology of K3 surfaces (that in the middle being the only significant piece) *is* through a Clifford algebra from the cohomology of an (nonobvious) Abelian variety with appropriate Tate twists. Indeed, actual descent to give Frobenius action in positive characteristic requires a rigidying argument. This uses endomorphisms from the Clifford algebra. So, the eigenvalue weights come from Weil's theorem on Abelian varieties.

Yet – has anyone considered the question in print earlier – [8] shows the middle cohomology of most complete intersections is not expressible from Abelian varieties.

Question 3.6. How, in general, would you distinguish those Chow motive pieces that do come from Abelian varieties?

The archetype for such a consideration is the classically considered use of minimal models to describe appropriate Hasse–Weil zeta functions of elliptic curves, one

that includes factors for the primes of bad reduction. I say this even though Denef–Löser stratification replacement uses resolution of singularities in 0 characteristic. That is, their method won't directly touch the primes of the incidental stratification.

For good reason, I should consider how my viewpoint can properly and practically tackle the motivic aspects; also what [34] gains using Voevodsky's motives.

4 Motivations for extending Galois stratification

In each of the remaining subsections, I put a classical veneer – appropriate to areas of specific researchers – on papers that extend the Galois stratification procedure.

4.1 The author's motivations

Very special problems motivated surprisingly general approaches to diophantine equations. Well-known examples drove the author's motivations and his many conversations with the principles involved. We have already mentioned Artin's conjecture. One short breathless paragraph summarizes another.

The intense work on zeta functions at *all* primes for an elliptic curve, E, over $\mathbb{Q}$ aimed at getting a functional equation for the associated Hasse–Weil zeta function $Z(E)$ that – combined with the Eichler–Shimura congruence formula for the Frobenius on modular curves and the Shimura–Taniyama–Weil conjecture – motivated the use of $Z(E)$ to conjecture – a la Birch–Swinnerton–Dyer – what is the rank of $E(\mathbb{Q})$ points.

Indeed, less breathlessly, what natural diophantine statements could be expected to have such Euler factors at *every* prime? Before Ax and Kochen, most number theorists seemed convinced of the naturalness of the Artin conjecture: such a close analog of Chevalley's theorem! It, however, no longer appears to be so natural. The Langlands Program is quite dependent on specifics – of which there are few examples – in the paragraph above. Yet the belief in Hasse–Weil zeta functions rest on the naturalness of the Langlands Program.

Section 3.2 concludes – via a cohomology component – with a strong uniformity statement, giving an object over $\mathbb{Q}$ from which Euler factors for almost all primes p come from reductions moduli those primes. Then Section 3.3, based on Denef–Löser, strengthens the cohomology analogy with the work on elliptic curves.

Finding semiclassical examples of such diophantine problems was this author's motivation. Exceptional covers, as in Example 3.2, is one of a general type, which like the Artin conjecture, starts as a simple quantified statement. Generalizing the property of exceptional covers and *Davenport pairs* (as in [17] and [18]) tied many seemingly disparate diophantine considerations together.

Toward the virtues of translating diophantine statements to Chow motives one property seems aesthetically significant: Separating zeta coefficients that are rational functions in Tate motives – and, therefore, on evaluating Fr_q at them, give rational functions in q – from those that are coefficients involving more complicated cohomology.

In Proposition 3.3, (3.7) expresses that for infinitely many Poincaré factors, the coefficients that appear are rational functions in Tate motives. Expression (3.8) suggests that usually there will also be infinitely many Chow motive coefficients that are linear combinations of cohomology not composed from Tate motives.

Remark (3.5), about elliptic curves over $\mathbb{Q}$ with the GL_2 property in Serre's Open Image, poses one of the most simply stated unsolved problems for the Langland's program. This concludes my summary of my many years ago motivations.

4.2 Generalizing Felgner to Frobenius vector problems

Denote an algebraic closure of $\mathbb{F}_q$ by $\bar{\mathbb{F}}_q$. We regard such a generalization as considering whether there could be analogous results in quantifying our variables according to Definition 4.1.

Definition 4.1. For a given prime-power q, and $d_1, \ldots, d_m$ an m-tuple of integers, refer to $\mathrm{Fr}_q^{\boldsymbol{d}} \stackrel{\mathrm{def}}{=} (\mathrm{Fr}_q^{d_1}, \ldots, \mathrm{Fr}_q^{d_m})$ as a *Frobenius vector*. Then $\mathrm{Fr}_q^{\boldsymbol{d}}$ acts on $\bar{\mathbb{F}}_q^m$ coordinatewise, allowing us to speak of the elements $\boldsymbol{x} \in \bar{\mathbb{F}}_q^m$ fixed by a Frobenius vector. Denote these $\bar{\mathbb{F}}_q^m(\mathrm{Fr}_q^{\boldsymbol{d}})$.

Notice we may write the elements – in $\mathbb{F}_{p^2}$, referenced by Felgner's problem, as $\bar{\mathbb{F}}_q^m(\mathrm{Fr}_q^{\boldsymbol{d}})$ with $\boldsymbol{d} = (2, \ldots, 2)$. Frobenius vectors allow generalizing elementary statements (and Galois stratifications) to where variables have values in differing extensions of $\mathbb{F}_q$. With the notation above, consider two Frobenius vectors: $\mathrm{Fr}_q^{\boldsymbol{d}}$ of length m; and $\mathrm{Fr}_q^{\boldsymbol{e}}$ given by $\boldsymbol{e} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k)$ with $\boldsymbol{e}_i$ of length n_i. Thus, $\boldsymbol{e}$ has length $N_k = \sum_{i=1}^k n_i$.

Here are some basic questions:

(4.1a) Are the corresponding Poincaré series of Galois stratifications, with no quantified variables, rational?

(4.1b) Are there Bombieri–Dwork bounding degrees of the involved rational functions?

(4.1c) Does the Galois stratification procedure generalize to give rational functions? That is, given quantifiers, can we eliminate them to be at (4.1a) with bounds from (4.1b)?

(4.1d) If the above, when are these Poincaré series new; not expressed from series associated to our previous Galois stratifications?

Wan's zeta functions

There are no Galois stratifications or quantified variables in [38]. Just the definition of a zeta function defined by a Frobenius vector. That is coefficients defined from counting points $\boldsymbol{x} \in \bar{\mathbb{F}}_q^m(\mathrm{Fr}_q^{\boldsymbol{d}})$ as above, a la Dwork, on an affine variety in $\mathbb{A}^m$.

Here are its contributions to (4.1a). Following a preliminary result by Faltings, documented in [37], [38, Theorem 1.4] shows the zeta is a rational function.

Then it considers (4.1b) on the total degree of the zeta function akin to what Theorem 3.1 uses. There is a preliminary result for special Frobenius vectors with $d_1|d_2|\cdots|d_m$ (consecutive ds dividing the next in line) in [24] based on Katz's explicit bound for ℓ-adic Betti numbers in [30]. Then [38, Conjecture 1.5] has a conjecture that there is an explicit total degree bound in general.

4.3 Logicians Chatzidakis, Hrushovski, and Tomasovic

[5], [28], and [36] deal with the theory of difference fields (or schemes), and the generalization of Galois formulas (akin to [22]) over difference rings. Difference fields include the field $\bar{\mathbb{F}}_q$ together with a symbol for the Frobenius automorphism. The generalization includes the addition of a symbol σ for a distinguished automorphism of a scheme.

Galois stratification abstracted

Recall, the main device of [19] was – necessarily – going beyond the original set of questions, to the richer set of *Galois stratifications*. A short statement of what then happened: a quantifier elimination (as in Section 2.3) gave a primitive recursive procedure for the theory of *first-order definable sets in the language of schemes over finite fields* equipped with powers of the Frobenius automorphism (and related theories).

That statement is similar to the [36, p. 2, Theorem 1.1] main result. Except, the word *twisted* appears in front of *Galois stratification*, and *difference schemes* replace *schemes*.

Further, [36, Theorems 1.1 and 1.2] are – with these enhancements – the exact analog of results in [19], and proceed on the same basis. That is, you start with a Galois stratification over a *difference* scheme (X, σ). Here, the data is a stratification of (X, σ) into difference subscheme pieces $(X_i, \sigma), i \in I$, where each X_i is equipped with an étale Galois cover $Z_i \to X_i$ with group G_i and conjugacy classes $\mathbf{C}_i$ in G_i.

Defining, though, appropriate Galois covers has subtleties about extending σ and assuring with the additional data that "difference field specializations" of points of X_i lift to corresponding extending specializations in the Galois cover.

The idea is just like a Galois formula attached to affine space based on achieving Frobeniuses in the conjugacy classes attached to points of X running over finite fields,

subject to quantifiers placed on some of the variables, or for counting such achievements.

Additional complications

More general here is that σ is added structure on X from an endomorphism of X. Also, everything is in the language of difference schemes. Many notions are not defined in this paper, but in a longer paper called "Twisted Lang–Weil" [35].

The essence of [36] is to define the direct image, $\mathcal{B}$, of the twisted Galois stratification $\mathcal{A}$, given an étale morphism $(X, \sigma) \to (Y, \sigma')$, so that their underlying Galois formulas are appropriately related: an exact analog of [19] and Section 2.2.

Elimination of quantifiers follows from applying this to projection of affine n-space onto affine m-space $m < n$. Until inductively, you are at a quantifier free statement on a (almost certainly) complicated Galois stratification. Complicated or not, a Chebotarev analog applies to decide if almost all Frobeniuses end up in the associated conjugacy classes.

The difficulty is in extending the definitions to include difference schemes in both pieces. Defining a Galois cover of difference schemes [35], so that there is a Chebotarev analog, and then proving the endomorphisms survive, in appropriate shape, the process of direct image.

Both apply to (4.1c) on eliminations of quantifiers, though neither goes after the Poincaré series. There is even the notion of an existentially closed difference field, [5], the analog of the (existentially closed – **PAC**) Frobenius fields of Section 2.2 as in (2.9a).

Certainly, the very long [28] preprint and [36] cover territory akin and in ways more general than this paper, though without the attempt to consider the relation with zeta functions. To give a hint as to what it does include, we conclude with a few comments on it starting with the *Twisted Lang–Weil Theorem*. Look back at Section 2.3 for the extensive use made of variants on the original Lang–Weil theorem.

Here is an example of what Hrushovsky and Tomašiç are considering.

Example 4.2. Suppose $\varphi : A \to A$ is a nontrivial endomorphism of an Abelian variety over $\mathbb{F}_q$. Denote its graph by S.

With Γ_{Fr_q} the graph of the Frobenius map $\boldsymbol{a} \mapsto \boldsymbol{a}^q$, consider $|S(\bar{\mathbb{F}}_q) \cap \Gamma_{\mathrm{Fr}_q}(\bar{\mathbb{F}}_q)|$. The cardinality of this set is the same as that of the elements in $\bar{\mathbb{F}}_q$ for which $\varphi(\boldsymbol{a}) = \boldsymbol{a}^q$.

In generality, replace A by any variety X and replace S by any correspondence with finite projections to each factor of $X \times X$ (there are details to state this completely correctly). The nonobvious estimate for this cardinality is essentially the same as for Lang–Weil, and both authors recount the history behind the result initiated (apparently) by a conjecture of Deligne in [9], with an ultimate proof dependent on his proof of the Weil conjectures. The case with S the diagonal exactly gives Lang–Weil.

As [28] exposits, this is a key ingredient in the theory of $\bar{\mathbb{F}}_q$ with its Frobenius. Again, they are each thinking of the generalization to difference schemes for which

the key word is endomorphism of X. So, [36] is considering "Galois covers" $\hat{\varphi} : \hat{C} \to X$ with X equipped with an endomorphism σ and $\hat{C}$ equipped with endomorphisms Σ, above it, that can be regarded as closed under something akin to conjugation. For this situation they must prove the analog of the Chebotarev theorem counting the number of times the Frobenius is achieved within Σ for points $X(\bar{\mathbb{F}})$.

For logicians the key question: Does this gives a theory of various kinds of difference fields. That would include $\mathbb{F}_q$ with the action of the Frobenius Fr_q? The answer is Yes! That is what their papers show. In the Frobenius case, how does that work with our theme of uniformity in p, an algebraic-geometric property sort-of compatible with their logic frameworks?

Question 4.3. What has this to do with Denef–Löser and with Wan?

Is this difference equation approach like the production of Poincaré series with those Chow motive coefficients? Especially since Felgner's simple question offered so many peeks at the value of a decision procedure.

For example, [28] considers specific diophantine applications to a very interesting field: The cyclotomic closure, $\mathbb{Q}_{\mathrm{cyc}} \stackrel{\mathrm{def}}{=} \bigcup_{n=1}^{\infty} \mathbb{Q}(e^{\frac{2\pi i}{n}})$. This field is especially interesting because it is a characteristic 0 analog of the algebraic closure of a finite field. Consider its mysteries in light of the Fried–Völklein conjecture [23], a strengthening of its main theorem.

Conjecture 4.4. *If $K \leq \bar{\mathbb{Q}}$ is Hilbertian (Hilbert's irreducibility theorem holds in K) and its absolute Galois group, G_K, is projective (among profinite groups), then G_K is pro-free.*

The special case $\mathbb{Q}_{\mathrm{cyc}}$ is a conjecture of Shafarevich. If true, in lieu of the theory of Frobenius fields, this connects $\mathbb{Q}_{\mathrm{cyc}}$ in an even deeper way with the theory of finite fields. That, by the way, is a big piece of the motivation for [22].

Bibliography

[1] E. Bombieri, On exponential sums in finite fields, II, *Invent. Math.*, **47** (1978), 29–39.
[2] J. W. S. Cassals and A. Frölich, Algebraic Number Theory, Acad. Press, London, 1967.
[3] Z. Chatzidakis, L. van den Dries, and A. Macintyre, Definable sets over finite fields, *J. Reine Angew. Math.*, **427** (1992), 107–135.
[4] C. Chevalley, Demonstration d'une hypothèse de E. Artin, *Abh. Math. Semin. Hamb.*, **11** (1936), 73–75.
[5] Z. Chatzidakis and E. Hrushovski, Model theory of difference fields, *Trans. Am. Math. Soc.*, **351**(8) (1999), 2997–3071.
[6] H. Davenport and D. J. Lewis, Notes on congruences (I), *Q. J. Math. Oxf. (2)*, **14** (1963), 51–60.
[7] P. Deligne, La conjecture de Weil pour les surfaces K3, *Invent. Math.*, **15** (1972) 206–226.
[8] P. Deligne, Les intersections complètes de niveau de Hodge, *Invent. Math.*, **15** (1972) 237–250.
[9] P. Deligne, La Conjecture de Weil I, *Publ. Math. IHES*, **43** (1974), 273–307.

[10] J. Denef and F. Loeser, Motivic Igusa zeta functions, *J. Algebraic Geom.*, **7**(3) (1998), 505–537.
[11] J. Denef and F. Loeser, Motivic integration and the Grothendieck group of pseudo-finite fields, In: Proc. Int. Congress of Math. (ICM 2002), Vol. 2, 2002, pp. 13–-23.
[12] B. Dwork, On the rationality of the zeta function of an algebraic variety, *Am. J. Math.*, **82** (1960), 631–648.
[13] B. Dwork, On the zeta function of a hypersurface III, *Ann. Math.*, **83** (1966), 457–519.
[14] U. Felgner, Posed at the Model Theory Conference in Oberwolfach, January 1990.
[15] M. D. Fried, On Hilbert's irreducibility theorem, *J. Number Theory*, **6** (1974), 211–231.
[16] M. D. Fried, Galois groups and complex multiplication, *Trans. Am. Math. Soc.*, **235** (1978), 141–162.
[17] M. D. Fried, The place of exceptional covers among all diophantine relations, *J. Finite Fields*, **11** (2005) 367–433, arXiv:0910.3331v1 [math.NT].
[18] M. D. Fried, Variables separated equations: strikingly different roles for the branch cycle lemma and the finite simple group classification, *Sci. China Math.*, **55** (2012), 1–72, arXiv:1012.5297v5 [math.NT], 10.1007/s11425-011-4324-4.
[19] M. D. Fried and G. Sacerdote, Solving diophantine problems over all residue class fields of a number field and all finite fields, *Ann. Math.*, **104** (1976), 203–233.
[20] M. D. Fried, D. Haran, and M. Jarden, Galois stratification over Frobenius fields, *Adv. Math.*, **51** (1984), 1–35.
[21] M. D. Fried, Effective counting of the points of definable sets over finite fields, *Isr. J. Math.*, **85** (1994), 103–133.
[22] M. D. Fried and M. Jarden, In: Field Arithmetic, Ergebnisse der Mathematik III, Vol. 11, 1st Edition, Springer Verlag, Heidelberg, 1986 (455 pgs); 2nd Edition 2004 (780 pgs) ISBN 3-540-22811-x. Our cites here are from the 2nd Edition, 1986 is similar.
[23] M. D. Fried and H. Völklein, The embedding problem over an Hilbertian-PAC field, *Ann. Math.*, **135** (1992), 469–481.
[24] L. Fu and D. Wan, Total degree bounds for Artin L-functions and partial zeta functions, *Math. Res. Lett.*, **10** (2003) 33–41.
[25] R. Guralnick, P. Müller, and J. Saxl, The rational function analoque of a question of Schur and exceptionality of permutations representations, *Mem. Am. Math. Soc.*, **162** 773 (2003), ISBN 0065-9266.
[26] T. Hales, What is motivic measure?, the version WhatIsMotivicMeasure.pdf, arXiv:math/0312229, with an appendix discussing Galois stratification.
[27] R. Hartshorne, Algebraic Geometry, Grad. Texts in Math., Vol. 52, Springer-Verlag, 1977.
[28] E. Hrushovski, The Elementary Theory of Frobenius Automorphisms, preprint July 24, 2012, pp. 1–149.
[29] N. M. Katz, Local-to-global extensions of representations of fundamental groups, *Ann. Inst. Fourier*, **36**(4) (1988), 69–106.
[30] N. Katz, Sums of Betti numbers in arbitrary characteristic, *Finite Fields Appl.*, **7** (2001) 29–44.
[31] E. Kiefe, Sets definable over finite fields and their zeta functions, *Trans. Am. Math. Soc.*, **223** (1976), 171–194.
[32] S. Lang and A. Weil, Number of points of varieties in finite fields, *Am. J. Math.*, **76** (1954), 819–827.
[33] D. Mumford, Introduction to Algebraic Geometry; The Red Book, Harvard Lect. Notes, 1966.
[34] J. Nicaise, Relative motives; theory of pseudo-finite fields, *Int. Math. Res. Pap.*, (2007), rpm001, 70 pp. doi:10.1093/imrp/rpm001.
[35] I. Tomašiç, A twisted theorem of Chebotarev, *Proc. Lond. Math. Soc. (3)*, **108**(2) (2014), 291–326.

[36] I. Tomašiç, Twisted Galois stratification, *Nagoya Math. J.*, **222**(1) (2016), 1–60. doi:10.1017/nmj.2016.9.
[37] D. Wan, Partial zeta functions of algebraic varieties over finite fields, *Finite Fields Appl.*, **7** (2001) 238–251.
[38] D. Wan, Rationality of partial zeta functions, *Indag. Math. N.S.*, **14**(2) (2003), 285–292.

Dorian Goldfeld and Giacomo Micheli

The algebraic theory of fractional jumps

Abstract: In this paper, we start by briefly surveying the theory of fractional jumps and transitive projective maps. Then we show some new results on the absolute jump index, on projectively primitive polynomials, and on compound constructions.

Keywords: Finite fields, projective automorphisms, pseudorandom number generation, fractional jumps, full orbit sequences

MSC 2010: 11B37, 15B33, 11T06, 11K45, 65C10, 37P25

1 Introduction

Generating sequences of pseudorandom numbers is of great importance in applied areas and especially in cryptography and for Monte Carlo methods (e. g., to compute integrals over the reals). The task of generating streams of pseudorandom numbers is closely related to the study of dynamical systems over finite fields, which have been of great interest recently [11, 12, 13, 14, 20, 21, 18, 19]. More in general, for an interesting survey on open problems in arithmetic dynamics, see [4]. Constructions of pseudorandom number generators are studied, for example, in [6, 7, 8, 9, 10, 16, 17, 25, 26]. This paper focuses on one of the most recent ones, provided in [1]. In a nutshell, [1] provides a new construction of pseudorandom number sequences using the theory of transitive projective maps. From an applied point of view, the interest of this new construction relies on the fact that it costs asymptotically less to compute than the classical Inversive Congruential Generator (ICG) sequence [1, Section 7] and also achieves the same discrepancy bounds as the ICG (see [1, Section 6]). From a purely mathematical perspective, the theory of fractional jumps is intimately connected with different areas of mathematics such as finite projective geometry, field theory, additive and analytic number theory, and can turn it into a very rich area of research.

The main task of this paper is to summarize the theory of the fractional jump (FJ) construction and complete some mathematical aspects which were left open in the previous papers. Finally, we also show that the compound construction for the inver-

Acknowledgement: The first author was partially supported by Simons Collaboration Grant 567168. The second author was supported by the Swiss National Science Foundation grant number 171249. The second author would like to thank the department of Mathematics of Columbia University for hosting him in September 2018, as most of these ideas were developed during this stay. The authors would like to thank Philippe Michel, Alessandro Neri, and Violetta Weger for interesting discussions and suggestions.

Dorian Goldfeld, Columbia University, New York, USA
Giacomo Micheli, University of South Florida, Tampa, USA, e-mail: gmicheli@usf.edu

https://doi.org/10.1515/9783110621730-009

sive congruential generator nicely extends to FJs. Also, we leave some open questions at the end of the paper.

Notation
Let q be a prime power, n a positive integer, and $\mathbb{F}_q$ be the finite field of order q. Let $\mathbb{A}^n$ be the affine space over $\mathbb{F}_q$ (for the purposes of this paper, this can be simply identified with $\mathbb{F}_q^n$). Let $\mathbb{P}^n$ be the projective space of dimension n over $\mathbb{F}_q$. Fix the standard projective coordinates $X_0, \dots, X_n$ on $\mathbb{P}^n$. Let $\mathrm{GL}_{n+1}(\mathbb{F}_q)$ be the group of invertible matrices over $\mathbb{F}_q$ and $\mathrm{PGL}_{n+1}(\mathbb{F}_q)$ be the group of projective automorphisms of $\mathbb{P}^n$. For the entire paper, we fix the canonical decomposition

$$\mathbb{P}^n = U \cup H,$$

where

$$U = \{[X_0 : \dots : X_n] \in \mathbb{P}^n : X_n \neq 0\} \cong \mathbb{A}^n,$$
$$H = \{[X_0 : \dots : X_n] \in \mathbb{P}^n : X_n = 0\} \cong \mathbb{P}^{n-1}.$$

For a group G and an element $g \in G$, we denote by $o(g)$ the order of g. Let $\Psi \in \mathrm{PGL}_{n+1}(\mathbb{F}_q)$. We can write Ψ as $[M]$ for some $M = (m_{i,j})_{i,j} \in \mathrm{GL}_{n+1}(\mathbb{F}_q)$. Let us denote by $\mathrm{DeHom}(\Psi)$ the n-tuple of rational functions

$$(f_1, \dots f_n) = \left(\frac{m_{1,n+1} + \sum_{j=1}^n m_{1,j} x_j}{m_{n+1,n+1} + \sum_{j=1}^n m_{n,j} x_j}, \dots, \frac{m_{n,n+1} + \sum_{j=1}^n m_{n,j} x_j}{m_{n+1,n+1} + \sum_{j=1}^n m_{n,j} x_j} \right).$$

When we have an n-tuple f of rational functions of degree 1 with the same denominator b, we say that b is the denominator of f. Unless otherwise stated all the logarithms are in basis 2.

2 The theory of fractional jumps

In this section, we survey the ingredients needed to construct transitive fractional jumps and give new results on projective primitivity.

2.1 Transitive projective maps

The first ingredient needed is a transitive automorphism of the projective space. We start by recalling the definition of projectively primitive polynomials, which are closely related to transitive projective automorphisms.

Definition 2.1. A polynomial $\chi \in \mathbb{F}_q[x]$ of degree m is said to be *projectively primitive* if the two following conditions are satisfied:

(i) χ is irreducible over $\mathbb{F}_q$;
(ii) for any root α of χ in $\mathbb{F}_{q^m} \cong \mathbb{F}_q[x]/(\chi)$, the class $[\alpha]$ of α in the quotient group $G = \mathbb{F}_{q^m}^*/\mathbb{F}_q^*$ generates G.

Remark 2.2. Clearly, any primitive polynomial is also projectively primitive.

A characterization can be derived from [2, Lemma 2] with $e = 1$.

Proposition 2.3. *An irreducible polynomial $\chi \in \mathbb{F}_q[x]$ of degree m is projectively primitive if and only if $x^{q-1} \in \mathbb{F}_q[x]/(\chi)$ has order $(q^m - 1)/(q - 1)$.*

In [1], transitive projective maps were characterized, we report the result here for completeness.

Theorem 2.4. [1, Theorem 3.4] *Let Ψ be an automorphism of $\mathbb{P}^n$ with $\Psi = [M] \in \mathrm{PGL}_{n+1}(\mathbb{F}_q)$. Then Ψ is transitive on $\mathbb{P}^n$ if and only if the characteristic polynomial $\chi_M \in \mathbb{F}_q[x]$ of M is projectively primitive.*

Remark 2.5. Theorem 2.4 also implies that to find a transitive projective automorphism of $\mathbb{P}^n$ one can simply fix $\Psi = [M_f] \in \mathrm{PGL}_{n+1}(\mathbb{F}_q)$, where M_f is the companion matrix (or any of its conjugates) of a projectively primitive polynomial f.

The following result shows that one can in principle always construct a primitive polynomial from a projectively primitive one.

Theorem 2.6. *A polynomial $f \in \mathbb{F}_q[x]$ is projectively primitive if and only if there exists $\lambda \in \mathbb{F}_q^*$ such that $f(x/\lambda)$ is primitive.*

Proof. If there exists $\lambda \in \mathbb{F}_q^*$ such that $f(x/\lambda)$ is primitive, then it is obvious that f is projectively primitive. Let us now show the other implication. Let α be a root of f in its splitting field $\mathbb{F}_{q^{\deg(f)}}$. We have to find λ such that $\lambda\alpha$ has order $q^{\deg(f)} - 1$. Recall that for an element $\beta \in \mathbb{F}_{q^{\deg(f)}}^*$ we denote by $[\beta]$ its reduction in the quotient group $G = \mathbb{F}_{q^{\deg(f)}}^*/\mathbb{F}_q^*$.

First, observe that for any $\lambda \in \mathbb{F}_q^*$, we have that $N = (q^{\deg(f)} - 1)/(q - 1)$ divides $o(\lambda\alpha)$ because $N = o([\alpha]) = o([\lambda\alpha])$. So if we can find $\lambda \in \mathbb{F}_q^*$ such that $(\lambda\alpha)^N$ has order $q - 1$ we are done.

Choose a multiplicative generator g of $\mathbb{F}_q^*$ and write $\alpha^N = \mu = g^e$ for some positive integer e. Moreover, assume that the choice of g is also such that e is minimal. First, observe that all the prime factors of e divide $q - 1$ as otherwise if p is a prime factor of e that does not divide $q - 1$, one can rewrite $(g^p)^{e/p} = \mu$, and g^p is again a generator for $\mathbb{F}_q^*$, contradicting the minimality of e.

We now want to prove that $\gcd(N, e) = 1$. Suppose the contrary and let p be a prime factor of $\gcd(N, e)$. Consider $\gamma = \alpha^{N/p}$, if we show that $\gamma^{q-1} = 1$ we get the contradiction by the definition of N (N is the smallest integer such that $\alpha^N \in \mathbb{F}_q^*$). But this is obvious:

$$\gamma^{q-1} = \alpha^{N(q-1)/p} = \left(g^e\right)^{(q-1)/p} = \left(g^{e/p}\right)^{(q-1)} = 1.$$

Since we want that $(\lambda\alpha)^N$ has order $q-1$, we have to select λ such that $(\lambda\alpha)^N$ is a multiplicative generator of $\mathbb{F}_q^*$. Write $\lambda = g^s$ for some $s \in \mathbb{N}$, then we can write

$$(\lambda\alpha)^N = g^{sN}\alpha^N = g^{sN}\mu = g^{sN+e}.$$

Since N and e are coprime, the Dirichlet theorem on arithmetic progressions applies, therefore, we can select $\bar{s}$ such that $P = \bar{s}N + e$ is a prime larger than $q-1$. The claim follows by observing that if g is a generator for $\mathbb{F}_q^*$, then g^P is a generator of $\mathbb{F}_q^*$. □

A direct consequence of the result above is that when q is small, the problems of finding a primitive polynomial or a projectively primitive one are equivalent.

Corollary 2.7. *Given a monic projectively primitive polynomial f over $\mathbb{F}_q$, constructing a primitive polynomial costs $O(q\log(q)\log(\deg(f)))$ operations in $\mathbb{F}_q$.*

Proof. We first factor $q-1$ as a precomputation, which costs less than $O(\sqrt{q})$. Given a monic projectively primitive polynomial f and one of its roots $\alpha \in \mathbb{F}_{q^{\deg(f)}} = \mathbb{F}_q[x]/(f(x))$, we simply test (for any λ in $\mathbb{F}_q$) if $\beta_\lambda = (\lambda\alpha)^{\frac{q^{\deg(f)}-1}{q-1}}$ has order $q-1$. The cost is then as follows. Observe that the norm of α is given by the degree zero coefficient of f, so $\beta = N(\alpha) = \alpha^{\frac{q^{\deg(f)}-1}{q-1}}$ does not have to be computed. Since β lives in $\mathbb{F}_q$, for any $\lambda \in \mathbb{F}_q^*$, we check if $\lambda^{\frac{q^{\deg(f)}-1}{q-1}}\beta = \lambda^{\deg(f)}\beta = \beta_\lambda$ has order $q-1$ in $\mathbb{F}_q^*$. To do that, we simply compute $\beta_\lambda^{(q-1)/r}$, where r runs over all prime divisors of $q-1$, which are at most $O(\log(q))$. The total number of $\mathbb{F}_q$-operations is then $O(q\log(q)\log(\deg(f)))$, where $O(\log(\deg(f)))$ is the cost of computing $\lambda^{\deg(f)}$. □

Remark 2.8. I. Shparlinski observed that using an adaptation of the method in [23] it is probably possible to improve the result in Corollary 2.7 to a running time that is soft-proportional to $q^{1/4}$.

We recall now the definition of fractional jump index.

Definition 2.9. Let Ψ be an automorphism of $\mathbb{P}^n$. Let $U = \{[X_0, X_1, \dots, X_{n-1}, 1] : \forall i \in \{0,\dots,n-1\} X_i \in \mathbb{F}_q\} \subseteq \mathbb{P}^n$ and $P \in U$. The *fractional jump index of Ψ at P* is

$$\mathfrak{J}_{P,\Psi} = \min\{k \geq 1 : \Psi^k(P) \in U\}.$$

The *absolute fractional jump index $\mathfrak{J}$ of Ψ* is the quantity

$$\mathfrak{J}_\Psi = \max\{\mathfrak{J}_{P,\Psi} : P \in U\}.$$

In [1], it is shown that for a transitive projective map, the absolute jump index cannot be larger than $n+1$.

Proposition 2.10 ([1, Corollary 4.3]). *Let $\Psi \in \mathrm{PGL}_n(\mathbb{F}_q)$ be transitive. The absolute jump index of $\mathfrak{J}_\Psi$ of Ψ is less than or equal to $n+1$.*

We can actually prove a stronger result

Theorem 2.11. *Let $\Psi \in \mathrm{PGL}_{n+1}(\mathbb{F}_q)$ be transitive. Then $\mathfrak{J}_\Psi = n+1$.*

Proof. The direction $\mathfrak{J}_\Psi \leq n+1$ is given by Proposition 2.10. Let us show that $\mathfrak{J}_\Psi \geq n+1$. Recall that $H = \{[X_0 : \ldots : X_n] \in \mathbb{P}^n : X_n = 0\} \cong \mathbb{P}^{n-1}$. Let L be the largest integer such that there exists a point $\overline{P} \in \mathbb{P}^n$ such that

$$\{\Psi(\overline{P}), \Psi^2(\overline{P}), \ldots \Psi^L(\overline{P})\} \subseteq H,$$

so that $\mathfrak{J}_\psi = L+1$. Observe that we can always choose $\overline{P}$ in U because Ψ is transitive: in fact, consider the smallest ℓ such that $P' = \Psi^{-\ell}(\overline{P}) \in U$ (this is possible as Ψ is transitive). Then

$$\{\Psi(P'), \Psi^2(P'), \ldots \Psi^{L+\ell}(P')\} \subseteq H.$$

This forces $\ell = 0$ and, therefore, $\overline{P} \in U$.

Set

$$T = \{P \in \mathbb{P}^n : \Psi^i(P) \in H \ \forall i \in \{1, \ldots, L\}\}.$$

It is easy to see that T is nonempty by the choice of L, and is a projective subspace of $\mathbb{P}^n$ that intersects U, because $\overline{P} \in U$. We want to show that the dimension of T is zero, so it consists only of one point. Consider $\Psi^{L+1}(T)$ (that has the same dimension of T) and assume by contradiction that its dimension is greater than or equal to 1. Then its intersection with H is nonempty as H is a projective hyperplane, so let $Q \in \Psi^{L+1}(T) \cap H$. Set $R = \Psi^{-L-1}(Q)$ and observe that $\Psi^i(R) \in H$ for any $i \in \{1, \ldots, L\}$ as $R \in T$, but also $\Psi^{L+1}(R) \in H$ by construction, which is a contradiction by the maximality of L. This forces $\dim \Psi^{L+1}(T) = \dim T = 0$ which forces $T = \{\overline{P}\}$. Now, since $\dim T \geq n - L$ (each of the conditions $\Psi^i(T) \subseteq H$ imposes an equation), this forces $L \geq n$. Therefore, $\mathfrak{J}_\Psi \geq n+1$. □

Remark 2.12. Transitivity is necessary for the result above to hold: consider for example the nontransitive map of $\mathbb{P}^1$ given by $[X, Y] \mapsto [X+Y, Y]$. The absolute jump index is 1 (no point at finite is mapped at infinite).

2.2 Constructing a transitive fractional jump

The fractional jump of a projective map can be formally defined as follows.

Definition 2.13. Let $U = \{[X_0, X_1, \ldots, X_{n-1}, 1] : \forall i \in \{0, \ldots, n-1\} X_i \in \mathbb{F}_q\} \subseteq \mathbb{P}^n$ and

$$\begin{aligned} \pi : \mathbb{A}^n &\longrightarrow U \\ (x_1, \ldots, x_n) &\mapsto [x_1, \ldots, x_n, 1]. \end{aligned}$$

The *fractional jump of* Ψ is the map

$$\begin{aligned} \psi : \mathbb{A}^n &\to \mathbb{A}^n \\ x &\mapsto \pi^{-1} \Psi^{\mathfrak{J}_{\pi(x)}} \pi(x). \end{aligned}$$

Remark 2.14. The fractional jump is clearly well-defined but its definition depends on the point where it is evaluated, which might be an issue if one wants to describe the map globally. Theorem 2.16 ensures that this is not the case.

Obviously, if one starts with a transitive projective automorphism one will get a transitive fractional jump. Interestingly enough, the converse implication is also true, apart from two degenerate cases; see [2, Theorem 2] where this issue is settled. We report the result here for completeness.

Theorem 2.15. *Let Ψ be an automorphism of $\mathbb{P}^n$ and let ψ be its fractional jump. Then Ψ acts transitively on $\mathbb{P}^n$ if and only if ψ acts transitively on $\mathbb{A}^n$, unless q is prime and $n = 1$, or $q = 2$ and $n = 2$, with explicit examples in both cases.*

In [1], an explicit global description of a fractional jump was given.

Theorem 2.16 ([1, Section 5] or [2, Theorem 1]). *Let Ψ be a transitive automorphism of $\mathbb{P}^n$, and let ψ be its fractional jump. Then, for $i \in \{1, \dots, n+1\}$ there exist*

$$a_1^{(i)}, \dots, a_n^{(i)}, b^{(i)} \in \mathbb{F}_q[x_1, \dots, x_n]$$

of degree 1 such that, if

$$\begin{aligned} U_1 &= \{x \in \mathbb{A}^n : b^{(1)}(x) \neq 0\}, \\ U_i &= \{x \in \mathbb{A}^n : b^{(i)}(x) \neq 0, \text{ and } b^{(j)}(x) = 0, \ \forall j \in \{1, \dots, i-1\}\}, \\ &\quad \text{for } i \in \{2, \dots, n+1\}, \\ &\quad \text{and} \\ f^{(i)} &= \left(\frac{a_1^{(i)}}{b^{(i)}}, \dots, \frac{a_n^{(i)}}{b^{(i)}}\right), \\ &\quad \text{for } i \in \{1, \dots, n+1\}, \end{aligned}$$

then $\psi(x) = f^{(i)}(x)$ if $x \in U_i$. Moreover, the rational maps $f^{(i)}$ can be explicitly computed.

Remark 2.17. Observe that the datum of a fractional jump ψ is equivalent to the datum of the vector of degree 1 polynomials $(a^{(1)}, \dots a^{(n+1)}; b^{(1)}, \dots, b^{(n)})$ where $a^{(i)} = (a_1^{(i)}, a_2^{(i)}, \dots, a_n^{(i)})$.

3 Fractional jumps in practice

In this section, we describe some aspects of the practical implementation of fractional jumps.

3.1 Compact description

In this section, we give a compact description of a fractional jump. We first need an ancillary lemma.

Lemma 3.1. *Let $\Psi = [M] \in \mathrm{PGL}_{n+1}(\mathbb{F}_q)$ be transitive. For $i \in \{1,\dots,n+1\}$, set $f^{(i)} = \mathrm{DeHom}(M^i)$ and set $b^{(i)}$ to be the denominator of $f^{(i)}$. Then $b^{(1)} \neq 0$ and for any $i \in \{2,\dots,n+1\}$ we have that $b^{(i)} \not\equiv 0 \bmod b^{(1)},\dots,b^{(i-1)}$.*

Proof. First, observe that since Ψ is transitive we have that:
- the characteristic polynomial of M is irreducible and equal to the minimal polynomial μ_M
- $b^{(1)}$ is different from 1, as otherwise no point at finite is mapped at infinity and, therefore, the map cannot be transitive on $\mathbb{P}^n$.

Let j be the smallest integer such that $b^{(j)} \equiv 0 \bmod (b^{(1)},\dots,b^{(j-1)})$. Of course, we can assume $j \leq n+1$. By degree reasons, there exist $\lambda_1,\lambda_2,\dots,\lambda_{j-1} \in \mathbb{F}_q$ such that $(\sum_{k=1}^{j-1}\lambda_k b^{(k)}) - b^{(j)} = 0$. But this implies that the matrix

$$N = \left(\sum_{k=1}^{j-1}\lambda_k M^k\right) - M^j = M\left(\left(\sum_{k=1}^{j-1}\lambda_k M^{k-1}\right) - M^{j-1}\right)$$

has the last row identically zero, so it is not invertible. But since the characteristic polynomial of M is irreducible, any matrix in $\mathbb{F}_q[M]\backslash\{0\}$ is invertible. This forces $N = 0$. But then the polynomial $g = (\sum_{k=1}^{j-1}\lambda_k X^{k-1}) - X^{j-1}$ is zero at M and, therefore, divisible by the minimal polynomial μ_M. But since $j-1 \leq n$ and μ_M has degree $n+1$, we must have $g = 0$, which is a contradiction because g has degree $j-1$. □

We are now ready to provide a compact description of a fractional jump.

Algorithm 1. Fractional Jump Generation Algorithm

Input: a projectively primitive morphism $\Psi = [M] \in \mathrm{PGL}_{n+1}(\mathbb{F}_q)$
Output: the fractional jump of Ψ

$M^{(1)} \leftarrow M$ ▷ $m^{(1)}_{h,k}$ is the hth row, kth column entry of the matrix $M^{(1)}$.
for $h \in \{1,\dots n\}$ **do**
  $a^{(1)}_h \leftarrow m^{(1)}_{h,n+1} + \sum_{k=1}^{n} m^{(1)}_{h,k}x_k$
$b^{(1)} \leftarrow m^{(1)}_{n+1,n+1} + \sum_{k=1}^{n} m^{(1)}_{n+1,k}x_k$
$a^{(1)} \leftarrow (a^{(1)}_1,\dots,a^{(1)}_n)$
for $i \in \{2,\dots n+1\}$ **do**
  $M^{(i)} \leftarrow M^i$ ▷ $m^{(i)}_{h,k}$ is the hth row, kth column entry of the matrix $M^{(i)}$.
  $b^{(i)} \leftarrow m^{(i)}_{n+1,n+1} + \sum_{k=1}^{n} m^{(i)}_{n+1,k}x_k \bmod b^{(1)}, b^{(2)},\dots b^{(i-1)}$
  for $h \in \{1,\dots n\}$ **do**
    $a^{(i)}_h \leftarrow m^{(i)}_{h,n+1} + \sum_{k=1}^{n} m^{(i)}_{h,k}x_k \bmod b^{(1)}, b^{(2)},\dots b^{(i-1)}$
  $a^{(i)} \leftarrow (a^{(i)}_1,\dots a^{(i)}_n)$
return $(a^{(1)}, a^{(2)},\dots,a^{(n+1)}), (b^{(1)}, b^{(2)},\dots,b^{(n+1)})$

Theorem 3.2. *Storing a fractional jump requires at most* $\lceil \log(q) \rceil (n+1)^2(n+2)/2$ *bits.*

Proof. Algorithm 1 produces a fractional jump from a transitive projective automorphism. Now observe that the bit size of $(a^{(1)}, b^{(1)})$ is the same as the bit size of M, which is $(n+1)^2\lceil \log(q) \rceil$. The bit size of $(a^{(2)}, b^{(2)})$ is $(n+1)n\lceil \log(q) \rceil$ as we were able to use the relation $b^{(1)} = 0$. More in general, the bit size of $(a^{(i)}, b^{(i)})$ is $(n+1)(n+2-i)\lceil \log(q) \rceil$ as we can use the relation $b^{(1)} = b^{(2)} = \cdots = b^{(i-1)} = 0$. The process terminates and it is well-defined because of Lemma 3.1. Adding everything up, we get

$$\sum_{i=1}^{n+1} (n+1)(n+2-i)\lceil \log(q) \rceil = \lceil \log(q) \rceil (n+1)^2(n+2)/2. \qquad \square$$

3.2 Expected cost of evaluation

Evaluating a fractional jump is a very easy task, as it involves only one inversion in the base field. In this section, we compute the *expected cost* of evaluating a fractional jump, essentially weighting the computational cost with the probability that a random point in $\mathbb{F}_q^n$ is selected.

Definition 3.3. Let ψ be a map on $\mathbb{F}_q^n$. We define the *expected cost* of computing ψ on $\mathbb{F}_q^n$ to be

$$\mathbb{E}[\psi] = q^{-n} \sum_{x \in \mathbb{F}_q^n} \mathrm{Cost}(\psi, x),$$

where $\mathrm{Cost}(\psi, x)$ denotes the number of binary operations needed to evaluate ψ at x.

We now compute the expected complexity of evaluating a fractional jump sequence in the large field regime, which is the one for which we have the nice discrepancy bounds in [1, Section 8].

Algorithm 2. Fractional Jump Evaluation Algorithm

Input: a fractional jump ψ and a point $y \in \mathbb{F}_q^n$.
Output: $\psi(y)$.

1: **for** $i \in \{1, \dots n+1\}$ **do**
2: **if** $b^{(i)}(y) \neq 0$ **then**
3: $c \leftarrow b^{(i)}(y)^{-1}$
4: $v \leftarrow a^{(i)}(y)$
5: **return** cv

Theorem 3.4. *Let q be a prime, $\Psi = [M] \in \mathrm{PGL}_{n+1}(\mathbb{F}_q)$ be a transitive projective automorphism, and ψ be its fractional jump. Suppose that $[M]$ has a representative in*

$\mathrm{GL}_{n+1}(\mathbb{Q})$ *having entries in* $\{-1,0,1\}$. *Suppose that* $q \geq n^3$. *The expected cost of evaluating a fractional jump is* $O((n+\log\log(q))\log(q)\log\log(q)\log\log\log(q)+n^2\log(q))$.

Proof. We want to estimate the average cost of Algorithm 2. As usual, set $U^{(1)} = \{x \in \mathbb{A}^n(\mathbb{F}_q) : b^{(1)}(x) \neq 0\}$ and for $i \in \{2,\dots n+1\}$ set

$$U^{(i)} = \{x \in \mathbb{A}^n(\mathbb{F}_q) : b^{(i)}(x) \neq 0, \text{and } b^{(1)}(x) = b^{(2)}(x) = \cdots = b^{(i-1)}(x) = 0\},$$

and

$$\mathbb{E}[\psi] = q^{-n}\sum_{i=1}^{n+1}\sum_{x\in U^{(i)}}\mathrm{Cost}(\psi,x).$$

For $x \in U^{(1)}$, by the fact that M has small coefficients, evaluating $a^{(1)}$ and $b^{(1)}$ involves at most $O(n^2)$ sums. Therefore, we have that $\mathrm{Cost}(\psi,x) = O(I(q)+nM(q)+n^2S(q))$, where $I(q)$ is the cost of an inversion, $M(q)$ is the cost of a multiplication in $\mathbb{F}_q$ and $S(q)$ the cost of an addition in $\mathbb{F}_q$. For $x \in U^{(i)}$ and $i \geq 2$, evaluating $a^{(i)}$ and $b^{(i)}$ becomes more expensive, as it might involve also $n-1$ multiplications by elements of $\mathbb{F}_q$ for each component (the coefficients $m_{h,k}^{(i)}$). The final cost of evaluating at $x \in U^{(i)}$ is then $\mathrm{Cost}(\psi,x) = O(I(q)+(n+n^2)M(q)+n^2S(q))$. Since there are $q^n - q^{n-1}$ elements in $U^{(1)}$ and q^{n-1} in the union of the rest of the $U^{(i)}$'s, we have that

$$\mathbb{E}[\psi] = O\left(I(q)+nM(q)+n^2S(q)+\frac{I(q)+(n+n^2)M(q)+n^2S(q)}{q}\right).$$

Since $q > n^3$ and $I(q), M(q), S(q)$ are all polynomial time operations in $\log(q)$, we have that $\frac{I(q)+(n+n^2)M(q)+n^2S(q)}{q} = O(1)$ and then

$$\mathbb{E}[\psi] = O(I(q)+nM(q)+n^2S(q)).$$

Observe that if one uses fast Fourier transform for multiplication [22] and the Schönhage algorithm for inversions [15, Remark 11.1.99] we have that

$$I(q) = M(q)\log\log(q)$$

and

$$M(q) = \log(q)\log\log(q)\log\log\log(q).$$

Adding two integers modulo q simply costs $O(\log(q))$, from which we get the final claim. □

Example 3.5. Fix, for example, $p = 3868562622766813359059780 3$ and $f = x^3 - x - 1 \in \mathbb{F}_p[x]$. One can check with a computer algebra system (e. g., SAGE [24]) that $(p^3-1)/(p-1)$ is a prime number and that f is an irreducible polynomial. It follows directly from

Definition 3.3 that f is projectively primitive and, therefore, the projective map produced by its companion matrix (see Remark 2.5) verifies the hypothesis of Theorem 2.4, and thus it generates a transitive fractional jump verifying the hypothesis of Theorem 3.4. Computationally, it is very easy to produce projectively primitive polynomials, but it would also be interesting to give a systematic way to construct them (such as the one using Artin–Schreier jumps in [2]).

Remark 3.6. In terms of expected complexity (and whenever the coefficients are carefully chosen), fractional jumps behave better than ICGs, as we are about to explain. In fact, let us now compare the result of Theorem 3.4 for $n > 1$ with $n = 1$ which is essentially the case of the ICG (see [1, Example 2.4]). Evaluating an ICG having small coefficients costs one inversion $O(I(q))$ whether evaluating a fractional jump with small coefficients costs averagely $O(I(q)+nM(q)+n^2S(q))$. Notice now that if q is a large prime and n is relatively small we have that $I(q) + nM(q) + n^2S(q) \sim I(q)$. On the other hand, an ICG only generates one pseudorandom point at each iteration, whether instead the fractional jump construction generates n-pseudorandom points.

3.3 Compound generator for fractional jumps

In this subsection, we show that the compound generator construction for the inversive congruential generator easily extends to a fractional jump and provide an example.

Theorem 3.7. *Let ℓ and n be positive integers and $\{p_1, p_2, \dots, p_\ell\}$ be ℓ distinct primes. For any $i \in \{1, \dots \ell\}$, let Ψ_i be a transitive projective automorphism of $\mathbb{P}^n(\mathbb{F}_{p_i})$ and $\psi_i : \mathbb{A}^n(\mathbb{F}_{p_i}) \longrightarrow \mathbb{A}^n(\mathbb{F}_{p_i})$ be its fractional jump.*

Let $N = p_1 \cdots p_\ell$ and $R = \mathbb{Z}/N\mathbb{Z}$. There exists a transitive map ψ on R^n such that, for any $i \in \{1, \dots n\}$, its reduction modulo p_i is ψ_i.

Proof. Let

$$v_i = \prod_{\substack{j=1 \\ j \neq i}}^{\ell} p_j$$

and r_i be a representative modulo N of the inverse of v_i modulo p_i. Set $u_i = v_i r_i$ and L_i the map which takes as input an element of $\mathbb{F}_{p_i}^n$ and outputs its canonical representative in $\{0, \dots p_i - 1\}^n \subseteq R^n$. Consider the map

$$\psi : R^n \longrightarrow R^n$$
$$x \mapsto \sum_{i=1}^{\ell} u_i \overline{\psi}_i(x)$$

where

$$\overline{\psi}_i(x) = L_i(\psi_i(x \bmod p_i)).$$

First, observe that ψ is well-defined, as it is a sum of well-defined maps. We have now to prove that ψ is a bijection. To see this, notice that we have the following diagram:

$$\begin{array}{ccc} R^n & \xrightarrow{\psi} & R^n \\ {\scriptstyle \pi_i}\downarrow & & {\scriptstyle \pi_i}\downarrow \\ \mathbb{F}_{p_i}^n & \xrightarrow{\psi_i} & \mathbb{F}_{p_i}^n \end{array}$$

where π_i is the natural reduction of R^n modulo p_i. The diagram is commutative thanks to the choice of u_i, which is zero modulo p_j for any $j \neq i$, and modulo p_i is equal to 1. We want to prove first that ψ is surjective. Let $z \in R^n$ and consider $z_i = \pi_i(z)$. Since ψ_i is bijective, there exists $x_i \in \mathbb{F}_{p_i}^n$ such that $\psi_i(x_i) = z_i$. By the Chinese remainder theorem, we can find $x \in R^n$ such that $x \equiv x_i \bmod p_i$ for all $i \in \{1, \dots, \ell\}$. It is now immediate to see that $\psi(x) = z$. So ψ is surjective and therefore bijective as R^n is a finite set.

We have now to show that ψ is transitive. To see this, we will show that the order of an element $\overline{x} \in R^n$ is zero modulo p_i^n for any $i \in \{1, \dots, \ell\}$, so the claim will follow as the order of ψ at $\overline{x}$ is at most N^n. Suppose that d is a positive integer such that $\psi^d(\overline{x}) = \overline{x}$, then applying π_i on both sides and using the commutativity of the diagram we have that

$$\pi_i(\psi^d(\overline{x})) = \psi_i^d(\pi_i(\overline{x})) = \pi_i(\overline{x}),$$

from which it follows that d must be divisible by p_i^n as ψ_i is transitive. □

Remark 3.8. Notice that also other lifts L_i to R^n would be suitable for the compound generator, not only the canonical one $\mathbb{F}_{p_i}^n \to \{0, 1, \dots, p_i - 1\}^n \subseteq R^n$.

Example 3.9. To fix the ideas for our constructions, we produce here a small toy example for $R = \mathbb{Z}/15\mathbb{Z}$ and $n = 2$. Let us construct first a transitive projective map over $\mathbb{P}^2(\mathbb{F}_5)$. For this, consider the polynomial $x^3 + 3x + 3 \in \mathbb{F}_5[x]$ and its companion matrix

$$M = \begin{pmatrix} 0 & 0 & 3 \\ -1 & 0 & 3 \\ 0 & -1 & 0 \end{pmatrix}.$$

To compute the fractional jump of $\Psi_1 = [M] \in \mathrm{PGL}_3(\mathbb{F}_5)$, we also need the matrices M^2 and M^3:

$$M^2 = \begin{pmatrix} 0 & -3 & 0 \\ 0 & -3 & -3 \\ 1 & 0 & -3 \end{pmatrix} \quad M^3 = \begin{pmatrix} 3 & 0 & 1 \\ 3 & 3 & 1 \\ 0 & 3 & 3 \end{pmatrix}.$$

The fractional jump of $[M]$ is then

$$\psi_1(x_1, x_2) = \begin{cases} (\frac{2}{x_2}, \frac{x_1-3}{x_2}) & \text{if } x_2 \neq 0 \\ (0, \frac{2}{x_1+2}) & \text{if } x_2 = 0 \text{ and } x_1 \neq 3 \\ (0,0) & \text{if } x = (3,0) \end{cases}$$

We now need a projectively primitive polynomial of degree 3 over $\mathbb{F}_3$. We select $x^3 + 2x + 1 \in \mathbb{F}_3[x]$. Its companion matrix is

$$M = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 2 \\ 0 & -1 & 0 \end{pmatrix}.$$

Analogously, one computes the fractional jump of $\Psi_2 = [M] \in \mathrm{PGL}_3(\mathbb{F}_3)$ obtaining

$$\psi_2(x_1, x_2) = \begin{cases} (-\frac{1}{x_2}, \frac{x_1-2}{x_2}) & \text{if } x_2 \neq 0 \\ (0, -\frac{1}{x_1+1}) & \text{if } x_2 = 0 \text{ and } x_1 \neq 2 \\ (0,0) & \text{if } x = (2,0) \end{cases}$$

The compound generator of ψ_1 and ψ_2 is then

$$\psi : R^2 \longrightarrow R^2$$
$$\psi(x_1, x_2) = 6 \cdot L_1(\psi_1(x_1 \bmod 5, x_2 \bmod 5)) + 10 \cdot L_2(\psi_2(x_1 \bmod 3, x_2 \bmod 3))$$

where L_1 (resp., L_2) is the obvious map lifting $\mathbb{F}_5$ (resp., $\mathbb{F}_3$) to $\{0,1,2,3,4\}$ (resp., $\{0,1,2\}$) in $\mathbb{Z}/15\mathbb{Z}$. One can check directly that ψ is in fact transitive on R^2.

4 Some ideas to achieve unpredictability from a fractional jump sequence

Since we already have nice (provable) distributional properties of FJs given by the results in [1] (which make fractional jumps suitable for Monte Carlo methods, e. g.), in this section we would like to provide some modifications of the fractional jump construction that could be of use for pseudorandom number generation in settings where unpredictability is a critical property (such as cryptography). In this setting, we have an opponent observing the stream of pseudorandom numbers and he must not be able to reconstruct the generator, or predict next values of the stream.

Remark 4.1. We would like to observe that the main issue we encounter when we want to use the basic fractional jump construction for pseudorandom number generation in a cryptographic setting is the following: when the base field $\mathbb{F}_q$ is large, on most of the points of $\mathbb{F}_q^n$ we act as n rational functions in n variables of degree 1 (more precisely in

the notation of Theorem 2.15 we act as $f^{(1)}$ on all points of U_1, which are $q^n - q^{n-1}$). Therefore, for each pseudorandom number we observed, we get a system of linear equations in the coefficients of the rational functions defining $f^{(1)}$. It is therefore expected that in $(n+1)^2$ points we can reconstruct $f^{(1)}$ by solving a linear system (assuming that all the points in the iteration lie all in U_1, which is a reasonable assumption as it has size comparable with q^n).

In what follows, we describe some constructions which seem to avoid the issue presented in the remark above.

4.1 Secret prime q

Here, we follow the ideas of [3]. Choose two large odd primes p, q with the property that $p < q$ and $q = kp + 2$ if $p \neq 2$ is odd. The designer keeps q secret, constructs a secret full orbit fractional jump $\psi : \mathbb{F}_q^n \longrightarrow \mathbb{F}_q^n$, and chooses a secret starting point $u_0 \in \mathbb{F}_q^n$. Consider now the canonical lift $L : \mathbb{F}_q^n \longrightarrow \{0, 1, \ldots, q-1\}^n$. The pseudorandom sequence is then produced as $L(\psi^m(u_0)) \bmod p$. To avoid the small biases given by the reduction, one can use rejection sampling by skipping elements of the sequence $\psi^m(u_0) \bmod q$ that have components that are congruent to $q-2$ or $q-1$ modulo q. Of course, p should be chosen relatively small compared with q.

4.2 Forcing jumps

Let $\psi : \mathbb{F}_q \longrightarrow \mathbb{F}_q$ be a fractional jump, T be a subset of $\mathbb{F}_q^n$ roughly of size $(q^n - 1)/2$, T^c be its complement. Define the map

$$\phi(x) = \begin{cases} \psi(x) & \text{if } x \in T \\ \psi(\psi(x)) & \text{if } x \in T^c \end{cases}$$

The designer keeps ψ, T, T^c, and ϕ, secret and outputs the sequence $\phi^m(0)$. If one wants to reconstruct the fractional jump ψ, according to Remark 4.1, one would need to observe at least $(n+1)^2$ iterations of ψ. But in this construction, either ψ or ψ^2 is used with probability 1/2, therefore, in order to reconstruct ψ the attacker has $2^{(n+1)^2}$ systems to solve, one of which will lead to the reconstruction of ψ. Notice that with this construction the orbit of ϕ starting at any point is bounded from below by $q^n/2$.

5 Further research

In this section, we list some questions arising from the theory of fractional jumps.

Of course, any primitive polynomial is also projectively primitive. Moreover, we saw in Corollary 2.7 that whenever q is small, finding a primitive polynomial or a projectively primitive polynomials are equivalent problems.

Question 5.1. For a fixed degree (e. g., 3), can one produce algorithms to find projectively primitive polynomials similar to the one in [5]?

Also, it would be very interesting to see attacks to the constructions in Section 4.

Question 5.2. Are there (nontrivial) attacks to the constructions in the Subsections 4.1 and 4.2?

Finally, we ask to compute the linear complexity of fractional jump sequences, i. e., if $\{v_i\}$ is the sequence in $\mathbb{F}_q^n$ compute good lower bounds for the minimal N such that there exist $c_1, \dots, c_{N-1} \in \mathbb{F}_q$ such that for all $i \in \{0, \dots, q^n - 1\}$ we have $v_{N+i} = \sum_{j=0}^{N-1} c_j v_{i+j}$.

Question 5.3. What is the linear complexity of fractional jump sequences produced using the methods described in this paper?

Theorem 3.4 ensures that computing a fractional jump sequence arising from a transitive projective automorphism having a representative matrix with small coefficients has small computational cost (in comparison with the inversive congruential generator, e. g.). Theorem 2.4 implies that the projective automorphism obtained using the companion matrix of a projectively primitive polynomial is transitive. It is therefore natural to ask the following.

Question 5.4. In which cases one can construct a projectively primitive polynomial with small coefficients?

For example, the results in [2] ensure that this is always possible in degree p over the finite field $\mathbb{F}_p$ using $x^p - x + a$.

Bibliography

[1] Federico Amadio Guidi, Sofia Lindqvist, and Giacomo Micheli, Full orbit sequences in affine spaces via fractional jumps and pseudorandom number generation, *Math. Comput.*, 2018.

[2] Federico Amadio Guidi and Giacomo Micheli, Fractional jumps: complete characterisation and an explicit infinite family, In: International Workshop on the Arithmetic of Finite Fields, Springer, 2018, pp. 250–263.

[3] Michael Anshel, Dorian Goldfeld, et al., Zeta functions, one-way functions, and pseudorandom number generators, *Duke Math. J.*, **88**(2) (1997), 371–390.

[4] Robert Benedetto, Laura DeMarco, Patrick Ingram, Rafe Jones, Michelle Manes, Joseph H Silverman, and Thomas J Tucker. Current trends and open problems in arithmetic dynamics, arXiv preprint arXiv:1806.04980, 2018.

[5] W.-S. Chou, On inversive maximal period polynomials over finite fields, *Appl. Algebra Eng. Commun. Comput.*, **6**(4) (1995), 245–250.

[6] Jürgen Eichenauer-Herrmann, Inversive congruential pseudorandom numbers avoid the planes, *Math. Comput.*, **56**(193) (1991), 297–301.

[7] Jürgen Eichenauer-Herrmann, Inversive congruential pseudorandom numbers: a tutorial, *Int. Stat. Rev.*, (1992), 167–176.
[8] Jürgen Eichenauer-Herrmann, Statistical independence of a new class of inversive congruential pseudorandom numbers, *Math. Comput.*, **60**(201) (1993), 375–384.
[9] Jürgen Eichenauer-Herrmann, Eva Hermann, and Stefan Wegenkittl, A survey of quadratic and inversive congruential pseudorandom numbers, *Lect. Notes Stat.*, **127** (1998), 66–97.
[10] Edwin D. El-Mahassni and Domingo Gómez-Pérez, On the distribution of nonlinear congruential pseudorandom numbers of higher orders in residue rings, In: Applied Algebra, Algebraic Algorithms, and Error-Correcting Codes, Springer, 2009, pp. 195–203.
[11] Andrea Ferraguti, Giacomo Micheli, and Reto Schnyder, On sets of irreducible polynomials closed by composition, In: International Workshop on the Arithmetic of Finite Fields, Lecture Notes in Comput. Sci., Springer, 2016, pp. 77–83.
[12] Andrea Ferraguti, Giacomo Micheli, and Reto Schnyder, Irreducible compositions of degree two polynomials over finite fields have regular structure, arXiv preprint arXiv:1701.06040, 2017.
[13] Domingo Gómez-Pérez, Alina Ostafe, and Igor E. Shparlinski, Algebraic entropy, automorphisms and sparsity of algebraic dynamical systems and pseudorandom number generators, *Math. Comput.*, **83**(287) (2014), 1535–1550.
[14] David Rodney Heath-Brown and Giacomo Micheli, Irreducible polynomials over finite fields produced by composition of quadratics, arXiv preprint arXiv:1701.05031, 2017.
[15] Gary L Mullen and Daniel Panario, Handbook of Finite Fields, Chapman and Hall/CRC, 2013.
[16] Harald Niederreiter and Igor E. Shparlinski, Recent advances in the theory of nonlinear pseudorandom number generators, In: Monte Carlo and Quasi-Monte Carlo Methods 2000, Springer, 2002, pp. 86–102.
[17] Harald Niederreiter and Igor E. Shparlinski, Dynamical systems generated by rational functions, In: International Symposium on Applied Algebra, Algebraic Algorithms, and Error-Correcting Codes, Springer, 2003, pp. 6–17.
[18] Alina Ostafe, Pseudorandom vector sequences derived from triangular polynomial systems with constant multipliers, In: International Workshop on the Arithmetic of Finite Fields, Lecture Notes in Comput. Sci., Springer, 2010, pp. 62–72.
[19] Alina Ostafe, Elena Pelican, and Igor E. Shparlinski, On pseudorandom numbers from multivariate polynomial systems, *Finite Fields Appl.*, **16**(5) (2010), 320–328.
[20] Alina Ostafe and Igor E. Shparlinski, On the degree growth in some polynomial dynamical systems and nonlinear pseudorandom number generators, *Math. Comput.*, **79**(269) (2010), 501–511.
[21] Alina Ostafe and Igor E. Shparlinski, On the length of critical orbits of stable quadratic polynomials. *Proc. Am. Math. Soc.*, **138**(8) (2010), 2653–2656.
[22] Arnold Schönhage and Volker Strassen, Schnelle Multiplikation großer Zahlen. *Computing*, **7**(3) (1971), 281–292.
[23] Igor Shparlinski, On finding primitive roots in finite fields. *Theor. Comput. Sci.*, **157**(2) (1996), 273–275.
[24] The Sage Developers, SageMath, the Sage Mathematics Software System (Version 7.4), 2016, http://www.sagemath.org.
[25] Alev Topuzoğlu and Arne Winterhof, Pseudorandom sequences, In: Topics in Geometry, Coding Theory and Cryptography, Springer Netherlands, Dordrecht, 2006, pp. 135–166.
[26] Arne Winterhof, Recent results on recursive nonlinear pseudorandom number generators, In: SETA, Springer, 2010, pp. 113–124.

Gary McGuire and Daniela Mueller

Some results on linearized trinomials that split completely

Abstract: Linearized polynomials over finite fields have been the subject of many papers over the last several decades. Recently, there has been a renewed interest in linearized polynomials because of new connections to coding theory and finite geometry. We consider the problem of calculating the rank or nullity of a linearized polynomial $L(x) = \sum_{i=0}^{d} a_i x^{q^i}$ (where $a_i \in \mathbb{F}_{q^n}$) from the coefficients a_i. The rank and nullity of $L(x)$ are the rank and nullity of the associated $\mathbb{F}_q$-linear map $\mathbb{F}_{q^n} \longrightarrow \mathbb{F}_{q^n}$. McGuire and Sheekey [6] defined a $d \times d$ matrix A_L with the property that

$$\text{nullity}(L) = \text{nullity}(A_L - I).$$

We present some consequences of this result for some trinomials that split completely, i. e., trinomials $L(x) = x^{q^d} - bx^q - ax$ that have nullity d. We give a full characterization of these trinomials for $n \leq d^2 - d + 1$.

Keywords: Linearized polynomials, finite field, ECDLP, elliptic curves, cryptography

MSC 2010: 11T55

1 Introduction

Let $\mathbb{F}_{q^n}$ be the finite field with q^n elements, where q is a prime power. Let

$$L(x) = a_0 x + a_1 x^q + a_2 x^{q^2} + \cdots + a_d x^{q^d}$$

be a q-linearized polynomial with coefficients in $\mathbb{F}_{q^n}$. The roots of $L(x)$ that lie in the field $\mathbb{F}_{q^n}$ form an $\mathbb{F}_q$-vector space, which can have dimension anywhere between 0 and d.

The dimension of the space of roots of L that lie in $\mathbb{F}_{q^n}$ is equal to the nullity of L considered as an $\mathbb{F}_q$-linear map from $\mathbb{F}_{q^n}$ to $\mathbb{F}_{q^n}$. McGuire and Sheekey [6] defined a $d \times d$ matrix A_L with the property that

$$\text{nullity}(L) = \text{nullity}(A_L - I_d).$$

The entries of A_L can be computed directly from the coefficients of L.

Acknowledgement: We thank Christophe Petit and John Sheekey for helpful conversations. The research of the second author was supported by a Postgraduate Government of Ireland Scholarship from the Irish Research Council.

Gary McGuire, Daniela Mueller, School of Mathematics and Statistics, University College Dublin, Ireland, e-mail: daniela.mueller@ucdconnect.ie

https://doi.org/10.1515/9783110621730-010

In this paper, we focus on the case of largest possible nullity, i. e., the case that $L(x)$ has all its roots in $\mathbb{F}_{q^n}$. In this case, nullity$(L) = d$, and so $A_L - I_d$ has rank 0 and is therefore the zero matrix. Thus we will be studying when $A_L = I_d$. This case of largest possible nullity was also obtained in [4].

We also restrict to trinomials. When computing the rank or nullity, we may assume without loss of generality that $L(x)$ is monic. We will study polynomials of the form

$$L(x) = x^{q^d} - bx^q - ax \in \mathbb{F}_{q^n}[x]$$

where q is a prime power and $n \geq 1$. We want to find $a, b \in \mathbb{F}_{q^n}$ such that L splits completely over $\mathbb{F}_{q^n}$, i. e., L has q^d roots in $\mathbb{F}_{q^n}$. Thus, the problem becomes finding $a, b \in \mathbb{F}_{q^n}$ such that $A_L = I_d$. We will provide a full characterization of this situation for $n \leq d(d-1)+1$. Our results are summarized and stated in the following theorem.

Theorem 1.1.

1. *If $n \leq (d-1)d$ and d does not divide n, then there is no polynomial $L = x^{q^d} - bx^q - ax$ with $a, b \in \mathbb{F}_{q^n}$ that splits completely over $\mathbb{F}_{q^n}$.*
2. *Let $n = id$ with $i \in \{1, \ldots, d-1\}$. Let $L = x^{q^d} - bx^q - ax \in \mathbb{F}_{q^n}[x]$. Then L has q^d roots in $\mathbb{F}_{q^n}$ if and only if $a^{1+q^d+\cdots+q^{(i-1)d}} = 1$ and $b = 0$.*
3. *Let $n = (d-1)d+1$. Let $L = x^{q^d} - bx^q - ax \in \mathbb{F}_{q^n}[x]$. Then L has q^d roots in $\mathbb{F}_{q^n}$ if and only if all the following hold:*
   - *$N(a) = (-1)^{d-1}$*
   - *$b = -a^{qe_1}$ where $e_1 = \sum_{i=0}^{d-1} q^{id}$*
   - *$d-1$ is a power of the characteristic of $\mathbb{F}_{q^n}$*

   where $N(a) = a^{1+q+\cdots+q^{(d-1)d}} = a^{(q^n-1)/(q-1)}$.

We will prove part 1 in Section 2, part 2 in Section 3 and part 3 in Sections 4 and 5. In Section 6, we present a possible application to elliptic curve cryptography.

Our result generalizes a result of Csajbók et al. [3] which states that $a_0x + a_1x^q + a_3x^{q^3}$ (where $a_i \in \mathbb{F}_{q^7}$) cannot have q^3 roots in $\mathbb{F}_{q^7}$ if q is odd. This is the $d = 3$ case of our theorem. Also, in that paper, the authors give one example of a trinomial that does split completely when $d = 3$, $n = 7$, and q is even. Our theorem characterizes fully the trinomials that split completely and allows us to count their number (for each a of norm $(-1)^q$ there is one polynomial, so there are $\frac{q^n-1}{q-1}$ such trinomials).

One can trivially obtain some results of this type by taking qth powers. For example, when $n = 2d-2$, the trinomial $x^{q^d} - bx^q - ax$ cannot have q^d roots in $\mathbb{F}_{q^n}$, which follows by taking the q^{d-2} power of the trinomial. This is for a particular value of n, whereas our theorem extends this to a larger range of values of n.

One recent application of calculating the rank of linearized polynomials concerns rank metric codes and MRD codes; see [9]. In particular, we would obtain an $\mathbb{F}_{q^n}$-linear MRD code from a space of linearized polynomials of dimension kn over $\mathbb{F}_q$, with the

property that every nonzero element has rank at least $n-k+1$. For example, in the case $k=3$, we would obtain an MRD code from the set of all trinomials $cx^{q^d}-bx^q-ax$ ($a,b,c\in\mathbb{F}_{q^n}$) if all of them have nullity ≤ 2.

Linearized polynomials in the context of coding theory have also been studied recently in [7] (list decoding of Gabidulin codes) and [1] (subspace codes for network coding). Another application of linearized polynomials to affine dispersers is in [2].

Finally, we set the scene for our results. We are seeking $n\geq 1$ and $a,b\in\mathbb{F}_{q^n}$ such that $L=x^{q^d}-bx^q-ax$ splits over $\mathbb{F}_{q^n}$. The companion matrix C_L of $L(x)=x^{q^d}-bx^q-ax$ as defined in [6] is the $d\times d$ matrix

$$\begin{pmatrix} 0 & 0 & \dots & 0 & a \\ 1 & 0 & \dots & 0 & b \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

We define $A_L=A_{L,n}=C_LC_L^q\cdots C_L^{q^{n-1}}$, where C^q means raising every matrix entry to the power of q. As stated above, L splits completely over $\mathbb{F}_{q^n}$ if and only if $A_L=I_d$.

2 Fixed d not dividing n and $n\leq(d-1)d$

In this section, we will prove the first part of Theorem 1.1.

Theorem 2.1. *If $n\leq(d-1)d$ and d does not divide n, then there is no polynomial $L=x^{q^d}-bx^q-ax$ with $a,b\in\mathbb{F}_{q^n}$ that splits completely over $\mathbb{F}_{q^n}$.*

Proof. We will write A_n instead of $A_{L,n}$ as L is fixed throughout the proof.

If $n=1$, then $A_1=C_L\neq I_d$. Indeed, if $n\leq d-1$, then the $(1,1)$ entry of A_n is 0, so $A_n\neq I_d$.

Note that $A_n=A_{n-1}C_L^{q^{n-1}}$. But the first column of $C_L^{q^{n-1}}$ is $(010\dots0)^T$. Thus, the $(1,1)$ entry of A_n is the $(1,2)$ entry of A_{n-1}. If $n\geq d$ then the $(1,1)$ entry of A_n is also the $(1,d)$ entry of A_{n-d+1}.

Let M_k denote the $(1,d)$ entry of A_k. Then $M_1=a$, and $M_k=0$ for $k=2,\dots,d-1$, since $M_k=(0\dots0M_1\dots M_{k-1})\cdot\left(a^{q^{k-1}}b^{q^{k-1}}0\dots0\right)^T$.

Set $M_0=0$. Then for $k\geq d$, we have a recursive formula, which follows directly from matrix multiplication:

$$M_k=M_{k-d}a^{q^{k-1}}+M_{k-d+1}b^{q^{k-1}}. \tag{2.1}$$

Claim. *$M_j=0$ for $j=id+2,\dots,(i+1)d-(i+1)$ and $i=0,\dots,d-3$.*

Proof of Claim. We prove the claim by induction on i. The base case $i = 0$ was done above. Note that if $M_{k-d} = M_{k-d+1} = 0$ then $M_k = 0$. So if the claim is true for i, then we have $M_{id+2+d} = 0, \dots, M_{(i+1)d-(i+1)+d-1} = 0$, i. e., the claim is true for $i+1$. This completes the proof of the claim.

Note that when $i = d-3$, then $id+2 = (i+1)d-(i+1)$, so the claim is not true for $i = d-2$.

For the remaining n not divisible by d, we will show that the $(1,1)$ entry of A_n cannot be 1 if the $(1,j)$ entry is 0 for some $j \in \{2,\dots,d\}$, and thus A_n cannot be the identity matrix. Note that the $(1,j)$ entry of A_n is M_{n-d+j}.

For $i = 1,\dots,d-2$, we have $M_{(i-1)d+2} = 0$, and thus

$$M_{id+1} = M_{(i-1)d+1}a^{q^{id}}.$$

Since $M_1 = a$, we have $M_{id+1} = a^{1+q^d+\cdots+q^{id}}$. But M_{id+1} is the $(1,(i+1)d+1-n)$ entry of A_n for $n = id+1,\dots,(i+1)d-1$. If $A_n = I_d$ then the $(1,(i+1)d+1-n)$ entry must be 0, so we must have $M_{id+1} = a^{1+q^d+\cdots+q^{id}} = 0$, and thus $a = 0$.

Recall that the $(1,1)$ entry of A_n is M_{n-d+1}. But M_{n-d+1} must be either 0 or a multiple of a, since all initial values $M_0,\dots,M_{d-1}$ of the recursive formula are either a or 0. Therefore, if $a = 0$, we have $M_{n-d+1} = 0$ and so $A_n \neq I_d$. □

Remark 2.1. The proof is not valid when d divides n. If $n = id$ with $i \in \{1,\dots,d-1\}$, the $(1,1)$ entry of A_{id} is $M_{(i-1)d+1} = a^{1+q^d+\cdots+q^{(i-1)d}}$ for $i \geq 1$, and so we have the equation $a^{1+q^d+\cdots+q^{(i-1)d}} = 1$ and cannot deduce that $a = 0$.

The recursive formula (2.1) established in the proof of Theorem 2.1 is valid in greater generality: Set $M_{l,l-d} = 1$, and $M_{l,k} = 0$ for $k \leq 0$ and $k \neq l-d$. For $1 \leq l \leq d$ and $k \geq 1$, let

$$M_{l,k} = M_{l,k-d}a^{q^{k-1}} + M_{l,k-d+1}b^{q^{k-1}}. \tag{2.2}$$

Then $M_{l,k}$ is the (l,d) entry of $A_{L,k}$. Furthermore, the (l,j) entry of $A_{L,k}$ is $M_{l,k-d+j}$.

3 Fixed d dividing n and $n \leq (d-1)d$

In the case that d divides n, we have a solution, namely $a = 1$ and $b = 0$, i. e., the polynomial $x^{q^d} - x$ splits completely because $\mathbb{F}_{q^n}$ has a subfield $\mathbb{F}_{q^d}$. We now characterize exactly which polynomials split completely.

Theorem 3.1. *Let $n = id$ with $i \in \{1,\dots,d-1\}$. Let $L = x^{q^d} - bx^q - ax \in \mathbb{F}_{q^n}[x]$. Then L has q^d roots in $\mathbb{F}_{q^n}$ if and only if $a^{1+q^d+\cdots+q^{(i-1)d}} = 1$ and $b = 0$.*

Proof. By Remark 2.1, if L splits completely, we have $a^{1+q^d+\cdots+q^{(i-1)d}} = 1$. Now the $(1,d+1-i)$ entry of A_{id} is $M_{1,i(d-1)+1}$. For $i = 1$, this is $M_{1,d} = ab^{q^{d-1}}$. For $i \geq 2$, we have

$M_{1,i(d-1)+1} = M_{1,(i-1)d-(i-1)}a^{q^{i(d-1)}} + M_{1,(i-1)(d-1)+1}b^{q^{i(d-1)}}$. But by the claim in the proof of Theorem 2.1, we have $M_{1,(i-1)d-(i-1)} = 0$ for $i = 2, \dots, d-1$. Thus

$$\begin{aligned} M_{1,i(d-1)+1} &= M_{1,(i-1)(d-1)+1}b^{q^{i(d-1)}} \\ &= M_{1,(i-2)(d-1)+1}b^{q^{(i-1)(d-1)}+q^{i(d-1)}} \\ &= \cdots \\ &= ab^{q^{d-1}+q^{2(d-1)}+\cdots+q^{i(d-1)}}. \end{aligned}$$

But if $A_{id} = I_d$, then $M_{1,i(d-1)+1} = 0$, and since $a \neq 0$, we must have $b = 0$.

The converse is a well-known result, however we will exhibit a new proof using a result of [4]. Assume that $a^{1+q^d+\cdots+q^{(i-1)d}} = 1$ and $b = 0$. Then the $(l,1)$ entry of A_{id} is

$$\begin{aligned} M_{l,(i-1)d+1} &= M_{l,(i-2)d+1}a^{q^{(i-1)d}} \\ &= \cdots \\ &= M_{l,1-d}a^{1+q^d+\cdots+q^{(i-1)d}} \\ &= M_{l,1-d} \\ &= \begin{cases} 1 & \text{for } l = 1, \\ 0 & \text{for } l = 2, \dots, d. \end{cases} \end{aligned}$$

By [4, Corollary 3.2], this implies that $A_{id} = I_d$. □

4 Fixed d and $n = (d-1)d + 1$

In this section, we will prove some preliminary results which are part of the proof of Theorem 1.1 part 3.

4.1 Assuming L splits completely

If $n = (d-1)d + 1$, then, the $(1,j)$ entry of $A_{L,n}$ is $M_{1,(d-2)d+j+1}$ (where $j = 1, \dots, d$). So to get $A_{L,n} = I_d$, the following system of equations has to be satisfied for $l = 1, \dots, d$

$$\begin{cases} M_{l,(d-2)d+l+1} = 1 \\ M_{l,(d-2)d+j+1} = 0 \quad \text{for } j = 1, \dots, l-1, l+1, \dots, d \end{cases} \tag{4.1}$$

Lemma 4.1. *$M_{1,(d-2)d+2} = ab^{e_2}$ where $e_2 = \frac{q^{(d-1)d}-q^{d-1}}{q^{d-1}-1}$.*

Proof. By the recursive formula (2.2), $M_{1,(d-2)d+2} = M_{1,(d-3)d+2}a^{q^{(d-2)d+1}} + M_{1,(d-3)d+3} \times b^{q^{(d-2)d+1}}$. But it follows from the claim in the proof of Theorem 2.1 that $M_{1,(d-j-1)d+j} = 0$

for $j = 2, \ldots, d-1$. Thus

$$\begin{aligned} M_{1,(d-2)d+2} &= M_{1,(d-3)d+3} b^{q^{(d-2)d+1}} = M_{1,(d-4)d+4} b^{q^{(d-2)d+1}+q^{(d-3)d+2}} = \cdots \\ &= M_{1,d} b^{q^{(d-2)d+1}+q^{(d-3)d+2}+\cdots+q^{2d-2}} \\ &= ab^{q^{(d-2)d+1}+q^{(d-3)d+2}+\cdots+q^{2d-2}+q^{d-1}} \\ &= ab^{\sum_{i=1}^{d-1} q^{i(d-1)}} \\ &= ab^{e_2}. \end{aligned}$$

□

Lemma 4.2. *$M_{1,(d-1)d+1} = a^{e_1} + ab^{e_2+q^{(d-1)d}}$ where $e_1 = \frac{q^{d^2}-1}{q^d-1}$ and $e_2 = \frac{q^{(d-1)d}-q^{d-1}}{q^{d-1}-1}$.*

Proof. By the recursive formula (2.2),

$$M_{1,(d-1)d+1} = M_{1,(d-2)d+1} a^{q^{(d-1)d}} + M_{1,(d-2)d+2} b^{q^{(d-1)d}}.$$

By Lemma 4.1, $M_{1,(d-2)d+2} = ab^{e_2}$. Also $M_{1,(d-2)d+1} = a^{1+q^d+\cdots+q^{(d-2)d}}$ as established in the proof of Theorem 2.1.

Therefore,

$$\begin{aligned} M_{1,(d-1)d+1} &= a^{\sum_{i=0}^{d-1} q^{id}} + ab^{e_2+q^{(d-1)d}} \\ &= a^{e_1} + ab^{e_2+q^{(d-1)d}}. \end{aligned}$$

□

Theorem 4.1. *Let $n = (d-1)d + 1$. Let $L = x^{q^d} - bx^q - ax \in \mathbb{F}_{q^n}[x]$. If L has q^d roots in $\mathbb{F}_{q^n}$, then:*

1. $a^{1+q+\cdots+q^{(d-1)d}} = (-1)^{d-1}$ *and*
2. $a^{1+qe_1e_2} = (-1)^{d-1}$ *and*
3. $b = -a^{qe_1}$;

where $e_1 = \frac{q^{d^2}-1}{q^d-1}$ and $e_2 = \frac{q^{(d-1)d}-q^{d-1}}{q^{d-1}-1}$.

Proof. If $A_{L,n} = I_d$, then system (4.1) has to be satisfied. By Lemma 4.1, we have $ab^{e_2} = 1$ (the $(1,1)$ entry of $A_{L,n}$), and by Lemma 4.2, we have $a^{e_1} + ab^{e_2+q^{(d-1)d}} = 0$ (the $(1,d)$ entry of $A_{L,n}$). But if $ab^{e_2} = 1$, then $a^{e_1} + ab^{e_2+q^{(d-1)d}} = a^{e_1} + b^{q^{(d-1)d}}$, and thus we have $b^{q^{(d-1)d}} = -a^{e_1}$. Raising both sides to the power of q gives us $b^{q^n} = (-1)^q a^{qe_1}$. Since q is a prime power, $(-1)^q = -1$ in $\mathbb{F}_{q^n}$. Thus, $b = -a^{qe_1}$ which proves the third conclusion.

Lemma 4.1 says $ab^{e_2} = 1$ which now implies

$$a^{-1} = b^{e_2} = \left(-a^{qe_1}\right)^{e_2} = (-1)^{e_2} a^{qe_1e_2},$$

and so $a^{1+qe_1e_2} = (-1)^{e_2}$. (Note that $a \neq 0$ since $ab^{e_2} = 1$.)

Recall that $e_2 = \sum_{i=1}^{d-1} q^{i(d-1)}$. So if q is even, then e_2 is even. If q is odd, then $q^{i(d-1)}$ is odd for all $i = 1, \dots, d-1$. So if $d-1$ is even, then e_2 is an even sum of odd numbers, and thus even, and if $d-1$ is odd, then e_2 is an odd sum of odd numbers and thus odd. Thus, $(-1)^{e_2} = (-1)^{q(d-1)}$. Since $(-1)^q = -1$ in $\mathbb{F}_{q^n}$ we have $(-1)^{e_2} = (-1)^{d-1}$.

By [6, Corollary 1], if L splits, then $N(-a) = (-1)^{nd}N(1)$, where N is the norm function over $\mathbb{F}_{q^n}$. So we have the additional condition $N(a) = (-1)^{n(d-1)}$ or $a^{\frac{q^n-1}{q-1}} = (-1)^{n(d-1)}$. But $n = (d-1)d+1$, so n is always odd. Consequently, $(-1)^{n(d-1)} = (-1)^{d-1}$.

Hence, a satisfies the equations

$$\begin{cases} a^{1+qe_1e_2} = (-1)^{d-1} \\ a^{\frac{q^n-1}{q-1}} = (-1)^{d-1}. \end{cases} \tag{4.2}$$

□

In the next section, we will show that conclusion 1 of this theorem actually implies conclusion 2.

4.2 GCD of $x^k \pm 1$ and $x^l \pm 1$

The GCD of $x^k - 1$ and $x^l - 1$ is well known to be $x^{\gcd(k,l)} - 1$, but we are interested in the GCD of $x^k + 1$ and $x^l + 1$. The following is surely well known, but we include a proof.

Theorem 4.2. *The GCD of $x^k + 1$ and $x^l + 1$ is $x^{\gcd(k,l)} + 1$ if $\frac{k}{\gcd(k,l)}$ and $\frac{l}{\gcd(k,l)}$ are both odd, and 1 otherwise.*

Proof. Let $d = \gcd(k,l)$ and let s, t be Bézout coefficients for k and l, i. e., $sk + tl = d$. Let $g = \gcd(x^k+1, x^l+1)$. Then $x^k \equiv -1 \bmod g$ and $x^l \equiv -1 \bmod g$. Thus $x^{sk+tl} \equiv (-1)^{s+t} \bmod g$. So g divides $x^{sk+tl} - (-1)^{s+t} = x^d - (-1)^{s+t}$. We need to check if $x^d - (-1)^{s+t}$ divides $x^k + 1$ and $x^l + 1$.

Let $e = \frac{k}{d}$ and $f = \frac{l}{d}$. Then $x^k + 1 = x^{ed} + 1 = ((-1)^{s+t})^e + 1 \bmod x^d - (-1)^{s+t}$ and similarly, $x^l + 1 = ((-1)^{s+t})^f + 1 \bmod x^d - (-1)^{s+t}$. So we need to have $(-1)^{(s+t)e} + 1 = 0$ and $(-1)^{(s+t)f} + 1 = 0$, i. e. $e, f, s+t$ all need to be odd. But $sk + tl = d$ implies $se + tf = 1$, so e, f odd implies $s + t$ odd. Thus if e, f are odd, then $g = x^d - (-1)^{s+t} = x^d + 1$. □

Remark 4.1. Similarly, one can show that $\gcd(x^k - 1, x^l + 1) = x^{\gcd(k,l)} + 1$ if $\frac{k}{\gcd(k,l)}$ is even and $\frac{l}{\gcd(k,l)}$ is odd.

Lemma 4.3. *Let $n = (d-1)d + 1$ and let $e_1 = \frac{q^{d^2}-1}{q^d-1}$ and $e_2 = \frac{q^{(d-1)d}-q^{d-1}}{q^{d-1}-1}$. Then*

$$\gcd\left(1 + qe_1e_2, \frac{q^n - 1}{q-1}\right) = \frac{q^n - 1}{q-1}.$$

Proof. We first show that $\frac{q^n-1}{q-1} = 1 + q(\frac{q^{d^2}-1}{q^d-1})(\frac{q^{(d-1)d}-q^{d-1}}{q^{d-1}-1}) \bmod q^n - 1$. Recall that $\frac{q^n-1}{q-1} = \sum_{i=0}^{n-1} q^i = 1 + q + q^2 + \cdots + q^{n-1}$ and $n = d^2 - d + 1$. Then

$$1 + q\left(\frac{q^{d^2}-1}{q^d-1}\right)\left(\frac{q^{(d-1)d}-q^{d-1}}{q^{d-1}-1}\right) = 1 + q\left(\sum_{i=0}^{d-1} q^{id}\right)\left(\sum_{j=1}^{d-1} q^{j(d-1)}\right)$$
$$= 1 + \sum_{i=0}^{d-1}\sum_{j=1}^{d-1} q^{id+j(d-1)+1}.$$

We claim that $id + j(d-1) + 1 \bmod n$ with $i = 0, \ldots, d-1$ and $j = 1, \ldots, d-1$ gives us exactly the numbers $\{1, \ldots, n-1\}$. Assuming the truth of this claim, $1 + q(\frac{q^{d^2}-1}{q^d-1})(\frac{q^{(d-1)d}-q^{d-1}}{q^{d-1}-1}) \bmod q^n - 1 = 1 + q + q^2 + \cdots + q^{n-1} = \frac{q^n-1}{q-1}$. Since $\frac{q^n-1}{q-1}$ divides $q^n - 1$, $\frac{q^n-1}{q-1}$ divides $1 + q(\frac{q^{d^2}-1}{q^d-1})(\frac{q^{(d-1)d}-q^{d-1}}{q^{d-1}-1})$, and thus $\gcd(1 + q(\frac{q^{d^2}-1}{q^d-1})(\frac{q^{(d-1)d}-q^{d-1}}{q^{d-1}-1}), \frac{q^n-1}{q-1}) = \frac{q^n-1}{q-1}$ and the result is proved.

It remains to prove the claim. To see this, we will show that the sets

$$\{(i+j)d - (j-1) \mid i = 0, \ldots, d-1;\ j = 1, \ldots, d-1\}$$

and

$$\{kd - m \mid m = 0, \ldots, d-2;\ k = m+1, \ldots, m+d\}$$

are equal, and it is easy to see that all values in the second set are distinct.

Fixing j and varying $i = 0, \ldots, d-1$ gives us the numbers

$$jd - (j-1), (j+1)d - (j-1), \ldots, (j+d-1)d - (j-1).$$

When $i + j \le d - 1$, then $(i+j)d - (j-1) \le n$ and all these numbers are of the form $kd - m$ where $m \in \{0, \ldots, d-2\}$ and $m < k \le d - 1$.

When $i + j \ge d$, then $(i+j)d - (j-1) > n$ and we subtract n to get $(i + j - d + 1)d - j$. Now $i + j - d + 1 \le j$ since $i \le d - 1$, and thus $(i + j - d + 1)d - j$ is not of the above form $kd - m$ with $m < k \le d - 1$. □

Corollary 4.1. *Let $n = (d-1)d + 1$ and let $e_1 = \frac{q^{d^2}-1}{q^d-1}$ and $e_2 = \frac{q^{(d-1)d}-q^{d-1}}{q^{d-1}-1}$. Then*

$$\gcd\left(x^{1+qe_1e_2} + (-1)^d, x^{\frac{q^n-1}{q-1}} + (-1)^d\right) = x^{\frac{q^n-1}{q-1}} + (-1)^d.$$

Proof. If q is even, then both $1 + qe_1e_2$ and $\frac{q^n-1}{q-1}$ are odd. Recall that $\frac{q^n-1}{q-1} = \sum_{i=0}^{n-1} q^i$. So if q is odd, then $\frac{q^n-1}{q-1}$ is odd if n is odd, and even if n is even. But $n = (d-1)d + 1$ is always odd, so $\frac{q^n-1}{q-1}$ is odd. We have already established in the proof of Theorem 4.1 that if q is odd, then e_2 is odd if $d - 1$ is odd, and even if $d - 1$ is even. Now $e_1 = \sum_{i=0}^{d-1} q^{id}$ is odd if

q and d are odd and even if q is odd but d is even. But either d or $d-1$ is always even, so e_1e_2 is even. Thus $1+qe_1e_2$ is odd. Consequently, by Theorem 4.2

$$\gcd(x^{1+qe_1e_2}+(-1)^d, x^{\frac{q^n-1}{q-1}}+(-1)^d) = x^{\gcd(1+qe_1e_2, \frac{q^n-1}{q-1})}+(-1)^d$$

for any q, d. □

Corollary 4.2. *In the conclusions of Theorem* 4.1*, conclusion 1 implies conclusion 2.*

5 The main result

In this section, we will prove the third part of the main theorem as stated in the introduction. The following lemma is surely well known but we include a short proof.

Lemma 5.1. $\binom{n}{i} = 0 \bmod p$ *for all* $i = 1, 2, \ldots, n-1$ *if and only if* n *is a power of* p.

Proof. If n is a power of p, then the above binomial coefficients are divisible by p. On the other hand, we claim that if $n = p^k w$, where $p \nmid w$, $w > 1$ and $k \geq 0$, then $\binom{p^k w}{p^k}$ is not divisible by p. First, note that $\binom{n}{m} = \frac{n!}{m!(n-m)!} = \frac{\prod_{i=1}^{n} i}{(\prod_{i=1}^{m} i)(\prod_{i=1}^{n-m} i)} = \frac{\prod_{i=n-m+1}^{n} i}{\prod_{i=1}^{m} i} = \frac{\prod_{i=0}^{m-1}(n-i)}{\prod_{i=1}^{m} i} = \frac{n}{m}\prod_{i=1}^{m-1}\frac{n-i}{i}$. Thus $\binom{p^k w}{p^k} = w\prod_{i=1}^{p^k-1}\frac{p^k w-i}{i}$. Now write $i = lp^j$ with $p \nmid l$. Then $\frac{p^k w-i}{i} = \frac{p^k w-lp^j}{lp^j} = \frac{(p^{k-j}w-l)p^j}{lp^j} = \frac{p^{k-j}w-l}{l}$ which is not divisible by p. □

Finally, we present the last part of the proof of the main theorem.

Theorem 5.1. *Let* $n = (d-1)d+1$ *and* $e_1 = \frac{q^{d^2}-1}{q^d-1}$. *Let* $L = x^{q^d} - bx^q - ax \in \mathbb{F}_{q^n}[x]$. *Then* L *has* q^d *roots in* $\mathbb{F}_{q^n}$ *if and only if each of the following holds:*

1. $a^{1+q+\cdots+q^{(d-1)d}} = (-1)^{d-1}$
2. $b = -a^{qe_1}$
3. $d-1$ *is a power of the characteristic of* $\mathbb{F}_{q^n}$.

Proof. Recall that the $(l,1)$ entry of $A_{L,n}$ is $M_{l,n-d+1}$. We will first show that

$$M_{l,n-d+1} = \begin{cases} 1 & \text{for } l = 1, \\ 0 & \text{for } l = 2, \ldots, d \end{cases}$$

whenever the three conditions of the theorem are fulfilled. By [4, Corollary 3.2], this implies that $A_{L,n} = I_d$.

Let $k \geq d+1$. By the recursion (2.2),

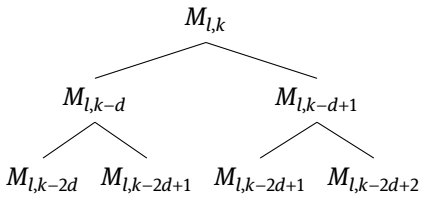
$$\begin{aligned} M_{l,k} &= M_{l,k-d}a^{q^{k-1}} + M_{l,k-d+1}b^{q^{k-1}} \\ &= (M_{l,k-2d}a^{q^{k-d-1}} + M_{l,k-2d+1}b^{q^{k-d-1}})a^{q^{k-1}} \end{aligned}$$

$$\begin{aligned}
&+ (M_{l,k-2d+1}a^{q^{k-d}} + M_{l,k-2d+2}b^{q^{k-d}})b^{q^{k-1}} \\
&= M_{l,k-2d}a^{q^{k-d-1}+q^{k-1}} + M_{l,k-2d+1}(a^{q^{k-1}}b^{q^{k-d-1}} + a^{q^{k-d}}b^{q^{k-1}}) \\
&\quad + M_{l,k-2d+2}b^{q^{k-d}+q^{k-1}}. \qquad (5.1)
\end{aligned}$$

Since $b = -a^{qe_1} = -a^{1+\sum_{i=0}^{d-2} q^{id+1}} \bmod a^{q^n} - a$ (condition 2 in the statement of the theorem), we have

$$\begin{aligned}
a^{q^{k-1}}b^{q^{k-d-1}} &= -a^{q^{k-1}+q^{k-d-1}(1+\sum_{i=0}^{d-2} q^{id+1})} \\
&= -a^{q^{k-1}+q^{k-d}+q^{k-d-1}+\sum_{i=1}^{d-2} q^{k+(i-1)d}} \\
&= -a^{q^{k-1}+q^{k-d}+q^{k-d-1}+\sum_{i=0}^{d-3} q^{k+id}} \\
&= -a^{q^{k-1}+q^{k-d}+q^{k+(d-2)d}+\sum_{i=0}^{d-3} q^{k+id}} \bmod a^{q^n} - a \\
&= -a^{q^{k-d}+q^{k-1}(1+\sum_{i=0}^{d-2} q^{id+1})} \\
&= a^{q^{k-d}}b^{q^{k-1}}.
\end{aligned}$$

So the coefficient of $M_{l,k-2d+1}$ in (5.1) that comes from expanding $M_{l,k-d}$ is the same as the coefficient that comes from expanding $M_{l,k-d+1}$.



Let $c_{2,0} = a^{q^{k-1}+q^{k-d-1}}$, $c_{2,1} = a^{q^{k-d}}b^{q^{k-1}}$, and $c_{2,2} = b^{q^{k-1}+q^{k-d}}$. Thus (5.1) is saying that

$$M_{l,k} = c_{2,0}M_{l,k-2d} + 2c_{2,1}M_{l,k-2d+1} + c_{2,2}M_{l,k-2d+2}.$$

One can see Pascal's triangle emerging. We claim that

$$M_{l,k} = \sum_{i=0}^{j} \binom{j}{i} c_{j,i}M_{l,k-jd+i}$$

for all $j = 0, \ldots, \lfloor \frac{k-1}{d} + 1 \rfloor$, where $c_{j,i}$ are expressions in a and b, determined by the following recursion:

$$c_{j,i} = \begin{cases} 1 & \text{for } j = i = 0, \\ c_{j-1,0}a^{q^{k-(j-1)d-1}} & \text{for } i = 0, \\ c_{j-1,i}a^{q^{k-(j-1)d+i-1}} = c_{j-1,i-1}b^{q^{k-(j-1)d+i-2}} & \text{for } 0 < i < j, \\ c_{j-1,j-1}b^{q^{k-(j-1)d+j-2}} & \text{for } i = j. \end{cases}$$

We have shown the statement for $j = 2$. Assume that the statement is true for any index less than j. Then

$$\begin{aligned}
M_{l,k} &= \sum_{i=0}^{j-1} \binom{j-1}{i} c_{j-1,i} M_{l,k-(j-1)d+i} \\
&= \sum_{i=0}^{j-1} \binom{j-1}{i} c_{j-1,i} (M_{l,k-jd+i} a^{q^{k-(j-1)d+i-1}} + M_{l,k-jd+i+1} b^{q^{k-(j-1)d+i-1}}) \\
&= \binom{j-1}{0} c_{j-1,0} a^{q^{k-(j-1)d-1}} M_{l,k-jd} + \binom{j-1}{j-1} c_{j-1,j-1} b^{q^{k-(j-1)d+j-2}} M_{l,k-jd+j} \\
&\quad + \sum_{i=1}^{j-1} M_{l,k-jd+i} \left(\binom{j-1}{i-1} c_{j-1,i-1} b^{q^{k-(j-1)d+i-2}} + \binom{j-1}{i} c_{j-1,i} a^{q^{k-(j-1)d+i-1}} \right).
\end{aligned}$$

Let $m = k - (j-2)d + i - 1$. Then $a^{q^{m-1}} b^{q^{m-d-1}} = a^{q^{m-d}} b^{q^{m-1}}$, i. e.,

$$a^{q^{k-(j-2)d+i-2}} b^{q^{k-(j-1)d+i-2}} = a^{q^{k-(j-1)d+i-1}} b^{q^{k-(j-2)d+i-2}},$$

and hence

$$c_{j-2,i-1} a^{q^{k-(j-2)d+i-2}} b^{q^{k-(j-1)d+i-2}} = c_{j-2,i-1} a^{q^{k-(j-1)d+i-1}} b^{q^{k-(j-2)d+i-2}}. \tag{5.2}$$

Then

$$c_{j-1,i-1} b^{q^{k-(j-1)d+i-2}} = c_{j-2,i-1} a^{q^{k-(j-2)d+i-2}} b^{q^{k-(j-1)d+i-2}}$$

and

$$c_{j-1,i} a^{q^{k-(j-1)d+i-1}} = c_{j-2,i-1} b^{q^{k-(j-2)d+i-2}} a^{q^{k-(j-1)d+i-1}}$$

and by (5.2), these two expressions are equal.

Thus

$$\begin{aligned}
M_{l,k} &= \binom{j}{0} c_{j-1,0} a^{q^{k-(j-1)d-1}} M_{l,k-jd} + \binom{j}{j} c_{j-1,j-1} b^{q^{k-(j-1)d+j-2}} M_{l,k-jd+j} \\
&\quad + \sum_{i=1}^{j-1} \binom{j}{i} c_{j-1,i} a^{q^{k-(j-1)d+i-1}} M_{l,k-jd+i}
\end{aligned}$$

as desired. This completes the proof of the claim.

We now have

$$\begin{aligned}
M_{l,n-d+1} = M_{l,(d-2)d+2} &= \sum_{i=0}^{d-1} \binom{d-1}{i} c_{d-1,i} M_{l,2-d+i} \\
&= \sum_{i=0}^{d-2} \binom{d-1}{i} c_{d-1,i} M_{l,2-d+i} + c_{d-1,d-1} M_{l,1}
\end{aligned}$$

$$= \begin{cases} c_{d-1,d-1}M_{l,1} & \text{for } l = 1, \\ \binom{d-1}{l-2}c_{d-1,l-2}M_{l,l-d} + c_{d-1,d-1}M_{l,1} & \text{for } l \geq 2 \end{cases}$$

since $M_{l,2-d+i} = 0$ when $i \neq l-2$ (and $i \leq d-2$).

As before, let $e_1 = \frac{q^{d^2}-1}{q^d-1}$ and $e_2 = \frac{q^{(d-1)d}-q^{d-1}}{q^{d-1}-1}$.

Then

$$\begin{aligned} c_{d-1,d-1} &= b^{q^{k-1}+q^{k-d}+q^{k-2d+1}+\cdots+q^{k-(d-2)d+d-3}} \\ &= b^{q^{(d-2)d+1}+q^{(d-3)d+2}+\cdots+q^{d-1}} \quad \text{for } k = (d-2)d+2 \\ &= b^{q^{d-1}+q^{2d-2}+\cdots+q^{(d-1)(d-1)}} \\ &= b^{e_2} \\ &= \left(-a^{qe_1}\right)^{e_2} \\ &= (-1)^{d-1}a^{qe_1e_2} \\ &= (-1)^{d-1}a^{\frac{q^n-1}{q-1}-1} \bmod a^{q^n}-a \quad \text{by Lemma 4.3.} \end{aligned}$$

Also

$$\begin{aligned} c_{d-1,0} &= a^{q^{k-1}+q^{k-d-1}+\cdots+q^{k-(d-2)d-1}} \\ &= a^{q^{(d-2)d+1}+q^{(d-3)d+1}+\cdots+q} \quad \text{for } k = (d-2)d+2. \end{aligned}$$

Thus

$$\begin{aligned} c_{d-1,d-1}M_{l,1} &= (-1)^{d-1}a^{\frac{q^n-1}{q-1}-1}(aM_{l,1-d} + bM_{l,2-d}) \\ &= (-1)^{d-1}a^{\frac{q^n-1}{q-1}}M_{l,1-d} + (-1)^d a^{\frac{q^n-1}{q-1}+\sum_{i=0}^{d-2}q^{id+1}}M_{l,2-d} \\ &= (-1)^{d-1}(-1)^{d-1}M_{l,1-d} + (-1)^d(-1)^{d-1}c_{d-1,0}M_{l,2-d} \\ &= M_{l,1-d} - c_{d-1,0}M_{l,2-d} \end{aligned}$$

since $a^{\frac{q^n-1}{q-1}} = (-1)^{d-1}$ (condition 1 in the statement of the theorem).

Hence,

$$\begin{aligned} M_{l,n-d+1} &= \begin{cases} M_{l,1-d} - c_{d-1,0}M_{l,2-d} & \text{for } l = 1, \\ \binom{d-1}{l-2}c_{d-1,l-2}M_{l,l-d} + M_{l,1-d} - c_{d-1,0}M_{l,2-d} & \text{for } l \geq 2 \end{cases} \\ &= \begin{cases} 1 & \text{for } l = 1, \\ 0 & \text{for } l = 2, \\ \binom{d-1}{l-2}c_{d-1,l-2} & \text{for } l \geq 3 \end{cases} \end{aligned} \tag{5.3}$$

since $M_{l,l-d} = 1$ and $M_{l,k} = 0$ when $k \neq l-d$ and $k \leq 0$.

So far we have only used conditions 1 and 2 in the statement of the theorem (so note for later that conditions 1 and 2 imply (5.3)). Assume now that condition 3 holds. By Lemma 5.1, $M_{l,n-d+1} = 0$ for all $l \geq 3$ because $\binom{d-1}{l-2} = 0$. This completes the proof that if the three conditions in the statement hold, then L splits completely.

Now we complete the proof of the theorem by showing the converse, i. e., we show that if L splits completely then the three conditions in the statement hold. Theorem 4.1 and Corollary 4.2 show that if L splits completely, then conditions 1 and 2 of the theorem hold. Because conditions 1 and 2 hold, we know that (5.3) holds.

On the other hand, since L splits completely, $M_{l,n-d+1} = 0$ for all $l \geq 3$. Therefore $\binom{d-1}{l-2} c_{d-1,l-2} = 0$ for all $3 \leq l \leq d$. We now use the fact that $c_{d-1,l-2}$ is a power of a, and is therefore nonzero because a is nonzero. We are forced to conclude that $\binom{d-1}{l-2} = 0$ for all $3 \leq l \leq d$. This implies that $d-1$ is a power of the characteristic of $\mathbb{F}_{q^n}$ by Lemma 5.1. □

6 Possible application to cryptography

6.1 Quasi-subfield polynomials

The recent work [5] explored the use of quasi-subfield polynomials to solve the Elliptic Curve Discrete Logarithm Problem (ECDLP). They define quasi-subfield polynomials as polynomials of the form $x^{q^d} - \lambda(x) \in \mathbb{F}_{q^n}[x]$ which divide $x^{q^n} - x$ and where $\log_q(\deg(\lambda)) < d^2/n$. For appropriate choices of n and d, linearized polynomials have a chance of being quasi-subfield polynomials. We first observe that the polynomials in Theorem 5.1 are quasi-subfield polynomials.

Lemma 6.1. *The linearized polynomial $L = x^{q^d} - bx^q - ax \in \mathbb{F}_{q^{(d-1)d+1}}[x]$ is a quasi-subfield polynomial when all the following conditions are satisfied:*

1. $a^{1+q+\cdots+q^{(d-1)d}} = (-1)^{d-1}$
2. $b = -a^{qe_1}$
3. $d-1$ *is a power of the characteristic of* $\mathbb{F}_{q^n}$.

Proof. Here, $\log_q(\deg(\lambda)) = 1$ and $d^2 > n = (d-1)d + 1$ so the condition $\log_q(\deg(\lambda)) < d^2/n$ is satisfied. By Theorem 5.1, $L(x)$ divides $x^{q^n} - x$. □

6.2 The ECDLP

Let E be an elliptic curve over a finite field $\mathbb{F}_q$, where q is a prime power. In practice, q is often a prime number or a large power of 2. Let P and Q be $\mathbb{F}_q$-rational points on E. The Elliptic Curve Discrete Logarithm Problem (ECDLP) is finding an integer l (if it exists) such that $Q = lP$. The integer l is called the discrete logarithm of Q to base P.

The ECDLP is a hard problem that underlies many cryptographic schemes and is thus an area of active research. The introduction of summation polynomials by [8] has led to algorithms that resemble the index calculus algorithm of the DLP over finite fields.

The algorithm to solve the ECDLP in [5] also uses summation polynomials, so we recall their definition.

Definition 6.1. [8] Let E be an elliptic curve over a field K. For $m \geq 1$, we define the summation polynomial $S_{m+1} = S_{m+1}(X_0, X_1, \ldots, X_m) \in K[X_0, X_1, \ldots, X_m]$ of E by the following property. Let $x_0, x_1, \ldots, x_m \in \overline{K}$, then $S_{m+1}(x_0, x_1, \ldots, x_m) = 0$ if and only if there exist $y_0, y_1, \ldots, y_m \in \overline{K}$ such that $(x_i, y_i) \in E(\overline{K})$ and $(x_0, y_0) + (x_1, y_1) + \cdots + (x_m, y_m) = \mathcal{O}$, where $\mathcal{O}$ is the identity element of E.

The summation polynomials S_m have many terms and have only been computed for $m \leq 9$.

[5] develop an algorithm to solve the ECDLP over the field $\mathbb{F}_{q^n}$ using a quasi-subfield polynomial $X^{q^d} - \lambda(X) \in \mathbb{F}_{q^n}[X]$ and the summation polynomial $S_{m+1}(X_0, X_1, \ldots, X_m) \in \mathbb{F}_{q^n}[X_0, X_1, \ldots, X_m]$. By [5, Theorem 3.2] (see also Appendix A1), their algorithm has complexity

$$m! q^{n-d(m-1)} \tilde{O}(m^{5.188} 2^{7.376m(m-1)} \deg(\lambda)^{4.876m(m-1)}) + mq^{2d}.$$

6.3 Linearized quasi-subfield polynomials

One of the problems outlined in [5] is to find suitable quasi-subfield polynomials that give optimal complexity in their algorithm. So in this section, we will investigate whether the linearized polynomials in this paper are a suitable choice (they are not).

In our notation, the field is $\mathbb{F}_{q^n}$ so brute force algorithms have $O(q^n)$ complexity and generic algorithms (Pollard Rho or Baby-Step-Giant-Step) have $O(q^{n/2})$ complexity.

If $n = (d-1)d + 1$ and we use $L = x^{q^d} - bx^q - ax \in \mathbb{F}_{q^{(d-1)d+1}}[x]$ as in Lemma 6.1 as our quasi-subfield polynomial, then we get complexity

$$m! q^{d^2 - dm + 1} \tilde{O}(m^{5.188} 2^{7.376m(m-1)} q^{4.876m(m-1)}) + mq^{2d}$$

for the algorithm in [5]. However, since $d^2 - dm + 1 + 4.876m(m-1) > n/2$ for any d, m, this will not beat generic discrete log algorithms. Thus it appears that the polynomials of Theorem 5.1 will not lead to an ECDLP algorithm that beats generic algorithms, although they can beat brute force algorithms.

Remark 6.1. We briefly discuss adding another term of small degree, for example, an x^{q^2} term. Suppose we have a linearized polynomial $L = x^{q^d} - cx^{q^2} - bx^q - ax \in \mathbb{F}_{q^n}[x]$ which splits completely. Assume $d^2 > 2n$ so that $L(x)$ is a quasi-subfield polynomial. Then the algorithm of [5] has complexity

$$m!q^{n-d(m-1)}\tilde{O}(m^{5.188}2^{7.376m(m-1)}(q^2)^{4.876m(m-1)}) + mq^{2d}.$$

To beat generic discrete log algorithms, we require at least $n - d(m-1) \leq n/2$ and $2d \leq n/2$, which implies $\frac{n}{2(m-1)} \leq d \leq \frac{n}{4}$ and, therefore, $m \geq 3$. As an example, if we choose $q = 2$ and $m = 4$ then we have $2^{7.376m(m-1)}(q^2)^{4.876m(m-1)} \approx 1.45 \cdot q^{205}$ inside the $\tilde{O}$. This means that the overall complexity can beat generic algorithms over $\mathbb{F}_{2^n}$ (for n sufficiently large). For example, a choice of d around $n/5$ when $q = 2$, $m = 4$, would give a complexity $O(q^{0.4n})$ for $n > 500$ (the 500 is a rough estimate). It remains to show that such polynomials exist.

To obtain an estimate for smaller field sizes, we may try $m = 3$, which implies that $d \approx \frac{n}{4}$. These choices would give us complexity

$$q^{n/2}\tilde{O}(3^{5.188}2^{44.256}q^{58.512}) + 3q^{n/2}$$

which is not better than generic algorithms. One example of a linearized polynomial which splits completely and matches these choices ($q = 2$, $d \approx \frac{n}{4}$) is $L = x^{1024} + x^4 + x \in \mathbb{F}_{2^{42}}[x]$.

7 Conclusion and open questions

We have provided necessary and sufficient conditions for $L = x^{q^d} - bx^q - ax \in \mathbb{F}_{q^{(d-1)d+1}}[x]$ to have all q^d roots in $\mathbb{F}_{q^{(d-1)d+1}}$.

The recursive formula that we found for trinomial linearized polynomials is valid for more general linearized polynomials too: Let $L = x^{q^d} - \sum_{i=0}^{d-1} a_i x^{q^i}$.

Set $M_{l,l-d} = 1$, and $M_{l,k} = 0$ for $k \leq 0$ and $k \neq l - d$. For $1 \leq l \leq d$ and $k \geq 1$, let

$$M_{l,k} = M_{l,k-d}a_0^{q^{k-1}} + M_{l,k-d+1}a_1^{q^{k-1}} + \cdots + M_{l,k-1}a_{d-1}^{q^{k-1}} = \sum_{i=0}^{d-1} M_{l,k-d+i}a_i^{q^{k-1}}. \tag{7.1}$$

Then $M_{l,k}$ is the (l,d) entry of $A_{L,k}$. Furthermore, the (l,j) entry of $A_{L,k}$ is $M_{l,k-d+j}$.

We are currently working on extending these results to this more general case, for example, to polynomials of the form $L = x^{q^d} - cx^{q^2} - bx^q - ax$.

Bibliography

[1] Eli Ben-Sasson, Tuvi Etzion, Ariel Gabizon, and Netanel Raviv, Subspace polynomials and cyclic subspace codes, *IEEE Trans. Inf. Theory*, **62**(3) (2016), 1157–1165.

[2] Eli Ben-Sasson and Swastik Kopparty, Affine dispersers from subspace polynomials. *SIAM J. Comput.*, **41**(4) (2012), 880–914.

[3] Bence Csajbók, Giuseppe Marino, Olga Polverino, and Yue Zhou, Maximum Rank-Distance Codes with Maximum Left and Right Idealisers, 2018.

[4] Bence Csajbók, Giuseppe Marino, Olga Polverino, and Ferdinando Zullo, A characterization of linearized polynomials with maximum kernel, *Finite Fields Appl.*, **56** (2019), 109–130.

[5] Ming-Deh A. Huang, Michiel Kosters, Christophe Petit, Sze Ling Yeo, and Yang Yun, Quasi-subfield polynomials and the elliptic curve discrete logarithm problem, In: MathCrypt 2018, 2018.

[6] Gary McGuire and John Sheekey, A characterization of the number of roots of linearized and projective polynomials in the field of coefficients, *Finite Fields Appl.*, **57** (2019), 68–91.

[7] Netanel Raviv and Antonia Wachter-Zeh, Some Gabidulin codes cannot be list decoded efficiently at any radius, *IEEE Trans. Inf. Theory*, **62**(4) (2016), 1605–1615.

[8] Igor Semaev. Summation polynomials and the discrete logarithm problem on elliptic curves. Cryptology ePrint Archive, Report 2004/031, 2004, http://eprint.iacr.org/2004/031.

[9] John Sheekey, MRD codes: constructions and connections, In Combinatorics and Finite Fields: Difference Sets, Polynomials, Pseudorandomness and Applications, Radon Series on Computational and Applied Mathematics, Vol. 23, Ed. by Kai-Uwe Schmidt and Arne Winterhof, De Gruyter, Berlin, 2019, pp. 255–286.

Tahseen Rabbani and Ken W. Smith

Non-Abelian orthogonal building sets

Extending orthogonal building sets to non-Abelian groups

Dedicated to the memory of Robert A. Liebler, who taught the second author how to wield the weapons of group representations in one's attack on combinatorial structures.

Abstract: We examine recent construction techniques of Hadamard difference sets in 2-groups, looking for an extension of orthogonal building sets to non-Abelian groups.

Keywords: Difference sets, Hadamard difference sets, extended building set, non-Abelian difference sets

MSC 2010: 05B10, 05B20

1 Difference sets in Hadamard 2-groups

Definition 1.1. The subset D of a group G of order v is a **difference set** with parameters (v, k, λ) if the size of D is k and for all nonidentity elements g in G, the equation $d_1 d_2^{-1} = g$ has exactly λ solutions $(d_1, d_2),\ d_1, d_2 \in D$.

We may replace the *set* D by the sum of its elements and view D as a member of the group ring $\mathbb{Z}[G]$. With this understanding, we write $D^{(-1)} := \sum_{d \in D} d^{-1}$ and observe that the equation in the definition above translates into the following group ring equation:

$$DD^{(-1)} = (k - \lambda) + \lambda G. \tag{1.1}$$

(Here, G represents the group ring element consisting of the sum of all the elements of the group G; equating a set with the sum of its elements is a typical "abuse of notation" in the study of difference sets.)

In this paper, we focus on difference sets with parameters $(2^{2d+2}, 2^{2d+1} \pm 2^d, 2^{2d} \pm 2^d)$ existing in 2-groups. Difference sets with parameters $(4N^2, 2N^2 \pm N, N^2 \pm N)$ are often called "Hadamard" difference sets because of their connections with Hadamard matrices of order $4N^2$. Groups of order 2^{2d+2} possessing a $(2^{2d+2}, 2^{2d+1} \pm 2^d, 2^{2d} \pm 2^d)$ difference set are designated as "Hadamard 2-groups." In this paper, we attempt to explain the existence of difference sets in some Hadamard 2-groups where general construction techniques; those using Abelian groups, do not appear to work.

Many of the concepts in this paper are an extension of those appearing in [4] and [3].

Tahseen Rabbani, University of Maryland, USA, e-mail: trabbani@math.umd.edu
Ken W. Smith, Sam Houston State University, Huntsville, TX, USA, e-mail: kenwsmith54@gmail.com

https://doi.org/10.1515/9783110621730-011

The recent conclusion of a search for difference sets in groups of order 256, [3], included a generalization of covering extended building sets to "orthogonal building sets" (OBS) in Abelian 2-groups. In this paper, we seek generalizations of orthogonal building sets to *non-Abelian* groups, building on the ideas from that paper.

1.1 Recent contributions to difference sets in 2-groups

In 1990, as part of the Director's Summer Program at the National Security Agency (NSA), a small team of undergraduate students, working under the direction of John Dillon, used the software package *CAYLEY* to complete a survey of groups of order 64, identifying those groups which had a $(64, 28, 12)$ difference set. Ken Smith, beginning a sabbatical year at the NSA, was also involved in this project. By August of 1990, it had been determined that 259 of the 267 groups of order 64 had difference sets. Those eight groups that did not have a difference set were ruled out by a result of Turyn and Dillon ([11], [5]) that prohibit difference sets in a group of order 2^{2d+2} if the group had a cyclic or dihedral image of order 2^{d+3}.

In our search for difference sets in 2-groups, the next larger sets of groups to examine were the 56,092 groups of order 256. At that time, those groups were stored in a sequence of *CAYLEY* libraries and any serious investigation into those groups appeared to be beyond the computer capabilities of that time.

In August 2012, Dillon communicated that most of the groups of order 256 had a difference set, with the status open for only 724 groups. In 2013, Taylor Applebaum, [1], then at the University of Richmond,[1] successfully constructed a difference set in 643 groups containing a normal subgroup isomorphic to $C_4 \times C_4 \times C_2$ (of a possible 649). After that, a team of students, John Clikeman, Gavin McGrew, and Tahseen Rabbani, led by Jim Davis, successfully constructed a difference set in the remaining six groups with such a normal subgroup. This construction technique, using the concept of extended building sets from [4], left 75 groups still in doubt. Difference sets in those remaining groups were then found using a variety of ad hoc techniques.

The number of groups of order 1024 is 49,487,365,422. At this time, there is no convenient library of these groups.

1.2 Extended building sets

We review a few essential concepts relating to the extended building set construction. Let G be a finite group and S a subset of G. As before, we identify S with the formal sum of its elements, so that $S := \sum_{s \in S} s$. If ϕ is an irreducible representation of G and $A = \sum_g a_g\, g$, an element in the group ring $\mathbb{C}[G]$, then $\phi(A) := \sum_g a_g \phi(g)$. We write $\mathrm{Irrep}(G)$

1 Now at Google.

for the set of irreducible representations of G. If G is Abelian then Irrep(G)= G^*, the set of characters of G.

Definition 1.2. A subset S of G has **modulus** m with respect to the character χ if $|\chi(S)| = m$.

Definition 1.3. The **support** of $A = \sum_g a_g\, g$ is the set of group elements g such that $a_g \neq 0$. The **dual support** of A is the set of irreducible representations ϕ such that $\phi(A) \neq 0$.

Definition 1.4. Elements $A, B \in \mathbb{C}[G]$ are **orthogonal** if the product $AB^{(-1)}$ is zero.

There is a classical lemma relating difference sets to the modulus.

Lemma 1.5. *Let D be a k-element subset of an Abelian group G of order v. D is a (v, k, λ) difference set in G if and only if D has modulus $\sqrt{k-\lambda}$ over all nonprincipal characters.*

Definition 1.6. A subset B of G is a **building block** with modulus m if for every nonprincipal character χ on G, $\chi(B) = 0$ or $|\chi(B)| = m$.

Definition 1.7. Let $a, t \in \mathbb{N}$, $m \in \mathbb{R}$. An $(a, m, t, \pm)$ **extended building set** (EBS) on a group G relative to a subgroup U is a set of t building blocks $\{B_1, B_2, \ldots, B_t\}$ with modulus m, such that $t-1$ blocks contain exactly a elements and one contains $a \pm m$ elements, determined by the fourth parameter, and such that for any nonprincipal character χ on U:

1. If χ is principal on U, there is only one B_i such that $\chi(B_i) \neq 0$, for $1 \leq i \leq t$.
2. If χ is nonprincipal on U, then $\forall i, \chi(B_i) = 0$.

If, for every nonprincipal character of G, only one block has a nonzero character sum, then we call the EBS *covering*, abbreviated cov EBS.

Theorem 1.8. *Let G be a finite Abelian group containing K as a subgroup with index t. For any $(a, m, t, \pm)$ cov EBS in K, there is a $(t|K|, at \pm m, at \pm m - m^2)$ difference set in G.*

Example 1.9. In an Abelian group of order 256, the above theorem tells us that one may lift a $(16, 8, 8, -)$ cov EBS in a subgroup isomorphic to $C_4^2 \times C_2$ up to a difference set in G. With presentation

$$K = C_4^2 \times C_2 = \langle x, y, z \mid x^4 = y^4 = z^2 = [x, y] = [x, z] = [y, z] = 1 \rangle,$$

the $(16, 8, 8, -)$ cov EBS, written as elements of a group ring, are:

$$\begin{aligned}
B_1 &= \langle x^2, y \rangle + x\langle x^2 z, y \rangle = [(1 + x^2) + x(1 + x^2 z)](1 + y)(1 + y^2) \\
B_2 &= \langle x^2, yz \rangle + x\langle x^2 z, yz \rangle = [(1 + x^2) + x(1 + x^2 z)](1 + yz)(1 + y^2) \\
B_3 &= \langle x, y^2 z \rangle + y\langle xy^2, y^2 z \rangle = [(1 + x) + y(1 + xy^2)](1 + y^2 z)(1 + x^2) \\
B_4 &= \langle xy, y^2 z \rangle + y\langle xy^3, y^2 z \rangle = (1 + x)(1 + y)(1 + y^2 z)(1 + x^2 y^2)
\end{aligned}$$

$$B_5 = \langle x, z\rangle + y\langle xy^2, z\rangle = [(1+x) + y(1+xy^2)](1+x^2)(1+z)$$
$$B_6 = \langle x^2, y^2, z\rangle = (1+x^2)(1+y^2)(1+z)$$
$$B_7 = \langle y, z\rangle + x\langle x^2y, z\rangle = [(1+y) + x(1+x^2y)](1+y^2)(1+z)$$
$$B_8 = \langle xy, z\rangle + x\langle x^3y, z\rangle = (1+x)(1+y)(1+x^2y^2)(1+z)$$

Any character of K, other than the principal character, maps exactly one of these sets to a nonzero complex number of modulus 8. These blocks are then subsequently translated by the coset representatives of G/K. If G is Abelian, this construction is guaranteed to produce a $(256, 120, 56)$ difference set. Furthermore, it should be evident that the ordering of the coset representatives is irrelevant if G is Abelian.

For non-Abelian groups with a normal $C_4^2 \times C_2$ subgroup, Applebaum conjectured there would always be some choice among of the 8! coset representatives which would result in a difference set in the 649 open groups with such a normal subgroup. This was eventually verified, programmatically with *GAP* [7], using the EBS above, by the end of 2013.

1.3 OBS, the ±1 version of EBS

We generalize covering EBS in two steps. First, the "Hadamard" parameters $(4N^2, 2N^2 \pm N, N^2 \pm N)$ allow us to replace group ring elements with coefficients from $\{0, 1\}$ with group ring elements whose coefficients are from $\{1, -1\}$. In this case, the difference set D is replaced by a ±1 version, $D^* := G - 2D$ and $D^{(-1)}$ is replaced by $D^{*(-1)} := G - 2D^{(-1)}$ so that equation (1.1) becomes

$$D^* D^{*(-1)} = 2^{2d+2}. \tag{1.2}$$

A building block is then a group ring element B, whose coefficients are from $\{-1, 0, 1\}$, with the property that for *every* character $\chi \in G^*$, $\chi(B)$ has modulus 2^{d+1} or is 0; an orthogonal building set (OBS) is then a collection of building blocks $\{B_i\}$ with the property that each character is zero on all but one element of the building set.

For example, replacing B_i by $B_i^* := K - 2B_i$, we have a set $\{B_i^*\}$ such that each character maps exactly one of the B_i^* to a complex number of modulus 16 and maps all others to zero. The principal character is now no different that any other character—this is the advantage of the ±1 viewpoint—in our case, above, the principal character maps B_6^* to 16 and maps the seven other B_i^* to zero. Here are the ±1 versions of the earlier blocks:

$$B_1^* = -(1 + x + x^2 + x^3z)(1+y)(1+y^2)(1-z)$$
$$B_2^* = -(1 + x + x^2 + x^3z)(1-y)(1+y^2)(1-z)$$
$$B_3^* = -(1 + x + y + xyz)(1-y^2)(1+x^2)(1-z)$$

$$
\begin{aligned}
B_4^* &= -(1+x)(1+y)(1-y^2)(1-x^2)(1-z)\\
B_5^* &= -(1+x+y+xy^3)(1+x^2)(1-y^2)(1+z)\\
B_6^* &= -(1-x-y-xy)(1+x^2)(1+y^2)(1+z)\\
B_7^* &= -(1+x+y+x^3y)(1-x^2)(1+y^2)(1+z)\\
B_8^* &= (1+x)(1+y)(1-x^2)(1-y^2)(1+z)
\end{aligned}
$$

For future reference (see Section 2), we note that each of these is a multiple of a group ring idempotent, specifically,

$$
\begin{aligned}
e_1 &:= \frac{1}{8}(1+y)(1+y^2)(1-z), \quad & e_2 &:= \frac{1}{8}(1-y)(1+y^2)(1-z),\\
e_3 &:= \frac{1}{8}(1-y^2)(1+x^2)(1-z), \quad & e_4 &:= \frac{1}{8}(1-y^2)(1-x^2)(1-z),\\
e_5 &:= \frac{1}{8}(1+x^2)(1-y^2)(1+z), \quad & e_6 &:= \frac{1}{8}(1+x^2)(1+y^2)(1+z),\\
e_7 &:= \frac{1}{8}(1-x^2)(1+y^2)(1+z), \quad & e_8 &:= \frac{1}{8}(1-x^2)(1-y^2)(1+z)
\end{aligned}
$$

1.4 The result of Drisko

The construction of OBS in Abelian groups is a powerful concept, especially if we link it to a result on transversals of a normal elementary Abelian subgroup.

Let $E \cong C_2^r$ be an elementary Abelian group of order 2^r, rank r. Suppose

$$E \trianglelefteq K \trianglelefteq G$$

and that K has a covering EBS consisting of 2^r building blocks. If the property $\chi(B_i) \neq 0$ gives a bijection between the building blocks B_i and the characters χ of E then a result of [6] (see also Theorem 1.7 in [3]) assures us that G has a difference set. In most of the examples in [3], K is an Abelian group.

Example 1.10. The extended building set in the first section, $\{B_1, B_2, \ldots, B_8\}$, used by Applebaum and others to find difference sets in groups of order 256, can be replaced by the following:

$$
\begin{aligned}
B_1^* &:= (1+x^2)(1+y^2)(1+z)(1-x-y-xy)\\
B_2^* &:= (1+x^2)(1+y^2)(1-z)(1-x-y-xy)\\
B_3^* &:= (1+x^2)(1-y^2)(1+z)(1-x-y-xy)\\
B_4^* &:= (1+x^2)(1-y^2)(1-z)(1-x-y-xy)\\
B_5^* &:= (1-x^2)(1+y^2)(1+z)(1-x-y-xy)\\
B_6^* &:= (1-x^2)(1+y^2)(1-z)(1-x-y-xy)
\end{aligned}
$$

$$B_7^* := (1-x^2)(1-y^2)(1+z)(1-x)(1-y)$$
$$B_8^* := (1-x^2)(1-y^2)(1-z)(1-x)(1-y)$$

This set has the advantage that these blocks are multiples of the idempotents

$$\frac{1}{8}(1\pm x^2)(1\pm y^2)(1\pm z)$$

existing in the group ring of the elementary Abelian group $E = \langle x^2, y^2, z\rangle$. Then any group of order 256, with $K \cong C_4^2 \times C_2$ a normal subgroup, has a difference set. This saves us from the computer search of Applebaum and others.

Of the fourteen groups of order 16, twelve have a difference set. Ten of these groups have difference sets constructed from an OBS in $C_2 \times C_2$. The other two groups with difference sets require a different construction.

Of the 267 groups of order 64, 259 have a difference set. 176 of these groups have difference sets constructed from an OBS in $C_2 \times C_2 \times C_2$ and 130 of these groups have difference sets constructed from an OBS in $C_4 \times C_4$, leaving 22 groups with difference sets *not* constructible by OBS.

Of the 56092 groups of order 256, there are 56049 possessing a difference set. Of those groups, 42268 have difference sets constructed from an OBS in $C_2 \times C_2 \times C_2 \times C_2$, 49165 of these groups have difference sets constructed from an OBS in $C_4 \times C_4 \times C_2$, and 684 of these groups have difference sets constructed from an OBS in $C_8 \times C_8$ giving a union of 54633 groups, leaving 1416.

The percentages of groups given by OBS is 83 %, 91.5 %, and 97.5 %, leaving 1416 groups with difference sets *not* constructible by OBS.

We summarize this in Tables 1 and 2.

Table 1: Success rate of Abelian OBS in groups of order 64.

Technique	# groups of order 64	%
OBS in C_2^3	176	65.9 %
OBS in $C_4 \times C_4$	130	48.7 %
ALL OBS in Abelian groups	237	88.8 %
other construction required	22	8.2 %
Turyn/Dillon bound	8	3.0 %

Table 2: Success rate of Abelian OBS in groups of order 256.

Technique	# groups of order 256	%
OBS in C_2^4	42268	75.35 %
OBS in $C_4 \times C_4 \times C_2$	49165	87.65 %
OBS in $C_8 \times C_8$	684	2.52 %
ALL OBS in Abelian groups	54633	97.40 %
other construction required	1416	2.52 %
Turyn/Dillon bound	43	0.08 %

2 Non-Abelian OBS

Since the existence of OBS in Abelian groups is so powerful, need we restrict OBS to Abelian groups? Our second step in generalizing extended building sets is to create OBS in non-Abelian groups.

Example 2.1. Let $Q = \langle x, y : x^4 = y^4 = 1, xyx^{-1} = y^{-1}, x^2 = y^2 \rangle$ be the quaternion group of order eight. Then the set

$$\{(1 - x - y - xy)(1 + x^2),\ (1 - x - y - xy)(1 - x^2)\}$$

acts as an OBS of two sets, with center $E = \langle x^2 \rangle$. Any group of order 16 with normal subgroup isomorphic to Q has a $(16, 6, 2)$ difference set.

Example 2.2. Consider the direct product $Q \times \langle z : z^2 = 1 \rangle$ of the quaternion group, above, and the cyclic group of order two. Then the set

$$\{(1 - x - y - xy)(1 \pm x^2)(1 \pm z)\}$$

acts as an OBS of four sets, with center $E = \langle x^2, z \rangle$. Any group of order 64 with normal subgroup isomorphic to this group has a $(64, 28, 12)$ difference set.

2.1 Characters and idempotents

The definition of building blocks and extended building sets in Abelian groups depends on interplay between subsets of an Abelian group and sets of characters of that group. We may generalize those definitions to non-Abelian groups by focusing on idempotents, associating a set of characters (or a set of irreducible representations) with a unique idempotent. Associated with each finite group G is a set of irreducible representations $\mathrm{Irrep}(G)$. We will assume that these representations are unitary, that is, $\phi(g^{-1}) = \overline{\phi(g)}^T$. Good references for group representations, characters and their associated idempotents include [8] (Chapter 3), [2] (Chapter 5), and the excellent small monograph [9].

Definition 2.3. An element e of a ring is an **idempotent** if $e^2 = e$. An element is **central** if it commutes with all elements of the ring. The **center** of a ring is the set of all central elements.

The center of the full matrix ring $\mathrm{Mat}_n(F)$ over a field F is the set of scalar matrices. This means that the only *central* idempotent of a full matrix ring is the identity matrix. Since the group ring $\mathbb{C}[G]$ of a finite group G over the complex numbers is isomorphic to the direct sums of the irreducible representations then, given any irreducible representation ϕ, there exists a unique element e_ϕ of the group ring $\mathbb{C}[G]$ such that

$$\phi(e_\phi) = \phi(1) \text{ and if } \phi' \in \mathrm{Irrep}(G), \phi' \neq \phi, \text{ then } \phi'(e_\phi) = 0.$$

Since, for any $\phi' \in \mathrm{Irrep}(G)$, $\phi'(e^2) = \phi'(e)\phi'(e)$, then $e_\phi \cdot e_\phi = e_\phi$ and so e_ϕ is an idempotent in $\mathbb{C}[G]$. The element e_ϕ is *the* central idempotent corresponding to the representation ϕ. Given a subset $S \subseteq \mathrm{Irrep}(G)$, we may define

$$e_S := \sum_{\sigma \in S} e_\sigma.$$

The element e_S is an idempotent and is the *unique* element of $\mathbb{C}[G]$ such that $\sigma(e_S) = \sigma(1)$ if and only if $\sigma \in S$.

The lattice of subsets of $\mathrm{Irrep}(G)$ gives rise to a lattice of central idempotents of $\mathbb{C}[G]$: If S, T are subsets of $\mathrm{Irrep}(G)$ then

$$e_{S \cap T} = e_S \cdot e_T$$

and if S, T are disjoints subsets of $\mathrm{Irrep}(G)$ then

$$e_{S \cup T} = e_S + e_T.$$

If S is a subset of $\mathrm{Irrep}(G)$ and S^C is the complement of S in $\mathrm{Irrep}(G)$, then

$$e_{S^C} = 1 - e_S.$$

This lattice of idempotents corresponds with the lattice of subsets of irreducible representations. In particular, if G is an Abelian group, any set of characters of G is associated with a unique idempotent and we may replace our emphasis on characters with a corresponding emphasis on idempotents. (The idempotents given by this lattice are all *central* in $\mathbb{C}[G]$. If G is a non-Abelian group then there will be idempotents that are not central.)

If χ is a linear representation of G then the idempotent e_χ is equal to

$$e_\chi = \frac{1}{|G|} \sum_g \overline{\chi(g)}\, g = \frac{1}{|G|} \sum_g \chi(g)\, g^{-1}.$$

(Note the need for the conjugate map.) Equating functions with group ring elements allows one to then write

$$\overline{\chi} = |G| e_\chi.$$

Definition 2.4. A set $\{e_1, e_2, \ldots, e_\ell\}$ of idempotents of $\mathbb{C}[G]$ is **complete** (or covering) if the idempotents are pairwise orthogonal and sum to 1.

If a set of *central* idempotents is complete, then the sets of irreducible representations corresponding the idempotents partition the set of all irreducible representations.

Idempotents provide an equivalent definition of building blocks.

Lemma 2.5. *Let G be an Abelian group. An element $B \in \mathbb{Z}[G]$ with coefficients from $\{-1, 0, 1\}$ is a building block of length m if and only if there is a nonzero idempotent e such that*

$$Be = B \quad \text{and} \quad BB^{(-1)}e = m^2 \cdot e.$$

Proof. A building block of modulus m in an Abelian group is a set B with the property that $|\chi(B)| = m$ for some characters χ and is equal to zero for all others. Let S be the dual support of B. Then, for $\chi \in S$,

$$\begin{aligned} |\chi(B)| = m &\iff \chi(B)\overline{\chi(B)} = m^2 \\ &\iff \chi(B)\chi(B^{(-1)}) = m^2 \iff \chi(BB^{(-1)}) = m^2\chi(1) \end{aligned}$$

Let e be the idempotent associated with the set S of characters. Then

$$Be_S = B \quad \text{and} \quad BB^{(-1)}e_S = m^2 e_S.$$

On the other hand, if B has the property that there exists an idempotent e such that $Be = B$ and $BB^{(-1)}e = m^2 e$ then let S be the set of all characters ϕ such that $\phi(e) = 1$. If $\phi \in S$, then

$$\phi(B)\overline{\phi(B)} = \phi(BB^{(-1)}) = m^2$$

and so $|\phi(B)| = m$. If $\phi \notin S$, then since $\phi(e) = 0$ we have

$$\phi(B) = \phi(Be) = \phi(B)\phi(e) = \phi(B) \cdot 0 = 0. \qquad \square$$

Lemma 2.6. *If G is Abelian, then $A, B \in \mathbb{C}[G]$ are orthogonal if and only if their dual supports are disjoint.*

Proof. Since each element in $\mathbb{C}[G]$ is determined by its image under the irreducible representations then

$$AB^{(-1)} = 0 \iff \forall \chi \in G^*, \chi(AB^{(-1)}) = 0.$$

But since G is Abelian, χ is a homomorphism and so $\chi(AB^{(-1)}) = \chi(A)\overline{\chi(B)}$. Since χ maps group ring elements into the complex numbers, then $\chi(A)\overline{\chi(B)} = 0$ implies that either $\chi(A) = 0$ or $\chi(B) = 0$, so χ is *not* in the dual support of either A or B. Therefore, $AB^{(-1)} = 0$ implies that the intersection of the dual supports of A and B are disjoint.

Conversely, if the dual supports of A and B are disjoint then for any $\chi \in G^*, \chi(A) = 0$ or $\chi(B) = 0$ and so $\chi(A)\overline{\chi(B)} = 0$. Since this is true for *all* $\chi \in G^*$, then $AB^{(-1)} = 0$. $\square$

2.2 Building blocks in non-Abelian groups

Let G be a finite group, possibly non-Abelian. We adapt the earlier definitions of Section 1 for the non-Abelian setting.

Definition 2.7. An element $B \in \mathbb{Z}[G]$ with coefficients from $\{-1, 0, 1\}$ is a **building block** of length m with respect to the nonzero idempotent e if

$$eB = B \quad \text{and} \quad BB^{(-1)} = m^2 \cdot e.$$

Fixing the positive real number m, a **building set** in G is a collection $\{B_1, B_2, \ldots, B_\ell\}$ of mutually orthogonal building blocks of length m with an associated set of idempotents, $\{e_1, e_2, \ldots, e_\ell\}$.

Definition 2.8. A building set is **covering** if the associated idempotents form a complete set of idempotents.

Lemma 2.9. *Suppose B_1, B_2 are building blocks of length m in a finite group G and suppose the associated idempotents e_1, e_2 are central. Then B_1, B_2 are orthogonal if and only if their dual supports are disjoint.*

Proof. Suppose the dual supports of B_1, B_2 are disjoint. Then for any representation $\phi \in \text{Irrep(G)}$, either $\phi(B_1) = 0$ or $\phi(B_2) = 0$. Since we may assume that ϕ is a unitary representation, then $\phi(B_2^{(-1)}) = \overline{\phi(B_2)}^T$ and so $\phi(B_2) = 0$ if and only if $\phi(B_2^{(-1)}) = 0$.

Therefore, $\phi(B_1 B_2^{(-1)}) = \phi(B_1)\phi(B_2^{(-1)}) = 0$. Since this is true for all irreducible representations, then $B_1 B_2^{(-1)} = 0$.

On the other hand, suppose $B_1 B_2^{(-1)} = 0$ but there is an irreducible representation ϕ in the intersection of the dual supports of B_1, B_2. Let e be the idempotent associated with ϕ. Then $B_1 B_1^{(-1)} e = m^2 e = B_2 B_2^{(-1)} e$. Since e is a central idempotent, $e = e^{(-1)}$ and $m^2 e = (m^2 e)^{(-1)} = (eB_2^{(-1)} B_2)$. Therefore,

$$\begin{aligned}(B_1 B_2^{(-1)})(B_1 B_2^{(-1)})^{(-1)} e &= B_1 (B_2^{(-1)} B_2) B_1^{(-1)} e = B_1 (B_2^{(-1)} B_2 e) B_1^{(-1)}. \\ &= B_1 (m^2 e) B_1^{(-1)} = (m^2 e) B_1 B_1^{(-1)} = (m^2 e)(m^2 e) = m^4 e^2 = m^4 e\end{aligned}$$

contradicting the claim that $B_1 B_2^{(-1)} = 0$. □

Example 2.10. The idempotents corresponding to the extended building set in example 1.9, were given immediately after that example. That set of idempotents, $\{e_1, \ldots, e_8\}$, is *complete*, as the sum of the eight idempotents is 1. In making this correspondence of idempotents with characters, we assign the principal character to the sixth building block, the one "short" block in the list, a set of size 8, not 16.

2.3 Families of (a, m, t) OBS

We define an (a, m, t) OBS and a product construction.

Definition 2.11. An (a, m, t) **orthogonal building set** (OBS) in a group K is a set $\{B_1, B_2, \ldots, B_t\}$ of t mutually orthogonal elements of $\mathbb{Z}[G]$ with coefficients from $\{-1, 0, 1\}$ such that each element B_i has support of size a and modulus m with respect to an idempotent e_i. The OBS is **covering** if the set $\{e_1, e_2, \ldots, e_t\}$ of associated idempotents is complete. The OBS is **elementary** if the order of K is equal to a and the idempotents are idempotents of a central elementary subgroup E.

Example 2.12. The two subsets of the quaternion group given in example 2.1 form an $(8, 4, 2)$ OBS while those in example 2.2 form a $(16, 8, 4)$ OBS.

A difference set in a group of order 2^{2d+2} gives, trivially, a $(2^{2d+2}, 2^{d+1}, 1)$ OBS. In order to construct a difference set in a group of order 2^{2d+2}, we seek OBS with parameters $(2^{2d+2-r}, 2^{d+1}, 2^r)$.

Lemma 2.13 (Product construction). *If $\mathcal{B}_1 = \{B_{1,1}, \ldots B_{1,i}, \ldots B_{1,t_1}\}$ is an (a_1, m_1, t_1) OBS in K_1 and $\mathcal{B}_2 = \{B_{2,1}, \ldots B_{2,i}, \ldots B_{2,t_2}\}$ is an (a_2, m_2, t_2) OBS in K_2 then $\mathcal{B}_1 \times \mathcal{B}_2 = \{B_{1,i} \times B_{2,j} : 1 \le i \le t_1, 1 \le j \le t_2\}$ is an (a_1a_2, m_1m_2, t_1t_2) OBS in $K_1 \times K_2$.*

The set of associated idempotents for $\mathcal{B}_1 \times \mathcal{B}_2$ is the direct product of the idempotents associated with $\mathcal{B}_1$ and $\mathcal{B}_2$.

Let $\mathcal{H}(a, m, t)$ represent the set of (isomorphism classes of) groups of order a with an elementary (a, m, t) OBS.

For example: $\mathcal{H}(2, 2, 2) = \{C_2\}$. $\mathcal{H}(4, 4, 4) = \{C_2^2\}$. $\mathcal{H}(4, 2, 1) = \{C_4, C_2^2\}$. $\mathcal{H}(8, 4, 2) = (\mathcal{H}(4, 2, 1) \times \mathcal{H}(2, 2, 2)) \cup \{Q_4\} = \{C_4 \times C_2, C_2^3, Q_4\}$.

Since the direct product of the set $\mathcal{H}(a_1, m_1, t_1)$ and $\mathcal{H}(a_2, m_2, t_2)$ is a subset of $\mathcal{H}(a_1a_2, m_1m_2, t_1t_2)$, we might seek groups with a "primitive" OBS, one not created by a product. (See 2.1 for a non-Abelian example.)

It is likely that there are more examples of covering building sets in non-Abelian groups of order 16. But it turns out that a "near miss" in $C_8 \times C_2$ will create OBS in important groups of order 32.

2.4 Complementary pairs of building sets

Some groups of order 64, with difference sets, do not appear to allow a covering building set from normal subgroups of order 16. But there are variations on the covering building set construction that are significant.

Example 2.14. The group $H = C_8 \times C_2$ does not have a covering building set of four building blocks of length eight. But it has an interesting "near miss." Consider the following four blocks in $H = C_8 \times C_2 = \langle x, y : x^8 = y^2 = [x, y] = 1\rangle$:

$$B_0 := (1 - x^4)(1 + x^2y)(1 - x - y - xy)$$
$$B_1 := (1 + x^4)(1 + y)(1 - x - x^2 - x^3)$$

$$B_2 := (1-x^4)(1-x^2y)(1-x-y-xy)$$
$$B_3 := (1+x^4)(1-y)(1-x-x^2-x^3)$$

The elements B_1 and B_3 are mutually orthogonal building blocks of length eight, with respect to the idempotents $\frac{(1+x^4)(1\pm y)}{4}$. Indeed, all pairs of blocks are mutually orthogonal *except* the pair B_0, B_2. Noting that $x^8 = 1, y^2 = 1$, we can compute

$$B_0 B_0^{(-1)} = 16(1-x^4) = B_2 B_2^{(-1)}$$

and

$$B_0 B_2^{(-1)} = 8x^2(1-x^4) = -B_2 B_0^{(-1)}$$

These relations imply that with appropriate choice of the elements g_i, the blocks B_0 and B_2 can be made to complement each other within the group ring equation so that $D = \sum g_i B_i$ will be a difference set.

The group $H = C_8 \times C_2$ only has three involutions (elements of order two), x^4, y and x^4y. The element x^4 is the only involution which is itself a square and so it is fixed by all automorphisms. Thus, if G is a larger group with H as a normal subgroup, then the idempotents $\frac{1-x^4}{2}$, $\frac{(1+x^4)(1-y)}{4}$ and $\frac{(1+x^4)(1+y)}{4}$ are all fixed by conjugation by elements of G.

Let G be a group of order 64 with H as a normal subgroup. Let $g_0 = 1, g_1, g_2, g_3$ be a left transversal of H in G. Set

$$D := \sum_j g_j B_j.$$

Then

$$DD^{(-1)} = 64 + 8x^2(1-x^4)g_2^{-1} - 8g_2x^2(1-x^4)$$

We have a difference set if and only if

$$8x^2(1-x^4)g_2^{-1} - 8g_2x^2(1-x^4) = 0.$$

Since $8(1-x^4)$ is in the center of the group ring, it can be effectively ignored. So we require the group element $x^2g_2^{-1}$ to be equal to g_2x^2. When this is true, the set $\{B_0 + g_2B_2, B_1 + g_2B_3\}$ is a $(32, 8, 2)$ OBS in a group $H' = \langle g_2, x, y\rangle$ of order 32. Any group of order 64 containing the subgroup H' then has a difference set of the form $D = B_0 + g_2B_2 + g'(B_1 + g_2B_3)$ where $g' \notin H'$.

There are eight groups of order 32 containing a normal copy of $H = C_8 \times C_2$ and an element $g \notin H$ such that $x^2g^{-1} = gx^2$. There are 118 groups of order 64 with a subgroup of order 32 isomorphic to one of these eight. So these 118 groups have a difference set

constructible in this manner. Included in this list are 13 groups not covered by the earlier Abelian or non-Abelian OBS.

This reduces the open groups to SmallGroup(64, cn)[2] where $cn \in \{42, 46, 48, 49, 51\}$. These remaining five groups all have $C_8 \times C_2$ as a normal subgroup but the "building block" structure appears to be more complicated. In these groups a more subtle argument, probably using the inner automorphisms of the group of order 64 acting on $C_2 \times C_8$, will be necessary.

Example 2.15. Let $M_{64} := \langle x, y : x^{32} = y^2 = 1, yxy^{-1} = x^{17} \rangle$, the modular group of order 64. The element x^{16} is a central involution (indeed the center of G is $\langle x^2 \rangle$) and $\langle x^{16}, y \rangle$ is a subgroup isomorphic to C_2^2. However, y is not central. The group ring $\mathbb{C}[M_{64}]$ has a complete set of four idempotents $e_{\pm\pm} := \frac{1}{4}(1 \pm x^{16})(1 \pm y)$. Led by these idempotents, we consider the following four blocks. (Here, g_1, g_2, g_3, g_4 are group elements to be chosen later.)

1. $B_{++} := [(1 + x^8) + g_1(1 - x^8)](1 + x^{16})(1 + y)$
2. $B_{+-} := [(1 + x^8) + g_2(1 - x^8)](1 + x^{16})(1 - y)$
3. $B_{-+} := [(1 + x^8) + g_3(1 - x^8)](1 - x^{16})(1 + y)$
4. $B_{--} := [(1 + x^8) + g_4(1 - x^8)](1 - x^{16})(1 - y)$

Observe that

$$B_{++}B_{++}^{(-1)}(e_{++}) = 16(1 + x^{16})(1 + y) \quad \text{and} \quad B_{+-}B_{+-}^{(-1)}(e_{+-}) = 16(1 + x^{16})(1 - y)$$

so the sets B_{++} and B_{+-} are orthogonal building blocks. However, the other two sets do not work quite as nicely, due to the fact that y is not in the center of G. But one can tinker with the values of the group elements g_1, g_2, g_3, g_4 and, just as in the $C_8 \times C_2$ example, create complementary building sets. The first difference set found in this group (Δ in [10]) has form (as a ± 1 DS)

$$\Delta^* = A(1 + x^{16}) + B(1 - x^{16})$$

where

$$A = x^3[(1 + y) - x^2(1 - y)](1 + x^8) + x^4[(1 + y) + x^4(1 - y)](1 - x^8)$$

and

$$B = x^{-2}(1 + x)(1 - y)(1 + x^8) - x(1 - x)(1 + y)(1 - x^8).$$

2 "SmallGroup(64, cn)" references a group of order 64 defined in the GAP ([7]) "SmallGroup" library.

2.5 Final thoughts

One might wonder if identification of OBS in various non-Abelian groups is sufficient to construct difference sets in *all* groups that have them. However, there are two groups of order 64 with the property that *all* normal subgroups in those groups also occur in groups that do *not* have differences sets. (One of these two groups is M_{64}, above.) In groups of order 256, there are ten groups with difference sets, yet all their normal subgroups are normal subgroups of groups that do *not* have difference sets. (One of these ten groups is the "most difficult" group of order 256, the group $C_{64} \rtimes_{47} C_4$; see [12] for a construction.) Thus the OBS construction cannot work directly in those groups. However, in these groups, we often found a "near miss" as described above in $C_8 \times C_2$; we found four sets where one pair were orthogonal building sets and the other pair were complementary, not quite building sets of the same modulus, but offsetting each other in just the right way.

On the other hand, *every* difference set in 2-groups has *some* further connection to OBS.

Theorem 2.16. *Suppose D is a ± 1 difference set in a 2-group G. Let g be a central involution in G. Then there are group ring elements $A, B \in \mathbb{Z}[G]$ such that*

$$D = A(1+g) + B(1-g).$$

The set $\{A(1+g), B(1-g)\}$ is then a $(2^{2d+1}, 2^{d+1}, 2)$ OBS in this group of order 2^{2d+2}.

2.6 Summary

It is possible that all the groups with difference sets can be explained in three waves of construction:

1. Use OBS in Abelian groups and selected non-Abelian groups to dispatch most (> 99 %) of the groups.
2. Use OBS in selected non-Abelian groups to dispatch most of the remaining groups.
3. Use complementary paired OBS (as above in examples 2.14 and 2.15) to complete construction of difference sets in the final more difficult groups.

All difference sets can be constructed from building blocks, in *some* way, but it is not clear if there is a recursive process that covers every case. The most difficult groups appear to be those that skate close to the Turyn/Dillon exponent bound ([11], [5]) described in Section 1.1 and, in addition, have a cyclic center, preventing a convenient elementary Abelian subgroup.

The use of *non-Abelian* groups for the construction of difference sets is important. Although large families of difference sets can be constructed from building blocks in Abelian groups, this does not cover all groups and if we are to find difference sets in all Hadamard 2-groups, we *must* develop a theory of non-Abelian building blocks.

In drafting this attempt at a general theory of covering building sets, a number of questions occur. (Any of these might be a good place to start an undergraduate research project or a masters thesis.)

1. Explain the difference sets occurring in the five "difficult" groups of order 64. Is there an elegant explanation using the normal subgroup $C_8 \times C_2$?
2. Identify OBS in groups of order 16 and 32. (Must we always require that the idempotents come from a central elementary Abelian subgroup E?)
3. Lift the OBS and near-OBS in groups of order 16 to OBS in groups of order 32.
4. Construct the difference sets in the "difficult" groups of order 256, those that do not have a convenient OBS. Can we generalize this construction to *families* of groups with cyclic centers?

Bibliography

[1] T. Applebaum, Difference Sets in Non-Abelian 2-Groups, Honors Thesis, University of Richmond, 2013.
[2] C. W. Curtis and I. Reiner, Representation Theory of Finite Groups and Associative Algebras, 1962.
[3] Davis, Jedwab, et al., Constructions of difference sets in nonabelian 2-groups, submitted, 2019.
[4] J. Davis and J. Jedwab, A unifying construction for difference sets, *J. Comb. Theory, Ser. A*, **80**(1) (1997), 13–78.
[5] J. F. Dillon, Variations on a scheme of McFarland for noncyclic difference sets, *J. Comb. Theory, Ser. A*, **40** (1985), 9–21.
[6] A. Drisko, Transversals in row-Latin rectangles, *J. Comb. Theory, Ser. A*, **84** (1998), 181–195.
[7] The GAP Group, GAP-Groups, Algorithms, and Programming, Version 4.9.1, 2018.
[8] D. Gorenstein, Finite Groups, 2nd ed., AMS Chelsea Publishing, 1980.
[9] W. Ledermann, Introduction to Group Characters, Cambridge University Press, 1977.
[10] R. A. Liebler and K. W. Smith, On difference sets in certain 2-groups, In: Coding Theory, Design Theory, Group Theory, Ed. by D. Jungnickel, S. Vanstone, Wiley, NY, 1993, pp. 195–212.
[11] R. Turyn, Character sums and difference sets, *Pac. J. Math.*, **15** (1965), 319–346.
[12] W. Yolland, Existence of a difference set in the last group of order 256, summer research report, Simon Fraser University, 2016.

Satoru Fukasawa and Katsushi Waki

Examples of plane rational curves with two Galois points in positive characteristic

Abstract: We present four new examples of plane rational curves with two Galois points in positive characteristic, and determine the number of Galois points for three of them. Related to these results, a theorem of the first author in algebraic geometry is applied to a group-theoretic problem on projective linear groups.

Keywords: Galois point, plane curve, Galois group, projective linear groups

MSC 2010: 14H50, 20G40

1 Introduction

Let $C \subset \mathbb{P}^2$ be an irreducible plane curve over an algebraically closed field k of characteristic $p \geq 0$ with $k(C)$ as its function field. For a point $P \in \mathbb{P}^2$, if the function field extension $k(C)/\pi_P^* k(\mathbb{P}^1)$ induced by the projection π_P is Galois, then P is called a Galois point for C. This notion was introduced by Yoshihara ([7, 9]). Furthermore, if a Galois point P is a smooth point of C (resp., a point in $\mathbb{P}^2 \setminus C$), then P is said to be inner (resp., outer). The associated Galois group at P is denoted by G_P. In this paper, we present four new examples of plane rational curves with two Galois points, which update the tables in [11].

Here, we assume that $p \geq 3$ and $q \geq 5$ is a power of p.

Theorem 1. *Let C_1 be the plane curve of degree q which is the image of the morphism*

$$\varphi_1 : \mathbb{P}^1 \to \mathbb{P}^2;\ (s:t) \mapsto \left(s^{\frac{q+1}{2}} t^{\frac{q-1}{2}} : (s-t)t^{q-1} : s^q - st^{q-1}\right).$$

Then:

(a) *The point $P_1 := (0:1:0) = \varphi_1(0:1)$ is an inner Galois point with $G_{P_1} \cong D_{q-1}$, where D_{q-1} is the dihedral group of order $q-1$.*

(b) *The point $P_2 := (1:0:0) = \varphi_1(1:1)$ is an inner Galois point with $G_{P_2} \cong \mathbb{Z}/(q-1)\mathbb{Z}$.*

(c) *The number of inner Galois points on C_1 is exactly two.*

Acknowledgement: The first author was partially supported by JSPS KAKENHI Grant Numbers 16K05088 and 19K03438.

Satoru Fukasawa, Katsushi Waki, Department of Mathematical Sciences, Faculty of Science, Yamagata University, Kojirakawa-machi 1-4-12, Yamagata 990-8560, Japan, e-mails: s.fukasawa@sci.kj.yamagata-u.ac.jp, waki@sci.kj.yamagata-u.ac.jp

https://doi.org/10.1515/9783110621730-012

Theorem 2. *Let C_2 be the plane curve of degree q which is the image of the morphism*

$$\varphi_2 : \mathbb{P}^1 \to \mathbb{P}^2;\ (s:t) \mapsto \left(s^{\frac{q+1}{2}} t^{\frac{q-1}{2}} : (s-t)^{\frac{q+1}{2}} t^{\frac{q-1}{2}} : s^q - st^{q-1}\right).$$

Then:
(a) *The point $P_1 := (0:1:0) = \varphi_2(0:1)$ is an inner Galois point with $G_{P_1} \cong D_{q-1}$.*
(b) *The point $P_2 := (1:0:0) = \varphi_2(1:1)$ is an inner Galois point with $G_{P_2} \cong D_{q-1}$.*
(c) *The number of inner Galois points on C_2 is exactly two.*

For the case where the characteristic is zero and C is rational, Yoshihara [10, Lemma 13] asserts that cyclic and dihedral groups do not appear as Galois groups of outer Galois points at the same time. To the contrary, in positive characteristic, we present the following example.

Theorem 3. *Let C_3 be the plane curve of degree $q+1$ which is the image of the morphism*

$$\varphi_3 : \mathbb{P}^1 \to \mathbb{P}^2;\ (s:t) \mapsto \left(s^{\frac{q+1}{2}} t^{\frac{q+1}{2}} : (s+t)^{q+1} : s^{q+1} + \gamma t^{q+1}\right),$$

where $\gamma \in \mathbb{F}_q \setminus \{0, \pm 1\}$. Then:
(a) *The point $P_1 := (0:1:0)$ is an outer Galois point with $G_{P_1} \cong D_{q+1}$.*
(b) *The point $P_2 := (1:0:0)$ is an outer Galois point with $G_{P_2} \cong \mathbb{Z}/(q+1)\mathbb{Z}$.*
(c) *The number of outer Galois points for C_3 is exactly two.*

Theorem 4. *Let C_4 be the plane curve of degree $q+1 > 6$ which is the image of the morphism*

$$\varphi_4 : \mathbb{P}^1 \to \mathbb{P}^2;\ (s:t) \mapsto \left(s^{\frac{q+1}{2}} t^{\frac{q+1}{2}} : (s+t)^{\frac{q+1}{2}} (s+\gamma t)^{\frac{q+1}{2}} : s^{q+1} - \gamma t^{q+1}\right),$$

where $\gamma \in \mathbb{F}_q \setminus \{\pm 1\}$ and $\gamma^{\frac{q-1}{2}} = 1$. Then:
(a) *The point $P_1 := (0:1:0)$ is an outer Galois point with $G_{P_1} \cong D_{q+1}$.*
(b) *The point $P_2 := (1:0:0)$ is an outer Galois point with $G_{P_2} \cong D_{q+1}$.*

Although we do not use a criterion [3, Theorem 1] of the first author for the proof of our four theorems, the criterion was useful to find our embeddings. For rational curves, by the criterion [3, Theorem 1] and Lüroth's theorem, the problem on the existence of two Galois points is translated into the following group-theoretic problem on projective linear groups.

Problem 1. Let $X = \mathbb{P}^1(k)$ be the projective line over k. We consider the following two conditions for a pair (H_1, H_2) of different finite subgroups H_1 and $H_2 \subset \mathrm{PGL}(2,k)$:
(a) $H_1 \cap H_2 = \{1\}$.
(b) There exist different points P_1 and $P_2 \in X$ such that

$$\{\sigma(P_2) \mid \sigma \in H_1 \setminus \{1\}\} = \{\tau(P_1) \mid \tau \in H_2 \setminus \{1\}\}$$

(with multiplicities).

When does a pair (H_1, H_2) with (a) and (b) exist?

The following application of the criterion to this problem is presented.

Theorem 5. *Assume that k is an algebraically closed field of characteristic zero and $n \geq 3$ is an integer. Then the following conditions are equivalent:*

(1) *There exists a pair (H_1, H_2) of cyclic groups of order n with conditions (a) and (b) in Problem* 1.

(2) $n \leq 5$.

2 Proof of Theorems 1 and 2

We assume that $\alpha \in \mathbb{F}_q$ is a primitive element. The following two lemmas are easily proved.

Lemma 1. *Let σ and $\tau \in \mathrm{Aut}(\mathbb{P}^1) \cong \mathrm{PGL}(2, k)$ be represented by the matrices*

$$A_\sigma = \begin{pmatrix} 1 & 0 \\ 0 & \alpha^2 \end{pmatrix} \quad \text{and} \quad A_\tau = \begin{pmatrix} 0 & 1 \\ \alpha & 0 \end{pmatrix}$$

respectively, that is, $\sigma(s,t) = (s,t)A_\sigma$ and $\tau(s,t) = (s,t)A_\tau$. Then:

(a) *The group H generated by σ and τ is isomorphic to D_{q-1}.*

(b) *The rational function $f(s) = s^{\frac{q-1}{2}} - s^{-\frac{q-1}{2}} \in k(\mathbb{P}^1) = k(s)$ is invariant under the action by H.*

Lemma 2. *Let $\eta \in \mathrm{Aut}(\mathbb{P}^1)$ be represented by the matrix*

$$A_\eta = \begin{pmatrix} 1 & 0 \\ \alpha - 1 & \alpha \end{pmatrix}.$$

Then:

(a) *The order of η is $q - 1$.*

(b) *The rational function $g(s) = \frac{s-1}{s^q - s} \in k(\mathbb{P}^1) = k(s)$ is invariant under the action by η^*.*

When R_1 and R_2 are different points in $\mathbb{P}^2$, the line passing through R_1 and R_2 is denoted by $\overline{R_1R_2}$.

Proof of Theorem 1. The composite (rational) map $\pi_{P_1} \circ \varphi_1$ is given by

$$(s : t) \mapsto \left(s^{\frac{q+1}{2}} t^{\frac{q-1}{2}} : s^q - st^{q-1}\right) = \left(s^{\frac{q-1}{2}} t^{\frac{q-1}{2}} : s^{q-1} - t^{q-1}\right),$$

since π_{P_1} is represented by $(X : Y : Z) \mapsto (X : Z)$. By this expression, the degree of $\pi_{P_1} \circ \varphi_1$ is $q - 1$, and hence, φ_1 is birational onto the image C_1. Note that

$$\pi_{P_1} \circ \varphi_1(s : 1) = \left(s^{\frac{q-1}{2}} : s^{q-1} - 1\right) = \left(1 : f(s)\right),$$

where $f(s)$ is the rational function as in Lemma 1. By Lemma 1, the rational function $f(s)$ is invariant under the action by $H \cong D_{q-1}$. Therefore, $k(s)/k(f(s))$ is a Galois extension, and hence, P_1 is a Galois point. In this case, $G_{P_1} \cong D_{q-1}$. Assertion (a) in Theorem 1 follows. Further,

$$\pi_{P_2} \circ \varphi_1(s:1) = \big(s-1 : s^q - s\big) = \big(g(s) : 1\big),$$

where $g(s)$ is the rational function as in Lemma 2. By Lemma 2, the rational function $g(s)$ is invariant under the action by η^*. Therefore, $k(s)/k(g(s))$ is a Galois extension, and hence, P_2 is a Galois point. In this case, $G_{P_2} \cong \mathbb{Z}/(q-1)\mathbb{Z}$. Assertion (b) in Theorem 1 follows.

We prove that the set of all inner Galois points for C_1 is equal to $\{P_1, P_2\}$. Let $Q := \varphi_1(1:0) = (0:0:1)$. Then the composite map $\pi_Q \circ \varphi_1$ is given by

$$(s:t) \mapsto \big(s^{\frac{q+1}{2}} t^{\frac{q-1}{2}} : (s-t)t^{q-1}\big) = \big(s^{\frac{q+1}{2}} : (s-t)t^{\frac{q-1}{2}}\big).$$

Since the degree of $\pi_Q \circ \varphi_1$ is $(q+1)/2$, the point Q is a singular point of C_1 with multiplicity $(q-1)/2$. The ramification indices of $\pi_Q \circ \varphi_1$ at $(0:1)$ and at $(1:0)$ are equal to $(q+1)/2$ and $(q-1)/2$, respectively. For the function $h(t) = (1-t)t^{\frac{q-1}{2}}$,

$$h'(t) = -\frac{1}{2} t^{\frac{q-3}{2}} (1+t).$$

Therefore, three points $(0:1)$, $(1:0)$ and $(1:-1)$ are all ramification points for $\pi_Q \circ \varphi_1$. Furthermore, the ramification index at $(1:-1)$ is 2. Note that $\varphi_1^{-1}(Q)$ consists of a unique point $(1:0)$. Therefore, for each point $R \in C_1 \setminus \{Q\}$, the map $\pi_R \circ \varphi_1$ is ramified at $(1:0)$ with index $\geq (q-1)/2$. Assume that R is an inner Galois point. It follows from [8, III. 7.2] that $\pi_R \circ \varphi_1$ is ramified at $(1:0)$ with index $(q-1)/2$ or $q-1$. If the index at $(1:0)$ is $q-1$, then $R = P_2$. Assume that the index at $(1:0)$ is $(q-1)/2$. Then there exists a ramification point $\hat{S} \in \mathbb{P}^1$ with index $(q-1)/2$ such that $\varphi_1(\hat{S}) \neq Q$ and $\varphi_1(\hat{S}) \in \overline{RQ}$. Considering $\pi_Q \circ \varphi_1$, $\hat{S} = (0:1)$ or $(1:-1)$. If $\hat{S} = (0:1)$, then $R = \varphi_1(0:1) = P_1$, since $C_1 \cap \overline{P_1 Q} = \{P_1, Q\}$. Assume that $\hat{S} = (1:-1)$. Then $(q-1)/2 = 2$, and hence, $q = 5$ and $R \in \overline{Q\varphi_1(1:-1)}$. However, this is a contradiction, because the point $\varphi_1(1:-1) = (1:2:0) = \varphi_1(2:1)$ is a singular point and $C \cap \overline{Q\varphi_1(1:-1)} = \{Q, \varphi_1(1:-1)\}$. We complete the proof of Theorem 1(c). □

Proof of Theorem 2. Assertion (a) in Theorem 2 is similar to assertion (a) in Theorem 1. The composite map $\pi_{P_2} \circ \varphi_2$ is given by

$$(s:1) \mapsto \big((s-1)^{\frac{q+1}{2}} : s^q - s\big) = \big(1 : f(s-1)\big).$$

Then the induced field extension is $k(u)/k(f(u))$, where $u = s-1$. This is a Galois extension, similar to assertion (a) in Theorem 1. Assertion (b) in Theorem 2 follows.

We prove that the set of all inner Galois points for C_2 is equal to $\{P_1, P_2\}$. Let $Q := \varphi_2(1:0) = (0:0:1)$. Then the composite map $\pi_Q \circ \varphi_2$ is given by

$$(s:t) \mapsto \left(s^{\frac{q+1}{2}} t^{\frac{q-1}{2}} : (s-t)^{\frac{q+1}{2}} t^{\frac{q-1}{2}}\right) = \left(s^{\frac{q+1}{2}} : (s-t)^{\frac{q+1}{2}}\right).$$

Since the degree of $\pi_Q \circ \varphi_2$ is $(q+1)/2$, the point Q is a singular point of C_2 with multiplicity $(q-1)/2$. Further, by this expression, $\pi_Q \circ \varphi_2$ is a cyclic covering and all ramification points are $(0:1)$ and $(1:1)$ with index $(q+1)/2$. Note that $\varphi_2^{-1}(Q)$ consists of a unique point $(1:0)$, and the intersection multiplicity of $C_2 \cap L$ at Q is at most $\frac{q-1}{2} + 1$ for any line L passing through Q. Therefore, for each point $R \in C_2 \setminus \{Q\}$, the map $\pi_R \circ \varphi_2$ is ramified at $(1:0)$ with index $(q-1)/2$ or $(q+1)/2$. Assume that R is an inner Galois point. It follows from [8, III. 7.2] that $\pi_R \circ \varphi_2$ is ramified at $(1:0)$ with index $(q-1)/2$. Then there exists a ramification point $\hat{S} \in \mathbb{P}^1$ with index $(q-1)/2$ such that $\varphi_2(\hat{S}) \neq Q$ and $\varphi_2(\hat{S}) \in \overline{RQ}$. Considering $\pi_Q \circ \varphi_2$, $\hat{S} = (0:1)$ or $(1:1)$. Then $R = P_1$ or P_2, since $C_2 \cap \overline{QP_i} = \{Q, P_i\}$ for $i = 1, 2$. We complete the proof of Theorem 2(c). □

3 Proof of Theorems 3 and 4

Proof of Theorem 3. Let $Q := \varphi_3(1:0) = (0:1:1)$. Then the composite map $\pi_Q \circ \varphi_3$ is given by

$$(s:t) \mapsto \left(s^{\frac{q+1}{2}} t^{\frac{q+1}{2}} : (s+t)^{q+1} - \left(s^{q+1} + \gamma t^{q+1}\right)\right) = \left(s^{\frac{q+1}{2}} t^{\frac{q-1}{2}} : s^q + st^{q-1} + (1-\gamma)t^q\right).$$

Since $\gamma \neq 1$, the degree of $\pi_Q \circ \varphi_3$ is q, and hence, φ_3 is birational onto the image C_3.

We consider the point P_1. Since $\gamma \neq 0$, $P_1 \in \mathbb{P}^2 \setminus C_3$. The composite map $\pi_{P_1} \circ \varphi_3$ is given by

$$(s:1) \mapsto \left(s^{\frac{q+1}{2}} : s^{q+1} + \gamma\right) = \left(1 : s^{\frac{q+1}{2}} + \gamma s^{-\frac{q+1}{2}}\right).$$

The rational function $s^{\frac{q+1}{2}} + \gamma s^{-\frac{q+1}{2}}$ is invariant under the actions $s \mapsto \delta s^{-1}$ and $s \mapsto \zeta s$, where δ is a $(q+1)/2$-th root of γ and ζ is a $(q+1)/2$-th root of unity. Therefore, the extension $k(\mathbb{P}^1)/\pi_{P_1}^* k(\mathbb{P}^1)$ is a Galois extension, and hence, P_1 is a Galois point. In this case, $G_{P_1} \cong D_{q+1}$. Assertion (a) in Theorem 3 follows. We consider the point P_2. Since $\gamma \neq -1$, $P_2 \in \mathbb{P}^2 \setminus C_3$. The composite map $\pi_{P_2} \circ \varphi_3$ is given by

$$(s:1) \mapsto \left((s+1)^{q+1} : s^{q+1} + \gamma\right).$$

Note that

$$(1+\gamma)\left(s^{q+1} + \gamma\right) - \gamma(s+1)^{q+1} = (s-\gamma)^{q+1}.$$

Therefore, the extension $k(\mathbb{P}^1)/\pi_{P_2}^* k(\mathbb{P}^1)$ is a Galois extension, and hence, P_2 is a Galois point. In this case, $G_{P_2} \cong \mathbb{Z}/(q+1)\mathbb{Z}$. Assertion (b) in Theorem 3 follows.

We prove that the set of all outer Galois points for C_3 is equal to $\{P_1, P_2\}$. Note that the rank of the matrix

$$\begin{pmatrix} \varphi_3 \\ \frac{d\varphi_3}{ds} \end{pmatrix} = \begin{pmatrix} s^{\frac{q+1}{2}} & (s+1)^{q+1} & s^{q+1}+\gamma \\ \frac{q+1}{2}s^{\frac{q-1}{2}} & (s+1)^q & s^q \end{pmatrix}$$

is two for each s, that is, the differential map of φ_3 is injective at each point $(s:1) \in \mathbb{P}^1$. We consider the Hessian matrix H of φ_3:

$$\begin{pmatrix} \varphi_3 \\ \frac{d\varphi_3}{ds} \\ \frac{d^2\varphi_3}{ds^2} \end{pmatrix} = \begin{pmatrix} s^{\frac{q+1}{2}} & (s+1)^{q+1} & s^{q+1}+\gamma \\ \frac{q+1}{2}s^{\frac{q-1}{2}} & (s+1)^q & s^q \\ \frac{q^2-1}{4}s^{\frac{q-3}{2}} & 0 & 0 \end{pmatrix}.$$

Note that $\det H = 0$ if and only if $s = 0, -1, \gamma$. It follows that all flexes of C_3 are $(1:0)$, $(0:1)$, $(-1:1)$ and $(\gamma:1)$ (see [5, Section 7.6]). If R is an outer Galois point such that the order of G_R is at least five, then there exists a ramification point with index at least three (see, e. g., [5, Theorem 11.91]). Then R is contained in the tangent line at a flex. If $R \neq P_2$, then R is contained in the line defined by $X = 0$, which passes through points $\varphi_3(1:0)$, $\varphi_3(0:1)$ and P_1. It follows from [8, III. 7.2, 8.2] that there exist subgroups $H_1 \subset G_{P_1}$ and $H_2 \subset G_R$ of order $\frac{q+1}{2}$ fixing the point $(1:0)$. Then H_1 and H_2 are normal subgroups of G_{P_1} and G_R, respectively, and hence, they fix points $(1:0)$ and $(0:1)$. By a property of $\mathrm{PGL}(2,k)$, it follows that $H_1 = H_2$, that is, $G_{P_1} \cap G_R \neq \{1\}$. This is a contradiction (see, e. g., [1, Lemma 7]). We complete the proof of Theorem 3. □

Proof of Theorem 4. Let $Q := \varphi_4(1:0) = (0:1:1)$. Then the composite map $\pi_Q \circ \varphi_4$ is given by

$$(s:t) \mapsto \left(s^{\frac{q+1}{2}}t^{\frac{q+1}{2}} : (s+t)^{\frac{q+1}{2}}(s+\gamma t)^{\frac{q+1}{2}} - \left(s^{q+1} - \gamma t^{q+1}\right)\right).$$

Note that the coefficients of t^{q+1} and $s^q t$ for the function $(s+t)^{\frac{q+1}{2}}(s+\gamma t)^{\frac{q+1}{2}} - (s^{q+1} - \gamma t^{q+1})$ are $\gamma^{\frac{q+1}{2}} + \gamma = 2\gamma \neq 0$ and $\frac{q+1}{2}(\gamma+1) \neq 0$, respectively. The degree of $\pi_Q \circ \varphi_3$ is q, and hence, φ_4 is birational onto the image C_4.

We consider the point P_1. Since $\gamma \neq 0$, $P_1 \in \mathbb{P}^2 \setminus C_4$. The composite map $\pi_{P_1} \circ \varphi_4$ is given by

$$(s:1) \mapsto \left(s^{\frac{q+1}{2}} : s^{q+1} - \gamma\right) = \left(1 : s^{\frac{q+1}{2}} - \gamma s^{-\frac{q+1}{2}}\right).$$

The rational function $s^{\frac{q+1}{2}} - \gamma s^{-\frac{q+1}{2}}$ is invariant under the actions $s \mapsto \delta s^{-1}$ and $s \mapsto \zeta s$, where δ is a $(q+1)/2$-th root of $-\gamma$ and ζ is a $(q+1)/2$-th root of unity. Therefore, the extension $k(\mathbb{P}^1)/\pi_{P_1}^* k(\mathbb{P}^1)$ is a Galois extension, and hence, P_1 is a Galois point. In this case, $G_{P_1} \cong D_{q+1}$. Assertion (a) in Theorem 4 follows. We consider the point P_2. Since

$(-1)^{q+1}-\gamma \neq 0$ and $(-\gamma)^{q+1}-\gamma = \gamma(\gamma-1)^q \neq 0$, it follows that $P_2 \in \mathbb{P}^2 \setminus C_4$. The composite map $\pi_{P_2} \circ \varphi_4$ is given by

$$(s:1) \mapsto \left((s+1)^{\frac{q+1}{2}}(s+\gamma)^{\frac{q+1}{2}} : s^{q+1}-\gamma\right).$$

Note that

$$\frac{1}{1-\gamma}\left\{-\gamma(s+1)^{q+1}+(s+\gamma)^{q+1}\right\} = s^{q+1}-\gamma.$$

Let

$$u := \frac{s+\gamma}{s+1} \quad \text{and} \quad h(u) := -\gamma u^{-\frac{q+1}{2}} + u^{\frac{q+1}{2}}.$$

Then $k(\mathbb{P}^1)/\pi_{P_2}^* k(\mathbb{P}^1) = k(u)/k(h(u))$. This is a Galois extension, similar to assertion (a) in Theorem 4. In this case, $G_{P_2} \cong D_{q+1}$. Assertion (b) in Theorem 4 follows. □

4 Application of the criterion by the first author to group theory

Proof of Theorem 5. Assume condition (1). We can assume that $n \geq 5$. It follows from [3, Theorem 1] that there exists a morphism

$$\varphi : \mathbb{P}^1 \to \mathbb{P}^2$$

of degree $|H_1|+1 = n+1$ and different inner Galois points $\varphi(P_1)$ and $\varphi(P_2)$ exist for $\varphi(\mathbb{P}^1)$ such that φ is birational onto $\varphi(\mathbb{P}^1)$ and $G_{\varphi(P_i)} = H_i$ for $i = 1, 2$. Note that H_i has exactly two fixed points on $\mathbb{P}^1$. Similar to the proof of a theorem [6, Theorem 1] of Miura, $\varphi(\mathbb{P}^1)$ does not have a cusp and the inequality

$$2 \times (n-2) \times \frac{2n-2}{n-1} \leq 3(n+1) - 6$$

holds for the number of flexes (see, e. g., [5, Theorem 7.55]). This implies $n \leq 5$.

Assume condition (2). For $n = 4, 5$, the existence of a pair of cyclic groups with (a) and (b) in Problem 1 is proved in [3, Theorem 2(1)(4)]. For $n = 3$, it is proved by [6, Example 1] and [3, Theorem 1]. □

Remark 1.

(1) Problem 1 can be generalized to the case where k is any field.

(2) As a corollary of Theorem 5, for any field k of characteristic zero, the same claim "(1) ⇒ (2)" as in Theorem 5 is proved.

Remark 2. In the case where $p > 0$, the following examples of pairs (H_1, H_2) with (a) and (b) in Problem 1 are already known, according to the existence of two Galois points ([4, 2]) and the criterion:
(1) $H_1 \cong H_2 \cong \mathbb{Z}/(q-1)\mathbb{Z}$.
(2) $H_1 \cong H_2 \cong (\mathbb{Z}/p\mathbb{Z})^{\oplus e}$, where $e \geq 1$ is an integer.

Bibliography

[1] S. Fukasawa, Classification of plane curves with infinitely many Galois points, *J. Math. Soc. Jpn.*, **63** (2011), 195–209.
[2] S. Fukasawa, Galois points for a non-reflexive plane curve of low degree, *Finite Fields Appl.*, **23** (2013), 69–79.
[3] S. Fukasawa, A birational embedding of an algebraic curve into a projective plane with two Galois points, *J. Algebra*, **511** (2018), 95–101.
[4] S. Fukasawa and T. Hasegawa, Singular plane curves with infinitely many Galois points, *J. Algebra*, **323** (2010), 10–13.
[5] J. W. P. Hirschfeld, G. Korchmáros, and F. Torres, Algebraic Curves over a Finite Field, Princeton Univ. Press, Princeton, 2008.
[6] K. Miura, Galois points on singular plane quartic curves, *J. Algebra*, **287** (2005), 283–293.
[7] K. Miura and H. Yoshihara, Field theory for function fields of plane quartic curves, *J. Algebra*, **226** (2000), 283–294.
[8] H. Stichtenoth, Algebraic Function Fields and Codes, Universitext, Springer-Verlag, Berlin, 1993.
[9] H. Yoshihara, Function field theory of plane curves by dual curves, *J. Algebra*, **239** (2001), 340–355.
[10] H. Yoshihara, Galois points for plane rational curves. *Far East J. Math.*, **25** (2007), 273–284; Errata, ibid. **29** (2008), 209–212.
[11] H. Yoshihara and S. Fukasawa, List of problems, available at: http://hyoshihara.web.fc2.com/openquestion.html.

Charles J. Colbourn and Violet R. Syrotiuk

Covering strong separating hash families

Abstract: In principle, detecting arrays provide test suites with few tests for complex engineered systems with many interacting factors, each taking on one of a small set of levels, provided that there are few significant interactions, each involving few factors. In practice, explicit construction of detecting arrays for real-world testing remains challenging. To address this, covering strong separating hash families, which are certain arrays whose entries are vectors over a finite field, are introduced. Covering strong separating hash families are shown to underlie a very compact representation for detecting arrays for a variety of testing scenarios. The consequences of this compact representation for asymptotic existence bounds, and prospects for constructive algorithms, are discussed.

Keywords: Covering array, detecting array, perfect hash family, strong separating hash family

MSC 2010: 05B15, 05B40, 11T06, 51E20

1 Introduction

Our concern is with the effective construction of test suites for large complex engineered systems. Combinatorial arrays for this purpose have been defined [15], falling in a general category of *arrays meeting t-restrictions* [3]. Among the most widely studied t-restrictions are hash families involving "separation" of t-sets of columns and covering arrays involving "coverage" of t-sets. Although initially pursued for different applications, hash families have been widely used in the construction of variants of covering arrays [12]. In this paper, we strengthen a connection between hash families and detecting arrays, both constructed over a finite field $\mathbb{F}_q$. We start by providing a substantial amount of context and formal definitions for both.

1.1 Hash families

A *perfect hash family* $\mathsf{PHF}(N; k, v, t)$ is an $N \times k$ array on v symbols, in which in every $N \times t$ subarray, at least one row consists of distinct symbols. Perfect hash families were

Acknowledgement: This work is supported in part by the U.S. National Science Foundation grants No. 1421058 and No. 1813729, and in part by the Software Test and Analysis Techniques for Automated Software Test program by OPNAV N-84, U.S. Navy. Thanks to Randy Compton for initial discussions to formulate the combinatorial objects discussed here.

Charles J. Colbourn, Violet R. Syrotiuk, School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ, USA, e-mails: Charles.Colbourn@asu.edu, syrotiuk@asu.edu

https://doi.org/10.1515/9783110621730-013

introduced as a tool for storage and retrieval of frequently used information [32]. Perfect hash families arise in the construction of cover-free families, separating systems, group testing algorithms, and secure frameproof codes (for example, in [7, 41, 42]), and in numerous cryptographic applications such as [9, 24, 6]. An older survey on PHFs is given in [20]; a more recent survey appears in [22]. In the latter, the emphasis is on a generalization requiring more separation: A $\mathsf{PHF}_\delta(N; k, v, t)$ is an $N \times k$ array on v symbols, in which in every $N \times t$ subarray, at least δ rows consist of distinct symbols.

A different generalization, discussed in [2, 36, 40], for example, follows; we extend the usual definition from $\delta = 1$ to arbitrary values of δ. A *separating hash family*, $\mathsf{SHF}_\delta(N; k, v, \{w_1, w_2, \ldots, w_t\})$, is an $N \times k$ array on v symbols so that for any $C_1, C_2, \ldots, C_t \subseteq \{1, 2, \ldots, k\}$ such that $|C_1| = w_1$, $|C_2| = w_2$, ..., $|C_t| = w_t$, and $C_i \cap C_j = \emptyset$ for every $i \neq j$, there exist at least δ rows in which whenever $c \in C_i$, $c' \in C_j$, and $i \neq j$, the entries in columns c and c' are not the same. The multiset $\{w_1, w_2, \ldots, w_t\}$ is the *type* of the SHF, and it is often written in exponential notation: $(s_1)^{a_1} \cdots (s_\ell)^{a_\ell}$ indicates that there are a_i classes of size s_i for each $1 \leq i \leq \ell$. Separating hash families of type $1^t d^1$ (i. e., $\mathsf{SHF}_\delta(N; k, v, 1^t d^1)$) are *strong* SHFs; existence results and bounds have been examined more thoroughly in this special case, at least when $\delta = 1$ [25, 26, 34].

1.2 Arrays for interaction testing

Combinatorial testing [28, 33] addresses the design and analysis of test suites to evaluate correctness and performance of complex engineered systems. There are k factors $F_1, \ldots, F_k$. Each factor F_i has a set $S_i = \{v_{i1}, \ldots, v_{is_i}\}$ of s_i possible *levels*. A *test* is a k-tuple $(v_1, \ldots, v_k)$ with $v_i \in S_i$ for $1 \leq i \leq k$. A *test suite* is a collection of tests. Each test, when executed, provides a *response*. When $\{i_1, \ldots, i_t\} \subseteq \{1, \ldots, k\}$ and $\sigma_{i_j} \in S_{i_j}$, the set $\{(i_j, \sigma_{i_j}) : 1 \leq j \leq t\}$ is a *t-way interaction*. The *strength* of the interaction is t. Each test *covers* $\binom{k}{t}$ t-way interactions. A test suite can be viewed as an $N \times k$ array $A = (\sigma_{i,j})$ in which $\sigma_{i,j} \in S_j$ when $1 \leq i \leq N$ and $1 \leq j \leq k$. The test suite has *size* N and *type* $(s_1, \ldots, s_k)$.

Let $A = (\sigma_{i,j})$ be a test suite of size N and type $(s_1, \ldots, s_k)$. Let $T = \{(i_j, \sigma_{i_j}) : 1 \leq j \leq t\}$ be a t-way interaction. Denote the set $\{r : a_{ri_j} = \sigma_{i_j}, 1 \leq j \leq t\}$ of rows of A in which T is covered by $\rho_A(T)$. Extending to a set $\mathcal{T}$ of interactions, $\rho_A(\mathcal{T}) = \bigcup_{T \in \mathcal{T}} \rho_A(T)$. When $\rho_A(T) = \emptyset$, the potential effect of T cannot be observed, so one normally insists that $|\rho_A(T)| \geq 1$ (and hence that the test suite be a covering array, defined soon). This requirement does not suffice to determine which interaction(s) affect the response significantly, so further requirements were developed in [15, 16, 30], and extended to treat issues in practical testing in [1, 19, 37, 38].

Now we define the test suites of interest here. Let A be a test suite of size N and type $(s_1, \ldots, s_k)$. Let $\mathcal{I}_t$ be the set of all t-way interactions for A. An $N \times k$ array A of type $(s_1, \ldots, s_k)$ is *(d, t, δ)-detecting* if $|\rho_A(T) \setminus \rho_A(\mathcal{T})| < \delta \Leftrightarrow T \in \mathcal{T}$ whenever $\mathcal{T} \subseteq \mathcal{I}_t$, and $|\mathcal{T}| = d$. The notation $\mathsf{DA}_\delta(N; d, t, k, (s_1, \ldots, s_k))$ indicates that the array is *mixed*.

The array is *uniform* when all factors have the same number v of levels, and the notation simplifies to $\mathsf{DA}_\delta(N; d, t, k, v)$. The value of δ is the *separation* [38] (the original definition [15] sets $\delta = 1$). Rows in $\rho_A(T) \setminus \rho_A(\mathcal{T})$ are *witnesses* for T with respect to $\mathcal{T}$.

When $d = 0$, $\mathcal{T} = \emptyset$ and $\rho_A(\emptyset) = \emptyset$, and so a $\mathsf{DA}_\delta(N; 0, t, k, (s_1, \ldots, s_k))$ is a *covering array* $\mathsf{CA}_\delta(N; t, k, (s_1, \ldots, s_k))$. The notation $\mathsf{CA}_\delta(N; t, k, v)$ is used when the type is uniform. Covering arrays have been extensively studied and applied in testing, primarily when $\delta = 1$ [10, 12, 27, 28, 33].

To develop a final parameter of concern, we follow the presentation in [17]. The motivation arises from the desire to make mixed detecting arrays from uniform ones by (repeatedly) replacing two levels of a factor by one (i. e., by *fusion*). Any sequence of fusions applied to a covering array yields a covering array, but for detecting arrays with $d \geq 1$ this need not be true. Let A be a $\mathsf{DA}_\delta(N; d, t, k, (s_1, \ldots, s_k))$. Let $T = \{(i_j, \sigma_{i_j}) : 1 \leq j \leq t\}$ be a t-way interaction for A. Let $D = \{a_i : 1 \leq i \leq d\}$ be a set of d column indices of A with $\{i_1, \ldots, i_t\} \cap \{a_1, \ldots, a_d\} = \emptyset$. A set system $\mathcal{S}_{A,T,D}$ on the ground set $\{(c, \sigma) : c \in D, \sigma \in S_c\}$ contains the collection of sets

$$\{\{(a_1, v_1), \ldots, (a_d, v_d)\} : T \cup \{(a_1, v_1), \ldots, (a_d, v_d)\} \text{ is covered in } A\}.$$

If after carrying out fewer than s fusions within each column in D, no set X of elements of $\mathcal{S}_{A,T,D}$ with $|X| \leq d$ can be deleted so as to leave fewer than δ sets, then (T, D) has *corroboration* s in A. (Note that the corroboration is defined with respect to a specified separation δ.) When every choice of (T, D) has corroboration (at least) s, $\mathsf{DA}_\delta(N; d, t, k, (s_1, \ldots, s_k))$ has *corroboration* s, and is denoted by $\mathsf{DA}_\delta(N; d, t, k, (s_1, \ldots, s_k), s)$.

When $d = t = 1$, detecting arrays can be constructed via a correspondence with a "Sperner partition system" [8, 29, 31]. When $t = 1$ and $d \geq 1$, probabilistic and algorithmic methods are explored in [17] using an algebraic approach that we generalize in this paper. A related algebraic approach when $t = 2$, $k = 2q$, and $d \geq 1$ is introduced in [18]. See [16] for pointers to other work on detecting arrays.

1.3 Weaving the threads together

Naturally, one might ask how the well-studied areas of hash families and arrays for interaction testing meet. The use of hash families in constructing covering arrays and related objects has been well studied [12, 22]: Using hash families whose symbols are in one-to-one correspondence with the columns of a covering or detecting array, one can replace each symbol of the hash family with the corresponding column of the interaction testing array to construct a larger testing array. But what is the role of finite fields?

In Section 2, we explore a further connection, by describing an analogue of perfect hash families whose entries are vectors over a finite field. We outline how these generate covering arrays. Then in Section 3 we generalize this approach to define covering

strong separating hash families, and demonstrate that they yield detecting arrays. Furthermore, they provide effective mechanisms to obtain specified separation and corroboration. In Section 4, we employ probabilistic methods to obtain bounds for the existence of such hash families, and hence also bounds on the sizes of detecting arrays.

2 Covering perfect hash families and covering arrays

Sherwood et al. [39] developed a construction of covering arrays using an analogue of perfect hash families whose entries are vectors over a finite field. We employ a generalization of their definitions from [14]. Let q be a prime power, and let $\mathbb{F}_q$ be the finite field of order q. Let $\mathcal{R}_{t,q} = \{\mathbf{r}_0, \ldots, \mathbf{r}_{q^t-1}\}$ be the set of all (row) vectors of length t with entries from $\mathbb{F}_q$, and let $\mathcal{T}_{t,q}$ be the set of all column vectors of length t with entries from $\mathbb{F}_q$, not all 0. As noted in [39], when $\mathcal{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_t\}$ satisfies $\mathbf{x}_i \in \mathcal{T}_{t,q}$ for $1 \le i \le t$, the array $A = (a_{ij})$ in which $a_{ij} = \mathbf{r}_i\mathbf{x}_j$ is a $\mathsf{CA}(q^t; t, t, q)$ if and only if the $t \times t$ matrix $X = [\mathbf{x}_1 \cdots \mathbf{x}_t]$ is nonsingular.

For any nonzero $\mu \in \mathbb{F}_q$, substituting $\mu\mathbf{x}_i$ for $\mathbf{x}_i$ does not alter the non-singularity. Define $\langle \mathbf{x} \rangle = \{\mu\mathbf{x} : \mu \in \mathbb{F}_q, \mu \neq 0\}$. Let $\mathcal{V}_{t,q}$ be a set of representatives of the column vectors in $\mathcal{T}_{t,q}$. Then $|\mathcal{V}_{t,q}| = \frac{q^t-1}{q-1} = \sum_{i=0}^{t-1} q^i$.

A *covering perfect hash family* $\mathsf{CPHF}(n; k, q, t)$ is an $n \times k$ array $C = (\mathbf{c}_{ij})$ with entries from $\mathcal{V}_{t,q}$ so that, for every set $\{\gamma_1, \ldots, \gamma_t\}$ of distinct column indices, there is at least one row index ρ of C for which $[\mathbf{c}_{\rho\gamma_1} \cdots \mathbf{c}_{\rho\gamma_t}]$ is nonsingular; this is a *covering t-set* and the t-set of columns is *covered*.

Suppose that C is a $\mathsf{CPHF}(n; k, q, t)$. Then there exists a $\mathsf{CA}(n(q^t - 1) + 1; t, k, q)$. The proof is straightforward [14]: Replace each entry $\mathbf{c}_{ij}$ of C by the column vector obtained by multiplying $\mathbf{c}_{ij}$ by each $\mathbf{r}_\ell \in \mathcal{R}_{t,q}$. This produces a $\mathsf{CA}(nq^t; t, k, q)$; because the all-zero row appears (at least) n times, $n - 1$ copies can be removed.

Because this CPHF construction of covering arrays employs a compact representation of certain covering arrays as covering perfect hash families, in practice it makes feasible the explicit construction of covering arrays for "large" k and t [14, 43, 44]. Surprisingly, the imposition of such structure yields covering array numbers among the best known at present. Indeed CPHFs lead to the best current asymptotic existence upper bounds for covering array numbers with fixed strength t [14, 21] and to efficient (and practical) algorithms to construct covering arrays realizing the given bounds [13, 14]. Specific constructions of CPHFs also arise from constructions in projective geometries [35, 45].

The success of this approach in constructing covering arrays from hash families leads to natural questions: Can we generalize CPHFs to construct detecting arrays? Can we control the separation and corroboration? In the first exploration of detecting arrays [15], it was shown that a covering array $\mathsf{CA}(N; t + d, k, v)$ is a $\mathsf{DA}(N; d, t, k, v)$ provided that $v > d$. This generalizes: When $v > d$, a $\mathsf{CA}_\delta(N; t + d, k, v)$, A, is a

$\mathsf{DA}_{\delta(v-d)v^{d-1}}(N;d,t,k,v)$. The corroboration of A can be determined by noting that when $T = \{(i_j, \sigma_{i_j}) : 1 \le j \le t\}$ is a t-way interaction for A, and $D = \{a_i : 1 \le i \le d\}$ is a set of d column indices of A with $\{i_1, \ldots, i_t\} \cap \{a_1, \ldots, a_d\} = \emptyset$, $\mathcal{S}_{A,T,D}$ contains each of the possible v^d sets at least δ times. When $1 \le s < v - d - 1$, it follows that A is a $\mathsf{DA}_{\delta(v-d-s-1)(v-s)^{d-1}}(N;d,t,k,v,s)$.

At first it appears that we have what we sought, a method to make detecting arrays with larger separation and corroboration. But a closer look reveals that when d and v are large, the route via covering arrays results in far too many rows. This leads us finally to our central direction, to generalize the notion of CPHFs to make detecting arrays directly. We treat this next.

3 Covering strong separating hash families

Let $d \ge 0$ and $t \ge 2$. As before, let $\mathcal{R}_{t+1,q} = \{\mathbf{r}_0, \ldots, \mathbf{r}_{q^{t+1}-1}\}$ be the set of all (row) vectors of length $t+1$ with entries from $\mathbb{F}_q$. For a subset S of $\mathbb{F}_q$ of cardinality g, let $\mathcal{R}_{t+1,q,g}$ be the set of all vectors in $\mathcal{R}_{t+1,q}$ whose last entry belongs to S. Let $\mathcal{W}_{t+1,q}$ be a set of representatives of all column vectors of length $t+1$ with entries from $\mathbb{F}_q$, so that at least one of the first t coordinates is nonzero. (The latter condition eliminates precisely one vector from $\mathcal{V}_{t+1,q}$.) When $\mathbf{v} \in \mathcal{W}_{t+1,q}$, denote by $\partial(\mathbf{v})$ the column vector of length t obtained by deleting the last entry in $\mathbf{v}$.

A *covering strong separating hash family* $\mathsf{CSSHF}_\delta(n;k,q,d,t)$ is an $n \times k$ array $C = (\mathbf{c}_{ij})$ with entries from $\mathcal{W}_{t+1,q}$ so that, for every set $\{\gamma_1, \ldots, \gamma_t\} \cup \{a_1, \ldots, a_d\}$ of distinct column indices, there are at least δ row indices ρ of C for which:

1. $[\partial(\mathbf{c}_{\rho\gamma_1}) \cdots \partial(\mathbf{c}_{\rho\gamma_t})]$ is nonsingular, and
2. $[\mathbf{c}_{\rho\gamma_1} \cdots \mathbf{c}_{\rho\gamma_t} \mathbf{c}_{\rho a_i}]$ is nonsingular whenever $1 \le i \le d$.

Some easy consequences of the definition follow.

Proposition 3.1. *Let $C = (\mathbf{c}_{ij})$ be an $n \times k$ array with entries from $\mathcal{W}_{t+1,q}$. Then:*

1. *When $d \ge 1$, C is a $\mathsf{CSSHF}_\delta(n;k,q,d-1,t)$ if C is a $\mathsf{CSSHF}_\delta(n;k,q,d,t)$, but the converse need not hold;*
2. *When $d \ge 1$, C is a $\mathsf{CPHF}_\delta(n;k,q,t+1)$ if C is a $\mathsf{CSSHF}_\delta(n;k,q,d,t)$, but the converse need not hold; and*
3. *C is a $\mathsf{CSSHF}_\delta(n;k,q,0,t)$ if and only if $\partial(C) = [\partial(\mathbf{c}_{ij})]$ is a $\mathsf{CPHF}_\delta(n;k,q,t)$.*

Theorem 3.2. *Let q be a prime power, and let t, d, and k be integers with $t+d \le k$ and $1 \le d < q$. Whenever a $\mathsf{CSSHF}_\delta(n;k,q,d,t)$ exists and $d < g \le q$, a $\mathsf{DA}_{\delta(g-d)}(ngq^t;d,t,k,q)$ exists.*

Proof. Let $\mathcal{R} = \mathcal{R}_{t+1,q,g} = \{\mathbf{r}_1, \ldots, \mathbf{r}_{gq^t}\}$ be a set of gq^t row vectors. Let $C = (\mathbf{c}_{ij})$ be a $\mathsf{CSSHF}_\delta(n;k,q,d,t)$. Form a $ngq^t \times k$ array A with rows indexed by $\{1, \ldots, gq^t\} \times \{1, \ldots, n\}$

and columns indexed by $\{1, \ldots, k\}$. In the cell in row (σ, ρ) and column κ of A, place the entry $\mathbf{r}_\sigma \mathbf{c}_{\rho,\kappa}$. Then A has ngq^t rows, k columns, and q symbols. Denote by A_ρ the $gq^t \times k$ array generated by row ρ of C. We claim that $A = \begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix}$ is a $\mathsf{DA}_{\delta(g-d)}(ngq^t; d, t, k, q)$.

To verify this, let T be a t-way interaction $\{(\gamma_1, \nu_1), \ldots, (\gamma_t, \nu_t)\}$. Suppose that $\mathcal{T} = \{T_1, \ldots, T_d\}$ is a set of d other t-way interactions. Without loss of generality, we can assume that no T_i is on exactly the same columns as T, so let c_i be a column index of T_i that does not appear in T for $1 \le i \le d$. (We do not assume that $\{c_1, \ldots, c_d\}$ are all distinct.) We must show that $|\rho_A(T) \setminus \rho_A(\mathcal{T})| \ge \delta(g - d)$. There is a set of δ row indices $\{\ell_1, \ldots, \ell_\delta\}$ of C so that for each $\ell \in \{\ell_1, \ldots, \ell_\delta\}$, $[\partial(\mathbf{c}_{\ell\gamma_1}) \cdots \partial(\mathbf{c}_{\ell\gamma_t})]$ is nonsingular, and $[\mathbf{c}_{\ell\gamma_1} \cdots \mathbf{c}_{\ell\gamma_t} \mathbf{c}_{\ell c_i}]$ is nonsingular whenever $1 \le i \le d$. When $\ell \in \{\ell_1, \ldots, \ell_\delta\}$, it suffices to show that $|\rho_{A_\ell}(T) \setminus \rho_{A_\ell}(\mathcal{T})| \ge g - d$.

Because $[\partial(\mathbf{c}_{\ell\gamma_1}) \cdots \partial(\mathbf{c}_{\ell\gamma_t})]$ is nonsingular, A_ℓ contains exactly g rows that cover T, as a consequence of the fact that there is a unique solution $\mathbf{r}$ to

$$\mathbf{r}[\partial(\mathbf{c}_{\ell\gamma_1}) \cdots \partial(\mathbf{c}_{\ell\gamma_t})] = (\nu_1, \ldots, \nu_t),$$

each leading to g solutions for $\mathcal{R}$. For each column c_i, the g rows that cover T in A_ℓ contain g distinct symbols in column c_i, as a consequence of the fact that $[\mathbf{c}_{\ell\gamma_1} \cdots \mathbf{c}_{\ell\gamma_t} \mathbf{c}_{\ell c_i}]$ is nonsingular. Hence $\mathcal{T}$ can cover at most d of the g rows in A_ℓ, and we have the desired conclusion. □

Increased separation can arise both by selecting larger δ or by increasing g (when possible). To address corroboration, consider a t-way interaction T and a set $D = \{a_1, \ldots, a_d\}$ of d other (distinct) columns. In the construction, $\mathcal{S}_{A,T,D}$ is a collection of (at least) δg sets of size d. But more is true. The g sets arising from a particular row ℓ of C are disjoint (they form a partial parallel class). Hence, no matter how f fusions are performed in each column of D, there remain at least $g - fd$ disjoint sets from each partial parallel class. It follows that whenever $0 \le s \le \lfloor \frac{g-d-1}{d} \rfloor$, a $\mathsf{DA}_{\delta(g-(s+1)d)}(ngq^t; d, t, k, q, s)$ exists.

These observations concerning corroboration appear to be most effective when d is small. However, consider the specific case of a $\mathsf{CSSHF}_1(n; k, q, 1, t)$. Whenever $s + \alpha \le q - 1$, we can form a $\mathsf{DA}_\alpha(n(1 + s + \alpha)q^t; 1, t, k, q, s)$. Hence in this case we can increase corroboration by one, while decreasing separation by one, without changing the detecting array.

4 Existence of CSSHFs

Theorem 3.2 is useful only when suitable CSSHFs can be found. What does suitable mean here? Our concern is to construct detecting arrays with few rows and many

columns, to support testing for many factors with few tests. Earlier research on the more restrictive situation for CPHFs [14] establishes asymptotic existence results using probabilistic methods and employs these to provide efficient algorithms for the generation of CPHFs and covering arrays. These approaches lead not only to the best known asymptotic existence results for covering arrays, but also to many instances of the fewest rows known for covering arrays for practical sizes [11, 14].

Here, we explore first steps to construct CSSHFs, in order to provide a foundation for probabilistic and algorithmic methods, Choose elements of an $n\times k$ array uniformly at random from $\mathcal{W}_{t+1,q}$. Now consider a set T of t distinct columns, and a set D of d distinct columns disjoint from T. Consider the selections on columns in T one at a time (in any order), ensuring that after σ columns are treated, the $t\times\sigma$ matrix thus far has rank σ. The σ columns generate a space of q^σ columns, none of which can be adjoined while maintaining nonsingularity. Hence the next column treated has probability $\frac{q^t-q^\sigma}{q^t-1}$ of maintaining nonsingularity. Once the $t\times t$ matrix on T is ensured to be nonsingular, only the last coordinates in the vectors chosen on D matter, and each can be treated independently. For each one, the entry chosen in the last coordinate leads to nonsingularity of the corresponding $(t+1)\times(t+1)$ matrix with probability $\frac{q-1}{q}$ (only one choice from $\mathbb{F}_q$ can make the matrix singular). Hence the probability that both conditions are met for T and D in a single row is

$$\psi_{d,t,q} = \left(\frac{\prod_{i=0}^{t-1}(q^t-q^i)}{(q^t-1)^t}\right)\left(\frac{q-1}{q}\right)^d.$$

Because rows are selected independently, the probability that the conditions are met *fewer than* δ times within n rows is

$$\sum_{i=0}^{\delta-1}\binom{n}{i}(1-\psi_{d,t,q})^i(\psi_{d,t,q})^{n-i}.$$

Using linearity of expectations, the basic probabilistic method ensures that when

$$\binom{k}{t+d}\binom{t+d}{d}\left[\sum_{i=0}^{\delta-1}\binom{n}{i}(1-\psi_{d,t,q})^i(\psi_{d,t,q})^{n-i}\right] < 1,$$

a $\mathsf{CSSHF}_\delta(n;k,q,d,t)$ exists. Equivalently,

$$\binom{k}{t+d}\binom{t+d}{d}(\psi_{d,t,q})^n\left[\sum_{i=0}^{\delta-1}\binom{n}{i}\left(\frac{1-\psi_{d,t,q}}{\psi_{d,t,q}}\right)^i\right] < 1. \tag{4.1}$$

This generalizes the statement for CPHFs (with $d=0$ and $\delta=1$) [14], in which we have

$$\binom{k}{t}(\psi_{0,t,q})^n < 1.$$

Fix d, t, q, and δ. Then taking logarithms in (4.1) explicitly determines a constant c (depending only on the fixed values of d, t, q, and δ) so that $n \leq c \log k + o(\log k)$. (There does not seem to be an easily stated expression for the constant c, except in the simplest cases.) With this in hand, improvements using the Lovász local lemma [4, 23] or oversampling (postprocessing [5]) can be obtained, analogous to the methods for CPHFs [14]. We leave this extension to subsequent work.

5 Concluding remarks

The explicit construction of detecting arrays for a wide variety of parameter sets can be useful in testing applications, but previous algorithmic techniques have been effectively limited to treating small values of the parameters. Here, we have shown that by using arithmetic over the finite field, CSSHFs provide a remarkably compact representation for uniform detecting arrays with different values of t, d, and k, at least when the number of symbols is a prime power. CSSHFs are easily extended to support larger separation. Practical concerns dictate the need for constructing mixed detecting arrays, not just uniform ones. We have outlined how larger corroboration in a detecting array can support fusion of levels to produce mixed arrays, and shown that CSSHFs can provide larger corroboration. This provides a means to make detecting arrays that are nearly uniform, not a general technique for handling all distributions of numbers of levels of factors. When some factors have few levels and some have many, different approaches are needed.

The basic probabilistic analysis developed here provides a means to obtain a useful upper bound on the number of tests in a detecting array. Algorithms to realize the sizes given by the probabilistic bounds provide explicit, efficient construction methods. In subsequent work, we plan to explore the effectiveness of these algorithms for a wide range of parameters.

Having reduced the problem of finding uniform detecting arrays for a broad class of parameters to that of finding CSSHFs, we resorted to probabilistic methods to establish the existence of the CSSHFs. We do not expect that such probabilistic methods lead to arrays that are optimal (fewest rows for a specified number of columns), despite their being within a constant factor of the optimal number of rows. Hence it would be of substantial interest to find explicit constructions of CSSHFs using the algebra of the finite field or the associated projective geometry, particularly if these outperform the probabilistic guarantee, and provided that they can be explicitly and efficiently generated.

Bibliography

[1] A. N. Aldaco, C. J. Colbourn, and V. R. Syrotiuk, Locating arrays: a new experimental design for screening complex engineered systems, *SIGOPS Oper. Syst. Rev.*, **49**(1) (2015) 31–40.

[2] N. Alon, G. Cohen, M. Krivelevich, and S. Litsyn, Generalized hashing and parent-identifying codes, *J. Comb. Theory, Ser. A*, **104** (2003), 207–215.

[3] N. Alon, D. Moshkovitz, and S. Safra, Algorithmic construction of sets for k-restrictions, *ACM Trans. Algorithms*, **2** (2006), 153–177.

[4] N. Alon and J. H. Spencer, The Probabilistic Method, John Wiley & Sons, Inc., Hoboken, NJ, 2008.

[5] E. van den Berg, E. Candès, G. Chinn, C. Levin, P. D. Olcott, and C. Sing-Long, Single-photon sampling architecture for solid-state imaging sensors, *Proc. Natl. Acad. Sci.*, **110**(30) (2013), E2752–E2761.

[6] S. R. Blackburn, M. Burmester, Y. Desmedt, and P. R. Wild, Efficient multiplicative sharing schemes, *Lect. Notes Comput. Sci.*, **1070** (1996), 107–118.

[7] D. Boneh and J. Shaw, Collusion-secure fingerprinting for digital data, *IEEE Trans. Inf. Theory*, **44** (1998), 1897–1905.

[8] Y. Chang, C. J. Colbourn, A. Gowty, D. Horsley, and J. Zhou, New bounds on the maximum size of Sperner partition systems. *Eur. J. Comb.* (to appear).

[9] B. Chor, A. Fiat, M. Naor, and B. Pinkas, Tracing traitors, *IEEE Trans. Inf. Theory*, **46** (2000), 893–910.

[10] C. J. Colbourn, Combinatorial aspects of covering arrays, *Matematiche*, **58** (2004), 121–167.

[11] C. J. Colbourn, Covering array tables: $2 \le v \le 25$, $2 \le t \le 6$, $t \le k \le 10000$ (2005-19), www.public.asu.edu/~ccolbou/src/tabby.

[12] C. J. Colbourn, Covering arrays and hash families, In: Information Security and Related Combinatorics, NATO Peace and Information Security, IOS Press, 2011, pp. 99–136.

[13] C. J. Colbourn and E. Lanus, Subspace restrictions and affine composition for covering perfect hash families, *Art Discrete Appl. Math.*, **1** (2018), #P02.03.

[14] C. J. Colbourn, E. Lanus, and K. Sarkar, Asymptotic and constructive methods for covering perfect hash families and covering arrays, *Des. Codes Cryptogr.*, **86** (2018), 907–937.

[15] C. J. Colbourn and D. W. McClary, Locating and detecting arrays for interaction faults, *J. Comb. Optim.*, **15** (2008), 17–48.

[16] C. J. Colbourn and V. R. Syrotiuk, On a combinatorial framework for fault characterization, *Math. Comput. Sci.*, **12**(4) (2018), 429–451.

[17] C. J. Colbourn and V. R. Syrotiuk, Detecting arrays for main effects, *Lect. Notes Comput. Sci.*, **11545** (2019), 112–123.

[18] C. J. Colbourn and V. R. Syrotiuk, There must be fifty ways to miss a cover, In: 50 Years of Combinatorics, Graph Theory, and Computing, Ed. by F. Chung, R. L. Graham, F. Hoffman, R. C. Mullin, L. Hogben, and D. B. West, Chapman and Hall/CRC Press, 2019, pp. 319–334.

[19] R. Compton, M. T. Mehari, C. J. Colbourn, E. De Poorter, and V. R. Syrotiuk, Screening interacting factors in a wireless network testbed using locating arrays, In: IEEE INFOCOM International Workshop on Computer and Networking Experimental Research Using Testbeds (CNERT), 2016.

[20] Z. J. Czech, G. Havas, and B. S. Majewski, Perfect hashing, *Theor. Comput. Sci.*, **182** (1997), 1–143.

[21] S. Das and T. Mészáros, Small arrays of maximum coverage, *J. Comb. Des.*, **26**(10) (2018), 487–504.

[22] R. E. Dougherty and C. J. Colbourn, Perfect hash families: the generalization to higher indices, In: Discrete Mathematics and Applications, Springer, to appear. https://doi.org/10.1007/978-3-030-55857-4_7.

[23] P. Erdős and L. Lovász, Problems and results on 3-chromatic hypergraphs and some related questions, In: Infinite and Finite Sets, North-Holland, Amsterdam, 1975, pp. 609–627.
[24] A. Fiat and M. Naor, Broadcast encryption, *Lect. Notes Comput. Sci.*, **773** (1994), 480–491.
[25] C. Guo and D. R. Stinson, A tight bound on the size of certain separating hash families, *Australas. J. Comb.*, **67** (2017), 294–303.
[26] C. Guo, D. R. Stinson, and Tran Van Trung, On tight bounds for binary frameproof codes, *Des. Codes Cryptogr.*, **77**(2–3) (2015), 301–319.
[27] A. Hartman, Software and hardware testing using combinatorial covering suites, In: Interdisciplinary Applications of Graph Theory, Combinatorics, and Algorithms, Ed. by M. C. Golumbic, I. B. A. Hartman, Springer, Norwell, MA, 2005, pp. 237–266.
[28] D. R. Kuhn, R. Kacker, and Y. Lei, Introduction to Combinatorial Testing, CRC Press, 2013.
[29] P. C. Li and K. Meagher, Sperner partition systems, *J. Comb. Des.*, **21**(7) (2013), 267–279.
[30] C. Martínez, L. Moura, D. Panario, and B. Stevens, Locating errors using ELAs, covering arrays, and adaptive testing algorithms, *SIAM J. Discrete Math.*, **23** (2009/2010), 1776–1799.
[31] K. Meagher, L. Moura, and B. Stevens, A Sperner-type theorem for set-partition systems, *Electron. J. Comb.*, **12** (2005), 20.
[32] K. Mehlhorn, Data Structures and Algorithms 1: Sorting and Searching, Springer-Verlag, Berlin, 1984.
[33] C. Nie and H. Leung, A survey of combinatorial testing, *ACM Comput. Surv.*, **43**(2) (2011), **11**.
[34] X. Niu and H. Cao, Constructions and bounds for separating hash families, *Discrete Math.*, **341**(9) (2018), 2627–2638.
[35] S. Raaphorst, L. Moura, and B. Stevens, A construction for strength-3 covering arrays from linear feedback shift register sequences, *Des. Codes Cryptogr.*, **73**(3) (2014), 949–968.
[36] P. Sarkar and D. R. Stinson, Frameproof and IPP codes, In: Progress in Cryptology—INDOCRYPT 2001 (Chennai), Lecture Notes in Computer Science, Vol. 2247, Springer, Berlin, 2001, pp. 117–126.
[37] S. A. Seidel, M. T. Mehari, C. J. Colbourn, E. De Poorter, I. Moerman, and V. R. Syrotiuk, Analysis of large-scale experimental data from wireless networks, In: IEEE INFOCOM International Workshop on Computer and Networking Experimental Research Using Testbeds (CNERT), 2018, pp. 535–540.
[38] S. A. Seidel, K. Sarkar, C. J. Colbourn, and V. R. Syrotiuk, Separating interaction effects using locating and detecting arrays, In: International Workshop on Combinatorial Algorithms, 2018, pp. 349–360.
[39] G. B. Sherwood, S. S. Martirosyan, and C. J. Colbourn, Covering arrays of higher strength from permutation vectors, *J. Comb. Des.*, **14** (2006), 202–213.
[40] J. N. Staddon, D. R. Stinson, and R. Wei, Combinatorial properties of frameproof and traceability codes, *IEEE Trans. Inf. Theory*, **47** (2001), 1042–1049.
[41] D. R. Stinson, Tran Van Trung, and R. Wei, Secure frameproof codes, key distribution patterns, group testing algorithms and related structures, *J. Stat. Plan. Inference*, **86** (2000), 595–617.
[42] D. R. Stinson and R. Wei, Combinatorial properties and constructions of traceability schemes and frameproof codes, *SIAM J. Discrete Math.*, **11** (1998), 41–53.
[43] J. Torres-Jimenez and I. Izquierdo-Marquez, Covering arrays of strength three from extended permutation vectors, *Des. Codes Cryptogr.*, **86** (2018), 2629–2643.
[44] J. Torres-Jimenez and I. Izquierdo-Marquez, A simulated annealing algorithm to construct covering perfect hash families, *Math. Probl. Eng.*, **2018** (2018), 1860673.
[45] G. Tzanakis, L. Moura, D. Panario, and B. Stevens, Constructing new covering arrays from LFSR sequences over finite fields, *Discrete Math.*, **339**(3), 1158–1171 (2016).

Eduardo Camps, Hiram H. López, Gretchen L. Matthews, and Eliseo Sarmiento

Monomial-Cartesian codes closed under divisibility

Abstract: A monomial-Cartesian code closed under divisibility is an evaluation code defined by a set of monomials that is closed under divisibility, evaluated over a Cartesian product. In this work, we prove that the dual of a monomial-Cartesian code closed under divisibility is monomially equivalent to a code that belongs to this same family of codes. Then we describe the length, the dimension and the minimum distance of these codes in terms of the minimal generating set of monomials.

Keywords: Cartesian codes, monomial codes, monomial-Cartesian codes, decreasing codes

MSC 2010: 14G50, 11T71

1 Introduction

Evaluation codes form an important family of error-correcting codes, including Cartesian codes, algebraic geometry codes, and many variants finely tuned for specific applications, such as the locally recoverably codes defined by Tamo, Barg, and Vladut [11]. In this paper, we consider a particular class of evaluation codes, called monomial-Cartesian code closed under divisibility. Monomial-Cartesian codes closed under divisibility generalize Reed–Solmon and Reed–Muller codes, as we will see.

A monomial-Cartesian code closed under divisibility is defined using the following concepts. Let $K := \mathbb{F}_q$ be a finite field with q elements and $R := K[x_1, \dots, x_m]$ be the polynomial ring over K in m variables. Let $\mathcal{M} \subseteq R$ be a set of monomials such that $M \in \mathcal{M}$ and M' divides M, then $M' \in \mathcal{M}$. We say that such a set is **closed under divisibility**. Let $L(\mathcal{M})$ be the subspace of polynomials of R that are K-linear combinations of monomials of $\mathcal{M}$:

$$L(\mathcal{M}) := \operatorname{Span}_K\{M : M \in \mathcal{M}\} \subseteq R.$$

Acknowledgement: The first and fourth author were partially supported by SIP-IPN, project 20195717, and CONACyT. The third author is partially supported by NSF DMS-1855136.

Eduardo Camps, Eliseo Sarmiento, Escuela Superior de Física y Matemáticas, Instituto Politécnico Nacional, Mexico City, Mexico, e-mails: ecfmd@hotmail.com, esarmiento@ipn.mx
Hiram H. López, Department of Mathematics, Cleveland State University, Cleveland, OH, USA, e-mail: h.lopezvaldez@csuohio.edu
Gretchen L. Matthews, Department of Mathematics, Virginia Tech, Blacksburg, VA, USA, e-mail: gmatthews@vt.edu

https://doi.org/10.1515/9783110621730-014

Fix nonempty subsets $S_1, \ldots, S_m$ of K. Define their **Cartesian product** as

$$\mathcal{S} := S_1 \times \cdots \times S_m \subseteq K^m.$$

In what follows, $n_i := |S_i|$, the cardinality of S_i for $i \in [m] := \{1, \ldots, m\}$, and $n := |\mathcal{S}|$, the cardinality of $\mathcal{S}$. Fix a linear order on $\mathcal{S} = \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$, $\boldsymbol{s}_1 \prec \cdots \prec \boldsymbol{s}_n$. We define the **evaluation map**

$$\begin{array}{rccc} \mathrm{ev}_{\mathcal{S}}\colon & L(\mathcal{M}) & \to & K^n \\ & f & \mapsto & (f(\boldsymbol{s}_1), \ldots, f(\boldsymbol{s}_n)). \end{array}$$

From now on, we assume that the degree of each monomial $M \in \mathcal{M}$ in x_i is less than n_i. In this case, the evaluation map $\mathrm{ev}_{\mathcal{S}}$ is injective; see [5, Proposition 2.1]. The **complement** of $\mathcal{M}$ in $\mathcal{S}$ denoted by $\mathcal{M}^c_{\mathcal{S}}$, is the set of all monomials in R that are not in $\mathcal{M}$ and their degree respect i is less than n_i.

Definition 1.1. If $\mathcal{M} \subseteq R$ is closed under divisibility, then the image $\mathrm{ev}_{\mathcal{S}}(L(\mathcal{M})) \subseteq K^n$ is called the **monomial-Cartesian code closed under divisibility** associated to $\mathcal{S}$ and $\mathcal{M}$. We denote it by $C(\mathcal{S}, \mathcal{M})$.

The length and the dimension of a monomial-Cartesian code $C(\mathcal{S}, \mathcal{M})$ are given by $n = |\mathcal{S}|$ and $k = \dim_K C(\mathcal{S}, \mathcal{M}) = |\mathcal{M}|$, respectively [5, Proposition 2.1]. Recall that the **minimum distance** of a code C is given by

$$\delta(C) = \min\{|\operatorname{Supp}(\boldsymbol{c})| : 0 \neq \boldsymbol{c} \in C\},$$

where $\operatorname{Supp}(\boldsymbol{c})$ denotes the support of $\boldsymbol{c}$, that is the set of all nonzero entries of $\boldsymbol{c}$. Unlike the case of the length and the dimension, in general, there is no explicit formula for $\delta(C(\mathcal{S}, \mathcal{M}))$ in terms of $\mathcal{S}$ and $\mathcal{M}$.

The **dual** of a code C is defined by

$$C^\perp = \{\boldsymbol{w} \in K^n : \boldsymbol{w} \cdot \boldsymbol{c} = 0 \text{ for all } \boldsymbol{c} \in C\},$$

where $\boldsymbol{w} \cdot \boldsymbol{c}$ represents the **Euclidean inner product**. The code C is called a **linear complementary dual** (**LCD**) [8] if $C \cap C^\perp = \{\boldsymbol{0}\}$, and is called a **self-orthogonal** code if $C^\perp \subseteq C$.

Instances of monomial-Cartesian codes closed under divisibility for particular families of Cartesian products $\mathcal{S}$ and monomials sets that are closed under divisibility $\mathcal{M}$ have been previously studied in the literature. For example, a Reed–Muller code of order r in the sense of [12, p. 37] is monomial-Cartesian code closed under divisibility $C(K^m, M_r)$, where M_r is the set of monomials of degree less than r. An **affine Cartesian code** of order r is the monomial-Cartesian code closed under divisibility $C(\mathcal{S}, M_r)$. This family of affine Cartesian codes appeared first time in [4] and then independently in [6]. In [1], the authors studied the case when the finite field K is $\mathbb{F}_2$ and the set of monomials satisfy some decreasing conditions; then their results were generalized in [2] for $K = \mathbb{F}_q$ and monomials associated to curve kernels. The case when the set of

monomials $\mathcal{M}$ is a tensor product, the minimum distance of the associated code can be computed using the same ideas that [9].

It is important to note that some families of monomial-Cartesian codes are not closed under divisibility. For instance, the family of codes given in [10], which is well known for its applications to distributed storage, are not closed under divisibility because these are subcodes of Reed–Solomon codes where some monomials are omitted. To be precise, fix $r \geq 2$ with $r+1|n$. Set

$$V := \left\langle g(x)^j x^i : 0 \leq j \leq \frac{k}{r} - 1, 0 \leq i \leq r-1 \right\rangle$$

where $g(x) \in \mathbb{F}_q[x]$ has $\deg g = r+1$ and $\mathbb{F}_q = A_1 \dot{\cup} \cdots \dot{\cup} A_{\frac{n}{r+1}}$ with $|A_j| = r$ for all j so that $\forall \beta, \beta' \in A_j$,

$$g(\beta) = g(\beta').$$

Then $C(\mathbb{F}_q, V)$ is not closed under divisibility as $g(x)^j x^i \in V$ and x divides $g(x)^j x^i$ but $x \notin V$.

This notion of divisibility will be restricted to codes defined by sets of monomials as defined above. Recall that given a curve X over a finite field $\mathbb{F}$ and a divisor G on X, the space of rational functions associated with G, sometimes called the Riemann–Roch space of G, is

$$\mathcal{L}(G) := \{f \in \mathbb{F}(X) : (f) + G \geq 0\} \cup \{0\}$$

where (f) denotes the divisor of f. In general $\mathcal{L}(G)$ is not closed under divisibility, meaning $f \in \mathcal{L}(G)$ and $f = gh$ does not imply $g, h \in \mathcal{L}(G)$. For instance, if one considers the Hermitian curve X given by $y^q + y = x^{q+1}$ over $\mathbb{F}_{q^2}$, then $y \in \mathcal{L}((q+1)P_\infty)$. However, $y = x\frac{y}{x}$, but $(\frac{y}{x}) = (y) - (x) = qP_{00} - P_\infty - \sum_{b \neq 0} P_{0b} \not\geq -(q+1)P_\infty$. Hence, $\frac{y}{x} \notin \mathcal{L}((q+1)P_\infty)$.

In the next section, we prove that the dual of a monomial-Cartesian code closed under divisibility is also a code of the same type. Then we describe its basic parameters in terms of the minimal generating set. For more information about coding theory, we recommend [7, 13]. For algebraic concepts not described in this notes, we suggest to the reader [14]. We close this section with a bit of notation that will be useful in the remainder of this paper. We will use $K^* := K \setminus \{0\}$ to denote the multiplicative group of K. Given a point $\boldsymbol{a} = (a_1, \ldots, a_m) \in \mathbb{Z}_{\geq 0}^m$, $\boldsymbol{x^a}$ is the corresponding monomial in R; i. e., $\boldsymbol{x^a} := x_1^{a_1} \cdots x_m^{a_m}$.

2 Basic parameters

In this section, we continue with the same notation as in Section 1, in particular $\mathcal{M} \subseteq R$ is a set of monomials that is closed under divisibility, $\mathcal{S}$ represents a Cartesian set $\mathcal{S} = S_1 \times \cdots \times S_m$, $n_i = |S_i|$, for $i \in [m]$, $n = |\mathcal{S}|$ and $C(\mathcal{S}, \mathcal{M})$ represents the decreasing monomial-Cartesian code associated to $\mathcal{S}$ and $\mathcal{M}$.

A **monomial matrix** is a square matrix with exactly one nonzero entry in each row and column. Let C_1 and C_2 be codes of the same length over K, and let G_1 be a generator matrix for C_1. Then C_1 and C_2 are **monomially equivalent** provided there is a monomial matrix M with entries over the same field K so that G_1M is a generator matrix of C_2. Monomially equivalent codes have the same length, dimension, and minimum distance.

Definition 2.1. For $\boldsymbol{s} = (s_1, \dots, s_m) \in \mathcal{S}$ and $f \in R$, define the **residue** of f at $\boldsymbol{s}$ as

$$\operatorname{Res}_{\boldsymbol{s}} f = f(\boldsymbol{s}) \left(\prod_{i=1}^{m} \prod_{s_i' \in S_i \setminus \{s_i\}} (s_i - s_i') \right)^{-1},$$

and the **residues vector** of f at $\mathcal{S}$ as

$$\operatorname{Res}_{\mathcal{S}} f = (\operatorname{Res}_{\boldsymbol{s}_1} f, \dots, \operatorname{Res}_{\boldsymbol{s}_n} f).$$

Theorem 2.2. *The dual of the code $C(\mathcal{S}, \mathcal{M})$ is monomially equivalent to a monomial-Cartesian code closed under divisibility. Even more, the set*

$$\Delta := \left\{ \operatorname{Res}_{\mathcal{S}} \frac{x_1^{n_1-1} \cdots x_m^{n_m-1}}{M} : M \in \mathcal{M}^c \right\}$$

forms a basis for the dual $C(\mathcal{S}, \mathcal{M})^\perp$, meaning

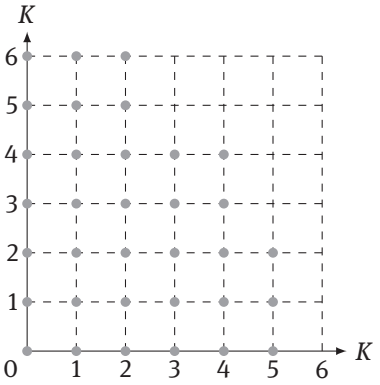
$$C(\mathcal{S}, \mathcal{M})^\perp = \operatorname{Span}_K(\Delta).$$

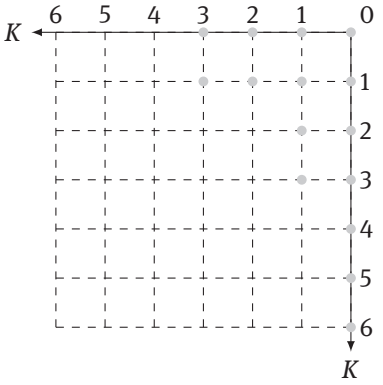
Proof. We start by proving that the set

$$\Delta' := \left\{ \frac{x_1^{n_1-1} \cdots x_m^{n_m-1}}{M} : M \in \mathcal{M}_{\mathcal{S}}^c \right\}$$

is closed under divisibility. Let $M \in \mathcal{M}_{\mathcal{S}}^c$ and $\boldsymbol{x^a}$ a divisor of $\frac{x_1^{n_1-1} \cdots x_m^{n_m-1}}{M}$. Then there exists a monomial $\boldsymbol{x^b}$ in R such that $\frac{x_1^{n_1-1} \cdots x_m^{n_m-1}}{M} = \boldsymbol{x^a x^b}$. As $M \in \mathcal{M}^c$ and $\mathcal{M}$ is closed under divisibility, then $\boldsymbol{x^b} M \in \mathcal{M}^c$ and $\boldsymbol{x^a} = \frac{x_1^{n_1-1} \cdots x_m^{n_m-1}}{\boldsymbol{x^b} M} \in \Delta'$. This proves that the set Δ' is closed under divisibility. Due to [5, Theorem 2.7] and its proof, Δ is a basis for the dual $C(\mathcal{S}, \mathcal{M})^\perp$. Finally, it is clear that $\operatorname{Span}_K\{\boldsymbol{c} : \boldsymbol{c} \in \Delta\}$ is monomially equivalent to $\operatorname{ev}_{\mathcal{S}}(\Delta')$, which is a monomial-Cartesian code closed under divisibility. □

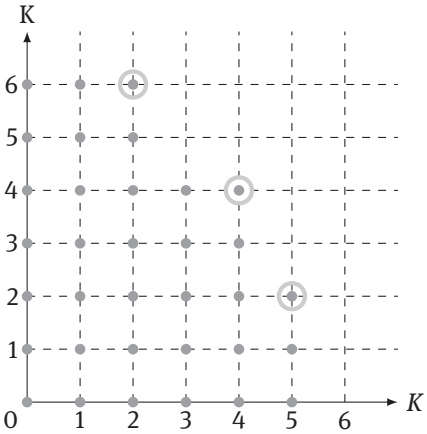
Example 2.3. Let $K = \mathbb{F}_7$, $\mathcal{S} = K^2$ and $\mathcal{M}$ the set of monomials of $K[x_1, x_2]$ whose exponents are the points in the left picture below. Then the code $C(\mathcal{S}, \mathcal{M})$ is generated by the vectors $\operatorname{ev}_{\mathcal{S}}(M)$, where M is a monomial whose exponent is a point in the left picture below and the dual $C(\mathcal{S}, \mathcal{M})^\perp$ is generated by the vectors $\operatorname{Res}_{\mathcal{S}}(M)$, where M is a monomial whose exponent is a point in the right picture below.





Definition 2.4. A subset $\mathcal{B}(\mathcal{M}) \subseteq \mathcal{M}$ is a **generating set** of $\mathcal{M}$ if for every $M \in \mathcal{M}$ there exists a monomial $B \in \mathcal{B}(\mathcal{M})$ such that M divides B. A generating set $\mathcal{B}(\mathcal{M})$ is called **minimal** if for every two elements $B_1, B_2 \in \mathcal{B}(\mathcal{M})$, B_1 does not divide B_2 and B_2 does not divide B_1.

Example 2.5. Let $K = \mathbb{F}_7$, $\mathcal{S} = K^2$ and $\mathcal{M}$ the set of monomials of $K[x_1, x_2]$ whose exponents are the points in the left picture of Example 2.3. The circles in the following picture are the exponents of the monomials that belong to the minimal generating set of $\mathcal{M}$.



From now on, $\mathcal{B}(\mathcal{M})$ denotes the minimal generating set of $\mathcal{M}$. We are going to describe properties of the code $C(\mathcal{S}, \mathcal{M})$ in terms of $\mathcal{B}(\mathcal{M})$. The following proposition says how to find a generating set of $\mathcal{M}^c_{\mathcal{S}}$ in terms of $\mathcal{B}(\mathcal{M})$.

Proposition 2.6. *Given a monomial* $M = x_1^{a_1} \cdots x_m^{a_m} \in \mathcal{B}(\mathcal{M})$, *define the monomials* $P(M) := \{\frac{x_1^{n_1-1} \cdots x_n^{n_m-1}}{x_i^{a_i-1}} : i \in [m], \text{ and } n_i - a_i - 2 \geq 0\}$. *The set*

$$\gcd(P(M))_{M \in \mathcal{B}(\mathcal{M})}$$

is a generating set of $\mathcal{M}^c$. *The set* gcd *is defined by induction, if* M_1, M_2, *and* M_3 *are elements of* $\mathcal{B}(\mathcal{M})$, *then*

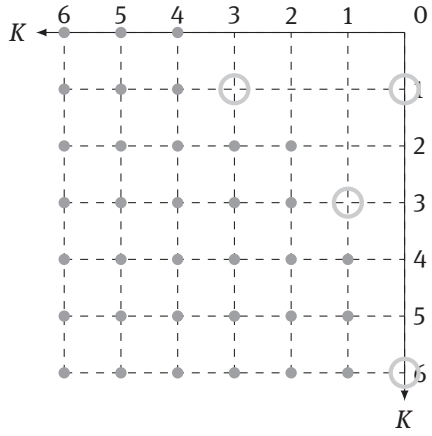
$$\gcd\big(P(M_1), P(M_2), P(M_3)\big) = \gcd\big(\gcd\big(P(M_1), P(M_2)\big), P(M_3)\big),$$

where $\gcd(P(M_1), P(M_2)) = \{\gcd(M_1', M_2') : M_1' \in M_1, M_2' \in M_2\}$.

Proof. It is clear that for every monomial $M = x_1^{a_1} \cdots x_m^{a_m} \in \mathcal{B}(\mathcal{M})$ the set $P(M)$ is a minimal generating set for $\{M\}^c$. Given any two monomials M_1 and M_2, the set $\{\gcd(M_1, M_2)\}$ is a minimal generating set for the set of monomials that divide M_1 and M_2, thus the result follows. □

It is important to note that the set $\gcd(P(M))_{M \in \mathcal{B}(\mathcal{M})}$ from Proposition 2.6 is not always a minimal generating set, as the following example shows.

Example 2.7. Let $K = \mathbb{F}_7$, $\mathcal{S} = K^2$ and $\mathcal{M}$ the set of monomials of $K[x_1, x_2]$ whose exponents are the points in the left picture of Example 2.3. The circles in the picture of Example 2.5 are the exponents of the monomials that belong to $\mathcal{B}(\mathcal{M})$. The circles below are the exponents that belong to $\gcd(P(M))_{M \in \mathcal{B}(\mathcal{M})}$. It is clear that it is not a minimal generating set.



Theorem 2.8. *Let* P_i *be the subsets of size* i *of* $\mathcal{B}(\mathcal{M})$. *Then:*

(i) *The length of* $C(\mathcal{S}, \mathcal{M})$ *is given by* $\prod_{i=1}^{m} n_i$.

(ii) *The dimension of the code* $C(\mathcal{S}, \mathcal{M})$ *is*

$$\sum_{i=1}^{|\mathcal{B}(\mathcal{M})|} \left((-1)^{i-1} \sum_{T \in P_i} \prod_{j=1}^{n} (t_j + 1) \right),$$

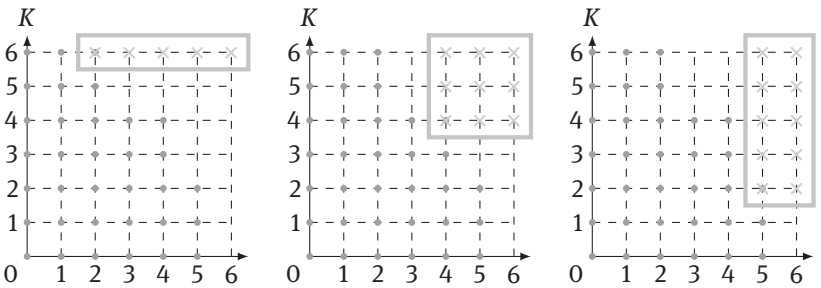
where $(t_1, \ldots, t_m)$ *is the exponent of the* gcd *of the elements of* T.

(iii) *The minimum distance of* $C(\mathcal{S}, \mathcal{M})$ *is given by*

$$\min\left\{ \prod_{i=1}^{m} (n_i - a_i) : x_1^{a_1} \cdots x_m^{a_m} \in \mathcal{B}(\mathcal{M}) \right\}.$$

Proof. (i) It is clear because $\prod_{i=1}^{m} n_i$ is the cardinality of $\mathcal{S}$; (ii) Given two monomials M and M', we see that $\gcd(M, M')$ is the minimal generating set of the set of monomials that divide to M and also to M'. For any monomial $M = x_1^{t_1} \cdots x_m^{t_m}$, $\prod_{j=1}^{n}(t_j + 1)$ is the number of monomials that divide M. Thus the dimension follows from the inclusion exclusion theorem; (iii) Let $\prec$ be the graded-lexicographical order and take $f \in \mathrm{Span}_K\{M : M \in \mathcal{M}\}$. If $M = x_1^{b_1} \cdots x_m^{b_m}$ is the leading monomial of f, in [3, Proposition 2.3] the author proved that $|\mathrm{Supp}(\mathrm{ev}_{\mathcal{S}} f)| \geq \prod_{i=1}^{m}(n_i - b_i)$. As $\mathcal{B}(\mathcal{M})$ is a minimal generating set of $\mathcal{M}$, there exists $M' = x_1^{a_1} \cdots x_m^{a_m} \in \mathcal{B}(\mathcal{M})$ such that M divides M'. Thus $|\mathrm{Supp}(\mathrm{ev}_{\mathcal{S}} f)| \geq \prod_{i=1}^{m}(n_i - a_i)$ and $\delta(C(\mathcal{S}, \mathcal{M})) \geq \min\{\prod_{i=1}^{m}(n_i - a_i) : x_1^{a_1} \cdots x_m^{a_m} \in \mathcal{B}(\mathcal{M})\}$. Assume for $i \in [m]$, $S_i = \{s_{i1}, \ldots, s_{in_i}\}$. Let $x_1^{\alpha_1} \cdots x_m^{\alpha_m} \in \mathcal{B}(\mathcal{M})$ such that $\prod_{i=1}^{m}(n_i - \alpha_i) = \min\{\prod_{i=1}^{m}(n_i - a_i) : x_1^{a_1} \cdots x_m^{a_m} \in \mathcal{B}(\mathcal{M})\}$. Define $f_\alpha := \prod_{i=1}^{m} \prod_{j=1}^{\alpha_i}(x_i - s_{ij})$. As $|\mathrm{Supp}(\mathrm{ev}_{\mathcal{S}} f_\alpha)| = \prod_{i=1}^{m}(n_i - a_i)$, and $f_\alpha \in \mathrm{Span}_K\{M : M \in \mathcal{M}\}$ because all monomials that appear in f_α divide $x_1^{\alpha_1} \cdots x_m^{\alpha_m}$, then we have $\delta(C(\mathcal{S}, \mathcal{M})) \leq \min\{\prod_{i=1}^{m}(n_i - a_i) : x_1^{a_1} \cdots x_m^{a_m} \in \mathcal{B}(\mathcal{M})\}$ and the result follows. □

Example 2.9. Let $K = \mathbb{F}_7$, $\mathcal{S} = K^2$ and $\mathcal{M}$ the set of monomials of $K[x_1, x_2]$ whose exponents are the points in the left picture of Example 2.3. The length of the code is 49, which is the total number of grid points in $\mathcal{S}$. The dimension is 34, which is the total number of points in the left picture of Example 2.3. The minimal generating set $\mathcal{B}(\mathcal{M})$ is $\{x_1^2 x_2^6, x_1^4 x_2^4, x_1^5 x_2^2\}$. By Theorem 2.8 $|\mathrm{Supp}(\mathrm{ev}_{\mathcal{S}} x^2 y^6)| \geq 5$, which is the number of grid points between the point $(2, 6)$ and the point $(6, 6)$. See first picture (from left to right) below. In a similar way, $|\mathrm{Supp}(\mathrm{ev}_{\mathcal{S}} x_1^4 x_2^4)| \geq 9$ and $|\mathrm{Supp}(\mathrm{ev}_{\mathcal{S}} x_1^5 x_2^2)| \geq 10$. See second and third picture (from left to right) below. As $\min\{5, 9, 10\} = 5$, the minimum distance $\delta(C(\mathcal{S}, \mathcal{M}))$ is 5.



Bibliography

[1] M. Bardet, V. Dragoi, A. Otmani, and J. P. Tillich, Algebraic properties of polar codes from a new polynomial formalism, In: 2016 IEEE International Symposium on Information Theory, 2016, pp. 230–234.

[2] E. Camps, E. Martínez-Moro, and E. Sarmiento, Vardøhus Codes: Polar codes based on castle curves kernels, *IEEE Trans. Inf. Theory* (2019), doi:10.1109/TIT.2019.2932405.

[3] C. Carvalho, On the second Hamming weight of some Reed–Muller type codes, *Finite Fields Appl.*, **24** (2013), 88–94.
[4] O. Geil and C. Thomsen, Weighted Reed–Muller codes revisited, *Des. Codes Cryptogr.*, **66**(1–3) (2013), 195–220.
[5] H. H. López, G. L. Matthews, and Ivan Soprunov, Monomial-Cartesian codes and their duals, with applications to LCD codes, quantum codes, and locally recoverable codes, https://arxiv.org/pdf/1907.11812.pdf.
[6] H. H. López, C. Rentería-Márquez, and R. H. Villarreal, Affine Cartesian codes, *Des. Codes Cryptogr.*, **71**(1) (2014), 5–19.
[7] F. J. MacWilliams and N. J. A. Sloane, The Theory of Error-Correcting Codes, North-Holland, 1977.
[8] James L. Massey, Linear codes with complementary duals, *Discrete Math.*, **106–107** (1992), 337–342.
[9] I. Soprunov and J. Soprunova, Bringing toric codes to the next dimension, *SIAM J. Discrete Math.*, **24**(2) (2010), 655–665.
[10] I. Tamo and A. Barg, A family of optimal locally recoverable codes, *IEEE Trans. Inf. Theory*, **60**(8) (2014), 4661–4676.
[11] A. Barg, I. Tamo, and S. Vladut, Locally recoverable codes on algebraic curves, *IEEE Trans. Inf. Theory*, **63**(8) (2017), 4928–4939.
[12] M. Tsfasman, S. Vladut and D. Nogin, Algebraic Geometric Codes: Basic Notions, Mathematical Surveys and Monographs, Vol. 139, American Mathematical Society, Providence, RI, 2007.
[13] J. H. Van Lint, Introduction to Coding Theory, third edition, Graduate Texts in Mathematics, Vol. 86, Springer-Verlag, Berlin, 1999.
[14] R. H. Villarreal, Monomial Algebras, second edition, Monographs and Research Notes in Mathematics, 2015.

De Gruyter Proceedings in Mathematics

Mahmoud Filali (Ed.)
Banach Algebras and Applications. Proceedings of the International Conference held at the University of Oulu, July 3–11, 2017
ISBN 978-3-11-060132-9, e-ISBN 978-3-11-060241-8

Ioannis Emmanouil, Anargyros Fellouris, Apostolos Giannopoulos, Sofia Lambropoulou (Eds.)
First Congress of Greek Mathematicians. Proceedings of the Congress held in Athens, Greece, June 25–30, 2018
ISBN 978-3-11-066016-6, e-ISBN 978-3-11-066307-5

Theodora Bourni, Mat Langford (Eds.)
Mean Curvature Flow. Proceedings of the John H. Barrett Memorial Lectures held at the University of Tennessee, Knoxville, May 29–June 1, 2018
ISBN 978-3-11-061818-1, e-ISBN 978-3-11-061836-5

Paul Baginski, Benjamin Fine, Anja Moldenhauer, Gerhard Rosenberger, Vladimir Shpilrain (Eds.)
Elementary Theory of Groups and Group Rings, and Related Topics. Proceedings of the Conference held at Fairfield University and at the Graduate Center, CUNY, November 1–2, 2018
ISBN 978-3-11-063673-4, e-ISBN 978-3-11-063838-7

Galina Filipuk, Alberto Lastra, Sławomir Michalik, Yoshitsugu Takei, Henryk Żołądek (Eds.)
Complex Differential and Difference Equations. Proceedings of the School and Conference held at Będlewo, Poland, September 2–15, 2018
ISBN 978-3-11-060952-3, e-ISBN 978-3-11-061142-7

Mohammad Ashraf, Vincenzo De Filippis, Syed Tariq Rizvi (Eds.)
Algebra and Its Applications. Proceedings of the International Conference held at Aligarh Muslim University, 2016
ISBN 978-3-11-054092-5, e-ISBN 978-3-11-054240-0

Alexander Katz (Ed.)
Topological Algebras and their Applications. Proceedings of the 8th International Conference on Topological Algebras and their Applications, 2014
ISBN 978-3-11-041433-2, e-ISBN 978-3-11-041355-7

www.degruyter.com